Review: Noson S. Yanofsky (2013): *The Outer Limits of Reason.* What Science, Mathematics, and Logic Cannot Tell Us. Cambridge (Mass), London: The MIT Press. 403pp + xiv. ISBN 978-0-262-01935-4*

Tim Räz†

July 1, 2014

Yanofsky's book explores the limitations of empirical science, mathematics, logic, and philosophy. Yanofsky's goal is to discuss paradigmatic examples in some depth, not to provide the reader with a mere compendium of curiosities. A wealth of cases is grouped into eight substantive chapters, bound together by an introductory chapter with an outline of coming attractions, and a synthesis in the final chapter. The chapters are intended to be largely self-contained, and each chapter closes with suggestions for further reading. The book is written for a wide audience; Yanofsky attaches great importance to a clear and accessible presentation, and most formulas are omitted.

In the introductory chapter 1, Yanofsky characterizes his enterprise in general terms and discusses the different themes of the book. A special emphasis is on the notion of paradox: an argument based on (seemingly) reasonable premisses and leading to a contradiction or falsehood. Further themes are reductions – a technique to establish connections between impossibility results –, the language dependence of contradictions, and the notion of reason.

While Yanofsky's exposition is clear, the thematic arrangement did not strike me as particularly instructive; I would have preferred a more traditional synopsis. However, there is usually an overview at the beginning of each chapter.

Chapter 2 is an exposition of "Language Paradoxes". Yanofsky first discusses the liar paradox, and various solutions. Yanofsky's preferred solution is to accept

*This is the pre-peer reviewed version of the following article: Räz, T. (2015): Noson S. Yanofsky, The Outer Limits of Reason. What Science, Mathematics, and Logic Cannot Tell Us, Cambridge (MA)/London: The MIT Press, 2013, xiv + 403 pp., US$ 29.95, £20.95 (Hardback), ISBN 9780262019354, Dialectica 69 (2): 248-254, which has been published in final form *here*. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

†E-mail: tim.raez@gmail.com. Address: FB Philosophie, Universität Konstanz, 78457 Konstanz, Germany. Acknowledgement: I thank Tilman Sauer for comments on an earlier draft of this review.

that the liar sentence is contradictory, as it is a product of the human mind. He then discusses several self-referential paradoxes, such as the barber paradox, Grelling's paradox and Russell's paradox. Yablo's paradox presents an exception to the view that all paradoxes of language are self-referential.

The presentation of the individual paradoxes in this chapter is successful, and Yanofsky's goal of providing the reader with a gentle introduction is laudable, but the original context of some of the paradoxes is lost. For example, Russell's paradox has its origin in set theory, and I found it to be out of place in the context of paradoxes of natural language.

Chapter 3 present a variety of "Philosophical Conundrums", limitations of our knowledge discovered by philosophers rather than scientists. Yanofsky discusses the ship of Theseus, problems of (personal) identity, the acquisition of (abstract) concepts, realism and nominalism concerning abstract and concrete objects, Zeno's paradoxes of motion, time travel, paradoxes of vagueness, the Monty Hall problem, and others.

This chapter will not satisfy the more philosophically-minded reader. Yanofsky defends a position he dubs "extreme nominalism", an antirealism concerning abstract and concrete objects, of both everyday experience and physics, and finds that "[w]hat do exist are physical stimuli" (p. 40). However, if there are stimuli – what are they so reliably caused by? The best explanation, it seems to me, is that they are caused by objects.

Chapter 4, about "Infinity Puzzles", is a whirlwind tour of set theory. Yanofsky begins with finite sets and cardinalities, and introduces counterintuitive properties of infinite sets via Hilbert's hotel. He shows that some infinite sets, such as the rational numbers, are as large as the natural numbers, while others, such as the real numbers, are larger than the natural numbers. Russell's paradox prompts an axiomatization of set theory via the Zermelo-Fraenkel axioms. Finally, topics such as the independence of the Zermelo-Fraenkel axioms from the continuum hypothesis and from the axiom of choice are explored.

The learning curve in this chapter is quite steep for an uninitiated reader; however, this is due to the subject matter. Yanofsky succeeds in presenting the beginnings of set theory in an accessible manner. I have only minor quibbles. First, the Zermelo-Fraenkel axioms are stated without further elaboration; some remarks as to their significance would have been nice. Second, the importance, and impact, of Russell's paradox should be stressed more.

Chapter 5 deals with computational complexity, limitations on the usefulness of algorithms due to efficiency. We are first introduced to some examples of easy problems that can be solved in a reasonable, polynomially bounded amount of time, such as addition and multiplication, and Euler's solution to the Königsberg bridges problem. We then turn to hard problems, such as the Traveling Salesman, Hamiltonian cycles, and satisfiability of propositional formulas. The amount of time it takes to solve these problems can grow exponentially in the size of input – these are NP problems. Yanofsky explains how NP problems are related via reduction, mentions the Cook-Levin theorem and the famous P = NP problem. Finally, he briefly discusses approximation algorithms and problems that are even harder than NP.

This chapter is a good exposition of the basic ideas of computational complexity. I liked the idea of covering computational complexity before the more principled limitations of computability. Even though the limitations are practical, they are very real and deserve the philosopher's attention; the interested reader could turn to an excellent essay by Scott Aaronson (2013) on why philosophers should care about computational complexity. At times, I would have liked to be given more explanation as to why some problems that superficially look alike, such as Euler cycles and Hamiltonian cycles, can nevertheless fall in different complexity classes.

Chapter 6 covers the limitations of computability. Yanofsky first introduces basic notions such as algorithms and programs. The most important algorithmically undecidable problem is the Halting Problem: Does an algorithm terminate on a certain input or not? Yanofsky proves its undecidability. We learn about further undecidable problems. Rice's theorem shows that no interesting property of an algorithm is decidable. Some problems are not solvable even if we are given an oracle that solves the Halting Problem. Finally, Yanofsky briefly discusses the ramifications of (un-)computability for the human mind, mentioning the works of Kurt Gödel, Roger Penrose, and Douglas R. Hofstadter.

The exposition in this chapter is, again, accessible; Yanofsky conveys a feel for computability and its limitations. At times, he covers the ground too fast. For example, the correspondence between programs and numbers, or Rice's theorem, could have been explained in more depth. Finally, the discussion of computers and minds is too brief to be satisfactory. Many philosophers are skeptical as to the ramifications of these results for the human mind, and some critical voices should have been mentioned.

Chapter 7 discusses limitations of science in three sections. In the first section, on limitations of predictability, Yanofsky distinguishes predictability and determinism, and emphasizes the importance of the former notion in science. There are three kinds of limitations. First, in chaotic systems, such as the Lorentz system, the temporal evolution depends sensibly on initial conditions. Second, some systems do not have a simple description in terms of elementary mathematical operations; the three-body problem is an example. Finally, for some many-particle systems, a statistical description is more feasible than precise bookkeeping.

The second section introduces counterintuitive features of Quantum Mechanics (QM), which are traced back to the so-called "Wholeness Postulate": "[t]he outcome of an experiment depends on the *whole* setup of the experiment" (p. 176). We are initiated in results such as the Kochen-Specker theorem, nonlocality as demonstrated by Bell's theorem, and quantum eraser experiments. Yanofsky also discusses four interpretations of QM: The Copenhagen interpretation, the Multiverse (many-worlds) interpretation, the Hidden Variables (de Broglie-Bohm) interpretation, and Quantum Logic.

The third section discusses aspects of relativity theory. According to Yanofsky, "[t]he central idea of relativity theory is that properties of the physical universe depend on how they are measured." (p. 214) Special Relativity is introduced via two postulates: First, "observers at constant speed must observe the same laws of motion" (p. 216); second, "observers will always view the speed of light at the same rate" (p. 219). Consequences of these postulates are length contraction, time dilation, relativity of simultaneity, and the equivalence of mass and energy. General Relativity is introduced via the elevator thought experiment; features such as the curvature of space-time and the bending of light are mentioned. The chapter closes with an outlook on the problem of reconciling General Relativity and QM.

I found the sections on QM and relativity disappointing. There are, first, some inaccuracies. Here are some examples – the expert will find more. Quantum superpositions are described as follows: "Usually an object has a *position* [...] [t]he phenomenon of being in more than one place at one time is called *superposition*" (p. 178, emphasis in original). However, superposition is not tied to the position of particles – rather, superposition means that if a system can be in two states, then a third state, which is a combination – superposition – of the first two, is also a possible state. The characterizations of the measurement problem, and of QM as indeterministic (for the most part), are also problematic. Turning to relativity, the "central idea" of relativity, quoted above, can be seen to be false by confronting it with the second postulate. Finally, claims such as "[a]ll is relative" (p. 221) are not warranted by relativity theory. Yanofsky should have abstained from this kind of rhetoric, as he distances himself from a "New Age" orientation in the introduction of the book. The discussion is also disappointing from a philosophical point of view: Most of the work on the foundations of physics of the last twenty years or so has been disregarded, and taking it into account might have led to a more nuanced discussion of consciousness, free will, or the possibility of an observer-independent reality.

Chapter 8, on "Metascientific Perplexities", deals with issues from philosophy of science, the applicability of mathematics, and Anthropic Principles in three sections. The first section is about some classical problems of philosophy of science, such as the problem of induction, Hempel's paradox, the role of theoretical considerations in theory choice, Karl Popper's idea of falsifiability, Thomas Kuhn's theory of theory change, and the question whether we have reached the end of science.

In the second section, Yanofsky discusses "Wigner's puzzle", the "unreasonable effectiveness of mathematics": how can mathematics be so useful in science, even though it has a high degree of autonomy from science? Yanofsky goes through several purported examples from the history of science, and discusses solutions. He observes that, firstly, mathematics is not applied in many branches of science outside of physics, and many branches of mathematics never get applied; secondly, many parts of mathematics are directly rooted in science. He thinks that the combination of these two observations dissolves the problem.

Yanofsky turns to an even deeper set of puzzles in section three. First, why does the universe exhibit structure and regularities at all? Second, why is our universe such that life can exist? Third, why are there living creatures in the universe that are capable of understanding the universe? One answer to these questions is the so-called "Anthropic Principle". It its weak form, the Anthropic Principle says that the universe has to have properties that make it possible for us to observe it. Finally, Yanofsky discusses possible explanations of the weak Anthropic Principle.

The short section on philosophy of science covers only a tiny fraction of the field, and what is covered is not entirely even-handed, for example when Yanofsky writes that Kuhn "felt that objective truth does not really exist" (p. 249). On the other hand, I liked the discussion of Wigner's puzzle. The choice of going through many examples of the unreasonable effectiveness is refreshing, and the proposed dissolution of the puzzle quite convincing.

Chapter 9 deals with "Mathematical Obstructions". First, we learn about classes of numbers and their limitations: For example, the irrational numbers cannot be written as a ratio of two whole numbers. Galois theory is discussed as a tool to determine which polynomial equations can be solved. Then a connection between mathematical problems, such as the Tiling Problem, and limitations of computability – the Halting Problem – is established. The former are reducible to the latter, and therefore unsolvable. Turning to logic, we are introduced to Peano Arithmetic, symbolization, the arithmetization of syntax, and the "fixed-point machine". With these ingredients, Yanofsky walks us through logical paradoxes, such as Tarski's Theorem and Gödel's First Incompleteness Theorem. Goodstein's Theorem serves as an example of a "natural" theorem that can only be proved using infinitary methods. Finally, Yanofsky discusses Gödel's Second Incompleteness Theorem.

The exposition in this chapter is uneven. The introduction of the problems is accessible, but Yanofsky should have tried to give us more insight into how some of the results are established. The geometrical proof of the irrationality of $\sqrt{2}$ is nice, but I found it harder to grasp than the usual algebraic proof. Then, the discussion of the logical paradoxes, and Gödel's First Incompleteness Theorem in particular, proceeds too fast. For example, more explanation of why and how we can write down a proof predicate would have been desirable.

The final chapter 10 offers a synthesis. Yanofsky presents a different typology of the various limitations: physical, mental-construct (encompassing linguistic and formal paradoxes), practical (such as from complexity), and of our intuition, as in QM and relativity theory. Self-referential paradoxes are systematized in a table. Yanofsky returns to the notion of reason, which he defines as "the set of processes or methodologies that do not lead to contradictions or falsehoods". The book ends with personal observations.

All in all, Yanofsky's project succeeds to a large degree. He has a gift for presenting the technical issues in an accessible and light-hearted way; the book should be accessible to readers who are not deterred by elementary mathematics. The pleasure Yanofsky derives from mathematics, logic, and science is apparent throughout the book and will capture the inclined reader. Some chapters and sections, in particular on computer science and mathematics, are a pleasure to read. The subject matter demands an interdisciplinary approach, and Yanofsky's efforts to bring together results from different fields is fruitful.

The book also has its limitations. First, the selection of sciences discussed is biased: Non-formal and non-fundamental sciences, such as biology, economics, psychology, etc. are left out. Even if the limitations of these sciences were fundamentally different from those of physics, it would have been interesting to hear more on, say, limitations of social sciences due to the use of statistical methods, or the limitations of neoclassical and behavioral approaches in economics.

Second, Yanofsky has a tendency to mention important qualifications of general claims only briefly, or to relegate them to endnotes, while the main text proceeds as if these qualifications could be ignored. For example, he informs us that QM has a deterministic interpretation – Bohmian mechanics, – yet insists, for the most part, on the indeterminism of QM. This omission will not satisfy philosophers. It is legitimate for a popular account of science to ban technical or minor issues from the main text, but here, it affected the quality of the book.

Third, Yanofsky's classification of limitations, outlined in the first and last chapter, is unsatisfactory. This is probably due to the fact that these field diverge as to defy a unified treatment. Consequently, the book works well as a collection of paradoxes and counterintuitive results, but it does not provide a convincing, unified perspective on the limitations of science, mathematics, logic, and philosophy.

In sum, some of the chapters, in particular on set theory, complexity, and computability, may serve as useful introductions to these issues. As a whole, the book is unsuitable for use in philosophy classes, because the philosophical discussion lacks depth, and most of the philosophical progress of the last decades is not taken into account. However, the book was not written for philosophers, and I can recommend it as a popular account of the limitations of our knowledge.

References

Aaronson, S. 2013. Why Philosophers Should Care About Computational Complexity. In B. J. Copeland, C. Posy, and O. Shagrir, eds., *Computability: Turing, Gödel, Church, and Beyond*. MIT Press, pp. 261–328.