BOOK REVIEWS

BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 35, Number 1, January 1998, Pages 69–75
S 0273-0979(98)00735-6

Vicious circles: On the mathematics of non-wellfounded phenomena, by Jon Barwise and Larry Moss, CSLI Publications, Stanford, CA, 1996, iv + 390 pp., \$24.95, ISBN 1-57586-008-2

Circularity and the axioms of set theory

In recent years, various kinds of circular phenomena have been studied by researchers from different disciplines. Examples are the study of automata in the science of computation, which generally contain cycles of state transitions; recursive domain equations in computer science, which play a role in the semantics of concurrent programming languages with recursion; theories of truth and (self-)reference in philosophy; and terminological cycles in artificial intelligence.

Let us look at an example, taken from computation science, and treated in *Vicious circles* in all detail. *Streams* are infinite sequences of elements over an (alphabet) set A. Let a and b be in A and let s be the stream that informally is defined as consisting of an a followed by a b, again followed by an a and a b, and so on. Because of the fact that when the first two elements of s are removed, the same stream s is obtained again, it seems natural to model it by using the ordinary pairing operator: $s = \langle a, \langle b, s\rangle\rangle$, or, using t to denote the stream obtained by removing the first a from s, by the following system of equations: $s = \langle a, t\rangle$ and $t = \langle b, s\rangle$. This description is satisfactory in that it seems to capture our intuition well. However, it is inconsistent with the standard axioms of set theory.

The most familiar universe of sets is the so-called iterative or cumulative universe. It is described by the axioms of basic set theory, ZFC$^-$, to which the Foundation Axiom (FA) is added. This axiom states that the membership relation $\in$ on any set is wellfounded. An equivalent formulation is that the universe of all sets consists of a union of layers, corresponding to the various stages of construction. Starting with the empty set, at each stage new sets are built from already existing ones by means of the set-theoretical operations. It follows that the universe determined by ZFC (= ZFC$^-$ + FA) does not contain sets Ω that satisfy $\Omega = \{\Omega\}$. The stream s of our example above does not belong to this universe either: In the construction of s, the stream t is used which, at its turn, is constructed using s. This kind of circle is in the universe of ZFC to be considered 'vicious'.

1991 *Mathematics Subject Classification.* Primary 03-XX, 04-XX, 68-XX, 90-XX, 94-XX.

Dropping FA

If one wants to consider sets like Ω above, then FA, with which they are in contradiction, has to be dropped from ZFC. The following two quotations, which I have taken from Kunen's book on set theory [Kun80], illustrate that this is less dramatic than it may seem at first sight: "Likewise, our adopting the Axiom of Foundation does not comment on whether there are *really* (whatever that means) any x such that $x = \{x\}$; we are simply refraining from considering such x" [Kun80, page 95]. And: "Unlike the other axioms of ZFC, Foundation has no application in ordinary mathematics, since accepting it is equivalent to restricting our attention to **WF**[1], where all mathematics takes place anyway" [Kun80, page 101].

In view of the latter remark, omitting FA and replacing it, as we shall see below, with another axiom may be understood as adapting the axioms of set theory to the needs of mathematical practice, which seems, at least partly, to be taking place *outside* the realm of wellfounded sets.

The anti-foundation axiom

With FA out of the way, the axioms of set theory (ZFC^-) no longer prevent Ω and s from being members of the universe of sets. Still, none of the axioms asserts the existence of such sets. A richer universe of sets, to which *non-wellfounded* sets such as Ω and the stream s above do belong, can be obtained by replacing FA with another axiom, which can be thought of as a strong version of its negation. This axiom was first explored by Forti and Honsell [FH83] and later by Aczel [Acz88], who gave it the name AFA, for Anti-Foundation Axiom. The resulting collection of axioms, $\text{ZFC}^- + \text{AFA}$, is called ZFA, and the elements of the universe it describes are called *hypersets*.

There are many equivalent ways of formulating AFA. The original formulation of Forti and Honsell is that for every relational structure, there is a homomorphism onto a transitive set. Aczel's reformulation of it is very similar, and reads: every graph has a unique decoration. In the present book, Barwise and Moss adapt yet another version of AFA which they call the *Solution Lemma*, since it is based on a lemma of that name already present in [Acz88]. It will be explained in some detail below.

The objectives of *Vicious circles*

The objectives of *Vicious circles* are twofold. One is to study the *use* of hypersets in modelling 'real life' (possibly circular) phenomena. In particular, hypersets are used to deal with circularity in computer science (e.g., automata, transition systems, self-applicative programs), with circularity in philosophy (e.g., common knowledge and the Conway Paradox), and with semantical and logical paradoxes (such as the Liar Paradox and the Hypergame Paradox). The second goal of the book is to contribute to the study of the universe of hypersets. Taking Aczel's book as the point of departure, various new results are obtained, notably on bisimulation and coinduction, greatest fixed points, and corecursion.

The intended reader of this book is a researcher in one of the several fields just mentioned who might want to use hypersets for modelling certain phenomena. The reader need not be a mathematician, but some background in mathematics

[1] The class of all wellfounded sets.

and specifically set theory is required. Although some of the basic concepts of set theory are reviewed in the book, the reader is referred to textbooks on elementary set theory to learn or review this material. In particular, two recent textbooks are recommended by the authors: [Dev93] and [Mos94] (which are among the first textbooks to present hypersets).

The Solution Lemma

The Solution Lemma asserts that every system of equations of a certain form (which I shall not explain in detail but illustrate by means of some examples) (1) has a solution, which (2) is unique.

As one might expect, (1) gives rise to the existence of non-wellfounded sets. Consider for example the following simple system, consisting of only one equation $x = \{x\}$, in one indeterminate x. According to the Solution Lemma, there exists a set, say Ω, which satisfies the equation in the sense that $\Omega = \{\Omega\}$. Thus the existence of a non-wellfounded set has been established. For another example, let p and q be two given sets. They are used as 'constants' in the following equations, with indeterminates x, y, and z: $x = \{x, y\}$, $y = \{p, q, y, z\}$, $z = \{p, x, y\}$. Again, there exist sets, say a, b, and c, with $a = \{a, b\}$, $b = \{p, q, b, c\}$, and $c = \{p, a, b\}$. A third example, where the right side of the equations has some more structure, is $x = \langle a, y \rangle$ and $y = \langle b, x \rangle$. (As usual, a pair of sets $\langle p, q \rangle$ is a shorthand for the set $\{\{p\}, \{p, q\}\}$.) The solution of the latter system of equations consists of two (non-wellfounded) sets s and t, which satisfy $s = \langle a, t \rangle$ and $t = \langle b, s \rangle$ and which, as we saw above, represent two infinite streams.

Part (2) of the Solution Lemma gives us a way of proving equality. Consider, for instance, two sets a and b satisfying $a = \{a\}$ and $b = \{b\}$. Intuitively, they are the same, but note that the usual way of proving their equality, namely by looking at the elements of the two sets and showing that they are the same, is of no avail here. The equality of a and b follows, however, by (2) from the observation that both sets are a solution of the equation $x = \{x\}$.

Using the Solution Lemma

Although part (1) of the Solution Lemma makes it possible to define all the (non-wellfounded) sets one might want to have, including functions between them (which, of course, are themselves sets), and although part (2) is sufficient to prove all the equalities one needs, neither of them is very practical as soon as more complicated examples are considered. Therefore, the authors present a more abstract scheme for defining functions, which is derived from (1). It is called *corecursive*, since it is dual to the familiar scheme of recursion (although this is not made precise in the book). Similarly, a more practical proof method is derived from (2), based on the notion of *bisimulation relation*.

This machinery is then applied in the book to a number of examples, amongst which are graphs, modal logic, games, some semantical paradoxes, streams, and fixed points. Both in the development of the corecursion definition principle and the (bisimulation) proof method, and in the treatment of the examples just mentioned, the authors make new contributions to the theory of hypersets. These include various generalizations of Aczel's original solution lemma, a detailed analysis of so-called uniform and smooth operators on classes and their role in a formalization of corecursion, and various observations on modal logic.

An example: Streams

I shall briefly describe the application of the Solution Lemma (and its derived proof principle) to the example of streams, which I believe to be representative.

Let the set A^∞ of all infinite streams over a set A be defined as the collection of the solutions of all (possibly infinite) systems of equations of the form $\{ x_i = \langle a_i, x_j \rangle \mid i, j \in I,\, a_i \in A\}$, where I is some (countable) index set, and $\{x_i \mid i \in I\}$ is the set of indeterminates. It is shown in the book that, equivalently, A^∞ is the largest non-empty set satisfying the equation $x = A \times x$. The assumption of AFA (the Solution Lemma) is crucial here. Without AFA—and irrespective of whether FA is assumed or not—the equation has a unique solution, namely, the empty set.

Thus streams $s \in A^\infty$ are pairs of the form $s = \langle a, t\rangle$, for an element a in A and t again in A^∞. Using the Solution Lemma—or by applying the derived principle of corecursion—one can also define functions on streams. Without giving the details of a formal definition, there exists, for instance, a function $zip : A^\infty \times A^\infty \to A^\infty$, so that for all a and b in A, and s and t in A^∞, $zip(\langle a, s\rangle,\, \langle b, t\rangle) = \langle a, \langle b, zip(s, t)\rangle\rangle$.

Next we would like to prove properties of such functions, such as the following property of *zip*:

$$\text{(P)}:\ \forall a \in A\, \forall s \in A^\infty\, \forall t \in A^\infty,\ \ zip(\langle a, s\rangle, t) = \langle a, zip(t, s)\rangle.$$

The proof will make use of the 'bisimulation proof principle' announced above.

The concept of bisimulation was discovered several times in different places. The word bisimulation is due to Milner [Mil80], who used it for a notion of Park [Par81] (which the latter originally intended to call 'mimicry'). Since Aczel and Mendler's paper [AM89], we know that there exist infinitely many different kinds of bisimulation. The present book deals with a number of them, including bisimulation relations on sets, on systems of equations, on labelled graphs, on transition systems, and on streams. Here we look at bisimulation relations on streams.

A relation R on the set A^∞ of streams is a bisimulation if for all a and b in A and s and t in A^∞,

$$(\langle a, s\rangle,\, \langle b, t\rangle) \in R \Rightarrow \begin{cases} a = b & \text{and} \\ (s, t) \in R. \end{cases}$$

Using part (2) of the Solution Lemma, one can prove the following proof principle for streams: for all s and t in A^∞: if there exists a bisimulation relation R with $(s, t) \in R$, then $s = t$.

Thus in order to prove the equality of two streams, it is sufficient to establish the existence of a bisimulation between them. This (kind of) proof principle is nowadays often referred to as *coinduction proof principle*, although in the present book this term is reserved for a different (but equivalent) principle. Intuitively, the coinduction proof principle is valid: the first elements of two streams that are related by a bisimulation are equal. Because the remainders are again related, also their first elements are the same, that is, the second elements of the original streams. Continuing this way, the streams can be seen to be equal.

Property (P) of *zip* mentioned above can now be proved as follows. Let S be defined as

$$S = \{(zip(\langle a, s\rangle, t),\ \langle a, zip(t, s)\rangle) \mid a \in A,\ s, t \in A^\infty\}$$

and let R be defined as the union of S with its converse. The statement now follows from the fact that R is a bisimulation, which is shown next. Suppose

(s,t) is in R, say in S (the other case is similar). Then there are a, s', and t' such that $s = zip(\langle a, s'\rangle, t')$ and $t = \langle a, zip(t', s')\rangle$. Suppose $t' = \langle b, t''\rangle$. Note that, by the definition of zip, $s = zip(\langle a, s'\rangle, t') = \langle a, \langle b, zip(s', t'')\rangle\rangle$. Also $t = \langle a, zip(\langle b, t''\rangle, s')\rangle$. Because the first element of both s and t is a, and because $(\langle b, zip(s', t'')\rangle,\ zip(\langle b, t''\rangle, s'))$ is an element of the converse of S, and hence of R, both conditions of a bisimulation relation are fulfilled.

Wellfounded alternatives?

We saw that in ordinary set theory, without the assumption of AFA, and with or without the assumption of FA, the empty set is the only solution of the equation $x = A \times x$. If, however, we are willing to be satisfied with a set x for which there exists a bijective function between x and $A \times x$, denoted by $x \cong A \times x$, then many more solutions exist. Among those, there are two sets of special interest: the empty set and the set S defined by $S = \{s \mid s : \{0, 1, 2, \dots\} \to A\}$. The latter can be seen to satisfy the equation because of the existence of the bijection $\pi : S \to A \times S$, which maps a function s in S to the pair $\langle s(0), s'\rangle$, where the function s' is defined by $s'(n) = s(n+1)$. Both solutions can be characterized very precisely, using the language and insights from *category theory* (for which, e.g., [ML71] is a basic reference): the empty set is an *initial algebra* of the equation, and S a *final coalgebra*.

The notion of (final) coalgebra was already around before the publication of Aczel's book, but it was only afterwards that it was taken more seriously. One of the reasons has been that Aczel uses a final coalgebra as a model of ZFA. This construction, described by Aczel in a subsequent paper [AM89], was further generalized by Barr in 'wellfounded set theory' [Bar93]. Later it was realized that, still independent of the assumption of AFA, in final coalgebras both corecursion schemes and coinduction proof principles similar to the ones above are valid.

And so, one could also propose to take the above set S as the set-theoretic model of streams—in fact, this is what is usually done—and still apply the kind of reasoning as before. For instance, the definition of a bisimulation relation can be easily adapted: a relation R on S is a bisimulation if for all s and t in S: if $(s,t) \in R$ then $\pi(s)_1 = \pi(t)_1$ and $(\pi(s)_2, \pi(t)_2) \in R$ (where the subscripts refer to the first and second component of the pairs in $A \times S$). And the coinductive proof principle still holds: two elements s and t in S are equal if they are related by a bisimulation on S. (Its proof does not need AFA.)

Paraphrasing Barwise and Moss: "Choosing the best mathematical model of a phenomenon is something of an art, depending both on taste and technique." Assuming AFA allows one to reason about streams satisfying equalities like $s = \langle a, s\rangle$, which is from a set-theoretic perspective more elegant and simpler than an equation like $\pi(s) = \langle a, s\rangle$. The use of category theory, notably initial algebras and final coalgebras, gives an alternative approach to circular behaviour, which lives inside the world of wellfounded sets, and which allows for the development of a theory that is in many ways similar to that of the present book. These categorical methods have been used and further developed in computer science to find solutions of recursive domain equations, which serve as semantical models for concurrent and nondeterministic programming languages with recursion. Some pointers to the literature are [SP82] and [AJ94], where partially ordered sets are used; [BZ82] and

[BV96], which use metric spaces instead; and [AM82], [MA86], [Bar93] and [RT94], where plain (wellfounded) sets are used.

Conclusions and recent work by the authors

Vicious circles is a book of great importance for the theory of hypersets: not only does it offer a thorough yet very readable introduction to the subject, giving many often new and original examples of the use of hypersets for the mathematical modelling of non-wellfounded phenomena, the book also makes many substantial contributions to the further development of the theory itself. For both reasons, *Vicious circles* is mandatory for anyone with a serious interest in the subject.

The book ends by discussing a number of open problems and topics for future research. Some of these are treated in the following references to recent work by Barwise and Moss: in [MD97], the main issue is to compare the treatment of corecursion in *Vicious circles* to Aczel's Special Final Coalgebra Theorem; [Mos97] is a further improvement of the chapter in the book on modal logic from operators (without using AFA, though); [BM97] is a spinoff of the chapter on modal logic from the book, and gives a version of the modal correspondence theory for Kripke models (not frames); and [MS96] uses AFA and bisimulation in getting models of a simple fragment of situation theory.

Let me—in the spirit of these modern times of hypersets and hyperlinks—conclude with an Internet address that combines both: www.phil.indiana.edu/barwise/vccorrections.html. It is called 'Vicious Circles Home Page', and contains amongst others a (modest) list of corrections.

References

[Acz88] P. Aczel. *Non-well-founded sets.* Number 14 in CSLI Lecture Notes. Center for the Study of Languages and Information, Stanford, 1988. MR **89j:**03039

[AJ94] S. Abramsky and A. Jung. Domain theory. In S. Abramsky, D.M. Gabbay, and T.S.E. Maibaum, editors, *Handbook of Logic in Computer Science*, volume 3. Oxford Univ. Press, 1994. MR **97b:**68121

[AM82] M.A. Arbib and E.G. Manes. Parametrized data types do not need highly constrained parameters. *Information and Control*, 52(2):139–158, 1982. MR **84i:**68014

[AM89] P. Aczel and N. Mendler. A final coalgebra theorem. In D.H. Pitt, D.E. Ryeheard, P. Dybjer, A. M. Pitts, and A. Poigne, editors, *Proceedings category theory and computer science*, Lecture Notes in Computer Science, pages 357–365, 1989. MR **91f:**18001

[Bar93] M. Barr. Terminal coalgebras in well-founded set theory. *Theoretical Computer Science*, 114(2):299–315, June 1993. Some additions and corrections were published in TCS Vol. 124, 1994, pp. 189–192. MR **94m:**18006a; MR **94m:**18006b

[BM97] J. Barwise and L.S. Moss. Modal correspondence for models. *Journal of Philosophical Logic*, 1997. To appear.

[BV96] J.W. de Bakker and E. de Vink. *Control Flow Semantics.* Foundations of Computing Series. The MIT Press, 1996. MR **96h:**68074

[BZ82] J.W. de Bakker and J.I. Zucker. Processes and the denotational semantics of concurrency. *Information and Control*, 54(1/2):70–120, July/August 1982. MR **85e:**68056

[Dev93] K. Devlin. *The Joy of Sets.* Undergraduate Texts in Mathematics. Springer-Verlag, 1993. MR **94e:**03001

[FH83] M. Forti and F. Honsell. Set theory with free construction principles. *Annali Scuola Normale Superiore di Pisa*, X(3):493–522, 1983. MR **85f:**03054

[Kun80] K. Kunen. *Set Theory*, volume 102 of *Studies in logic and the foundations of mathematics.* North-Holland, 1980. MR **82f:**03001

[MA86] E.G. Manes and M.A. Arbib. *Algebraic approaches to program semantics.* Texts and Monographs in Computer Science. Springer-Verlag, 1986. MR **88d:**68069

[MD97] L.S. Moss and N. Danner. On the foundations of corecursion. *Logic Journal of the IGPL*, pages 231–257, 1997. CMP 97:08

[Mil80] R. Milner. *A Calculus of Communicating Systems*, volume 92 of *Lecture Notes in Computer Science*. Springer-Verlag, Berlin, 1980. MR **82g:**68001

[ML71] S. Mac Lane. *Categories for the Working Mathematician*, volume 5 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1971. MR **50:**7275

[Mos94] Y.N. Moschovakis. *Notes on set theory*. Undergraduate Texts in Mathematics. Springer-Verlag, 1994. MR **95a:**04001

[Mos97] L.S. Moss. Coalgebraic logic. *Annals of Pure and Applied Logic*, 1997. To appear.

[MS96] L.S. Moss and J. Seligman. Situation theory. In J. van Benthem and A. ter Meulen, editors, *Handbook of Logic and Language*, chapter 4, pages 239–309. Kluwer Acad. Publishers, 1996.

[Par81] D.M.R. Park. Concurrency and automata on infinite sequences. In P. Deussen, editor, *Proceedings 5th GI conference*, volume 104 of *Lecture Notes in Computer Science*, pages 15–32. Springer-Verlag, 1981.

[RT94] J.J.M.M. Rutten and D. Turi. Initial algebra and final coalgebra semantics for concurrency. In J.W. de Bakker, W.-P. de Roever, and G. Rozenberg, editors, *Proceedings of the REX School/Symposium 'A decade of concurrency'*, volume 803 of *Lecture Notes in Computer Science*, pages 530–582. Springer-Verlag, 1994. FTP-available at `ftp.cwi.nl` as pub/CWIreports/AP/CS-R9409.ps.Z. MR **96d:**68141

[SP82] M.B. Smyth and G.D. Plotkin. The category-theoretic solution of recursive domain equations. *SIAM Journal of Computation*, 11(4):761–783, November 1982. MR **84d:**68052

J.J.M.M. Rutten
Centrum voor Wiskunde en Informatica (CWI)
E-mail address: `janr@cwi.nl`