[11-01-12-H]

This book is dedicated to Prabir Roy, Louis McAuley, Jonathan Seldin, Anil Nerode, and Maurice Boffa, my teachers, and to W. V. O. Quine and R. B. Jensen, the founders of this style of set theory.

Cahiers du Centre de logique
Volume 10

Contents

Cahiers du Centre de logique
Volume 10

Chapter 1

Introduction: Why Save the Universe?

This book is intended for one of two uses. It could be used as an introduction to set theory. It is roughly parallel in structure to Halmos's classic *Naive Set Theory*, though more topics have been added. The book contains exercises in most chapters, in line with its superficial character of being an elementary set theory text, but no representation as to pedagogical soundness is made. Some of the exercises are very hard (these are marked).

The other possible use of the book is to demonstrate the naturalness and effectiveness of an alternative set theory, Jensen's corrected version NFU of W. V. O. Quine's system "New Foundations" (NF), so-called because it was proposed in Quine's paper *New foundations for mathematical logic* in 1937. No introduction to set theory based on Quine's approach has appeared (to my knowledge) since J. B. Rosser's *Logic for Mathematicians*, which came out in 1953 (second edition 1978). As our title implies, NFU is a set theory in which there is a universal set.

Quine's "New Foundations" has a bad reputation. Its consistency relative to ZFC remains an open question. In 1953 (immediately after Rosser published his book *Logic for Mathematicians*, which was based on "New Foundations"), E. Specker proved that "New Foundations" proves the negation of the Axiom of Choice! No one has been able to derive a con-

tradiction from NF, but the failure of AC makes it an unfriendly theory to work in.

None of this should be allowed to reflect on NFU. NFU is known to be consistent since the work of Jensen in 1969. It can be extended with the Axiom of Infinity (this is implicit in our Axiom of Projections). It is consistent with the Axiom of Choice. It can be extended with more powerful axioms of infinity in essentially the same ways that ZFC can be extended; we introduce an Axiom of Small Ordinals which results in a theory at least as strong as ZFC (actually considerably stronger, as has been shown recently by Robert Solovay). We believe that if Jensen's result had been given before Specker's, the subsequent history of interest in this kind of set theory might have been quite different.

These considerations are not enough to justify the use of NFU instead of ZFC. What positive advantages do we claim for this approach? The reason that we believe that NFU is a good vehicle for learning set theory is that it allows most of the natural constructions of genuinely "naive" set theory, the set theory of Frege with unlimited comprehension (it can be claimed that Cantor's set theory always incorporated "limitation of size" in some form, even before it was formalized). The universe of sets is actually a Boolean algebra (there is a universe; sets have complements). Finite and infinite cardinal numbers can be defined as equivalence classes under equipotence, following the original ideas of Cantor and Frege. Ordinal numbers can be defined as equivalence classes of well-orderings under similarity. The objects which cause trouble in the paradoxes of Cantor and Burali-Forti (the cardinality of the universe and the order type of the ordinals) actually exist in NFU, but do not have quite the expected properties. Many interesting large classes are actually sets: the set of all groups, the set of all topological spaces, etc. (most categories of interest, actually). The reason that the paradoxes are avoided is a restriction on the axiom of comprehension. These paradoxes, and the paradox of Russell, are discussed in the text.

The restriction on the axiom of comprehension which NFU shares with "New Foundations" is not very easy to justify to a naive audience, unless one starts with a presentation of the Theory of Types. We do not introduce NFU in this way; we use a finite axiomatization (Hailperin first showed that NF is finitely axiomatizable) to introduce NFU in a way quite analogous to the way that ZFC is usually introduced, as allowing the con-

struction of sets and relations using certain basic operations. The basic operations used are quite intuitive. Set-builder notation is introduced early, and the impossibility of unrestricted comprehension is pointed out in the usual way using Russell's paradox. After the arsenal of basic constructions is assembled, the schema of stratified comprehension is proved as a theorem.

Another point for those familiar with the usual treatment of NF and NFU is that we use notation somewhat more similar to the notation usually used in set theory, avoiding some of the unusual notation going back to Rosser or Principia Mathematica which has become traditional in NF and related theories. Our notation does have some eccentricities, which are discussed in the section titled "Parentheses, Braces and Brackets" (p. 76).

The usual set theory of Zermelo and Fraenkel is not entirely neglected; there is an introduction to the usual set theory as an alternative, motivated in the context of NFU by a study of the isomorphism types of well-founded extensional relations, in section 19.2.

There is a study of somewhat more advanced topics in set theory at the end, including the proof of Robert Solovay's theorem that the existence of inaccessible cardinals follows from our Axiom of Small Ordinals. There is also a discussion of the analogous system of "stratified λ-calculus" in which the notion of *function* rather than the notion of *set* is taken as primitive.

All references are deferred to the section of Notes at the end (Chapter 24, p. 217).

Exercise

Find and read Quine's original paper *New foundations for mathematical logic*.

Cahiers du Centre de logique
Volume 10

Chapter 2

The Set Concept. Extensionality. Atoms

If the reader glances at the Introduction, he may expect an unusual treatment of set theory and its use as a foundation for mathematics in these pages. If our premise is correct, this will not be the case. The basic fact that mathematics is here founded on the undefined concepts of *set* and *membership* is unchanged. The notion of *ordered pair* is also treated as primitive, but we will indicate how it could be defined using the set concept. The techniques which are used to achieve this end differ only in technical detail from the techniques used in a more familiar treatment. We believe that the constructions given here are if anything more natural than the traditional constructions; that is why this book was written.

What is a set? This is not a question that we will answer directly. "Set" is an undefined notion. The reader will have to count on her intuitive notion of what a "set" or "collection" is at the start. She may find that some of the properties of the set concept that we develop here will be somewhat unfamiliar.

A set or collection has members. The basic relationship between objects in our theory, written $a \in b$, can be translated "a is an element of b" or "a is a member of b" or "a belongs to b". One way in which it should *not* be translated is "a is a part of b". The subset or inclusion relation, to be introduced later, is a better translation for the intuitive relation of

part to whole. This is the major intuitive pitfall with set theory; a set does not have its elements as parts. An important difference between the membership relation and the relation of part to whole is that the latter is transitive but the former is not: if A is a part of B and B is a part of C, then A is a part of C; but it is not the case that given $A \in B$ and $B \in C$, $A \in C$ follows. We will see counterexamples to this shortly.

A set is exactly determined by its members. This can be summarized in our first axiom:

Axiom of Extensionality. *If A and B are sets, and for each x, x is an element of A if and only if x is an element of B, then $A = B$.*

The Axiom of Extensionality can be paraphrased in more colloquial English: "Sets with the same elements are the same".

Not all objects in our universe are sets. Objects which are not sets are called "atoms". You can think of ordinary physical objects, for instance, as being atoms. We certainly do not think of them as being sets! Atoms have no elements, since they are not sets:

Axiom of Atoms. *If x is an atom, then for all y, $y \notin x$ (read "y is not an element of x").*

An advantage of the presence of atoms is that we can suppose that the objects of any theory (or the objects of the usual physical universe) are available for discussion, even if we do not know how to describe them as sets or do not believe that they are sets. It turns out that our axioms will allow us to prove the existence of atoms, which is a rather surprising result!

Observe that the converse is not necessarily true: if an object has no elements, we cannot conclude that it is an atom. It is possible for there to be sets with no elements. What we can prove using the Axiom of Extensionality is that there is no more than one set with no elements:

Theorem. *If A and B are sets, and for all x, x is not an element of A and x is not an element of B, then $A = B$.*

PROOF. — For all x, if x is an element of A, it is an element of B and

vice versa (a false statement implies anything...). By Extensionality, A and B must be equal. □

We will state axioms shortly which will guarantee the existence of a set with no elements. This unique set is called *the empty set* and denoted by the symbol $\{\ \}$.

A natural way to specify a set is to take the collection of all objects with some property. For example, we could consider the collection of all prime numbers greater than 17. This is a technique of specifying sets which has pitfalls (we will see some of these in later chapters), and we will not use it as our technique of choice for building sets at first, but it does inspire a very useful notation for sets. The set of all prime numbers greater than 17 can be written $\{x \mid x \text{ is a prime number and } x \text{ is greater than } 17\}$; in general, the collection of all objects with property ϕ can be written $\{x \mid x \text{ has property } \phi\}$, which is read "The set (or class, or collection) of all x such that x has property ϕ". The choice of the letter representing elements of the proposed set is indifferent: the expressions $\{x \mid x \text{ has property } \phi\}$, $\{y \mid y \text{ has property } \phi\}$ and $\{A \mid A \text{ has property } \phi\}$ are exactly the same (as long as appropriate substitutions of y or A for x are made in the description of the property ϕ: $\{x \mid x \text{ is a prime number and } x \text{ is greater than } 17\}$ is equivalent to $\{A \mid A \text{ is a prime number and } A \text{ is greater than } 17\}$).

It will turn out that not every set of the form $\{x \mid x \text{ has property } \phi\}$ actually exists; there are properties which *cannot* define sets. We will introduce such a property eventually. As a consequence, we will not use the general technique of collecting all objects with a given property to build sets; we will use a number of basic (and natural) constructions which experience indicates are safe to build sets. We will eventually use properties of these basic constructions to prove a theorem showing that a large class of properties (the "stratified" properties) do in fact define sets, and after the proof of this theorem we will use sets of the form $\{x \mid x \text{ has property } \phi\}$ much more freely. Some of the basic constructions which we use are incorrectly considered to be "dangerous" because they lead to problems in the context of the usual set theory; for instance, this set theory has a universal set. Illustrating the fact that such constructions are not dangerous is one of the aims of this work.

If the reader feels that arbitrary collections of objects of our theory *must*

exist in some sense, he can understand the sentence "the set $\{x \mid x$ has property $\phi\}$ does not exist" as meaning not that there is no collection of all objects x of our theory such that ϕ, but that this collection cannot be regarded as a set in our theory; such collections which are not sets will sometimes be discussed (we call them "proper classes").

Cahiers du Centre de logique
Volume 10

Chapter 3

Boolean Operations on Sets

We will now introduce the first of the basic set constructions promised in the last chapter. Two particular sets which we might expect to find are $\{\ \}$, the empty set mentioned in the previous chapter, and V, the universe, the set which contains everything. Notation for these sets might be $\{x \mid x = x\}$ or $\{x \mid \text{True}\}$ for V, and $\{x \mid x \neq x\}$ or $\{x \mid \text{False}\}$ for $\{\ \}$. We state an axiom:

Axiom of the Universal Set. *$\{x \mid x = x\}$, also called V, exists.*

We will not need a special axiom for $\{\ \}$, for reasons which will become evident shortly. Be warned: this is our first departure from the usual set theory ZFC; in ZFC, the Axiom of the Universal Set is *false*; the universe does not exist. The reasons for this will be made clear in a later chapter.

Given a set A, an obvious set which might leap to mind would be the set of all things not in A. If A were the set of beautiful things, the complementary set which would come to mind is the set of things which are not beautiful. We assert an axiom providing the existence of such sets (the symbol $\notin$ is read "is not an element of"):

Axiom of Complements. *For each set A, the set $A^c = \{x \mid x \notin A\}$, called the complement of A, exists.*

Two things to note here: the Axiom of the Universal Set and the Axiom of Complements together imply that the empty set $\{\ \}$ exists: the empty set can be constructed as the complement of the universe, i.e., $\{\ \} = V^c$; the Axiom of Complements is uniformly false in ZFC, where *no* set has a complement. We believe that the presence of the universe and of complements together actually give the system of set theory presented here a more intuitive flavour than the usual set theory.

When we have two sets A and B, a natural set to consider is the set which contains all the elements of both A and B. If A is the set of green objects and B is the set of red objects, the set we are interested is the set of all objects which are either green or red. In this case, the two sets do not overlap; if A were the set of college professors and B were the set of absent-minded people, the set we are interested in would be the set of people who are either college professors or absent-minded or both. We assert an axiom to provide for this kind of set:

Axiom of (Boolean) Unions. *If A and B are sets, the set*

$$A \cup B = \{x \mid x \in A \text{ or } x \in B \text{ or both}\},$$

called the (Boolean) union *of A and B, exists.*

These are the only new primitive constructions and axioms required for most of this chapter. There are several derived set operations and relations among sets which we will need to define.

Given two sets A and B, another set which we might want to consider is the set of objects in the "overlap" between A and B; if A were the set of college professors and B were the set of absent-minded persons, the set of interest would be the set of absent-minded college professors. We can *prove* that such sets exist for every A and B:

Theorem. *For each set A and B, the set*

$$A \cap B = \{x \mid x \in A \text{ and } x \in B\},$$

called the (Boolean) intersection *of A and B, exists.*

PROOF. — $A \cap B = (A^c \cup B^c)^c$. An object is an element of $(A^c \cup B^c)^c$ exactly if it is *not* an element of $A^c \cup B^c$. An object is *not* an element of

$A^c \cup B^c$ exactly if it is not the case either that it is an element of A^c or that it is an element of B^c; equivalently, if it is not the case either that it is not an element of A or that it is not an element of B. But this is true exactly if it is an element of A and an element of B. □

Note the possible confusion caused by two different uses of the word "and" in English: an object belongs to $A \cap B$ if and only if it belongs to A *and* belongs to B, but the set made up of the elements of A *and* the elements of B is $A \cup B$. We simply have to be careful: these two different uses of "and" are logically totally different from one another.

In some contexts, the "working universe" is not the whole of V. This causes problems when complements are to be taken. For instance, if we are working in arithmetic we think of the complement of the set of even numbers as being the set of odd numbers, rather than the set of all odd numbers and non-numbers. It is convenient to define a new concept, introduced in the following:

Theorem. *For each pair of sets A, B, the set*

$$B - A = \{x \mid x \in B \text{ and } x \notin A\},$$

called the relative complement of A with respect to B, *exists.*

Proof. — $B - A = B \cap A^c$. □

A final operation on sets (only occasionally used):

Definition. *For A, B sets, we define the* symmetric difference $A \Delta B$ *as* $(B - A) \cup (A - B)$.

A very important relation between sets is the "subset" relation, or "inclusion". We say that "A is a subset of B" or that "A is included in B" if every element of A is also an element of B. We give a

Definition. *$A \subseteq B$ (A is a* subset *of B, A is* included *in B) exactly if A and B are sets and for every x it is the case that if x is an element of A, then x is an element of B.*

$A \subset B$ (A is a proper subset *of B, A is properly included in B) exactly if A is a subset of B and A is not equal to B.*

$A \supseteq B$ and $A \supset B$ are the converse relations, read "A is a superset of B or A contains B", "A is a proper superset of B or A properly contains B".

Theorem. *$\{\} \subseteq A$ for every set A.*

PROOF. — If x is an element of $\{\}$... (anything, including "x is an element of A", follows). □

We can see that inclusion is not equivalent to membership; $\{\} \subseteq \{\}$ by the Theorem, but it is not the case that $\{\}$ is an element of $\{\}$! The relation between inclusion and the operations above is expressed in the following:

Theorem. *For all sets A, B, $A \subseteq B$ exactly if $A \cup B = B$.*

A formal proof is left to the reader. It should be obvious that adding all the elements of a subset of B to B will not change B!

We define another important relationship between sets:

Definition. *Sets A and B are said to be* disjoint *exactly if for every x, either x is not an element of A or x is not an element of B or both (i.e., no x belongs to both A and B).*

Sets are disjoint if they do not overlap; this insight is equivalent to the following result, which we state without proof:

Theorem. *Sets A, B are disjoint exactly if $A \cap B = \{\}$.*

There is no conventional symbol for the relationship of disjointness, so the result of this Theorem is used to represent disjointness symbolically. It is not correct to say that disjoint sets A and B have "no intersection"; they do have an intersection, namely the empty set, but this intersection has no elements. The notion of disjointness extends to more than two sets: for example, sets A, B, and C are said to be disjoint if there is no object which belongs to more than one of them. The condition $A \cap B \cap C = \{\}$ is weaker; this merely asserts that there is no object which belongs to all three sets, and allows for the possibility that an object may belong to two of them. For this reason, collections of disjoint sets with more than two

elements are often referred to as *pairwise* disjoint sets; the real condition in the case above is that $A \cap B = A \cap C = B \cap C = \{\ \}$.

The operations and relations defined in this chapter comprise an interpretation of the field known as "Boolean algebra"; we will distinguish the operations defined here as "Boolean operations". We list some useful facts about them which are also axioms or theorems of Boolean algebra, expressed in our notation:

commutative laws: $A \cap B = B \cap A$
$A \cup B = B \cup A$

associative laws: $(A \cap B) \cap C = A \cap (B \cap C)$
$(A \cup B) \cup C = A \cup (B \cup C)$

distributive laws: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

identity laws: $A \cup \{\ \} = A$
$A \cap V = A$

idempotence laws: $A \cup A = A$
$A \cap A = A$

cancellation laws: $A \cup V = V$
$A \cap \{\ \} = \{\ \}$

De Morgan's laws: $(A \cap B)^c = A^c \cup B^c$
$(A \cup B)^c = A^c \cap B^c$

double complement law: $(A^c)^c = A$

other complement laws: $A^c \cap A = \{\ \}$
$A^c \cup A = V$
$V^c = \{\ \}$
$\{\ \}^c = V$

inclusion principles: $A \subseteq B$ exactly if $A = A \cap B$
also exactly if $B = B \cup A$.

These laws should remind one to some extent of the ordinary rules of algebra for addition and multiplication; but notice the exact symmetry between the two operations, and that nothing in the usual algebra corresponds to the complement operation (in particular, not the additive inverse!)

The operations of union and intersection as we have defined them allow us to combine finitely many sets, or to find the common part of finitely many sets, by repeated application of the binary Boolean operations. We will sometimes want to define unions and intersections of collections of sets which are not necessarily finite; this is provided by the following

Axiom of Set Union. *If A is a set all of whose elements are sets, the set $\bigcup[A] = \{x \mid$ for some B, $x \in B$ and $B \in A\}$, called the (set)* union of *A, exists.*

The construction of set intersections will require the assistance of axioms not yet provided. Since we chose Boolean union as a primitive and used it to define Boolean intersection, we will follow precedent and patiently wait until we have the required additional machinery needed to construct set intersections.

Exercises

(a) The smallest Boolean algebra consists of the sets $\{\ \}$ and V. Develop an interpretation of this Boolean algebra in mod 2 arithmetic, giving definitions of the operations of Boolean algebra in terms of the operations of the arithmetic. What Boolean algebra operations correspond to your arithmetic operations? Explain why there are two different nontrivial interpretations.

(b) Verify that the symmetric difference operation is commutative and associative, and that intersection distributes over it.

(c) Look up "Venn diagrams" in another source and verify some of the axioms given at the end of the chapter or the results of the previous exercise using Venn diagrams.