CHAPTER 11

Computing Machinery and Intelligence (*1950*)

Alan Turing

Introduction

Jack Copeland

Together with 'On Computable Numbers' (Chapter 1), 'Computing Machinery and Intelligence' forms Turing's best-known work. This elegant and sometimes amusing essay was originally published in 1950 in the leading philosophy journal *Mind*. Turing's friend Robin Gandy (like Turing a mathematical logician) said that 'Computing Machinery and Intelligence'

> 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.[1]

The quality and originality of 'Computing Machinery and Intelligence' have earned it a place among the classics of philosophy of mind.

The Turing Test

'Computing Machinery and Intelligence' contains Turing's principal exposition of the famous 'imitation game' or Turing test. The test first appeared, in a restricted form, in the closing paragraphs of 'Intelligent Machinery' (Chapter 10). Chapters 13 and 14, dating from 1951 and 1952 respectively, contain further

[1] R. Gandy, 'Human versus Mechanical Intelligence', in P. Millican and A. Clark (eds.), *Machines and Thought: The Legacy of Alan Turing*, vol. i (Oxford: Clarendon Press, 1996), 125.

Copyright 2004. Clarendon Press.
All rights reserved. May not be reproduced in any form without permission from the publisher, except fair uses permitted under U.S. or applicable copyright law.

EBSCO Publishing : eBook Academic Collection (EBSCOhost) - printed on 7/26/2023 9:07 PM via UNIVERSIDAD DEL ROSARIO
AN: 176999 ; B. Jack. Copeland.; The Essential Turing
Account: s4300083.main.eds

discussion and amplification; unpublished until 1999, this important additional material throws new light on how the Turing test is to be understood.[2]

The imitation game involves three participants: a computer, a human interrogator, and a human 'foil'.[3] The interrogator attempts to determine, by asking questions of the other two participants, which of them is the computer. All communication is via keyboard and screen, or an equivalent arrangement (Turing suggested a teleprinter link). The interrogator may ask questions as penetrating and wide-ranging as he or she likes, and the computer is permitted to do everything possible to force a wrong identification. (So the computer might answer 'No' in response to 'Are you a computer?' and might follow a request to multiply one large number by another with a long pause and a plausibly incorrect answer.) The foil must help the interrogator to make a correct identification.

The ability to play the imitation game successfully is Turing's proposed 'criterion for "thinking"' (pp. 442, 443). He gives two examples of the sort of exchange that might occur between an interrogator and a machine that plays successfully. The following is from p. 452.

Interrogator: In the first line of your sonnet which reads 'Shall I compare thee to a summer's day', would not 'a spring day' do as well or better?
Machine: It wouldn't scan.
Interrogator: How about 'a winter's day'? That would scan all right.
Machine: Yes, but nobody wants to be compared to a winter's day.
Interrogator: Would you say Mr Pickwick reminded you of Christmas?
Machine: In a way.
Interrogator: Yet Christmas is a winter's day, and I do not think Mr Pickwick would mind the comparison.
Machine: I don't think you're serious. By a winter's day one means a typical winter's day, rather than a special one like Christmas.

Did Turing Propose a Definition?

Turing is sometimes said to have proposed a definition of 'thinking' or 'intelligence'; and sometimes his supposed definition is said to be an 'operational' or 'behaviourist' definition. For example:

An especially influential behaviorist definition of intelligence was put forward by Turing.[4] (Ned Block)

[2] This additional material was first published in B. J. Copeland (ed.), 'A Lecture and Two Radio Broadcasts on Machine Intelligence by Alan Turing', in K. Furukawa, D. Michie and S. Muggleton (eds.), *Machine Intelligence 15* (Oxford University Press, 1999). See also B. J. Copeland, 'The Turing Test', *Minds and Machines*, 10 (2000), 519–39 (reprinted in J. H. Moor (ed.), *The Turing Test* (Dordrecht: Kluwer, 2003)).

[3] The term 'foil' is from p. 40 of B. J. Copeland, *Artificial Intelligence: A Philosophical Introduction* (Oxford: Blackwell, 1993).

[4] N. Block, 'The Computer Model of the Mind', in D. N. Osherson and H. Lasnik (eds.), *An Invitation to Cognitive Science*, vol. iii (Cambridge, Mass.: MIT Press, 1990), 248.

EBSCOhost - printed on 7/26/2023 9:07 PM via UNIVERSIDAD DEL ROSARIO. All use subject to https://www.ebsco.com/terms-of-use

[Turing] introduced ... an operational definition of 'thinking' or 'intelligence'... by means of a sexual guessing game.[5] (Andrew Hodges)

The Turing Test [was] originally proposed as a simple operational definition of intelligence.[6] (Robert French)

There is no textual evidence to support this interpretation of Turing, however. In 'Computing Machinery and Intelligence' Turing claimed to be offering only a '*criterion* for "thinking"' (emphasis added). Moreover, in his discussion of the Turing test in Chapter 14, Turing says quite specifically that his aim is not 'to give a definition of thinking' (p. 494).

In fact, Turing made it plain in 'Computing Machinery and Intelligence' that his intention was not to offer a definition, for he said:

The game may perhaps be criticised on the ground that the odds are weighted too heavily against the machine. If the man were to try and pretend to be the machine he would clearly make a very poor showing. He would be given away at once by slowness and inaccuracy in arithmetic. May not machines carry out something which ought to be described as thinking but which is very different from what a man does? (p. 442)

A computer carrying out something that 'ought to be described as thinking' would nevertheless fail the Turing test if for any reason it stood out in conversation as very different from a man. It follows that 'thinking' cannot be defined in terms of success in the imitation game. Success in the game is arguably a sufficient condition for thinking; but success in the imitation game is not also a *necessary* condition for thinking. (Someone's breathing spontaneously is a sufficient condition for their being alive, but it is not also a necessary condition, for someone may be alive without breathing spontaneously.)

The Male–Female Imitation Game

Turing introduced his criterion for 'thinking' by first describing an imitation game involving a human interrogator and two *human* subjects, one male (A) and one female (B). The interrogator must determine, by question and answer, which of A and B is the man. A's object in the game is to try to cause the interrogator to make the wrong identification. Having introduced the imitation game in this way, Turing said:

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?' (p. 441)

[5] A. Hodges, *Alan Turing: The Enigma* (London: Vintage, 1992), 415.

[6] R. French, 'The Turing Test: The First 50 Years', *Trends in Cognitive Sciences*, 4 (2000), 115–22 (115).

EBSCOhost - printed on 7/26/2023 9:07 PM via UNIVERSIDAD DEL ROSARIO. All use subject to https://www.ebsco.com/terms-of-use

Some commentators have suggested, on the basis of this passage, that Turing's criterion for thinking is that the computer in the Turing test be able to impersonate a woman.[7] Later in the article, however, Turing described matters differently, saying that the part of A is taken by a machine and 'the part of B . . . by a man' (p. 448). This runs contrary to the suggestion that the computer is supposed to imitate a woman (rather than a man or a woman). Moreover in Chapter 14 Turing says that '[t]he idea of the test is that the machine has to try and pretend to be a man . . . and it will pass only if the pretence is reasonably convincing' (p. 495). In Chapter 13 Turing presents the test in a starkly ungendered form: here the point of the test is to determine whether or not a computer can 'imitate a brain' (p. 485). On balance, then, it seems rather unlikely that Turing's intention in 'Computing Machinery and Intelligence' was to put forward a test in which the computer must impersonate a woman.

The role of the man-imitates-woman game is frequently misunderstood. For example, Hodges claims that this game is irrelevant as an introduction to the Turing test—indeed, it is a 'red herring'.[8] However, the man-imitates-woman game forms part of the protocol for scoring the test. Will interrogators decide wrongly as often in man-imitates-woman imitation games as they do in computer-imitates-human games? This question, Turing said, replaces 'Can machines think?'

The Current Status of the Turing Test

Section 6 of 'Computing Machinery and Intelligence', entitled 'Contrary Views on the Main Question', occupies nearly half of the article. It contains no fewer than nine objections to Turing's position, together with Turing's rebuttal of each. One of them, the 'Mathematical Objection', is also discussed in Chapters 10 and 12 (the introduction to Chapter 12 gives some further information about this important and controversial objection).

Since 'Computing Machinery and Intelligence' first appeared, Turing's test has received considerable attention from philosophers, computer scientists, psychologists, and others, and numerous additional objections have been raised to the test, some of them ingenious indeed. Nevertheless, it seems to me that none of these objections is successful (see my chapter in Moor's *The Turing Test* in the list of further reading). A discussion of one such objection, called here the Shannon–McCarthy objection, will give something of the flavour of the debate that still rages over the Turing test. Another form of objection—the 'Fiendish Expert' objection—is discussed in the introduction to Chapter 14.

[7] See, for example, S. G. Sterrett, 'Turing's Two Tests for Intelligence', *Minds and Machines*, 10 (2000), 541–59; S. Traiger, 'Making the Right Identification in the Turing Test', *Minds and Machines*, 10 (2000), 561–72 (both reprinted in J. H. Moor (ed.), *The Turing Test* (Dordrecht: Kluwer, 2003)).

[8] Hodges, *Alan Turing*, 415.

EBSCOhost - printed on 7/26/2023 9:07 PM via UNIVERSIDAD DEL ROSARIO. All use subject to https://www.ebsco.com/terms-of-use

The Shannon–McCarthy Objection

This objection envisages a hypothetical computer that is able to play the imitation game successfully, for any set length of time, in virtue of incorporating a very large—but nevertheless finite—'look-up' table. The table contains *all* the exchanges that could possibly occur between the computer and the interrogator during the length of time for which the test is run. The number of these is astronomical—but finite. For example, the exchange displayed earlier concerning sonnets and Mr Pickwick forms part of this (imaginary) table.

Clearly an interrogator would have no means by which to distinguish a computer using this table from a human respondent. Yet presumably the computer—which does nothing but search the table provided by its (hypothetical) programmers—does not think. In principle, therefore, an unthinking, unintelligent computer can pass the test.

Claude Shannon and John McCarthy put the objection forward in 1956:

> The problem of giving a precise definition to the concept of 'thinking' and of deciding whether or not a given machine is capable of thinking has aroused a great deal of heated discussion. One interesting definition has been proposed by A. M. Turing: a machine is termed capable of thinking if it can, under certain prescribed conditions, imitate a human being by answering questions sufficiently well to deceive a human questioner for a reasonable period of time. A definition of this type has the advantages of being operational, or, in the psychologists' term, behavioristic. ... A disadvantage of the Turing definition of thinking is that it is possible, in principle, to design a machine with a complete set of arbitrarily chosen responses to all possible input stimuli... Such a machine, in a sense, for any given input situation (including past history) merely looks up in a 'dictionary' the appropriate response. With a suitable dictionary such a machine would surely satisfy Turing's definition but does not reflect our usual intuitive concept of thinking.[9]

This objection has been rediscovered by a number of philosophers, and it is in fact usually credited to Block, who published a version of it in 1981.[10] (It is sometimes referred to as the 'blockhead' objection to the Turing test.)

What might Turing have said in response to the objection? A hint is perhaps provided by the following exchange between Turing and Newman (Chapter 14, p. 503):

> Newman: It is all very well to say that a machine could... be made to do this or that, but, to take only one practical point, what about the time it would take to do it? It would only take an hour or two to make up a routine to make our Manchester machine analyse all possible variations of the game of chess right out, and find the best move that way—*if* you

[9] C. E. Shannon and J. McCarthy (eds.), *Automata Studies* (Princeton: Princeton University Press, 1956), pp. v–vi.

[10] N. Block, 'Psychologism and Behaviorism', *Philosophical Review*, 90 (1981), 5–43.

EBSCOhost - printed on 7/26/2023 9:07 PM via UNIVERSIDAD DEL ROSARIO. All use subject to https://www.ebsco.com/terms-of-use

didn't mind its taking thousands of millions of years to run through the routine. Solving a problem on the machine doesn't mean finding a way to do it between now and eternity, but within a reasonable time. ...

Turing: To my mind this time factor is the one question which will involve all the real technical difficulty.

The Shannon–McCarthy objection establishes only that the

Turing Test Principle If x plays Turing's imitation game satisfactorily, then x thinks

is false in *some possible world*. The objection directs our imagination toward a possible world that is very different from the actual world—a world in which an astronomically large look-up table can be stored in a computer's memory and searched in a reasonable time—and points out that the Turing test principle is false in *that* world. However, there is no textual evidence to indicate that Turing was claiming anything more than that the Turing test principle is *actually* true, i.e. true in the actual world. Nor did he need to claim more than this in order to advocate the imitation game as a satisfactory real-world test.

Had Turing been proposing a definition of 'thinking', then he would indeed have had to say, consistently, that the Turing test principle is true in *all* possible worlds. (To take a more obvious case, if 'bachelor' is defined as 'unmarried male of marriageable age', then it is true not only in the actual world but in every possible world that if x is an unmarried male of marriageable age, then x is a bachelor.) At bottom, then, the Shannon–McCarthy objection depends on the interpretational mistake of taking Turing to be proposing a definition.

There is further discussion of the Turing test in Chapters 13, 14, and 16.

Learning Machines

The discussion of learning begun in Chapter 10 is continued in the iconoclastic Section 7 of 'Computing Machinery and Intelligence', entitled 'Learning Machines'. Turing poses the rhetorical question: 'Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's?' (p. 460). The child's mind may contain 'so little mechanism' that 'something like it can be easily programmed'. If this child-machine 'were then subjected to an appropriate course of education one would obtain the adult brain'. These remarks are of a piece with Turing's suggestion in Chapter 10 that 'the cortex of an infant is an unorganised machine, which can be organised by suitable interfering training' (p. 424).

Turing mentions in Section 7 that he has 'done some experiments with one such child-machine, and succeeded in teaching it a few things, but the teaching method was too unorthodox for the experiment to be considered really success-

EBSCOhost - printed on 7/26/2023 9:07 PM via UNIVERSIDAD DEL ROSARIO. All use subject to https://www.ebsco.com/terms-of-use

ful' (p. 461). Here he is probably referring to the experiments with an unorganized machine that are described in Chapter 10, where he says that he has 'succeeded in organising such a (paper) machine into a universal machine', but that the technique used 'is not sufficiently analogous to the kind of process by which a child would really be taught' (pp. 427-8).

Situated AI

AI traditionally has attempted to build disembodied intelligences carrying out abstract activities—e.g. chess-playing—and whose only way of interacting with the world is by means of a screen or printer. An alternative approach now called 'situated AI' aims at building embodied intelligences situated in the real world. 'Computing Machinery and Intelligence' ends with a characteristically far-sighted statement in which Turing sketches each of these two approaches to AI. He contrasts research that focuses on 'abstract activity, like the playing of chess' with research aiming 'to provide the machine with the best sense organs that money can buy, and then teach it to understand and speak English' (p. 463). Turing recommended that 'both approaches should be tried' (ibid.; compare Chapter 10, pp. 420–1).

Rodney Brooks, a modern pioneer of situated AI and Director of the MIT Artificial Intelligence Laboratory, pointed out that although Turing proposed both these 'paths toward his goal of a thinking machine', Artificial Intelligence for a long time 'all but ignored' the situated approach.[11] Now the tables have turned and there is huge interest in situated AI.

One of Brooks's experimental robots, Herbert—named after Herbert Simon—searched the offices and work-spaces of the MIT AI Lab for empty soda cans, picking them up and carrying them to the trash.[12] Herbert, unlike previous generations of experimental robots, operated in real time in a busy, cluttered, and unpredictably changing real-world environment. Brooks's humanoid learning robot Cog—from 'cognizer'—has four microphone-type 'ears' and saccading foveated vision provided by cameras mounted on its 'head'.[13] Cog's legless torso is able to lean and twist. Strain gauges on the spine give Cog information about posture, while heat and current sensors on the robot's motors provide feedback concerning exertion. Cog's arm and manipulating hand are coated with electrically conducting rubber membranes providing tactile information. Those working in situated AI regard Cog as a milestone on the road toward the realization of Turing's dream.

[11] R. Brooks, 'Intelligence without Reason', in L. Steels and R. Brooks (eds.), *The Artificial Life Route to Artificial Intelligence* (Hillsdale, NJ: Erlbaum, 1995), 34. See also R. Brooks, *Cambrian Intelligence: The History of the New AI* (Cambridge, Mass.: MIT Press, 1999).

[12] R. Brooks, 'Elephants Don't Play Chess', *Robotics and Autonomous Systems*, 6 (1990), 3–15.

[13] R. A. Brooks and L. A. Stein, 'Building Brains for Bodies', *Autonomous Robots*, 1 (1994), 7–25.

EBSCOhost - printed on 7/26/2023 9:07 PM via UNIVERSIDAD DEL ROSARIO. All use subject to https://www.ebsco.com/terms-of-use

Further reading

Block, N., 'Psychologism and Behaviorism', *Philosophical Review*, 90 (1981), 5–43.

Dennett, D. C., 'Can Machines Think?', in his *Brainchildren: Essays on Designing Minds* (Cambridge, Mass.: MIT Press, 1998).

French, R., 'The Turing Test: The First 50 Years', *Trends in Cognitive Sciences*, 4 (2000), 115–22.

Michie, D., 'Turing's Test and Conscious Thought', *Artificial Intelligence*, 60 (1993), 1–22. Reprinted in P. Millican and A. Clark (eds.), *Machines and Thought: The Legacy of Alan Turing* (Oxford: Clarendon Press, 1996).

Moor, J. H. (ed.), *The Turing Test* (Dordrecht: Kluwer, 2003).

—— 'An Analysis of the Turing Test', *Philosophical Studies*, 30 (1976), 249–57.

Provenance

What follows is the text of the original printing of 'Computing Machinery and Intelligence' in *Mind*.[14] (Unfortunately Turing's typescript has been lost.)

[14] Footnotes have been renumbered consecutively. All footnotes not marked 'Editor's note' appeared in *Mind*. Where the text contains numbers referring to pages of *Mind* these have been replaced by the numbers of the corresponding pages of the present edition, enclosed in square brackets. Not all cross-references in Turing's article were dealt with correctly by the editor of *Mind*—some of the numbers appearing in *Mind* presumably refer to pages of Turing's original typescript. These also have been replaced by the numbers of the corresponding pages of this volume.

EBSCOhost - printed on 7/26/2023 9:07 PM via UNIVERSIDAD DEL ROSARIO. All use subject to https://www.ebsco.com/terms-of-use