Set Theory

Basic concepts

In some respects, basic set theory is just a rewriting of basic logic. Looking at the list of "Set Identities" on p. 89 of Rosen, one has to have the feeling, "I've seen this before." It is remarkably similar to the list of "Logical Equivalences" on p. 24.

When we replace sentence variables (such as p) by set membership statements (such as "$x \in A$"), some of the logical connectives naturally give rise to *set operations*:

$x \in A \vee x \in B$	$\iff$	$x \in A \cup B$
$x \in A \wedge x \in B$	$\iff$	$x \in A \cap B$
$x \in A \oplus x \in B$	$\iff$	$x \in A \oplus B$ or $A \,\Delta\, B$ (*symmetric difference*)
$\neg[x \in A]$ or $x \notin A$	$\iff$	$x \in \overline{A}$
$x \in B \wedge \neg[x \in A]$	$\iff$	$x \in B - A$
F or 0	$\iff$	$x \in \emptyset$ (always false)
T or 1	$\iff$	$x \in \mathcal{U}$ (always true)

Other logical connectives, however, more naturally give rise to *set relations*:

$\forall x\, [x \in A \;\rightarrow\; x \in B]$	$\iff$	$A \subseteq B$
$\forall x\, [x \in A \;\longleftrightarrow\; x \in B]$	$\iff$	$A = B$

Remark: Some authors write $\subset$ instead of $\subseteq$ for the basic relation of set inclusion, but most now reserve $\subset$ for *proper inclusion*:

$$A \subset B \quad \text{means} \quad A \subseteq B \wedge A \neq B.$$

Note that, given $A \subseteq B$, the extra condition $A \neq B$ can be expressed as $\exists x\, [x \in B \wedge x \notin A]$; but in general circumstances, $A \neq B$ could also be true because of existence of an x such that $x \in A \wedge x \notin B$. In general,

$$A \neq B \;\longleftrightarrow\; A \oplus B \neq \emptyset.$$

Venn diagrams can be very helpful in elementary reasoning about sets. Beware, however, that if more than 3 sets are involved (say N sets), it is hard to draw a Venn diagram that exhibits a region (and only one) for each of the 2^N possibilities of set membership. Note also that the empty set, $\emptyset$, is not located "somewhere" in Venn land. Instead, $\emptyset \subseteq A$ for all A. (Incidentally, $\emptyset$ is not a Greek letter ϕ. It was originally modeled on a Scandinavian vowel letter.)

Proof that $\emptyset \subseteq A$: Consider an arbitrary x. Then $x \in \emptyset$ is false, by definition. Therefore, $x \in \emptyset \;\rightarrow\; x \in A$. Since x was arbitrary, we have proved (UG)

$$\forall x\,[x \in \emptyset \;\rightarrow\; x \in A]\,;$$

but that is equivalent to the theorem, by definition of $\subseteq$.

Remark: It is important to understand that two sets are *equal* if they have the same members, even if conceptually the sets are specified by completely unrelated definitions. Thus "the set of Presidents whose fathers were also President" is equal to "the set of Presidents who had to be chosen by either the House of Representatives or the Supreme Court instead of by a clear-cut vote of the Electoral College", although there is no causal connection between the two conditions. This (so-called *extensional*) concept of sethood is similar to two other things we have seen earlier: (1) For logical propositions, $p \;\longleftrightarrow\; q$ simply means that p and q have the same truth value, so we arrive at seemingly strange statements like "China is in Asia if and only if $2+2=4$." (2) If a and b are names of the same individual element, then $a=b$ even if a does not mean the same thing as b, and even if some people think that they are two different things; the classic example is "Morning Star" and "Evening Star".

Ways to prove set identities

[Notes to be written later.]

Cartesian product; Power set; "Set builder" axiom

[Notes to be written later.]

Russell's paradox; Noncomputable functions

So far we have had two ways of dealing symbolically with sentences like "John is a student." First we introduced *predicates*, or propositional symbols with arguments, like this:

> Let $s(x)$ mean that x is a student. Then $s(\text{John})$.

Then we started working with sets:

> Let S be the set of all students. Then $\text{John} \in S$.

You may have wondered why we need both of these logical structures. Doesn't every predicate or property correspond to a set, and vice versa?

The problem stems from the fact that when we think of sets as objects in their own right, we expect to be able to write quantifiers over sets and "set-builder" definitions with variables for sets: things like

$$\exists x\,[\text{John} \in x \wedge x \subset S] \qquad \text{and} \qquad A = \{x \colon x \subset S \wedge \text{John} \in x\}.$$

In the very early 20th century it was discovered that overly expansive thinking about sets could lead to inconsistencies. The central observation is called **Russell's Paradox**:

Define $R = \{x \colon x \notin x\}$. Is $R \in R$?

On the one hand, if $R \in R$, then (by definition of R) $R \notin R$. So by *reductio ad absurdum* $R \notin R$. On the other hand, if $R \notin R$, then (by definition of R) $R \in R$. So we have proved

$$R \in R \wedge R \notin R,$$

a statement of the form $p \wedge \neg p$, in violation of one of the most elementary laws of logic. (Since $p \wedge \neg p \;\rightarrow\; q$ is a tautology, we could now prove any statement q whatsoever, which is absurd.) Something is seriously wrong.

Historically there were two approaches taken to defuse this situation. The first one, propounded by Russell himself, was to say that there is something wrong with the whole question of whether a set is a member of itself; "$x \notin x$" should be rejected as a meaningless grammatical error that can't be either true or false. If our original universe of discourse is the integers, then all our sets are (originally) sets of integers, not sets of sets. "The set of all integers such that ... " is not an integer, so it can't be a member of a set of integers. However, it *is* a new type of object, and there should be no problem in talking about sets containing that kind of object — sets of sets. But those sets contain only sets of integers, not sets of sets. Sets of sets can be members only of a new type of object, sets of sets of sets. Thus the universe grows in steps, and one is not allowed to mix integers and sets in the same set, or to mix sets (of integers) with sets of sets in the same set. In this way the paradox never arises. This hierarchical theory of sets was called the *theory of types*. Unfortunately, it seems to be very complicated and needlessly restrictive.

The second approach (which eventually won out) was to pin the blame on the assumption that every set-builder construction one can write down actually corresponds to a set. In this doctrine, one cannot assume existence of a set unless an axiom justifies it. In particular, set-building is not legitimate except to carve out a piece of a larger set that we already know exists. Set-builder definitions are required to be of the form

$$A = \{x \in S \colon p(x)\}$$

(where $p(x)$ is some open sentence). S is allowed to be left tacit as a universe of discourse, provided that it is a legitimate set. Russell's paradox is now avoided, because $R = \{x \in S \colon x \notin x\}$ simply turns out to be not a member of S, so $R \notin R$ and "$R \in R$" is false — there's no contradiction. In order to have any sets at all, we must have some other axioms that (cautiously) assume the existence of some sets. The two most powerful axioms are that

1. If A is a set, then the collection of all subsets of A is a set (the power set of A).
2. If A and B are sets, then their union $A \cup B$ is a set.

Then, starting from a set of elementary objects like the integers (or even from the empty set!) the axioms applied repeatedly create increasingly big and complicated sets. This system is called the *Zermelo–Frankel set theory*. Quine showed that it can be obtained from the theory of types by combining each type with all the types below it in the hierarchy; thus there's nothing illegal about writing $\{\text{John}, \{\text{John}\}\}$, but the paradox is still avoided because the set-builder $\{x \in T : x \notin x\}$ can only be applied to one particular type-level T at a time.

Note that in both theories there is no ultimate universe $\mathcal{U}$ comprising "everything in the world". There is no *set of all sets*.

Russell's paradox is not an isolated curiosity. It is typical of a large class of arguments that show that certain "universal" objects can't exist. A famous example from computer science is **the unsolvability of the halting problem**:

Let the universe of discourse be computer files, and let $P(Q)$ stand for the result of applying program P to data file Q. Q may itself be a program. (For definiteness one can think of "programs" as either Pascal source files or Unix binaries, for example.) For example, if P is a program that counts the number of bytes in a file, then the action $P(P)$ is to print the number of bytes in the byte-count program file.

Now suppose that there is a program P_0 that solves the halting problem. This means that the action $P(Q)$ is to print whichever of these statements is correct:

1. Q is not a valid program file.
2. $Q(Q)$ executes and eventually halts.
3. $Q(Q)$ executes and runs forever (for example, by going into an infinite loop).

Note that $P_0(Q)$ could be the special case $P_1(Q,Q)$ of a program $P_1(P,Q)$ for diagnosing $P(Q)$ in this way. (The latter is a perfectly reasonable question to ask and doesn't itself involve the peculiar notion of applying a program to itself.)

Now let P_2 be P_0 modified in this way: Instead of printing "$Q(Q)$ executes and eventually halts", it deliberately goes into an infinite loop. Nobody would want to write such a program, but *if P_1 exists, then someone COULD write P_2*.

The killer question is: What happens when you issue the command $P_2(P_2)$? Well, P_2 is a valid program file, so the first alternative is out. If $P_2(P_2)$ runs forever, then $P_2(P_2)$ prints "$P_2(P_2)$ runs forever" and then stops. That's a contradiction, so the second alternative can't happen. But if $P_2(P_2)$ ever halts, then P_2 is designed so that $P_2(P_2)$ goes into a loop. That's also a contradiction!

Conclusion: The program P_2 can't exist in the first place. Therefore, P_0 can't exist. And therefore, P_1 can't exist! It is impossible to write a computer program that solves the halting problem for all program and data pairs, (P,Q). More generally, *not all meaningful functions are computable, even in principle*.