Editorial: Alan Turing and Artificial Intelligence

VAROL AKMAN
Department of Computer Engineering, Bilkent University, 06533 Ankara, Turkey
E-mail: akman@cs.bilkent.edu.tr; http://www.cs.bilkent.edu.tr/~akman

PATRICK BLACKBURN
Computerlinguistik, University of Saarland, D-66041 Saarbrücken, Germany
E-mail: patrick@coli.uni-sb.de; http://www.coli.uni-sb.de/~patrick

(Received ..... ; Accepted in final form .....)

> His [Turing's] point was that we should not be species-chauvinistic, or anthropocentric, about the insides of an intelligent being, for there might be inhuman ways of being intelligent.
> –Daniel C. Dennett

Alan Mathison Turing (23 June 1912–7 June 1954) was one of the most eminent scientists of the 20th century (Figure 1). His research was a central catalyst of the computer revolution. The concept of a Turing machine, which he developed in the 1930s, is still one of the most widely used models of computation in theoretical computer science, but this monumental contribution was only the first of many. His biographer Andrew Hodges — author of (A. Hodges, 1992) and (A. Hodges, 1997), and maintainer of the "Alan Turing Home Page"[1] — divides Turing's publications into five areas: mathematical logic, mechanical intelligence, pure mathematics, morphogenesis, and cryptanalysis. Moreover, in Sir Roger Penrose's words, Turing was also "a deep and influential philosopher in addition to his having made contributions to mathematics, technology and code-breaking that profoundly contribute to our present-day well being" (A. Hodges, 1998).

In a landmark article published in October 1950 in the philosophy journal *Mind* (Figure 2) Turing made a famous assertion. He predicted that by the year 2000 it would be feasible to write a program that would, after five minutes of questioning, have at least a 30% chance of fooling an average conversational partner into believing it was a human being (Turing, 1950).[2]

[1] See http://www.turing.org.uk/turing/biblio.html.

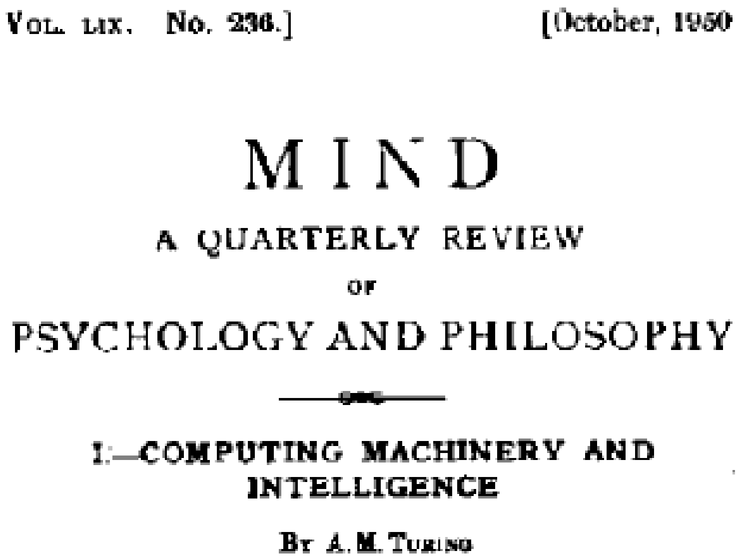
[2] As Charniak and McDermott (1985, p. 10) remark "Actually, the [*Mind*] paper makes it sound as if Turing had in mind the computer pretending to be a woman in the man/woman game, but the point is not completely clear, and most have assumed that he intended the test to be a person/computer one, and not woman/computer." See (Saygin et al., 1999) for an attempt at clarification.

Figure 1. Alan Turing in 1946. This is detail from a larger photograph which shows Turing with other members of an athletic club in Surrey. A serious runner, Turing achieved world-class marathon standards. Courtesy of Andrew Hodges.

Fifty years later, on January 28–29, 2000, *The Future of the Turing Test: The Next Fifty Years* was held at Dartmouth College. The purpose of this conference was to evaluate the merits and problems of the Turing Test, and the competition for the 2000 Loebner Prize (Anon., 1992) was held on January 28. Six programs participated, with Richard Wallace's *Alice* taking the first place. The overall results did not confirm Turing's prediction: it is reported that "[i]n this test judges were 91% correct after 5 minutes and 93% after 15 minutes. No computer was mistaken for a human."[3]

This special issue of *JoLLI* commemorates the golden anniversary of Turing's *Mind* article, and centers on the following question: just how influential have the ideas of Turing been in AI — and more importantly, just how influential *should* they be? There has long been controversy about this. For some the relevance of the Turing test to AI is unproblematic: in fact, Ginsberg (1993, p. 8) defines AI as follows: "Artificial intelligence is the enterprise of constructing a physical symbol system that can reliably pass the Turing test." Others find the the link problematic. For example, Luger and Stubblefield (1997, p. 12) argue that "The Turing test, in spite of its intuitive appeal, is vulnerable to a num-

[3] See http://www.dartmouth.edu/~phil/events/TuringTestConference.html. Also cf. "The Turing Test Page" at http://cogsci.ucsd.edu/~asaygin/tt/ttest.html and consult (Saygin et al., 1999) for a thorough review of Turing Test.


VOL. LIX. No. 236.] [October, 1950

MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND INTELLIGENCE

BY A. M. TURING

Figure 2. Turing's 1950 paper is one of the most cited in philosophical AI literature. Copyright © Oxford University Press.

ber of justifiable criticisms. One of the most important of these is aimed at its bias toward purely symbolic problem-solving tasks. It does not test abilities requiring perceptual skill or manual dexterity, even though these are important components of human intelligence." Others have raised fundamental objections to the test (the best known of these is Searle's Chinese Room Argument (Searle, 1984)) and concluded that the AI enterprise is ill-founded.

The papers in this special issue take as their point of departure the Turing Test (henceforth TT) as a criterion for thought and raise a variety logical, syntactic, semantic, pragmatic, or information-theoretic arguments to argue that any program that actually passes the test would — or would not — be a thinking entity. With the exception of the paper by Copeland and Proudfoot (which closes the volume) the contributions are arranged alphabetically by author.

The first contribution, *Animals, Zombanimals, and the Total Turing Test: The Essence of Artificial Intelligence*, by Bringsjord, Caporale, and Noel, argues that the TT has sent AI down a blind alley, at least as far as human intelligence is concerned. Indeed, they argue this not merely for the ordinary TT, but for Harnad's Total Turing Test (TTT), a more stringent version in which the judge can look at the contestants to determine which is a person and which a robot. In their view, the TTT accurately reflects the working assumptions of current agent-based approaches to AI. But while this Turing-inspired methodology opens the door to the construction of artificial animals, it lacks the power to produce genuine human intelligence.

Edmonds's *The Constructibility of Artificial Intelligence (as Defined by the Turing Test)* views the TT as a test of social competence. Edmonds argues that the ability to successfully enter into human social dynamics is not something that can be successfully programmed in advance: a considerable learning period is required, and many aspects of intelligence would have to be obtained as a result of interaction with the environment. In his view, the (potential) merit of the TT is that it could encourage a shift of attention from design-oriented paradigms back to consideration of the social roots of intelligence.

In *Minds, Machines and Turing: The Indistinguishability of Indistinguishables*, Harnad discusses a hierarchy of Turing Tests. Starting from the standard Test (dubbed T2), he moves onto a robotic Turing Test (T3) which extends the fundamentally symbolic nature of the T2 with external sensorimotor functionality.[4] The higher levels are T4, where the internal micro-functions are identical, and T5 where everything empirically detectable is the same. Harnad argues that T3 is the appropriate level for research in Cognitive Science, and discusses the implications of this stance.

The starting point of Hernandez-Orallo's *Beyond the Turing Test* is that the TT is unnecessarily anthropomorphic. What is really required, he argues, is an explicit computational formalization of what it means to comprehend something — and he makes a concrete proposal. Using Kolmogorov complexity (algorithmic information theory) as his tool, he proposes the C-test (C for "comprehension") as a criterion for understanding. While the C-test is only a necessary condition for comprehension, he argues that it signposts the way to what will be needed as research on cognition advances: progressively more accurate tests of cognitive abilities.

Rapaport's *How to Pass a Turing Test: Syntactic Semantics, Natural-Language Understanding, and First-Person Cognition*, is a proposal for

[4] Towards the end of his elegant little book (W. Hodges, 1977, p. 252), Wilfrid Hodges makes a proposal for an intelligent robot:

> [L]et us imagine ourselves constructing a robot who is to behave exactly like a human being. Being a robot, he does exactly what his internal instructions (known as his *programs*) tell him to do.

He then observes (1977, p. 253):

> It might be objected that any such program would have to be hideously complicated. The answer is that the program would need to be no more or less complicated than whatever it is in a human being which records the meaning of a word which he understands.

what he dubs "syntactic semantics." Contra Searle, he argues that syntax does suffice for real thinking. In essence, semantics is the process of modelling one domain in terms of another, a process which needs to be understood recursively. Suitably programmed computers will eventually pass the TT, and there is no barrier to claiming that such machines will think.

We close the special issue with Copeland and Proudfoot's *What Turing Did After He Invented the Universal Turing Machine*. Unlike the other contributions, it is not focussed on the TT: rather it documents the entire range of Turing's work in Computer Science and AI. Drawing on published and unpublished material (including correspondence with Turing's co-workers) the authors chronicle Turing's work on the first stored-program digital computers; his work on AI, connectionism, oracle machines, and Artificial Life; and the influence of Turing on Wittgenstein and vice-versa (both were fellows of Cambridge colleges in the 1930s). As their article makes clear, the TT is not the brain-child of a theoretician insulated from practical computational realities, nor the creation of a mathematically-brilliant but philosophically-uninformed mind. It is the considered product of one of the most creative figures of the 20-th century — and the problem space it so neatly delineates remains the focus of ongoing research.

The papers you will find here develop letter and spirit of Turing's original contributions. The do not lazily fall back into the same old sofa, but follow — or question — the inspiring ideas of a great man in the search for new, more precise, conclusions. It is refreshing to know that the fertile landscape created by Alan Turing remains a source of novel ideas.

Acknowledgements

We appreciate Dag Westerståhl's encouragement to go ahead with this special issue. We are grateful to our referees for all their hard work. Figure 1 comes from Andrew Hodges' "The Alan Turing Home Page" (http://www.turing.org.uk/turing). The opening quotation appears in (Dennett, 1998, p. 13).

References

Anon., 1992, "Can machines think?" *AI Magazine* **13**(2), 80–95.

Charniak, E., and McDermott, D., 1985, *Introduction to Artificial Intelligence*, Reading, MA: Addison-Wesley.

Dennett, D., 1998, "Can machines think?" pp. 3–20 in his *Brainchildren: Essays on Designing Minds*, Cambridge, MA: MIT Press.

Ginsberg, M., 1993, *Essentials of Artificial Intelligence*, San Mateo, CA: Morgan Kaufmann.
Hodges, A., 1992, *Alan Turing: the Enigma*, London: Vintage, Random House. Originally published in 1983.
Hodges, A., 1997, *Turing: a Natural Philosopher*, London: Phoenix.
Hodges, A., 1998, "Oration at Alan Turing's birthplace." Available at http://www.turing.org.uk/turing/bio/oration.html.
Hodges, W., 1977, *Logic*, London: Penguin.
Luger, G.F. and Stubblefield, W.A., 1997, *Artificial Intelligence: Structures and Strategies for Complex Problem Solving*, Reading, MA: Addison-Wesley Longman. Third edition.
Saygin, A.P., Cicekli, I, and Akman, V., 1999, "Turing Test: 50 years later," Department of Computer Engineering, Technical report BU–CEIS–9905, Bilkent University, Ankara. Available at http://cogsci.ucsd.edu/~asaygin/papers/tt.pdf.
Searle, J.R., 1984, *Minds, Brains and Science*, Cambridge, MA: Harvard University Press.
Turing, A., 1950, "Computing machinery and intelligence," *Mind* **59**(236), 433–460. Available at http://cogprints.soton.ac.uk/abs/comp/199807017.