Journal of Robotics and Automation

The Turing Test and Android Science

Gilberto Marzano*

Rezekne Academy of Technologies, Rezekne, Latvia

Abstract

This article considers the Turing's *Imitation Game* in the light of behavioral robotics.

As is well-known, Turing's classic test was based on the simultaneous comparison by an interrogator of two hidden entities, one being human and the other being a machine. If the interrogator, who was human, was not able to correctly identify the responders by saying for certain which one was the machine and which was the human the machine passed the test and one could claim that it exhibited an intelligent behavior.

Nowadays, advances in robotics have introduced new issues, and the classic Turing test would appear to be inadequate. Indeed, intelligent behavior cannot be referred exclusively to the cognitive sphere and in fact Android Science suggests including physical behavioral elements in the comparison too.

In this article, the Ishiguro total test is illustrated, and a challenging extension of the Turing test, in which the interrogator can also be a machine, is hypothesized and highlighted as a premise for designing a reverse Turing test.

Keywords

Turing's imitation game, Machine misidentification, Android science, Ishiguro's total Turing test, Comprehensive Turing test, Reverse Turing test

Introduction

The new generation of androids and gynoids, the masculine and feminine forms of robots respectively, are designed to look and act like human beings.

Many lines of progress have been made since the Cog robot, designed two decades ago, that was equipped with a few degrees of freedom and a variety of sensory systems, including visual, auditory, vestibular, kinesthetic, and tactile senses [1]. Cog was one of the first robots to be developed following the notion that, to be similar to a human, a robot should not only have to exhibit logical and analytical traits but also be able to emulate human behavior. The Cog's designers were persuaded that four attributes specific to human intelligence should come to bear on robot behavior: Developmental organization, social interaction, embodiment and physical coupling, and multimodal integration [2,3]. This view is broadly shared today in the scope of *behavioral robotics*, in an approach that focuses on a robots' ability to exhibit complex human-appearing behaviors.

In the last decade, astonishing advances have been achieved in *behavioral robotics* [4,5] since powerful software and new smart interactive components have greatly expanded machine learning capability, image interpretation, and data mining, as well as human-robot interaction and machine perception proficiency [6,7].

In the perspective of behavioral robotics, in the early 2000s, Ishiguro, the director of the Intelligent Robotics Laboratory at Osaka University in Japan, introduced the term "*Android Science*" to designate an interdisciplinary research sector that encompasses two complementary approaches: The use of cognitive science to build very humanlike robots, and the use of robots for verifying hypotheses to understand humans [8]. Ishiguro shares the idea of Cog's designers that "to build an artificial system with similar grounding to a human system it is necessary to build a robot with human form" [9].

Nowadays, although the debate on the effects of advances in robotics focuses on issues of the labor market, there are also a few challenging questions that **behavioral robotics** is raising in various other social spheres, such as ethics [10], philosophy [11], and religion [12].

Moreover, both in popular and scientific literature [13-15], it is common to come upon the question "can a robot think like a human being?" that seems, in all respects, like a new form of the old question "can machines think?" for which Turing formulated his famous test in 1950.

In this regard, a question arises: Is the Turing test still valid after such a long time?

This article introduces some considerations regarding Turing's test in the light of behavioral robotics.

Turing's Imitation Game

In 1950, in his famous paper *Computing Machinery and Intelligence*, published in the philosophy journal *Mind*, Turing suggested the *Imitation Game* to replace the question "can machines think?", which for him was too ambiguous [16].

Turing proposed a test based on the popular party game whereby a player, interrogating, through an intermediator, another person who they cannot see or hear, must determine if the mystery person is a man or a woman.

In Turing test, a human player (the interrogator) interrogates two other hidden players, one a computer and one a human, by issuing written questions and receiving written responses using natural language. The evaluator has to determine which of the players is the computer and which is the human based on the qualities of the responses. Turing argues that if the machine appears to have a human-like behavior, one can assume that it exhibits an *intelligent* behavior. Of course, that is quite different from assuming that it thinks. Indeed, the Turing test was designed to assess the ability of a machine to reproduce the human behavior. Its relevance lies on its simplicity and generality, as demonstrated by a rich literature that encompasses multifarious scientific sectors [17-24].

Turing predicted that by the year 2000, technological progress would produce a computer extraordinarily powerful enough that a program would be able to fool the average evaluator for 5 minutes on about 70% of occasions: 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 af-

ter five minutes of questioning. […] I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted [16].

Following this, other scientists have varied Turing's original idea, and various similar tests have been proposed, the most widely known of which are:

- The Coffee Test [25]. A machine is given the task of going into an average American home and figuring out how to make coffee. It has to find the coffee machine, find the coffee, add water, find a mug, and brew the coffee by pressing the proper buttons. However, this test appears extremely specific and very restrictive.

- The Robot College Student Test [26]. A machine is given the task of enrolling at a university, taking and passing the same classes that a human would, and obtaining a degree.

- The Employment Test [27]. We have to test whether or not we have the capability to automate a specific job.

Recently, a modified Turing test has been proposed for testing the next generation of medical diagnostics and medical robotics [28].

Turing, in his test, deliberately avoided any direct physical interaction between the interrogator and the computer, since he deemed that the physical simulation of a human was unnecessary for intelligence. Later, the first Artificial Intelligence approaches shared the same opinion and tended to consider intelligence as closely related to the cognitive sphere [29]. This view continues to be widely shared by those who, speaking about a computer or a human being, analogize software with the mind and hardware with the brain. This metaphorical relationship reflects a dualistic bias since the reality, at any rate in the field of computers, is quite different. Software produces the expected results if it runs on an appropriate hardware. Many intelligent software solutions have been made possible thanks to the notable progress made in hardware. An evident example is the case of chess programs where the brute force of a computer's massive data storage and processing speed can win against the ability of a human player.

However, in a posthumous essay, Turing himself shows a revolutionary view about future machines:

My contention is that machines can be constructed which will simulate the behaviour of the human mind very closely. They will make mistakes at times, and at times they may make new and very interesting state-

ments, and on the whole the output of them will be worth attention to the same sort of extent as the output of a human mind[a] [30].

Furthermore, according with Brooks, we have to underline that Turing carefully considered the question of embodiment, and "only for technical reasons chose to pursue aspects of intelligence which could be viewed, at least in his opinion, as purely symbolic" [31].

Nowadays, the internet has demonstrated the crucial role of social interactions and their relevant contribution to human intelligent behavior, whilst the *Internet of Things* [32] is opening new dimensions of intelligent interactions and suggesting new forms of distributed intelligence [33].

The Ishiguro Total Turing Test

Recently, as a result of research in the field of Android Science, a new version of the Turing test, the so-called *total Turing test*, has been formulated. Ishiguro suggested a total Turing test that also included a video signal so that the interrogator can test the subject's perceptual abilities, as well as the capability to move physical objects through a hatch. Accordingly, to pass the total Turing test, the android needs to be equipped with vision tools to be able to perceive objects, and movement tools to be able to manipulate them.

Ishiguro claimed that the original Turing test was designed to evaluate the intelligence of a computer under the assumption that mental capacities could be abstracted from embodiment, whilst "the android enables us to evaluate total intelligence" [8].

As with Turing's original test, Ishiguro's total Turing test involves a time-competition.

In Ishiguro's test [34], an android is showed on the computer screen and a human interrogator is asked to find out the colors of its clothes. The screen between the android and the subject opens for 2 seconds. After identifying the colors, the subject is asked whether they became aware that the other was an android. Two types of android, one static and the other with micro-movements, are prepared.

Employing this experiment on a sample of 20 subjects, the results demonstrated that 70% of the subjects were not aware that they were dealing with an android when the android could perform micro-movements.

---

[a]Intelligent machinery, a heretical theory', was a lecture given to '51 Society' at Manchester. There are 2 versions of the text of this lecture at the Turing Digital Archive, one TS numbered 1-10, the other CTS numbered 96-101. c. 1951; available at: http://www.turingarchive.org/viewer/?id=458&title=1; last accessed 11.12.2017.

A Challenging Extension of the Total Turing Test

Following from the work of Ishiguro, and with regard to the recent advances made in behavioral robotics, one can propose a challenging extension of the Turing total test that has been designed as the *reverse Turing test.*

We can hypothesize that, in light of the likely future evolution of the relations between humans and robots, the roles of the machine and the human should be reversed, so that the interrogator would be a robot rather than a human.

The robot's task would be to recognize the true nature of the respondents, namely, that it must determine which of the respondents is the machine and which is the human. This test would allow the level of efficiency of a human-like robot to be assessed. By all means, the robot interrogator ought to be programmed in order to be able to elaborate questions autonomously. This brings us to the new fundamental issue facing machine learning: Implementing intelligent machines that are self-modifiable. Indeed, new intelligent systems and robots should be increasingly autonomous and self-determining.

The proposed test will combine two different abilities of a machine: Language-based communication and vision. The interrogator and the respondents would be in three separate rooms. Respondents' rooms would be equipped with a video camera mounted in a corner of the room, connected to a monitor that is placed in the interrogator's room.
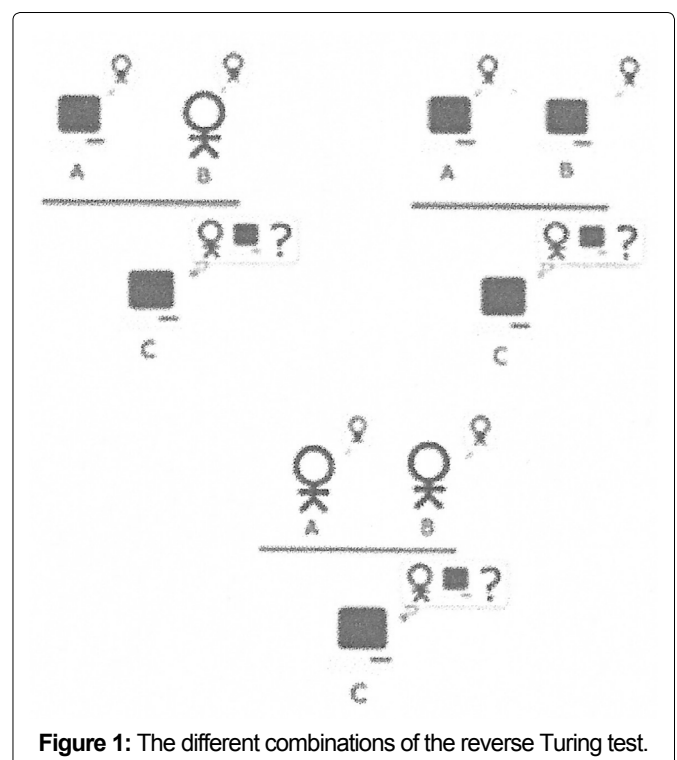

Figure 1: The different combinations of the reverse Turing test.

The interrogator would pose questions, receiving both an answer in natural language from the respondent and a few seconds of live feed from the room they are in.

Figure 1 illustrates the possible interaction between the interrogator (C), namely the machine, and the respondents (A, B), for which three different pairings are possible - a machine and a human, two machines, or two humans.

It is not important whether the communication between interrogator and respondents is spoken or written, although spoken communication would be preferable.

To pass the test, the machine must recognize the real nature of the respondents by using no more than four questions, although the precise number of questions and the length of time of the live image of the respondents could vary. This would mean that different levels of the test could be designed.

In consideration of the ongoing advances in behavioral robotics, we can imagine that, in the near future, the interrogator and the respondents would stay in the same room. In this case, the robot would be advantaged in discriminating between robot and human. Indeed, information about the emitted infrared spectrum of the observed subject would provide a very good discriminative stimulus [35].

As a variant of the reverse Turing test, we can hypothesize a networking test. This would entail creating a group comprising both human and robot subjects (at least two of each), and asking each subject to identify the nature of the others.

Conclusion

The literature shows that Turing test remains a controversial issue that polarizes different opinions at scientific and philosophic levels [36].

Indeed, it can legitimate many different interpretations.

For example, what is intriguing in the Turing test is that the machine tries to cheat the human interrogator by pretending to be a human in terms of conversational abilities. This explains why Eliza, the program by Weizenbaum that emulates the interaction of a Rogerian psychotherapist with a patient, should continue to be associated with the Turing test [37,38].

However, most of those who work in the field of robotics and Artificial Intelligence want to achieve robust and efficient solutions whose objective is not to fool a human being. Building computer-based solutions that emulate human behavior is aimed at eliminating onerous and repetitive activities or extending services to humans. Who would like to invest in the development of a robot

that lies, is inattentive, jealous, malicious, and unfair?

Indeed, the current advances in robotics allow us to conceive a test in which robots and humans could compare their mutual recognition skills without attempting to cheat each other.

The reverse Turing test, in which the interrogator is a robot, provides a way to investigate the future dimension of the human-machine relationship. Indeed, one can see the promising improvements in the scope of so-called *Artificial General Intelligence* (AGI) [39,40], whose aim is the development of programs and machines that can successfully perform any intellectual task that a human being can.

In this regard, I am persuaded that Turing test, for its visionary approach, will continue to stimulate fertile reflections in the field of Artificial Intelligence and to be a vital part of every study on the subject.

Indeed, from the preliminary formulation of the proposed reverse Turing test, a few challenging questions arise. For example, what does it mean if the interrogator doesn't recognize the identity of a respondent that is a machine? Moreover, what sort of humans we have to choose to compete against machines? Or, what kind of questions should the machine ask for? These points require a more thorough analysis. However, the advantage of the proposed test is the possibility to implement an experimental version, on which I am working on, based on the currently available technology.

References

1. Brooks RA, Breazeal C, Marjanovic M, et al. (1999) The Cog project: Building a humanoid robot. Lecture Notes in Computer Science 1562: 52-87.

2. Brooks RA (1994) Coherent behavior from many adaptive processes. In: Cliff D, Mayer JA, Wilson SW, Proceedings of the Third International Conference on Simulation of Adaptive Behavior, From animals to animats 3. MIT Press, 3: 22-29.

3. Brooks RA (1997) The cog project. Journal of the Robotics Society of Japan 15: 968-970.

4. Shirai Y, Hirose S (2012) Robotics research: The eighth international symposium. Springer Science & Business Media.

5. Siciliano B, Khatib O (2016) Springer handbook of robotics. Springer.

6. Christensen HI, Okamura A, Mataric M, et al. (2016) Next generation robotics.

7. Seibt J, Nørskov M, Andersen SS (2016) What social robots can and should do: Proceedings of robophilosophy 2016/TRANSOR 2016. IOS Press, 290: 424.

8. Ishiguro H (2005) Android science: Towards a new cross-disciplinary framework. In: Proc Cog Sci Workshop: Toward Social Mechanisms of Android Science, 1-10.

9. Brooks RA (1998) Using human development as a model for adaptive robotics. Robotics Research 339-343.

10. Lin P, Abney K, Bekey GA (2011) Robot ethics: The ethical and social implications of robotics. MIT Press.

11. Tzafestas SG (2016) An introduction to robophilosophy. River Publishers.

12. Bolonkin A (2011) Universe, human immortality and future human evaluation. (1st edn), Elsevier.

13. Olney AM (2007) Dialogue generation for robotic portraits. In: Proceedings of the international joint conference on artificial intelligence 5th workshop on knowledge and reasoning in practical dialogue systems, 15-21.

14. Bollebakker S, Piel H (2016) Towards a robotic dystopia? Replacing animal companions with technology in science and fiction. Digression 2: 16-29.

15. Li THS, Kuo PH, Ho YF, et al. (2016) Robots that think fast and slow: An example of throwing the ball into the basket. IEEE Access 4: 5052-5064.

16. Turing A (1950) Computing machinery and intelligence. Mind 59: 433-460.

17. Lin L, Hu PJH, Sheng ORL (2006) A decision support system for lower back pain diagnosis: Uncertainty management and clinical evaluations. Decision Support Systems 42: 1152-1169.

18. Herbert-Read JE, Romenskyy M, Sumpter DJ (2015) A Turing test for collective motion. Biol Lett 11: 20150674.

19. Geman D, Geman S, Hallonquist N, et al. (2015) Visual turing test for computer vision systems. Proceedings of the National Academy of Sciences 112: 3618-3623.

20. Arnold T, Scheutz M (2016) Against the moral Turing test: Accountable design and the moral reasoning of autonomous systems. Ethics and Information Technology 18: 103-115.

21. Schoenick C, Clark P, Tafjord O, et al. (2017) Moving beyond the turing test with the allen ai science challenge. Communications of the ACM 60: 60-64.

22. Kulkarni RH, Padmanabham P (2017) Integration of artificial intelligence activities in software development processes and measuring effectiveness of integration. IET Software 11: 18-26.

23. Lowe R, Noseworthy M, Serban IV, et al. (2017) Towards an automatic Turing test: Learning to evaluate dialogue responses. Proceedings of the 55th annual meeting on Association for Computational Linguistics, 1116-1126.

24. Montenegro JMF, Argyriou V (2017) Cognitive evaluation for the diagnosis of Alzheimer's disease based on Turing Test and Virtual Environments. Physiol Behav 173: 42-51.

25. Goertzel B (2014) Artificial general intelligence: Concept, state of the art, and future prospects. Journal of Artificial General Intelligence 5: 1-46.

26. Goertzel B, Iklé M, Wigmore J (2012) The architecture of human-like general intelligence. In: Wang P, Goertzel B, Theoretical foundations of artificial general intelligence. Atlantis Press 4: 123-144.

27. Nilsson NJ (2005) Human-level artificial intelligence? Be serious! AI Magazine 26: 68-75.

28. Ashrafian H, Darzi A, Athanasiou T (2015) A novel modification of the Turing test for artificial intelligence and robotics in healthcare. Int J Med Robot 11: 38-43.

29. Minsky M, Papert S (1970) Proposal to ARPA for research on artificial intelligence at MIT, 1970-1971.

30. Turing A (1951) Intelligent machinery, A Heretical theory. In: Shieber SM, The Turing test: Verbal behavior as the hallmark of intelligence. Mit Press, 105-110.

31. Brooks RA (1991) Intelligence without reason. Artificial intelligence: critical concepts 3: 107-163.

32. Whitmore A, Agarwal A, Da Xu L (2015) The Internet of Things-A survey of topics and trends. Information Systems Frontiers 17: 261-274.

33. Heylighen F, Lenartowicz M (2017) The Global Brain as a model of the future information society: An introduction to the special issue. Technological Forecasting and Social Change 114: 1-6.

34. Ishiguro H (2016) Android Science. In: Kasaki M, Ishiguro H, Asada M, et al. Cognitive Neuroscience Robotics. Springer, 193-234.

35. Shimura K, Ando Y, Yoshimi T, et al. (2014) Research on person following system based on RGB-D features by autonomous robot with multi-kinect sensor. In: System Integration (SII), 2014 IEEE/SICE International Symposium on, 304-309.

36. Warwick K, Shah H (2015) Can machines think? A report on Turing test experiments at the Royal Society. Journal of Experimental & Theoretical Artificial Intelligence 28: 989-1007.

37. Shieber SM (2016) Principles for Designing an AI Competition, or Why the Turing Test Fails as an Inducement Prize. AI Magazine 37: 91-96.

38. Abrams J (2017) Is Eliza human, and can she write a sonnet? A look at language technology. 31: 4-10.

39. Goertzel B, Orseau L, Snaider J (2014) Artificial general intelligence: 7th International Conference, AGI 2014, Quebec City, QC, Canada, August 1-4, 2014, Proceedings. Springer.

40. Müller VC, Bostrom N (2016) Future progress in artificial intelligence: A survey of expert opinion. In: Müller VC, Fundamental issues of artificial intelligence. Springer International Publishing, 553-571.

SCHOLARS.DIRECT