# THE IMITATION GAME

MARK BISHOP

This issue of the Kybernetes journal is concerned with the philosophical question - Can a Machine Think? Famously, in his 1950 paper 'Computing Machinery and Intelligence' [11], the British mathematician Alan Turing suggested replacing this question - which he found "too meaningless to deserve discussion" - with a simple [behaviourial] test based on an imagined Victorian-esque pastime he entitled the 'imitation game'. In this special issue of Kybernetes a selection of authors with a special interest in Turing's work (including those who participated in the 2008 AISB[1] symposium on the Turing test[2]) have been invited to explore and clarify issues arising from Turing's 1950 paper on the imitation game; now more widely known as the Turing test.

As early as 1941 Turing was thinking about machine intelligence [3] - specifically how computing machines could solve problems by searching through the space of possible problem solutions guided by heuristic principles. And in 1947 Turing gave what is perhaps the earliest public lecture on machine intelligence at the Royal Astronomical Society, London. Subsequently, in 1948, following a year's sabbatical at Cambridge, Turing completed a report for the UK's National Physical Laboratory on his research into machine intelligence, entitled *Intelligent Machinery* [10]. Although not published contemporaneously, the report is notable for predicting many core themes which eventually emerged from the yet nascent science of machine intelligence: Expert Systems; Connectionism; Evolutionary Algorithms; but most intriguingly of all in the context of this special issue, the report offers perhaps the earliest version of the imitation game / Turing test. Turing presents this original version as follows:

---

[1]The (British) society for the study of Artificial Intelligence and the Simulation of Behaviour.

[2]To coincide with the Loebner Prize competition held at the University of Reading (UK) on the 12th October 2008, the AISB elected to sponsor a one day 'invited-speaker' symposium to present an alternative, formal, academic critique of issues around the Turing test. The day commenced with talks from three eminent speakers (Baroness Greenfield; Michael Wheeler & Selmer Bringsjord) who offered a personal context to, and their perspective on, the Turing test. These presentations were followed in the afternoon session with the invitation to four subsequent speakers (Andrew Hodges; Luciano Floridi; Margaret Boden; Owen Holland) to address specific matters related to the Turing test (e.g. definitional; adequacy; tests in other modalities; technical/computational issues). The day ended with a short round-table discussion regarding some of the issues raised during the day

*"The extent to which we regard something as behaving in an intel-
ligent manner is determined as much by our own state of mind and
training as by the properties of the object under consideration. If
we are able to explain and predict its behaviour or if there seems to
be little underlying plan, we have little temptation to imagine intel-
ligence. With the same object therefore it is possible that one man
would consider it as intelligent and another would not; the second
man would have found out the rules of its behaviour.*

*It is possible to do a little experiment on these lines, even at the
present stage of knowledge. It is not difficult to devise a paper
machine which will play a not very bad game of chess. Now get
three men as subjects for the experiment A, B, and C. A and C
are to be rather poor chess players, B is the operator who works the
paper machine. (In order that he should be able to work it fairly fast
it is advisable that he be both mathematician and chess player.) Two
rooms are used with some arrangement for communicating moves,
and a game is played between C and either A or the paper machine.
C may find it quite difficult to tell which he is playing.*

*(This is a rather idealized form of an experiment I have actually
done.[3])*

Subsequently, in the initial exposition of the imitation game presented in the
1950 paper [11], Turing called for a human interrogator (C) to hold a conversation
with a male and female respondent (A and B) with whom the interrogator could
communicate only indirectly by typewritten text. The object of this game was for
the interrogator to correctly identify the gender of the players (A and B) purely
as a result of such textual interactions; what makes the task non-trivial is that
(a) the respondents are allowed to lie and (b) the interrogator is allowed to ask
questions ranging over the whole gamut of human experience. At first glance it is
perhaps mildly surprising that, even after many such textual interactions, a skilled
player can determine (more accurately than by chance) the correct gender of the
respondents.[4]

---

[3]Some commentators (e.g.Whitby [1] and Shah & Warwick (herein)) have suggested that Tur-
ing didn't intend the imitation game to be the specification of some fully operational procedure
to be performed by future machine intelligence researchers as a yardstick with which to evaluate
their wares, but merely as a thought experiment, a "philosophical ice-breaker" (ibid), "attempt-
ing to deal with the ill-definition .. of the question .. can machines think?" (Wiggins [15]). The
fact that Turing personally enacted this first version of the imitation game offers perhaps partial
evidence against this interpretation.

[4]Here Turing's Victorian-esque parlour game describes a scenario perhaps not unfamiliar to
that many twenty-first century video gamers encounter when participating in a large multi-user
virtual world - such as World of Warcraft or Second Life - where in-game avatars controlled by
real-world players may often fail to reflect the gender they claim to be; the controller may be
female and the avatar male or vice versa.

Turing then asked the question - What will happen when a machine takes the part of (A) in this game? Would the interrogator decide wrongly as often as when playing the initial imitation game? In this flavour of the imitation game / Turing test - which has become known as the 'standard interpretation' - a suitably programmed computer takes the part of either player (A) or player (B) (i.e. the computer plays as either the man or the woman) and the interrogator (C) simply has to determine which respondent is the human and which is the machine[5].

However, a close reading of the 1950 paper reveals several other possible interpretations other than the standard version outlined above. For example it is possible to interpret Turing when he says: "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?" as meaning:

> (a) literally what he says - that the computer must pretend to be a woman, and the other participant in the game actually is a woman (see Genova [6] and Traiger [9]);
>
> (b) that the computer must pretend to be a woman, and the other participant in the game is a man who must also pretend to be a woman[6].

Although in a very literal sense, the above present valid alternative interpretations of the imitation game, the core of Turing's 1950 article (and material in other articles that Turing wrote at around the same time) strongly support the claim that Turing actually intended the standard interpretation[7].

In the 1950 paper Turing confidently predicted that by the year 2000 there would be computers with 1G of storage, (which turned out to be a relatively accurate prediction), which would be able to perform the [standard] Turing test such that the average interrogator would not have more than 70% chance of making the right identification after five minutes of questioning; the latter claim being slightly ambiguous :- did Turing intend the imitation game to be played out over five

---

[5]Although it is implicit in this 1950 version of the imitation game that the interrogator knows at least one of the respondents is a machine, a subsequent version - presented in a radio discussion in 1952 [12] - describes a 'jury' of interrogators questioning a number of entities seriatim; some entities being computers, some being human. Clearly, during each interrogation in this version of the test, the jury does not know if they are interacting with a human or a machine. Similarly when Colby et al tested PARRY, they did so by assuming that the interrogators did not need to know that one or more of those being interviewed was a computer during the interrogation [2]. Copeland, in commenting on the revised 1952 test [5], argues that the 1950 version is the better, as the single interview mode is open to a "biasing effect which disfavours the machine".

[6]Towards the end of section (5) of the 1950 paper [11] Turing, perhaps rather confusingly suggests, [the computer] "can be made to play satisfactorily the part of (A) in the imitation game, the other part being taken by a man".

[7]See Copeland [4], Piccinini [8], and Moor [7].

minutes of questioning in total or did he mean five minutes of questioning per respondent?

Furthermore, although Turing specifically describes playing the imitation game with 'average' interrogators, some commentators - perhaps remembering Kasparov's titanic series of games against chess playing machines - hint at a 'strong' version of the imitation game; where the interrogator is an expert interrogator, the game is played as an open ended conversation and the test is for full 'human indistinguishability' (cf. Hugh Loebner's specification for the gold medal prize in his version of the Turing test[8]).

In 2008 the organisers of the annual Loebner bronze medal Prize elected to put Turing's 1950 prediction to the test in the first - least demanding - manner, by enacting a set of five minute Turing tests for the bronze medal prize; specifically, each interrogator was allowed a total of five minutes to respond to both entities (the human and the computer). As a consequence the expected interaction time with the computer program was just two and a half minutes. However, in 2008 even this minimal 'five minute' claim proved optimistic as Elbot - evaluated as the best computer program in this competition - achieved a maximum deception rate of 25% over two and a half minutes of interaction; still 5% short of the 30% deception rate Turing had predicted in 1950.

Nonetheless it seems very likely that in the next few years Turing's predictions for a 'time limited' Turing test will be met; whether that means at that juncture "general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted" (as Turing asserted) is very doubtful as, in the 50 plus years since the paper was first published, the status of the Turing test as a definitive measure of machine intelligence and understanding has been extensively critiqued[9].

In 2008 the AISB sponsored an invited speaker symposium on the Turing test at the University of Reading in the hope of eliciting further clarity in the interpretation of the test, further insight into its implications and further reflection

---

[8]In 1990 Hugh Loebner agreed with 'The Cambridge Center for Behavioural Studies' to underwrite a contest designed to implement a Turing-style test. Dr. Loebner pledged (i) a 'Grand Prize' of $100,000 and a 'solid 18 carat gold medal' for the first computer program whose responses were *indistinguishable from a human's* and (ii) an annual prize - $3,000 in 2010 - and bronze medal to be awarded to the most *human-like* computer program (i.e. the best entry relative to other entries that year, irrespective of how good it is in an absolute sense). A comprehensive description of the 2008 competition by Huma Shah and Kevin Warwick is presented herein.

[9]Perhaps the best known criticism of 'a Turing style test of machine understanding' comes from John Searle. In the Chinese Room Argument (CRA) [14] Searle endeavours to show that even if a computer behaved in a manner fully indistinguishable from a human (when answering questions about a simple story) it cannot be said to genuinely understand its responses and hence the computer cannot be said to genuinely think (for recent discussion of the CRA see Preston & Bishop [13]).

as to its status as a [practical] measure of machine intelligence[10]. However the breadth and depth of the material presented in this special issue clearly illustrate that Turing's imitation game continues to present novel insights into mind and machine and any hope for a 'near final' word on the imitation game remains as far off as ever. On behalf of the AISB and Kybernetes, I would like to thank all the expert contributors to both the 2008 symposium and this volume for making time to address these ever fascinating issues.

- Mark Bishop (Chair elect AISB and Professor of Cognitive Computing at Goldsmiths, University of London), November 2009.

## References

[1] Whitby, B. (1996) The Turing Test: AI's Biggest Blind Alley? In: Millican, P. & Clark, A. (eds) (1996) Machines and Thought: The Legacy of Alan Turing. Mind Association Occasional Series 1, pp. 53-62. New York & Oxford: Oxford University Press.

[2] Colby, K.M., Hilf, F.D., Weber, S. & Kraemer, H.C. (1972) Turing-like indistinguishability tests for the validation of a computer simulation of paranoid processes, Artificial Intelligence 3: 199–221

[3] Copeland, B.J. & Proudfoot, D. (2005) Turing and the Computer, in Copeland, B.J. (ed) (2005) Alan Turing's Automatic Computing Engine, Oxford & New York: Oxford University Press.

[4] Copeland, B.J. (2000), The Turing Test, Minds and Machines, 10: 519–39.

[5] Copeland, B.J. (ed) (2004) The Essential Turing. Oxford: Clarendon Press: 488.

[6] Genova, J. (1994) Turing's Sexual Guessing Game, Social Epistemology, 8: 313–26.

[7] Moor, J. (2001), The Status and Future of the Turing Test, Minds and Machines, 11: 77-93.

[8] Piccinini, G. (2000), Turing's Rules for the Imitation Game, Minds and Machines, 10: 573-85.

[9] Traiger, S. (2000) Making the Right Identification in the Turing Test, Minds and Machines, 10: 561-572.

[10] Turing, A.M. (1948) Intelligent Machinery. National Physical Laboratory Report, 1948, in Copeland, B.J. (ed) (2004) The Essential Turing, Oxford & New York: Oxford University Press.

[11] Turing, A.M. (1950) Computing Machinery and Intelligence, Mind 59, pp. 433–460.

[12] Turing, A.M., Braithwaite, R., Jefferson, G. & Newman, M. (1952) Can Automatic Calculating Machines Be Said To Think? in Copeland, B.J. (ed) (2004) The Essential Turing. Oxford: Clarendon Press: 487–506.

[13] Preston, J. & Bishop, M. (eds) (2002) Views into the Chinese Room: New Essays on Searle and Artificial Intelligence, Oxford & New York: Oxford University Press.

[14] Searle, J. (1981) Minds, Brains, and Programs, Behavioral and Brain Sciences, 3: 417-57.

[15] Wiggins, G. (2007) Personal communication, Geraint Wiggins, Professor of Computational Creativity, Goldsmiths, University of London, UK.

---

[10]The AISB will host a second Turing symposia at their spring convention 2010 to continue discussion of this question http://www.cse.dmu.ac.uk/~aayesh/TuringTestRevisited/Welcome.html