II

Foundations

Because, finally, he said:
"This is really great stuff!
"And I guess the old alphabet
"ISN'T enough!"
——DR. SEUSS†

§1 SETS, CLASSES, AND CONGLOMERATES

In the introductory chapter, we have seen that in category theory we are confronted with extremely large "collections" such as "all sets", "all groups", or "all topological spaces". The reader with some set-theoretical background knows that these entities cannot be regarded as sets. For instance, if $\mathscr{U}$ were the set of all sets, then the subset $A = \{x \mid x \in \mathscr{U} \text{ and } x \notin x\}$ of $\mathscr{U}$ would have the property that $A \in A$ if and only if $A \notin A$ (Russell's paradox). A mathematician working, for example, in group theory or topology usually isn't (and needn't be) bothered with these set-theoretical difficulties. However, it is essential that those who work in category theory be able to deal with "collections" like those mentioned above. It is also advantageous that certain "naturally arising" categorical constructions not be outlawed simply because of foundational considerations. There have been several attempts to find a suitable foundation for category theory. Brief sketches of some of the currently used foundations appear in the appendix. Each of them has its advantages and disadvantages and it remains an open problem to design a foundation that is free of serious disadvantages. What we desire at the moment is just a foundation that is sufficiently flexible so as not to unduly inhibit our categorical inquiry and one that we can be reasonably sure will not lead to paradoxes. Below we provide a brief outline of the features we require of such a foundation. That this foundation can indeed be realized is shown in the appendix. Thus, the reader with a good background in axiomatic set theory should next read the appendix. The following account is for those who, by necessity, must approach the subject on a more naïve level.

† From *On Beyond Zebra*, © 1955 by Random House, Inc.

1.1 SETS

Sets can be thought of as the usual sets of intuitive set theory (or of Zermelo-Fraenkel or Gödel-Bernays-von Neumann set theory). In particular, we require that the following constructions can be performed with sets.

(1) For each set X and each "property" P, we can form the *set* $\{x \mid P(x)$ and $x \in X\}$ of all members of X having property P. (If $P(x)$ is the property "$x \neq x$", this yields the **empty set**, denoted by $\varnothing$.)

(2) For each set X, we can form the *set* $\mathscr{P}(X)$ of all subsets of X.

(3) For any sets X and Y, we can form the following *sets*:

- (i) The set $\{X, Y\}$ whose members are exactly X and Y.
- (ii) The **(ordered)**† **pair** $(X, Y) = \{\{X\}, \{X, Y\}\}$ with first coordinate X and second coordinate Y.
- (iii) The **union set** $X \cup Y$.
- (iv) The **intersection set** $X \cap Y$.
- (v) The **complement set** $X - Y$.
- (vi) The **cartesian product set** $X \times Y$.
- (vii) The *set* Y^X of all functions from X to Y.††

(4) For any set I and any family $(X_i)_{i \in I}$††† of sets, we can form the following *sets*:

- (i) The **image set** $\{X_i \mid i \in I\}$ of the indexing function.
- (ii) The **union set** $\bigcup_{i \in I} X_i$.
- (iii) The **intersection set** $\bigcap_{i \in I} X_i$, if $I \neq \varnothing$.
- (iv) The **cartesian product set** $\prod_{i \in I} X_i$.
- (v) The **disjoint union set** $\coprod_{i \in I} X_i \left(= \bigcup_{i \in I} (X_i \times \{i\})\right)$.

(5) The usual "collections":

N of all natural numbers
Z of all integers
Q of all rational numbers
R of all real numbers

are sets. So is each ordinal number and each cardinal number.

† From now on, the word "pair" will mean "ordered pair", "triple" will mean "ordered triple", and so forth.

†† A **function from** X **to** Y is defined to be a triple (X, f, Y) where $f \subset X \times Y$ and for each $x \in X$ there is exactly one $y \in Y$ such that $(x, y) \in f$. The uniquely specified y is usually denoted by $y = f(x)$ or by $x \mapsto y$. Sometimes the triple (X, f, Y) is denoted by $f\colon X \to Y$ or occasionally (and inaccurately) by f alone. A function is called **injective** if for each $y \in Y$ there is at most one $x \in X$ such that $f(x) = y$. It is called **surjective** if for each $y \in Y$ there is at least one $x \in X$ such that $f(x) = y$; and it is called **bijective** if it is both injective and surjective. If $f\colon X \to Y$, $A \subset X$ and $B \subset Y$, then $f[A] = \{y \mid$ for some $x \in A, f(x) = y\}$ and $f^{-1}[B] = \{x \mid$ for some $y \in B, f(x) = y\}$.

††† A **family** $(X_i)_{i \in I}$ (sometimes denoted by $(X_i)_I$ or even by (X_i)) is actually a function f with domain I such that for each $i \in I$, $f(i) = X_i$.

From the above considerations, we see, for example, that each group, each topological space, and each lattice is a set. However, by means of the above constructions, we cannot form "the *set* of all sets", or "the *set* of all groups".

1.2 CLASSES

To handle "large collections" such as "all sets", we require that:

(1) For each "property" P we can form a "collection" whose members are exactly those *sets* that have property P. We call this the **class of all sets with property** P and denote it by $\{x \mid x \text{ is a set and } P(x)\}$ (or more briefly, by $\{x \mid P(x)\}$). For example, we can form the "*class* of all sets", the "*class* of all ordinal numbers", and the "*class* of all groups". Obviously, classes are precisely the "subcollections" of the class $\mathcal{U}$ of all sets. We will call $\mathcal{U}$ the "**universe**".

(2) For convenience of expression, we also wish to regard sets as special classes. (This can easily be achieved by adding to the above requirements for sets, the additional requirement that each member of a set be a set.)† Those classes which are not sets are called **proper classes.** Often sets are referred to as **small classes** and proper classes are called **large classes.** This distinction between "small" and "large" will turn out to be essential in many categorical investigations. Notice that the universe $\mathcal{U}$ of all sets is a proper class and that Russell's paradox now translates into the harmless statement that the class of all sets that are not members of themselves, is not a set but is a proper class.

(3) Given classes A and B, we can form the following *classes*:

(i) The **union class** $A \cup B = \{x \mid x \in A \text{ or } x \in B\}$.
(ii) The **intersection class** $A \cap B = \{x \mid x \in A \text{ and } x \in B\}$.
(iii) The **complement class** $A - B = \{x \mid x \in A \text{ and } x \notin B\}$.
(iv) The **cartesian product class** $A \times B = \{(a, b) \mid a \in A \text{ and } b \in B\}$.
(v) The **disjoint union class** $A \uplus B = A \times \{\varnothing\} \cup B \times \{\{\varnothing\}\}$.

Hence, we can define functions between classes, equivalence relations on them, and so forth.

(4) We will need the "class form" of the choice axiom; namely:

Axiom of Choice for Classes.

Every equivalence relation on a class has a system of representatives.
Or, equivalently:

There is a function $C: \mathcal{U} \to \mathcal{U}$ such that for each non-empty set X, $C(X) \in X$.

Notice that this implies the usual "set form" of the Axiom of Choice.

1.3 CONGLOMERATES

If A is a proper class, then there exists no class that has A as a member (since every member of a class must be a set). However, we will occasionally need to consider "collections" of classes. For this reason, we introduce the broader concept of "conglomerate". Roughly speaking, conglomerates are

† This also implies that if the pair (x, y) is a set, then so are x and y.

"collections" having classes or conglomerates as members. In particular, we require that:

(1) Every class is a conglomerate.

(2) Conglomerates are closed under the usual set-theoretic constructions outlined above (1.1); i.e., they are closed under the formation of pairs, unions, products, etc.

Thus we can effectively treat conglomerates in the same manner that we treated sets; we can construct functions between them, equivalence relations on them, etc. There is the temptation, of course, to form the "cartel" of all conglomerates. However, assuming that our primary interest lies with "usual categories", such as the category of all sets or the category of all topological spaces, we will not need to consider such an entity.