Making the Right Identification in the Turing Test[1]

SAUL TRAIGER
Department of Philosophy and Cognitive Science Program, Occidental College, Los Angeles, California, 90041, USA; E-mail: traiger@oxy.edu

Abstract. The test Turing proposed for machine intelligence is usually understood to be a test of whether a computer can fool a human into thinking that the computer is a human. This standard interpretation is rejected in favor of a test based on the Imitation Game introduced by Turing at the beginning of "Computing Machinery and Intelligence."

Key words: Artificial Intelligence, Imitation Game, Turing Test

Alan Turing's 1950 paper, "Computing Machinery and Intelligence," introduces the question, "Do machines think?" and the problem of interpreting that question. The interpretation of the question as "Do people believe that machines think?" is quickly dismissed as absurd, though Turing doesn't immediately say what is absurd about it. Turing proposes a reformulation of the question in terms of what he calls "The Imitation Game," taking considerable effort to describe the game and its employment in answering the original question and in presenting a prediction. It's surprising, then, that most of the literature on what has come to be called the "Turing Test" completely discounts the role of the Imitation Game in Turing's proposal and prediction. I argue that Turing's invocation of the Imitation Game forms the very basis of Turing's Test for machine intelligence. In order to understand Turing's test and his prediction, we must examine the text with care, and attempt to interpret what Turing actually said. I begin by documenting what I call the Standard Interpretation of the Turing Test. Then I argue that the Standard Interpretation does not square with the text of "Computing Machinery and Intelligence." Fortunately an interpretation that does square with the text is available. I close by explaining why it matters to get the interpretation right.

The Standard Interpretation takes the Turing Test to be as follows: There is a computer and two humans. One human is the interrogator. She or he communicates with the computer and the other human using a teletype or computer terminal. The interrogator is told that one of the two individuals it is communicating with is a computer, and that the other is a human. The computer's goal is to fool the interrogator into thinking that it is the human. The other human's goal is to convince the interrogator that he or she is the human. The computer attempts to achieve the deception by imitating human verbal behavior. If an interrogator does not make the *right identification*, where a "right identification" is identifying the computer as a computer, then the computer passes the test.

There are a couple of observations to make before documenting the widespread acceptance of this reading of the Turing Test. First, it makes no reference to the

Minds and Machines **10:** 561–572, 2000.

Imitation Game, though it does describe the computer's method as imitating the behavior of a human being. Finally, on the Standard Interpretation, the characterization of passing or failing the test makes reference to the computer.[2] The computer passes the test if the interrogator does not identify it as a computer. I turn now to establishing that what I've called the Standard Interpretation is indeed just that.

Ned Block describes the Turing Test as follows:

> His [Turing's] version of behaviorism formulates the issue of whether machines could think or be intelligent in terms of whether they could pass the following test: a judge in one room communicates by teletype (This was 1950!) with a computer in a second room and a person in a third room for some specified period. (Let's say an hour.) The computer is intelligent if and only if the judge cannot tell the difference between the computer and the person. (Block, 1990, p. 48)

This formulation is an unhappy one for several reasons, the most important of which is that it makes the test both necessary and sufficient for intelligence. On Block's reading, only those who have passed the test are intelligent. Perhaps Block sees the test as a test only of *computer* intelligence. If so, it's hardly in the spirit of Turing's attempt to characterize thought and intelligence in a way which does not make reference to what Turing took to be the contingent features of the agent, such as it's size, the stuff of which it's made. As Block construes it, the Turing Test is a test that only a computer could play. Finally, it isn't clear what counts as passing the Block's version of the Turing Test. Block says that the test is passed when the judge can't tell the difference between the computer and the person. Does that mean that the judge can't identify two distinct communicative agents, or more likely, does it mean that the quality of the discourse is indistinguishable? If the latter, how would the judge determine that the two sets of discourse are qualitatively identical?

The formulation of the test by Dyer is close to Block's. Like Block, Dyer describes the testing conditions as including a set of interrogators, a computer, and a "person." I take it that the person is a human being, though Dyer himself does not specify. The interrogator knows that one of the two interlocutors is a computer, the other a person/human being. The computer passes the test if the interrogators "are about evenly divided on which teletype" is connected to the computer and which to the person.[3] (Dyer, 1994, p. 174).

The Standard Interpretation is as entrenched in the AI community as it is in philosophy. Armer characterizes the test as one in which an "interrogator, who can communicate with a human and a machine via teletype... does not know which is which, is to decide which is the machine" (Armer, 1960, p. 390). Jason L. Hutchens, the runner-up for the 1998 Loebner Prize, a prize awarded on the basis of actual machine performance in a Standard Version Turing Test, summarizes the test as requiring the interrogator to "decide which is a human being and which is a computer program. The human being helps the interrogator to make the correct identification, while the computer program attempts to trick the interrogator into making the wrong identification. If the latter case occurs, the computer program is said to be exhibiting intelligence" (Hutchens, unpublished).

The Standard Interpretation is so entrenched that it even appears in (mis)quotations of Turing's own formulation. Donald Michie quotes Turing as follows:

> I believe that in about fifty years' time it will be possible, to programme computers, with a storage capacity of about 10^9, to make them play the imitation game so well that an average interrogator will not have more than 70 per cent chance of making the right identification [as between a human and computer] after five minutes of questioning. (Michie, 1996, p. 29)

Michie offers this up as Turing's original text, but it is not. Michie has taken the liberty of adding the parenthetical phrase "as between a human and a computer," a phrase which does not appear in the original text. In his zeal to help us understand what Turing means, Michie has misrepresented what Turing actually said. One might argue that although this is indeed scholarly sloppiness on Michie's part, what else could Turing have meant than the phrase that Michie inserted? This is a fair question, and one that I will answer below.

A few commentators offer a short formulation of the Turing Test that conforms to the Standard Interpretation, while acknowledging that Turing's actual test was different. Donald Davidson's initial description of the Turing Test is representative. Davidson writes, "... Turing predicted that by the year 2000 it will be possible to build a computer that will have at least a 30 per cent chance of fooling an average interrogator into thinking it is a person" (Davidson, 1990, p. 1). While Davidson later describes the Imitation Game, he thinks that Turing may have thought the interrogator would be faced with the task of choosing between a computer and a woman, that is, that he is presented with the task as deciding which interlocutor is the woman and which is the computer. I'll argue below that Turing's test cannot credibly be so interpreted.

William Rappaport writes, Turing rejected the question "Can machines think?" in favor of the more behavioristic question "Can a machine convince a human to believe that it (the computer) is a human?" (Rappaport, 1994, p. 227). Rappaport thinks that a necessary condition for convincing the interrogator is that the computer understand a natural language. It's not at all obvious that the test, so construed, requires natural language understanding. In a footnote, however, Rappaport refines the characterization, noting that the computer's job is to convince the interrogator that it is a woman. Rappaport does not see that the very difference to which he has drawn our attention makes a difference to the necessity claim. It's much more plausible that natural language understanding is required to convince the interrogator that its interloeutor is a woman than to convince it that it is a human being. So I shall argue below.

Moor refers to the "standard version of the imitation game" which "involves a man, a computer, and a human interrogator." The interrogator's task is to "decide which respondent is the man and which respondent is the woman" (Moor, 1976, p. 249). In an endnote Moor adds that Turing's text is ambiguous, that it underdetermines the standard interpretation. Moor doesn't offer alternative interpretations, and he makes no reference to the original imitation game. In a later article (Moor, 1987), Moor appreciates that the game described under the Standard Interpretation is a transformation from the original imitation game, and he observes that Turing did not explicitly say whether the computer replacing the man is instructed to play the part of a woman. Moor concludes, "In any event these niggling details are probably unimportant to Turing's main claim that questions about how well a computer does in the imitation game should replace the original question "Can machines think?" (Moor, 1987, p. 1126).

I hope this is enough of a sample to establish the widespread interpretation of Turing's refraining of "Can machines think?" as "Can machines deceive a human into thinking that the computer is a human being?"[4] I've mentioned several ancillary infelicities such as the conflation of "person" with "human being" by Dyer because I think Turing was struggling to come up with a formulation that would provide some clarity. I think he would be dismayed by the level of confusion displayed in the cited commentaries. I have yet to establish that the Standard Interpretation is a misinterpretation, and it is to that task that I now turn.

To understand Turing's Test, we begin with the Imitation Game, the game Turing elaborately presented at the beginning of his paper. Let's first note that the Imitation Game is introduced as a game involving three *people*, that there are no computers mentioned in the description of the game. The three people are an interrogator, a man and a woman. Turing labels the man "A", the woman "B", and the interrogator "C." The interrogator, however, has his or her own labels for the other two individuals, who are in a different room. The interrogator uses the labels "X" and "Y." Communication occurs by the passing of typewritten answers, passing sheets of paper or using a teleprinter. (Turing does not say that the questions from the interrogator must be typewritten.) The interrogator is charged with determining which player is the man and which is the woman. In Turing's own words:

> The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game he says either "X is A and Y is B" or "X is B and Y is A." (Turing, 1950, p. 433)

The male's goal is to "cause C to make the wrong identification." The female's goal is to aid C in making the right identification.

That's the Imitation Game. The game itself makes no reference to computers, and Turing does not say that it requires deception on the part of the male. The male is to try to cause the interrogator to make the wrong identification. One can certainly cause someone to have a false belief without engaging in deception, where deception includes the intention to bring about a false belief in another. Why is this important? When Turing later asks us to imagine a computer playing this game, given this formulation of the game, he does not have to imagine the computer being given the instruction to deceive the interrogator. In fact, while Turing gives an example of a possible exchange among contestants that does involve deception, it's only an example, and not a feature of the description of the game. So even when the game is first described in terms of human players, Turing minimizes the metaphysical overhead.

I want to draw attention to a second detail, before moving to the introduction of machines into the game. Turing uses the labels "A" for the male, "B" for the female and "C" for the interrogator. The form of the interrogator's final judgment is "X is A and is B" or "X is B and Y is A." Now Turing may have used these abbreviations to reduce verbal clutter, but I speculate that there was another purpose. As many commentators have noted, there's nothing special about the game determining whether someone is a male or a female. The game could involve other characteristics, though what is to be included in the range of "replacement" characteristics is an interesting question, and one to which we will return. If this is right, "A" and "B" could be placeholders for whatever characteristics may be used in different versions of the game. Turing's formulation invites generalization.

Machines enter the mix immediately after the description of the game. Turing writes:

> 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?" (Turing, 1950, p. 434)

The interpretation of this paragraph is crucial. We are to substitute a machine, specifically a digital computer, for A in the game. But substituting one player for another in a game presupposes the identity of the game. So nothing about the game has changed. We merely have a different individual, one who happens to be a digital computer, taking on the role of A. The machine, as A, is to cause C to believe that A is a woman. Playing the Imitation Game, it would make no sense for the computer to try to deceive the interrogator into thinking that the computer is a human being, because the interrogator is told that one of the two players is a man and one is a woman. Again, Turing says "The object of the game for the interrogator is to determine which of the other two is the man and which is the woman."[5]

There's a vast difference between the way Turing situates the machine and the situation described in the Standard Interpretation. In Turing's version the interrogator's belief, that the individual who happens to be a machine is a human, and hence intelligent, is a belief which can be defeated by the poor showing of the machine. On the Standard Interpretation, the interrogator believes that one or the other interlocutor is a computer, (where there are two interlocutors) and so does not believe in advance that an arbitrary participant with which it communicates is a human. The interrogator brings to the game his or her beliefs about computers and machine intelligence. Similarly on the Standard Interpretation the computer and the human competitor may also exploit beliefs about machines to try to trick or help the interrogator.

The opening paragraph of "Computing Machinery and Intelligence" makes it clear that Turing was not interested in whether or not a population might have beliefs which would incline them to regard machines as intelligent. In light of the novelty of digital computers at the time Turing was writing, he knew that most people would be disposed to regard computers as incapable of intelligence. Interrogators in such a population would be very likely to underestimate the capabilities of computers, due to their lack of familiarity with them. A clever programmer could exploit that ignorance in such a way as to trick interrogators into thinking that the computer is a human being.

One strategy which has been implemented in a number of attempts to pass the Standard Interpretation Turing Test is to make use of a text-base of typical greetings, remarks and responses which are often encountered in casual conversation. A program that functions in this way is RACTER.[6] Here's an opening exchange between RACTER and a human:

Hello, I'm Racter. Are you Saul?
>Yes.
You are Saul Smith?
>No.
What's your name then?
>Saul Traiger.
I may call you Saul, then?
>Yes.
Hello, Saul. I believe you're interviewing me. Are you continuing Saul Smith's interview?
>No.
Then well forget about Saul Smith and start over. What would you like to know? Excuse me a minute while I attend to some business – I'll be right back. I'm back.

A little later RACTER asks:

Hmm. Are you related to Louella Traiger?

RACTER clearly has codified some of the opening gambits of ordinary casual conversations. In runs of the program, novice users are often very surprised by such opening conversational moves. They are surprised at the appropriateness of RACTER's opening conversational moves. When the emphasis is on deceiving a human being into thinking that it is conversing with a human rather than a computer, strategies such as this one are fair game. A programmer who can string out a series of conversational gambits such as the ones used by RACTER may very well succeed in stalling a human's recognition that it is conversing with a computer. The Imitation Game, in contrast, has a point that constrains the participants to get beyond chit-chat and move directly to content driven discourse. A computer that succeeds in the Imitation Game will have to process ordinary natural language conversations which go well beyond polite conversation. A computer might succeed in the Standard Interpretation game without facing such a challenge.

Turing's formulation attempts to level the playing field, by putting the machine in the same circumstances we would place a human player. Those circumstances are not just the use of teletype across different rooms. They go to the heart of the setup of the game itself. The interrogator decides wrongly when it judges A to be the woman. The reframed question is simply whether the interrogator judges wrongly as often with a machine as A it does with a man as A. What counts as deciding rightly, or, as Turing later puts it, "making the right identification?" That's simple. The interrogator makes the right identification when it identifies Y as B, that is, when it identifies the individual who is the woman as the woman. Now the interrogator might infer the right identification from the prior conclusion that A is a computer. But it need not. The individual who occupies the role of A, whether human or computer, loses the game when the interrogator makes the right identification. It still loses if the interrogator fails to identify the computer as a computer, because that identification is no part of the Imitation Game.

What level of performance in the Imitation Game counts as success in the game? In the passage quoted above, Turing says that the question is whether the interrogator will decide wrongly as often when A is the computer as when A is a man. It is this question that explicitly replaces the original question, "Do machines think?" We need to distinguish this from Turing's *prediction*, later in the paper, where he writes:

> I believe that in about fifty years' time it will be possible, to programme computers, with a storage capacity of about 10^9 to make them play the imitation game so well that an average interrogator will not have more than 70 per cent chance of making the right identification after five minutes of questioning. (Turing, 1950, p. 442)

Taking the phrase, "making the right identification" to mean "identifying the woman as a woman," I interpret this passage as presupposing that interrogators in the original male/female imitation game also have a 70 percent chance of making the right identification, though Turing never explicitly says this. Having introduced the computer in the role of the male, Turing's question is "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?" (Turing, 1950, p. 434), It's natural to interpret the prediction as answering in the affirmative, particularly since Turing follows with the further prediction that common usage at the end of the twentieth century would allow one "to speak of machines thinking without expecting to be contradicted" (Turing, 1950, p. 442). If this is right, the prediction about the Imitation Game should be understood as follows: Within 50 years, when computers replace the male in the Imitation Game, interrogators will have the same (70 percent) chance they had in the original male/female game of making the right identification. It's worth noting, however, that if this is right, Turing can be criticized for providing no grounds for the 70 percent estimate for interrogator performance in the original Imitation Game.

Why has the Standard Interpretation taken hold so firmly, and does it matter if it has? At the end of section 5, after explaining the universality of digital computers, Turing writes:

> We may now consider again the point raised at the end of §3. It was suggested tentatively that the question, "Can machines think?" should be replaced by "Are there imaginable digital computers which would do well in the imitation game?" If we wish we can make this superficially more general and ask "Are there discrete-state machines which would do well?" But in view of the universality property we see that either of these questions is equivalent to this, "Let us fix our attention on one particular digital computer C. Is it true that by modifying this computer to have anadequate 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?" (Turing, 1950, p. 442)

This passage is a bit confusing because Turing now uses the label "C" for the computer while earlier in the paper "C" referred to the interrogator. But the crucial phrase here is found at the end of the last sentence in the passage quoted, "the part of B being taken by a man." The Standard Interpretation takes this to be a change in the conditions of the game. The part of B is now the role of a human being, played by a male human being. The computer and the male are thus charged with playing the role of a human being, not the role of a female human being.

While it is admittedly somewhat natural to read the passage in this way, such a reading makes several questionable moves.[7] First, one must read "man" in "the part of B being taken by a man" as "male" *and at the same time*, one must read "the part of B" to mean "the part of a human player" rather than "the part of a woman." Let's take each phrase in turn. First, Turing, like his contemporaries, in "Computing Machinery and Intelligence" other published papers, typically used the term "man" as we would now the gender neutral "person" or the explicit "man or woman." Second, the phrase, "the part of B" must refer to the role B plays in the original imitation game, because Turing describes no other role, here or elsewhere. He never says that the role of B has changed from that of a woman to that of a human being, man or woman. Thus, the proper reading is the one which is consistent with the description of the Imitation Game Turing has given up to this point, namely, that a computer replaces the male, and the role of the female, is played by a human, either male or female.

Another passage that appears to support the Standard Interpretation occurs in section 2, "Critique of the new problem," where Turing makes the important point that the game puts to one side the actual physical attributes of the participants. He thinks that such attributes are irrelevant to intelligence. Turing says that even if a chemist were able to make artificial skin, "we should feel there was little point in trying to make a "thinking machine" more human by dressing it up in such artificial flesh." This may suggest that Turing has moved from the Imitation Game to the question of whether a machine, which is physically unlike a human, can imitate human intelligence. That sense is heightened by a set of questions and possible answers which Turing lists as illustrating the advantage of "the proposed criterion." Those questions, including requesting a sonnet, asking for a sum, and making queries about appropriate chess moves, seem too open-ended to be relevant to the Imitation Game. They seem instead to be questions one would ask to determine whether something was a machine rather than a human. So it may seem Turing has moved away from the Imitation Game to the question of whether a machine imitate a human in a conversation conducted via teletype.[8]

The questions of section 2 are the appropriate questions to ask in the Imitation Game. The interrogator is not restricted to posing obvious questions about gender; indeed, if the other players are at all clever, the interrogator will need to develop a patternof clues about their respective genders. One of the great strengths of the reframed question is that conversations are open-ended. The interrogator is free to ask any question that will advance his or her inquiry. It's also true that when a computer takes the place of a human male, the computer needs to guard against revealing its non-humannature. But guarding against revealing itself as a computer can be just a small part of the computer's strategy, if it hopes to perform at a comparable level to a human. The fact that the interrogator initially believes that the conversation is only among humans will help keep this aspect of the computer's strategy in the background.

The test as described in the Standard Interpretation loses this key virtue of Turing's Imitation Game, that the kinds of questions posed by the interrogator will be those that would typically be posed to human participants. The interrogator in the Imitation Game will not begin with the hypothesis that one participant is a computer, and hence the computer participant, like its human counterpart, can devote its energies to answering the same kinds of questions that would typically be posed to a human.

Another way one might defend the Standard Interpretation is to admit that the Turing Test is not couched in terms of the original Imitation Game, but that the Imitation Game was just an example of a type of game involving imitation, and that when Turing introduces the machine into the game, he's really introducing a new imitation-type game, rather than introducing the computer into the original Imitation Game. The key idea of the Imitation Game, that A is trying to deceive the interrogator by pretending to be something he's not, is retained. The only difference is that instead of a man imitating a woman, we have a computer imitating a human.

In response to this defense of the Standard Interpretation, I must emphasize again that the text of Turing's paper does not support such a reading. Turing refers to one and the same Imitation Game throughout. He never suggests that there are other games he wants to consider. There is, however, an independent reason why Turing would have rejected the specification of the game in the terms favored by the Standard Interpretation. We've noted that on the Standard Interpretation, the playing of the game makes essential reference to the computer. The game is the game of a computer trying to imitate a human being. The interrogator is told that it may or may not be communicating with a computer, or that one of two interlocutors is a computer. In order for the game to make any sense, we must attribute some idea of what a computer is to the interrogator. But this already introduces complications. What conversational strategy does one take when one suspects that one may be conversing with a computer? Shouldn't we expect different results with persons who have different conceptions of computers? In fact, such a game would take us very far afield from ordinary conversation. But it was facility in ordinary conversation that Turing took to be the "criterion" in his reframed question.

Turing thought it absurd to interpret the question "Can machines think" as "Do people believe that machines think." The absurdity stems in part from the fact that there would not be a univocal notion of "machine" in the population. So a poll could not capture an accepted view on the matter. The same difficulty presents itself on Standard Interpretation Turing Tests. Interrogators setting out to unmask a computer will mobilize and rely on their on conceptions of what computers are and what they can and can't do, just as the computer program may successfully exploit those very beliefs in order to appear human-like. And this makes the test dependent on such conceptions of computation and mechanism, precisely the uninteresting question bypassed by Turing's Imitation Game.[9]

The Imitation Game is a more stringent game than the game of machine imitation of human behavior envisioned by the Standard Interpretation. In the former, a player who fails to be seen as human will quickly lose the game.[10] But being seen as a human player is not sufficient for winning the game. The computer in the Imitation Game must convince the interrogator that it is a female human being. If the Standard Interpretation were correct, we'd need an explanation of why Turing, having fully spelled out the Imitation Game, then gives a weaker game for the computer to play. But Turing emphasizes that his scenario is a difficult one for a computer, while acknowledging that a computer might fail it and still be intelligent.[11]

Having emphasized the importance of Turing's Imitation Game, I hasten to add that I am not suggesting that the Imitation Game is the only test Turing would have admitted as sufficient for establishing machine thinking. My criticism of the Standard Interpretation emphasizes two key aspects of the Imitation Game that the Standard Interpretation's tests miss. First, the subject matter of the game must be open-ended. Conversations should lead from the subject introduced by the game to other subjects easily. Second, the playing of the game must not turn on the participants' views about the nature of computers. There are other necessary conditions, of course, such as the requirement that participants use the same natural language, as well as the conditions specifying the mode of communication. Another restriction, suggested but not made explict by Turing, is some sort of cultural alignment among participants. An Imitation Game is unlikely to work, even with just human players, if the players come from very different cultures. Thus, assuming a shared cultural background, the game could be one of determining which of A or B is a resident of a large U.S. city, or which of A or B is a Democrat, or which is a member of the American Rifle Association. I'm inclined to think that none of these alternatives are as good starting points as Turing's Imitation Game, because concepts related to gender are so pervasive in so many cultures.

French argues that the admission that a test for intelligence must appeal to culturally-based cognition, or what French calls "subcognitive processes" is a devastating blow to the test (French, 1990). If the intelligent discourse depends on contingent psychological associations made by an agent embedded in a culture, a computer created in isolation from the culture of its interlocutors lacks the relevant subeognitive processes, and will be unmasked in any Imitation Game.

This objection, like the Standard Interpretation, depends on overemphasizing the role of imitation in Turing's test. French is surely right that any successful participant in the Imitation Game will have to have to share many of the cultural associations of its fellow players. It does not follow that a computer could not acquire those associations by a different route from its human counterparts. In fact, while Turing expects that the best way to have the computers acquire cultural knowledge is by placing them in the same kind of learning environments humans typically experience, he does not commit himself to the claim that a computer that passes the test will provide a model of the psychological reality of human cognition. That is, the internal processes of the computer may not match up with human internal cognitive processes, even where both humans and computers learn the relevant associations through the same environmental stimuli.[12] As we've noted, the interrogator's presumption in the original Imitation Game is that the other two participants are human beings. Thus, if there are empirically discoverable subcognitive differences, a virtue of placing a computer the original Imitation Game is that those differences need not surface. That's a virture, if, as Turing holds, such differences don't make a difference to the question of intelligence.

It is worth acknowledging that some commentators who invoke the Standard Interpretation explicitly allow that the interpretation deviates from Turing's formulation. Such authors sometimes refer to "revised" or "restricted" Turing Tests. On my interpretation, the kinds of tests that count as revised or restricted Turing Tests are not worthy of the name. They are not alternatives to Turing's Imitation Game, but separate and weaker tests, tests that Turing would not have endorsed as sufficient for intelligence.

Notes

[1]An earlier version of this paper was presented at Turing 2000: The Dartmouth Conference on the Future of the Turing Test, January 28–30, 2000. I'd like to thank Jack Copeland, Paul Hovda, James Moor, and Gualtiero Picciini for helpful comments and suggestions.

[2]Some versions of the Standard Interpretation characterize the test as involving oniy two participants, the computer and an interrogator. The computer's performance is not compared to another player. Instead the computer passes the test when the interrogator fails to identify it as a computer.

[3]Dyer's interpretation is unusual because he posits more than one interrogator, and he calls for a split decision from the group of interrogators in one case. Most readers interpret Turing as advocating many runs of the test, with one interrogator per test, but different interrogators for each run.
[4]Shieber (1994) is another example.
[5]See Moor (1987) for a constrasting view. Moor thinks that role the computer plays is left undetermined by Turing. This is a weaker position than the Standard Interpretation.
[6]Cf. Kenner (1986).
[7]See Piccinini (2000) for the Standard Interpretation reading of this passage.
[8]This is Hodges' defense of the Standard Interpretation. See Hodges (1999, pp. 37 ff).
[9]Turing saw ordinary views about computers as an obstacle to progress in computer science generally (Turing, 1947).
[10]This point depends on the assumption that the game is set up in such a way that the interrogator assumes that the human players are normal speakers of the language. If Turing did not make this assumption, neither version of the Test would have a point.
[11]"May not machines carry out something which out to be described as thinking but "which is very different from what a man does? This objection is a very strong one"
[12]In the 1947 lecture, Turing cites Shannon's work on using heuristics, and he says that he wouldn't consider game-playing success with the use of heuristics significant. This suggests that Turing thought that the coding of human cognitive processes was not the best way to design intelligent machines.

References

Armer, Paul (1995), 'Attitudes towards intelligent Machines', in Edward A. Feigenbaum and Julian Feldman, eds., *Computers and Thought*, MIT Press.

Block, Ned (1990), 'The Computer Model of the Mind', in D. Osherson and E. Smith, eds., *An Invitation to Cognitive Science*, Volume 3, MIT Press.

Davidson, Donald (1990), 'Turing's Test', in K.A. Mohyeldin Said, W.H. Newton-Smith, R. Viale and K.V. Wilkes, eds., *Modelling the Mind*, Clarendon Press.

Dietrich, Eric (ed.) (1994), *Thinking Computers and Virtual Persons*, Academic Press.

Dyer, Michael G. (1994), 'Intentionality and Computationalism', in Dietrich.

Hodges, Andrew (1999), *Turing*, Routledge.

Hutchens, Jason L. 'Introducing Megahal', http://ciips.ee.uwa.edu.au/~ hutch/hal/Ta1k.html

Kenner, Hugh (1986), 'Racter: Artificial Intelligence or Canned Lunacy?' *Byte*, May, 1986, p. 289.

Michie, Donald (1993), 'Turing's Test and Conscious Thought', *Artificial Intelligence* 60, pp. 1–22, reprinted in Millican and Clark (1996).

Millican, Peter and Clark, Andy (1996), *Machines and Thought*, Oxford University Press.

Moor, J.H (1976), An Analysis of the Turing Test', *Philosophical Studies* 30, pp. 249–257.

Moor, J.H. (1987), 'Turing Test', in Stuart C. Shapiro, ed., *Encyclopedia of Artificial Intelligence*. New York: John Wiley and Sons, pp. 1126–1130.

Piccinini, Gualtiero (2000), 'Turing's Rules for the Imitation Game', *Minds and Machines* 10, pp. 573–582

Rappaport, William J. (1994), 'Syntactic Semantics', in Dietrich.

Shieber, Staurt M. (1994), 'Lessons from a Restricted Turing Test', *Communications of the Association for Computing Machinery*, 37(6), pp. 70–78.

Turing, A.M (1996), 'Lecture to the London Mathematical Society on 20 February, 1947' in B.E. Carpenter and R.W. Doran, eds., *A.M Turing's Ace Report of 1946 and Other Papers*, MIT Press.

Turing, A. M. (1950), 'Computing Machinery and Intelligences', *Mind* 59, pp. 433–460.