

Turing's three philosophical lessons and the philosophy of information

BY LUCIANO FLORIDI^{1,2,3,4,*}

¹*Research Chair in Philosophy of Information, and ²UNESCO Chair in Information and Computer Ethics, University of Hertfordshire, de Havilland Campus, Hatfield AL10 9AB, UK*

³*Faculty of Philosophy, University of Oxford, 10 Merton Street, Oxford OX1 4JJ, UK*

⁴*Department of Computer Science, University of Oxford, Wolfson Building, Parks Road, Oxford OX1 3QD, UK*

In this article, I outline the three main philosophical lessons that we may learn from Turing's work, and how they lead to a new philosophy of information. After a brief introduction, I discuss his work on the method of levels of abstraction (LoA), and his insistence that questions could be meaningfully asked only by specifying the correct LoA. I then look at his second lesson, about the sort of philosophical questions that seem to be most pressing today. Finally, I focus on the third lesson, concerning the new philosophical anthropology that owes so much to Turing's work. I then show how the lessons are learned by the philosophy of information. In the conclusion, I draw a general synthesis of the points made, in view of the development of the philosophy of information itself as a continuation of Turing's work.

Keywords: fourth revolution; level of abstraction; philosophical anthropology; philosophy of information; Shannon

1. Introduction

When one looks at Turing's philosophical legacy, there seems to be two risks. One is to reduce it to his famous test [1]. This has the advantage of being clear-cut. Anybody can recognize the contribution in question and place it within the relevant debate, the philosophy of artificial intelligence. The other risk is to dilute it down into an all-embracing narrative, making Turing's ideas the seeds of anything we do and know today. This has the advantage of acknowledging the greatness of this genius. In both cases, however, we are less likely to identify which conceptual contributions by Turing have helped to shape our contemporary philosophical discourse, and which can direct its future development. In order to avoid both risks, in the following pages, I shall concentrate on three specific philosophical lessons, which seem to be particularly significant in view of the emergence of the philosophy of information and its

*l.floridi@herts.ac.uk

One contribution of 18 to a Theme Issue 'The foundations of computation, physics and mentality: the Turing legacy'.

subsequent development. I shall offer not a philological or scholarly analysis, but a minimalist, hermeneutical exercise. It is part of Turing's extraordinary genius that other interpreters will learn more and different lessons from his intellectual legacy.

The three philosophical lessons to which I wish to attract the reader's attention are: how Turing's work on the method of levels of abstraction (LoA) can teach us to ask philosophical questions properly; what philosophical questions are most pressing today, as a consequence of Turing's work; and, finally, Turing's influence in shaping our new philosophical anthropology, what I shall call the fourth revolution. I will then connect these lessons to the development of the philosophy of information.

2. Lesson 1: fixing the level of abstraction or how to ask philosophical questions

Imagine the following scenario. You ask the price of an item, let us say a second-hand car, and you receive the following answer: 5000. The question concerned a variable, the price x of the car in question, and you received an exact numerical value for x , yet something is missing. You still have no idea about the price because you do not know the *type* of the variable: is it British pounds, US dollars, euros, etc.? Of course, the context usually helps. If you are in the UK and you are asking a car dealer, your question should be understood as concerning the price in British pounds and so should the answer. This is trivial, you may think. Grice's conversational rules obviously apply. It is, and they do. But this is also a crucial assumption, easily forgotten. On November 1999, NASA lost the \$125 m Mars Climate Orbiter (MCO) because the Lockheed Martin engineering team used English (also known as Imperial) units of measurement, while the agency's team used the metric system for a key spacecraft operation. As a result, the MCO crashed into Mars.¹ Assuming that contexts will always disambiguate the types of your variables paves the way to costly mistakes. So what has all this got to do with Turing? Quite a lot, as it turns out. To show you why, let me introduce a slightly abstract model of factual information.²

We can treat factual information of the kind illustrated above by the price of the second-hand car as a compound of question + answer. If some theoretical simplification is allowed, the question may be reduced to a Boolean one, followed by a yes or a no answer. In the original version of our example, the price of the second-hand car then becomes: [is the price of this car 5000 + yes]. You see immediately that the problem is not in the answer, but in the question: it contains no indication of the type of the variable being handled. The correct piece of information is of course: [is the price of this car £5000? yes]. We have just introduced the correct level of abstraction or LoA, represented by the symbol for British pounds, not, for example, by the symbol € for euros. Now Turing was the first to understand the crucial importance of expressing the LoA at which sensible questions may be asked. It might seem amazingly obvious, but the second example above, regarding the MCO, shows how easy and dangerous it is to forget about implicit LoAs. The importance of being clear about one's

¹'Mars Climate Orbiter Mishap Investigation Board Phase I Report' (Press release). NASA. See ftp://ftp.hq.nasa.gov/pub/pao/reports/1999/MCO_report.pdf

²For an introduction, see [2]; for a full philosophical analysis, see [3].

own level of abstraction was as obvious as the fact that the Earth is round, and that America was just there to be discovered. Yet it took Turing's genius to bring it to light. Of course, Turing's contribution was not that of introducing the concept of typed variables, or that of establishing the need for frames of reference. These ideas were already common at his time. His lesson was to make clear for the first time how philosophical and conceptual questions too could be answered only by fixing the LoA at which it would then make sense to receive an answer. This is one of the greatest and lasting contributions of his famous test [1], far more important than the wrong predictions about when machines would pass it or what consequences one should draw if they did pass it [4]. It is sometimes forgotten that Turing refused even to try to provide an answer to the question 'can a machine think?', because he considered it a problem 'too meaningless to deserve discussion'. Using our simple example, it would be like asking the price of the second-hand car in absolute figures, insisting that no currency is used in order to express it. Nonsense. Likewise, Turing objected that the question involved vague concepts such as 'machine' and 'thinking'. In other words, it lacked a clear level of abstraction. So he suggested replacing it with the Imitation Game, which is exactly more manageable and less demanding because it fixes a rule-based scenario easily implementable and controllable [5]. By so doing, he specified a LoA—the 'currency' he chose for the game was human intelligence, but it could have been something else, from animal intelligence to human creativity, as many other versions of the Turing imitation game have shown—and asked a new question, which may be summed up thus: 'may one conclude that a machine is thinking, at the Level of Abstraction represented by the imitation game?' After half a century, philosophy is still learning such a crucial lesson.³

3. Lesson 2: focusing on the most important problems or which philosophical questions to ask

On 23 April 2010, Bill Gates gave a talk at the Massachusetts Institute of Technology, Cambridge, MA, in which he asked: 'are the brightest minds working on the most important problems?' By 'the most important problems' he meant 'improving the lives of the poorest; improving education, health, nutrition'. Unfortunately, the list should probably include improving peaceful interactions, human rights, environmental conditions, living standards, etc., and this is only the beginning [8]. Clearly, the brightest *philosophical* minds should not be an exception, but turn their attention to such pressing challenges. Of course, one may stop philosophizing and start doing something about this messy world instead. We may, in other words, close down our philosophy departments and never corrupt our brightest youths philosophically. Yet, such solution smacks of self-defeat. It would be like deciding to burn the wicker basket in which we are travelling, because our hot air balloon is descending too quickly. Philosophy is what you want to keep in a good world, not what you want to get rid of in a bad one. Athens is a better place with Socrates. So there must be a different way forward. The fact is that philosophy can be extremely helpful, for it is

³On the use of the method of levels of abstraction in philosophy, see [6] and [3]. On Turing's crucial role in the development of the method, see [7]. The second lesson will require a much longer premise.

philosophy, understood as a conceptual design, that forges the new ideas, theories, perspectives and more generally the intellectual framework that can then be used to understand and deal with the ultimate questions that challenge us so pressingly. In the team effort made by the brightest minds, the philosophical ones can contribute insights and visions, analyses and syntheses, heuristics and solutions that can empower us to tackle 'the most important problems'. Every little effort helps us in the battle against idiocy, obscurantism, intolerance, fanaticisms and fundamentalisms of all kinds, bigotry, prejudice and mere ignorance. If this sounds self-serving recall that the longer the jump forward is, the longer the run-up to it should be. Or, with a different metaphor, philosophy takes care of the roots, so that the rest of the plant might grow more healthily. Suppose we accept all this as a reasonable assumption. Which ideas, theories, perspectives and, more generally, which intellectual framework should philosophers be designing now and for the foreseeable future, so that their contribution will be timely and helpful? Which philosophical questions should they be addressing? The answer would be inconceivable without Turing's legacy, for it lies in the conceptual threads that run across so many of our 'most important problems'. In a global information society, virtually any of the crucial challenges that we are facing is linked to information and communication technologies, in terms of causes, effects, solutions, scientific investigations, actual improvements, conceptual resources needed to understand them, or even just the wealth required to tackle them, as Bill Gates' example clearly shows. Obviously, information resources, technologies and sciences are not a panacea, but they are a crucial and powerful weapon in our fight against so many evils. The second lesson to be learnt from Turing therefore concerns the sort of questions that the brightest philosophical minds should be addressing. We live in an infosphere in which behind the most important problems often lies a Turing machine. It is a new world in which we have begun to re-conceptualize ourselves, a third lesson we have learnt from Turing, as I shall argue in §4.

4. Lesson 3: developing a new philosophical anthropology or from which perspective to approach philosophical questions

Oversimplifying, science has two fundamental ways of changing our understanding. One may be called *extrovert*, or about the world, and the other *introvert*, or about ourselves. Three scientific revolutions have had great impact both extrovertly and introvertly. In changing our understanding of the external world and how we can interact with it, they also modified our conception of who we are and may expect to become. After Copernicus, the heliocentric cosmology displaced the Earth and hence humanity from the centre of the universe. Darwin showed that all species of life have evolved over time from common ancestors through natural selection, thus displacing humanity from the centre of the biological kingdom. And following Freud, we acknowledge nowadays that the mind is also unconscious and subject to the defence mechanism of repression, thus displacing it from the centre of pure rationality, a position that had been assumed as uncontroversial at least since Descartes. The reader who, like Popper and myself, would be reluctant to follow Freud in considering psychoanalysis a strictly scientific enterprise like astronomy or evolutionary theory might yet be willing to concede that contemporary neuroscience is a probable candidate for such a revolutionary role. Either way,

the result is that today we acknowledge that we are not immobile, at the centre of the universe (Copernican revolution), we are not unnaturally separate and diverse from the rest of the animal kingdom (Darwinian revolution), and we are very far from being Cartesian minds entirely transparent to ourselves (Freudian or neuroscientific revolution).

One may easily question the value of this classic picture. After all, Freud [9] himself was the first to interpret these three revolutions as part of a single process of reassessment of human nature [10]. His hermeneutic manoeuvre was, admittedly, rather self-serving. But it does strike a reasonable note. In a similar way, when we now perceive that something very significant and profound has happened to human life after the computer revolution, I would argue that our intuition is once again perceptive, because we are experiencing what may be described as a fourth revolution, in the process of dislocation and reassessment of humanity's fundamental nature and role in the universe. This has been going on since the 1950s and Turing is undoubtedly the representative figure of such revolution. Computer science and the issuing technological applications have exercised both an extrovert and an introvert influence. They have not only provided unprecedented epistemic and engineering powers over natural and artificial realities; by doing so they have also cast new light on who we are, how we are related to the world and hence how we understand ourselves. Today, we are slowly accepting the idea that we are not standalone and unique entities, but rather informationally embodied organisms (*inforgs*), mutually connected and embedded in an informational environment, the *infosphere*, which we share with both natural and artificial agents similar to us in many respects. Turing has changed our philosophical anthropology as much as Copernicus, Darwin and Freud. This has had a significant impact on what it means to do philosophy after Turing, the last point to which I wish to call the reader's attention.

5. Lessons learnt: establishing a new philosophy of information or how to make sense of the world today

What can enable humanity to make sense of our contemporary world, respect it and improve it responsibly, and hence help in solving 'the most important problems'? The answer seems quite simple: a new philosophy of information. Among our mundane and technical concepts, information is currently one of the most important, widely used yet least understood. The brightest *philosophical* minds should turn their attention to it in order to design the philosophy of our time properly conceptualized *for* our time. This is a quick and dirty way of introducing the philosophy of information (PI) as a much needed development in this history of philosophy. Let me now sketch the longer story that links it to Turing.

Admittedly, it would be too much of a stretch to attribute to Turing the foundation or even the beginning of a new philosophy of information. After all, he never focused on the concept of information itself, or on problems about communication understood as *information flow* or *transmission*, despite the fact that he and Shannon knew each other's work. Thus, the index in Turing's *The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence and Artificial Life, Plus the Secrets of Enigma* [11] does not even

contain an entry for 'information' and a book like that by Luenberger [12] mentions Turing only once, in relation to Bletchley Park. And yet, I would argue that, without Turing, his groundbreaking work on *information processing*, the scientific and technological consequences of it, and the three lessons outlined above, contemporary interest in the philosophy of information would be very hard to explain. Turing shares with Shannon and Wiener the merit of having called our philosophical attention to the world of information and its dynamics. Without his three lessons, there would be no philosophy of information. The fact that nowadays we are more likely to treat computers as communication machines than powerful calculators only indicates how deep the influence of Turing's work has been on our world.

6. Conclusion

The development of new philosophical ideas seems to be akin to economic innovation. For, when Schumpeter adapted the idea of 'creative destruction' in order to interpret economic innovation, he might as well have been talking about intellectual development. Philosophy flourishes by constantly re-engineering itself. Nowadays, its pulling force of innovation is represented by the world of information, computation and communication phenomena, their corresponding sciences and technologies, and the new environments, social life, as well as the existential, cultural, economic and educational issues that they are bringing about. It is a new scenario that owes very much to Turing's work and intellectual legacy. In the previous pages, I have sketched three philosophical lessons that we should learn from Turing. I suggested that the philosophy of information, insofar as it brings to fruition Turing's legacy, can present itself as an innovative paradigm that opens up a very rich, helpful and timely area of conceptual investigations. Academically speaking, PI is the philosophical field concerned with the critical investigation of the conceptual nature and basic principles of information, including its dynamics, utilization and sciences, and with the elaboration and application of information-theoretic and computational methodologies to philosophical problems. More concretely, PI appropriates an explicit, clear and precise interpretation of the classic 'ti esti' question, namely 'what is information?', the clearest hallmark of a new field. As with any other field-question, this too only serves to demarcate an area of research, not to map its specific problems in detail. PI seeks to expand the frontier of our philosophical understanding, by providing innovative methodologies to address our most important problems from a contemporary perspective. It relies on Turing's intuition of the crucial importance of the method of abstraction to ensure that such problems are addressed at the right level.

The scientific revolution made seventeenth century philosophers redirect their attention from the nature of the knowable object to the epistemic relation between it and the knowing subject, and hence from metaphysics to epistemology. The subsequent growth of the information society and the appearance of the infosphere, as the environment in which millions of people spend their lives nowadays, have led contemporary philosophy to privilege critical reflection first on the domain represented by the memory and languages of organized knowledge, the instruments whereby the infosphere is managed—thus moving from epistemology to philosophy of language and logic—and then on the nature of its very fabric

and essence, information itself and its dynamics, including communication, flows and processing. Information has thus arisen as a concept as fundamental and important as Being, knowledge, life, intelligence, meaning or good and evil—all pivotal concepts with which it is interdependent—and so equally worthy of autonomous investigation. It is also a more impoverished concept, in terms of which the others can be expressed and interrelated, when not defined. This is why the philosophy of information may explain and guide the purposeful construction of our intellectual environment, and provide the systematic treatment of the conceptual foundations of contemporary society.

The future of PI depends on how well we shall engage with Turing's intellectual legacy, with 'the most important problems' of our time and with classic philosophical issues. I am optimistic. Thanks also to Turing, the Baconian–Galilean project of grasping and manipulating the alphabet of the universe has begun to find its fulfilment in the computational and informational revolution, which is affecting so profoundly our knowledge of reality and how we conceptualize it and ourselves within it. Informational narratives possess an ontic power, not as magical confabulations, expressions of theological logos or mystical formulae, but immanently, as building tools that can describe, modify and implement our environment and ourselves. From this perspective, the philosophy of information can be presented as the study of the informational activities that make possible the construction, conceptualization, semanticization (giving meaning to) and finally the moral stewardship of reality, both natural and artificial, both physical and anthropological. The philosophy of information enables humanity to make sense of the world and construct it responsibly. It promises to be one of the most exciting and beneficial areas of philosophical research of our time. Its development will be an appropriate way to continue Turing's work and honour his legacy in philosophy.

References

- 1 Turing, A. M. 1950 Computing machinery and intelligence. *Mind* **59**, 433–460. (doi:10.1093/mind/LIX.236.433)
- 2 Floridi, L. 2010 *Information—a very short introduction*. Oxford, UK: Oxford University Press.
- 3 Floridi, L. 2011 *The philosophy of information*. Oxford, UK: Oxford University Press.
- 4 Floridi, L., Taddeo, M. & Turilli, M. 2009 Turing's imitation game: still a challenge for any machine and some judges. *Minds Mach.* **19**, 145–150. (doi:10.1007/s11023-008-9130-6)
- 5 Moor, J. 2003 *The Turing test: the elusive standard of artificial intelligence*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- 6 Floridi, L. 2008 The method of levels of abstraction. *Minds Mach* **18**, 303–329. (doi:10.1007/s11023-008-9113-7)
- 7 Floridi, L. In press. Turing test and the method of levels of abstraction. In *Alan Turing—his work and impact* (eds S. B. Cooper & J. van Leeuwen). Amsterdam, The Netherlands: Elsevier.
- 8 Floridi, L. 2010 Information ethics. In *The Cambridge handbook of information and computer ethics* (ed. L. Floridi), ch. 5. Cambridge, UK: Cambridge University Press.
- 9 Freud, S. 1917 A difficulty in the path of psycho-analysis. In *The complete psychological works of Sigmund Freud, the Standard Edition*, vol. XVII (ed. J. Strachey), pp. 135–144. London, UK: Hogarth Press.
- 10 Weinert, F. 2009 *Copernicus, Darwin, and Freud: revolutions in the history and philosophy of science*. Oxford, UK: Blackwell.
- 11 Turing, A. M. 2004 *The essential Turing: seminal writings in computing, logic, philosophy, artificial intelligence and artificial life, plus the secrets of enigma*. Oxford, UK: Clarendon Press.
- 12 Luenberger, D. G. 2006 *Information science*. Princeton, NJ: Princeton University Press.