Designing behavior: interaction, cognition, biology and AI

Simon PENNY

Encyclopedia Of New Media Art
Bloomsbury UK (in press)

Keywords: cognition, interaction, Artificial Intelligence, Artificial Life, spectactor

List of suitable images
Petit Mal (2 images, supplied by author)
Traces (2 images, supplied by author)

What business has an artist in critiquing the rhetorics of computing and artificial intelligence? This is a question I might have been asked in 1987 when I published my first essay on the subject, entitled *Simulation, Digitisation, Interaction: the impact of computing on the Arts*. As indicated by this title, my preoccupation as an artist was (and would be for the ensuing two decades) with the phenomenon of real-time computing as a new technological context that would support a new kind of aesthetic engagement. I wanted to build systems that were '*lively*', not *lifelike* in an anthropomorphic or biomorphic sense (animatronics and similar pursuits always seemed theoretically naïve to me). I wanted to make *designing behavior* a central part of my work, and the experience of "lifelike" behavior to be central to the experience of the work. (Documentation of such works, including Petit Mal, Fugitive and Traces can be found at simonpenny.net). The behavioral dimension a behaving artwork was my central concern (so much so that questions of imagery - so central to much 'new media art' - sometimes felt to me like a tedious distraction).

Such a remit would, one might assume, predispose one to wholeheartedly pursue the agendas of conventional AI, a topic of much public interest and concern, then as now. In my case the opposite was true. I never drank the AI kool-aid. I was under no misapprehension that computers would achieve human-like cognition, in my lifetime at least, if ever. I still believe this, and I find arguments to the contrary vapid. Neither was I under any misapprehension that the computational behaviors of my (or any systems) systems constituted life or living. This position relates to a key debate in the Artificial Life movement that succeeded (the collapse of) first generation Artificial Intelligence, in the late 1980s. In the AI community, there had been a distinction between 'Hard AI' and 'Soft AI'. Hard AI asserted that AI was intelligence, Soft AI held that AI only simulated intelligence. A similar distinction was made in Artificial Life circles in the 90s, between Hard Alife (digital 'life forms' were alive) and Soft Alife (digital systems could simulate some of the qualities of living systems). This notion of (non-anthropomorphic) *liveliness* provided me a perspective from which to critique discourses of artificial intelligence and of cognition that have now informed my critical inquiry for decades (see Penny 2017).

In the second-order cybernetic sense captured succinctly by Humberto Maturana when he quipped '*everything said is said by an observer'* - and given that art *is* experience (Dewey) - then *liveliness* is an attribution applied by someone who has an experience. I saw my task as

building a technological infrastructure to facilitate such experience, utilising real-time computational technology that was reactive, based on real-time sensor data. That experience, as I conceived it, always involved active, embodied participation by the party formerly known as 'viewer' or 'audience', and whom I came to refer to, using Augusto Boal's neologism, a *spectactor*.

It hit me as an epiphany in the mid 90s that this mutuality of embodied interaction destabilises the aesthetic calculus based on the binary distinction between observer and observed, between subject (and all that implies) and object. A viewer looks at painting - the viewer is the active party, the painting is passive. In my work, the spectactor acts, the work responds, and so on. A spectactor engages in a dance or a conversation (as opposed to a contemplation): a sequence of events in which each is the consequence of the last and precursor of the next. This scenario I realized, stood outside conventional conception of the work-viewer relationship, at least in the plastic arts.

My commitment, as an artist, to the primacy of sensorial experience, led me to reject unquestioned aspects of computer culture, both artifactual and theoretical. I rejected the general idea that in order to interact with a computer, one had to pre-process one's human qualities into streams of symbolic input (keystrokes etc). Keyboards pointers and screens were anathema. My systems had to interpret human bodily dynamics that other humans interpret as meaningful. There would be no spoon-feeding symbols to the computer.

In insisting on sensorial immediacy, I'd set myself a technical research agenda developing sensor systems that could adequately fulfill the task. The same goal also led me into a long critical analysis of the nature of the technology, its capabilities and what it excluded or made difficult, and why. It also begged inquiry about the claims made about the technology and to what degree they were justified - many of which were part of the larger debates around AI. This paper discusses these topics, along with whether (and if so, how) these things have changed in the intervening decades. (This essay has been written in parallel with my essay *Of aesthetics and computation – contextualising first generation Media Arts* in ENMA volume (Thomas). That essay traces an autobiographical, culturally contextualised story of my practice. In this essay I attempt to operate in a more academic register. I encourage readers to read both.)

Artificial Intelligence and biological intelligence

The term Artificial Intelligence must be understood as being, from the outset, an hysterical and perhaps disingenuous marketing strategy. I have noted that a more honest name for the field might be *automated reasoning*, and if this had been the case, much of the public consternation, and much of the sensational press, would never have occurred (and much of the research funding might not have been forthcoming). The technology that we are encouraged to call *artificial intelligence* is predicated on a long tradition of humanist/rationalist thought, specifically the idea that intelligence is a matter of manipulation and communication of symbols. Words and numbers are, in this assessment, (the) vehicles in which intelligence rides. From poetry to statistics, intelligence is purported, under this interpretation, to be *representational*, it comprises the assembling and manipulation of symbols that say something about the world, about experience.

AI is thus a product of a long philosophical tradition of reasoning on representations, of logical operations on symbols.

Like so many terms in this territory, *intelligence* has multiple meanings. The philosophical tradition in which AI is grounded - enlightenment rationalism more or less - privileges conscious logical operations, propositional knowledge. With respect to the range of way that animals succeed in the world, this seems a very narrow conception. When a coyote stalks a rabbit, what is it thinking about? Any 'planning' is of the order of 'staying hidden from the rabbit while moving as close as possible'. When (if) it makes the kill, is it planning how to tear apart the animal? Is its experience of the death bite and blood a computation? When a frog snaps its tongue at (what it takes to be) a fly, its brain functions as little but a fast-switching reflex action coordinator. But to represent this process as: 'there is a fly, due south, 15 degrees elevation, travelling west at some angular speed' is to process 'phenomena as sensed' into representational language. When a person identifies and enjoys the scent of narcissus in the late winter garden, is intelligence occurring? When we are afraid of the sound of thunder or afraid of swimming in rough surf, are we to call this 'intelligence'?

The question of the relationship between artificial intelligence and the arts begs questions raised above, because the arts are about experiences, and AI has never been about experiences. Indeed, the idea that (an) AI might be capable of having something we might fairly call an experience is currently technically impossible certainly, and not conceivable within the capabilities of the technology nor, many would say any technology based on current practices.

Practices like my own function as a crucible - an experimental environment - for exploring such issues. Another way of understanding this question is to see it as capturing the differing viewpoints of Anglo-American analytical philosophy and 'continental' phenomenological philosophy, or between the internalist, mentalist, cognitivist, computationalist cognitive science of the latter part of the twentieth century and the 'postcognitivist' approaches captured by the descriptors embodied, embedded, enactive, situated and distributed (and perhaps 'extended' - along with 'ecological' psychology, material engagement, intersubjectivity and other neologisms). Why all these neologisms? Because, the paradigms of internalist cognition that have shape language for three hundred years, are crumbling.
Throughout the C20th in western scholarship, there have been (roughly) two opposing teams that have understood 'cognition' or intelligence through wildly different lenses. On one hand, the area roughly defined as philosophy, specifically 'philosophy of mind', and more generally as enlightenment rationalist thought, has hewn to a symbolic orientation. Its more logico-mathematical branch has led to the automation of symbolic logic in/as digital computing. Such schools of thought assume an industrial production-line model of cognition, with thinking machinery grinding along, occasionally receiving 'input' and processing it into 'output' (sometimes in 'packets'!). This *big idea* about cognition implies three separate assumptions. That thinking occurs in a 'black box' isolated from the world. That cognition is sequential, and that perception occurs as a preliminary (and peripheral) input stage. That perception entails the pre-processing of sensation into symbolic tokens, which are the medium of cognition. All of these (mechanistic) assumptions are dubious, and deploy metaphors from pre-existing technological formations, as if such models pertained to or had explanatory relevance to the neurochemical soup of flows and the complex interactions that characterise biological existence.

On the other hand, thinkers whose practices have been based in observations of nature, from Jakob von Uexküll, to J.J. Gibson to Humberto Maturana and Francisco Varela, along with various philosophers such as Gilbert Ryle and Michael Polanyi, Martin Heidegger and Maurice Merleau-Ponty, have looked at intelligence as what animals do to have success in the world. This orientation has two major aspects. First, to paraphrase Maturana and Varela, living is cognizing, and anything that lives cognizes – down to lowly bacteria. This point of view defuses the human-exceptionalism of the enlightenment tradition, as well as the separation of thinking from physicality, as defined in Cartesian Dualism (see below). Secondly, it asserts that thinking is immersed in, is engaged with and is about *the world* – the dynamic, processual, enactive experience of living (or, if that thinking is not successful, of coming to not-live). This then allows that catching a pretty scent or avoiding danger are kinds of thinking. One might fairly object that throwing the cognitive net so wide means that nothing is left outside cognition, which for some would render the concept of marginal value. There are two sides to this objection: leaving nothing outside cognition serves salutary purposes outlined above. But the word itself is defined within an enlightenment rationalist framework.

Which brings us back to art. Art is experiential. Very little is more basic to the rhetoric of the arts, at least in the modern period. Whether its punk rock or Shakespeare, *being there* is deemed to be crucially important. Likewise, we want to see a painting 'in the flesh', that's why we have architectural oddities called museums – so we can *muse*. Cinema, of course, provides a problematic middle case - even if you go to the movie theatre, what you're looking at is not really there. The same can be said for recorded music (this doesn't seem to stop us from liking it). So, from this perspective, doing art with artificial intelligence; trying to harness one to the other; or attempting to splice these two together: these attempts bring down upon our heads this implicit schism in western culture, between what we experience and what we say about it; between what *is* and how it is described; between ontology and epistemology, between *know-how* and *know-that* (Ryle).

That is not to say that art - the arts - does not often entail a substantial amount of 'saying about' - symbolic content. We, and the arts, are immersed in western culture after all. This does make the matter more murky. And when we experience art, we seem compelled to talk about it, with our students or friends – that is, we construct (linguistic) representations that we purport to have some relevance to our experience, and by which we attempt to inculcate, in our listener or audience, an experience analogous to our own (which we - implicitly or explicitly - are probably recommending as a better or more authoritative kind of experience).

Deploying computational technologies and techniques in the pursuit of art-making is therefore by no means a trivial deployment of technology, like using a hammer to drive a nail. As Theodor Rosak, in1986, presciently observed: *"Embodied in the machine there is an idea of what the mind is and how it works. The idea is there because scientists who purport to understand cognition and intelligence put it there. No other teaching tool has ever brought intellectual baggage of so consequential a kind to it*. While marketers of such technologies would have us believe that these technologies are simple and straightforward in their functions, this is clearly not so. The very rhetoric of enhancing productivity and efficiency that informs so much marketing rhetoric gives the lie to this. And when juxtaposed with the 'pleasure' marketing angle

– expanded opportunities for fun, play, sociality, and …*freedom*, the armatures of a capitalist value system become exposed as baldly as rocks at low tide. There is much to be deconstructed, and there is much to be constructed, if computing is to work effectively in the context of creative culture without being a *Trojan horse.* Or perhaps, in these pandemic days, the appropriate metaphor is one of infection, and the diverse, debilitating (and unexpected) long-term effects suffered by many.

The question of artificial intelligence

The perspective of *liveliness* provides a particular purchase on the question of artificial intelligence, because it insists on providing sensorimotor experience of something that gives the impression of, if not intelligence, then awareness. "*It knows I'm here*" is a more visceral, more *creaturely* sense of intelligence than that defined by the automation of propositional logic, ie abstract reasoning. The conception of reason as concerned with abstract rules was fundamental to symbolic AI, as the mechanisation of symbolic logic. So, while Hubert Dreyfus and crew were applying phenomenology to dismantle the presumptions of AI, especially the idea that 'common-sense' would be amenable to propositional logic, I was trying to provide my systems with a modicum of 'common sense'. This was consistent with the notion of situated bottom-up robotics as proposed by Rodney Brooks and others in the Artificial Life movement of the 90s. Brooks was an originator of reactive robotics and its theoretical validation. In asserting '*the world is its own best model*' he disputed the preoccupation with mapping, planning and representation that was central to AI-based robotics (Brooks 1991).

The challenges faced by robotics and AI in those days (the late 1980s and early1990s) were not only due to the shortcomings of computational logic vis-a-vis common-sense: the assumption in the AI community was that common-sense *could* be represented symbolically, it was just a matter of developing the right technique. But increasingly it came to be understood that common sense was a problem-in-principle for algorithmic computing – computers would never 'get it'. The fundamental erroneous assumption was that common sense was the kind of thing that *could be* represented. The community had clearly not comprehended Polanyi's arguments about 'tacit' knowledge. But further, the very idea that embodied experiences could be usefully 'abstracted' was brought into question by Varela, Thompson and Rosch, when they noted, in 1991 (in a what amounts to a postscript on the common-sense crisis): *...if we wish to recover commonsense, then we must invert the representationalist attitude by treating context-dependent know-how not as a residual artefact that can be progressively eliminated by the discovery of more sophisticated rules but as, in fact, the very essence of creative cognition.* (Varela, Thompson and Rosch. 1991). Here the authors succinctly capture the *big lie* of enlightenment/Platonist thought: that embodied physical experiences are little but pointers to some abstract truth in the realm of ideals. Here, one might object to the elision of three separate ideas - abstractions in some immaterial 'heaven'; symbolic operations in the brain; and representations in computer code. I would counter that the latter two are structured according to the assumption of that (ancient) first, revealing the persistent dualisms that structure western thought: embodied experience is the ground of awareness, as opposed to: embodied experience is illusory, fallible or simply pointers to some abstract truth.

Within the internalist paradigm cognition happens *in the brain*, as reasoning on symbols, the idea that cognition might occur as dynamic interaction of an agent fully immersed in a world is simply unthinkable. Hence, the idea of connecting computer to world had secondary importance, and sensor, or 'machine perception' research was thus devalued, and lagged behind research into the manipulation of symbols. It was always assumed, in the AI community, that it was ok to simply 'make up' or simulate the assumed symbolic output of those imagined sensors, because the purported interpretation of sensory experience was going to amount to some kind of simple translation, a lookup table, when it finally occurred. Ha! They could not have been more wrong. The very idea that sensing is just a more or less interchangeable attachment, like a headlight on a bicycle, turns out to be entirely incorrect, as they might have realised, had they read McCulloch, Pitts, Lettvin and Maturana famous paper *What the frog's eye tells frog's brain* (Lettvin et al, 1940). The frogs' eye does not provide image data for the frogs' brain so that said brain can construct an image, infer the presence and position of a fly, then calculate tongue trajectory and instruct tongue muscles. Why? Because by then, the fly would be long gone and this particular frog would starve, thereby winning the Darwin Award for frogs by taking its genes out of the evolutionary pool.

Having no sensors to provide knowledge of the world, AI was not 'situated' in any real way. The world an AI program 'knew' was the world that had been described (proscribed) to it by its programmers, as *representations*: symbols in code. AI was never any*where*, it was nowhere. It was a make-believe world: '*Imagine there is a place like this, where events of kind X are contingent on events of kind Y. Now – make the right decisions*.' When seen in this way, the whole exercise seems absurd, or, at the very best, an abstract logical game, like chess. This is still largely the case. Even in modern machine learning, the systems draw upon symbolic content put there (in the vast databases of Google, Facebook, Amazon, et al) *by people*, who dutiously identify faces, make locations on maps and disambiguate graphics.

In the western tradition, the study of mind is embedded in a humanist tradition (rooted in Cartesianism) that has enforced both a fictitious res cogitans/res extensa dualism and equally dubious human exceptionalism, by formulating, essentially, that intelligence is reason - hence (symbolic) AI was automated reasoning. Philosophy of mind, as defined by Cartesianism, was prohibited from accessing empirical reasoning because mind/soul was immaterial, and therefore not amenable to empiricism. This point may require some unpacking for non-philosopher. In short – Descartes' deal with the Pope was that the person would be subdivided into two non-intersecting parts, the *res cogitans* and the *res extensa*, the material part and the thinking part. Science, and thus empiricism, would apply to all things material, and hence not to things pertaining to the immaterial soul. Philosophy of mind was thus prohibited from accessing empirical reasoning due to the res cogitans/res extensa binary that rendered what we now call cognition as immaterial, and therefore in the territory of religious faith that explicitly resides beyond the reach of empiricism, which related only to material things. The (rather laughable) result was that philosophy of mind, as a discipline, was forced to remain secluded in a medieval scholasticism, while science flourished around it.

It is double ironic then, that (symbolic)AI, in parallel with its rejection of cybernetics, threw its philosophical lot in with the already retrogressive Cartesian attitude, and took Putnam's *functionalism* to defend the idea that intelligence was immaterial, and might just as easily occur

in digital electronics as in neurons. Ethology, on the other hand, was concerned with the kinds of awareness of other-than-human species, and so was free of the constraints of both human exceptionalism and the mind-body dualism, and was free to deploy empirical techniques. In the early years of the C20th, Pioneer ethologist Jakob von Uexküll developed his (profound) concept of the Umwelt: the subjective experience-world that a creature is given (and constrained) by the kinds of senses it has. The upshot being that creatures can share the same physical space but have entirely different 'worlds'. This perspective provided me with ways of thinking the interactive quasi-animals I made, that made their way in their worlds by virtue of their own awareness. (I developed infrared video sensing and used other sensing modalities (ultrasonic, pyroelectric, etc) that humans do not possess, for my systems.

Creatures, human and non-human, live in contingent enactive temporal flow, immersed in (always changing) contexts, always adapting. Many creatures, certainly higher animals appear to possess this ineffable quality of *awareness*, that has no equivalent in any technology. As philosopher John Haugeland pithily noted "*The trouble with artificial intelligence is that computers don't give a damn*" (Haugeland,1998) which neatly demolishes that kind of apocalyptic AI dread - his point being that the kind of intelligence we have is predicated on a sense of selfhood, and we have little conception of how to put such a quality into a machine, or even whether that would be possible. This does beg the question of how it is we, as sentient beings, come to have the assurance that we are in fact 'here'. It was such questions that drew me, increasingly, into consideration of matters of cognition. That is not to say that there are not very justified fears regarding deployment of increasingly sophisticated AI systems, its just that this is not one of them.

Cognition and Embodiment

This critique of AI as fundamentally abstract and ungrounded was thrown into high relief due to my training in the arts. The sensorial immediacy of art provides a very different, if seldom articulated, conception of intelligence. In this context, intelligence is not the manipulation of abstract symbols in some nonphysical nowhere. In doing the skilled bodily practices of an artist or artisan, it seems self-evident that cognition abides at the intersection between self and world. It is intelligent action, or intelligent perception in the world that matters in both the making of, and the experiencing of, art. This sense of intelligence is consistent with a non-human-exceptionalist, evolutionary view. For every creature before, say, Plato, intelligence was what assured survival. This logic made von Uexküll's work immediately accessible to me, and likewise made much of the work in Alife (Artificial Life) immediately relevant. (Artificial Life was the now almost forgotten flourishing of biologically inspired computing research that arose in the 90s as a result of the collapse of first generation symbolic AI, due the aforementioned common sense problem see below).

The phenomenologically informed critique of AI had another, related, effect in the field of cognition, because the Dreyfusian argument hinged on the idea that reasoning on symbols just isn't the way people think, and indeed, there are diverse aspects of human being-in-the-world that cannot be easily shoehorned into a conventional dualistic conception of thinking. As J.J. Gibson put it: *What a thing is and what it means are not separate, the former being physical and the latter mental, as we are accustomed to believe.* (Gibson, 1982). This line of thinking opens

up a critique about the assumptions AI made about the nature of human cognition – assumptions that, unsurprisingly, were based in enlightenment rationalism and philosophy of mind – that bankrupt pursuit.

The Embodied Mind, by Varela, Thompson and Rosch was representative of such critiques of internalist, mentalist, computationalist conceptions of human cognition that arose during the common-sense crisis, in parallel with the emergence (heh) of Artificial Life. This kind of work resonated with my home-baked ideas, and I was thus drawn into critical cognitive science, and discovered the range of approaches called enactive, embedded, extended, situated, and distributed. I was introduced to Gibson's ecological psychology and to Maturana and Varela's autopoietic biology, from which sprang enactive cognition. This work not only reinforced my distrust of the assumptions underlying AI and computational discourses, but provide leverage for a substantive critique of HCI. Perhaps most importantly, it provided me with a new vocabulary for thinking about the intelligences of the arts, that did not demand an assumed division of mind from body. For nearly a century, artists inquiring into the cognitive dimensions of their practice, if they were philosophically or scientifically inclined, had no language to use except a language of internalism - unless they were lucky enough to discover Ryle or Polanyi. Understanding embodied practices through the lens of internalism results in interminable and exhausting mental gymnastics. The paradigms of embodied, enactive, distributed and situated cognition, conversely, provide a more parsimonious ways of thinking about skilled practice qua cognition.

Alife and Alife Art

Through the 90s, I was informed and encouraged by an interdisciplinary field that for a while was topical in technocultural circles but is now all but forgotten, except among a rarefied group of enthusiasts - Artificial Life or Alife. Artificial Life sprang up as an alternative to the seemingly failed project of symbolic AI. Symbolic AI had set out to establish that something like human intelligence could be realized in computer programs. Symbolic AI was little more than the automation of symbolic logic. As a matter of principle (Putnam's functionalism) it ignored the embodied, enactive and phenomenological dimensions of human being, and rejected deployment of biological models that had been stock-in-trade for cybernetics, in which context neural networks were first developed (McCulloch and Pitts). In retrospect, it is little wonder that Good Old Fashioned AI (Haugeland) was found to be entirely incompetent when confronted with real-life (the 'common-sense' problem).

Fractals, complexity theory, dynamical systems theory, chaos theory and concepts bottom-up or emergent order informed the movement, along with an examination of the 'intelligences' of biological systems, particularly those of the social insects. As symbolic AI collapsed, many researchers turned to biological models that AI had staunchly rejected, and in the process reconsidered much work that had occurred under the banner of cybernetics – a field that was taboo in the AI community - the return of the repressed. Brooks proposed that AI should stand for Artificial Insects – this was part of an argument regarding the hubris of AI: how could one presume to make synthetic human intelligence when we could not make an artificial insect? As a result, neural networks, redubbed connectionism, gained renewed interest – and this work directly gave rise to contemporary machine learning (see below). It is therefore a curious historical irony that contemporary AI is based in technologies first generation AI rejected.

Is the 'new AI' old wine in new bottles?

The crisis of symbolic AI (GOFAI) was relieved in two ways, both ironic. Firstly, contemporary AI finds its roots in of neural nets - that symbolic AI rejected as biomorphic – rediscovered or resuscitated in the period of Artificial Life that replaced (symbolic) AI. The kind of AI that is currently popular, or regarded as most productive, is sometimes called *deep learning*, that utilises a technique called Multi Level Perceptrons (MLP). This is turn builds on *representation learning*, *machine learning* and *logistic regression*. Such work has deep roots in early cybernetics, specifically the neural networks of McCulloch and Pitts 1943, through the work of Hebb (1949) to Rosenblatt (1958) and the short-lived period of parallel computing called connectionism (Rumelhart's back propagation (1986) and the massively parallel, flat (as opposed to hierarchical) hardware architecture of the Connection Machine (Danny Hillis, 1987).

There are, then, ways in which the new AI is diametrically opposed to GOFAI, and there are ways in which it is same old same old. The new AI is not post-symbolic, it still resides in the milieu of the representational – although it is finally, for good or bad, coming to know the world directly via surveillance systems and the internet of things (IoT) – a point I will return to below.

Contemporary AI is essentially, statistical inference on vast amounts of collected data. One way to characterize the difference between GOFAI and contemporary AI is that in the old days (1980s) researchers deployed sophisticated algorithms on very limited datasets, while today, vast amounts of data are worked over by comparatively trivial statistical processes, and the results are better. But again, this improvement is not so much a result of better techniques, but of better hardware - vastly more storage, vastly more computational muscle, faster networking, and the existence of the internet itself. Google, the internet and server-farms have provided qualities of (organised) data undreamed-of in the 1980s, when the word terabyte, or even gigabyte, were science-fiction. During the heyday of symbolic AI, it was observed by iconoclastic computer engineers that all the gains of AI could be attributed to advances in processor speeds and memory capacity (cf Moores Law) – an opinion that was stifled in AI circles for understandable reasons. When Goodfellow, Bengio and Courville say of contemporary AI "*Deep learning has become more useful as the amount of training data has increased"* (Goodfellow Bengio and Courville 2016) one has a strong sense of déjà vu or "the more things change…". A good argument can be made then that the effectiveness of contemporary AI correlates with rapid processing over vast amounts of data. Crucially for our argument, the source of that 'training' data is *people*. People identifying faces in images, corelating things in the world with maps, and making value judgements – *good burgers here*. Vast surveillance systems are 'trained on' data disambiguated by people. People name the world, they make labels. From tagging photos in Facebook to reviews in Yelp, people provide the perceptual front-end for AI, turning the world into data.

I am not disputing the effectiveness of deep learning for the extraction of 'knowledge' from vast databases already online. The point is that the process of turning experience of the world into data, into symbols, demands a process that has largely been performed by humans, mostly voluntarily or unknowing (mechanical Turk, crowdsourcing, recaptcha). AI has no more common-sense now than it did then. The new AI sidesteps the common-sense problem because

the data it accesses is already symbolic. It simply draws upon the knowledge of the world as understood by people (common sense) and converted by them into data, stored in vast server farms. Humans are the informational microbiome for the internet organism, converting undigestible experiences into usable informational molecules. Contemporary AI cannot be understood as a standalone technology, but only as a cyborgian complex with human sensory membranes like oscillating cilia feeling out the world.

Goodfellow, Bengio and Courville also remark "…*computational units (that) become intelligent only via their interactions with each other is inspired by the brain*" (sic). (Goodfellow Bengio and Courville 2016). This statement begs the questions: which aspects of the brain? And: whose idea of the brain are we talking about? More perniciously, the problem here is the fundamental internalist fallacy: *as if intelligence happens exclusively in the brain*. As Francisco Varela put it, succinctly in 1999: *The mind cannot be separated from the entire organism...the organism as a meshwork of entirely co-determining elements makes it so that our minds are literally inseparable, not only from the external environment, but also from what Claude Bernard already called the milieu intérieur, the fact that we have not only a brain but an entire body.* (Varela, 1999).

It appears to have escaped the notice of many cognitive scientists and most AI researchers that the brain is an organ of a body, and the two are richly cross-connected, not simply by nerves, but by endocrine, immune and neurotransmitter modalities that are still only dimly understood. Indeed, to speak of 'two' only reinforces the problem. Suffice to say, as an example, that every red blood cell has receptors for 28 neurotransmitters. This implies that so-called neurotransmitters are in fact body-wide messenger molecules (as are hormones) that do different things in different locations. This stands to reason from an evolutionary standpoint – it is a general understanding of evolutionary biology that new functions occur by repurposing existing resources: as Elizabeth Bates memorably observed, *language is a new machine made of old parts*. (Bates and MacWhinney 1988). As those new activities accord advantages, they are selected for.

A background assumption of (neo)Cartesianism, is - to put it colloquially - that the brain grew a body to play with. I have observed previously that the opposite is true: the body grew a brain. early neural tissue (in jellyfish) coordinate locomotion, in sea urchins it regulates the movement of the gut, of digestion. Much later it coordinates sensory capacities like vision hearing and vibroception, and only in the 11th hour, evolutionarily speaking, does language occur. Neural tissue exists to help an animal get along in the world. Getting from there to logic and language (representation) is a long haul, but it is clear that the middle term is sensing or more accurately sensorimotor functions. An example being the elementary pattern recognition in light sensitive tissue, such as shadow detection in molluscs (described by von Uexküll) implying the presence of a mobile predator between it at the sun. Jump forward, evolutionarily, to the frog, and the link between its eye (light sensitive neural tissue) and the activation of its tongue. (Lettvin et al 1940)

Things or processes?

There is a nounishness to the AI approach, which is usually characterized in terms of logical *processes* operating on data 'objects'. There is a comfort in the (presumed) stability of objects. Contrarily, an understanding of the fundamental processual dynamism of things was central to

cybernetics, as it was central to a major force in postcognitivist theories of cognition – *enactivism*. Importantly, enactivism arose out of biological research, and was informed along the way by phenomenology and aspects of Buddhist philosophy – aspects of Buddhist philosophy that is experientially engaged, embodied and non-dualist, with its attention to processes of arising and passing away, of growth and decay.

Once one sets out on the path of representation – the idea that symbols adequately stand in for things – not only is dualism explicit, but it is assumed that representations are *sufficient*. This was the lesson that Borges illustrated in his (very) short 1946 story *On Exactitude in Science*, in which a nation obsessed with cartography created a 1:1scale map. (Borges 1946) A map is a representation, and any representation is a reduction, a simplification. It captures features that are deemed to be salient and leaves others out. A map cannot be a representation and simultaneously contain all the information in the source phenomenon of which it is a model. X may mark the spot, but the details of specific nature of that spot, the kind of soil, the varieties of vegetation, is lost. Nor is that X in the world – except in absurd comedy. The argument remains, as it did for critiques of GOFAI, that these systems have no connection with the phenomenal world. I know of no such system that can operate on phenomena themselves – as opposed to the representation of such.

Representation and the relationship of the biological to the world.

To return to the question we began with, this all turns on whether intelligence works on descriptions (symbols or representations) of the thing, or whether (biological) intelligence has access to the thing, the phenomenon - like the scent of narcissus – itself. It seems absurd to say "*I am smelling a representation of the scent of narcissus*". This is the fundamental fallacy upon which AI stumbles, it was the case in symbolic AI (GOFAI) and it is still the case in the neural network derived contemporary practices (which, it must be remembered, were violently opposed by symbolic AI partisans Minsky and Papert). To take a Foucauldian perspective: while opposed, these two approaches are both undergirded by a common commitment to the idea of *representation*. And why not? The notion of representation is fundamental to math, to logic, and to humanistic scholarship in general. Indeed, the work of the academy, across the board – is (largely) work upon representations. The clinical diagnostician in the hospital, the chemist at the lab bench, the trumpeter and the sculptor pouring molten metal – they experience a performative reality of the *dance of agency* (Pickering) among a network of agents, human and non-human. But such experiential irregularities are hammered out, or chased out, as texts are wrought - as actions in the changing world are fossilized as (textual) records. (Note here the metaphorisation of embodied experience – 'chasing' is a technique in metalsmithing, in which a punch, repeatedly hit with a hammer, moves, shapes or flattens ridges in sheetmetal. Hammering needs no explanation).

McCulloch and Pitts et al stayed close to neurochemical and neuroelectrical phenomena - and stalwartly resisted identifying these as symbols or representations. Theirs was the language of analog electronics, a language of dynamical, resonant phenomena in the temporal domain: there were '*signals*' at best. (The question of when a 'signal' become a 'symbol' gets us into deep epistemological water, and was part of the debate about whether Brook's *subsumption architecture* contained representations). Computer scientists were not usually so careful, or even

aware of the morass they were always already enmired in, being trained in symbolic traditions of mathematics and mathematical engineering (cf denaturalisation). For them, the building blocks of intelligence were the building blocks of 'reason' – symbols (abstract objects) operated upon by (logical) procedures. Hence the forced rhetorical back-formation inherent in the idea that brains use algorithms. Such an idea, if not nonsense, demands careful articulation with respect to particular specific cases with clear recognition that such explanations are just explanations, and are external to the throbbing, pulsing jellies they refer to. The entire sub-field of computational neuroscience may be guilty of this fundamental attribution error. It may be possible to metaphorically describe the operations of neural clusters in a language of such formal reasoning procedures, but it does not prove that such operations are actually occurring 'in the flesh'.

This is the kind of logical fallacy that Humberto Maturana emphatically expounded upon for most of his career. This position was developed from a combination of biologically materialist theory and second order cybernetics: "*Everything said is said by an observer*" is perhaps the most quoted of Maturana's aphorisms, for good reason. It points to both the separation between the observer and the observed and the idea that descriptive or explanatory labels applied by the observer to help 'explain' what they see is of the order of analogy only. It is not in or of the observed phenomenon. To take a classically autopoietic example, to ascribe to DNA the notion of information, and to ascribe to RNA the process of transcription of information is fallacious – there is no *information* in the cell. As Francisco Varela asserted, memorably, in biological systems, for instance the molecular processes inside a cell "*information, sensu stricto, does to exist*". Katherine Hayles has endorsed this, saying "*Autopoiesis also changes the explanation of what circulates through the system to make it work as a system. . . Indeed, one could say either that information does not exist in this paradigm or that it has sunk so deeply into the system as to become indistinguishable from the organizational properties defining the system as such*" (Hayles 1999).

In terms more pertinent to AI, Maturana emphasized the role of the observer and the map/territory problem: "*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 provide or create by themselves an understanding of the phenomena that an observer explains to him or herself through them*". (Maturana 2002). That is to say, the math is entirely in the observers' head, it is a story she tells herself about that which is observed, a mnemonic: *the thigh bone is connected to the knee bone….* The point that is here relevant to the current discussion is twofold. First, all of the categorisation of data in AI (GOFAI or the contemporary flavors) is of the order of '*things said by an observer'*: representations, mathematical formalisms. Second, none of the 'intelligence' that is gained about the data is intelligence about the world, because the data is always something '*said by an observer'*.

Conclusion

"What business has an artist in critiquing the rhetorics of computing and artificial intelligence?" Based on the forgoing, we might conclude that: inasmuch as art is an example of *non-artificial* intelligence of a materially engaged form; and that this kind of intelligence is inseparable from our human existence as biologically viable creatures: it is an ideal perspective from which to conduct such a critique. Cultures of the arts implicitly if not explicitly affirm conceptions of intelligence and cognition that are inherently holistic and embodied – as is regular human being-

in-the-world. Art practices offer, in object lessons, counterarguments to many of the notions rhetorically affirmed by computer culture and specifically AI. Attempting to realise coherent projects that combine these different cultures transpires to creates an excellent laboratory for working through the ramifications of this cultural collision.

I hope to have shown that computing technologies are not the kind of simple tool that can just be uncritically deployed. They are complex, and impose value systems that can undermine the goals of a particular project, or generally, of art. That is not to say that one cannot effectively use such technologies in artmaking, nor that new and interesting things cannot be achieved. My long history of technical development and problem solving in the service of what might be called *aesthetic* goals attests to that. But without the application of a *critical technical practice* (Agre), such work is vulnerable to, at best, naivete, and at worst a kind of theoretical undermining.

My argument is that a deep engagement both with the 'nuts and bolts' of the technology, and the ideas that inform it, can lead to a more complex kind of practice that does not simply deploy pre-made tools, but originates tools and interrogates the value-systems inherent in them. This may produce kinds of practices that cannot be identified as art, because they do not fit neatly into pre-existing and institutionalised genres – most first generation 'new media art' existed in such a space. It may be that such work is never recuperated as 'art' and into the annals of 'art history' - so be it. (I coined the term *Free Range Intellectual* to describe practices like mine. A friend wittily rejoindered "Its what's for dinner".) Such practices play in a larger field. A preoccupation with whether a thing is 'art' or worse, with a desire to pursue something identified by others as a career as an artist, seems trivial and beside the point.

References

Agre, Philip E. "Toward a Critical Technical Practice: Lessons Learned Trying to Reform AI." *Social Science, Technical Systems, and Cooperative Work: Beyond the Great Divide*, edited by Geoffrey Bowker, Susan Leigh Star, Les Gasser, and William Turner, 131–157. Mahwah, NJ: Erlbaum. 1997.

Bates, Elizabeth, and Brian MacWhinney. "What Is Functionalism?" *Papers and Reports on Child Language Development* 27:137–152. 1988.

Boal, Augusto. *Theatre of the Oppressed*. Translated by Charles A. McBride and Maria-Odilia Leal McBride. New York: Theatre Communications Group. 1985

Borges, J. L. On exactitude in science. P. 325, In, Jorge Luis Borges, *Collected Fictions* (Trans. Hurley, H.) Penguin Books. 1998.

Brooks, Rodney, A. Intelligence without Representation. *Artificial Intelligence 47*:139–159. 1991

Dewey, John. *Art as experience*. Minton, Balch & Company, New York. 1934

Gibson, James J. “Notes on Affordances.” *Reasons for Realism: Selected Essays of James J. Gibson*, edited by Edward Reed and Rebecca Jones, 401–418. Hillsdale, NJ: Lawrence Erlbaum. 1982

Goodfellow Ian, Yoshua Bengio and Aaron Courville, *Deep Learning*, MIT Press, 2016.

Hayles, N. Katherine. *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago: University of Chicago Press 1999.

Haugeland, John. *Having Thought* (47–61). Harvard University Press 1998.

Lettvin, Jerome Y., Humberto R. Maturana, Warren S. McCulloch, and Walter H. Pitts. “What the Frog’s Eye Tells the Frog’s Brain.” *Proceedings of the Institute of Radio Engineers* Nov 1940.

Maturana H. R. Everything said is said by an observer. Thompson W. I. (ed.) *Gaia: A way of knowing*. Lindisfarne Press, New York 1987 pp65–82.

Maturana, Humberto R. “Autopoiesis, Structural Coupling and Cognition: A History of These and Other Notions in the Biology of Cognition.” *Cybernetics and Human Knowing: A Journal of Second-Order Cybernetics, Autopoiesis and Cyber-Semiotics* 9 (3–4) 2002. 5–34.

McCulloch, Warren, and Walter Pitts. A logical calculus of the ideas immanent in nervous activity -*The bulletin of mathematical biophysics*. Springer 1943

Pickering, Andrew. *The Mangle of Practice: Time, Agency, and Science.* Chicago: University of Chicago Press *1995*

Penny, Simon. Simulation, Digitization, Interaction: The impact of computing on the arts. *Artlink V7 #3,4*. 1987 Art+Tech issue.

Penny, Simon. *Making Sense: Cognition, Computing, Art and Embodiment*. MIT Press 2017

Polanyi, Michael. *The Tacit Dimension.* Chicago: University of Chicago Press 1966

Roszak, Theodore*. The Cult of Information: The Folklore of Computers and the True Art of Thinking.* New York: Pantheon. *1986*

Ryle, Gilbert*. Knowing How and Knowing That: The Presidential Address. Proceedings of the Aristotelian Society 46:1–16 1946*

Varela, F. J., Thompson, E. and Rosch, E. *The Embodied Mind: Cognitive science and human experience*, Cambridge, MA & London: MIT Press. 1991 (p148)

Varela, F.J. 'Steps to a Science of Inter-being: Unfolding the Dharma Implicit in Modern Cognitive Science'. *The Psychology Of Awakening*. Watson, Batchelor and Claxton, (eds.) 1999 pp71-89.