Computing Aspects of Set Theory

by

Erwin Engeler, ETH Zürich

For C. S. Tang in friendship and appreciation on the occasion of his 70th birthday

In today's talk I would like to address a topic that lies close to the heart of most logicians, namely the foundations of set theory. In a sense, this topic stood at the origin of the technical development of mathematical logic in the hands of Frege, Peano and its other founding fathers. In fact, without it, mathematical logic may not have been around as a developed discipline when it became necessary – we love to believe – as an integral part for the foundations of computer science.

Logic in computer science is a many-faceted ongoing enterprise, as is apparent from the proceedings of the yearly conferences which occur under this heading. The logics (plural) that are more or less gainfully employed within this enterprise are very diverse: from modal logics – where Professor Tang's contributions are attached – to infinitary logics – originally my own domain.

Among the many data structures with which present programming languages operate, we also find the concept "set". It is generally relegated to a very minor role, quite contrary to that reductionist philosophy of mathematics which would like to see all of mathematics reduced to sets (in my opinion a most unattractive prospect, although doubtlessly do-able). I do not propose here to transform that role into a central one, although logicians have often been responsible for contributions to the design of programming languages. Rather this talk concentrates on the reverse question: What can set theory learn from the foundations of computer science? In particular, what can we say about the existence and nature of non-sets?

The "classes" of von Neumann-Bernays-Gödel set theory (vN–G–B) were introduced as a refuge for those collections of sets which could not properly be called sets because of their antinomial behaviour: the ordinal numbers, the collection of all sets, etc. In his seminal paper [12] von Neumann laid the blame squarely on their size and made any logically defined collection a proper class ("ein II-Ding") if it turned out to be "too big". He was quite aware of the provisional, and not too convincing, character of this distinction, but he lacked any hint towards deeper insight into the problematics of the antinomies. Quine's "New Foundations" [10] proposed a syntactic way of avoiding non-sets, but it lacked even more the quality of mathematical intuitiveness, and moreover led to rather unexpected consequences (e.g. Specker [11]). This is not the place to review the manifold developments of this question, suffice it to say that, in my opinion, it left us far from any universally accepted solution to the problem of sundering sets from non-sets.

There is one aspect in von Neumann's finite axiomatization of set theory, which he himself discounted as a minor technicality, and which was removed in the subsequent modifications by Bernays [1] and Gödel [6]: the basic entities are not "really" sets but are intuitively thought of as "functions", and so named before they are formally separated into I-things and II-things ("I-Dinge" and "II-Dinge"). Thus, von Neumann's axiomatization is in a sense also an axiomatization of the notion of function and application of function to arguments.

I have repeatedly held (e.g. in [2], chapter 3) that the proper axiomatization of the general notion of function and application is Combinatory Logic. The purpose of this talk is to develop the first few steps of the study of this programmatic point of view in the context of set theory. Since combinatory logic is one of the possible foundations of computer science (especially if enriched with types it serves very well, e.g. as typed lambda calculus), this point of view corresponds clearly to the intention formulated in the introduction.

The models of combinatory logic are called combinatory algebras; they consist of a nontrivial (at least two elements) set D and a binary operation "$\cdot$" called application. The axioms can either be given by a scheme ("combinatory completeness"):

For every term $t(x_1, \ldots, x_n)$ built up from variables among $x_1, \ldots, x_n$ and using application, there exists an element T in D, for which $T \cdot x_i \cdots x_n$ equals $t(x_1, \ldots, x_n)$, {application parenthesizes to the left}.

Or, equivalently, by singling out two elements $\boldsymbol{S}$ and $\boldsymbol{K}$ in D with the two axioms

$$\begin{gathered}\boldsymbol{K} \cdot x \cdot y = x, \\ \boldsymbol{S} \cdot x \cdot y \cdot z = x \cdot z \cdot (y \cdot z).\end{gathered}$$

If A is a non-empty set, we can construct a combinatory algebra from it by using sets of lists (of lists, ...), defined recursively as follows:

$$\begin{aligned} G_0(A) &= A, \\ G_{i+1}(A) &= G_i(A) \cup \{\alpha \to a : a \in G_i(A), \alpha \subseteq G_i(A) \text{ finite}\}, \end{aligned}$$

where $\alpha \to a$ denotes the list with head a and tail α. Then, taking

D_A as the set of subsets of $\bigcup_i G_i(A)$ we define the binary operation of application on D_A by

$$M \cdot N = \{a : \exists \alpha \subseteq N.\ \ \alpha \to a \in M\}.$$

With this operation, as is turns out, D_A *constitutes a combinatory algebra* $\mathcal{D}_A$. It is very rich because we have taken all subsets of $\bigcup_i G_i(A)$ indiscriminately, (while we could have restricted ourselves e.g. to the recursively enumerable sets). More on this later.

Already in my first paper on this topic [3], I observed that the richness mentioned above allows the isomorphic embedding of any algebraic structure (of limited cardinality) into $\mathcal{D}_A$. We also mentioned the possibility of doing the same with relational structures, without giving a statement or proof. Since this remark was never taken up in subsequent developments (for an overview of this research programme see [4]), we now state it as a theorem.

Representation theorem (for relational structures). *Let* $\mathcal{A} = \langle A, R\rangle$ *be a relational structure,* R *being a binary relation over the nonempty set* A*. Then there exist an injection* $f : A \to D_A$ *with the property that for any* $a, b \in A$ *we have* $\langle a, b\rangle \in R$ *iff* $f(a) \cdot f(b) = f(b)$.

The proof is straightforward: Let f_n be defined recursively by

$$f_0(a) = \{a\},$$

$$f_{n+1}(a) = f_n(a) \cup \{\{b\} \to x : x \in f_n(b), \langle a, b\rangle \in R\}.$$

Then $f(a) = \bigcup_n f_n(a)$. This f is injective, because $f(a) \cap A = \{a\}$ for all $a \in A$. The relation R is indeed represented by virtue of the proposed correspondence. Namely, take any $\langle a, b\rangle \in R$. Then

$$x \in f(a) \cdot f(b) \text{ iff } \exists \alpha \subseteq f(b).$$

$$\alpha \to x \in f(a), \text{ iff } \exists n.\ x \in f_n(b), \text{ iff } x \in f(b).$$

Conversely, assume $f(a) \cdot f(b) = f(b)$. Since $b \in f(b)$, we have $b \in f(a) \cdot f(b)$

and therefore

$$\exists \alpha \subseteq f(b).\ \alpha \to b \in f(a),$$

which implies

$$\exists n.\{b\} \to b \in f_n(a) \wedge \langle a, b\rangle \in R$$

by definition of f.

Let us now take a model of ZFC (Zermelo-Fraenkel set theory including the axiom of choice), say $\mathcal{M} = \langle M, \in\rangle$ and take the above representation, using $\langle a, b\rangle \in R$ for $b \in a$. Thus, $b \in a$ is interpreted by the equation $f(a) \cdot f(b) = f(b)$ inside $\mathcal{D}_M$. The following remarks serve to point out the strikingly transparent correspondence between the structure of the membership relation on a set in M (and its elements) and the set of objects in $\mathcal{D}_M$ onto which is mapped by $f(a)$. Taking a closer look at $f(a)$, we observe

$$\begin{aligned} f(a) \quad = \quad & \{a\} \\ & \cup\{\{b\} \to b : b \in a\} \\ & \cup\{\{b\} \to c(\{c\} \to) : c \in b \in a\} \\ & \cup\{\{b\} \to (\{c\} \to (\{d\} \to d)) : d \in c \in b \in a\} \\ & \cup \ldots \end{aligned}$$

Thus, parenthesizing arrows to the right and dropping the braces around the singletons, $f(a)$ consists of a and all arrow-chains $a_1 \to a_2 \to a_3 \to \ldots \to a_n \to a_n$ for the chains of elementhood $a_n \in a_{n-1} \in \ldots \in a_2 \in a_1 \in a$. Since $\mathcal{M}$ satisfies the axiom of foundation, no "a_i" occurs twice in the $\in$-chain, and all chains are finite.

Of course, considering sets as trees of elements-of-elements is far from the original; this point of view has been taken and used already by Fraenkel in the 1920s in connection with his treatment of the axiom of choice (without foundation). What is new here is the fact that the collection of sets has now been embedded in combinatory algebra and what becomes interesting is to watch the action of combinators on the embedded objects.

For example, the identity combinator $\boldsymbol{I}$, having the property $\boldsymbol{I} \cdot x = x$ for all elements x in the combinatory algebra, and which in $\mathcal{D}_M$ has the form

$$\boldsymbol{I} = \{\{x\} \to x : x \in G(A)\},$$

can serve as an object corresponding to the "set of all sets", since obviously $\boldsymbol{I} \cdot f(a) = f(a)$ for all sets a. Clearly, it is not the only such object. But there is no such object which is (the f-image of) a set. It simply does not have the right shape!

The classical proof of the consistency of set theory with classes relative to ZF, [9] operates with collections of (model-) sets as classes, defined by their properties. It makes use of the fact, used also for finite axiomatizability of (vN–G–B), that these properties compose from a short list of binary predicates on sets; these gave rise to Gödel's constructors $F_1 - F_8$. It seems a straightforward matter to identify the elements of the combinatory algebra $\mathcal{D}_M$ that correspond to these classes, resp. operations, and it is to be hoped that further insight into the shapes of non-sets arise in this manner.

By embedding any structure in a combinatory algebra, we automatically provide it with a notion of programs executable in this structure. This was of course first noticed with natural numbers and partial recursive functions (in lambda calculus; for a historical survey see Kleene [7]), but the fact is quite general and thus also applies to (embedded) set theory. Already without any special concern about computability, just by using the functional character of elements of $\mathcal{D}_M$ (viewed as left-multipliers) we gain an additional viewpoint. In particular, it would be interesting to delineate the possibilities of the replacement axiom if this principle is reinterpreted not by (functional) classes of pairs but by some – probably restricted – set of elements f in $\mathcal{D}_M$. These f would have to satisfy

$$\forall a \exists b \forall x. \quad a \cdot x = x \quad \Leftrightarrow \quad b \cdot (f \cdot x) = f \cdot x.$$

Of course, not only the class-constructions (used for the axioms of separation and replacement) are to be investigated as combinatory objects, but also the proper constructions on sets, such as pairing, sum- and powerset-constructions. Indeed, it is to be expected that by putting them into this framework, an algebraic version of set theory inside combinatory algebra will emerge.

The emerging "algebraic from set theory" could also profit from some newer developments that in some sense start form a similar idea: While we took a prefabricated framework, combinatory algebras, for an abstract notion of computability on sets, Moss [8] develops such a theory *ab initio* as a theory of power set recursive functions.

I have well-founded hopes that at least a part of the research programme sketched above will be realized in the near future: luckily one of my PhD students, Mr. Darms, has taken an interest in the subject and is expected to carry it to term.

References

[1] P. Bernays, A system of axiomatic set theory I – VII. J. of Symbolic Logic **1** (1937) – **19** (1954). Reprinted in: G. H. Müller (ed.), Sets and Classes, North-Holland 1976, pp. 1–119.

[2] E. Engeler, Metamathematik der Elementarmathematik. Springer 1983. Engl. Translation: Foundations of Mathematics, Springer 1993.

[3] E. Engeler, Algebras and Combinators. Algebra **13** (1981), pp. 339–392. Reprinted in: E. Engeler, Algorithmic Properties of Structures. World Scientific Publ. Co. 1993, pp. 185-188.

[4] E. Engeler et al., The Combinatory Programme. Birkhäuser 1995.

[5] E. Engeler, Algorithmic Properties of Structures. World Scientific Publ. Co. 1993.

[6] K. Gödel, The Consistency of the Continuum Hypothesis. Princeton Univ. Press, 1940.

[7] S. C. Kleene, Origins of revursive function theory. Ann. Hist. Comput. **3** (1981), pp. 52–67. In particular: S. C. Kleene, A theory of positive integers in formal logic. Am. J. of Math. **57** (1935), pp. 153–173 and 219–244.

[8] L. S. Moss, Power set recursion. Annals of Pure and Applied Logic **71** (1995), pp. 247–306.

[9] I. L. Novak, A construction for models of consistent systems. Fund. Math. **37** (1951), pp. 87–110.

[10] W. V. Quine, New foundations for mathematical logic. Am. Math. Monthly **44** (1937), pp. 70–80. Reprinted in: W. V. Quine, From a logical point of view. Harvard University Press 1953, pp. 80–101.

[11] E. P. Specker, The axiom of choice in Quine's New Foundations for Mathematical Logic. Proc. Nat. Acad. Sci. (USA), **39** (1953), pp. 972–975. Reprinted in: Ernst Specker Selecta, (ed. by G. Jäger et al.). Birkhäuser 1990, pp. 71–74.

[12] J. von Neumann, Eine Axiomatisierung der Mengenlehre. J. f. Math. **154** (1925), pp. 219–240; corr. **155** (1926), pp. 128. Engl. Transl. in: J. van Heijenoort, From Frege to Gödel. Harvard University Press 1967, pp. 339–342.