The Turing Test

Verbal Behavior as the Hallmark of Intelligence

edited by Stuart M. Shieber

A Bradford Book
The MIT Press
Cambridge, Massachusetts
London, England

Immediate Responses

The centrality of Turing's test in the philosophy of mind is undisputed. The nature of thought is the key question of the field and the reduction of thought to mechanism is currently the key approach to that question. Turing, in proposing his Test, had packaged in one easily graspable form many of the central problems of philosophy of mind that had exercised people for centuries: the mind-body problem, how mental states relate to the world, the problem of the existence of other minds.

There is a certain irony in the centrality of the Turing Test in philosophy. As we have seen, Turing himself was of, at best, mixed mind as to the role of the Test, of what it is a test *of*. Is it a test of thinking, or intelligence, or intelligent behavior, or none of these? According to Turing, passing of the Test was not intended as *definitional* of intelligence; the Test was intended to *replace* the question of machine intelligence, not to answer it. He thought the latter question too imprecise and woolly to be a respectable topic of inquiry, "too meaningless to deserve discussion". Block (1990) notes the direct analogy with Turing's earlier work on the Turing machine. Computability by a Turing machine was intended to serve as a precise replacement for the vaguer notion of mechanical computability. In the same way, Turing hoped that passing of the Turing Test could serve as a precise replacement for the vaguer notion of mechanical thought. But as Moor points out,

> if Turing intends that the question of the success of the machine at the imitation game replace the question about machines thinking, then it is difficult to understand how we are to judge the propriety and adequacy of the replacement if the question being replaced is too meaningless to

Premise 1: Humans are intelligent.

Premise 2: The conversational verbal behavior of humans reveals that (human) intelligence.

Premise 3: If an agent has behavior of a type that can reveal intelligence and that is indistinguishable from that of an intelligent agent, the former agent is itself intelligent.

Premise 4: Any agent that passes the Turing Test has conversational verbal behavior indistinguishable from that of humans.

Conclusion: Therefore, any agent that passes the Turing Test is intelligent.

Figure 1
The Turing syllogism

deserve discussion. Our potential interest in the imitation game is aroused not by the fact that a computer might learn to play yet another game, but that in some way this test reveals a connection between possible computer activities and our ordinary concept of human thinking. (Moor 1976, *chapter 17*)

Thus, philosophers have been inexorably led to the question of the relationship between a machine's passing of the Test and its thinking capacity, the Big Question.

Turing's view notwithstanding, it is possible to reason from a machine's passing of the Test to its intelligence according to a kind of syllogism, presented in figure 7.1, which is implicitly assumed by all philosophers investigating the ramifications of the Turing Test beyond the limited confines that Turing himself proposed it within. Turing himself never explicitly supports this syllogism. Indeed, he finds the conclusion of the syllogism to be meaningless. The nearest he comes to approving of it is an implication that passing the Test "ought to be described as thinking", in his discussion of the fact that the Test is not a necessary condition, of which more below.

The Turing Test Not a Necessary Condition

If the syllogism works at all, it works in only one direction; the converse of the conclusion—that is, "Any intelligent agent can pass the Turing Test"—need not follow from the premises. The test is in a certain sense too hard. A machine that "ought to be described as thinking" might fail the Turing Test for all kinds of incidental reasons. Perhaps it gives itself away by typing its answers too quickly or accurately, for instance.[1]

The Turing syllogism is thus an argument for the Turing Test being a sufficient, but not necessary, condition of intelligence. Even Turing admitted that the Test would not serve as a necessary condition. Turing found the objection to be "very strong, but at least we can say that if, nevertheless, a machine can be constructed to play the imitation game satisfactorily, we need not be troubled by this objection." (Turing 1950, 435, *chapter 4*) However, the problem with sufficient conditions is that nothing prevents them from being so excessively burdensome as to be useless. If we make it a sufficient condition to being a lifeguard that the candidate be able to swim 100 kilometers carrying a 300 kilogram weight, we are sure to hire only qualified lifeguards, but we are equally sure to hire too few.

For this reason, Turing himself attempted to eliminate from the Test certain incidental properties of machines that seemed irrelevant for determining intelligent behavior. For example, possessing skin should not be indicative one way or another of possessing intelligence (says Turing), so the Test does not allow the interrogator to view the participants while conversing. Rather, the interlocution proceeds by means of typewritten messages. Despite these attempts by Turing to eliminate incidental properties, Robert French (1990, *chapter 13*) argues that the Turing Test is, like the super-lifeguard test, so difficult that it is essentially useless as a test of intelligent behavior. Rather, he says, it is inherently a test of culturally-oriented, human intelligence, and as such may be irrelevant to the question of whether machines can exhibit intelligent behavior.

---

1 In the science-fiction movie *Blade Runner* (Scott 1982), the character Deckard and other "blade runners" are specially trained police whose job is to identify human-like machine "replicants". They use a "Voight-Kampff" test, a kind of über-Turing-Test, a dialogue with the subject in which the tell-tale sign is not the verbal behavior at all, but subtle coincident movements of the iris detectable only by special equipment. We can take this to be the paradigmatic example of an incidental distinction.

The Turing Test Not a Sufficient Condition

On the other hand, it might be argued that the Test is not even a sufficient condition for determining intelligent behavior. This is tantamount to denying the syllogism in one (or more) of its premises.

Premise 1 is relatively uncontroversial (although its falsehood is raised as a possibility by Purtill [1971, *chapter 10*] and denied explicitly by Sampson [1973, *chapter 11*]).

Premise 2 is essentially Descartes's observation. Denial of it amounts to a form of solipsism.

Premise 4 constitutes the very definition of the notion of "passing the Turing Test", and is thus true by fiat.

Premise 3 is by far the most popular of the premises to deny. In order to show that the Test is not a sufficient condition, that it is somehow too easy, it can be demonstrated that an artifact that is patently unintelligent is capable of passing the Test (though perhaps not likely to do so). This is surprisingly easy. Imagine a machine that responds to the interrogator's queries by emitting a random sequence of keystrokes. (The idea is conventionally implemented using monkeys and typewriters.) There is some (admittedly astronomically small) probability that these keystrokes will fortuitously spell out perfectly plausible responses to the queries, and the interrogator would therefore be fooled into confusing the random keystroke generator with a human. If one holds the incontrovertible stance that the random typing responses were not true intelligent behavior—how could they be, as they were not produced by an intelligent being?—then the *mere possibility* of such an occurrence, by itself, demonstrates that passing the Turing Test is not a sufficient condition for intelligent behavior, at least in the strongest sense of the term "sufficient".

Attempts to drive a wedge between the two concepts "agent that could pass the Turing Test" and "agent exhibiting intelligent behavior" much more sophisticated than this have been provided in the philosophical literature. (See the chapter The Wedge and The Spark on page 147 below.) Keith Gunderson's "toe-stepping machine" (1964b, *chapter 9*), Ned Block's "Aunt Bertha" (1981, *chapter 15*), and John Searle's "Chinese Room" (1980, *chapter 14*) are crisp, if controversial, examples.

In the face of such examples, three responses are possible. First, one might find fault with the examples, as many of the critics of Searle do. (See, for example, Hofstadter [1981], Dennett [1987a].)

Second, one might back off from such a strong notion of sufficient condition, as Moor (1976, *chapter 17*) does. Rather than viewing passing the Turing Test as a *guarantee* of intelligent behavior, we might embrace a slightly weaker notion of passing the Turing Test as *evidence* of intelligent behavior. We thereby move our view of the Test from a criterion to an indication of intelligence; because of the difficulty of the task, however, the evidence can be of almost arbitrary strength (Shieber 2004).

Finally, one might deny the distinction in the first place, as Turing himself seems to have wanted to do. Determining the relationship between the two concepts assumes that they have independent status. But the original point of the Turing Test, contra its usage as a philosophical thought experiment, was not to serve in the determination of the property of behaving intelligently, but to serve as a replacement for the notion of intelligence and intelligent behavior, which are such ill-defined notions that no arguments concerning them can have any import. This gambit, taken by Chomsky (*chapter 20*), is internally consistent, and certainly solves Turing's problem, but it may do little to further an understanding of the philosophical issues Turing raises. And perhaps that is all for the best.

The papers in this final part of the collection cover the range of replies to and views of the Turing Test from the philosophical community. They fall into two classes, with some overlap: first, essentially the entire responsa from 1950 on to the article published in *Mind*, comprising the papers by Pinsky, Gunderson, Purtill, Sampson, Millar, and French; second, a selection of the most important papers from throughout the philosophical literature covering the full range of view on the Big Question. These include the papers by Gunderson, French, Block, Searle, Dennett, and a new, previously unpublished paper by Chomsky.

No two respondents to Turing's proposals share the same interpretation of the Turing Test. In part, this is what makes the Test such a fascinating thought experiment. But all would agree that it provides a lens with which to focus discussion on the central issue in philosophy of mind, the characterization of thought itself.