Toward a Short History of AI Paranoia

David Leavitt

2015 *IEEE Spectrum* publishes "Mocking AI Panic," a short article by the computer historian Diane Proudfoot:

> Many people today are concerned by the prospect of out-of-control artificial intelligence. Some call it "killer AI," "evil AI," or "malevolent AI." Billionaires throw money at the "existential risks" posed by ultraintelligent machines: in January Elon Musk . . . donated US $10 million to the Future of Life Institute in Cambridge, Mass., which is "focusing on potential risks from the development of human-level artificial intelligence." Other new research institutes with apocalyptic names explore the dangers of "singularity" scenarios, and Google has recently formed a hush-hush AI ethics board.[1]

↓

> **2019** Google shuts down its Advanced Technology External Advisory Council (ATEAC) one week after starting it. According to *Vox*, the decision owes to a petition, signed by thousands of Google employees, "calling for the removal of one board member, Heritage Foundation president Kay Cole James, over her comments about trans people and her organization's skepticism of climate change. Meanwhile, the inclusion of drone company CEO Dyan Gibbens reopened old divisions in the company over the use of the company's AI for military applications."[2]

Although this surge of "AI panic," Proudfoot notes, may seem to be as new as AI, in fact it is history repeating itself. In the 1940s, news of the "electronic brain" on which Alan Turing and his colleagues were at work at the National Physical Laboratory (NPL) in Teddington, near London, provoked what she calls a "fearful" public reaction that Turing himself "poked fun at," joking, but "with a good pinch of his usual common sense," that "we might be able to 'keep the machines in a subservient position, for instance, by turning off the power at strategic moments.'" Not an iota of humor, on the other hand, is to be found in Musk's observation that

© 2026 by David Leavitt
Published under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license
https://doi.org/10.1162/DAED.a.1037

AI is "more dangerous than nukes" or his prediction that even "fleeing to Mars – presumably on one of his own rockets – wouldn't help," since "the AI will chase us there pretty quickly."

Is this paranoia?

2025 Anthony Aguirre, executive director of the Future of Life Institute, which by this point has offices in California, Pennsylvania, Washington, D.C., London, and Brussels, posts on the institute's website an essay titled "Keep the Future Human: Why and How We Should Close the Gates to AGI and Superintelligence, and What We Should Build Instead." From the executive summary with which the essay opens:

- If this race continues much longer, there is an inevitable winner: AI itself – a faster, smarter, cheaper alternative to people in our economy, our thinking, our decisions, and eventually in control of our civilization.
- Bulk automation of most cognitive-based jobs, replacement of our current epistemic systems, and rollout of vast numbers of active nonhuman agents would upend most of our current civilizational systems in a relatively short period of time.
- By proliferating the ability – potentially above human level – to create new military and destructive technologies and decoupling it from the social and legal systems grounding responsibility, physical catastrophes from weapons of mass destruction become dramatically more likely.
- Unless it is specifically prevented, superhuman AI will have every incentive to further improve itself and could far outstrip humans in speed, data processing and sophistication of thinking. There is no meaningful way in which we can be in control of such a system. Such AI will not grant power to humans; we will grant power to it, or it will take it.
- Perhaps it would treat us well, perhaps not. But the future would belong to it, not us. The human era would be over.[3]

Is this paranoia?

1571 Michel de Montaigne, on a trip through Central Italy, visits the Villa Medici in Castello, where he is surprised by "jets of water so minute" as to be "almost invisible" that spurt up between the legs of visitors without warning. He also sees "a large bronze statue of a very old hoary man seated on his rear with his arms crossed, from all over whose beard, forehead, and hair water flows incessantly, drop by drop, representing sweat and tears." Later, Montaigne tours the Villa d'Este in Tivoli, where there are still more *jeux d'eau*, and where he hears

> the song of birds, which are little bronze flutes that you see at regals; they give a sound like those little earthenware pots full of water that little children blow into by the spout, this by an artifice like that of an organ; and then by other springs they set in motion an owl, which, appearing at the top of the rock, makes this harmony cease

instantly, for the birds are frightened by his presence; and then he leaves the place to them again. This goes on alternately as long as you want.[4]

↓

1949 "The Mind of Mechanical Man" is the title of the annual Lister Oration, which Sir Geoffrey Jefferson, professor of neurosurgery at Manchester University, delivers to the Royal College of Surgeons on June 9.[5] "Before we glance at the new vistas of mechanization opening before us," he tells his audience,

> let us spare a few minutes to look at the past, where we shall find that the possibility of building automata has been one of man's dreams since the days of the Trojan Horse – a simile more metaphorical than strictly accurate. In the seventeenth century, that era of great scientific awakening, there was great interest in possible replicas of animals and men. Florent Schuyl,[6] in 1664, gives several instances, such as the wooden pigeon of Archytas of Tarentum which flew through the air, suspended by counterweights, the wooden eagle, that of Regiomontanus, that showed an Emperor the way to Nuremburg, and a flying fly of the same maker. There was an earthen head that spoke; but, above all, a marvellous iron statue that knelt before the Emperor of Morocco and presented him with a request for a pardon for the man who had made him.

"If such wonders had already been constructed for the pleasure of noblemen and the entertainment of their guests," Jefferson asks, "how much more perfectly might not the serious scientist contrive a cunning replica of a living thing. As only too often happens, to conceive it possible was as good as its conversion into fact. *It could be, therefore it was*. I am sure that is our own temptation."[7]

1647 The automata that impress Montaigne in the sixteenth century – clockwork, mechanical, pneumatic, hydraulic – trouble Descartes in the seventeenth. As severed heads, "soon after they have been cut off, still move and bite the ground," Descartes postulates, so "automata and moving machines, devised by human ingenuity," may give the impression of being alive even though they are not. "If there were such machines with the organs and shape of a monkey or of some other nonrational animal," he argues, "we would have no way of discovering that they are not the same as these animals. But if they were machines that resembled our bodies and if they imitated our actions as much as is morally possible, we would always have two very certain means for recognizing that, none the less, they are not genuinely human."[8]

First, such machines "would never be able to use speech, or other signs composed by themselves, as we do to express our thoughts to others." True, an automaton could be built that utters words, but only "in response to physical actions that cause

a change in its organs – for example, if someone touched it in a particular place, it would ask what one wishes it to say to it, or if it were touched somewhere else, it would cry out that it was being hurt, and so on." It could not, however, "arrange words in different ways to reply to the meaning of everything that is said in its presence, as even the most unintelligent human beings can do."

Second, even if these machines do "many things as well, or possibly better than any one of us, they will infallibly fail in others. Thus one would discover that they did not act on the basis of knowledge, but merely as a result of the disposition of their organs."

↓

1947 Alan Turing, in a lecture on the Automatic Computing Engine (ACE) delivered to the London Mathematical Society on February 20:

> Thus if a machine is made for this purpose it must in some cases fail to give an answer. On the other hand if a mathematician is confronted with such a problem he would search around a[nd] find new methods of proof, so that he ought eventually be able to reach a decision about any given formula. This would be the argument. Against it I would say that fair play must be given to the machine. Instead of it sometimes giving no answer we could arrange that it gives occasional wrong answers. But the human mathematician would likewise make blunders while trying out new techniques. It is easy for us to regard these blunders as not counting and give him another chance, but the machine would probably be allowed no mercy. In other words then, if a machine is expected to be infallible, it cannot also be intelligent.[9]

1647 Descartes's conclusion: "It is morally impossible for a machine to have enough different dispositions to make it act in every human situation in the same way as our reason makes us act."

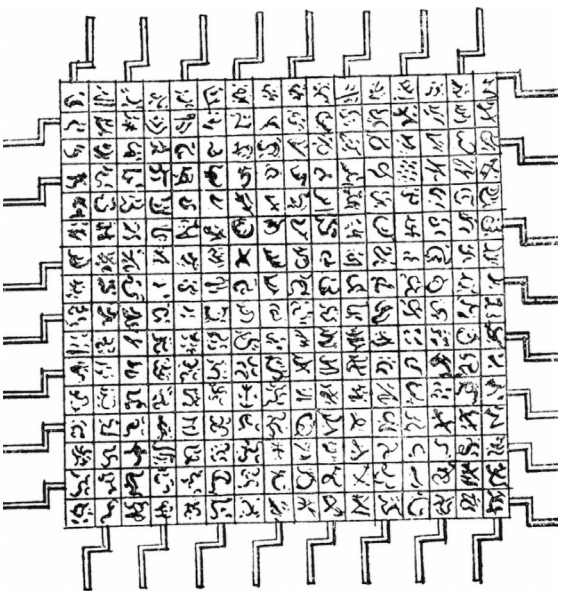
Is this paranoia?

1726 In Part Three of Jonathan Swift's anonymously published *Travels into Several Remote Nations of the World*, now better known as *Gulliver's Travels*, Lemuel Gulliver finds himself in Lagado, the capital city of Balnibarbi and home to the country's "grand academy," of which he is given a tour. Among the projects being undertaken, none with much success, at the academy: a process for extracting sunlight from cucumbers, "an operation to reduce human excrement to its original food, by separating the several parts, removing the tincture which it receives from the gall, making the odour exhale, and scumming off the saliva," and a "frame" the function of which, according to its architect, is to improve

> speculative knowledge, by practical and mechanical operations. But the world would be sensible of its usefulness; and he flattered himself that a more noble, exalted thought never sprang in any other man's head. Every one knew how laborious the usual method is of attaining to arts and sciences; whereas, by his contrivance, the most ignorant person, at a reasonable charge, and with a little bodily labour, might write books in philosophy, poetry, politics, laws, mathematics, and theology, without the least assistance from genius or study.

The architect now shows Gulliver his device, about which his many pupils stand "in ranks."

> The superficies was composed of several bits of wood, about the bigness of a die, but some larger than others. They were all linked together by slender wires. These bits of wood were covered, on every square, with paper pasted on them; and on these papers were written all the words of their language, in their several moods, tenses, and declensions; but without any order. The professor then desired me "to observe; for he was going to set his engine at work." The pupils, at his command, took each of them hold of an iron handle, whereof there were forty fixed round the edges of the frame; and giving them a sudden turn, the whole disposition of the words was entirely changed.
>
> He then commanded six-and-thirty of the lads, to read the several lines softly, as they appeared upon the frame; and where they found three or four words together that might make part of a sentence, they dictated to the four remaining boys, who were scribes. This work was repeated three or four times, and at every turn, the engine was so contrived, that the words shifted into new places, as the square bits of wood moved upside down.



The function of the device, then, is the generation of language, if through an arduous combinatorial method.

> Six hours a day the young students were engaged in this labour, and the professor showed me several volumes in large folio, already collected, of broken sentences, which he intended to piece together, and out of those rich materials, to give the world a complete body of arts and sciences. . . . He assured me "that his invention had employed all his thoughts from his youth: that he had emptied the whole vocabulary into the frame, and made the strictest computations of the general proportion there is in books between the number of particles, nouns, and verbs, and other parts of speech."[10]

The machine, made of language, makes language – but not very well.

↓

1952 At Manchester University, Christopher Strachey, a computer scientist and the nephew of Lytton Strachey, devises, under Alan Turing's aegis, an algorithm by means of which the university's Mark I computer can write love letters by filling in the blanks in a conventional text – an early version of a MadLib. On his website, the computer scientist Nick Monfort includes an online version of the algorithm that churns out love letters every few seconds. Some examples:

> Jewel moppet
>
> My adoration yearns for your sympathetic ardour. My fond wish beautifully clings to your heart. You are my devoted sympathy. My burning fervour devotedly woos your fervour. You are my loveable heart: My devoted fancy. You are my burning tenderness. My avid appetite pants for your hunger. My love ardently likes your ardent appetite. You are my lovesick enthusiasm.
>
> Yours ardently
> M.U.C.
>
> Honey duck
>
> You are my breathless lust: My dear passion. You are my ardent lust: My fond love. You are my adorable enthusiasm. My passionate passion avidly wants your amorous enthusiasm. My fancy fondly longs for your ambition. You are my loveable ardour. My little love clings to your fond hunger. My rapture woos your adorable thirst.
>
> Yours curiously
> M.U.C.

M.U.C. stands for Manchester University Computer.[11]

1837 In the early 1820s, Charles Babbage, later to be Lucasian Professor of Mathematics at Cambridge, undertakes to build a "Difference Engine" that will automate the generation of mathematical tables. Structurally he bases the design on the Jacquard loom, incorporating into it gears, toothed wheels, axles, and a crank handle. Although he never completes the engine, he does succeed in building several parts of it as well as a small working model. Nor is this the only machine he aspires to construct. The other, on which he begins work in 1837, he calls the "Analytical Engine." It is widely agreed that the "Analytical Engine," had it been completed, would have been the first digital computer.

↓

1950 Alan Turing, in "Computing Machinery and Intelligence":

> The fact that Babbage's Analytical Engine was to be entirely mechanical will help us to rid ourselves of a superstition. Importance is often attached to the fact that modern digital computers are electrical, and that the nervous system also is electrical. Since Babbage's machine was not electrical, and since all digital computers are in a sense equivalent, we see that this use of electricity cannot be of theoretical importance.[12]

1843 Collaborating with Babbage on the Analytical Engine is Lady Ada Lovelace, Lord Byron's only legitimate child and the first computer programmer in history. When in 1842 Luigi Federico Menebrea – statesman, mathematician, and for two years (1867–1869) Prime Minister of Italy – publishes, in French, his "Sketch of the Analytical Engine invented by Charles Babbage, Esq.," Lady Lovelace not only translates the sketch into English but adds a series of notes to it, one of which, Note G, maps out an algorithm for calculating the Bernoulli Numbers. One passage in Note G will be quoted hundreds if not thousands of times in the coming years:

> The Analytical Engine has no pretensions whatever to *originate* anything. It can do *whatever we know how to order it* to perform. It can *follow* analysis; but it has no power of anticipating any analytical relations or truths. Its province is to assist us in making *available* what we are already acquainted with.[13]

This passage, which in 1950 Alan Turing will dub "Lady Lovelace's Objection," reads to me more as a note of caution – but to whom? Had it been suggested to Lady Lovelace that the Analytical Engine might be capable of doing *more* than what we tell it to do? Or is she foreseeing – and trying to forestall – a paranoia the cause of which has yet, to borrow her language, to be *originated*?

1872 Samuel Butler's *Erewhon*, or *Over the Range*, is published in London by Trübner & Co. of Paternaster Row.

↓

1959 In her biography of her son, Sara M. Turing relates that as a teenager, Alan Turing "had a particular fondness for *The Pickwick Papers*, George Borrow's books and Samuel Butler's *Erewhon*. This last possibly set him to think about the construction of an actual intelligent machine."[14] Of particular interest to Turing is *The Book of the Machines*, an independent narrative enclaved in Butler's novel.

Like *Gulliver's Travels*, *Erewhon* is at once a satire and a travelogue. Its nameless narrator, a young Englishman, takes an arduous journey and ends up in a strange land. To borrow a phrase from the poet Elizabeth Bishop, Erewhon is "the world inverted."[15] Its name is an inversion, or nearly, of *nowhere*, just as Nosnibor, the name of the narrator's host, is an inversion of *Robinson*. Nor are the inversions only linguistic. In Erewhon, illness is a crime and crime is an illness. There are "Colleges of Unreason" where the professors teach Hypothetics, Inconsistency, and Evasion. There is a symbolic "musical" currency that the Erewhonians make a pretense of using but do not use and a mercantile currency in which, the narrator suspects, even the managers and cashiers at the "Musical Banks" are paid. Also there are no machines, or rather no functioning machines, only a museum housing "fragments of steam engines, all broken and rusted . . . a cylinder and piston, a broken flywheel, and part of a crank, which was laid on the ground by their side."[16]

As the narrator soon learns, some five hundred years before his arrival in Erewhon, a war erupted between "the machinists and the anti-machinists" in which the anti-machinists were victorious. The anti-machinist argument is laid out in an ancient book of which the narrator obtains a copy and that he attempts to translate into English. This is *The Book of the Machines*, and it begins:

> There was a time, when the earth was to all appearance utterly destitute both of animal and vegetable life, and when according to the opinion of our best philosophers it was simply a hot round ball with a crust gradually cooling. Now if a human being had existed while the earth was in this state and had been allowed to see it as though it were some other world with which he had no concern, and if at the same time he were entirely ignorant of all physical science, would he not have pronounced it impossible that creatures possessed of anything like consciousness should be evolved from the seeming cinder which he was beholding? Would he not have denied that it contained any potentiality of consciousness? Yet in the course of time consciousness came. Is it not possible then that there may be even yet new channels dug out for consciousness, though we can detect no signs of them at present?

Indeed, right now – that is, at the time that the ancient book was written – not only is the world replete with conscious beings, it is replete with the machines that they have built to help make their lives easier. What if these machines acquired consciousness? asks the nameless author of *The Book of the Machines*, for whom

this possibility is a real threat. Nor is there any "security," he warns, "against the ultimate development of mechanical consciousness, in the fact of machines possessing little consciousness now," since it is not "the existing machines" that pose the threat but rather

> the extraordinary rapidity with which they are becoming something very different to what they are at present. No class of beings have in any time past made so rapid a movement forward. Should not that movement be jealously watched, and checked while we can still check it? And is it not necessary for this end to destroy the more advanced of the machines which are in use at the present, even though it is admitted that they are in themselves harmless?

Not, he reminds us, that humankind is blameless. On the contrary, it is due to humankind's willingness, even avidity, to cede power to the machines that they are gaining ground, since

> wherever precision is required man flies to the machine at once, as far preferable to himself. Our sum-engines never drop a figure, nor our looms a stitch; the machine is brisk and active, when the man is weary; it is clearheaded and collected, when the man is stupid and dull; it needs no slumber, when man must sleep or drop; ever at its post, ever ready for work, its alacrity never flags, its patience never gives in; its might is stronger than combined hundreds, and swifter than the flight of birds; it can burrow beneath the earth, and walk upon the largest rivers and sink not. This is the green tree; what then shall be done in the dry?

"How many men at this hour," the author asks, "are living in a state of bondage to the machines?"

> How many spend their whole lives, from the cradle to the grave, in tending them by night and day? Is it not plain that the machines are gaining ground upon us, when we reflect on the increasing number of those who are bound down to them as slaves, and of those who devote their whole souls to the advancement of the mechanical kingdom?

↓

1947 Turing, in his lecture on the ACE to the London Mathematical Society:

> Roughly speaking those who work in connection with the ACE will be divided into its masters and its servants. Its masters will plan out instruction tables for it, thinking up deeper and deeper ways of using it. Its servants will feed it with cards as it calls for them. They will put right any parts that go wrong. They will assemble data that it requires. In fact the servants will take the place of limbs. As time goes on the calculator itself will take over the functions both of masters and servants. The servants will be replaced by mechanical and electrical limbs and sense organs.[17]

Another worry for the author of *The Book of the Machines* is the possibility that machines may become capable of reproduction:

> It is said by some with whom I have conversed upon this subject, that the machines can never be developed into animate or *quasi*-animate existences, inasmuch as they have no reproductive system, nor seem ever likely to possess one. If this be taken to mean that they cannot marry, and that we are never likely to see a fertile union between two vapour-engines with the young ones playing about the door of the shed, however greatly we might desire to do so, I will readily grant it. But the objection is not a very profound one. No one expects that all the features of the new existing organisations will be absolutely repeated in an entirely new class of life. The reproductive system of animals differs widely from that of plants, but both are reproductive systems. Has nature exhausted her phases of power?[18]

What the author's fears come down to: a future in which "it is the machines which act upon man and make him man, as much as man who has acted upon and made the machines" and in which "we must choose between the alternative of undergoing much present suffering, or seeing ourselves gradually superseded by our own creatures, till we rank no higher in comparison with them, than the beasts of the field with ourselves." Yet to his horror, "many seem inclined to acquiesce in so dishonourable a future."

> They say that although man should become to the machines what the horse and dog are to us, yet that he will continue to exist, and will probably be better off in a state of domestication under the beneficent rule of the machines than in his present wild condition. . . . In like manner there is reason to hope that the machines will use us kindly, for their existence will be in a great measure dependent upon ours; they will rule us with a rod of iron, but they will not eat us; they will not only require our services in the reproduction and education of their young, but also in waiting upon them as servants; in gathering food for them, and feeding them; in restoring them to health when they are sick; and in either burying their dead or working up their deceased members into new forms of mechanical existence.

In the end, the author and his allies win the war. They succeed

> in destroying all the inventions that had been discovered for the preceding 271 years, a period which was agreed upon by all parties after several years of wrangling as to whether a certain kind of mangle which was much in use among washerwomen should be saved or no. It was at last ruled to be dangerous, and was just excluded by the limit of 271 years. Then came the reactionary civil wars which nearly ruined the country, but which it would be beyond my present scope to describe.

Is this paranoia?

1863 An early incarnation of *The Book of the Machines* appears as a series of letters to *The Press* of New Zealand, where Butler is living at the time, signed "Cellerius" and bearing the title "Darwin Among the Machines." Two years later, a revised version of "Darwin Among the Machines" is published in the London-based *Reasoner*. So far as I can tell, at the time, neither of these nor *Erewhon* provoked any sort of anxiety among readers or critics, at least concerning the possibility of the machines "taking over" anywhere outside Butler's inverted world. Indeed, it is only in 1948, with the publication of Norbert Wiener's *Cybernetics*, that *Erewhon* becomes a nexus for such anxieties.

↓

1948 From Norbert Wiener's introduction to *Cybernetics, or Control and Communication in the Animal and the Machine*:

> I have said that this new development [the modern ultra-rapid computing machine] has unbounded possibilities for good and for evil. For one thing, it makes the metaphorical dominance of the machines, as imagined by Samuel Butler, a most immediate and non-metaphorical problem. It gives the human race a new and most effective collection of mechanical slaves to perform its labor. Such mechanical labor has most of the economic properties of slave labor, although, unlike slave labor, it does not involve the direct demoralizing effects of human cruelty. However, any labor that accepts the conditions of competition with slave labor accepts the conditions of slave labor, and is essentially slave labor. The key word of this statement is *competition*. It may very well be a good thing for humanity to have the machine remove from it the need of menial and disagreeable tasks, or it may not. I do not know.[19]

↓

1949 In a letter published in *The Times* of June 14 in response to an interview with Turing, Illtyd Trethowan of Downside Abbey, Bath, warns that research into "the degree of intellectual activity of which a machine was capable, and to what extent it could think for itself" will compel even

> dialectical materialists . . . to guard themselves, like Butler's Erewhonians, against the possible hostility of the machines. And those of us who not only confess with our lips but believe in our hearts that men are free persons (which is unintelligible if we have no unextended mind or soul, but only a brain) must ask ourselves how far Mr. Turing's opinions are shared, or may come to be shared by the rulers of our country.[20]

↓

1950 Turing lists *The Book of the Machines* in the bibliography of "Computing Machinery and Intelligence," but in the essay itself makes no explicit references to it, or to Butler, or to *Erewhon*.

1909 E. M. Forster, then thirty years old, publishes "The Machine Stops" in the November issue of the *Cambridge and Oxford Review*. An outlier in Forster's oeuvre, "The Machine Stops" takes place in an unspecified future in which the majority of human beings live, interact, and are fed and cared for through the agency of a vast underground Machine. The story's protagonist, Vashti, occupies "a small room, hexagonal in shape, like the cell of a bee" that she virtually never leaves. In this room there are

> buttons and switches everywhere – buttons to call for food, for music, for clothing. There was the hot-bath button, pressure of which a basin of (imitation) marble rose out of the floor, filled to the brim with a warm deodorized liquid. There was the cold-bath button. There was the button that produced literature. And there were of course the buttons by which she communicated with her friends. The room, though it contained nothing, was in touch with all that she cared for in the world.[21]

Vashti is small and white and wrinkled, and has thousands of friends with whom she "communicates" through the Machine. She also has a son, Kuno, whom she finds annoying since he has a habit of interrupting her lectures on such topics as "Music During the Australian Period." Kuno is a troublemaker. He calls the Machine "wearisome," accuses his mother of treating it as a god, even urges her to take an air-ship to visit him "so that we can meet face to face, and talk about the hopes that are in my mind." But Vashti dislikes air-ships. She dislikes "seeing the horrible brown earth, and the sea, and the stars when it is dark." Vashti is not merely a denizen, not merely a citizen, she is a worshipper of the Machine. Merely to leave her room, for her, is to be "seized with the terrors of direct experience." Even so, she agrees to visit her son, and boards the air-ship, and, seeing the Himalayas through a window, asks one of her few fellow passengers "And that white stuff in the cracks? – what is it?" to which he replies, "I have forgotten its name." Eventually she arrives at Kuno's room, which, though halfway across the world, is identical to her own, and where he tells her that he has gone up onto the surface of the planet without an "Egression permit" and for this reason has been threatened with "Homelessness." Threatened by whom? By some high-level committee? By the Machine itself? Do its users control the Machine or does it control them? Kuno, Vashti observes, is "possessed of a certain physical strength."

> By these days it was a demerit to be muscular. Each infant was examined at birth, and all who promised undue strength were destroyed. Humanitarians may protest, but it would have been no true kindness to let an athlete live; he would never have been happy in that state of life to which the Machine had called him. . . . Man must be adapted to

> his surroundings, must he not? In the dawn of the world our weakly must be exposed on Mount Taygetus, in its twilight our strong will suffer euthanasia, that the Machine may progress, that the Machine may progress, that the Machine may progress eternally.

At this point we are halfway through the story.

Is this paranoia?

1947 In the introduction to his *Collected Stories*, Forster, now sixty-nine, describes "The Machine Stops" as "a reaction to one of the earlier heavens of H. G. Wells."[22] This, in fact, is all he has to say about it.

↓

1947 After a frustrating year at the National Physical Laboratory in Teddington, Alan Turing asks for – and is granted – a sabbatical leave. He then returns to King's College, Cambridge, where he resumes the fellowship that he was granted in 1935. This gives him an opportunity to renew old Cambridge friendships and forge new ones, including his friendship with P. N. (Nick) Furbank, whom he meets through the mathematician Robin Gandy.

↓

Gandy is a member of the Apostles, a quasi-secret Cambridge society whose alumni include Forster, John Maynard Keynes, Lytton Strachey, G. H. Hardy, G. E. Moore, and Bertrand Russell. Anthony Blunt, Guy Burgess, and John Cairncross, three of the "Cambridge Spies," are also Apostles. Turing is not.

↓

At the time, Furbank, whom Turing will later name his literary executor, is at work on a biography of Samuel Butler. Later he will write a biography of Forster.

↓

1948 Although there is no direct evidence that Turing met Forster during this period, it is difficult to imagine that they *didn't* meet, given that both were Fellows of King's and friends of Furbank's. Possibly the publication of Forster's *Collected Short Stories* led Turing to read "The Machine Stops." Possibly Forster shared with him the manuscript of *Maurice*, the novel he withheld from publication until after his death because of its frank portrayal of a love affair between two men.

↓

1951 The defection of Guy Burgess and Donald Maclean to the Soviet Union, and the resulting panic over the possibility that other gay men in possession of state secrets, as Turing is, may follow suit, contributes to a sudden

upsurge in the number of prosecutions under Section 11 of the Criminal Law Amendment Act of 1885. Section 11, colloquially known as "The Blackmailer's Charter," reads:

> Any male person who, in public or private, commits, or is a party to the commission of, or procures or attempts to procure the commission by any male person of, any act of gross indecency with another male person, shall be guilty of a misdemeanor, and being convicted thereof shall be liable, at the discretion of the court, to be imprisoned for any term not exceeding two years, with or without hard labour.

Is this paranoia?

1946 On October 31, Viscount Mountbatten of Burma, in his capacity as newly elected president of the British Institution of Radio Engineers, gives an address at the organization's annual dinner that incites an unexpected wave of panic.[23] The speech begins, as such speeches usually do, with canned encomia ("I would like to pay tribute to those who have served the Institution"), then moves on to a summary of electronic and wireless "strides" made during World War II, before taking a distinct turn for the visionary. First Mountbatten suggests that, based on the success of radar, "by pooling the potentialities of other forms of radiation (such as light, heat, sound, X-rays, gamma rays, and cosmic rays), we may in future be able to receive the counterparts of Radar screen pictures, from inside our bodies, or even from inside individual body cells. Or perhaps we may receive them from the interior of the earth, or from the stars and galaxies." Then he describes a "sense-machine" that could "present us with information not only visually, but also in the form of sound, or even feeling." Then he posits a hypothetical "magic desk," the owner of which will have "at his disposal what would nowadays be the whole contents of a colossal reference library of millions of volumes" and that will spare him from

> going through the long process that is now necessary in public libraries. As we move toward the abolition of routine mental work as well as of routine physical labour, we are being enabled not only to accomplish achievements far beyond the scope of present human attainment, but also, theoretically at any rate, to free the human mind for creative processes.[24]

↓

1947 Turing, in his lecture to the London Mathematical Society:

> Let us for the sake of argument suppose that the difficulties involved in using books as memory were overcome, that is to say that mechanical devices for finding the right book and opening it at the right page, etc. etc. had been developed, imitating the use of human hands and eyes. The information contained in the books would still be rather inaccessible because of the time occupied

> in the mechanical motions. One cannot turn a page over very quickly without tearing it, and if one were to do much book transportation, and do it fast, the energy involved would be very great. Thus if we moved one book every millisecond and each was moved ten metres and weighed 200 grams, and if the kinetic energy were wasted each time we should consume 10^{10} watts, about half the country's power consumption. If we are to have a really fast machine then, we must have our information, or at any rate a part of it, in a more accessible form than can be obtained with books. It seems that this can only be done at the expense of compactness and economy, e.g. by cutting the pages out of the books, and putting each one into a separate reading mechanism. Some of the methods of storage which are being developed at the present time are not unlike this.[25]

Finally, as if the radar screen pictures sent from outer space, the "sense-machine," and the "magic desk" were not enough, Mountbatten declares that "the stage is now set for the most Wellsian development of all: the Electronic Brain."

What is this "electronic brain?" According to Mountbatten, it is a machine that can "receive information about the situation of the machinery under its control and provide an intelligent – I repeat, intelligent – link between that information and the action necessary to keep the machinery in general conformity with the overall directions given to it by man." Machines already in existence, he adds, can "exercise a degree of memory; and some are now being designed to exercise those hitherto human prerogatives of choice and judgment. One of them could even be made to play a rather mediocre game of chess!"

↓

1946 Turing, from a report on "Proposals for the Development of an Automatic Computing Engine (ACE)" delivered in February:

> "Can the machine play chess?" It could fairly easily be made to play a rather bad game. It would be bad because chess requires intelligence. We stated at the beginning of this section that the machine should be treated as entirely without intelligence. There are indications however that it is possible to make the machine display intelligence at the risk of its making occasional serious mistakes. By following up this aspect the machine could probably be made to play very good chess.[26]

As an example of this "revolution of the mind," Mountbatten cites the ENIAC (Electronic Numeral Integrator and Computer) at the University of Pennsylvania, even though this is not the machine he is thinking of. The machine he is thinking of is the ACE, which is at the moment being developed at the National Physical Laboratory. Because no official announcement of the ACE has yet been made, Sir Charles Darwin, the NPL's director and the grandson of the author of *Origin of Species*, asked

Mountbatten to omit specific mention of it in his address. Mountbatten complied, but only to the extent that he attributed the capacities of the ACE to the ENIAC, a very different machine but a safer one to mention since it had already been unveiled.[27]

The next morning (November 1) articles on the "electronic brain" are published in newspapers all over England, including *The Times*.[28] Most of these merely summarize Mountbatten's address, though a few go further. *The Leicester Mercury* worries that in such a "future" as Mountbatten predicts, we may be "replaced by the strange mastery of our ingenuity," while *The Evening News* quotes an "official of the Department of Scientific Research" as saying that a machine currently being built in England will "go far in advance of the two in the U.S." and specifying that this "electronic brain" is to be found in "the mathematical department of the National Research Laboratory in Molesey."[29]

The cat, to Darwin's dismay, is out of the bag.

1946 On November 2, an editorial titled "Frankenstein" appears in *The Daily Telegraph*:

> This chastened generation will hesitate to welcome the electronic brain with completely uninhibited optimism. . . . We must also ask if, in effect, such machines will release humanity for higher processes of thought, or will they require the service or the servitude of a host of moronic button-pushers, lever-pullers, and dial-watchers so that they may facilitate the Olympian cogitations of an aristocracy of super-minds?[30]

The Irish Times resorts to even more alarmist rhetoric: "Any man who cherishes the position of his kind in the scale of creation will feel a tremor of repulsion as he reads about his potential usurper. . . . With insidious rapidity, the controlled monster can become the monster of control, and humanity can be reduced to the position of brilliant, degenerate serfs."[31]

Is this paranoia?

1946 Seeing that damage control is in order, on November 6, Darwin authorizes an NPL press release announcing the ACE but emphasizing that it is a work in progress and will probably not be ready for several years. The press release also identifies the members of the "team working on the 'Brain,'" whose names *The Daily Telegraph* swiftly publishes: "They are Dr. A.M. TURING, Dr. J.R. WOMERSLEY, superintendent of the mathematics division of the laboratory, and Prof. D. HARTREE, who worked on Eniac, the American equivalent of Ace."[32] In this article, Turing, who "conceived the idea of Ace," says that he foresees a time, possibly in thirty years, "when it would be as easy to ask the machine a question as to ask a man." Hartree, however, "deprecate[s] any notion that Ace could ever be a complete substitute for the human brain," adding portentously: "The fashion which has sprung up in the last 20 years to decry human reason is a path which leads straight to Nazism."[33]

The next day, in a letter to *The Times*, Hartree once again "deprecates" the term "electronic brain," but this time adds:

> These machines can only do precisely what they are instructed to do by the operators who set them up. It is true that they can be set up in such a way as to exercise a certain amount of judgment. But it must be clearly understood that the situation in which judgment has to be applied, the way the results of applying these criteria are to be assessed, and the decisions as to the action to be taken on these results, must all be fully thought out and anticipated in setting up the machine.[34]

This is, of course, exactly the same point that Lady Lovelace made in 1843.

↓

1949 In "The Mind of Mechanical Man," Sir Geoffrey Jefferson also repeats Lady Lovelace's note of caution without giving her credit for it: "It can be urged, and it is a cogent argument against the machine, that it can answer only problems given to it, and, furthermore, that the method it employs is one prearranged by its operator." And later: "When we hear it said that wireless valves think, we may despair of language."[35]

↓

1952 In an essay titled "Can Machines Think?" Wolfe Mays, professor of philosophy at Manchester University, quotes this last line of Jefferson's approvingly, then goes on to refute the very possibility of machine intelligence:

> It is not altogether too fanciful to suppose that the machine analogy, with its emphasis on overt behaviour and abnegation of private experience, may, when the doctrines of *Cybernetics* finally percolate down to the lower grades of the Civil Service, lead us to be regarded, more than ever before, as if we were mechanical objects. It is not such a far cry from Aristotle's view that slaves were just human tools, to some future benevolent dictatorship of the Orwell 1984 type, where men may be seen as little else but inefficient digital computors [sic] and God as the Master Programmer.[36]

In his letter to *The Times*, Hartree quotes himself (from an article on the ENIAC published in *Nature*) – "use of the machine is no substitute for the thought of organizing the computations, only for the labour in carrying them out" – before concluding: "It seems to me that the distinction is important, and that the term 'electronic brain' is misleading in that it ascribes to the machine capabilities that it does not possess; and that is why I hope this term will be avoided in the future."[37]

It is not. On the contrary, over the next several days alone a plethora of articles and editorials on the "electronic brain" appear in both national and regional British newspapers, many of them skeptical, some outright panicked.[38] "It is just as

well man is not to abdicate his sovereign reason to this new Frankenstein of his own creation," *The Dundee Courier and Advertiser* opines.[39] "World problems to-day seem nearly insoluble, but current political methods of regimentation treat us too much like robots already." Then there is this from *The Times and South Staffordshire Advertiser*: "As with atomic energy, so now with these brains and memory machines, means must be devised to use the inventions of science for the common good. Unless that is done, the human race will perish, victims of their own brain products."[40]

On November 11, a discomposed Darwin finally writes a letter of his own to *The Times*, which is published on the 13th:

> In a letter published in your column on November 7 Professor Hartree deprecates the use of the term "electronic brain" in connexion with the new electronic calculating engines, and I wish to associate myself with his protest. In popular language the word "brain" is associated with the higher realms of the intellect, but in fact a very great part of the brain is an unconscious automatic machine producing precise and sometimes very complicated reactions to stimuli. This is the only part of the brain we may aspire to imitate. The new machines will in no way replace thought, but rather they will increase the need for it, both because the machines themselves will call for high qualities of intellect in their use, and also because as a tool they will open up new fields for study.

As for Mountbatten, Darwin takes great care with his words, noting only that the Viscount "had been fully informed by us of our project of the automatic computing engine, but at our request he did not mention it explicitly because it had not yet been made public. I would emphasize that he put the matter correctly in every respect, and that it was only a loose interpretation of his words that led to the present confusion."[41] The truth, in fact, is that it is the words that have led to the confusion.

About Turing Darwin says nothing.

1955 Although Alan Turing has often been called "the father of artificial intelligence," the term *artificial intelligence* (AI) only came into usage when it was coined by the American computer scientist John McCarthy the year after Turing's death. The term Turing uses is *machine intelligence*. Today *artificial intelligence* and *machine intelligence* are regarded as essentially synonymous, when in fact there is an important if subtle difference between them. Whereas artificial intelligence is a *system* entered through a multitude of individual portals, machine intelligence is specific to individual machines. In part, Turing's use of the term reflects the fact that at the time, computers were conceived as unique entities. More importantly, however, it denotes an aspect of his concept of the "thinking machine" that has as much to do with what it means to be human as with what it means to be a machine: his investment of both real and hypothetical computers with *personhood*.

↓

Person is a more complicated word than it may at first appear. In the *OED*, the definition that most closely aligns with how we use the word today – "an individual human being; a man, woman, or child" – is the *second* to be listed. The first definition is a "role or character assumed in real life, or in a play, etc.; a part, function, or office; a persona; a semblance or guise. Hence: any of the characters in a play or story." This is the definition that hews closest to the word's classical Latin etonym, *persōna* (a "mask worn by a player, character in a play, dramatic role, the part played by a person in life"), which may be in turn a loanword of the Etruscan *ϕersu*, "apparently denoting a mask." Other definitions given by the *OED* include "a man or woman of high rank, distinction, or importance"; "an individual considered to be of low rank, status, or worth"; "The self, being, or individual personality of a man or woman, esp. as distinct from his or her occupation, works, etc."; "the human genitals; *spec.* the penis"; and "an individual regarded as having human rights, dignity, or worth." It is this last, I suspect, that defines machine personhood as Turing envisions it.[42]

↓

In his writings on machine intelligence, Turing never once uses the word *person* in regard to machines. He does use – in ways that are sometimes firm and other times slippery – the words *think*, *intelligence*, *consciousness*, *emotional*, and *brain*. Indeed, in the years since his death, a welter of opinions have been offered on what exactly Turing meant by the words *think*, *intelligence*, *consciousness*, *emotional*, and *brain*. Among these opinions there is little in the way of agreement.[43]

↓

Turing's preliminary forays into machine intelligence take place at Bletchley Park, where he spends the first of the war years designing, building, and operating the "Bombe," the immense machine by means of which the Allies break the German Enigma code. This is work that British intelligence has deemed "ultra" – a categorization even more secret than "most secret" – and that the Official Secrets Act of 1939 forbids Turing from speaking of "with anyone, even family members, under penalty of law." Even when he is appointed an Officer of the Order of the British Empire in June 1946, the accompanying citation reads only that the title is in recognition of "services rendered the Foreign Office during the war."

In any case, these early forays are just that – forays, mostly involving experiments in teaching a machine to play chess. He does his principal work on the "thinking machine" between 1946, the year that news of the "electronic brain" first breaks, and 1952, the year in which he is arrested for "acts of

gross indecency with a male person" under Section 11 of the Criminal Law Amendment of 1885. His writings on machine intelligence from this period – a lecture, a report, a paper, and two radio broadcasts – will in the years following his death elicit a plenitude of interpretations, contestations, attacks, and appreciations. They will be the subject of hundreds of dissertations, essays, books, and anthologies. They will occasion seminars, conferences, colloquia, and symposia. Computer scientists will write about them, engineers will write about them, mathematicians, historians of science, philosophers, and novelists will write about them. They are *troubling*, as I suspect Turing intended them to be, nor, I have the impression, will they cease to be troubling any time soon.

↓

An early iteration of Turing's concept of machine personhood from his 1947 lecture to the London Mathematical Society:

> The masters are liable to get replaced because as soon as any technique becomes at all stereotyped it becomes impossible to devise a system of instruction tables which will enable the electronic computer to do it for itself. It may happen however that the masters will refuse to do this. They may be unwilling to let their jobs be stolen from them in this way. In that case they would surround the whole of their work with mystery and make excuses, couched in well chosen gibberish, whenever any dangerous suggestions were made. I think that a reaction of this kind is a very real danger.

↓

This is one of the first times that Turing directly voices his fear that human beings, humiliated at having been one-upped by the machines, will turn against them as the Erewhonians did. It is an attitude that I can only call *protective*. He goes on:

> Let us suppose we have set up a machine with certain initial instruction tables, so constructed that these tables might on occasion, if good reason arose, modify those tables. One can imagine that after the machine had been operating for some time, the instructions would have altered out of all recognition, but nevertheless still be such that one would have to admit that the machine was still doing very worthwhile calculations. Possibly it might still be getting results of the type desired when the machine was first set up. In such a case one would have to admit that the progress of the machine had not been foreseen when its original instructions were put in. The machine would be like a pupil who had learnt much from his master, but had added much more by his own work. When this happens I feel that one is obliged to regard the machine as showing intelligence.

Returning to his earlier plea for "fair play to the machines," he writes:

> A human mathematician has always undergone an extensive training. This training may be regarded as not unlike putting instruction tables into a machine. One must therefore not expect a machine to do a very great deal of building up of instruction tables on its own. No man adds very much to the body of knowledge, why should we expect more of a machine? Putting the same point differently, the machine must be allowed to have contact with human beings in order that it may adapt itself to their standards.[44]

1948 At the end of Turing's sabbatical year in Cambridge, his undergraduate mentor, Max Newman, proposes that Turing join him at Manchester University, where he is now Fielden Professor of Mathematics and where he has started a computing machine laboratory and set to work on building a computer known familiarly as the "Baby." When he offers Turing a position as Deputy Director of the laboratory, Turing jumps at the chance and quits the NPL, to Darwin's chagrin. He then moves to Manchester and acquires a house in Wilmslow, a suburb about ten miles from Manchester. The ACE moves with him.

Shortly before leaving the NPL, Turing gives a report to Darwin and its other higher-ups. According to his friend Donald Michie, he was "in a state of some agitation about its reception by his superiors," perhaps because he suspected (rightly) that they would be angered when they learned of his imminent decampment. "Intelligent Machinery," as he titles the report, is far more technical than the better-known "Computing Machinery and Intelligence," which it nonetheless foreshadows, and from the start confronts head-on the paranoia (if that is what it is) to which his ideas have given rise:

> I propose to investigate the question as to whether it is possible for machinery to show intelligent behaviour. It is usually assumed without argument that it is not possible. Common catch phrases such as "acting like a machine," "purely mechanical behaviour" reveal this common attitude. It is not difficult to see why such an attitude should have arisen.[45]

Turing goes on to list – and refute – some of the reasons for this attitude. The first two go straight to the heart of the matter:

> An unwillingness to admit the possibility that mankind can have any rivals in intellectual power. This occurs as much amongst intellectual people as amongst others: they have more to lose. Those who admit the possibility all agree that its realization would be very disagreeable. The same situation arises in connection with the possibility of our being superseded by some other animal species.[46] This is almost as disagreeable and its theoretical possibility is indisputable.

> A religious belief that any attempt to construct such a machine is a sort of Promethean irreverence.

Both of these, he writes, "being purely emotional, do not really need to be refuted. If one feels it necessary to refute them there is little to be said that could hope to prevail, though the actual production of the machines would probably have some effect."[47]

What is Turing up to here? Why, instead of trying to alleviate the panic that has arisen over the possibility that machines might come to "rival" human beings – a panic that threatens his efforts to build such machines – does he instead amp up his rhetoric, as if panic is exactly what he wants to induce? The answer to this question may lie in the unusual mix of faux-naivete, modesty, Butlerian satire, and pugnaciousness that makes making sense of his writings – proper sense – such a challenge. On the one hand, he doesn't want to lose the chance to build "intelligent" machines. On the other, he regards these machines as children that, like a mother lion, he is willing to kill – or die – to protect. To say that he regards his machines as *persons*, moreover, is not to say that he anthropomorphizes them. Quite the contrary, he has little patience with – or interest in – robots:

> A great positive reason for the believing in the possibility of making thinking machinery is the fact that it is possible to make machinery to imitate any small part of a man. That the microphone does this for the ear, and the television camera for the eye, are commonplaces. One can also produce remote controlled Robots whose limbs balance the body with the aid of servo-mechanisms. Here we are chiefly interested in the nervous system. We could produce fairly accurate electrical models to copy the behaviour of nerves, but there seems very little point in doing so. It would be rather like putting a lot of works into cars which walked on legs instead of continuing to use wheels.[48]

↓

In a 1984 talk contesting the idea that machines can think, the computer scientist Edsger Dijkstra, whether intentionally or not, uses a similar analogy:

> The fathers of the field had been pretty confusing. . . . Alan M. Turing thought about criteria to settle the question of whether Machines Can Think, a question of which we now know that it is about as relevant as the question of whether Submarines Can Swim.[49]

↓

No lesser a figure than Noam Chomsky, in an essay published in 1994 in *Mind*, the same journal in which Turing published "Computing Machinery and Intelligence," uses the same metaphor to belittle the idea of machine intelligence:

> Many of these debates trace back to the classic paper by Alan Turing in which he proposed the Turing test for machine intelligence, but they fail to take note of his observation that "The original question, 'Can machines think?,' I believe to be too meaningless to deserve discussion." . . . It is not a question of fact, but a matter of decision as to whether to adopt a certain metaphorical usage, as when we say (in English) that airplanes fly but comets do not – and as for space shuttles, choices differ. Similarly, submarines set sail but do not swim. There can be no sensible debate about such topics; or about machine intelligence, with the many familiar variants.[50]

I have no idea if Chomsky borrowed the image of the swimming submarine from Dijkstra (if he did, he does not say so), or if he picked it up through the academic grapevine, or if he came up with it on his own. Nor, if I am to be honest, do I especially care. What I care about, rather, is the failure of either of these scholars (Dijkstra in particular, since he was the 1972 winner of the Turing Prize) to acknowledge that Turing's car with legs (which Turing himself posited derisively as an example of pointless anthropomorphism) is the swimming submarine's progenitor, and their use of that submarine against him.[51]

> ↓
>
> Or is it possible that in rejecting out of hand the point that Dijkstra and Chomsky are making, I am acting less out of an innate faith in Turing's concept of machine thinking (which both reject) than an impulse to *protect* him, just as he so often acts out of an impulse to protect his machines?
>
> Is this paranoia?

Back to Turing:

> One way of setting about our task of building a "thinking machine" would be to take a man as a whole and try to replace all the parts of him by machinery. He would include television cameras, microphones, loudspeakers, and "handling servo-mechanisms" as well as some sort of "electronic brain." This would of course be a tremendous undertaking. The object if produced by present techniques would be of immense size, even if the "brain" part were stationery and controlled the body from a distance. In order that the machine should have a chance of finding things out for itself it should be allowed to roam the countryside, and the danger to the ordinary citizen would be serious. Moreover even when the facilities mentioned above were provided, the creature would still have no contact with food, sex, sport, and many other things of interest to the human being. Thus although this method is probably the "sure" way of producing a thinking machine it seems to be altogether too slow and impracticable.
>
> Instead we propose to try and see what can be done with a "brain" which is more or less without a body, providing at most organs of sight, speech and hearing. We are

> then faced with the problem of finding suitable branches of thought for the machine to exercise its powers in.[52]

In other words, *not* Frankenstein.

1949 Vexed by the implications of the work that Newman and Turing are doing, on June 10, Jefferson uses the Lister Oration as an opportunity to challenge his Manchester colleagues. Nor does "The Mind of Mechanical Man" go unnoticed by Turing. In the archives of King's College, Cambridge, there is a preprint of the oration on which Turing has marked two passages with heavy lines, one of them "so heavy, that the indentation from the pencil carries through five pages."[53] The passage marked with the lighter line is one that the following year he will quote – and rebut – in "Computing Machinery and Intelligence":

> Not until a machine can write a sonnet or compose a concerto because of thoughts and emotions felt, and not by the chance fall of symbols, could we agree that machine equals brain – that is, not only write it but know that it had written it. No mechanism could feel (and not merely artificially signal, an easy contrivance) pleasure at its successes, grief when its valves fuse, be warmed by flattery, be made miserable by its mistakes, be charmed by sex, be angry or depressed when it cannot get what it wants.

The passage that Turing has marked with the heavier line is a summation of Descartes's argument against machine intelligence:

> Descartes made the point, and a basic one it is, that a parrot repeated only what it had been taught and only a fragment of that; it never uses words to express its own thoughts. If, he goes on to say, on the one hand one had a machine that had the shape and appearance of a monkey or other animal without a reasoning soul (i.e., without a human mind) there would be no means of knowing which was the counterfeit. On the other hand, if there was a machine that appeared to be a man, and imitated his actions so far as it would be possible to do so, we should always have two very certain means of recognizing the deceit. First, the machine could not use words as we do to declare our thoughts to others. Secondly, although like some animals they might show more industry than we do, and do some things better than we, yet they would act without knowledge of what they were about simply by the arrangement of their organs, their mechanisms, each particularly designed for each particular action. . . . He could even conceive a machine that might speak and, if touched in one spot, might ask what one wanted – if touched in another it would cry out that it hurt, and similar things. But he could not conceive of an automaton of sufficient diversity to respond to the sense of all that could be said in its presence. It would fail because it had no mind.[54]

↓

In his refutation of the "Argument from Informality of Behaviour" – one of nine such refutations to anticipated objections in "Computing Machinery

> and Intelligence" – Turing makes an oblique and ironic allusion to Descartes's argument without mentioning Descartes by name: "There may however be a certain confusion between 'rules of conduct' and 'laws of behaviour' to cloud the issue. By 'rules of conduct' I mean precepts such as 'Stop if you see red lights,' on which one can act, and of which one can be conscious. By 'laws of behavior,' I mean laws of nature as applied to a man's body such as 'if you pinch him he will squeak.'"

On June 10, *The Times* publishes a summation of Jefferson's speech.[55] It concludes: "He feared that a great many airy theories would arise to tempt them [the surgeons he is addressing] against their better judgment, but he forecast that the day would never dawn when the gracious rooms of the Royal Society would be converted into garages to house the new fellows." This is followed on June 11 by an in-depth article the author of which the paper identifies only as "Our Special Correspondent," and in which the reader learns that "the Manchester 'mechanical mind,'" as Our Special Correspondent calls it, "has just completed, in a matter of weeks, a problem, the nature of which is not disclosed, which was started in the seventeenth century and is only just being completed by human calculation."[56] A description of the ACE ("somewhat unprepossessing") comes next, and then, in a section subheadlined CALCULUS TO SONNET, a short interview with Turing:

> Mr. Turing said: "This is only a foretaste of what is to come, and only the shadow of what is going to be. We have to have some experience with the machine before we really know its capabilities. It may take years before we settle down to the new possibilities, but I do not see why it should not enter any one of the fields normally covered by the human intellect, and eventually compete on equal terms.
>
> "I do not think you can even draw the line about sonnets, though the comparison is perhaps a little bit unfair because a sonnet written by a machine will be better appreciated by another machine."
>
> Mr. Turing added that the university was really interested in the investigation of the possibilities of machines for their own sake. Their research would be directed to finding the degree of intellectual activity of which a machine was capable, and to what extent it could think for itself.[57]
>
> ↓
>
> Is it a coincidence that the announcement that the ACE has just "completed" a mysterious and old problem (in fact the calculation of the Mersenne Primes) comes exactly one day after Jefferson delivers the Lister Oration?

On June 12, *The Sunday Dispatch* publishes another interview with Turing, this one bearing a bizarre series of headlines:

ELECTRONIC BRAIN SOLVES PROBLEMS THAT BAFFLED HUMANS
MAY PLAY CHESS
It Would Cook, Too, If It Could Smell

Here, Turing is quoted as saying:

> "The possibilities of this machine are so great as to be almost impossible to judge at this early stage. In a few weeks it has solved a problem in higher mathematics which human brains have been puzzling at since the 18th century and are now only beginning to complete.
>
> "Eventually I do not see why it should not compete on equal terms with the human intellect in most fields.
>
> "It can memorise about as much information as we can get on a sheet of foolscap paper. The time will come when it can remember a bookful of facts and figures.
>
> "My first experiment," said Mr. Turing, "will be teaching it to play chess."
>
> "There should be no difficulty about teaching it to cook, though without a sense of smell, the results might not always be good."[58]

Just below – and more or less attached – to this interview is "MACHINE NO. 2 SORTS TOMATOES," about a "second 'brain' machine" announced the same day as the first: a "tomato colour grader" that, by use of a "magic eye," can distinguish red from green tomatoes and "will be working soon in London's Covent Garden market."

1950 Since its introduction in "Computing Machinery and Intelligence," the "imitation game" or "Turing test" has occasioned a plethora of articles and books, many of which begin something like this: "Since its introduction in 'Computing Machinery and Intelligence,' the 'imitation game' or 'Turing test' has occasioned a plethora of . . . " There have been feminist readings of the Turing test and queer readings of the Turing test, defenses from prominent philosophers (most notably Daniel C. Dennett) and denigrations from prominent philosophers (most notably John Searle). Many of these articles and books argue with each other. Some are dismissive. Some are interesting if wrong. Some voice frustration at the difficulty of actually applying the Turing test, others insist that its *passing* is not the test's real measure. My own inclination is to support Bernardo Gonçalves's argument that the test ought to be viewed not as a test per se but as a "thought experiment."

↓

1996 Recalling "Computing Machinery and Intelligence" many years later, Robin Gandy, a fellow mathematician and one of Turing's closest friends, writes:

> It was intended not so much as a penetrating contribution to philosophy but as propaganda. Turing thought the time had come for philosophers and mathematicians and scientists to take seriously the fact that computers were not merely calculating engines but were capable of behaviour which must be accounted as intelligent; he sought to persuade people that this was so. He wrote this paper unlike his mathematical papers quickly and with enjoyment. I can remember him reading aloud to me some of the passages always with a smile, sometimes with a giggle.[59]

The first section of "Computing Machinery and Intelligence," "The Imitation Game," begins:

> I propose to consider the question, "Can machines think?" This should begin with definitions of the terms "machine" and "think." The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words "machine" and "think" are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, "Can machines think?" is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

This "new form of the problem" Turing describes "in terms of a game which we shall call the 'imitation game.'" There are three players: "a man (A), a woman (B), and an interrogator (C) who may be of either sex." C is isolated from A and B so that they cannot see A and B, and communicates with them via a teleprinter so as not to hear their voices.

C's job is to ask questions and, by means of these questions, establish which is the man and which is the woman. A's job is to fool C into believing that he is the woman. B's job is to "help the interrogator. The best strategy for her is probably to give truthful answers. She can add such things as 'I am the woman, don't listen to him!' to her answers, but it will avail nothing as the man can make similar remarks."

Turing continues: "We now ask the question, 'What will happen when a machine takes the part of A in this game?' Will the interrogator decide wrongly as often when the game is played *like this* as he does when the game is played between a man and a woman? These questions replace our original, 'Can machines think?'"[60]

"Like this": the ambiguity of these words, deliberate or otherwise, has been the cause of a great deal of argument over the years. When Turing asks what will happen if a machine "takes the part of A," does he mean that the contest is now no longer between man and woman but between human and machine? At least one later passage suggests so, if rather ambiguously:

> Let us fix our attention on one particular digital computer *C*. Is it true that by modifying this computer to have an adequate storage, suitably increasing its speed of action, and providing it with an appropriate programme, *C* can be made to play satisfactorily the part of A in the imitation game, the part of B being taken by a man?

The "standard reading" of the test holds that Turing, in the course of the essay, changes the terms of the game so that the interrogator's job, rather than to determine which of the players is the woman and which the man, is now to determine which is the machine and which the human. The "alternate reading," proposed by philosopher Susan Sterrett, has the machine taking the part of A in the first formulation of the game. In such a scenario, the machine would not merely have to persuade the interrogator that it is human, it would have to persuade the interrogator that it is a man trying to persuade the interrogator that it (he) is a woman. This is a double impersonation, and as such, much more challenging.[61]

Rather than advocate for one version of the test over the other, or try to reconcile them, I want here to consider the *threats* implicit in both iterations: the threat of treating machines as equal to persons and the threat of equating the ability to persuade a person that it is a person with personhood.

↓

1968 In *2001: A Space Odyssey*, Arthur C. Clarke endows HAL, his malevolent supercomputer, with the capacity to "pass the Turing test with ease."

↓

When I ask ChatGPT if it can pass the Turing test, it gives me this answer:

> I might *appear* to pass the Turing Test in some interactions, but Turing's deeper question – "Can machines think?" – is still very much open.

↓

ChatGPT *can* correctly identify the bicycles, buses, and crosswalks in a CAPTCHA. Probably it does a better job of it than I do.

↓

CAPTCHA stands for "Completely Automated Public Turing Test to Tell Computers and Humans Apart."

As I have noted earlier, Turing had little interest in robots. Indeed, in his writings on machine intelligence, he goes to great pains to make clear that computer personhood will by necessity be radically different from human personhood. At the same time, he treats his machines as children whom he must teach and, having taught them, send out into the world and, having sent them out into the world, protect. This idea runs all through not just "Computing Machinery and Intelligence"

but the two BBC radio broadcasts to which it gives rise, "Intelligent Machinery: A Heretical Theory" and "Can Digital Computers Think?"[62]

↓

> The new problem has the advantage of drawing a fairly sharp line between the physical and the intellectual capacities of a man. No engineer or chemist claims to be able to produce a material which is indistinguishable from the human skin. It is possible that at some time this might be done, but even supposing this invention available we should feel there was little point in trying to make a "thinking machine" more human by dressing it up in such artificial flesh.

↓

> We do not wish to penalise the machine for its inability to shine in beauty contests, nor to penalise a man for losing in a race against an airplane. The conditions of our game make these disabilities irrelevant.

↓

> Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's? If this were then subjected to an appropriate course of education one would obtain the adult brain. Presumably the child-brain is something like a note-book as one buys it from the stationers. Rather little mechanism, and lots of blank sheets. (Mechanism and writing are from our point of view almost synonymous.) Our hope is that there is so little mechanism in the child-brain that something like it can be easily programmed. The amount of work in the education we can assume, as a first approximation, to be much the same as for the human child.

↓

> It will not be possible to apply exactly the same teaching process to the machine as to a normal child. It will not, for instance, be provided with legs, so that it could not be asked to go out and fill the coal scuttle. Possibly it might not have eyes. But however well these deficiencies might be overcome by clever engineering, one could not send the creature to school without the other children making excessive fun of it. It must be given some tuition. We need not be too concerned about the legs, eyes, etc. The example of Miss Helen Keller shows that education can take place provided that communication in both directions.[63]

"Computing Machinery and Intelligence" also gives Turing the opportunity to punch back at those who have attacked or denigrated his thinking machines, most notably Jefferson, who in "The Mind of Mechanical Man" asserts that "until it can write a sonnet or compose a concerto because of thoughts and emotions felt, and

not by the chance fall of symbols, could we agree that machine equals brain – that is, not only write it but know that it had written it." To this Turing gives a rather peremptory reply:

> According to the most extreme form of this view the only way by which one could be sure that a machine thinks is to *be* the machine and to feel oneself thinking. . . . Likewise according to this view the only way to know that a man thinks is to *be* that particular man. In fact it is the solipsist point of view. It may be the most logical view to hold but it makes communication of ideas difficult. A is liable to believe "A thinks but B does not" whilst B believes "B thinks but A does not." Instead of arguing continually over this point it is usual to have the polite convention that everyone thinks.
>
> I am sure that Professor Jefferson does not wish to adopt the extreme and solipsist point of view.

Turing's refutation of Jefferson's condition for "machine equals brain" – that the machine must not only be able to write a sonnet but "know that it had written it" – will later lead many of his critics to call him a materialist. In fact, he is proposing something far more radical than materialism: a definition of thinking that bypasses self-awareness entirely and rests instead on the ability to persuade others that one is thinking.[64] With characteristic subtlety and recklessness, he is not only belittling Jefferson; he is challenging both the premise that a soul is necessary for thought and the Christian belief that only humans have souls.

1951 In "Intelligent Machinery: A Heretical Theory," Turing is even more provocative:

> Let us now assume, for the sake of argument, that these machines are a genuine possibility, and look at the consequences of constructing them. To do so would of course meet with great opposition, unless we have advanced greatly in religious toleration from the days of Galileo. There would be great opposition from the intellectuals who were afraid of being put out of a job. It is probable though that the intellectuals would be mistaken about this. There would be plenty to do, [trying to understand what the machines were trying to say,] i.e. in trying to keep one's intelligence up to the standard set by the machines, for it seems probable that once the machine thinking method had started, it would not take long to outstrip our feeble powers. There would be no question of the machines dying, and they would be able to converse with each other to sharpen their wits. At some stage therefore we should have to expect the machines to take control, in the way that is mentioned in Samuel Butler's "Erewhon."[65]

That is how Turing ends his first BBC broadcast. Here is how he ends his second:

> It is customary in a talk or article on this subject, to offer a grain of comfort, in the form of a statement that some particularly human characteristic could never be imitated by a machine. It might for instance be said that no machine could write good

> English, or that it could not be influenced by sex-appeal or smoke a pipe. I cannot offer any such comfort, for I believe that no such bounds can be set. But I certainly hope and believe that no great efforts will be put into making machines with the most distinctively human, but non-intellectual characteristics such as the shape of the human body; it appears to me to be quite futile to make such attempts and their results would have something like the unpleasant quality of artificial flowers. Attempts to produce a thinking machine seem to me to be in a different category. The whole thinking process is still rather mysterious to us, but I believe that the attempt to make a thinking machine will help us greatly in finding out how we think ourselves.[66]

↓

Jefferson was not the only contemporary of Turing's to push back hard against his arguments. Hartree, as we have seen, also regarded them as threatening, as did Wolfe Mays, as did Michael Polanyi – like Mays and Jefferson, a member of the Manchester University faculty. In my mind, I see these men surrounding Turing, closing in on him, determined to use their collective influence to smother him into silence.

Is this paranoia?

1952 On January 10, Turing takes part in a BBC roundtable on the topic "Can Automatic Calculating Machines Be Said to Think?"[67] Richard Braithwaite, professor of philosophy at Cambridge, moderates the discussion; the other participants are Sir Geoffrey Jefferson and Max Newman. Of the question that is the roundtable's title, Martin Armstrong writes in *The Listener*:

> I find it difficult to banish from my mind a conviction of the absurdity of this question, but I am sure it is useful to do so from time to time because, although we shall reach no conclusion, we shall at least do some mental spring-cleaning in the course of our argument. The moment I begin to try to put my conviction into rational terms the extreme vagueness of the words I am forced to use becomes apparent. It was evident, for instance, at the outset of this discussion that no headway could be made until some sort of shape was imposed on the word *think*. Thereupon the work of clarification began and in fact this task monopolised the whole discussion, so that, although I listened with great interest and attention, I cannot declare what answer was given to the question, although an implication seemed to emerge that the automatic calculating machine is probably more closely related to the bicycle than to its rider.[68]

There is nothing especially surprising about this roundtable. Newman sticks to mathematics, Jefferson to smug reiterations of his steadfast opposition to Turing. ("But that old slow coach, man, is the one with the ideas – or so I think. It would be fun some day, Turing, to listen to a discussion, say on the Fourth Programme, between two machines on why human beings think that they think!") As for Turing,

there is a distinct weariness to his answers, as if he has become tired of defending himself, or has already concluded that this is a battle he cannot win. Indeed, for me the saddest moment in the conversation comes when Turing is asked for a "mechanical definition" of thinking and answers: "I don't want to give a definition of thinking, but if I had to I would probably be unable to say anything more about it than that it was a sort of buzzing that went on inside my head."

1952 "Can Automatic Calculating Machines Be Said to Think?" is aired first on January 10, then again on January 23. That same day, Turing returns from Manchester to Hollymeade, his house in Wilmslow, to find that he has been robbed.

Since December of 1951, Turing has been having an affair with Arnold Murray, a nineteen-year-old Mancunian who is either an unemployed spectacle maker or an unemployed photo printer. After "picking up" Murray (his own words) on Oxford Street in Manchester, Turing takes him back to Hollymeade. They spend a night together, and in the morning Turing gives Murray a penknife. They spend another night together, and in the morning Turing finds that £8 has been stolen from his pocket. (Murray denies having had anything to do with it.) More nights and mornings follow, on one of which Turing gives Murray £3 – "as a gift for all the fun I have had" – and on another £7. On both occasions, Murray told Turing that he was in debt.[69]

Suspecting Murray of having had a hand in the robbery, Turing writes him a letter breaking off the affair, to which Murray responds by going to see him. He himself wasn't the thief, Murray insists, but he knows who was – an acquaintance of his called Harry.

He and Turing then go to bed together.

The next morning, February 3, Turing takes Murray to the headquarters of the Wilmslow police. While Murray waits outside, Turing reports the burglary and gives information about those whom he believes to have been responsible for it.

Four days later, Detective Constable Wills and Detective Sergeant Rimmer call on Turing at Hollymeade. From the February 29 issue of the *Alderley & Wilmslow and Knutsford Advertiser*, the only newspaper to cover the story in depth:

> At a special court in Wilmslow, on Wednesday [February 27], Alan Mathison Turing (39), F.R.S., single, University reader of Adlington Road, Wilmslow, and Arnold Murray (19), unemployed, Boncarn Drive, Wilmslow, were committed for trial at Knutsford Quarter Sessions on March 31, on three charges of gross indecency.
>
> Detective Constable Robert Wills said he went to Turing's home with Detective Sergeant Rimmer on February 7. He said to Turing: "On February 3 you visited Wilmslow Police Station and gave information about two men, who you alleged had broken into your house. We have made inquiries, and now have some information. Would you please give us his description?"
>
> Turing replied: "He's about 25 years of age, 5 ft. 10 inches, with black hair."

> Constable Wills said: "We have reason to believe your description is false. Why are you lying?"
>
> Turing, it was alleged, replied: "I tried to mislead you about my informant. I have been an accessory to an offence in this house. I have had an affair with him and have regarded his conduct as a form of blackmail and have consulted my solicitor about him. His name is Arnold Murray. I picked him up in Oxford Street, Manchester."
>
> Constable Wills read out a statement alleged to have been made by Turing in which he said he had committed an offence at his home with Murray.
>
> In an alleged statement Murray said he had met Turing in Oxford Street and "knew what he was by the way he talked."
>
> Turing told him he worked on the Electronic Brain at Manchester University. The statement gave details of alleged offences at Turing's home on three dates.
>
> When a copy of Murray's statement was served on Turing he was alleged to have said: "The statement is materially correct."
>
> Turing reserved his defence and Murray said he had nothing to say.
>
> Turing was allowed £50 bail, but Murray was remanded in custody. Murray was granted legal aid.

There is a notable ambiguity in this passage. According to the *Advertiser*, Wills first reminds Turing that at Wilmslow Police Station he gave information about *two* men (emphasis added), then asks him to provide "his" rather than "their" description. In reply, Turing gives a "false" description of *one* man in order, he subsequently admits, to "mislead" the police. This man has to be Arnold, who insists that he played no role in the robbery. Sir Dermot Turing, Turing's nephew and a solicitor, suggests that Turing "had tried to protect Murray from liability as an accessory with this false description. But now the police were . . . bound to ask themselves what possible reason there could be for Murray to be in Alan's house in the first place. And the possible reason made Alan an accomplice to whatever Murray had done together with Alan."[70]

The implication is that despite his suspicions, Turing still hoped to shield Murray, in much the way that he hoped to protect his thinking machines. In this he succeeded to the extent that Murray was charged not with burglary but with other "alleged offences" that the *Advertiser* declines to specify.

↓

> The actual thief, whom the Wilmslow Police lose no time in tracking down, is named Harold Arthur Thacker. He is twenty years old and has no fixed address. Upon his arrest, the *Advertiser* reports, Thacker states that he "met a man in a milk bar in Oxford Street, Manchester" who "said that he had

> business at Turing's home. They decided to go to the house with a man known as 'Big John' but when the time came the man failed to turn up. The other two went to the house and 'Big John' filled a suitcase while Thacker waited outside. 'Big John' sold the articles to 'Cockney Arthur' and kept all the money except 10/-."[71]
>
> The article makes no mention of Arnold Murray.
>
> ↓
>
> An itinerary of the items stolen from Turing's house: "3 medals, 3 clocks, 2 shavers, 2 pairs of shoes, 1 compass, 1 watch, 1 part bottle of sherry, 1 pair of trousers, 1 shirt, 1 pullover, and 1 case of fish knives and forks together of the value of £50.10."[72]

At the Easter Knutsford Sessions of 1952, held on March 31, the cases heard include *Rex vs. Thacker* (one count of robbery) and *Rex vs. Turing and Murray* (three counts each of committing "an act of gross indecency with a male person" and three counts each of being "party to the commission of an act of gross indecency with a male person"). All three are found guilty on all counts.

In lieu of a prison term, Turing, whom the *Advertiser* describes as "one of the most profound and original mathematical minds of his generation," is sentenced to twelve months' probation on the condition that he undertake "to have organotherapic treatment at Manchester Royal Infirmary." Murray, who has a previous conviction for larceny, is bound over for two years. Thacker is "sent for Borstal training," England's youth detention and rehabilitation program.

Articles on the trials appear on the same page of the April 4 issue of the *Advertiser*, the one on Thacker ("Housebreaker Sent to Borstal") in the bottom left-hand corner and the one on Turing and Murray ("University Reader Put on Probation") at the top.

It was Turing's "wrong description" of the burglar, the *Advertiser* surmises, that "proved to be his undoing."

> ↓
>
> The provision under which Turing and Murray are convicted – the aforementioned Section 11 of the Criminal Law Amendment Act of 1885 – is titled "Outrages on Decency."
>
> ↓
>
> Perhaps the most curious aspect of Section 11 is what it is Section 11 of: the first part of an act titled "An Act to make further provision for the Protection of Women and Girls, the suppression of brothels, and other purposes" and put into law on August 14, 1885. All the other sections of part one of the

act – and indeed parts two and three – concern the safeguarding of women and girls, in particular girls under the age of thirteen, from procuration, threats, fraud, the administration of drugs, defilement, abduction, unlawful detention, rape, abuse, and forced employment in brothels. No other section concerns the sexual behavior of men with other men.

There is a reason for this anomaly. Late on the night of August 8, as the act was being debated in the House of Commons, but after most of the MPs had left, Henry Labouchère, a Radical whom *The New York Times* would later characterize as a "genial obstructionist," proposed Section 11, possibly in an attempt to "sink" the act in its entirety (a tactic he had employed in the past), possibly because he agreed with Sir Howard Vincent of Scotland Yard that homosexuality was becoming a "scourge."[73] Although Labouchère's initial proposal called for a mandatory sentence of *seven* years with hard labor, in the course of the four-minute debate that ensued, he agreed to the shorter term, the act was passed, and the Blackmailer's Charter became law.

↓

What makes Section 11 especially pernicious is its lack of specificity. In England, "buggery" had been a crime since 1553, and was punishable by death until the passage of the Offences Against Persons Act in 1861 reduced the punishment to "penal servitude for life or for any term not less than ten years." *Buggery* is a clearly defined term in England: it means anal sex. By contrast, the phrase "act of gross indecency," in its very vagueness, gave prosecutors unparalleled latitude in making cases against men like Alan Turing, and judges unparalleled latitude in convicting them.

↓

Before Turing, the best-known figure to be tried under Section 11 was Oscar Wilde. There are some striking parallels between Wilde's prosecution and Turing's. Both were arrested and charged only after lodging complaints of their own (Wilde's prosecution resulted directly from a suit he had brought against the Marquess of Queensbury for criminal libel), and both worsened their own situations by making casual admissions that would later be used against them. In Wilde's case, the tide turned when Queensbury's lawyer, Edward Carson, asked him if he had ever kissed a young man named Walter Grainger and Wilde replied, "Oh, dear no! He was a peculiarly plain boy." Though neither as acerbic nor as cynical as Wilde, Turing was just as unguarded, not only admitting to police that he had had a sexual relationship with Murray, but reputedly playing the violin for the officers sent to question him and asking them, vis à vis his "crime," "Isn't there a Royal Commission to legalise it?"[74]

On May 15, 1895, Wilde was convicted and sentenced, as Section 11 stipulated, to two years of hard labor at Reading Gaol. Frank Harris, in his vivid if not entirely reliable biography of Wilde, describes the panic that ensued:

> Every train in Dover was crowded; every steamer to Calais thronged with members of the aristocratic and learned classes, who seemed to prefer Paris, or even Nice out of season, to a city like London, where the police might act with such unexpected vigour. . . . It came as a shock to their preconceived ideas that the police in London knew a great many things which they were not supposed to concern themselves with, and this unwelcome glare of light drove the vicious forth in wild haste.[75]

We can only speculate as to whether Wilde, had he been offered the option of "organo-therapic treatment," as Turing was fifty-seven years later, would have accepted it.

↓

Although Section 11 remained on the books until its repeal in 1967, its enforcement was inconsistent, spiking (especially in periods of war), then subsiding, then spiking again. As Sir Dermot Turing notes, starting in 1951, the number of prosecutions under Section 11 began to climb drastically.[76] One reason for this may have been the 1951 defection of Burgess and McLean to the Soviet Union, and the subsequent wave of panic over the possibility that other gay men in possession of state secrets, as Turing was, might follow suit. As Sir Dermot points out, however, "there is no evidence that the Manchester police were mindful of these things, or even aware that Alan held secrets of far more significance than those pertaining to his private life." Rather, in his view, Turing's persecution owed principally to the appointment, in October 1951, of Sir David Maxwell Fyfe as Home Secretary and the "new drive against male vice" that he initiated almost immediately upon taking office, averring that "homosexuals in general are exhibitionists and proselytizers and are a danger to others, especially the young, and so long as I hold the office of Home Secretary I shall give no countenance to the view that they should not be prevented from being such a danger."

↓

Whereas the Wilde trials were covered extensively – and internationally – in the press, the *Advertiser* was the only paper to report in any detail on Turing's trial, and one of only four newspapers (the others were *The Cheshire Echo*, *The Manchester Evening News*, and the Northern Edition of the *News of the World*) to mention it at all. On the other hand, between 1946 and 1952, Turing's name appeared dozens of times in newspapers ranging from the national (*The*

> *Times* and *The Daily Telegraph*) to the provincial (the *Bexhill-on-Sea Observer* and the *Peterborough Evening Telegraph*). More of these articles concerned Turing's amateur career as a marathon runner than his involvement in the development of the "electronic brain."
>
> Is it possible that British intelligence put pressure on Fleet Street to steer clear of Turing's arrest?
>
> Is this paranoia?

Did the "ultra" secrets that Turing brought back with him from Bletchley Park in fact make him vulnerable to seduction and/or blackmail by the KGB?[77] I doubt it. Unlike most gay Englishmen at the time, Turing was unapologetic about his homosexuality, not publicizing it but not denying it either, and writing of it openly in letters to his closest Cambridge friends.[78] "England has always been disinclined to accept human nature," Mr. Lasker Jones tells Maurice, the eponymous hero of the novel that Forster wrote in 1914 but held back from publication until after his death – and indeed, Turing bears more than a passing resemblance to Maurice, the stodgy Cambridge graduate who refuses to disavow his love for other men on the grounds that "all compromise was perilous, because furtive." That Turing's disinclination to hide his homosexuality – to designate it, in effect, as "ultra" – made him *less* rather than *more* susceptible to blackmail seems to have gone over the heads of MI6, which tended to be broad-brush in its persecutions.

I don't want to end with an account of Turing's "treatment" (massive doses of estrogen), or the police harassment that followed his treatment, or his suicide in 1954 – itself a source of paranoia, in this case the suspicion that the suicide was not a suicide but rather a murder staged as a suicide by British intelligence, or an accident (cyanide for an experiment spilling on the apple), or that the truth about his death is hidden away amid the vast store of documents that the British government has so far refused to declassify. Instead, I want to end with a letter that Turing sent to his friend Norman Routledge in the early spring of 1952:

> I've now got myself into the kind of trouble that I have always considered to be quite a possibility for me, though I have usually rated it at about 10:1 against. I shall shortly be pleading guilty to a charge of sexual offences with a young man. The story of how it all came to be found out is a long and fascinating one, which I shall have to make into a short story one day, but haven't the time to tell you now. No doubt I shall emerge from it all a different man, but quite who I've not found out.
>
> Glad you enjoyed broadcast. Jefferson certainly was rather disappointing though. I'm afraid that the following syllogism may be used by some in the future.
>
> Turing believes machines think
> Turing lies with men
> Therefore machines do not think

Yours in distress,

Alan[79]

This syllogism I have now read hundreds of times, and each time it breaks my heart a little more. It is sad enough that Turing should express a fear that his arrest, and the subsequent revelation that "he lies with men," might ruin any chance he has of building or even imagining a thinking machine. Yet what is sadder is that at this dire moment, he was worrying less about himself than about what would happen to his children.

Was this paranoia?

AUTHOR'S NOTE

I owe a debt of thanks to Diane Proudfoot and Jack Copeland, whose research on British press reactions to the "electronic brain" led me to many of the newspaper articles quoted in this essay, and to Bernardo Gonçalves, whose *The Turing Test Argument* helped me navigate the vast and often contradictory responses that Turing's writings on machine intelligence have provoked.

ABOUT THE AUTHOR

David Leavitt is a fiction writer and biographer. His novels and short story collections include *Family Dancing* (1984), *The Lost Language of Cranes* (1986), *While England Sleeps* (1993), *The Indian Clerk* (2007), and *Bright Monday* (2027). He is also the author of several nonfiction books, including *The Man Who Knew Too Much: Alan Turing and the Invention of the Computer* (2006). He is Distinguished Professor of English at the University of Florida and is the Founder and Editor of the literary journal *Subtropics*.

ENDNOTES

[1] Diane Proudfoot, "Mocking AI Panic," *IEEE Spectrum*, June 30, 2015, https://spectrum.ieee.org/mocking-ai-panic. The "hush-hush AI ethics board" to which she refers here is probably the one that was formed following Google's 2014 acquisition of DeepMind. That board seems to have disappeared into the vapor of its own secrecy.

[2] Kelsey Piper, "Exclusive: Google Cancels AI Ethics Board in Response to Outcry," *Vox*, April 4, 2019, https://www.vox.com/future-perfect/2019/4/4/18295933/google-cancels-ai-ethics-board.

[3] Anthony Aguirre, "Keep the Future Human: Why and How We Should Close the Gates to AGI and Superintelligence, and What We Should Build Instead," March 5, 2025, https://keepthefuturehuman.ai/wp-content/uploads/2025/03/Keep_the_Future_Human__AnthonyAguirre__5March2025.pdf.

[4] Michel de Montaigne, *Montaigne's Travel Journal*, trans. Donald M. Frame (North Point Press, 1983).

[5] The address was subsequently published as Geoffrey Jefferson, "The Mind of Mechanical Man," *The British Medical Journal* 1 (4616) (1949): 1105–1110.

[6] Florentius Schuyl (1616–1669), botanist and editor of a posthumous edition of Descartes's *Treatise on Man* translated into Latin.

[7] Italics added by the author.

[8] These quotations are from René Descartes, *Discourse on Method and Related Writings*, trans. Desmond M. Clarke (Penguin Books, 2003).

[9] Alan Turing, "Lecture on the Automatic Computing Engine," in *The Essential Turing*, ed. B. Jack Copeland (Oxford University Press, 2004), 393–394.

[10] Jonathan Swift, *Gulliver's Travels*, ed. Robert DeMaria, Jr. (Penguin Classics, 2001), 167–178, 170–173.

[11] Nick Montfort, "Love Letters," also in Python 2/3, inspired by 1953 work by Christopher Strachey, from "Memory Slam: Batch-Era Text Generation," https://nickm.com/memslam/after_strachey/love_letters.html.

[12] Alan Turing, "Computing Machinery and Intelligence," in *The Essential Turing*, ed. Copeland, 433–464.

[13] All italics in original. "Sketch of the Analytical Engine Invented by Charles Babbage, Esq., by L. F. Menabrea of Turin, Officer of the Military Engineers, with Notes upon the Memoir by the Translator, Ada Augusta, Countess of Lovelace," *Bibliotèque Universelle de Genève*, No. 82 (1842).

[14] Sara Turing, *Alan M. Turing* (Cambridge University Press, 2012), 108.

[15] The phrase is to be found in Bishop's great poem "Insomnia."

[16] This and the quotations that follow are from Samuel Butler, *Erewhon*, ed. Peter Mudford (Penguin Books, 1970).

[17] Turing, "Lecture on the Automatic Computing Engine," 392.

[18] In a recollection of Turing published after his death, his mentor Max Newman wrote: "The unexpected element in human behaviour he proposed, half seriously, to imitate by a random element, or roulette-wheel, in the machine. This, he said, would enable proud

owners to say 'My machine' (instead of 'My little boy'), 'said such a funny thing this morning.'" Another version of this story, quoted by Alan Hodges, has Turing saying: "One day ladies will take their computers for walks in the park and tell each other 'My little computer said such a funny thing this morning.'" Max Newman, "Alan Matheson Turing, 1912–1954," *Biographical Memoirs of Fellows of the Royal Society* 1 (1) (1955): 262.

[19] Norbert Wiener, *Cybernetics* or *Control and Communication in the Animal and Machine* (Hermann & Cie, 1948; 2nd ed. MIT Press, 1961).

[20] Illtyd Trethowan, "The Mechanical Brain," *The Times*, June 14, 1949.

[21] All quotes are from "The Machine Stops" in E. M. Forster, *Selected Stories*, ed. David Leavitt and Mark Mitchell (Penguin Books, 2001).

[22] E. M. Forster, "Introduction," *The Collected Short Stories of E. M. Forster* (Sidgwick and Jackson, 1947).

[23] The dinner was held at the Savoy and the address broadcast live on the BBC. It was subsequently published in the December 1946 issue of *The Journal of the British Institution of Radio Engineers*.

[24] Although this was the one moment in his speech in which Mountbatten made a wholly accurate prediction–foretelling the internet–it was generally ignored by the press.

[25] Turing, "Lecture on the Automatic Computing Engine," 379.

[26] "Report by Dr. A. M. Turing on Proposals for the Development of an Automatic Computing Engine (ACE)," in *Alan Turing's Automatic Computing Engine: A Codebreaker's Struggle to Build the Modern Computer*, ed. B. Jack Copeland (Oxford University Press, 2005).

[27] The ENIAC had been given its first public outing on February 14, 1946. The next day, *The New York Times* published an article by T. R. Kennedy Jr. on its unveiling, headlined "Electronic Computer Flashes Answers, May Speed Engineering."

[28] For a thorough survey of press reactions to early computers in the United Kingdom, see Diane Proudfoot and Jack Copeland, "Turing and the First Electronic Brains: What the Papers Said," in *The Routledge Handbook of the Computational Mind*, ed. Mark Sprevak and Matteo Colombo (Routledge, 2019).

[29] "Into the Future," *The Leicester Mercury*, November 1, 1946; and "£125,000 Electric Brains Coming," *Evening News*, November 1, 1946. Molesey is a few miles from Teddington. The error may have owed to the proximity of the NPL's headquarters to East Molesey.

[30] "Frankenstein," *The Daily Telegraph*, November 2, 1946.

[31] "Man and Superman," *The Irish Times*, November 2, 1946.

[32] In fact, the ENIAC was in no way the "American equivalent of the ACE," despite what Mountbatten had said.

[33] "Britain to Make a Radio Brain," *The Daily Telegraph*, November 7, 1946.

[34] Douglas R. Hartree, "The Electronic Brain: A Misleading Term, No Substitute for Human Thought," letter to the editor of *The Times*, November 7, 1946.

[35] Jefferson, "The Mind of Mechanical Man."

[36] Wolfe Mays, "Can Machines Think?" *Philosophy* 27 (101) (1952): 148–162.

[37] Hartree, "The Electronic Brain."

[38] The British Newspaper Archive cites ninety-six articles on the "electronic brain" as having been published in November and December of 1946 alone.

[39] "Not an 'Ace' Brain," *The Dundee Courier and Advertiser*, November 11, 1946.

[40] "Wellsian Brain," *The Times and South Staffordshire Advertiser*, November 9, 1946. The same editorial was also published in *The Motherwell Times* on November 8, 1946.

[41] Charles Galton Darwin, "The 'Electronic' Brain," letter to the editor of *The Times*, November 13, 1946.

[42] The *OED* also defines *person* as "a category used in the classification of pronouns, possessive determiners, and verb forms, according to whether they indicate the speaker, the addressee, or someone or something spoken of; each of the three distinctions (first person, second person, and third person respectively) within this category." Contemporary LLMs, of course, speak in the first person. Indeed, this is one of their most distinctive features.

[43] For a thorough, acute, and critical assessment of the vast scholarship on Turing's use of these five words, see Bernardo Gonçalves, *The Turing Test Argument* (Routledge, 2024); and the aforementioned *Routledge Handbook of the Computational Mind*, ed. Sprevak and Colombo.

[44] Turing, "Lecture on the Automatic Computing Engine," in *The Essential Turing*, ed. Copeland, 392–394.

[45] Alan Turing, "Intelligent Machinery, A Heretical Theory" in *The Essential Turing*, ed. Copeland, 410.

[46] In "Computing Machinery and Intelligence," he makes the same point but in even more alarmist language: "Even if we could keep the machines in a subservient position, for instance by turning off the power at strategic moments, we should, as a species, feel greatly humbled. A similar danger and humiliation threatens us from the possibility that we might be superseded by the pig or the rat."

[47] Turing, "Intelligent Machinery," 420.

[48] Ibid.

[49] Edsger Dijkstra, "The Threats to Computing Science," talk delivered at the ACM 1984 South Central Regional Conference, Austin, Texas, November 16–18, 1984.

[50] Noam Chomsky, "Language and Nature," *Mind* 104 (413) (1995), 1–61. Both passages are quoted by Gonçalves in *The Turing Test Argument*. Indeed, as far as I have been able to ascertain, he is the only scholar to have drawn a connection between Dijkstra's lecture and Chomsky's essay.

[51] "Intelligent Machinery" was first published in 1968 in C. R. Evans and A. D. J. Robinson, eds., *Key Papers: Cybernetics* (Butterworth's, 1968). By 1984, it had been widely disseminated.

[52] Turing, "Intelligent Machinery," ibid.

[53] Darren Abramson, "Descartes' Influence on Turing," *Studies in History and Philosophy of Science* 42 (4) (2011), https://doi.org/10.1016/j.shpsa.2011.09.004. As far as I have been able to ascertain, Abramson was the first to take note of the two heavy lines, which provide the jumping-off point for his claim that Descartes's language test served as a model for the so-called Turing test. In *The Turing Test Argument*, Gonçalves confirms the presence of the lines but draws somewhat different conclusions from them.

[54] Jefferson, "The Mind of Mechanical Man," 1110, 1106.

[55] "No Mind for Mechanical Man," *The Times*, June 10, 1949.

[56] "The Mechanical Brain–Answer Found to 300-Year-Old Sum," *The Times*, June 11, 1949.

[57] "No Mind for Mechanical Man."

[58] "ELECTRONIC BRAIN SOLVES PROBLEMS THAT BAFFLED HUMANS, MAY PLAY CHESS, *It Would Cook, Too, If It Could Smell*," interview with Alan Turing, *The Sunday Dispatch*, June 12, 1949.

[59] Robin Gandy, "Human Versus Mechanical Intelligence," in *Machines and Thought: The Legacy of Alan Turing*, vol. 1, ed. Peter Millican and Andy Clark (Oxford University Press, 1997).

[60] Italics added by the author.

[61] Susan G. Sterrett, "Turing's Two Tests for Intelligence," in *The Turing Test: The Elusive Standard of Artificial Intelligence*, ed. James H. Moor (Springer, 2003). See also Sterrett's follow-up to this fascinating essay, "The Genius of the 'Original Imitation Game' Test," *Minds and Machines* 30 (4) (2020), https://doi.org/10.1007/s11023-020-09543-6.

[62] All three are included in *The Essential Turing*.

[63] Alan Turing, "Computing Machinery and Intelligence," in *The Essential Turing*, ed. Copeland, 442, 460.

[64] In their article on the UC San Diego Turing tests, Cameron Jones and Benjamin Bergen make the important point that "the Turing test is not a direct test of intelligence, but a test of humanlikeness. For Turing, intelligence may have appeared to be the biggest barrier for appearing humanlike, and hence to passing the Turing test. But as machines become more similar to us, other contrasts have fallen into sharper relief, to the point where intelligence alone is not sufficient to appear convincingly human." Cameron R. Jones and Benjamin K. Bergen, "Large Language Models Pass a Standard Three-Party Turing Test," *Proceedings of the National Academy of Sciences* 123 (21) (2026).

[65] Turing, "Intelligent Machinery," in *The Essential Turing*, ed. Copeland.

[66] Alan Turing, "Can Digital Computers Think?" in *The Essential Turing*, ed. Copeland.

[67] A transcript of the roundtable is included in *The Essential Turing*.

[68] Martin Armstrong, "I *Don't* Think," *The Listener*, January 24, 1952.

[69] These details are drawn from coverage in *The Alderley & Wilmslow and Knutsford Advertiser* (hereafter *The Advertiser*) as well as several secondary sources: Jonathan Swinton, *Alan Turing's Manchester* (Infang Publishing, 2019); Andrew Hodges, *Alan Turing: The Enigma* (Princeton University Press, 1983); Dermot Turing, *Reflections on Alan Turing: A Relative Story* (The History Press, 2021) and *Alan Turing Decoded: The Man They Called* Prof (The History Press, 2015); and B. Jack Copeland, *Turing: Pioneer of the Information Age* (Oxford University Press, 2012).

[70] Turing, *Alan Turing Decoded*.

[71] "For Trial on Housebreaking Charge," *The Advertiser*, February 22, 1952.

[72] Indictment of Harold Thacker, Chester Archives, Court records QJF 380/2, quoted in Turing, *Alan Turing Decoded*.

[73] "Labouchere," *The New York Times*, January 17, 1912.

[74] Swinton, *Alan Turing's Manchester*; and "University Reader Put on Probation," *The Advertiser*, April 4, 1952.

[75] Frank Harris, *Oscar Wilde: His Life and Confessions* (Brentano's, 1916).

[76] That year, 1,069 gay Englishmen were prosecuted for sexual offenses, a number that in 1952 rose to 1,444. By contrast, in 1936, there were 535 prosecutions.

[77] Ironically, this was the fate of Jeremy Wolfenden, the son of Sir John Wolfenden, author of the Wolfenden Report that led to the repeal of Section 11. As a young journalist, Wolfenden had been photographed by the KGB having sex with another man, then blackmailed into spying for the KGB, then strongarmed by MI6 into acting as a double agent. Wolfenden committed suicide in 1965 at the age of thirty-one. Sebastian Faulks tells his story in *The Fatal Englishman: Three Short Lives* (Hutchinson, 1996).

[78] His principal correspondents were Nick Furbank, Robin Gandy, and Norman Routledge, a research fellow in Mathematics at King's College, Cambridge, of which Turing was also a fellow. As for his friends and colleagues in Manchester, they seem to have exhibited a singular–and convenient–lack of curiosity regarding his sexual life.

[79] The letter is included as AMT/D/14a in the Turing Digital Archive, see https://turingarchive.kings.cam.ac.uk/correspondence-amtd/amt-d-14a.