THE TWO MEANINGS OF THE WORD “SET” IN MATHEMATICS

MATHIEU DUPONT/BRECKES

The word “set” has two uses in mathematics:

1. in a sentence like “a group is a *set* equipped with the following operations (...), such that the following axioms are satisfied (...)”, the word “set” denotes a base domain for mathematical structures, equipped with an equality relation (needed to express the axioms of group);
2. in the phrase “the *set* of even numbers”, the word “set” denotes a subdomain of the domain of natural numbers (which is a set in the first meaning).

One dogma of 20th century mathematics was that one cannot define sets. But this applies only to sets in the sense of objects of a so-called set theory, like ZF, not to sets in the two mathematical meanings, which *can* be defined.

We will work in a sufficiently rich typed language. A general type won't be assumed to have any structure, in particular we don't preequip types with an equality relation. We denote by $t\colon A$ the syntactical assertion that the term t is of type A (for example “$2 + 3\colon \mathbb{N}$” means that the type of the term $2 + 3$ is $\mathbb{N}$, and “let be $x\colon A$” is used to declare that the variable x is of type A). It is important not to use $\in$ for this relation between terms and types in the metalanguage, to avoid confusion with the membership relation between elements and sets^2 in the language of mathematics (see below). Remark: some people in type theory call types “sets”, which add to the confusion.

1. Meaning 1

We call *set*^{1} a type equipped with an equivalence relation (or *equality*), allowing to express that two objects of this type are equal or not. This is essentially the definition by Errett Bishop in *Foundations of constructive analysis*; it is usually used in formalisations of mathematics in proof assistants (in this context, it is sometimes called a *setoid*). As in Bishop's book, an *operation* f from a type A to a type B is given, for each $x\colon A$, by an object $f(x)\colon B$. A function $f\colon A \to B$ between two sets^1 is an operation preserving equality.

The only structure on the domain of a group needed to express the axioms of group is an equality, so the more general definition of a group one can give is a set^1 equipped with the usual operations satisfying the axioms of group. More generally, a model of a (let's say first-order) theory with equality is a set^1 equipped with some operations, relations, and satisfying some equations.

Some examples of sets^1 are:

- the set^1 of natural numbers, whose objects are the natural numbers with the ordinary equality of natural numbers;

Date: 2009/03/03.

- the set^1 of natural numbers modulo n, whose objects are the natural numbers and equality is equality modulo n;
- the set^1 of rational numbers, whose objects are the quotients $\frac{m}{n}$, where m and n are two integers such that $n \neq 0$, with the usual equality of rationals ($\frac{m}{n} = \frac{m'}{n'}$ iff $mn' = nm'$);
- the set^1 of functions $A \to B$ between two sets^1, with pointwise equality: by definition, two functions $f, g \colon A \to B$ are equal if, for every $x \colon A$, $f(x) = g(x)$;
- the empty set^1, which has no objects, the equality is trivial;
- the one-element set^1 which has only one object, equal to itself.

Let us recall that a *category* is a type $\mathcal{C}$ equipped, for each pair of objects A, B, with a set^1 $\mathcal{C}(A, B)$ (whose objects are called "morphisms from A to B"), with an associative composition $c_{A,B,C} \colon \mathcal{C}(A, B) \times \mathcal{C}(B, C) \to \mathcal{C}(A, C)$ and, for each object A, an identity $1_A \colon \mathcal{C}(A, A)$, which is a neutral element for the composition.

The right identity criterion for sets^1 and the constructions one can use to create new sets^1 follow from the following assertion:

sets^1 live in a category,

whose morphisms are the functions, denoted by Set. As a consequence:

(1) the identity criterion for objects in a category being isomorphism, two sets^1 are considered to be "the same" if there exists a bijection (an isomorphism in the category of sets^1) between them;
(2) the constructions one can do with sets^1 are the limits, colimits, etc. in the category of sets^1: cartesian product, coproduct ("disjoint union"), quotient of a set^1 by an equivalence relation, exponentiation, etc.

The first point (as well as the cardinality functor described below) suggests that sets^1 are the right formalisation of Cantor's notion of cardinal number (as Lawvere has noticed). This identity criterion is the right one for sets as bases of mathematical structures: if a set^1 A is equipped with a certain structure, and if it is in bijection with B, then we can transfer the structure from A to B. Two sets^1 equipped with the same kind of structure (e.g. two groups, two rings, etc.) are "the same" if there is a bijection between them preserving and reflecting the structure.

On the other hand, we cannot define two sets^1 to be equal if they have the same elements, because there is no way to compare the objects of two (syntactically) different sets.

2. Meaning 2

The second meaning of "set" in mathematics is probably the first, historically: this is a collection of objects of a same type, like sets of oranges, sets of natural numbers, sets of functions from A to B, etc. In a second-order theory, the "sets" over which one quantify are interpreted in a model as being the sets^2 of the domain of the model, which is a set^1.

Let us fix a set^1 A. We define a *set^2 of type* A as an injection $m \colon M \rightarrowtail A$. A common abuse of notation consists in writing just M instead of (M, m), which contributes to the confusion of the two meanings of "set". It is for this notion of set that a membership relation is available, between objects and sets^2 of type A: we say that $a \colon A$ is an *element* of $m \colon M \rightarrowtail A$ (and we write $a \in (M, m)$) if a is in the range of m (there exists $x \colon M$ such that $a = m(x)$). Here are a few examples:

- if we are given n different objects $a_1, \dots, a_n$ of A, there is a set^2
$$\{a_1, \dots, a_n\}\colon [n] \rightarrowtail A,$$
where $[n]$ is the set^1 with n elements (let us denote its elements by $1, \dots, n$) and where the injection maps i to a_i; its elements are exactly the objects equal to one of $a_1, \dots, a_n$;
- if we are given a property $P(x)$, with a free variable x of type A, we can construct the set^2 of the objects of A satisfying P: it is
$$\{x\colon A \mid P(x)\}\colon M \rightarrowtail A,$$
where the objects of the set^1 M are the objects of A satisfying P, the equality of M is defined as in A, and the inclusion maps such an object to itself.

We can define an order relation on sets^2 of type A: we say that $(M, m) \subseteq (M', m')$ if every element of (M, m) is an element of (M', m'). So

sets^2 of type A live in an order

(i.e. a type equipped with an order relation; sets^1 are the orders whose order relation is symmetric), which we denote by $\mathcal{P}(A)$.

As a consequence,

1. the criterion of identity for objects of an order being given by $a = b$ iff $a \leq b$ and $b \leq a$, two sets^2 $m\colon M \rightarrowtail A$ and $m'\colon M' \rightarrowtail A$ are "the same" if they have the same elements (so, for sets^2 of A we have extensionality);
2. it is with this notion of set that the usual constructions of "naïve set theory" are available: these are the typical constructions in an order, i.e. intersection, union, implication, the smallest set^2 (the empty set^2 of type A), the greatest set^2 (the identity on the whole set^1 A).

Every set^2 of type A can be equivalently described as a selection of some of the objects of A: the set^2 $m\colon M \rightarrowtail A$ is equal to the set^2 $\{x\colon A \mid x \in (M, m)\}$.

So, for each set^1 A, we get a model of a *genuine* set(^{2}) theory: this is a theory with two sorts (let's denote them by A and $\mathcal{P}(A)$), with a binary relation symbol $=$ on A (satisfying the axioms for equality), a binary relation symbol $\subseteq$ on $\mathcal{P}(A)$ (satisfying the axioms for order), and a binary relation symbol $\in$ from A to $\mathcal{P}(A)$, satisfying two axioms:

1. extensionality: $S \subseteq T$ iff for all $a\colon A$, $a \in S$ implies $a \in T$
2. comprehension: for any property $P(x)$, where $x\colon A$, there is $\hat{P}\colon \mathcal{P}(A)$ such that $x \in \hat{P} \Leftrightarrow P(x)$.

3. Comparison; cardinality

It is important to realise that these two notions of set are very different: they live in different structures and have different criteria of identity. But there is a link between these two notions: if we fix a set^1 A, there is a "forgetful" functor $\sharp\colon \mathcal{P}(A) \to \mathsf{Set}$, which maps a set^2 $M \rightarrowtail A$ to the set^1 M. This is Cantor's cardinality functor for sets^2 of type A (but of course at Cantor's time, the notion of functor didn't exist): it maps a set^2 of type A to the "number" of objects of this set^2, which is a set^1. For example, it maps the set^2 $\{1, 9, 8, 0\}\colon [4] \rightarrowtail \mathbb{N}$ to the four-elements set^1.

Of course, the cardinality functor can make different sets^2 identical. For example, in the order $\mathcal{P}(\mathbb{N})$, the set^2 of even numbers is strictly smaller than the set^2 of all natural numbers, but they have the same "number" of objects, i.e., in the category of sets^1, the set^1 whose objects are the even numbers is in bijection with $\mathbb{N}$, so they are the same set^1, when we forget the injection into $\mathbb{N}$ (we can even represent the set^2 of even numbers as the injection $2 \cdot - \colon \mathbb{N} \to \mathbb{N}$, since its elements are exactly the even numbers). And the set^2 $\{1, 5, 0\}$ in $\mathcal{P}(\mathbb{N})$ and the set^2 $\{\sin, \cos, \exp\}$ in $\mathcal{P}(\mathbb{R}^{\mathbb{R}})$ live in different orders, so we cannot even ask if they are the same, but if we forget the inclusions into $\mathbb{N}$ or $\mathbb{R}^{\mathbb{R}}$, we get just two different presentations of the three-element set^1.

4. Meaning 3?

There is a third meaning, which is not included in the title, since it is not in use in mathematics, but only in logic and philosophy of "mathematics", and sometimes, I'm afraid, in category theory. Moreover, it seems abusive to use the word "set" for this third notion. Originally, it is a special case of the second meaning, the special case we get if, in the theory of sets^2 given at the end of Section 3, we identify the two sorts A and $\mathcal{P}(A)$, so that there is only one kind of objects, which play both roles: they are elements and sets^2 at the same time. Of course, by Cantor's theorem this is inconsistent (this gives Russell's paradox). People tried to make consistent this inconsistent theory by restricting in a more or less elegant way the properties to which applies the scheme of comprehension. This has given various so-called "set theories" such as Zermelo-Fränkel, Bernays-Gödel, New Foundations, etc., formalising neither the notion of set^1, nor the notion of set^2, but some hybrids between elements and sets^2. Let us call *set*^{3} an object of such a theory.

If we have a model of such a theory (if such a model exists...), with U the base set^1 of all sets^3, to each set^3 $x \colon U$ corresponds a set^2 of type U, the set^2 of all sets^3 "belonging" to x, and so corresponds a set^1, by applying to it the cardinality functor. In this way, when we pretend to do mathematics in such a theory, sets^3 play the rôle both of sets^1 and sets^2 (this is probably the origin of the confusion between these two notions). Moreover, the "mathematics" one can develop in such a theory, once an encoding of the basic notions is given, not only admit meaningless assertions (for example, in ZF with the most used encoding of ordered pairs and natural numbers, we can express and even prove the proposition $(\mathbb{N} \times \mathbb{N}) \cap \mathcal{P}(\mathbb{N}) = (1, 2)$), but moreover are strongly limited:

1. all sets^1 are canonically sets^2 of a big set^1 U (which in some theories is not even represented by a set^3 [e.g. in Z.F., because of Russell's Paradox]), hence they share a universal equality (the equality of U), and so:
2. one cannot anymore speak of types not equipped with an equality (for example, we cannot define general categories, without equality at the level of objects, such as the category of sets^1).

5. Conclusion

As far as I am concerned, I always distinguish sets^1 and sets^2. I reserve the word "set" for sets^1 and use "subset of A" for sets^2 of type A (following the tradition in category theory). And, of course, I never use sets^3, since I've never seen any use for them.

6. Links

- For more details about the definitions, see Bishop’s set theory by Erik Palmgren (and Bishop’s book).
- In topos theory, one studies categories whose objects behave like sets(1). Then subobjects of a given object A are usually defined, like sets[2], as monomorphisms (which generalise injections) with codomain A. See Sets for mathematics by Lawvere and Rosebrugh, or Todd Trimble’s posts on his blog about the Elementary Theory of the Category of Sets for a semi-elementary approach to toposes. Note that, in a topos, for each object A, there is an object $\mathcal{P}(A)$ playing the rôle of the “object of all subobjects of A”; but I think that we shouldn’t assume that all subsets of a set(1) A form themselves a set(1), but rather an order, so I think that topos theory should be based on orders rather than on sets[1]; I intend to speak about that later.
- The paper *In the Search of a Naive Type Theory*, Lecture Notes in Computer Science 4941 (2008), by Agnieszka Kozubek and Paweł Urzyczyn (here are slides with the same title) advocates the same distinction; here is an excerpt of the introduction:

  > In fact, there are two very basic intuitions that are glued together into the notion of a “set”:
  >
  > – Set as a domain or universe;
  > – Set as a result of selection.
  >
  > We used to treat this identification as natural and obvious. But perhaps only because we were taught to do so. These two ideas are in fact different, and this very confusion is responsible for Russel’s paradox. In addition, ordinary mathematical practice often makes an explicit difference between the two aspects.

- See also Towards a categorical foundation of mathematics by Michael Makkai.