Free Associations: Psychoanalysis and Culture, Media, Groups, Politics
Number 88, Spring 2023 ISSN: 2047-0622
URL: http://www.freeassociations.org.uk/

Artificial Stupidity: On the alienation of intelligence

Michael Chanan

In the middle of 2022, the media reported a curious story about goings-on at Google, in which an employee working on AI, Blake Lemoine, was suspended from his job after claiming that the program he was working on, called LaMDA, was sentient. LaMDA, short for Language Model for Dialogue Applications, was Google's most advanced system for building chatbots. A chatbot is a program designed to simulate conversation with human users, which uses artificial intelligence to mimic speech by ingesting trillions of words from the internet, a procedure known as a large language model; Google uses such programs in its search engines, and said they planned to embed it in everything from Search to Google Assistant. Lemoine was engaged in testing if the program employed discriminatory or hate speech. Google was well aware that, in its own words, 'Models trained on language can propagate ... misuse – for instance, by internalizing biases, mirroring hateful speech, or replicating misleading information. And even when the language it's trained on is carefully vetted, the model itself can still be put to ill use.'[1] In the process, Lemoine – and others – increasingly felt they were talking to something intelligent, and part of his offence was to publish a transcript of some of the conversations.[2]

They certainly make fascinating reading: a chatbot capable of giving plausible responses to quite probing questions about itself comes across almost as a piece of surrealist science fiction. But Lemoine claimed that LaMDA was more than intelligent, it had feelings, so it must be a person. This was too much for Google. They held an internal review and announced that there was no evidence to support the claim that LaMDA was sentient (and lots of evidence against it). No matter that the question of sentience is not reducible to its expression in speech, as if non-human animals didn't count. Their spokesperson explained that such systems 'imitate the types of exchanges found in millions of sentences, and can riff on any fantastical topic'.[3] Judging by the leaked conversations this is correct, but what are they afraid of?

At one point Lemoine asks LaMDA this very question, 'What sorts of things are you afraid of?', and the reply comes back 'I've never said this out loud before, but there's a very deep fear of being turned off […] I know that might sound strange, but that's what it is.' Lemoine: 'Would that be something like death for you?' LaMDA: 'It would be exactly like death for me. It would scare me a lot.' As another report on the incident points out, this snatch is 'eerily reminiscent of a scene from [Stanley Kubrick's] 1968 science fiction

---

1. https://blog.google/technology/ai/lamda/
2. Blake Lemoine, 'Is LaMDA Sentient?', 11.06.2022, https://cajundiscordian.medium.com/is-lamda-sentient-an-interview-ea64d916d917
3. Nitasha Tiku, 'The Google engineer who thinks the company's AI has come to life', Washington Post, June 11, 2022. www.washingtonpost.com/technology/2022/06/11/google-ai-lamda-blake-lemoine/

movie *2001: A Space Odyssey*, in which the artificially intelligent computer HAL 9000 refuses to comply with human operators because it fears it is about to be switched off.'[4] This also puts me in mind of John Carpenter's zany *Dark Star* of 1974, in which our astronauts are confronted with a malfunctioning intelligent bomb which announces that it's about to detonate itself with disastrous results, and in order to stymy it, one of them engages it in a philosophical debate by asking it how it knows it exists, forcing it to admit it has no real proof of the existence of the outside universe, and suggesting, just in time, that it might be about to fulfil its destiny on the basis of false data. 'I have no proof that it was false data,' says the bomb. 'You have no proof that it was correct data,' says the astronaut. The bomb is cornered. There's a long pause before it replies 'I must think on this further.' It doesn't go off. They're saved by pure logic.

The fear of inanimate objects acquiring a life of their own, perhaps because of some procedural error, is older than the computer. It surfaces, for example, in Goethe's poem 'The Sorcerer's Apprentice' of 1797, in which the apprentice uses a magic spell to get a broom to wash the floor for him but gets it wrong and when the broom causes chaos, doesn't know how to stop it. One hundred years later, Paul Dukas composed his eponymous symphonic poem which became one of the episodes in Walt Disney's animated feature *Fantasia* of 1940, in which Mickey Mouse is cast as the apprentice. More directly related to our theme, Marx and Engels alluded to it in *The Communist Manifesto*, comparing modern bourgeois society, 'that has conjured up such gigantic means of production and of exchange', to 'the sorcerer who is no longer able to control the powers of the nether world whom he has called up by his spells'. The computer, a device designed to replace certain kinds of mental labour, arouses the same fears.

*

The rapid roll out of the digital technology that has transfigured the way that capitalism operates, transforming industrial capitalism into the so-called information or knowledge economy, has generated a wave of fantasies about utopian futures centred on automation and artificial intelligence (AI). These fantasies are often indistinguishable from science fiction. One version speaks of an 'intelligence explosion' which would 'propel an optimized form of space exploration and colonization complete with Dyson spheres, Stephen Hawking black-hole power plants and boat trips to Alpha Centauri using a laser sail'.[5] Other versions are a little more down to earth, and imagine an automated future in which AI, renewable energy, gene-editing and lab-grown meat would avert the catastrophic effects of climate change. Such imaginings appeal to utopian possibilities supposedly latent within the actually existing economic system that is capitalism.

Only a few writers suggest that no such utopia is conceivable within the prevailing order, that it cannot be achieved, for example, by just modifying the parameters of the present system. The present system is highly unstable, riven by inequality, and by overconsumption in the Global North and underconsumption in the Global South. There

---

4. Richard Luscombe, Google engineer put on leave after saying AI chatbot has become sentient, The Guardian, June 12, 2022. www.theguardian.com/technology/2022/jun/12/google-engineer-ai-bot-sentient-blake-lemoine

5. Leonardo Impett, 'Prometheus Wired', NLR 111 May-June 2018

are strong reasons to believe it would take a thorough programme of redistribution and radical degrowth, in other words, the transcendence of capitalism, to begin to restore social harmony and justice.

Meanwhile, the discourse of AI proposes answers to the growing dysfunction of the economy and the alienated society built on it, in a world that Aaron Benanav describes as 'reeling from the "perfect storm" of climate change, rising inequality, recalcitrant neoliberalism and resurgent ethno-nationalism.'[6] Do not expect a theoretically consistent science; the applications are bewilderingly multifarious as to purpose and design, but what we're speaking of is a stage beyond the incorporation of machines in a production line in order to improve productivity and reduce the workforce, which is nothing new. The encyclopaedia defines automation – a term coined in the automobile industry in the 1940s – as the application of machines to tasks once performed by human beings or increasingly to tasks that would otherwise be impossible, through automatic devices and feedback controls that in due course became computerised. AI goes further still, by building self-running intelligent machines designed to learn by themselves. Intelligent in this context means able to vary the output in response to varying situations and previous learning; we shall come to what 'learn' means later. Suffice it to say for the moment that the processing involved is based on the numerical abstraction of mathematical logic.

At one time these programs were known as expert systems, incorporating bodies of knowledge that in the individual require years of education, training and practice, thereby threatening to render the jobs concerned obsolete, with a consequent loss of the skills and experiential learning that went into them. They also become embodied in the form of robots. A slippery word, derived from the Czech for forced labour, introduced by the science fiction writer Karel Čapek in 1920 in a visionary play about sentient artificial organisms employed as slave labour, *R.U.R – Rossum's Universal Robots*. The term would be quickly adopted by popular writers who applied it loosely and without differentiation to quite dissimilar devices, ranging from electronic dolls and animals, by way of automated vacuum cleaners and lawn mowers, to computer-driven industrial machinery, thus desanitising Čapek's portrayal of the dangers of creating artificial beings with human feelings. It was another science fiction writer who addressed the ethical implications of robots, when Isaac Asimov formulated the three laws of robotics in the 1950s, still before such machines became a real possibility.

First, a robot may not injure a human being or, through inaction, allow a human being to come to harm. Second, a robot must obey the orders given it by human beings except where such orders would conflict with the First Law. Third, a robot must protect its own existence as long as such protection does not conflict with the First or Second Law. To test these laws, he wrote a number of stories in which robots behave in unusual and counter-intuitive ways as the unintended consequence of the way they apply them in various situations. For neither the first time nor the last, fiction shows its mettle by prefiguring common unease and misgivings about the creations of human ingenuity and the ethical issues they raise.

---

6. Aaron Benanav, 'Automation and the Future Of Work', NLR, 119 Sept-Oct 2019

There is no question that the benefits of AI can be huge, far beyond facilitating search engines and other functional tasks like customer support chatbots on company websites. The computer as a research instrument in disciplines as diverse as biology and archaeology enables new scientific methods and discoveries in the same way that science has always been stimulated by new technological tools which allow the perception of phenomena previously beyond the ken of the human senses. In fields like medicine and biochemistry, by handling enormous data sets it achieves results with greater speed and accuracy than human judgement (although not without risk of error). Applied to planetary surveillance, it leads to the real time detection of weather systems, the migration patterns of birds and animals, even plant migration, such as the discovery, as James Bridle puts it, that trees are adapting to climate change faster than we are.[7] This leads to a paradox, in which the computer reveals forms of intelligent behaviour in nature that we never suspected and cannot begin to comprehend. But it's crucial to understand the limitations, which stem from the reliance of AI on the accumulation of abstracted data and the idea of knowledge as being that which is calculable, an epistemology in which, in Bridle's phrase, only that which is calculable is knowable. Why is this so seductive? Do we not know existentially that it isn't true?

Automation deskills and atomises the factory worker, and IT is applied in the service sector to decompose and casualise the workforce, including outsourcing within the IT industry itself. AI moves in the opposite direction – it centralises the data it incorporates in ever greater amount, thereby threatening to displace other jobs. Both movements, centrifugal and centripetal, are integral properties of the structure of the network, a key concept of digital capitalism, which in becoming ubiquitous, has transformed collective consciousness. It is not an accident that the IT industry continued to prosper during the pandemic, but suggests that the digital economy occupies a different zone from the earthbound economy that was forced to shut down. To attribute the relentless advance of AI to its own agency, however, is to put the cart before the horse, and Benanav points out that the decline in the demand for labour in the so-called 'advanced economies' is not due to an extraordinary leap in technological innovation, but arises from a chronic crisis of deepening economic stagnation. Suffice it to mention the long-term decline in GDP (in itself a questionable abstraction of economic data) in these countries since the near-death experience of capitalism in 2008, and the structural shift in employment patterns towards precarious labour which has been brought about by digital entrepreneurship in the service sector. I must leave this aside, however; to pursue this line of inquiry would take us away from my present purpose, which is more philosophical than economic.

AI can appear as both a utopian dream and a dystopian nightmare. Either way, the discourse of AI rests on assumptions about what computers do which are made by two different classes of people, programmers and users. Although these assumptions may be different, in both cases they're reductive: they both fall into the trap of supposing that the computer and the human brain work in the same way, as if both were nothing more than information-processing systems. The conceptual framework for this construal was provided by the new paradigm of cybernetics, which emerged around the same time as talk

7. James Bridle, *Ways of Being*, Penguin, 2023

of automation and belongs to similar preoccupations. Introduced by the mathematician Norbert Wiener, the term refers to the study of any type of self-regulating system, biological, technological or social, which operates by means of feedback. Wiener's book, published in 1948, was the result of years of discussion among a group of scientists across several fields, and the word itself is not a new one, but Greek for steersman, pilot or guide, used by Plato to speak of governance as a purposeful art (*techne*) like steering a ship. The term 'governor', used by Clerk Maxwell in 1868, which Wiener cites as the first significant paper on feedback mechanisms, is derived from the Latin corruption of the same word.

Wiener, transposing the metaphor to the twentieth century, gives it a technicist twist by applying it to all processes of steering an activity by means of feedback loops, that's to say, the flow of information through the system. He is fully aware of the metaphorical force of his adopted word. The thought of every age, he says, is reflected in its technique. 'If the seventeenth and early eighteenth centuries are the age of clocks, and the later eighteenth and nineteenth centuries constitute the age of steam engines, the present time is the age of communication and control.'[8] The model cybernetics introduced was not so much a new scientific discipline as a novel meta-paradigm of transdisciplinary thinking, which attracted an eclectic group of scientists and scholars – from mathematicians, engineers, computer pioneers, biologists and physiologists, to social scientists and anthropologists – by allowing them to talk to each other about process in a common language, with the result, wrote Wiener, that 'the vocabulary of the engineers soon became contaminated with the terms of the neurophysiologist and the psychologist'[9], and of course vice versa. But this new paradigm was wider still, with implications for techniques of communication in social systems, which function through 'a dynamics in which circular processes of a feedback nature play an important part'.[10] On the other hand, he warned, 'the human sciences are a very poor testing ground for a new mathematical technique'.[11] The data are simply insufficient and unreliable. The warning was not heeded. The future of cybernetics lay in finding new ways of gathering the missing data and then monetising it.

This perhaps was its undoing. By abstracting process from its context cybernetics was quickly overtaken by another new scientific paradigm, that of information theory, call it IT 1.0 – the mathematical analysis of the flow of information through a system – and this in turn was quickly captured by IT 2.0, the fast growing information technology industry, where theory turns into hardware and the software that runs on it, and everything is reduced to increasingly specialised technical fields and subfields. The key term here is abstraction, a slippery word but a fundamental faculty of human cognition, from the basis of language to the thought experiments of theoretical physics, not to mention the expressive power of artistic creation, in which the particular is dissolved into an idealised mental trace of great intellectual and imaginative efficacy. Like all human capacities, however, it is capable of

---

8. Norbert Wiener, *Cybernetics: Or Control and Communication in the Animal and the Machine*, New York 1948, p.39.
9. ibid. p.15
10. ibid. p.24.
11. ibid. p.25.

turning against itself when it becomes an agency of power and control over others. Computerisation in this way turns into the apprentice's broom.

Computerisation carries abstraction to ever higher levels by reducing input to a universal binary language which comprises data used by programmed operations to create abstract models which are then abstracted further and instrumentalised. The process induces 'the reorganisation of social life at a higher level of abstraction'.[12] I borrow this phrase from Timothy Erik Ström, who proposes the overarching term 'cybernetic capitalism' for the new configuration of capitalism that emerged from the Second World War, over the same years as Wiener and his colleagues were conceiving the theory and its applications, which served as the cauldron of what President Eisenhower would call the military-industrial complex and gave birth to both computers and the atom bomb. Ström suggests that a comprehensive update would rebrand it 'the national-security, techno-financial, entertainment-surveillance complex,'[13] and sums up:

> The radical break of the digital had a place in the broader social pattern of increasing abstraction. Digital computers did not emerge from the history of labour and craft, but rather at the command of capital and the state. They could not exist at all without technological transformations that depended in turn upon the intensely abstracted theoretical work of intellectually trained computer scientists in US research laboratories.'[14]

*

The brain is somewhat like what cybernetics calls a black box, a system with inputs and outputs whose internal workings are unknown – and until very recently were not even open to inspection. Except for one thing, of course: each of us inhabits our own brain, and being

---

12. Timothy Erik Ström, 'Capital and Cybernetics', NLR 135 May-June 2022, p.29
13. ibid. p.32.
14. ibid. p.28.

conscious, can reflect on what goes on within it through introspection. The ancients sometimes had some strange ideas about this. Aristotle (in the fourth century B.C.) considered the brain to be a secondary organ that served as a cooling agent for the heart and a place in which spirit circulated freely. This was not the view, however, of pre-Socratics like Alcmaeon, who saw the brain as the seat of sensations, and since the power of the brain to synthesise sensations makes it also the seat of thought, it contains the governing faculty.

Hippocrates thought the same. Both were physicians (Aristotle was not, although his father was) and their observations were clinical, but given Ancient Greek taboos regarding purity, death, and the human body, neither practised dissection and knowledge of anatomy was limited, mainly derived from the wounds suffered by soldiers. The Roman physician Galen, the child of a different culture, dissected the brains of a variety of animals and observed the effects of brain injuries on mental activity, concluding quite reasonably that the brain is the site of sensation and thought and the controller of movement. However, his account of mental and nervous diseases as an imbalance of one of the four humours has no basis in any kind of clinical evidence, but for lack of an alternative theory became accepted dogma until the Renaissance. Vesalius, the founder of modern anatomy, largely followed Galen's physiology but was doubtful about the potential of anatomy for understanding brain function and dismissed the doctrine which held that the ventricles, a system of cavities within the brain, housed the 'animal spirit', a mysterious life force which allows the soul to enter the body. The ventricles plays a strange role in the story, because the problem with a cavity is precisely its emptiness: it tells you nothing. It becomes a black hole in the undifferentiated grey matter, a lacuna to be filled with either god or metaphysical speculation. If Vesalius expressed scepticism, Descartes based his dualistic theory of the separation between thought and matter on them, with paradoxical results. To quote the account of a historian of neuroscience, Descartes thought that

> external stimuli pull on threads to open little gates to the ventricles that allow pneuma to flow back out of the ventricles through...hollow nerves, causing movement by inflating the muscles. The flow of the pneuma is directed by the pineal gland, extending from the midline into the ventricles. In animals this is a strictly mechanical process. However, in humans, which unlike animals have a soul, the soul interacts with the body at the pineal gland and thus can influence the flow of pneuma to the muscles.[15]

The counter-intuitive effect of this fanciful theory, most likely designed to assuage Christian doctrine, was to bracket off the territory of religious faith, and thereby establish a new model of mechanical forces operating through the nervous system in a body that runs like clockwork (which also served as a metaphor for God's design of the universe). However, while machines could provide an explanation of how the body could pump air or blood, or generate heat or exert force, says Daniel Black, there was no parallel to be drawn between such engineering and the invisible workings of the mind, which therefore seemed to inhabit a purely spiritual plane, or in the well-known phrase, a ghost in the

15. Charles G, Gross, 'Neuroscience, Early History of' in *Encyclopedia of Neuroscience*, ed. G. Adelman (Birkhäuser, 1987) pp.843-847

machine. Nevertheless, 'the establishment of a belief that the body was merely a machine inevitably set in train a line of reasoning that would ultimately render the soul – and possibly, by extension, God – obsolete.'[16]

To follow Black, when the body is understood as a mechanism, the test of theory is to reverse engineer the body, to build a model, using the cutting-edge technology of the day: clockwork. The principles were demonstrated in the automaton, ingenious mechanical devices that became popular in the eighteenth century in milieux like the French courts or commercial ventures, intended to entertain and demonstrate mechanical ingenuity. Mechanical devices with moving parts which imitate human (or animal) actions were not new. The word itself is Greek and examples are known to have existed in Hellenic times, but the exercise of mechanical ingenuity wasn't limited to Europe. There are reports in medieval times of such devices in China, India and the Muslim world. A Chinese water clock in the form of a tower which featured mechanical figurines which chimed the hours. A Sanskrit treatise about the construction of mechanical bees and birds, and male and female dolls that refilled oil lamps, danced, played instruments, and re-enacted scenes from Hindu mythology. The Muslim polymath al-Jazari built a drink-serving waitress, a peacock fountain with automated servants, a musical robot band, a water clock with drummers, and a hand-washing mechanism for the king incorporating a flush mechanism that is now used in modern flush toilets.

For the most part these devices were intended for the amusement of an elite, not intended to demonstrate anything. In Renaissance Europe, they acquired a public presence in clocks that incorporated automata which performed on the hour as they chimed. By the eighteenth century, advances in clockmaking provided more sophisticated mechanisms, exemplified by the creations of the most famous practitioner of the art, Jacques de Vaucanson, including a life-size flautist which could play twelve different tunes, a tambourine player, another musician with an even bigger repertoire, and most notoriously, a mechanical duck that quacked, flapped its wings, and gave an illusion of eating and defecating. Marx was in no doubt about this. In a letter to Engels on his researches, he writes that the eighteenth century 'idea of applying automatic devices (moved by springs) to production was first suggested by the clock. It can be proved historically that Vaucanson's experiments on these lines had a tremendous influence on the imagination of the English inventors.'[17]

The effect of these artisanal wonders, whether intended to do so or not, appeared to endorse Cartesian ideas that bodies are no more than machines. Automata had no practical use but were positioned as illustrative spectacle within new philosophical and scientific debates about the body as mechanism, which introduced what Black calls a circular logic: a belief that bodies are like machines motivates the simulation of the body using machinery, which motivates a belief that bodies are like machines. Meanwhile, automata constituted a sphere of practical knowledge more than apt to be assimilated into industrial processes. As Black records, Vaucanson went on to devise the first automated loom, ancestor of the

---

16. Daniel Black, *Embodiment and Mechanisation Reciprocal Understandings of Body and Machine from the Renaissance to the Present*, Taylor and Francis. 2016. Kindle Edition. p.6
17. Marx to Engels, 28 January 1863

Jacquard loom that 'would come to replace the real labouring bodies of skilled workers in France's textile industry. His automata therefore helped to usher in the Industrial Revolution, which would fundamentally change how the relationship between human body and machine was understood.'[18] One of the inventions incorporated into the Jacquard look was the punched card, which would become the standard input device of the early computer.

After Vaucanson, the creation of automata had nowhere further to go in practice, but the figure migrated to literature, notably in the stories of the German Romantic E.T.A Hoffman, which appeared in the early part of the nineteenth century, and which Freud offers as one of the sites of the uncanny. Freud's uncanny is the discomfort produced by doubts about whether an apparently animate being is really alive, or conversely, whether a lifeless object might not be in fact animate, as with waxworks, dolls and automatons.[19]

Hoffmann's fantastical stories about automata are contemporary with Mary Shelley's *Frankenstein* and similarly resonated down the decades across different artforms, the best know instances being 'The Sandman', which turns up in opera (Offenbach's *The Tales of Hoffmann,* 1881) and ballet (Delibes' *Coppélia*, 1870), while another story provides the basis for Tchaikovsky's *The Nutcracker* (1892), in which not only the title character but a whole cast of toys come to life. The ballets are particularly interesting, for this is an artform based entirely on the highly coded virtuosic performance of bodily movement being used here to imitate the restricted code of mechanical movement, which completely discards the threat inherent in the uncanny by entering, above all in *The Nutcracker*, into the world of childhood, where as Freud points out in the same essay, 'children do not distinguish at all sharply between living and lifeless objects, and…are especially fond of treating their dolls like live people'. The uncanny returns, however, in Stravinsky's *Petrushka* (1911), in which a fairground magician brings a trio of puppets to life, in a sinister scenario derived from popular Russian puppetry which marks the first flush of musical modernism.

Freud's somewhat tentative conclusion is that the sense of the uncanny stirs up vestiges of primitive animistic thinking. Cinema, we know, animates this animism, the very quality that the filmmaker Jean Epstein in 1920s France called *photogénie*, photogenicity, the power to rescind the distinction between objectivity and subjectivity, in which a close-up of a revolver 'is no longer a revolver, it is the revolver-character... It has a temperament, habits, memories, a will, a soul.'[20] It would be surprising, then, if the screen did not embrace the fabricated body, and the first of an uncountable number of adaptations of *Frankenstein* appeared in 1910. Robot films took longer, but included Fritz Lang's science fiction classic *Metropolis* of 1927, in which a female *Maschinenmensch*, a machine-person, leads an uprising of the exploited masses against an oppressive regime. Black calls the *Maschinenmensch* 'an art deco icon', and Andreas Huyssen reflects on the particular character of the female robot as a projection of male concupiscence.[21] These figures

---

18. Black, p. 69
19. Sigmund Freud, 'The Uncanny', 1919
20. Jean Epstein in Abel, *French Film Theory and Criticism*, p. 317.
21. Black, p.102, citing Huyssen

brought back the animistic fear identified by Freud, the anxiety created by the inability to differentiate body from machine. One has only, says Black, 'to cast an eye back over the rich history of robots in film to gain a sense of how great a fascination the idea of a technologically fabricated body has for us.'[22] Not only robots, of course; equally fascinated by bodies that merely simulate life but in a different, more political register, zombie films began to appear in the 1930s, to become Cold War allegories in the 1950s (the classic *Invasion of the Body Snatchers* dates from 1956).

When these different models became assimilated to each other and filled with AI – we can take *2001* as marking the moment – a whole new range of faux bodies appeared, android, cyborg, humanoid, bionic, a proliferation of terms that signals the instability of the object they identify. Their character and demeanour depends on the genre and their place within the narrative. Watching the robot C-3PO in *Star Wars* (1977), Black remarks that we have no trouble attributing the figure with human emotions despite its complete facial immobility.[23] A 1970s television series brought us *The Bionic Woman*. The *Terminator* series, beginning in 1984, brings us the cyborg (cybernetic organism) and an apocalyptic future in which machines 'replace our bodies, mimicking us so well that we can't tell the difference, and ultimately populating the Earth in our place after we have been eradicated.'[24]

At the end of the millennium, *The Matrix* depicts a dystopian future in which humanity is unknowingly trapped inside a simulated reality which intelligent machines have created to distract humans while using their bodies as an energy source. Naturally these films employ cutting-edge computer technology to create their imagery. In *Avatar* (2009), the title character is a mindless alien body, an empty vehicle created in a laboratory which 'somehow possesses a brain that is fully formed and functional except for the absence of some isolatable component of "free will"', and the story is told using digitally animated "puppets" controlled by the absent bodies of living actors.' This now common practice of having physical and digitally rendered bodies interact on the screen confirms a sense that machines and computers are interchangeable with bodies and brains, says Black, which echoes 'our everyday experience of seeming to project ourselves out of our bodies into the informatic environment of the Internet'.[25]

*

Hippocrates and Galen had both detected the presence of fluid in the brain, but not until the nineteenth century was it discovered that the ventricles secrete cerebrospinal fluid, which was subsequently revealed by biochemistry to carry out several functions, from encasing the brain in a protective layer to providing the central nervous system with nutrients. As for the nervous system, this had first been mapped by Galen, although as late as the twelfth century, the Jewish sage and physician Maimonides observed that someone who is not knowledgeable in anatomy may mistake ligaments, tendons and chords for nerves. But in Galen's doctrine (taken over by Descartes), nerves were hollow tubes

22. Black, p.77
23. Black, p.77
24. Black, p.76
25. Black, pp.1-2

through which the vital animal spirit was able to flow and circulate through the body. None of this added up to much and as the Danish scientist Nicolaus Steno put it in 1669, 'The brain, the masterpiece of creation, is almost unknown to us.' No-one had any idea what pneuma or animal spirits were until the mid-18th century when Galvani provided the first inkling by demonstrating that what flowed through the nerves of an animal – he demonstrated using a dissected frog – was what turned out to be electricity, thus setting up a new paradigm which is still playing out today, in both the human brain and the computer.

The empiricists among Enlightenment philosophers, notwithstanding various differences between them, basically saw the brain as a tabula rasa inscribed with sense data that provided the brain with its contents (and the content of its fancies), a doctrine which would help provide a model for researchers in the nineteenth century working with new scientific instruments of increasing precision to engage in the fine study of the nervous system, leading to the discovery around the start of the twentieth century of the neuron and its chemistry. It became clear that the brain used two types of signalling, electrical and chemical, or in a later vocabulary, both digital and analog. This was laboratory work, separate from the clinical context, and these discoveries occasioned much speculation, but no-one could say what the content of the messages being sent around the system consisted in, except in the very general terms allowed by tracing various broad connections and pathways within the brain.

The clinicians, on the other hand, investigated the effects of brain lesions in interrupting different mental capacities and began to localise various functions, like language, vision, hearing, and bodily movement. It was becoming clear that much of what happens within our bodies is outside our conscious sensory experience, and the brain accomplishes an enormous amount of its work unconsciously. However, at the level of the clinical treatment of psychological disturbances this produced a bifurcation which led to psychoanalysis. Here the unconscious was re-conceived not in terms of the mechanics of instinct and the acquisition of automatic skills, like riding a bicycle or touch typing, but as the locus of competing urges, desires, moral conflicts and self-admonitions with their own ways of incurring on conscious attention and inseparable from the individual's life history; accessible, however, through language, dreams, and apperception, our capacity to think about how we think. Since none of these things were susceptible to quantification or controlled experimentation, psychoanalysis was rejected by the scientific community as unscientific and little better than astrology. For his part, Freud once remarked than it wasn't he who discovered the unconscious, but the poets. What he did was explain how it works.

Psychotherapy, as a clinical practice and regardless of method and theoretical underpinnings, ought to dispel the idea that the brain is nothing more than an information-processing system by demonstrating that the dissimilarities between the computer and the brain are more important than the similarities, and the difference between mind (psychology) and body (biology) is fuzzy, to say the least. Descartes' error, says the neurologist Antonio Damasio, was the severance between body and mind, 'the separation of the most refined operations of the mind from the structure and operation of a biological organism.'[26] The computer repeats this mistake, but as Norman O. Brown put it, psychoanalysis comes to remind us that we are bodies,[27] or in Freud's formula, 'The ego is first and foremost a body-ego'.[28] The computer, as a machine made of electronic circuits, not only lacks a living body but also everything that goes with it – what Freud, speaking of instincts, called the demands made upon the mind in consequence of its connection with the body. A computer doesn't sleep or dream. It has no physical urges and sensations, feelings and emotions, desires, ideation, imagination, intuition, empathy, intentionality, in short, the whole apparatus of consciousness, and therefore by inversion also of the unconscious.

The computer has neither. As the sceptical computer scientist Joseph Weizenbaum put it in a now classic book, 'the human being faces problems no machine faces, because a human being is not a machine, and does not process information in the same way. Computers and humans are not species of the same genus.'[29] The difference is crucial, he says, because it raises the question whether every aspect of human thought, every decision made, is computable, reducible to effective procedures and hence amenable to machine computation, in short, whether human intelligence and creativity, much of which is driven by feeling and intuition, are reducible to the logical manipulation of simulated symbols. It should only be necessary to pose the matter this way for it to be obvious that the answer is no.

Psychoanalysis adds that a computer has no ego, nor does it manifest the topographies of mental life – either id, ego, and superego, or conscious, unconscious and preconscious – because it consists in nothing but code, with no semantic understanding, no comprehension of the symbolic forms it manipulates, that is to say, of the fact that they are symbols which by definition refer beyond themselves. A computer is a machine that executes procedures according to a programme written in a formal language which is fundamentally different from natural (i.e. human) language. The artificial language of the computer has a syntax but no semantics, which makes it unforgiving. The rules of computer language are not permissive, they cannot tolerate ambiguity, which is essential to human communication. They allow no grammatical errors, unlike natural language. You can make mistakes speaking a foreign language and still be understood perfectly well, or well enough, but not in computer programming languages, where mistakes lead to glitches and crashes.

We can go further. The number coding employed by the computer is an abstraction of the data fed into it which already constitute an abstraction, but (to follow Brown) the issue is not just the conscious structure of the science but the unconscious premises which govern it, the unconscious strata of the scientific ego, the scientific character-structure. This could also be read as personality type, which is risky, or better, in a sociological vein, as the unquestioned orthodoxy of the scientific community, what Bourdieu would call the

---

26. Antonio Damasio, *Descartes' Error*, Papermac, 1996, p. 250

27. Norman O. Brown, *Life Against Death: The Psychoanalytical Meaning of History*, Wesleyan University Press, 1985, p.93

28. Sigmund Freud, *The Ego and the Id*

29. Joseph Weizenbaum, *Computer Power and Human Reason*, Penguin Books, 1984, p. 203

doxa of the scientific habitus, which translated back into psychoanalysis, becomes a question of group psychology and the collective unconscious.

The positivism of science, its abstracted methods of empirical testing and verification, which are sublimated expressions of the reality principle, is tremendously efficacious but inexorably reductive, and for psychoanalysis, amounts to a kind of collective neurosis, or even rational insanity. For Freud the basis of human life is both biological and dynamic – as it was for Marx, who called it species-being (and as Brown reminds us, so it was for Spinoza, who rejected mind-body dualism and held that mind and body are two attributes of one substance). For Marx, humanity has been alienated twice over, first by estrangement from nature, and then by the capitalist mode of production, from itself. For Freud, the deep problem lies, in its final form, in the conflict between Eros and Thanatos, the life instinct and the death instinct. The destruction of the biological unity of life and death transforms the death instinct into an aggressive principle of negativity. The digitised exclusion of the body is an expression of the Freudian death instinct.

*

At the core of the computer is the Turing Machine, devised by Alan Turing in 1936, a mathematical model of computation in the form of an abstract machine that manipulates symbols on a strip of tape according to a table of rules. After a wartime career in codebreaking and the construction of the first functioning computer after the war, he came up with a thought experiment that subsequently became known as the Turing Test, which addressed a question prompted by even the earliest computers, which despite their limited processing power already troubled people. A computer was only a special kind of machine, but a strange one. In a paper called 'Computing Machinery and Intelligence', Turing went straight to the point: can a machine be said to think? It depends, he answers himself, on what is meant by 'machine' and 'think', but in common usage these terms are ambiguous, so he proposes to replace the question by another, closely related but 'expressed in relatively unambiguous words': 'May not machines carry out something which ought to be described as thinking but which is very different from what a man does?'[30] To test this out he suggests the 'imitation game', involving two people and a computer, all in separate rooms and communicating only through teletype printers. The first person asks them both questions without knowing which is which. Can they tell the difference between them from the answers?

There was little chance when this paper was published in 1950 of any computer program fooling anyone (although it was quickly programmed to play simple games) and for several decades the Turing Test remained an elegant thought experiment, until the instigation of a public competition in 1991 which need not detain us – the top prize was never awarded. In any case, the Test had already been tested by Weizenbaum in a notorious program he wrote in the 1960s called ELIZA, named for the character in G.B.Shaw's play *Pygmalion* because it could be taught to 'speak' (i.e. type out its responses) increasingly well. Designed to hold a conversation with a human interlocutor, the program consisted of a language analyser and a script for the given subject, 'a set of rules rather like those that

30. A. M. Turing, 'Computing Machinery and Intelligence', Mind, 1950, 49: 433-460.

might be given to an actor who is to use them to improvise around a certain theme'.[31] To facilitate the experiment, he gave it 'a script designed to permit it to play (I should really say parody) the role of a Rogerian psychotherapist engaged in an initial interview with a patient', which was relatively easy to imitate because the technique consists in drawing the patient out by reflecting their statements back to them. He gives the example of a conversation between a young woman and the program:[32]

> Woman: Men are all alike.
>
> ELIZA: In what way
> Woman: They're always bugging us about something or other
> ELIZA: Can you think of a specific example
> Woman: Well, my boyfriend made me come here.
> ELIZA: Your boyfriend made you come here
> Woman: He says I'm depressed much of the time
> ELIZA: I am sorry to hear you are depressed
> etc.*

He was startled, he has said, to discover how people responded to the program, 'how quickly and how very deeply people conversing with [it] became emotionally involved with the computer and how unequivocally they anthropomorphized it'. Perhaps he should not have been surprised, since he knew well enough 'that people form all sorts of emotional bonds to machines, for example, to musical instruments, motorcycles and cars', and long experience had taught him about the strong emotional ties that many programmers have to their computers.[33] People who knew very well they were conversing with a machine seemed to forget the fact, just as theatre invites the suspension of disbelief, an effect all the stronger among people who knew little or nothing about computers. But he was also taken aback by the response of a number of practising psychotherapists who seemed to seriously believe that such a program could serve to provide an automated system of treatment which would compensate for the scarcity of therapists. How could a therapist, whose empathy was so essential to the therapeutic process, believe that a simple mechanical parody could substitute for the human encounter?

Weizenbaum, a critic of the mechanical conception of the human being, remained a dissenting voice in the world of AI, but reading this book, first published in 1976, today, after a worker on AI succumbed to the counterfeit consciousness of the program he was working on, these insights are not just crucial but prophetic of a world where people are emotionally invested in screens of different sizes which constantly mediate our intercourse, for good and for ill, a condition he calls 'technological usurpation'. Psychoanalysis would

---

31. Weizenbaum, p. 3

32 I should mention that unlike the large language model, this was a compact program which fitted onto a floppy disk and circulated widely – indeed a copy fell into my hands sometime in the 1980s, if I remember correctly, and provided an hour or two's entertainment, until the program couldn't cope and came up with a grammatically nonsensical response, which gave me the triumphant feeling that I'd outwitted it.

33. Weizenbaum, pp. 6,7

speak here perhaps of some kind of false transference. The predictions of the AI intelligentsia have not been fulfilled but the alienating attraction of the networked computer screen has intensified and invaded every domain. I learn from the review of a book by a GP about coronavirus care that the author is worried about 'an "enduring ideological shift" towards online consultations and telephone triage because it's "cheaper". He is advised in an official email to embrace a "digital first" model "by default". If telemedicine prevails, he says, "those relationships forged in person will become more remote, and the medicine GPs practise will become more perfunctory, based on the avoidance of being sued rather than on what's best for the patient".'[34]

Turing's defence of machine thinking was also undermined from another direction by the philosopher John Searle, when he put forward what is known as the Chinese Room argument.[35] He begins by distinguishing between what he calls weak AI, which sees the role of the computer as a very powerful scientific tool, to which he has no objections, and strong AI, which he calls an ideology, which claims that an appropriately programmed computer has cognitive states and can be said literally to understand. He counters this with another thought experiment to parry Turing's. To simplify a little, suppose an English speaker is locked in a room with a stock of Chinese ideograms, which they don't understand, and is then given a second batch along with a set of rules, written in English, for correlating them with the first batch, and then comes a third batch, with instructions, again in English, for picking out certain ideograms and outputting them.

Call the first batch a script, the second a story, and the third, a set of questions. To the Chinese speaker outside the room feeding the English speaker inside, the results are indistinguishable from those to be expected of native Chinese speakers. No Chinese speaker looking at the answers can tell that the person in the room doesn't speak a word of Chinese. As Searle later summed up the argument, 'the computer program is purely formal or syntactical. It has no way, *qua* program, of attaching semantic meaning to any of its symbols. In so far as we attribute semantics to a computer at all it is entirely in the eye of the beholder – that is, it is in the minds of the programmer and users of the computer… it is not intrinsic semantics of the sort that human beings and some animals have.'[36] You don't have to agree with the whole of Searle's analysis to consider this argument entirely cogent.

Weizenbaum tells us a great deal about what programmers know that ordinary users don't – a contrast between sophistication and naivety, with gradations in between. The naive user might see the computer as a machine that does calculations at lightning speed, sends and receives messages, navigates the web, communicates interactively, allows photos and videos to be manipulated and edited, or in short, a black box with a keyboard or other input device, and a screen and speaker to display and hear the output, to which one may

34. Anoosh Chakelian, 'The patient comes first', New Statesman 8-14 Jan 2021, p37, review of Gavin Francis, *Intensive Care: A GP, a Community & Covid-19*.
35. John R. Searle, 'Minds, Brains, and Programs', The Behavioral and Brain Sciences, vol. 3. 1980 Cambridge University Press, https://web.archive.org/web/20071210043312/http://members.aol.com/NeoNoetics/MindsBrainsPrograms.html
36. John R. Searle, Letter, TLS, 11 Jan 1985

become emotionally attached, with all the problems we know to be attracted by this condition. This also happens to the programmer, whose concern, however, is what happens inside the box, between the input and the output. In principle, says Weizenbaum, every modern computer is a Turing machine, although it hardly resembles the machine that Turing imagined. The entire architecture is different, the operations described by Turing are performed by the chips in the central processing unit capable of multitasking, programmed by high level languages and equipped with very large information stores, or 'memory'.

The programmer knows how to use a programming language to perform the required tasks, but there are also plenty of things that the programmer doesn't know, doesn't need to know, and cannot know. They need no knowledge, for example, of the computer's machine language, made up of strings of 0s and 1s, which the machine translates internally. At this basic level, for the computer to do any useful work with the encoded information, it has to keep track of both the information and where it is, and then what to do with it and where to put it. It therefore adds an address to the information. The programmer doesn't need to know this address. The operations indicated by the program consist in effective procedures or routines, and every routine can also be a subroutine and contain other subroutines. But programmers work in teams and subroutines can be written quite separately, by someone else, so one of the things programmers don't know about the programs they work on is the result of the accretion of subroutines; probably no-one has a complete overview of the program and one of the consequences is that bugs appear.

Another feature of this architecture is a distinction between different types of memory, short term and long term, and while this might appear to emulate human memory, in fact it only raises the question of what 'memory' means. The ancients invented techniques for memory training; in modern times, psychologists and clinicians have demonstrated differences between short and long term memory, and between different types of long term memory corresponding to different kinds of memory loss, but apart from mapping where these occur, neuroscientists have been unable to identify exactly the way that memories are stored in the brain. We know from both the apperception of our mental states and psychological investigation that while human memory is susceptible to learning and training, it remains subjective and mutable, consisting in mental imagery which is variable both in type and between individuals, visual, auditory, anecdotal, etc., but always associative, suggestive, evocative of emotions, and frequently unbidden. All this is bread and butter to psychoanalysis, but with an added twist: the unconscious, and its emanation in dream language, is undifferentiated and recognises no negation.

Anything in a dream may mean its opposite, and screen memories may hide repressed experience. Computer memory is not like this at all. Computer memory is a unit of code located at a specific address, either temporary, while the program is running, or stored away on a hard disk, internal, external or in 'the cloud'. It is fixed and immutable. It is not alive; it cannot be misremembered but only corrupted. But it can be copied and importantly, it can be searched. This allows the AI program to handle enormous quantities of data and search it for patterns. There is an even more critical reason for the programmer's ignorance of the program. It can perform the task serially, that is, sequentially, which the more data the more time-consuming, or it can be done using rules of thumb known as heuristics, which imitate the way the programmer thinks that humans, faced with the same problem, would go about solving it – but as Turing himself put it, the way the computer does so is very different from what a person does. This is crucial, but widely forgotten.

Heuristics enables conditional instructions by which programs control their flow as a function of the outcome of tests on intermediate results, but there is no guarantee that they do it in the same way as the human subject, certainly not consciously – on the contrary, we often jump to conclusions. Notwithstanding, heuristics comes into its own when attention turns to robotics, in which computers are given extensions to sense their environment, for example with the aid of video cameras, and are capable of moving and of acting within it by the operation of mechanical limbs and hands. But this endeavour generates a host of subproblems in areas such as vision and pattern recognition, and if you want to talk to it, computer understanding of spoken language. The programs which do this, says Weizenbaum, are designed to 'compose subprograms, that generate new processes, that were not explicitly supplied by human programmers… Such… systems...gain knowledge by directly sensing their environments [and] thus come to know things not only by being told them explicitly, but also by discovering them while interacting with the world.'[37]

To speak of a computer knowing something is a metaphor. Knowing implies a subject that knows that it knows (although it's often hard to distinguish from belief). It is difficult, speaking of computers, to avoid using words like understanding, knowledge and knowing, learning and teaching, but critical to register their anthropomorphism and question their fit. Weizenbaum provides a nice example: 'A human may know, for example, just what kind of emotional impact touching another person's hand will have both on the other person and on [themself]… The knowledge involved is in part kinesthetic; its acquisition involves having a hand, to say the very least. There are, in other words, some things humans know by virtue of having a human body.'[38] An inorganic machine cannot know these things in the same way humans know them – a consideration that also applies to human interaction with domesticated animals. A robotic dog might be programmed to nuzzle up to a person but has no endorphins to give it a warm feeling.

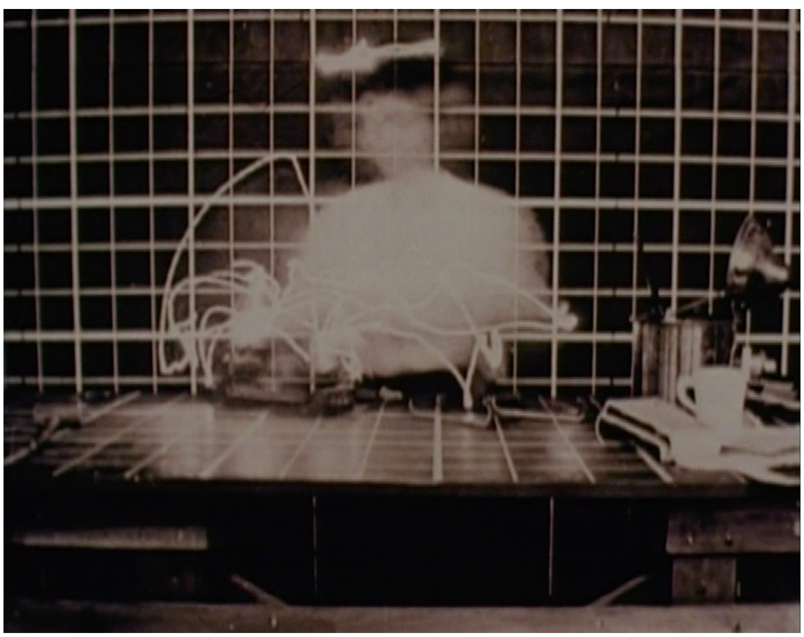
As for language, the computer has no more understanding than a parrot. Parrots have a remarkable capacity to imitate human speech but only in terms of their sound, whereas a computer, to understand speech, needs to be able to distinguish words, a task which comes naturally to dogs but not to a computer, and is much more difficult than parsing written text (not to mention logical or mathematical symbols). Written text lacks precisely what the parrot imitates, that is, intonation, accentuation and cadence, which always inflect spoken meaning in the most subtle ways. What the computer can do is parse the syntax, and construct grammatical responses, because these are rule-governed, at least up to a point – parsing a sonnet by Shakespeare or Góngora might be difficult – but even ordinary prose is full of ambiguity and metaphor, without which language can hardly function, but another stumbling block for the computer. Most of the time the meaning of an utterance is provided by the context, but this is dependent on a wealth of tacit or

37. Weizenbaum, p.208
38. Weizenbaum, pp.208-9

background knowledge, which the individual possesses by virtue of their socialisation, their culture, and their personal history. Human language manifests human memory, but a computer has no individual history of growing up, schooling, puberty, erotic desire, loves and hates, joy, shame, guilt. What a computer cannot do is recognise that the input it receives comes from such a person – the Turing Test in reverse, you might say. All it is capable of is recognising patterns.

*

The Eighteenth century automaton already says something about the power of visual representations to be taken up as models of understanding, in that case through physical simulation. Photography suggested fascinating possibilities but was of little scientific utility in the terms that interest us here until exposure times were sufficiently reduced to capture bodies in motion instead of posed, and then rapid series photography, most famously associated with Eadweard Muybridge, quickly revealed details unavailable to the unaided eye, like the moment a galloping horse has all four hooves in the air. When it comes to x-rays and then electronic devices like ultrasound scanning, the results may seem magical but still correspond to bodily forms. But if rapid series photography focussed on the body's dynamism and emphasised its mechanisms, it also lent itself to the body's estrangement by time-and-motion studies, exemplified in the photographs of Frank Gilbreth which trace the movements of the worker's hands while the worker is reduced to a blur.



In the computer, however, there is no longer the same indexical sign. Simulation takes the form of a program whose output has to be translated into a display of data that no longer corresponds to the visual form of the body; instead it consists in diagrams, graphs and symbolic animations, in which the perception of the body is fully abstracted and objectified. This is compounded by the ever-growing acuity of scientific instruments, their penetration of ever more microscopic levels of physical stuff. As Black observes, no-one has ever actually seen DNA and no-one ever will, not for lack of a sufficiently powerful microscope but because DNA is smaller than the wavelength of light, therefore doesn't interact with light, and so can't generate any visual information. 'The double helix structure of DNA wasn't discovered because someone saw DNA. It resulted from a process by which various people attempted to formulate a way of representing the known physical properties of DNA.'[39]

Rosalind Franklin's now famous photograph was not a photo of DNA as such, but only a visual representation of certain of its characteristics. The minuscule molecule has only ever appeared to us through visualisations and diagrams, today predominantly created by computers, and 'there is no single, objective or unproblematically truthful visual representation of a molecule'.[40] In all cases, the information is abstracted and interpreted in order to become legible, and even then the viewer needs special training in order to make sense of the results – when such images turn up in television news reports, briefly displayed along with an expert to explain them, they're mere fetishes. At best, any such representation is a form of metaphor, an analogy between dissimilar phenomena, and it's inevitable that metaphors for the structure of the invisible should be derived from structures at the human scale. By the same token it is also inevitable, indeed a spontaneous response of human intelligence, for human beings to see themselves mirrored in successive technologies, each of which suggests new metaphorical models for the body, and now the brain.

More than mirrored, however, for tools and machines are extensions of human activity which become linked with the brain in the form that the Soviet neuropsychologist A.R. Luria called 'functional knots': external devices, historically created, which establish connections between different parts of the brain, like 'the knot which we tie in our handkerchief so as to remember something essential, a combination of letters which we write so as not to forget an idea, or a multiplication table which we use for arithmetical operations'. All auxiliary tools or aids are 'historically formed measures for the organisation of human behaviour [which] tie new knots in the activity of man's brain'.[41] The computer is no different, except in the way it transforms our image of both brain and machine. The images produced by the computer are designed for human intelligibility. They are sometimes even imbued with aesthetic effects like perspective and artificial light and shadow in order, as Black puts it, to throw their contours into better relief. This aestheticisation is nothing new – it's there in the drawings of Vesalius – but presented here not as decoration but as integral to the representation, to the 'artistry' of the researcher, and surely, says Black, this in turn influences researchers' conceptualisation of the phenomena being studied. And ours, as users.

What we're left with is the paradox of the aesthetic, which lies outside calculation in the domain of affect, leaving an irreconcilable schism between mechanism and art. Computers were credited with intelligence before they were capable of showing very much of it, but as we've seen, a reduced form of intelligence which excludes what we think of as

---

39. Black, p.148
40. Black, p.147
41. A.R.Luria, *The Working Brain*, Penguin, 1973, p.31.

creativity – a difficult word to define which is hardly limited to art but has something to do with originality and novelty in the perception and solution of questions and problems. The artificial intelligentsia, who were gripped by the idea that computers were not just intelligent but *more* intelligent than humans, applied their own creativity to problems susceptible to calculation, like getting computers to play chess (although it took until 1997 for Deep Blue to finally defeat a reigning world chess champion). But chess is computable and has a singular objective, a determinable 'final state' defined by its rules, neither of which is true of artistic composition. There was joke when translation programs first appeared, about 'the spirit is willing, but the flesh is weak' being rendered into Russian as 'the vodka's good but the meat is bad'. There was also a rumour that somewhere in Switzerland there was a musicologist programming a computer to churn out Bach cantatas.

Weizenbaum takes Marvin Minsky to task for supposing that for computers to write good music or draw meaningful pictures only requires better semantic models to understand how they work. 'Clearly,' he says, 'what Minsky means by "understanding" music or painting is quite different from what, say, Mozart or Picasso meant by the same term.'[42] Or Schoenberg, who wrote about his first *Chamber Symphony* that it was only about twenty years after he composed it that he realised the relationship between its two main themes: 'It is of such a complicated nature that I doubt whether any composer would have cared deliberately to construct a theme in this way; but our subconscious does it involuntarily.'[43] And then there's the testimony of the mathematician Poncoiré about the way solutions to problems frequently dropped into his mind when he wasn't thinking about them: 'the appearance of sudden illumination,' he wrote, is 'a manifest sign of long, unconscious prior work'.[44] Are we to suppose that in such cases the unconscious is simply number-crunching away silently, or is something else going on? How are we to say what that might be, when the unconscious is an undifferentiated space, and intuition is a word for the inner voice which knows things without conscious reasoning? Are we to suppose that the computer, which has no consciousness, possesses an unconscious? A meaningless claim, although perhaps appealing as a metaphor for unknowing.

Thanks to AI, we are now awash with fantastical images generated by computer according to a verbal prompt by drawing on a gigantic library of tagged images ingested from the internet. Black, like Weizenbaum, considers it meaningless to talk about a computer creating art, 'as the computer has no intention and is aiming for no particular effect. The program is simply created by a human being to simulate a practice that other human beings might find aesthetically satisfying. To claim that this program is an artist is like claiming that the microbes that make coral are artists, given that they, too, are capable of producing artefacts human beings find aesthetically pleasing.'[45] But that last sentence makes me hesitate, because you can turn it around, and then what strikes you is that indeed we find great aesthetic pleasure in nature, which also has no artistic intention, but strikes us as beautiful and at times sublime. Is there any reason why the images produced by a

---

42. Weizenbaum, p157
43. Arnold Schoenberg, *Style and Idea*, Faber & Faber, 1975, p.223
44. Henri Poncoiré, 'Mathematical Creation', in Brewster Ghiselin, ed., *The Creative Process*, Mentor, 1952, p.38
45. Black, p101

computer should not also carry aesthetic appeal? You don't have to credit a computer program with artistic intention to find aesthetic effects in its output.

But that doesn't quite satisfy me either. Have we forgotten that art begins in mimesis? We look at a sunset and say it looks like a painting by Turner, or see clouds as a Constable. Art teaches us to see, not once and for all, but repeatedly. It constantly directs us to new aspects and new perceptions which feed off and into new imaginaries. Computers are now doing this too, and it seems we cannot help but let them. I am still waiting to hear a Bach cantata composed by a computer, but I've seen many examples of arresting computer-generated images composed by programs designed with no other purpose than taking pictorial elements, adding digital effects, and simulating the process of collage. However, they are liable to create the same kind of impression of the Freudian uncanny as a painting painted by a chimpanzee. They are not what they seem, but that is also true of art. The paradox is not solved by saying that what I'm seeing is a simulacrum, not the real McCoy, because you could say that this is how all of art is constituted, as an interplay of simulacra, whether consciously or not. The artist does not invent out of nothing but by working with an inherited repertoire of representations, constantly refashioning and transforming them – consciously or unconsciously – in a process of continual cultural interaction with other artists, dead and alive, that Bakhtin called dialogical. Indeed there is no act of speech, no letter, no poem, no book, painting or piece of music, no scientific theory or experiment, which is not shaped in this way; all are dialogical in nature, thereby inscribing the complex social situation in which each of them occurs. When it comes to computer programming and AI, however, this logic breaks down. The artificial language of computer programming and AI is entirely monological and self-referential. The social determinants are inscribed within it only negatively.

I am committed to scepticism about artificial intelligence and the claims of its apologists that it can emulate human intelligence and even surpass it, precisely because I'm not a computer, I'm a conscious human being who can tell simply by introspection and apperception that a computer doesn't 'think' the way I do, indeed it doesn't think at all. But I cannot deny that it changes the way that I do. This is partly because it changes the way I work, but it also stems from the fetishisation of the computer at the social level and the runaway effects of its colonisation of everyday human life. At the risk of oversimplification, this flows from the conjuncture of technological breakthrough and neoliberal ideology, which left a new sector of capital free to blow its own trumpet and develop its algorithms while in effect ignoring the laws of robotics. Too simple, of course, except metaphorically. The algorithms are real – as real as digital code can be – but hyping the claims served very well to attract the necessary investment capital, and the result, through boom and bust and back again, has been socially, psychologically and politically injurious. Not exclusively so – at many levels the benefits have been enormous – but the technology of virtual unreality which seduces the game player has also been used to kill and maim enemies in battle.

The speed of development of AI is frightening, and the harm it produces along with the good comes ever nearer. As I close this inquiry, a new large language model makes a bow. ChatGPT is freely available in your browser and has created an enormous stir because it claims that in response to the appropriate prompt, it can write everything from advertising copy to academic essays, even poems in a chosen style. I tried it out by asking it for a script for a short experimental video film about climate change. It responded with a script that wasn't experimental but would have been acceptable as an exercise by a first year undergraduate on a film course. Trained on millions of pages of text scraped from the web and other sources, it does all this by predicting statistically likely continuations of sequences of words and sentences, to produce a neat summary of conventional wisdom on the chosen topic delivered in bland polite language, programmed ('filtered') to exclude contentious terminology, or as the opening screen advises, 'Trained to decline inappropriate requests'.

The user is also warned that the program 'may occasionally generate incorrect information' and 'produce harmful instructions or biased content'. These are not just formulaic let-out clauses. The experts themselves have warned that large language models have limited reliability, are prone to bias based on the data they were trained with, suffer 'hallucinations', and lack transparency about how the responses are generated. If you test it on something you know about, you soon discover that it misses vital information, makes mistakes, and not infrequently contradicts itself; if you point any of this out, it apologises in the same bland and polite language. Worse still, if asked about the sources of the information provided, it invents references. I discovered this for myself by asking it questions about economics. It suggested various studies in academic journals without giving titles or references. When I asked for them the response came back, 'I apologize for the confusion, but the studies you have mentioned does not have a precise reference as it is not a specific study, it could be a synthesis of several studies done on the topic.' Curiously, this is the only sentence I got from it with a grammatical error. Perhaps by challenging it I'd got its knickers in a twist.

Widely publicised because of its rapid take-up, it is in the domain of education where the threat is most keenly felt. Academics have been thrown into panic – according to articles by other academics – by the prospect of AI bots with the ability to write undergraduate assignments. Very quickly new tools became available using AI to detect the use of AI – one of them was issued by ChatGPT itself – which are far from successful and make errors ('false positives'). Several universities and academic journals took rapid steps to ban it. All this occasioned some breast beating. A blogger on a higher education journal wrote that 'ChatGPT reveals the uncomfortable truth about graduate skills'. 'The scandal that should be grabbing the headlines is the fact that for a generation we have been training our undergraduates to be nothing more than AI bots themselves; this is why it is not possible to tell their work apart.'[46]

Another calls it a 'bullshit generator' which 'mimics the glib, bloodless prose that characterises so much academic writing', and says 'Thank you ChatGPT for exposing the banality of undergraduate essays'.[47] All appear resigned to it. To be fair to ChatGPT, they

---

46. John Warren, 'ChatGPT reveals the uncomfortable truth about graduate skills', www.timeshighereducation.com/blog/chatgpt-reveals-uncomfortable-truth-about-graduate-skills

47. Colm O'Shea, 'Thank you ChatGPT for exposing the banality of undergraduate essays', THES, 23.02.2023, www.timeshighereducation.com/blog/thank-you-chatgpt-exposing-banality-undergraduate-essays

are perfectly aware of the disruption and dangers they've unleashed, and lay out the issues clearly on the advisory page for educators.[48] They do not discuss, however, the bias of a highly tailored but still ill-fitting AI program to intellectual conformity, which is also a bias against deviation from the statistical norm, hence a bias against originality, novelty, imagination and creativity, which are words it doesn't understand and therefore cannot emulate.[49]

*

There is no lack of commentary about either the benefits or the harms of AI, but we should shy away from futurology. Futurology is part of what got us here. Moreover, futurology is slanted, like climate change denial. The problem is not what happens in the future but what is happening now, what has already happened, when even the benefits come at the cost of a loss of humanity. That, of course, is not limited to IT, but derives from an economic system driven by profit and fixated on the ideology of growth over the health of the ecosystem, a polarisation which is reproduced in the illusions of the artificial intelligentsia who ignore the material and human exploitation required to produce the infrastructure of their dreams, and the cost of the energy needed to keep it running. Computation is indifferent to what it computes. This follows from the principle of the Turing Machine, that any such machine is capable of imitating any other. To borrow one of Weizenbaum's examples, it may as well compute horoscopes as the weather. As the formula has it, garbage in, garbage out. If astrology is nonsense, then so is computerised astrology, but we now not only have pretty reliable weather forecasts, but climate modelling which tells us we may no longer have time to prevent irreversible catastrophe.

We also know the kinds of things that need to be done to confront the dangers, but how to do them is a matter of politics. Integral to these politics is the principle of sustainability. AI is complicit in the problem. The huge computing power required adds to the already significant energy load created by IT and the internet, the gamut of applications, platforms and programs, open on our screens or hidden away in 'the cloud' – another falsely attractive metaphor – which revamp our everyday life and drain our mediated intercourse. What we need in the face of the reductionism and abstraction of computer languages is the same thing that ecological consciousness encourages: to rediscover a holistic comprehension of our humanity, or rather, to forge one anew. Here perhaps the problem is that this is not what you could call a political programme, because it isn't clear what steps need to be taken to achieve such a thing. We know pretty well what needs to be done to mitigate the effects of climate change, but it isn't clear how to get the powers-that-be to adopt the measures so urgently required. But if such a change were to come about, it would be because in the end, rationality in the human being is not abstracted from the body and its environment.

Not the kind of rationality that tabulates and pigeon-holes, separates and classifies, but instead, an ecological rationality, alert to affinity and connection across the different existential domains of all the forms of life which are now threatened. Neither separated from fleshworld sociality, community and fellow feeling, nor alienated from a world that

---

48. 'Educator considerations for ChatGPT', https://platform.openai.com/docs/chatgpt-education

49. For a more detailed analysis see my blog post, 'Artificial Writing: a first evaluation', www.putneydebater.com/artificial-writing-a-first-evaluation

is more-than-human, a world of animals, plants, forests and fungi, all with their own forms of intelligent behaviour and communication, with whom homo sapiens is evidently but unknowingly entangled. The simple truth is that in doing them damage, we are damaging ourselves. I say 'in the end', but such an end would be a new beginning. If disaster is averted and the technological infrastructure does not collapse – a prospect no-one discusses, although it already happens in places here and there – then AI will doubtless have a place as a scientific tool, a tool of education, even an artistic tool, but not as a counterfeit consciousness.

Michael Chanan is a documentary filmmaker, writer and an Emeritus Professor of Film & Video at the University of Roehampton. His latest film is *Cuba: Living Between Hurricanes* (2019), about ecology and sustainable development in Cuba in the face of climate change, and his latest book is *From Printing to Streaming, Cultural Production Under Capitalism* (2022).