Review of John Stillwell, *Reverse Mathematics: Proofs from the Inside Out*

Benedict Eastaugh*
benedict@eastaugh.net

22 November, 2019

John Stillwell, **Reverse Mathematics: Proofs from the Inside Out**. Princeton, NJ: Princeton University Press, 2018, pp. 200. ISBN 978-0-69-117717-5 (hbk), 978-0-69-119641-1 (pbk), 978-1-40-088903-7 (e-book). $24.95.

Reverse mathematics is a programme in mathematical logic, initiated in the mid-1970s, which seeks to determine which axioms are necessary to prove theorems in areas of ordinary mathematics such as real analysis, countable abstract algebra, countably infinite combinatorics, and the topology of complete separable metric spaces. *Reverse Mathematics: Proofs From the Inside Out* is the first popular book on the subject, aimed at advanced undergraduates in mathematics, but also a good introduction for philosophers of mathematics. The time is certainly ripe for such a book, bringing this fascinating area of contemporary mathematical logic to a broader audience.

Stillwell motivates the study of reverse mathematics through the following extended analogy with geometry, developed in the first chapter. In Euclid's *Elements*, the fifth (parallel) postulate is used to prove the Pythagorean theorem. But is the use of the parallel postulate *necessary*? In fact, it is, but demonstrating this necessity in a rigorous manner requires two things that are at the heart of reverse mathematics. The first requirement is a *base theory* that does not prove the Pythagorean theorem, but that is compatible with its truth. The second is a *reversal*: an implication, provable in the base theory, from the theorem (the Pythagorean theorem, in this case) to the axiom (the parallel postulate). The base theory in the case of Euclidean geometry is provided by the first four of Euclid's postulates. The existence of non-Euclidean geometries that make the first four postulates true but the parallel postulate false demonstrates that the base theory does not prove the parallel postulate. This makes the reversal from theorem to axiom non-trivial, and shows that the axiom is indeed necessary in order to derive the theorem: whenever the theorem is true, the axiom must also be true.

Although the details are very different, in broad strokes reverse mathematics follows the same methodology in order to determine which axioms are necessary to prove theorems of ordinary mathematics. Instead of a language of points and lines, the base theory used in reverse mathematics, along with the axioms and theorems studied, is formulated in the language of *second-order arithmetic*. This is a two-sorted first-order system that extends the language of first-order Peano arithmetic by permitting quantification over both natural numbers, using variables of the first sort (called "number variables"), and over sets of natural numbers, using variables of the second sort (called "set variables").[1] One of Stillwell's main examples is the monotone convergence theorem, which states that every increasing, bounded sequence of real numbers converges to a limit. By formalising the notions of real number, sequence, limit, and convergence within second-order arithmetic, one can represent the monotone convergence theorem by a sentence MCT in the formal

*Munich Center for Mathematical Philosophy, LMU Munich. ORCID: 0000-0002-6629-3032.

[1] Since it is aimed at a fairly general audience, the book does not address the following important point about the semantics employed, namely that in reverse mathematics one always uses first-order semantics (Henkin semantics) to interpret the set variables, rather than interpreting them as ranging over the full powerset of the range of the number variables, as one does in the standard semantics for second-order logic [Shapiro 1991]. This means that, for example, if we take $\mathbb{N}$ to denote the standard model of arithmetic and REC to denote the set of computable sets of natural numbers, then $(\mathbb{N}, \mathrm{REC})$ is a model of the subsystem of second-order arithmetic RCA_0 described below. Such a model would not be admissible under the standard semantics for second-order logic, since not every set of natural numbers is computable, and so REC is not the full powerset of the natural numbers.

language. One may then formalise the standard proof of the monotone convergence theorem within a subsystem of second-order arithmetic known as ACA_0, where 'ACA' stands for the Arithmetical Comprehension Axiom, an axiom scheme stating that every set definable by a formula in the language of first-order arithmetic exists.

The other axioms of ACA_0 are the axioms of Robinson arithmetic Q, which give the usual properties of the successor, addition, and multiplication operations, plus the second-order induction axiom $\forall X((0 \in X \land \forall n(n \in X \to n+1 \in X)) \to \forall n(n \in X))$. Having shown that arithmetical comprehension is, in the presence of the other axioms, sufficient to prove the formalisation MCT of the monotone convergence theorem, one then seeks to show that it is also necessary. To do this one works in a weak base theory known as RCA_0. 'RCA' stands for Recursive Comprehension Axiom: a weakening of arithmetical comprehension that asserts that every *computable* (i.e. recursive) set exists. The other axioms of RCA_0 are those of Robinson arithmetic Q, plus the induction scheme for Σ^0_1 formulas. By adding the formalisation MCT of the monotone convergence theorem to the axioms of RCA_0, one can then derive any instance of the arithmetical comprehension scheme. This reversal from theorem (MCT) to axiom (ACA) shows that the axiom is indeed necessary in order to prove the theorem. Crucially, RCA_0 cannot prove all instances of arithmetical comprehension, and thus cannot prove MCT either. This illustrates the analogy with geometry, and also highlights the importance of the base theory, which must be weak enough so as to not prove outright the theorems being reversed, but strong enough to prove both directions of their equivalences.

Many hundreds of mathematical theorems have been studied in reverse mathematics. The vast majority have been found, somewhat surprisingly, to fall into just five equivalence classes: either they are provable in the base theory RCA_0, or they are equivalent to one of just four other axioms, listed in order of increasing strength (with the system name in brackets): weak König's lemma (WKL_0), the restriction of König's infinity lemma to trees formed from binary sequences; arithmetical comprehension (ACA_0); arithmetical transfinite recursion (ATR_0), which allows arithmetically definable operations to be iterated along any well-ordering; and Π^1_1 comprehension (Π^1_1-CA_0). The book focuses primarily on equivalences to arithmetical comprehension, with some discussion of equivalences to weak König's lemma in the penultimate chapter.

Stillwell's presentation of the methodology and goals of reverse mathematics follows closely the template laid down by Harvey Friedman and Stephen Simpson [Friedman 1975, Simpson 2009], as a programme of classifying theorems of analysis, algebra, logic, combinatorics, etc., in terms of set existence axioms of increasing strength. One *locus classicus* of this view is a passage in Simpson [2009, p. 2] which formulates the main question of reverse mathematics as "Which set existence axioms are needed to prove the theorems of ordinary, non-set-theoretic mathematics?". In articulating the received view of reverse mathematics, Stillwell writes (p. 24) that "The [set existence] axioms in question state there is a set of natural numbers n corresponding to each property $\varphi(n)$ in a certain class". The natural reading of this passage is that all set existence axioms are *comprehension principles*: axiom schemes asserting the existence of all sets of a certain definable class. Of the set existence principles that play a major role in reverse mathematics, recursive comprehension, arithmetical comprehension, and Π^1_1 comprehension are all of this form. But as noted by Dean and Walsh [2017], such a characterisation fails to include two of the five main subsystems of second-order arithmetic, WKL_0 and ATR_0.[2] As I argue in [Eastaugh 2019], the received view of reverse mathematics sketched above can only be maintained if one provides a more nuanced account of set existence principles that includes weak König's lemma and arithmetical transfinite recursion as well as comprehension principles, for example by taking set existence principles to be closure conditions on the powerset of the natural numbers.

After sketching the development of geometry and how it provides the basic methodology of base theory and reversals, Stillwell turns to the development of analysis, and in particular to the so-called "arithmetization of analysis" that took place during the nineteenth century. During this process, intuitive geometrical conceptions of the continuum were replaced by set-theoretic ones. The formalisation sketched in chapter 2 uses Dedekind's influential representation of real numbers as 'cuts' partitioning the rational numbers into upper and lower parts, with the real number thus represented corresponding to the point in the continuum where the rationals are divided by

[2] This point is much stronger than just the observation that weak König's lemma does not have the surface grammar of a comprehension principle. Rather, weak König's lemma is not equivalent over RCA_0 to any subscheme of the arithmetical comprehension scheme. For details see theorem 2.1 of [Eastaugh 2019, p. 160].

the cut. Rational numbers themselves are represented as pairs of integers (n, d) whose intended interpretation is $\frac{n}{d}$, while integers are represented as pairs of natural numbers (m, k) whose intended interpretation is $m-k$. These pairs can themselves be coded by single natural numbers, for example by letting the pair of natural numbers (a, b) be coded by the number $(a+b)^2 + a$. In this way the concepts of integer, rational, and real number are *arithmetized*: reduced to the two basic concepts available in the formal language of second-order arithmetic, namely natural numbers and sets of natural numbers.

The rest of chapter 2 covers the arithmetization of continuous functions, the language and axioms of Peano arithmetic, and an introduction to arithmetically definable sets of natural numbers. In many ways this makes this chapter the most difficult of the book, since it requires the reader to absorb several quite different developments in analysis, arithmetic, and computability theory, and begin to understand the connections between them. The decision to focus on theorems of analysis has obvious pedagogical advantages, since most readers in the target audience will have sufficient familiarity with the material to follow the general idea. Nevertheless, it does have its downsides, most obviously that the coding needed to arithmetize even the most basic concepts of analysis is complex and often opaque. The weakness of the formal systems involved also means that the choices involved in coding analytical objects must be made with great care. Stillwell's choice of Dedekind cuts, in particular, is not a suitable way of coding reals in RCA_0, because even simple manipulations of sequences of real numbers represented in this way often require stronger set existence principles than are available in the base theory. For this reason, working reverse mathematicians prefer to use fast-converging Cauchy sequences of rational numbers, a more computationally tractable coding scheme which avoids the pathologies of Dedekind cuts [Hirst 2007].

Chapter 3 surveys some classical topics in analysis, including limits, continuity and the intermediate value theorem, the Bolzano–Weierstraß theorem, and the Heine–Borel theorem. This is done with a focus on so-called *sequential* forms of these theorems which, because they quantify only over hereditarily countable objects, are amenable to being formalised in subsystems of second-order arithmetic. For example, one version of the Heine–Borel theorem states that if S is an infinite set of open intervals that covers $[0, 1]$, then there is some finite subset $F \subseteq S$ such that F also covers $[0, 1]$. An immediate corollary is the following sequential form: if $I = \langle I_0, I_1, \dots \rangle$ is a countable sequence of open covers, then there is a finite subsequence $F = \langle I_0, I_1, \dots, I_k \rangle$ which covers $[0, 1]$. In the first version of the theorem, S can be any set of open intervals, even an uncountable one. In the sequential form of the theorem, however, both the countable sequence of open covers I and the finite subsequence F are coded by a single real number or set of natural numbers. The sequential form can therefore, unlike the more general version, be formalised in second-order arithmetic by a statement HB that quantifies only over natural numbers and sets of natural numbers.

Chapter 4 reviews some basic results in computability theory, and then applies them to analysis. After sketching Hilbert's tenth problem and the need for a precise characterisation of the notion of algorithm or effective procedure, Stillwell gives some properties of computable functions, followed by an outline of an argument that the halting problem is undecidable, and a description of the computably enumerable sets. The existence of sets that are computably enumerable but not computable is then used to generate what are typically called *recursive counterexamples* in the literature on constructive and computable analysis. A stock example is that of *Specker sequences*: bounded increasing sequences of rational numbers whose least upper bound is a non-computable real number. The existence of Specker sequences shows that the monotone convergence theorem is not provable in RCA_0, because the axioms of RCA_0 do not prove the existence of non-computable sets.[3]

Chapter 5 is concerned primarily with how computability-theoretic notions can be arithmetized. Stillwell does this via Smullyan's *elementary formal systems* [Smullyan 1961], which are syntactic systems similar to those of Post. Elementary formal systems are an elegant way of exposing the computational aspects of formal systems, but they are rarely taught in contemporary logic courses, making this feel like an idiosyncratic choice. While it may help convey the ideas in a relatively compact way, the lack of good contemporary resources will make it hard for readers not

[3] This point relies implicitly on the use of the model $(\mathbb{N}, \mathrm{REC})$ described in footnote 1: the axioms of RCA_0 are true when we interpret their set variables as only ranging over computable sets, i.e. $(\mathbb{N}, \mathrm{REC}) \models \mathsf{RCA}_0$. Any Specker sequence S is computable, and hence a member of REC, but since the limit of S is non-computable, it is not in REC and thus the monotone convergence theorem is false in the model $(\mathbb{N}, \mathrm{REC})$.

already versed in computability theory to reconstruct the arguments in a rigorous way. After some background on Turing and Post's analyses of computation, and how they relate to one another and to elementary formal systems, the elementary formal systems are put to use in representing numbers and arithmetical relations. The converse procedure of representing elementary formal systems in arithmetic is done in a somewhat impressionistic manner, but it does convey the central ideas about coding sequences with Gödel's β function.

One way in which we can understand the hierarchy of systems studied in reverse mathematics is as examples of incompleteness that show the various ways in which non-computable sets are essential to mathematical practice. For example, if one starts out with an axiom system like RCA_0 that only proves the existence of computable real numbers, then one must add additional axioms in order to prove the least upper bound principle and all the theorems that rely on it. Reverse mathematics gives us a precise determination of which axiom must be added to RCA_0 to achieve this outcome, namely the arithmetical comprehension scheme (ACA). This is the subject of chapter 6, which brings us to reverse mathematics proper via a demonstration that different ways of expressing the sequential completeness of the real line (such as the sequential least upper bound principle, the monotone convergence theorem, and the sequential Bolzano–Weierstraß theorem) are all equivalent to one another and to ACA.

Chapter 6 also discusses another important set of concepts in computability theory and reverse mathematics, *finitely branching trees*. A tree is a set of finite sequences that is closed under subsequences. An important special case is where the elements of the tree are *binary sequences*, finite sequences of 0s and 1s. König's infinity lemma (KL) that every finitely branching tree with infinitely many nodes contains an infinite path has a computability-theoretic interpretation: there are computable, infinite, finitely branching trees with no computable paths, so König's lemma is computably false. However, all such trees contain paths which are computable relative to the set K of solutions to the halting problem. Viewed through the lens of reverse mathematics, this means that König's lemma implies arithmetical comprehension, and indeed is equivalent to it.

Restricting to trees of binary sequences produces a very different result. The principle that every infinite tree of binary sequences has an infinite path is known as weak König's lemma (WKL). This principle is, as discussed in chapter 7, equivalent to a wide variety of theorems in analysis including the sequential Heine–Borel covering theorem and the extreme value theorem ("Every continuous function on the closed unit interval attains a maximum"), and like the full König's lemma it is computably false. However, computable infinite trees of binary sequences always contain infinite paths that are less complex than the halting set K, thus showing that WKL is a strictly weaker principle than the full König's lemma.

The remainder of chapter 7 is devoted to a discussion of the base theory RCA_0, including giving a more precise account of how e.g. continuous functions can be represented in the base theory, and briefly describing the other so-called "Big Five" systems ATR_0 and $\Pi^1_1\text{-}\mathsf{CA}_0$, and some of their equivalences. However, as noted earlier, the main focus of the book is on equivalences with ACA. Although this is an understandable choice given the book's focus on theorems of analysis, and the technical complexities sometimes involved in equivalences to the stronger axioms of arithmetical transfinite recursion and Π^1_1 comprehension, it does mean that the book provides a rather partial impression of the field.

A similar issue arises in section 6.6, which concerns Ramsey's theorem. Stillwell does an excellent job of introducing the reader to the fundamentals of Ramsey theory, explaining in some detail how different weakenings of Ramsey's theorem are related to one another. A particularly important such weakening is the principle known as Ramsey's theorem for pairs (RT^2_2). As Stillwell explains, RT^2_2 is important in reverse mathematics because it is provable in ACA_0 but does not reverse to it. However, having reached a crucial point in the exposition, the book stumbles: it does not make clear that RT^2_2 is not provable in the base theory RCA_0—that Ramsey's theorem for pairs is not computably true—nor does it explain that RT^2_2 is actually incomparable with the intermediate system WKL_0, and thus lies outside the realm of the "Big Five" subsystems. This phenomenon is quite general, and the constellation of non-equivalent combinatorial principles that lie in the "Reverse Mathematics Zoo" between RCA_0 and ACA_0 provides much of the impetus behind contemporary research in reverse mathematics (see e.g. Hirschfeldt 2014). Since by this point in the book the reader would be in a good position to appreciate the importance of these results, Stillwell misses a golden opportunity to paint a richer and more complex picture of the field.

The short final chapter tries to place the results from the rest of the book into a larger context, relating reverse mathematics to constructive mathematics, the incompleteness phenomenon, and computability theory. While they are understandably quite compressed, these closing remarks are nevertheless to be commended for their references to both historical sources and more recent surveys, and they will hopefully encourage interested readers to delve more deeply. One unfortunate omission, given the focus on constructive mathematics in this chapter, is any mention of *constructive reverse mathematics*: the programme of proving equivalences between mathematical theorems and various non-constructive principles from within a constructively acceptable base theory [Ishihara 2006, Bridges and Palmgren 2018, §5]. This would have helped substantiate Stillwell's claim (p. 156) that Brouwer's fixed-point theorem "is not far outside constructive mathematics, since it is constructively equivalent to the weak König lemma". Reasoning within RCA_0 is not so far removed from constructive mathematical reasoning, but contrary to what this passage seems to imply, proofs of equivalences in RCA_0 are not in general constructively acceptable, since RCA_0 is a classical system that admits the law of excluded middle (for example, RCA_0 proves the intermediate value theorem, whose proof uses the law of excluded middle in an essential way, and which is thus not constructively valid). Nevertheless, the equivalence between WKL and Brouwer's fixed-point theorem can in fact be proven in a genuinely constructive way within constructive reverse mathematics [Hendtlass 2012, pp. 9–10].

Overall, *Reverse Mathematics: Proofs from the Inside Out* is an engaging and rewarding book which should excite the interests of anyone who has a familiarity with undergraduate-level real analysis and an interest in the foundations of mathematics. Readers with some background in mathematical logic (preferably including some familiarity with the basics of computability theory) who are interested in pursuing the subject further will find Denis Hirschfeldt's book *Slicing the Truth* [Hirschfeldt 2014] an excellent follow-up to this one, although Hirschfeldt focuses on combinatorics, rather than analysis. Working through the relevant sections (primarily in chapters I–IV) of Stephen Simpson's canonical reference work *Subsystems of Second Order Arithmetic* [Simpson 2009] would allow the motivated reader to grasp the finer details of the proofs sketched by Stillwell.

Such a reader should, however, be aware of a few imprecisions in the present volume. These include a typographical error (p. 40) in which one of the axioms of Peano arithmetic is given as the statement that for all m and n, if $S(m) \neq S(n)$ then $m \neq n$. Rather than an axiom of arithmetic, this is a simple validity of classical first-order logic with identity. The correct form of the axiom is the statement that for all m and n, if $S(m) = S(n)$ then $m = n$, expressing that the successor function S is injective. The relationship between subsystems of second-order arithmetic like RCA_0 and ACA_0 and their unsubscripted counterparts RCA and ACA is also misdescribed (p. 107, fn. 5, and p. 110, fn. 1). The '0' subscript denotes a system with restricted induction, thus tying the strength of the induction principle to the amount of set comprehension available in the system. The unsubscripted subsystems such as RCA and ACA, on the other hand, include the induction scheme for all formulas in the language of second-order arithmetic, i.e. with unrestricted use of both set and number quantifiers. Finally, a parenthetical remark (p. 146) states that weak König's lemma "implies recursive comprehension, so it is *the* set existence axiom for WKL_0". Taken literally, this is incorrect: weak König's lemma is a *conditional* set existence principle that relies on the presence of recursive comprehension, stated as an additional axiom, to guarantee that there are any binary trees to begin with [Eastaugh 2019, pp. 167–8]. However, since the base theory RCA_0 includes recursive comprehension, we can always assume that it is available, and in this sense weak König's lemma is indeed "the" set existence principle of WKL_0, since it is the sole axiom that distinguishes WKL_0 from the base theory RCA_0. Helpfully, a list of errata is available on the publisher's website at https://press.princeton.edu/titles/11143.html, although at the time of writing only the first of the issues mentioned above was included.

Funding

Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) – Projektnummer 390218268 / Funded by the German Research Foundation (DFG) – project number 390218268.

Acknowledgements

My thanks to Marianna Antonutti Marfori, Walter Dean, Damir Dzhafarov, Alice Eastaugh, and Alberto Naibo for helpful discussions.

References

D. Bridges and E. Palmgren. Constructive mathematics. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, summer 2018 edition, 2018.

W. Dean and S. Walsh. The prehistory of the subsystems of second-order arithmetic. *The Review of Symbolic Logic*, 10(2):357–396, 2017. doi:10.1017/S1755020316000411.

B. Eastaugh. Set existence principles and closure conditions: unravelling the standard view of reverse mathematics. *Philosophia Mathematica*, 27(2):153–176, June 2019. doi:10.1093/philmat/nky010.

H. Friedman. Some Systems of Second Order Arithmetic and Their Use. In R. D. James, editor, *Proceedings of the 17th International Congress of Mathematicians, Vancouver 1974*, volume 1, pages 235–242. International Mathematical Union, 1975. URL `https://www.mathunion.org/fileadmin/ICM/Proceedings/ICM1974.1/ICM1974.1.ocr.pdf`.

M. Hendtlass. Fixed point theorems in constructive mathematics. *Journal of Logic & Analysis*, 4 (10):1–20, June 2012. doi:10.4115/jla.2012.4.10.

D. R. Hirschfeldt. *Slicing the Truth: On the Computability Theoretic and Reverse Mathematical Analysis of Combinatorial Principles*. Number 28 in Lecture Notes Series, Institute for Mathematical Sciences, National University of Singapore. World Scientific Publishing Company, 2014. doi:10.1142/9208.

J. L. Hirst. Representations of reals in reverse mathematics. *Bulletin de L'Académie Polonaise Des Sciences: Série des sciences mathématiques, astronomiques, et physiques*, 55(4):303–316, 2007.

H. Ishihara. Reverse mathematics in Bishop's constructive mathematics. *Philosophia Scientiæ*, CS 6:43–59, 2006. doi:10.4000/philosophiascientiae.406. Special issue "Constructivism: Mathematics, Logic, Philosophy and Linguistics", edited by Gerhard Heinzmann and Giuseppina Ronzitti.

S. Shapiro. *Foundations Without Foundationalism: A Case for Second-Order Logic*. Number 17 in Oxford Logic Guides. Oxford University Press, Oxford, 1991. doi:10.1093/0198250290.001.0001.

S. G. Simpson. *Subsystems of Second Order Arithmetic*. Cambridge University Press, Cambridge, 2nd edition, 2009. doi:10.1017/cbo9780511581007.

R. M. Smullyan. *Theory of Formal Systems*. Princeton University Press, Princeton, 1961.