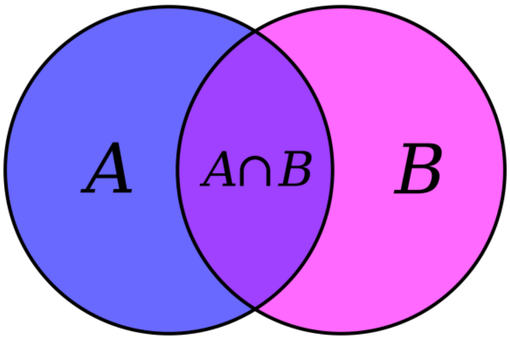
Critical Reasoning 18 - The Logic of Set Theory



Set theory and logic are inextricably linked. Indeed parts of set theory and logic can be defined in terms of the other. However due to its power to define nearly all mathematical objects, set theory has remained largely the province of mathematicians and logic that of philosophers, linguists, computer scientists and so on. Since its inception in 1874 by Georg Cantor in his seminal paper "On a Characteristic Property of All Real Algebraic Numbers", modern set theory has thrown up multiple philosophical problems and insights that are worthy of serious study by mathematicians and philosophers alike. As before we shall be following the outline of the topic in Copi's *Symbolic Logic* (1979, Ch. 8)

The Algebra of Classes

According to Copi, "The notion of a *class* [or ***set***] is too basic to be defined in terms of more fundamental concepts". Although mentioning synonyms such as: collection, aggregate, totality, set, and so on, does make the notion of a class more intuitive, they cannot define the term without circularity. Therefore the terms 'class' or 'set' are used as an undefined or primitive terms when discussed axiomatically. In this section, statements about classes or sets are conveniently expressed as equations and inequalities. To do so, three **operations on classes** are regarded as fundamental, which also allows the relation of class inclusion to be defined. (p. 170)

Using lower case Greek letters α, β, γ … to symbolise classes, other classes can be defined as:

$\alpha \cup \beta$ |the **sum** or **union** of α and β is the class to which α or β belong

$\alpha \cap \beta$ |the **product** or **intersection** of α and β, sometimes simply written as '$\alpha\beta$', is the class to which both α and β belong

$\overline{\alpha}$ |the **complement**, also written as '$-\alpha$', is the class of objects that do not belong to α

Equality of sets is represented using the '$=$' sign in its normal sense of equation such that:

$\alpha = \beta$|all members of α (if any) are members of β and, all members of β (if any) are members of α. (p. 170 -171)

The sum, product and complement classes exhibit the following properties that can be expressed as equations. Firstly, the sum and product of two classes are **commutative**, *i.e.* the order of the classes on which the operations are performed does not change the result, thus

$$\alpha \cup \beta = \beta \cup \alpha$$

$$\alpha \cap \beta = \beta \cap \alpha$$

They are also **associative**, *i.e.* rearranging the parentheses in an expression will not change the result, thus

$$(\alpha \cup \beta) \cup \gamma = \alpha \cup (\beta \cup \gamma)$$

$$(\alpha \cap \beta) \cap \gamma = \alpha \cap (\beta \cap \gamma)$$

And they are **distributive**, *i.e.* any object that belongs to either α or to both β and γ must belong to α or β and also to either α or γ, and conversley. Similarly, any object that belongs to α and to either β or γ must also belong to both α and β or to both α and γ, thus

$$\alpha \cup (\beta \cap \gamma) = (\alpha \cup \beta) \cap (\alpha \cup \gamma)$$

$$\alpha \cap (\beta \cup \gamma) = (\alpha \cap \beta) \cup (\alpha \cap \gamma)$$

The attentive reader may have noticed that there is a correspondence between the union operator ∪ to the logical operator OR and that there is a correspondence between the intersection operator ∩ and the logical operator AND. Similarly there is a correspondence between the complement operator – and the logical operator NOT. We shall come across further correspondences so that ultimately, almost all statements that can be expressed in the language of logic have a corresponding expression in the language of set theory; however Copi does not dwell on this.

Idempotence is the property of certain operations when applied multiple times not to change the result beyond the initial application. Two laws which express the idempotence of ∪ and ∩ resemble the principle of tautology for statements introduced in Critical Reasoning 07, thus

$$\alpha = \alpha \cup \alpha$$

$$\alpha = \alpha \cap \alpha$$

Following on from this is the **principle of absorption**, which resembles that of addition for statements, also introduced in Critical Reasoning 07, thus

$$\alpha = \alpha \cup (\alpha \cap \beta)$$

$$\alpha = \alpha \cap (\alpha \cup \beta)$$

Next Copi observes that, “since any object belongs to a given class if and only if it does not belong to the class of all objects that do not belong to the given class, the **complement of a complement** of a class is the class itself. We thus have a sort of double negative rule for complementation,” (p. 171) thus

$$\alpha = \overline{\overline{\alpha}}$$

Two versions of **De Morgan’s Theorem** are also true for sets because an object that does not belong to the sum or union of two classes cannot belong to either of them. Similarly, an object that does not belong to the product or intersection of two classes must belong to the complement of at least one of them, thus

$$\overline{\alpha \cup \beta} = \overline{\alpha} \cap \overline{\beta}$$

$$\overline{\alpha \cap \beta} = \overline{\alpha} \cup \overline{\beta} \qquad \text{(p. 172)}$$

Two special classes that need to be considered are the **empty set**, symbolised as '$\emptyset$', which has no members and the **universal set**, symbolised as 'U', to which all objects belong, including itself. It should be intuitively easy to see that the empty set is the complement of the universal set and *vice versa*, thus

$$\emptyset = \overline{U}$$

$$U = \overline{\emptyset}$$

Two immediate consequences of which are that the sum or union of any class and its complement is the universal set and that the intersection or product of any class and its complement is the empty set, thus

$$\alpha \cup \overline{\alpha} = U$$

$$\alpha \cap \overline{\alpha} = \emptyset$$

Copi mentions the following additional immediate consequences:

$$\alpha \cup \emptyset = \alpha,\ \ \alpha \cap U = \alpha,\ \ \alpha \cap \emptyset = \emptyset;\ \text{ and }\ \alpha \cup U = U$$

In the same way that a number or a variable can be designated by an expression with one or more terms, so a given class can be designated by an infinite number of **class-expressions** using equalities such as those above. For example, the class 'α' can be designated by,

'$\alpha \cap (\beta \cup \overline{\beta})$' (since $\beta \cup \overline{\beta} = U$ and $\alpha \cap U = a$).

The same 'α' can also be designated by,

$[\alpha \cap (\ \beta \cup \overline{\beta})] \cap (\gamma \cup \overline{\gamma})$ and so on.

By what Copi calls the '**Law of Expansion**' we can introduce any class symbol into a class expression in such a way that that the expanded class expression designates the same class. There is no sleight of hand in expanding expressions in this way. So long as the expanded class expression designates the same class as the original, we are not deriving something from nothing. Compare the application of logically equivalent rules of replacement introduced in Critical Reasoning 07.

Consider the class: $\alpha \cap (\beta \cup \overline{\beta})$. By the principle of distribution this is equal to $(\alpha \cap \beta) \cup (\alpha \cap \overline{\beta})$. To describe the form of classes of the latter sort, Copi introduces the phrase '**simple class term**' to refer to class symbols 'α', 'β', 'γ'… in contrast to other class expressions such as sums and products. "Now we can describe the expression '$(\alpha \cap \beta) \cup (\alpha \cap \overline{\beta})$' as a sum of distinct products, such that in each product only simple class terms or their complements appear and such that any simple class term which appears anywhere in the entire expression appears exactly once in every product." (p. 172)

Using just the equations presented above, any class expression can be transformed into another, perhaps simpler expression, that designates the same class. *E.g.*

$$\overline{\alpha \cap (\overline{\alpha} \cup \beta)}$$

$= \overline{\alpha} \cup (\overline{\overline{\alpha} \cup \beta})$ by De Morgan's Theorem

$= \overline{\alpha} \cup (\overline{\overline{\alpha}} \cap \overline{\beta})$ again by De Morgan's Theorem

$= \overline{\alpha} \cup (\alpha \cap \overline{\beta})$ by double negation

$= \overline{\alpha} \cap (\beta \cup \overline{\beta}) \cup (\alpha \cap \overline{\beta})$ by expansion

$= [(\overline{\alpha} \cap \beta) \cup (\overline{\alpha} \cap \overline{\beta})] \cup (\alpha \cap \overline{\beta})$ by distribution

$= (\overline{\alpha} \cap \beta) \cup (\overline{\alpha} \cap \overline{\beta}) \cup (\alpha \cap \overline{\beta})$ by association

Which, if we replace each product by one of the next three Greek lower case letters

$= \delta \cup \varepsilon \cup \zeta$ which is simply equal to $\alpha \cup \beta \cup \gamma$.

No doubt you have head of expressions like, "The are only two sorts of people: dog-lovers and non-dog-lovers" or "… people who've got rhythm and those who don't". While such statements are all false because the real world is not so neatly dichotomous, they do capture the way any class or set divides the universe into two mutually exclusive and jointly exhaustive subclasses or subsets. Thus for any class α,

$U = \alpha \cup \overline{\alpha}$ and $\alpha \cap \overline{\alpha} = \emptyset$

Any two classes, on the other hand, will divide the universe into four subclasses that are exclusive and exhaustive. Thus for any two classes α and β,

$$U = (\alpha \cap \beta) \cup (\alpha \cap \overline{\beta}) \cup (\overline{\alpha} \cap \beta) \cup (\overline{\alpha} \cap \overline{\beta})$$

Also, the product of any two of these four products is the empty set. If, for example the world were neatly dichotomous, then the product of the class of dog-lovers that have no rhythm with the class non-dog-lovers that have no rhythm is the empty set. Similarly, the product of the class of non-dog-lovers who have got rhythm with the class of non-dog-lovers that have no rhythm is also empty, and so on for the product of the other two combinations. In general, any n number of classes will divide the universe into 2^n subclasses which are exclusive and exhaustive. (p. 173)

Class or set notation allows for the expression of the four categorical proposition identified by Aristotle. Thus the proposition 'No α is β' (sometimes called **E** for short) asserts that α and β have no members in common, which means that their product is the empty set, *i.e.*

$$\alpha \cap \beta = \emptyset$$

The proposition that 'All α is β' (sometimes called **A**) asserts that there is nothing that belongs to α that does not belong to β, which means that the product of α and the complement of β is the empty set, *i.e.*

$$\alpha \cap \overline{\beta} = \emptyset$$

To symbolise the other two categorical propositions we need to introduce the symbol '$\neq$'that we encountered in Critical Reasoning 14. Thus if $\alpha \neq \beta$ then either α contains an object that is not in β

or β contains an object not in α. The proposition that 'Some α is β' (sometimes called **I**) which asserts that there is at least one member of α that is in β, means that the product of α and β is not empty, *i.e.*

$$\alpha \cap \beta \neq \emptyset$$

The proposition that 'Some α is not β' (sometimes called **O**) asserts that there is at least one member of α that is not in β, means that the product of α and the complement of β is not empty, *i.e.*

$$\alpha \cap \overline{\beta} = \emptyset$$

Expressed this way it is easy to see that the propositions **A** and **O** are contradictories, as are **E** and **I**. Another virtue of such class notation is that every proposition has exactly the same symbolisation as its obverse. For example 'All α is β' and 'No α is not β'are both symbolised as '$\alpha \cap \overline{\beta} = \emptyset$. Valid conversions meanwhile, such as 'Some α is β' symbolised '$\alpha \cap \beta \neq \emptyset$' and 'Some β is α' symbolised '$\beta \cap \alpha \neq \emptyset$' are equivalent by the principle of commutation since $(\alpha \cap \beta) = (\beta \cap \alpha)$. (p. 173 - 174)

When dealing with syllogisms involving categorical propositions, they can be divided into two kinds: those that contain only universal propositions (**A** and **E**) and those which contain at least one existential proposition (**I** and **O**)… All valid syllogisms of the first kind have the form,

$$\alpha \cap \overline{\beta} = \emptyset,\ \beta \cap \overline{\gamma} = \emptyset\ \therefore \alpha \cap \overline{\gamma} = \emptyset$$

The validity of this form can be derived using the algebra of classes already learned.

Since $\overline{\gamma} \cap \emptyset = \emptyset$ and $\alpha \cap \overline{\beta} = \emptyset$ is a premise, it follows that

$\overline{\gamma} \cap (\alpha \cap \overline{\beta}) = \emptyset$ which by association and commutation yields

$(\alpha \cap \overline{\gamma}) \cap \overline{\beta} = \emptyset$

Since $\alpha \cap \emptyset = \emptyset$ and $\beta \cap \overline{\gamma} = \emptyset$ is a premise, it follows that

$\alpha \cap (\beta \cap \overline{\gamma}) = \emptyset$ which by association and commutation yields

$(\alpha \cap \overline{\gamma}) \cap \beta = \emptyset$ Therefore,

$[(\alpha \cap \overline{\gamma}) \cap \beta] \cup [(\alpha \cap \overline{\gamma}) \cap \overline{\beta}] = \emptyset$ which by distribution yields

$(\alpha \cap \overline{\gamma}) \cap \left(\beta \cup \overline{\beta}\right) = \emptyset$ Now since

$\beta \cup \overline{\beta} = U$ and $(\alpha \cap \overline{\gamma}) \cap U = \alpha \cap \overline{\gamma}$ we have,

$\alpha \cap \overline{\gamma} = \emptyset$ (p. 174)

All valid syllogism of the second kind meanwhile have the form,

$$\alpha \cap \beta \neq \emptyset,\ \beta \cap \overline{\gamma} = \emptyset\ \therefore\ \alpha \cap \gamma \neq \emptyset$$

The validity of this form can also be derived using the algebra of classes already learned.

Since, $\alpha \cap \emptyset = \emptyset$, then if $\alpha \cap \beta \neq \emptyset$ then $\alpha \neq \emptyset$ and $\beta \neq \emptyset$

Since, $\alpha \cap \emptyset = \emptyset$ and $\beta \cap \overline{\gamma} = \emptyset$, which is a premise, then

$\alpha \cap (\beta \cap \overline{\gamma}) = \emptyset$ which by association yields,

$(\alpha \cap \beta) \cap \overline{\gamma} = \emptyset$

Since $(\alpha \cap \beta) = (\alpha \cap \beta) \cap U$ and $\gamma \cup \overline{\gamma} = U$ then

$\alpha \cap \beta = (\alpha \cap \beta) \cap (\gamma \cup \overline{\gamma})$ which by distribution yields,

$\alpha \cap \beta = [(\alpha \cap \beta) \cap \gamma] \cup [(\alpha \cap \beta) \cap \overline{\gamma}]$

But $[(\alpha \cap \beta) \cap \gamma] \cup \emptyset = (\alpha \cap \beta) \cap \gamma$ and we have already shown that $(\alpha \cap \beta) \cap \overline{\gamma} = \emptyset$,

Therefore, $\alpha \cap \beta = (\alpha \cap \beta) \cap \gamma$ however, since $\alpha \cap \beta \neq \emptyset$ is a premise, we know that

$(\alpha \cap \beta) \cap \gamma \neq \emptyset$ which by association and commutation yields,

$(\alpha \cap \gamma) \cap \beta \neq \emptyset$ from which it follows,

$\alpha \cap \gamma \neq \emptyset$

Clearly, the algebra of classes is not only capable of validating immediate inferences involving categorical propositions but is also capable of validating categorical syllogisms. (*l.c.*)

At this point Copi introduces the symbol '⊂' for **class inclusion**. Thus, '$\alpha \subset \beta$' asserts that all members of α, if any, are also members of β, which is an alternative formulation proposition **A**: 'all α is β'. There are several ways in which '$\alpha \subset \beta$' can be defined using only the symbols already introduced, either as

$$\alpha \cap \overline{\beta} = \emptyset \text{ or as } \alpha \cap \beta = \alpha \text{ or as } \alpha \cup \beta = \beta \text{ or as } \overline{\alpha} \cup \beta = U$$

all of which are equivalent. The relation ⊂ has a number of properties discussed in Critical Reasoning 14. These include reflexivity and transitivity as well as the property of transportation such that if $\alpha \subset \beta$ then $\overline{\beta} \subset \overline{\alpha}$. The latter can be shown by double negation and commutation by rewriting '$\alpha \subset \beta$' as '$\alpha \cap \overline{\beta} = \emptyset$ and '$\overline{\beta} \subset \overline{\alpha}$' as '$\overline{\beta} \cap \overline{\overline{\alpha}} \neq \emptyset$'. Its reflexiveness is obvious simply by rewriting '$\alpha \subset \alpha$' as '$\alpha \cap \overline{\alpha} = \emptyset$'. Its transitivity meanwhile has already been shown by Copi's algebraic proof of validity for categorical syllogisms containing only universal propositions, above. (*l.c.*)

Axioms for Class Algebra

Boolean algebra , after the English mathematician, philosopher and logician George Boole (1815 - 1864) is the algebra of two-valued logic (true or false) with only sentential connectives *and*, *or* and *not*. Equivalently the algebras of classes, under *product*, *sum* and *complement* can be set up as a formal deductive systems. Although there are many alternative postulation sets for Boolean algebra, we will be examining the one set out by Copi (1979, p. 175 - 176). The axioms and theorems that follow are reproduced essentially verbatim.

The special, unidentified primitive symbols of which are:

$\boldsymbol{C}, \cap, \cup, -, \alpha, \beta, \gamma, \ldots$

Axioms:

Ax. 1. If α and β are in $\boldsymbol{C}$, then $\alpha \cup \beta$ is in $\boldsymbol{C}$.

Ax. 2. If α and β are in $\boldsymbol{C}$, then $\alpha \cap \beta$ is in $\boldsymbol{C}$.

Ax. 3. There is an entity $\emptyset$ in $\boldsymbol{C}$ such that $\alpha \cup \emptyset = \alpha$ for any α in $\boldsymbol{C}$.

Ax. 4. There is an entity U in $\boldsymbol{C}$ such that $\alpha \cap U = \alpha$ for any α in $\boldsymbol{C}$.

Ax. 5 If α and β are in $\boldsymbol{C}$, then $\alpha \cup \beta = \beta \cup \alpha$.

Ax. 6. If α and β are in $\boldsymbol{C}$, then $\alpha \cap \beta = \beta \cap \alpha$.

Ax. 7. If α, β, γ are in $\boldsymbol{C}$, then $\alpha \cup (\beta \cap \gamma) = (\alpha \cup \beta) \cap (\alpha \cup \gamma)$.

Ax. 8. If α, β, γ are in $\boldsymbol{C}$, then $\alpha \cap (\beta \cup \gamma) = (\alpha \cap \beta) \cup (\alpha \cap \gamma)$.

Ax. 9. If there are unique entities $\emptyset$ and U satisfying Axioms 3 and 4, then for every α in $\boldsymbol{C}$ there is an $-\alpha$ in $\boldsymbol{C}$ such that $\alpha \cup -\alpha = U$ and $\alpha \cap -\alpha = \emptyset$

Ax. 10 There is an α in $\boldsymbol{C}$ and a β in $\boldsymbol{C}$ such that $\alpha \neq \beta$.

The system presented above is a formal deductive system rather than a logistic system. (See Critical Reasoning 16.) Although the primitive symbols used are undefined, their intended interpretation is that $\boldsymbol{C}$ the collection of all classes or sets, $\emptyset$ and U the empty and universal sets respectively, $\cup$ is class addition, $\cap$ is class multiplication and $-$ complementation. However another interpretation is possible, see below (p. 175)

Copi presents the following twenty theorems, some of which one may wish to derive as an exercise:

Th. 1. There is at most one entity $\emptyset$ in $\boldsymbol{C}$ such that $\alpha \cup \emptyset = \alpha$.

Th. 2. There is at most one entity U in $\boldsymbol{C}$ such that $\alpha \cap U = \alpha$.

Th. 3. $\alpha \cup \alpha = \alpha$

Th. 4. $\alpha \cap \alpha = \alpha$

Th. 5. $\alpha \cup U = U$

Th. 6. $\alpha \cap \emptyset = \emptyset$

Th. 7. $\emptyset \neq U$

Th. 8. If $\alpha = -\beta$ then $\beta = -\alpha$

Th. 9. $\alpha = - - \alpha$

Th. 10. If $\alpha \cap \beta \neq \emptyset$, then $\alpha \neq \emptyset$

Th. 11. $\alpha = (\alpha \cap \beta) \cup (\alpha \cap -\beta)$

Th. 12. $\alpha \cup (\beta \cup \gamma) = (\alpha \cup \beta) \cup \gamma$

Th. 13. $\alpha \cap (\beta \cap \gamma) = (\alpha \cap \beta) \cap \gamma$

Th. 14. $\emptyset = -U$

Th. 15. $\alpha \cup (\alpha \cap \beta) = \alpha$

Th. 16. $\alpha \neq -\alpha$

Th. 17. $-(\alpha \cap \beta) = -\alpha \cup -\beta$

Th. 18. $-(\alpha \cup \beta) = -\alpha \cap -\beta$

Th. 19. If $\alpha \cap -\beta = \emptyset$ and $\beta \cap -\gamma = \emptyset$ then $\alpha \cap -\gamma = \emptyset$

Th. 20. If $\alpha \cap \beta \neq \emptyset$ and $\beta \cap -\gamma = \emptyset$ then $\alpha \cap \gamma \neq \emptyset$

Copi hints that, "The methods of proof proceed largely by the substitution of equals for equals." (p. 176)

E.g. Th. 1: Consider any entities $\emptyset_1$ and $\emptyset_2$ in $\boldsymbol{C}$ such that $\alpha \cup \emptyset_1 = \alpha$ and $\alpha \cup \emptyset_2 = \alpha$.

Since α is any member of $\boldsymbol{C}$, $\emptyset_1 \cup \emptyset_2 = \emptyset_1$ and $\emptyset_2 \cup \emptyset_1 = \emptyset_2$.

Since $\emptyset_1 \cup \emptyset_2 = \emptyset_2 \cup \emptyset_1$ by axiom 5, we have

$\emptyset_1 \cup \emptyset_2 = \emptyset_2$ by substitution and

$\emptyset_1 = \emptyset_2$ by substitution again.

As mentioned above, Boolean Algebra is a formal deductive system amenable to more than one interpretation, one of which is the Algebra of Classes. However we can also assign to it a propositional interpretation on which '$\boldsymbol{C}$' is the collection of all propositions, with 'α, β, γ, …' as individual propositions and '∩', '∪' and '—' interpreted as conjunction, weak disjunction and negation respectively. Furthermore, if we interpret '=' as material equivalence then "all the axioms of Boolean Algebra become logically true propositions of the propositional calculus. Hence we can say that the propositional calculus is *a* Boolean Algebra." (p. 176.)

Zermelo-Fraenkel Set Theory (ZF) - The First Six Axioms

Cantor's set theory soon ran into some embarrassing contradictions such as Russell's paradox (discussed below). This prompted a number of early 20th century mathematicians and logicians including Ernst Zermelo and Abraham Fraenkel to develop an axiomatic system that would be consistent. According to Copi, "It was suspected that some unrecognized assumption was responsible for the contradictions. So it was natural to attempt to construct a deductive system for set theory, in which every assumption would be explicitly stated as an axiom." (p. 176 - 177)

According to Cantor's definition of a set given in his *Beiträge zur Begründung der transfiniten Mengenlehre*:

> A set is a gathering together into a whole of definite, distinct [or separate] objects of our perception [*Anschauung*] or of our thought—which are called **elements** of the set.

This conception involves at least three ideas: Firstly, what is meant by 'definite' is that there is some criterion for set membership, by which we may decide, at least in theory, whether or not some object is a member of the set in question. Secondly, what is meant by 'distinct' (or 'separate') is that any member can be recognised, again at least in theory, as different from any other member, so that no member gets counted more than once. Thus the set $\{1; 2; 2 - 1\}$ is counted as having only two members rather than three. Thirdly, that a set is 'gathering together into a whole' "indicates that sets themselves are objects and, therefore, are eligible to be members of other sets". (p. 177) Whether sets are also 'objects of our perception [*Anschauung*] or of our thought' is a psychologically contentious notion that need not detract from the present discussion.

While Cantor's definition of sets turned out to be inadequate for formal mathematics, the notion of a set in axiomatic set theory is instead taken as an undefined, primitive term with its properties defined by the Zermelo-Fraenkel axioms. (Wikipedia: Set (mathematics))

According to Copi, the fundamental relation in abstract set theory is that of membership, symbolised by the stylised Greek letter '∈'. Thus '$a \in A$' asserts that a is a member of (or belongs to) A, while '$b \notin A$' denies that b is a member of (or belongs to) A. Thus if the nation, France is regarded as the set of all its citizens, then

Pierre ∈ France

And since France is a member of the United Nations, then

France ∈ U.N.

But since the U.N. comprises of only member states, we have it that

Pierre ∉ U.N.

which shows that ∈ is not a transitive relation. (p. 177)

Although different attributes may pertain to the same object, such as the attribute of being a featherless biped pertaining to all and only rational animals (with the exception of Diogenes' plucked chicken); a set however is determined by its members. Thus the set of featherless bipeds, F is identical to the set of rational animals, R, written as

$$F = R$$

Whereas to deny that set A is the same as set B, is written

$$A \neq B$$

According to **Leibniz's Law** also known as the **Law of Identity** or **Identity of Indiscernibles**

$x = y$ if and only if every attribute of x is an attribute of y, and conversely.

Which, if we allow quantification over predicate variables, the identity relation can be defined as

$x = y = df(\forall F)(Fx \equiv Fy)$ (See Critical Reasoning 14.)

From this definition follow all the characteristics of the identity relation embodied in the rules of inference under the heading of 'Rules of Identity (Id.)' in Critical Reasoning 14. Although these rules are included in the logic used by set theory to derive conclusions from premises and theorems from axioms, the difference between sets and attributes remains: Although two attributes may belong to exactly the same distinct objects, two sets containing exactly the same members are identical. This characteristic of sets is stated as the first axiom of the **ZF** system, "not as a logical truth about the identity relation but as an assumption about sets". (p. 178) Thus,

ZF-1 $A = B = df(\forall x)(x \in A \equiv x \in B)$

This is known as the **Axiom of Extensionality**, which states that a set is determined or defined by its members.

If a set C has all its members as elements of set D, then C is said to be a **subset** of D, and C is said to be **included** in D. The subset or inclusion relation '⊂' can now be defined as

$C \subset D = df(\forall x)(x \in C \supset x \in D)$

According to this definition, every set is a subset of itself so that the relation '⊂' is reflexive. It is also transitive because if $C \subset D$ and $D \subset E$, then $C \subset E$. Where $C \subset D$ and $D \subset C$, then by the Principle of Extensionality, $C = D$. However where $C \subset D$ and $C \neq D$, then C is said to be a **proper subset** of D. In other words, a proper subset of a set A is a subset of A that is not equal to A. (*l.c.*)

In keeping with the set-builder notation learned in primary school, sets containing only a few members may be represented by listing their members within curly brackets or braces, separated by semicolons. Thus '$\{a; b\}$' or '$\{b; a\}$' contains just two members or elements, 'a' and 'b'. The order in which the members are written is irrelevant, since sets containing the same members are identical. A set containing just two members is known as a **pair set** or a **doubleton**. (*l.c.*)

Representing larger sets requires a variable (such as 'x'), a vertical bar separator (representing the concept of 'such that') and a logical predicate ('$\varphi(x)$') or propositional function that expresses some rule or condition that is satisfied by all and only those objects that belong to that set, all enclosed within braces as follows

$\{x|\Phi(x)\}$

All values of x in the universe of discourse, where $\Phi(x)$ is true are counted as in the set and those where $\Phi(x)$ is false are not in the set. Thus the set of all citizens of France would be written

$\{x|x\ is\ a\ citizen\ of\ France\}$

More generally, where Fx is satisfied by all those objects belonging to set S

$$S = \{x|Fx\}$$

So for any object, y

$$y \in \{x|Fx\} \equiv Fy$$

According to Copi, "This notation is in the spirit of Bertrand Russell's remark that '… a class [set] may be defined as all the terms [objects] satisfying some propositional function'". (p. 179)

Furthermore such notation subsumes the earlier method of notation in the way that, for example, the set $\{a; b; c\}$ can be symbolised as

$$\{x|x = a \text{ v } x = b \text{ v } x = c\}$$

According to Copi, "Cantor apparently believed - for some time at least - that *any* condition on objects could determine or define a set containing or comprehending just those objects satisfying that condition. Although Cantor did not formulate this belief it as a principle, it could be stated as

$$(\exists S)(\forall x)(x \in S \equiv \varphi x)$$

where 'φ' represents any predicate or condition and there is no free occurrence of 'S' in 'φx'". (p. 179) Unfortunately for Cantor and later Frege, this unrestricted **principle of comprehension** leads to a genuine contradiction, known as **Russell's paradox** after Bertrand Russell, who discovered it in 1901. Let φx be the propositional function that $x \notin x$, *i.e.* that x is not a member of itself. Then we have the set

$$R = \{x|x \notin x\}$$

Therefore for any y we have that

$$y \in R \equiv y \notin y$$

But substituting R for y yields

$$R \in R \equiv R \notin R$$

which is an obvious contradiction. If this is your first encounter with Russell's paradox you might feel that something is amiss. Only a very strange sort of set could be a member of itself! Not at all - we have already encountered the set of all sets, $\boldsymbol{U}$, which on a little reflection must be a member of itself. However what about the set of all sets which are not members of themselves? Is it a member of itself? Well, it is if it isn't and it is not if it is - a clear contradiction. (p. 179)

Another puzzle derived from Russel's paradox is the **barber paradox** in which the village barber shaves all those, and only those, who do not shave themself. Does the barber shave himself? Well he wouldn't shave himself because he only shaves those who do *not* shave themselves. Conversely, if the barber does not shave himself then he is among those who do not shave themself, whom he does shave. One solution to the puzzle is that such a barber does not exist. Russel was not entirely happy with the barber's paradox, which although attributed to him, was in fact only suggested to him by an unnamed person. Others have suggested that the barber must be a woman.

To avoid Russell's paradox requires that we place some restriction on the principle of comprehension. The one in use today, first proposed by Ernst Zermelo in 1908 and later augmented by Abraham Frankel in 1922 to produce the '**ZF**' system, requires that, instead of assuming that *any* condition on *any* objects defines the set of such objects, rather *given* any set A, any condition on members of A defines a *subset* of A that contains just those members of A that satisfy the condition. Zermelo called this restricted principle of comprehension the *Aussonderung* axiom. (p. 179) Also known as the **Axiom of Separation**, it may be stated as

ZF-2 $(\exists S)(\forall x)(x \in S \equiv x \in A \bullet \varphi x)$

where, again, 'φ' represents any predicate or condition and there is no free occurrence of 'S' in 'φx'. The axiom of separation avoids Russell's paradox as follows: Given any set A, if we let φx be the propositional function that $x \notin x$, *i.e.* that x is not a member of itself, then we have the set

$$R = \{x | x \in A \bullet x \notin x\}$$

Therefore for any y we have that

$$y \in R \equiv y \in A \bullet y \notin y$$

Now, substituting R for y yields

$$R \in R \equiv R \in A \bullet R \notin R$$

which is *not* a contradiction. This can be further simplified to

$$R \notin R \bullet R \notin A$$

from which it follows that

$$R \notin A$$

But, as Copi points out, "since A was any set whatever, and R's existence follows from the Axiom of Separation, it has been shown that, given any set, there is something that is not a member of it". In other words on **ZF** set theory, there is no universal set, as has been proved above. By 'universal' here we mean, in the context of a *universe of discourse*, a universal set would be one that contains all the objects that enter into that discussion. Contrast this with the Algebra of Classes in which a universal set is postulated. (p. 180)

Note that for the axiom of separation to produce a new set S there must already exist a set A of which S will be a subset, therefore we must assume that there exists at least one set. Using the same axiom and the propositional function $x \neq x$, which is not true for any x, we can prove that the **empty set** exists. Thus given the set A, we have

$$(\exists B)(\forall x)(x \in B \equiv x \in A \bullet x \neq x)$$

Thus,

$$B = \{x | x \in A \bullet x \neq x\}$$

And since there is no y such that $y \in A \bullet y \neq y$, it follows that for every y, $y \notin B$, and by the axiom of extensionality, that there can only be one set with no elements. (p. 181)

There are several ontological issues raised against the very idea of the empty set. If a set is to be interpreted extensionally by its members, of which there are none to define it, then it has no extension and fails to be anything at all. We cannot answer such objections here decisively, however consider the following: We do not agonise over the ontological status of the number zero the way that the ancient Greeks did who were perplexed as to how nothing could be something. Using the number zero to designate nothing does not mean that we have no concept of what we mean. We do! Similarly, we have already encountered examples of empty names in Classic Text 17 in which we have a meaningful concept but no real world reference that satisfies the truth function. The failure of such terms to refer does not mean that we fail to have a meaningful concept of them. And so, when we think of the empty set as defined extensionally by the absence of members, we are not failing to have a concept when actually we mean something quite explicit. And finally, "if there are any sets at all, the Axiom of Separation entails the existence of the empty set." (*l.c.*)

According to Copi's presentation of the **ZF** system, it is simpler to postulate the existence of the empty set directly as Axiom III here rather than the stronger assumption in Axiom VII that presupposes the existence of the empty set. Thus the **Empty Set Axiom** is stated as

ZF-3 There exists a set ∅ such that for any object x, $x \notin \emptyset$,
from which it is an immediate consequence that for any set S, $\emptyset \subset S$.

By 'objects', here Copi means 'sets of objects' exclusively, which can be the objects of mathematics including sets of lines, sets of points, sets of numbers and sets of functions of various kinds, though not individuals, which would unnecessarily complicate matters, although there are such formulations. The Empty Set Axiom shows up another fundamental difference between the Algebra of Classes, for which every class has an absolute complement, and set theory, for which there is no absolute complement of ∅. (p. 182)

According to the **Axiom of Pairing**:

ZF-4 Given any sets a and b, there exists a set S having just a and b as members.

Symbolically: $(\exists S)(\forall x)(x \in S \equiv x = a \vee x = b)$

For any sets x and y therefore, there exists a **pair set** denoted $\{x; y\}$, also called the **unordered pair**. Note that the Axiom of Pairing makes no assumption about the distinctness of the sets, thus from any set a we can form the unordered pair $\{a; a\}$, which can also be written as $\{a\}$ known as a **unit set** or **singleton**. (*l.c.*)

According to Copi, Russell was initially inclined to believe that $\{a\}$ was self-evidently the same as a; however he was later persuaded by the following argument, attributed to Frege, that a single term should be distinguished from the class of which it is its only member: "Let u be a set having more than one member; let $\{u\}$ be the set whose only member is u; then $\{u\}$ has one member, u has many members, hence $u \neq \{u\}$". An analogous argument may be made using the empty set: Because ∅ has no members and $\{\emptyset\}$ has one, it follows that $\emptyset \neq \{\emptyset\}$ and hence that $u \neq \{u\}$. (*l.c.*)

Apart from the non-existence of the universal set and absence of absolute complements, the **ZF** system conforms to the rules for the Algebra of Classes: Consider intersection first. Any two sets a and b have an intersection $a \cap b$ as per the Separation Axiom. If we let φx in the axiom be '$x \in b$' then we have

$$(\exists S)(\forall x)(x \in S \equiv x \in a \bullet x \in b)$$

And since S is unique by extensionality we may define the connective '∩' as

$$a \cap b = df\ \{x|x \in a \bullet x \in b\}$$

This process can be repeated to obtain $(a \cap b) \cap c$ if c is also a set, and $[(a \cap b) \cap c] \cap d$ if d is also a set, and so on. It can also be shown that $a \cap \emptyset = \emptyset$ and that ∩ is commutative, associative and idempotent. More generally, if A is an arbitrary *non-empty* collection of sets a_1, a_2, a_3, …, then there exists a set B that is the intersection of *all* the members of A. If we let a_i be any particular set in A and if we let φx be the condition that x belongs to every set in A, then by the Axiom of Separation

$$(\exists B)(\forall x)[x \in B \equiv x \in a_i \bullet (\forall a)(a \in A \supset x \in a)] \qquad \text{(p. 183)}$$

Note that B does not depend on the choice of any particular a_i over another since

$$B = \{x|(\forall x)(a \in A \supset x \in a)\}$$

And since B is unique by extensionality, Copi defines the symbol '⋂**'** (written in front of a set rather than between them,) as

$$\bigcap A = df\ \{x|(\forall a)(a \in A \supset x \in a)\}$$

Other notifications for $\bigcap A$ vary considerably, including '$\bigcap_i a_i$', '$\bigcap\{a \in A\}$', '$\bigcap\{x|x \in A\}$' and '$\bigcap_{x \in A} x$'. It should be obvious that general intersection must be the same in extensionality as regular intersection which we may symbolise as

$$\bigcap\{a; b\} = a \cap b \qquad \text{(p. 183)}$$

One of the stipulations in defining $\bigcap A$ was that A is non-empty. If that were not so then there would be no a_i whose membership is needed for the Axiom of Separation, so $\bigcap \emptyset$ is not defined. Nor could we define it as $\{x|(\forall a)(a \in \emptyset \supset x \in a)\}$ because, as Copi observes "every x satisfies the condition that $a \in \emptyset \supset x \in a$, since $a \notin \emptyset$ for every a. Thus the suggested definition would make $\bigcap \emptyset$ contain everything, whereas we have already shown that there is no universal set in **ZF**." So either we could leave it undefined or we could make $\bigcap \emptyset = \emptyset$ by definition, in order to make $\bigcap S$ defined for any S whatsoever. (*l.c.*)

Next consider union. Given any two sets a and b, we want them to have a union $a \cup b$ that contains all, and only, those members that are either in a or in b. However we cannot use the Axiom of Separation here, so Copi suggest that "we *could* introduce another axiom such that

$$(\exists B)(\forall x)(x \in B \equiv x \in a \vee x \in b)$$

where by Extensionality B is the unique set $\{x|x \in a \text{ v } x \in b\}$ for which we introduce the notation '$a \cup b$'. Then we could use it repeatedly to obtain $(a \cup b) \cup c$ if c is also a set and $[(a \cup b) \cup c] \cup d$ if d is also a set, and so on." If we *were* to do so we would see that "$a \cup \emptyset = a$, and that $\cup$ is commutative, associative and idempotent, and also that $\cap$ and $\cup$ are each distributive with respect to the other." However this method of union building 'one at a time' will be impractical for a set that contains every object that belongs to any set a_i of a collection A of sets, $\{a_1; a_2; a_3 \ldots; a_i; \ldots\}$ if the collection of A if large and impossible if it is infinite. A collection is said to be **infinite** if there is no natural number that is the number of its members. (p. 184)

So clearly the hypothetical axiom above will not do, however given a collection A of sets, $\{a_1; a_2; a_3 \ldots; a_i; \ldots\}$ the general **Axiom of Union** guarantees the existence of a set B that contains anything that belongs to any member of A, thus

ZF-5 $\quad (\forall A)(\exists B)(\forall x)[x \in B \equiv (\exists a)(x \in a \bullet a \in A)]$

Moreover, the Axiom of Extensionality guarantees the uniqueness of set B, so if $A = \{a_1; a_2; a_3 \ldots\}$, B can be symbolised as '$\bigcup A$' or variously as '$\bigcup_i a_i$'or as '$\{x|(\exists a)(x \in a \bullet a \in A)\}$' or as '$\bigcup\{a|a \in a\}$' or as '$\bigcup\{x|x \in A\}$' or as '$\bigcup_{x=A} x$'. (*l.c.*)

Clearly we want to be able to obtain the more modest $a \cup b$, given sets a and b. This can be done as follows: From a and b, by the Axiom of Pairing, we have $\{a; b\}$. Next by the Axiom of Union $\bigcup\{a; b\}$ is given by, $\{x|(\exists z)(x \in z \bullet z \in \{a; b\})\}$. But the only z's in $\{a; b\}$ are a and b, so $\bigcup\{a; b\} = \{x|x \in a \text{ v } x \in b\}$ which is simply $a \cup b$. According to Copi, "it is obvious that $\{x|x \in \emptyset\} = \emptyset$, so $\bigcup \emptyset = \emptyset$, and that $\bigcup\{x|x \in \{a\}\} = a$, so $\bigcup\{a\} = a$. It is also clear that $\bigcup\{x|x \in \{a; b\}\} = a \cup b$, so $\bigcup\{a; b\} = a \cup b$, and that $\{a\} \cup \{b\} = \{a; b\}$." (p. 184)

Although we do not have *absolute* complements in the **ZF** system, we can speak of a *relative* complement $A - B$ of B in A, which is the set of all members A that are not members of B. The existence and uniqueness this set is given by the Axioms of Separation and Extensionality, and defined as

$$A - B = df\ \{x|x \in A \bullet x \notin B\}$$

Again there are some obvious entailments: $A - \emptyset = A$, and $A - A = \emptyset$, and that if $A \cap B = \emptyset$ then $A - B = A$. Rather than being merely an academic distinction the difference between absolute and relative complements is quite useful in mathematics. When for example, we speak of the complement of even numbers as odd numbers, we mean the *relative* complement. The absolute complement of even numbers, by contrast, would have to include, not only odd numbers, but every other object that is not an even number, including *inter alia* every point, line, circle, plane, and function. (*l.c.*)

If, *in a particular discussion*, we are careful to stipulate that we are considering all sets a, b, c … as subsets of some specified set V, then we can symbolise the relative complements of a in V, b in V, c in V and so on, as we did with absolute complements, as: '$\overline{a}$', '$\overline{b}$', '$\overline{c}$' *etc.* And again, *in the context of that discussion*, we can rely on the laws of the Algebra of Classes presented at the beginning of this study unit. (p. 185)

We have already defined the concept of a subset, however there may be times when we wish to consider all the subsets within a given set. The totality of subsets within a given set E is stated by the **Power Axiom**:

ZF-6 $(\forall E)(\exists S)(\forall x)(x \in S \equiv x \subset E)$

The Axiom of Extensionality ensures the uniqueness of such a set and allows us to define the **power set** of E, symbolised as '$\wp E$', as

$$\wp E = df\ \{x | x \subset E\}$$

If E is a finite set, then $\wp E$ contains more members than E does. This may sound rather strange, after all how can a set have more subsets than the number of its members? Copi however provides the following examples: $\wp\emptyset = \{\emptyset\}$ which has one member; $\wp\{a\} = \{\emptyset; \{a\}\}$ which has two and $\wp\{a; b\} = \{\emptyset; \{a\}; \{b\}; \{a; b\}\}$ which has four. In general, it can be shown by **mathematical induction**, that if a given set E has n number of members then $\wp E$ has 2^n members, hence the term 'power set'. (*l.c.*)

Finally, for this section, Copi lists a number of obvious features of power sets: Since for any set E, we know that $E \subset E$ and $\emptyset \subset E$, it follows that $E \subset \wp E$ and $\emptyset \in \wp E$. From the latter it is clear that $\bigcap\{x | x \in \wp E\} = \emptyset$. "Other immediate consequences of our definitions and axioms are that $E \subset F$ if and only if $\wp E \subset \wp F$, that $\wp E \cup \wp F \subset \wp(E \cup F)$, that $\wp(E \cap F) = \wp E \cap \wp F$, and that, although $E = \bigcup \wp E$, in general we have only that $E \subset \wp \bigcup E$." (p. 185)

Mathematical Induction

Mathematical induction is a method for proving that a statement $P(n)$ is true for every natural number n, *i.e.* that the infinitely many cases $P(0)$ $P(1)$, $P(2)$, $P(3)$, ... all hold. Consider the following informal metaphors: If we can show that the first domino in a row will fall causing the next one to fall, then so will the second, and the third… and so on; alternatively if we can prove that we can climb on to the bottom rung of a ladder (the base case) and from each rung to the next one up (the induction step), then we can prove that we can climb up as high as we wish on the ladder. See the discussion of Peano's axiomatic system below.

Relation and Functions

Before exploring further implications of the Power Set Axiom, Copi takes a necessary digression into the set theory of relations and functions. In Critical Reasoning 14 we discussed the logic of relations and several of their attributes, including those of binary relations. Because not all binary relations are symmetrical we need a way to distinguish between, say, the relation Rab and Rba because Rab might be true while Rba might be false, so the *order* of the arguments a and b matters. One way to express this is to say that the relation R (or the propositional function Rxy is satisfied by) the ordered pair $(a; b)$ but not by $(b; a)$. (*l.c.*)

On pages 10 - 11 above we saw how a set can be expressed as the extension of a predicate, subject to the restrictions of the Separation Axiom. Thus, the set Fcontains all the objects for which the predicate or condition φx is true. Likewise, consider the collection of ordered pairs of objects $(x; y)$ between which the relation Rxy obtains. Although these members may determine a unique set, irrespective of the order in which they are arranged, it is possible to define order in purely set theoretic terms. According to Copi, "the essence of the notion of *order* is that the ordered pair $(x; y)$ is the same as the ordered pair $(a; b)$ if, and only if, both $x = a$ and $y = b$. This essence is captured by the Wiener-Kuratowski definition of an ordered pair." Thus,

$$(x; y) = df\ \{\{x\}; \{x; y\}\}$$

Given any x and y we can use the Axiom of Pairing to produce such a set, which will be unique by Extensionality. Obviously if both $x = a$ and $y = b$ then $(x; y) = (a; b)$. Conversely:

If $(x; y) = (a; b)$ then by the above definition $\{\{x\}; \{x; y\}\} = \{\{a\}; \{a; b\}\}$

Since $\{x\} \in \{\{x\}; \{x; y\}\}$ we have $\{x\} \in \{\{a\}; \{a; b\}\}$

Hence either $\{x\} = \{a\}$ or $\{x\} = \{a; b\}$

In either case $x = a$

And since $\{a; b\} \in \{\{a\}; \{a; b\}\}$ we have $\{a; b\} \in \{\{x\}; \{x; y\}\}$

Hence either $\{a; b\} = \{x\}$… ① or $\{a; b\} = \{x; y\}$ … ②

Similarly, either $\{x; y\} = \{a\}$… ③ or $\{x; y\} = \{a; b\}$ … ②

So, if ① and ③ are both true, then $x = a = y = b$

But if either ① or ③ are not true, then $\{x; y\} = \{a; b\}$ and $x = a$, as above.

Therefore $\{a; y\} = \{a; b\}$

Hence, if $a \neq b$ then $y = b$, but if $a = b$ then $y = b$

Therefore in every case $y = b$ (p. 186)

Given two sets A and B we can form an ordered pair $(x; y)$ such that $x \in A$ and $y \in B$. The collection of all such pairs is known as the **Cartesian Product** $(A \times B)$ of A and B. The Cartesian plane is one such example. If we consider the Euclidian plane with x and y axes established, then "if A is the set of all real numbers identified with points on the x-axis and B is the set of all real numbers identified with points on the y-axis, then the set of all ordered pairs $(x; y)$ represents the Cartesian plane itself - which accounts for the name". This collection of ordered pairs forms a set itself which can be shown as follows: Given sets A and B we can form their union $A \cup B$, which contains every $x \in A$ and every $y \in B$. Next we can form the power set of their union $\wp(A \cup B)$ which contains every set $\{x\}$ and ever set $\{x; y\}$ *inter alia*. But we can also form the power set of *that* power set $\wp\wp(A \cup B)$ which contains every set $\{\{x\}; \{x; y\}\}$, *inter alia*, which by the Axiom of Separation, yields just the desired subset. Thus,

$$(\exists S)(\forall z)\left[z \in S \equiv \left[\left(z \in \wp\wp(A \cup B)\right) \bullet (\exists x)(\exists y)[x \in A \bullet y \in B \bullet z = (x; y)]\right]\right]$$

(p. 186 - 187)

According to this formulation, for any sets A and B, there exists a set containing all and only those ordered pairs $(x; y)$ such that $x \in A$ and $y \in B$. This set is unique by Extensionality and is symbolised '$(A \times B)$'. By a similar argument involving unions instead of powers, it can be shown that if $(x; y) \in A$, then both x and y belong to $\cup\cup A$. Copi points out that for sets A and B it is not necessary that $A \neq B$ because "the product $A \times A$ is a perfectly good Cartesian product, that of A with itself." (p. 187)

Now any binary relation can defined simply as a set of ordered pairs such that R is a relation if and only if

$$(\forall x)\big[x \in R \supset (\exists u)(\exists v)[x = (u; v)]\big]$$

Such relations include the Cartesian product of any sets A and B, as well as any subset of a Cartesian product, including that of the empty set, *i.e.* the empty relation $\emptyset$. (p. 187)

The **domain of a relation** R, (dom R) is defined as the set of all first coordinates of R, while the **range of a relation** R, (ran R) is defined as the set of all second coordinates of R. Meanwhile, the **field of a relation** R, (fld R) is defined as the set of all coordinates of R. Given any relation R we are able to specify both its domain and range as follows. Using the Axiom of Union twice we can produce the set $\cup\cup R$ from which, by the Axiom of Separation, we get,

$$(\exists S)(\forall x)\left[x \in S \equiv \left[x \in \cup\cup R \bullet (\exists y)\big((x; y) \in R\big)\right]\right]$$

as the unique domain of R, and

$$(\exists T)(\forall x)\left[x \in T \equiv \left[x \in \cup\cup R \bullet (\exists z)\big((z; x) \in R\big)\right]\right]$$

as the unique range of R. It follows, for example, that dom $\emptyset$ = ran $\emptyset$ = $\emptyset$, that dom $(A \times B)$ = A and that ran $(A \times B)$ = B. It also follows that if R is the relation of equality of members of E, then dom R = ran R = E (*l.c.*)

Except for the empty relation, it is necessary to specify a set within, or over which, the relation is defined. This is not just an academic exercise. As Copi points out, if we were to define the identity relation $X = Y$ without specifying its domain (or range or field) we would have the collection Z of all ordered pairs $(x; y)$ such that $x = y$, which would contain every set $z = \big\{\{x\}; \{x; x\}\big\}$. But then the union of its union $\cup\cup Z$ would be the universal set, which has been proved, cannot exist. The same problem would arise if we were to use set membership ($x \in y$) or set inclusion ($x \subset y$) without restriction. (p. *l.c.*)

We are all familiar with functions in everyday life: The cost of a bag of potatoes of a certain grade is a function of its weight, and the tension in a given spring is a function of its length. In terms of set theory, a function is a one-to-one or many-to-one relation represented by a set of ordered pairs, such that no two of which have the same first coordinate. Thus if X and Y are sets, then a function *from* X (or *on* X) *to* Y (or *into* Y) is a relation f, such that dom f = X and such that for each x in X there is a unique y in Y with $(x; y) \in f$. Copi uses the following notation to indicate that f is a function from X to Y, thus: '$f: X \to Y$' which is defined as:

$$f \subset X \times Y \bullet (\forall x)\left[x \in X \supset (\exists y)\left[y \in Y \bullet (x; y) \in f \bullet (\forall z)\big((x; z) \in f \supset y = z\big)\right]\right]$$

The familiar way of representing such a function is simply $f(x)$ or $f(x) = y$ such that y is the corresponding value the function assumes given the argument x. The range of f meanwhile can be a proper subset of Y, however if it is equal to Y then f is said to map X onto Y. (p. 187 - 188)

Further definitions: If a function f from X to Y maps distinct elements in X onto distinct elements in Y then the function is said to be **one-to-one**. Moreover, "if $f: X \to Y$ is one-to-one and onto Y, then f effects a pairing of the elements of X with those of Y, and as such, is called a **one-to-one correspondence** between X and Y." In the case that there is a one-to-one correspondence between sets X and Y they are said to be **equivalent** or (**equinumerous**), symbolised as '$X \sim Y$'. (p. 188)

A set is defined as **finite** if there exists a set onto which a one-to-one function maps the set of all natural or counting numbers ℕ = {0; 1; 2; … } that are less than some natural number n, in which case the set is said to contain n elements. *E.g.* there is a one-to-one function that maps the set of all natural numbers less than n = 26 onto the set of letters of the English alphabet: 0 onto a; 1 onto b; 2 onto c; … 24 onto y and 25 onto z. An **infinite set** then is simply negatively defined as one that is not finite. (*l.c.*)

In the following paragraph Copi informally discuses some differences in equivalence between finite and infinite sets. Although, "no finite set is equivalent to any of its proper subsets [...] the matter is otherwise with infinite sets". Before the advent of set theory, Galileo observed counterintuitively that although not all numbers are perfect squares, yet there are as many perfect squares as there are numbers because they are as numerous as their roots, and all numbers are roots. Furthermore, there are as many whole numbers as there are rational fractions. We can show that there is a one-to-one correspondence between these two sets by the following rule which specifies a one-to-one function: Given two fractions, the one whose sum of the numerator and denominator is smallest, we enumerate first. Given two fractions whose sum of the numerator and denominator is the same, we count the one with the smallest numerator. (We do not enumerate equivalent fractions such as 2/4 and 4/8 because they will have already been counted as 1/2.) When we order the rational fractions accordingly, the following one-to-one correspondence becomes apparent:

$^1/_1$	$^1/_2$	$^2/_1$	$^1/_3$	$^3/_1$	$^1/_4$	$^2/_3$	$^3/_2$	$^4/_1$	$^1/_5$	$^5/_1$	$^1/_6$	…
0	1	2	3	4	5	6	7	8	9	10	11	…

(p. 188)

An **enumeration** of this sort is a one-to-one function from the natural numbers (0; 1; 2; …) onto the set being enumerated. Moreover, any set whose members can be mapped one-to-one onto the natural numbers is said to be **denumerable**, **countable** or **enumerable**. (p. 189)

Further definitions: "If there is a one-to-one correspondence between X and a proper subset of Y but no one-to-one correspondence between X and Y, then Y is said to be *larger than* X, or to have *more elements than* X, and X is said to be *smaller than* Y or to have *fewer elements than* Y." For example, it can be shown that the set of real numbers is larger than the set of natural numbers. Although Copi has not yet shown us how the natural numbers form a set, we can still appreciate Cantor's (1891) famous 'diagonal' proof which demonstrates that there are some infinite sets which cannot be put into one-to-one correspondence with the infinite set of natural numbers. In particular, on "any enumeration of real numbers there must be some real numbers that get left out". (*l.c.*)

Without loss of generality, we may confine ourselves to the interval from 0 to 1 and consider the following enumeration of real numbers in base 2 which Cantor used in his paper. Below we reproduce a table of just the first nine enumerations of the real numbers in the interval 0 to 1 to the 9th binary digit of the infinite array:

Enumeration	Binary Digit	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	... *j* th
1st	0.	**0**	0	0	0	0	0	0	0	0	...
2nd	0.	1	**1**	1	1	1	1	1	1	1	...
3rd	0.	0	1	**0**	1	0	1	0	1	0	...
4th	0.	1	0	1	**0**	1	0	1	0	1	...
5th	0.	1	1	0	1	**0**	1	1	0	1	...
6th	0.	0	0	1	1	0	**1**	1	0	1	...
7th	0.	1	0	0	0	1	0	**0**	0	1	...
8th	0.	0	0	1	1	0	0	1	**1**	0	...
9th	0.	1	1	0	0	1	1	0	0	**1**	...
...*i* th	0.	...	...	...	...	...	...	...	...	...	**...**

Because the array is ordered in a systematic way, we could fill in the corresponding table to any i^{th} enumeration of the real numbers to any *j*th digit we choose. Looking at the diagonal in red: by taking the complementary digits (swapping 0's for 1's and *vice versa*) of the diagonal, Cantor was able to construct a new sequence of digits (1; 0; 1; 1; 1; 0; 1; 0; 0…) that differed from any previous n enumerations because it would always differ at the n^{th} digit. Hence the number represented by the complementary sequence of digits in the diagonal cannot occur in the enumeration. (Wikipedia: Cantor's diagonal argument)

What is proved for the interval 0 to 1 will be true for all such intervals, therefore no enumeration of the real numbers will contain all of them. Hence the real numbers are non-denumerable, *i.e.* there are more real numbers than there are natural numbers (or integers). In fact, there are infinitely many more real numbers than there are natural numbers (or integers).

According to Copi, "Cantor's diagonal proof… is a special form of the proof of a more general theorem, which states that for any set A, its power set $\wp A$ is larger that A, itself". Obviously there is a one-to-one mapping from A onto the subset of singletons in $\wp A$. Thus for every a in A, there is a singleton $\{a\}$ in $\wp A$ and the function $s(a) = \{a\}$ is one-to-one. If we consider any one-to-one function f from A into $\wp A$ and form the subset A' of $\wp A$ that contains just those members a of A which are not members of $f(a)$ then the element of $\wp A$ onto which they are mapped is $A' = \{x : x \notin f(x)\}$. Now, for any a in A, if $a \in A'$ then $a \notin f(a)$ and $f(a) \neq A'$. However if $a \notin A'$ then $a \in f(a)$ and again $f(a) \neq A'$. But since a is *any* member of A, *no* member of A is mapped onto A' by f, which means that A' is left out of the mapping by f. But since f was *any* one-to-one function from A into $\wp A$ it follows that $\wp A$ is larger than A. (p. 189 - 190)

According to Copi, Cantor's unrestricted use of the principle of comprehension allowed for his set theory to contain the set of all sets S and by the argument above to contain $\wp S$, which paradoxically, must be larger than S when S is already the largest possible set. Fortunately the more limited Axiom of Separation of the **ZF** system circumvents this problem, allowing for ever larger sets to be formed using the Power Set Axiom repeatedly without ever reaching the absolute largest set. (p. 190)

Two further definitions for 'finite' and 'infinite sets' are as follows: An **infinite set** is one that is equivalent to one of its proper subsets, in which case a **finite set** is one that is not infinite. This pair of definitions can be shown to be equivalent to the one mentioned by Copi on p. 184 to wit that, a collection is said to be **infinite** if there is no natural number that is the number of its members. However this requires the Axiom of Choice to be introduced in the second to next section. Finally for this section, Copi mentions another definition of a finite set, attributed to the German mathematician Richard Dedekind *i.e.* "A set S is called *finite* if there exists a mapping of S into itself such that no proper subset of S is mapped into itself." (p. 190)

Natural Numbers and the Axiom of Infinity

Numbers are so much a part of our ordinary life that for most people the question of their existence never arises. In philosophy and mathematics, however nothing is taken for granted. In this section Copi sets about defining the natural (cardinal) numbers in terms of **ZF** set theory, postponing a discussion of ordinal numbers until the last section. Although mathematicians in the second half of the 19th Century carried out a program of "the arithmetization of analysis" in which complex, real and algebraic numbers as well as rational fractions and negative numbers were constructed on the base of natural numbers, it was not until Dedekind, and later Giuseppe Peano, that the natural numbers themselves were developed axiomatically.

Peano's axiomatic system relies on three undefined or primitive concepts: *zero*, *number* and *successor*. As Copi observes, "These are not defined explicitly *in* the system, but are rather implicitly defined *by* the axioms that make statements about them". Although there are nine axioms in total Copi only mentions five of them using x and x' to represent x and its successor respectively. Thus,

P1. 0 is a number.

P2. If x is a number, then x' is a number.

P3. If x and y are numbers and $x' = y'$, then $x = y$.

P4. There is no number x such that $x' = 0$.

P5. If $\varphi 0$ and for every number x, if φx then $\varphi x'$, then for every number x, φx. (p. 191)

Of course we can see here that what Peano meant by 'successor' - we can, informally and extra-systematically, say means 'adding 1'. The interpretation of P5 is also of special interest because it is a formulation of the principle of mathematical induction. However, not just the natural numbers but also many others sequences of numbers (including fractions) constructed from natural numbers also satisfy Peano's axioms as do systems other than the natural numbers. In the latter case we could just interpret them as alternative expressions of the natural numbers; however it is more desirable for us to have a way by which to actually *construct* the natural numbers and then back-check that they satisfy Peano's axioms. (p. 192)

Frege and later Russell, independently, did propose such a constructive definition based on the following: If A and B are equinumerous sets, *i.e.* between whose members there is a one-to-one correspondence, then A and B are said to have the same *number* of members. *E.g.* What is common to every pair set is that they have exactly *two* members. However there may be more similarities so that the number *two* cannot be defined exclusively in this way. Instead both Frege and Russell defined *two* to be the set of all pairs that is unique by Extensionality. Similarly, the number *zero* is the set whose only member is the empty set and *one* the set of all singletons, *three* the set of all triplets and so on. "More formally, a natural number is an equivalence class of finite sets under the equivalence relation of equinumerosity. This may appear circular, but it is not since equinumerosity can be defined without resort to the actual number of elements (for example, inductively)". (Wikipedia: Set-theoretic definition of natural numbers; Copi p. 192)

While the Frege-Russell definition of numbers is both intuitive and philosophically very satisfying, it cannot be used with **ZF** set theory directly, for the following reason: If there were a set of all singletons (or doubletons *etc.*) its union would be the universal set, which as we saw Russell's very paradox proved to be inconsistent. Instead Zermelo identified the natural numbers $0; 1; 2; 3$ … with the sets $\emptyset$; $\{\emptyset\}$; $\{\{\emptyset\}\}$; $\{\{\{\emptyset\}\}\}$ … Later, according to Copi, "John von Neumann proposed the use of an alternative and more convenient sequence of sets to indirectly define the natural numbers. By the Axioms of the Empty Set, Pairing and Union we can recursively define the following sequence of sets:

$$\emptyset, \{\emptyset\}, \{\emptyset; \{\emptyset\}\}, \{\emptyset; \{\emptyset\}; \{\emptyset; \{\emptyset\}\}\}, \ldots$$

where each set u is followed by its successor u^+ as the set $u \cup \{u\}$. Thus 0 is indirectly defined by the empty set, 1 by 0^+ which is the set $\{\emptyset\}$, 2 by 1^+ which is the set $\{\emptyset; \{\emptyset\}\}$, and so on. Note that instead of defining the number n directly, the Zermelo-von Neumann approach identifies the number n with a special *representative* n-numbered set that the **ZF** axioms guarantees to exist. (p. 192)

While this is not as elegant as the Frege-Russell definition, at least it is consistent. At any rate, we have it that:

$$0 = \emptyset$$
$$1 = \{0\}$$
$$2 = \{0; 1\}$$
$$3 = \{0; 1; 2\}\ldots$$

$$n^+ = \{0; 1; 2; 3; \ldots n\}\ldots$$

without limit, therefore infinitely many such sets must exist and by the next axiom there must exist a set containing them all.

ZF-7 **Axiom of Infinity**: There exists a set containing 0 and the successor of each of its elements.

As it stands the Axiom of Infinity provides that a set exists containing 0; 1; 2; 3…, what it does not guarantee is that such a set contains *all and only* such members, because there might be others. Copi therefore defines an **inductive set** as one that contains 0 and the successor of each of its members. But since 0; 1; 2; 3…, belongs to *every* deductive set, S, a subset of S must contain 0; 1; 2; 3…, so the intersection of all inductive subsets of S, call it ω (lower case omega), which exists by the Axiom of Separation and is unique by Extensionality, must contain *just* the natural numbers. Therefore we can define:

$$\omega = df \cap \{S_i : S_i \subset S \bullet S_i \text{ is an inductive set}\}$$

$$= \{0; 1; 2; 3; \dots\} \qquad \text{(p. 193)}$$

Now we can back-check that ω satisfies the Peano axioms: $0 \in \omega$ because ω is an inductive set and if $u \in \omega$ then $u^+ \in \omega$ for the same reason. This takes care of Axioms P1 and P2 respectively. Axiom P4 meanwhile which states that there is no number x such that $x' = 0$, follows directly from the fact that u^+ always contains u whereas 0, which is the empty set, does not contain any u. According to Copi, "because ω is the minimal inductive set, it follows that if any subset S of ω is an inductive set, then $S = \omega$. That is, if $S \subset \omega$, if $0 \in S$ and if $x^+ \in S$ whenever $x \in S$, then $S = \omega$. … which is a formulation of the principle of mathematical induction." (p. 193)

To prove P3 requires the introduction of the notion of a transitive set as well as two **lemmas** or subsidiary theorems about such sets. A **transitive set** is defined as one that contains every member of any of its members. Symbolically:

$$a \text{ is a transitive set} = df\ (\forall x)(\forall y)[(x \in y \bullet y \in a) \supset x \in a]$$

Transitive relations were discussed in Critical Reasoning 14 and in transitive sets the relation ∈ between a set's members and their members is a transitive one. *E.g.* The set of parts of plants forms a transitive set. Its members are roots, stems and leaves and these members also have members such as cuticles, epidermis, palisade layers and so on, which are also parts of plants…. *etc.* Alternative and equivalent ways of expressing 'a is a transitive set' include,

$$\cup a \subset a \qquad \text{or} \qquad (\forall x)(x \in a \supset x \subset a) \qquad \text{or} \qquad a \subset \wp a \qquad \text{(p. 194)}$$

Copi then proceeds with his two lemmas:

Lemma 1. *If a is a transitive set, then $\cup (a^+) = a$.*

Proof: $\cup (a^+) = \cup (a \cup \{a\})$ by defn. of a^+

$= \cup a \ \cup \ \cup \{a\}$

$= \cup a \ \cup \ a$

$= a$ because $\cup a \subset a$ for any transitive set a.

Lemma 2. *Every natural number n is a transitive set.*

Proof: By induction, let T be the set of natural numbers that are transitive sets, *i.e.*

$$T = \{x : x \in \omega \bullet \cup x \subset x\}$$

We observe that $0 \in T$, since for every x, $x \notin 0$, whence $x \in 0 \supset x \in T$.

Now for any natural number n, we must show $n \in T \supset n^+ \in T$

Suppose $n \in T$, then by Lemma 1, $\cup\,(n^+) = n$

Since $n^+ = n\ \cup \{n\}$, $n \in n^+$ and so $\cup\,(n^+) \in n^+$

Therefore $n^+ \in T$.

Now by induction , $T = \omega$ *i.e.* every natural number is a transitive set.

Having established his two lemmas, Copi is just steps away from proving P3.

Proof: If $n^+ = m^+$, then $\cup\,(n^+) = \cup\,(m^+)$

Since both n and m are in ω both n and m are transitive sets, by Lemma 2.

By Lemma 1, $\cup\,(n^+) = n$ and $\cup\,(m^+) = m$, therefore $n = m$ (p. 194)

Having back-checked that the von Neumann sets satisfy the Peano axioms we can verify that they have the same arithmetic properties of the natural numbers. As Copi points out, they are not philosophically more plausible than the Frege-Russell constructive definition of numbers, which cannot be part of the **ZF** system, but are *instances* or representatives of them. Thus, every number n is equivalent to every n membered set, for which they are “admirably suited for *counting* the number of members… indeed they have been called ‘counter sets [by Quine]”. (*l.c.*)

Cardinal Numbers and the Axiom of Choice

Copi begins this section with a reminder of the following notions established so far: In the preceding section we gave a set theoretic definition of the natural numbers 0, 1, 2, 3… Each natural number is the number of members in some or other finite set and *is itself* a finite set with just so many members. Furthermore, the Axiom of Infinity guarantees the existence of the infinite set ω. At the end of the previous to last section it was proved that ω has fewer members than its power set $\wp\omega$ and that, in general, any set A has fewer members than its power set $\wp A$. As Copi observes, “There are then, infinitely may sets, each containing a different finite number of members, and also infinitely many more sets, each containing infinitely many members. Thus there is an unending ascension of sets, each larger than any that proceeds it”. (p. 195)

Obviously, the size or magnitude of a set, whether finite or infinite, is of interest and the **cardinality** of a set, given by a cardinal number, is a measure of its number of elements. The cardinality of a given set A is variously denoted as: ‘$|A|$’, ‘$n(A)$’, ‘$\bar{\bar{A}}$‘, ‘$\#A$’, ‘card (A)’ or just ‘card A’ as in the text. According to Copi, The essential nature of a **cardinal number** is given by

card $A =$ card $B \equiv A \sim B$ (p. 195)

Recall that '$A \sim B$' denotes sets between which there is a one-to-one correspondence, and which are thus said to be equinumerous. Alternatively, equinumerous sets are said to have the same cardinality. While the Frege-Russell definition of a cardinal number as the set of all equivalent or equinumerous sets is intuitively very appealing, it cannot be used in the **ZF** system. Instead recall, in the previous section, we identified finite cardinal numbers with special *representative* sets that the ZF axioms guarantee to exist, from among equivalent sets. The same can be done for infinite cardinal numbers. From among the equivalent sets of all perfect squares, rational fractions *etc.* we can select the set of all natural numbers, ω, as representative of them so that card ω is the number of members in all these equivalent sets. Using, the Hebrew letter 'aleph' sub-script zero, pronounced 'aleph-null', we define

$$\aleph_0 = df \text{ card } \omega$$

such that $\aleph_0$ is the smallest infinite cardinal number. If we use the Power Set Axiom we can produce even larger sets, each uniquely representative of equivalent sets and itself, so that there are infinitely many cardinal numbers. (p. 196)

The arithmetic operations on infinite cardinal numbers are defined in the same way as for those of finite cardinal numbers; however the former lead to some familiar and unfamiliar outcomes. Consider addition: The sum of two numbers, m and n, for example, is the number p if and only if there exist M and N which are disjoint sets (having no elements in common) such that m = card M and n = card N and p = card ($M \cup N$). We can therefore define

$$m + n = df \text{ card } (M \cup N)$$

with the same provisos. Defined this way addition is independent of the particular disjoint sets M and N. Indeed we could have used a different pair of disjoint sets M' and N' such that if $M \sim M'$ and $N \sim N'$ then $(M \cup N) \sim (M' \cup N')$. The commutation and association laws for $+$ follow directly from those for ∪. This is true for arithmetic addition of both finite and infinite cardinals. For any finite numbers m and n, $m + n = m$ only if n = 0, for infinite cardinal numbers however this is not true. Take the set of natural numbers ω, which can be broken down into a finite part, {0; 1; 2; … n-1} and an infinite part, {n; n+1; n+2; … } such that ω = {0; 1; 2; … n-1} ∪ {n; n+1; n+2; … }. Now the first **summand** (a quantity to be added to another) has a cardinality of n, while the second has a cardinality of $\aleph_0$, hence for any n

$$n + \aleph_0 = \aleph_0 + n = \aleph_0 \quad \ldots \text{①} \qquad (\textit{l.c.})$$

A second difference is this: the sum of two finite numbers $\neq 0$ is always larger than either summand; however this is not the case with summands whose cardinality is $\aleph_0$. Take the odd numbers {1; 3; 5; 7… } whose cardinality is $\aleph_0$ and the even numbers {2; 4; 6; 8 … } whose cardinality is also $\aleph_0$. Their union is the set of natural numbers, ω whose cardinality is also $\aleph_0$. Thus, we have

$$\aleph_0 + \aleph_0 = \aleph_0 \quad \ldots \text{②} \qquad (\text{p. } 197)$$

Thirdly, for any finite numbers m and n, such that $m \leq n$, subtracting m from n leads to a unique result; however for infinite cardinal numbers this is not so. From ① and ② above we can show that

$\aleph_0 - \aleph_0$ can equal any cardinal number from 0 to $\aleph_0$. (*l.c.*)

Arithmetic addition is also defined for infinitely many summands. If $\{m_i\}$ is a set of infinitely many cardinal numbers and $\{M_i\}$ is a set of infinitely many sets, such that m_i = card M_i and $i \neq j \supset M_i \cap M_j = \emptyset$, then by definition, for every i and j

$$\sum_i m_i = \text{card} \bigcup_i M_i$$

Copi provides the following by way of example, "the denumerable set of all natural numbers can be decomposed into denumerably many sets, each of which contains denumerably many natural numbers, as in the following (diagonal) array:

1	2	4	7	…………….
3	5	8	………………………….	
6	9	……………………………………….		
10	………………………………………………………			
………………………………………………………………..				

which shows that

$\aleph_0 + \aleph_0 + \aleph_0 + \cdots = \aleph_0$ (p. 197)

The product of two numbers $m \bullet n$ (or simply mn) can be thought of in two ways: either by adding m to itself n times ($\sum_n m$) or by producing two sets M and N such that card M = m and card N = n and then using the Cartesian product to define

$m \bullet n = \text{card } (M \times N)$

This product is unique and independent of which sets M and N are used to form the product, which equals 0 if either M or N = $\emptyset$. The cardinal product $m \bullet n$ is also commutative (ab = ba), associative ($a(bc)$ = $(ab)c$) and distributive with respect to addition ($a(b$ + c) = ab + ac) for finite and infinite cases.

Copi however points out one complication in the case that multiplication involves infinitely may factors, even when the factors themselves are finite. Let $\mathfrak{M}$ be a set of infinitely many disjoint, nonempty, finite sets M_i, where m_i = card M_i and let $\prod_i m_i$ be the product we wish to define. "Intuitively this product is the number of distinct selection sets, μ, where each μ contains exactly one element from each of the sets M_i in $\mathfrak{M}$. In the finite case where [for example] $\mathfrak{M}$ = $\{\{a; b\}; \{c; d; e\}\}$, that is where M_1 = $\{a; b\}$ and M_2 = $\{c; d; e\}$, the distinct sets μ are $\{a; c\}$, $\{a; d\}$, $\{a; e\}$, $\{b; c\}$, $\{b; d\}$, $\{b; e\}$. There are six such sets, which is the expected product since 2 = card M_1 and 3 = card M_2". (p. 198)

The process that Copi used in forming distinct sets μ is straightforward: Take one member from M_1 and pair it with one member of M_2. That is our first pair set μ. Next, take one member from M_1 and pair it with a different member of M_2. That is our second pair set μ … and so on. The problem with infinite sets is that an arbitrary choice must be made from among the members of the infinitely may sets M_i of $\mathfrak{M}$ unless each M_i already has its members ordered in some way (or happens to be a singleton). Of course we could come up with any rule or algorithm by which to order sets, *e.g.*

alphabetically, numerically, by size *etc.* which is satisfactory for finite sets but "somewhat fanciful" (to use Copi's term) for making an infinite sequence of arbitrary choices. What is needed is another axiom that guarantees that there exists a selection set μ that contains exactly one member from each of any number of sets M_i. (p. 198)

ZF-8 **Axiom of Choice**: If $\mathfrak{M}$ is a set whose elements are all sets that are different from $\emptyset$ and mutually disjoint, its union $\cup\, \mathfrak{M}$ includes at least one subset μ having one and only one element in common with each element of $\mathfrak{M}$.

According to Wikipedia: Axiom of choice, if we think about set as bins and elements as objects that can be put in bins then, "informally put, the axiom of choice says that given any collection of bins, each containing at least one object, it is possible to make a selection of exactly one object from each bin. In many cases such a selection can be made without invoking the axiom of choice; this is in particular the case if the number of bins is finite, or if a selection rule is available: a distinguishing property that happens to hold for exactly one object in each bin".

Russell's ingenious example of a mythical millionaire who accumulated $\aleph_0$ pairs of shoes and $\aleph_0$ pairs of socks illustrates why the answers to the seemingly innocuous questions, "How many pairs of shoes does he own?" and "How many pairs of socks does he own?" could not be answered in the same way. In the case of shoes, they are distinguished left form right, so it is easy to make a selection of one out of each pair, even if there are infinitely many, by choosing all the left ones or all the right ones. Pairs of socks however are not so distinguished, so for an infinite collection of socks we cannot be sure that we have chosen one out of each pair unless we invoke the Axiom of Choice. (p. 198 - 199)

According to Russel,

> We may put the matter in another way. To prove that a class has $\aleph_0$ terms, it is necessary and sufficient to find some way of arranging its terms in a progression. There is no difficulty in doing this with the boots. The *pairs* are given as forming an $\aleph_0$, and therefore as the field of a progression. With each pair, take the left boot first and the right second, keeping the order of the pair unchanged; in this way we obtain a progression of all the boots. But with socks we shall have to choose arbitrarily, with each pair, which to put first; and an infinite number of arbitrary choices is an impossibility. Unless we can find a *rule* for selecting, *i.e.* a relation which is a selector, we do not know that a selection is even theoretically possible. (Reproduced in Copi, p. 199)

As soon as we have one subset μ we can generate another by taking the relative compliment of each set M_i - μ, in turn. Appealing to the Axiom of Choice again we can find another μ, and so on, until we have all the selection sets μ in $\cup\, \mathfrak{M}$. According to Copi, this legitimises the definition:

> If $\{m_i\}$ is a set of cardinal numbers, and if $\{M_i\}$ is a corresponding set of sets, such that card $M_i = m_i$ for each i, then

$$\prod_i m_i = \text{card } (\times_i M_i)$$

The same definition can be written in the multiply quantified logical notation with which we have become familiar, however this is rather cumbersome.[1] It is enough, however that we remember that the above formulation is actually just a shorthand symbolisation.

Copi introduces an alternative version of the Axiom of Choice as follows:

Axiom of Choice (Alternative version): For any set M there is a function f (a 'choice function' for M) such that the domain of f is the set of nonempty subsets of M and $f(S) \in S$ for every nonempty $S \subset M$.

If we wish to define the product of cardinal numbers m_i, this version of Axiom of Choice ensures that there exists a choice function f for the set $\cup \mathfrak{M}$ which selects an element x_i from each nonempty subset of M_i of $\cup \mathfrak{M}$; each of which determines a singleton $\{x_i\}$, such that the union of all of the latter is a selection set μ. (p. 199)

Although the Axiom of choice cannot be derived from the other **ZF** axioms, Kurt Gödel showed that it can be consistently added to the others and that its negation is not a theorem of **ZF**. According to Wikipedia: Axiom of choice, "Despite […] seemingly paradoxical facts, most mathematicians accept the axiom of choice as a valid principle for proving new results in mathematics". (See the section on Criticism and acceptance.)

Raising a cardinal number to an exponent in the finite case is analogous to multiplication: "just as multiplication involves the addition of equal summands, so **exponentiation** involves the multiplication of equal factors". From High School algebra we recall that for finite, positive cardinal numbers a, b, c and d: b^a is the same as b multiplied by itself a times over, from which it follows that

$b^{a+c} = b^a \cdot b^c$

$(b \cdot d)^a = b^a \cdot d^a$

$(b^a)^c = (b^c)^a = b^{a \cdot c}$

$b^0 = 1$ where $b \neq 0$ and

$0^a = 0$ where $a \neq 0$ (p. 200)

At this point Copi introduces the following notation for certain sets of functions.

$B^A = df\ \{f : f$ is a function on A to $B\}$

According to Copi, "This notation is intended to reveal and exploit the analogy with the familiar notation for exponentiation". The following examples reveal the analogy: If $B = \{b_1;\ b_2; \dots; b_n\}$ then n = card B and if we let $A = \{a_1\}$ then card A = 1. In this case there are exactly n functions in B^A: $f_1(a_1) = b_1$; $f_2(a_1) = b_2$; … ; $f_n(a_1) = b_n$. So in this case card B^A = (card B)$^{\text{card}\, A}$. In the case that we let $A = \{a_1;\ a_2\}$ for the same B, card A = 2 and there are n^2 functions in B^A. They are

[1] $\prod_i m_i = df$ card $\Big\{\mu: (\forall m_i)(\forall m_j)(\exists M_i)(\exists M_j)\Big[m_i \text{ card} = \text{card } M_i \bullet m_j \text{ card} = M_j \bullet (\forall i)(\forall j)\big(i \neq j \supset M_i \cap M_j = \emptyset\big) \bullet (\exists x)\Big(x \in M_i \bullet x \in \mu \bullet (\forall y)\big((y \in M_i \bullet y \in \mu) \supset y = x\big)\Big)\Big]\Big\}$

$f_{11}(a_1) = b_1$ $f_{12}(a_1) = b_1$ $f_{13}(a_1) = b_1$ … $f_{1n}(a_1) = b_1$

$f_{11}(a_2) = b_2$ $f_{12}(a_2) = b_2$ $f_{13}(a_2) = b_3$ … $f_{1n}(a_2) = b_n$

$f_{21}(a_1) = b_2$ $f_{22}(a_1) = b_2$ $f_{23}(a_1) = b_2$ … $f_{2n}(a_1) = b_2$

$f_{21}(a_2) = b_1$ $f_{22}(a_2) = b_2$ $f_{23}(a_2) = b_3$ … $f_{2n}(a_2) = b_n$

… … … … …

$f_{n1}(a_1) = b_n$ $f_{n2}(a_1) = b_n$ $f_{n3}(a_1) = b_n$ … $f_{nn}(a_1) = b_n$

$f_{n1}(a_2) = b_1$ $f_{n2}(a_2) = b_2$ $f_{n3}(a_2) = b_3$ … $f_{nn}(a_2) = b_n$

Here card B^A = (card B)^2 = (card B)$^{\text{card } A}$. In general therefore, for any nonempty A and B

card (B^A) = (card B)$^{\text{card } A}$ (p. 200)

From this conclusion Copi extends the definition of exponentiation to cover all cardinal numbers a and b, whether finite or infinite, as

$b^a = df$ card (B^A) where $a =$ card A and $b =$ card B

Copi points out that cardinal addition, cardinal multiplication and cardinal exponentiation are analogous in the following way, "Just as cardinal addition is based on the union of disjoint sets that have the same cardinal number being added, and just as cardinal multiplication is based on the Cartesian product of sets having the same cardinal numbers being multiplied, so cardinal exponentiation is based on mappings of one set onto another, where the range has the cardinal number that is raised to a power and the domain has the cardinal number that is the exponent". Therefore, for any positive infinite cardinal numbers a, b, c and d, the same equalities obtain as for finite cardinal numbers. Therefore $\aleph_0^0$ = 1 and for any positive finite a

$\aleph_0^a = \aleph_0 \cdot \aleph_0 \cdot \aleph_0 \cdot \ldots \cdot \aleph_0 = \aleph_0$ (*l.c.*)

A **characteristic function** (or indicator function) is a function defined on a set A that indicates membership of an element in a subset A' of A which has a value of 1 for all elements of A' and a value of 0 for all elements of A not in A'. (Wikipedia: Indicator function) As Copi explains, if the B in B^A is 2 = {1; 2} then we have all the set of all characteristic functions of subsets of A according to the following definition

For any subset $A' \subset A$ the characteristic function $F_{A'}$: $A \to 2$ is such that

$(\forall x)(x \in A' \supset f_{A'}(x) = 1)$ and

$(\forall x)(x \in (A - A') \supset f_{A'}(x) = 0)$

The characteristic function $f_{A'}$ for each subset A' of A can be thought of as a condition on members of A, such that every member x of A for which $f_{A'}$ = 1 satisfies the condition and for every member x of A for which $f_{A'}$ = 0 does not satisfy the condition. By the Axiom of Separation, any condition

$f_{A'}(x)$ on members of A determines a unique subset $A' \subset A$. Conversely, any subset $A' \subset A$ determines a characteristic function (or condition) $f_{A'}(x)$. (p. 201)

On reflection, it should be easy to see that the set of all characteristic functions of subsets of A is equivalent or equinumerous to all the subsets of A. Thus for any A, $2^A \sim \wp A$ and in particular $2^\omega \sim \wp\omega$. According to Copi, Cantor proved that for any set S, $S < \wp S$ and in particular that $\omega < \wp\omega$. Since we already know, via Cantor's diagonal proof, that $\omega < \{x : x \in \mathbb{R}$ such that $0 \leq x \leq 1\}$, so $\omega < \mathbb{R}$, where $\mathbb{R}$ is the set of real numbers. Therefore $\wp\omega$ and $\mathbb{R}$ are both greater than ω. (*l.c.*)

It can be proved (below) that the power set of natural numbers and the set of real numbers are equivalent or equinumerous, *i.e.* $\wp\omega \sim \mathbb{R}$. Moreover, counterintuitively as it may seem, there are as many real numbers over the interval 0 to 1 as there are real numbers altogether. So, just as with the diagonal proof, we can confine ourselves to the interval 0 to 1 in the following discussion, without loss of generality. Again, we can represent every such real number as a binary fraction $0.d_1d_2d_3$ ... where every d_i is one of two binary digits, 0 or 1. Thus, the number 1 is represented by 0.111…, 0 by 0.000…, ½ by both 0.100… and 0.011… and so on. (p. 201)

According to Copi, "Each binary fraction determines a unique subset of ω, namely, the subset that contains the natural number n, if and only if the [n^{th} digit] d_n = 1". Thus,

0.111… 1… determines the set ω

0.000… 0… determines the empty set

0.101… 1… determines the subset of ω that contains every natural number except 2

0.0101… 0101… determines the set of all even numbers where only every second d_n is a 1

and so on. Conversely, every subset A' of ω determines a binary fraction $0.d_1d_2d_3$ … , in which d_n = 1, if and only if, $n \in A'$." Thus we have a one-to-one correspondence between $\wp\omega$ on the one hand and $\{x : x \in \mathbb{R}$ such that $0 \leq x \leq 1\}$ on the other. Therefore $\wp\omega \sim \mathbb{R}$ and ($\wp\omega$) = card $\mathbb{R}$. The symbol for card $\mathbb{R}$, also known as the **cardinality of the continuum**, is '$\aleph$' (note: without a subscript). Since we know that $2^\omega \sim \wp\omega$ and that card (2^ω) = card ($\wp\omega$) and since card (2^ω) = (card 2)$^{\text{card}\omega}$ which is the same as $2^{\aleph_0}$, it follows that $2^{\aleph_0} = \aleph$. Finally, as with $\aleph_0$, we have it that

$$\aleph + \aleph = \aleph \cdot \aleph = \aleph \qquad \text{(p. 201 - 202)}$$

Ordinal Numbers and the Axioms of Replacement and Regularity

Copi begins this section with the following, seemingly obvious observation, "The elements of a set may be related to each other in a variety of ways: some may be smaller than others, some may be elements of others, some may be subsets of others, and so on". Any relation(s) that obtain between members can be thought of as imposing some kind of *order* on the set. Consider the following examples.

1. Let S be the power set $\wp\{a; b; c\}$ with the ordering relation of the set being *inclusion*, $\subset$.
2. Let P be the set of positive integers with the ordering relation being *divides without remainder*, d.

3. Let ω be the set of natural numbers with the ordering relation being *less than or equal to*, $\leq$.

We say that that each of these sets is *partially ordered* by the relation mentioned, where **partial order in a set X** is defined as being a reflexive, antisymmetric and transitive relation in X. (p. 202)

The relations of reflexivity, symmetry and transitivity were discussed in Critical Reasoning 14, however now we must draw a distinction between asymmetric and antisymmetric relations. Recall that a relation is **asymmetrical** if the following is true

$$(\forall x)(\forall y)(Rxy \supset {\sim}Ryx)$$

Paraphrasal: If x has a relation R to y, then y does not have a relation R to x.

E.g. "x is North of y", "x is the father of y" *etc.*

However a relation is **antisymmetric** if the following is true

$$(\forall x)(\forall y)\big[(x \in X \bullet y \in X) \supset \big((Rxy \ \bullet \ Ryx\ ) \supset x = y\big)\big]$$

Paraphrasal: There is no pair of distinct elements of X, each of which is related by R to the other.

E.g. "$x \leq y$" for x and $y \in \mathbb{R}$

Note that while every antisymmetric relation is also asymmetrical, not every asymmetrical relation is antisymmetric.

Of the three ordering relations above we can see that $\subset$ is a partial order in $\wp\{a; b; c\}$; d is a partial order in P; and $\leq$ is a partial order in ω. The first two sets however are *only* partially ordered by the relation mentioned because there are some distinct members that are not ordered by the relation. The third set, however is **totally ordered** or **simply ordered** because all distinct members are ordered by the relation mentioned. Note that all totally ordered sets are also partially ordered but not *vice versa.* (p. 203)

Copi defines a **well-ordered set** as a partially ordered set such that every non-empty subset of which has a first (least or smallest) element that is related to the ordering relation to every other element in the subset. The third set above is a well-ordered set under the ordering relation $\leq$; the first and second set are not well-ordered. In the first case, the subset of $\wp\{a; b; c\}$ comprising of only $\{a\}$ and $\{b\}$ does not contain a first element since neither $\{a\}$ contains $\{b\}$ nor $\{b\}$ contains $\{a\}$. In the second case, the subset of P comprising only of 2 and 3 does not contain a first member since neither 2 nor 3 divides the other without remainder. Of course, the first and second sets may be well-ordered under another ordering relation. In fact the first set may be well-ordered alphabetically, for example. This however only emphasises Copi's point that "a set may be well-ordered by one relation, but not by another." (*l.c.*)

In elaborating the theory of well-ordered sets Copi finds it convenient to work with the narrower relation of precedence $<$, rather than the disjunctive $\leq$. The former is irreflexive and asymmetrical, (rather than antisymmetric) but it is also transitive. It should be observed that well-ordered sets are *connected* by the relation that well-orders them. Thus, for any two distinct members ($x_1 \neq x_2$) of a well-ordered set either $x_1 < x_2$ or $x_2 < x_1$. The well ordering condition is met for any doubleton

$\{x_1;\ x_2\}$ because it is a nonempty subset and therefore contains a least (or smallest) element. (p. 203)

According to Copi, "Two partially ordered sets are said to be **similar** if there is a one-to-one correspondence between them that preserves order" In other words, X and Y are similar, symbolised $X \cong Y$, if they are both partially ordered sets and there is a one-to-one correspondence f from X onto Y, such that for any x_1 and x_2 in X $x_1 \leq x_2$ if, and only if, $f(x_1) \leq f(x_2)$ in Y. Of course, similarity $\cong$ entails one-to-one correspondence $\sim$ but not the other way round. Similarity may apply to an infinite, well-ordered set and a proper subset of itself. Consider, for example, the function $m = 2n$, where there is a one-to-one correspondence between the set of natural numbers and the set of even numbers that preserves order in both sets. (*l.c.*)

Copi points out that in any such mapping f, from one well-ordered set X onto a subset of itself,

> $x \leq f(x)$ for every x in X

This may be proved by the method of *reductio ad absurdum*.

> Suppose that there is an x in X such that $f(x) < x$
>
> Since X is well-ordered, there must exit a least (or smallest) element x_0 in X
>
> But x_0 could not be the least (or smallest) element in X because $f(x_0) \in X$ and $f(x_0) < x_0$
>
> For any y in X, if $y < x_0$ then $y \leq f(y)$
>
> But $f(x_0)$ is just such a y, so $f(x_0) \leq f\big(f(x_0)\big)$
>
> But since $f(x_0) < x_0$ and since f is an order preserving mapping, $f\big(f(x_0)\big) < f(x_0)$
>
> Now by transitivity, $f(x_0) < f(x_0)$ which is a contradiction because $<$ is irreflexive
>
> So the supposition that there is an x in X such that $f(x) < x$, must be false. (p. 204)

In order to explain an important corollary of the above proof Copi introduces a new term, "If X is a well ordered set and $a \in X$ then the subset $\{x{:}\ x \in X \bullet x < a\}$ is called the **initial segment** determined by a, usually denoted by $s(a)$". The corollary in question can now be stated as a theorem.

> No well-ordered set can be similar to one of its initial segments

the proof of which is straightforward: If X is a well-ordered set with a as an element and if f is a similarity function from X onto $s(a)$, then $f(a) \in s(a)$ and hence $f(a) < a$, which according to the above proof is impossible.

According to Copi, many theorems about well-ordered sets can be proved by the principle of **transfinite induction** which is analogous to that of mathematical induction, discussed earlier. In fact, for all but infinite sets the two are equivalent. In the case of infinite well-ordered sets however the following principle is applicable:

> If S is a subset of a well-ordered set X, and if an element x of X belongs to S whenever the initial segment determined by x is included in S, then $S = X$

the proof of which is again straightforward: If $X - S$ were non-empty, then it would contain a least (or smallest) element; call it x. Therefore, every element of the initial segment $s(x)$ must belong in S and, by the principle of transfinite induction, $x \in S$ also. But this is impossible because x cannot belong to both S and to $X - S$. Therefore, $X - S$ must be empty and since S is included in X ($S \subset X$), S must be identical to X, ($S = X$). When applied to the natural numbers this principle is known as **strong induction**. (p. 204)

At this point Copi introduces some further terminology, "If two simply ordered sets are similar, they are said to be **isomorphic** (with each other). Ordered sets that are isomorphic with each other are said to have the same **order type**." In other words, two ordered sets X and Y are said to have the same order type if they are order isomorphic *i.e.* $f: X \rightarrow Y$ such that both f and its inverse are **monotonic** (preserving the given orders of elements). The most important order types are those of well ordered sets.

In ordinary parlance 'ordinal number' is an adjective describing the numerical position of an object, *e.g.* first, second, third *etc.* In formal set theory, according to several authors, including one cited by Copi, "an **ordinal number** is defined as the order type of a well ordered set."[2] However there is a problem accommodating within the **ZF** system, an order type of a given, nonempty, well-ordered set defined as the set of all well-ordered sets that are isomorphic with the given set. The problem is a familiar one because the union of a set of all similar, well-ordered sets would be the universal set, which we have proven cannot exist. Compare the difficulty of accommodating the Frege-Russel definition of cardinal number within the **ZF** system. The workaround in that situation was to let an unspecified *representative set* of equivalent sets serve as the cardinal number of any set equivalent to it. Just "*which* set would serve as *the* cardinal number was specified only for the finite case. [However] in dealing with ordinal numbers, we *can* specify which well ordered set will serve as the ordinal numbered of any well ordered set isomorphic with it". (p. 204 - 205 original emphasis.)

Recall from the discussion of cardinals that every natural number n is a subset of ω, which is a well-ordered set. n is therefore also a well-ordered set containing all the natural numbers less than n. Also recall that for any two natural numbers m and n, $m < n$ if, and only if, $m \in n$. Each n is therefore the set such that $m \in \omega$ and $m < n$. The ordinal number of any *finite*, well-ordered set is thus defined as the natural number that is the number of elements in that set. "Given any [two] natural numbers m and n with $m \in n$, the initial segment $s(m)$ of n determined by m is $\{x: x \in n \bullet x < m\}$. That is, every natural number is a well ordered set such that the initial segment determined by each element in it is the same as that element". (p. 205)

Copi uses this representation of natural numbers as ordinal numbers to give a generalised definition of ordinal numbers including infinite ordinals. Thus,

> α is an ordinal number $= df$ α is a well-ordered set such that $s(\xi) = \xi$ for every ξ in α.

The lower case Greek letter ξ is pronounced 'xi' as in 'pixie'. Like the smallest infinite (a.k.a. transfinite) cardinal, the smallest infinite ordinal covered by Copi's definition is ω. Using the same

[2] See http://mathworld.wolfram.com/OrdinalNumber.html

definition of successor set given previously as $u^+ = u \cup \{u\}$ we can allow u to range over ordinals so that we can generate the sets ω^+, $(\omega^+)^+$ and so on. These successor sets are also ordinals according to the above definition. As Copi demonstrates, "if α is an ordinal number, then so is α^+, which is $\alpha \cup \{\alpha\}$. For if $\xi \in a^+$, then either $\xi \in \alpha$ or $\xi \in \{a\}$. In case $\xi \in \alpha$, then since α is an ordinal number, $s(\xi) = \xi$; and in case $\xi \in \{a\}$, $\xi = \alpha$, in which case $s(\xi) = \alpha$, that is, $s(\alpha) = \alpha$. So in either case $s(\xi) = \xi$ for any $\xi \in a^+$, and α^+ is an ordinal number. (p. 205)

Various arithmetical operations can be performed with ordinals. When combining well-ordered sets we put one directly after another, in order, but this relies on the sets being disjoint. Suppose however that sets A and B that we want to combine are not disjoint. We can generate sets A' and B' that are disjoint such that $A \cong A'$ and $B \cong B'$. We can let A' be the set of all ordered pairs $\{a; 0\}$ with a in A and B' be the set of all ordered pairs $\{b; 1\}$ with b in B. Thus there is a one-to-one correspondence between a and $\{a; 0\}$ and between b and $\{b; 1\}$, "with the order within A' and B' simply borrowed from A and B." Given that A and B are now characterised as disjoint, well-ordered sets, allows us to define the order in $A \cup B$ in such a way that pairs of elements in A and pairs of elements in B keep their order and let each element of A precede each element of B. The union $A \cup B$ is known as the **ordinal sum** of A and B and derives *its* well-ordered nature from that of A and that of B, respectively. (p. 205 - 206)

Building on this we can define addition for ordinals as follows. If α an β are ordinals, we can let A and B be disjoint, well-ordered sets such that α = ord A and β = ord β, and we can let C be the ordinal sum of A and B. Therefore the sum $\alpha + \beta$ = ord C and ord A + ord B = ord ($A \cup B$). Copi points out that "the sum $\alpha + \beta$ is independent of the particular choice of sets A and B. Any disjoint pair of similar sets would give the same result" (p. 206)

Some of the properties of ordinal addition are similar to those for cardinal addition. These include

$\alpha + 0 = \alpha = 0 + \alpha$

$\alpha + 1 = \alpha^+$

$\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$ (Association)

$\alpha < \beta$ if, and only if, there exists an ordinal $\gamma \neq 0$, such that $\beta = \alpha + \gamma$

However the Law of Commutation fails for ordinal addition when at least one of the ordinals is infinite. Thus

$1 + \omega = \omega$ but $\omega + 1 \neq \omega$

In the case that we put a 1 before the infinite sequence ω the result is similar to the original sequence, however in the case that we put a 1 after the infinite sequence ω, the similarity is "gone" because the new set has a last element that the original set did not have. (*l.c.*)

Multiplication of ordinals can be thought of as analogous to that of natural numbers. The product of two well-ordered sets A and B can be conceived of as adding A to itself B times. "To do so, we must have B disjoint sets, each similar to A. These can be produced as well ordered sets $A_b = A \times \{b\}$ for each b in B. Then the set of all these disjoint A_b's, $\{A_b : b \in B\}$, has, as its sum, $\cup \{A_b : b \in B\}$, ordered in the following way: $(a; b) < (a'; b')$ if and only if either $b < b'$ or ($b = b'$ and $a < a'$)".

Copi then defines the **ordinal product** of two well-ordered sets A and B as the Cartesian product $A \times B$, ordered as above. The ordinal product of two ordinal numbers α and β can be defined by introducing sets A and B, such that ord A = α and ord B = β and introduce a set C which we let be the ordinal product of A and B. Then the ordinal product of the two ordinal numbers α and β is defined as ord C. Copi's next point is so obvious that it might easily have been overlooked, had he not explicitly mentioned it. "The easiest well ordered sets to introduce and use here as A and B are the ordinals numbers α and β themselves, since each is a well ordered set whose ordinal number is itself". (p. 206)

Some of the properties of ordinal multiplication are similar to those for cardinal multiplication. These include

$\alpha 0 = 0 = 0\alpha$

$\alpha 1 = \alpha = 1\alpha$

$\alpha(\beta\gamma) = (\alpha\beta)\gamma$ (Association)

$\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$ (Left Distribution)

if $\alpha\beta = 0$ then either $\alpha = 0$ or $\beta = 0$

However the Law of Commutation fails for ordinal multiplication, where at least one of the ordinals is infinite. Thus

$2\omega = \omega$, because this product is an infinite sequence of ordered pairs, but

$\omega 2 \neq \omega$, because this product is an ordered pair of infinite sequences.

The Right Law of Distribution also fails because, in general

$(\alpha + \beta)\gamma \neq \alpha\gamma + \beta\gamma$

e.g. $(1 + 1)\omega = 2\omega = \omega$ but $1\omega + 1\omega = \omega + \omega = \omega 2 \neq \omega$ (*l.c.*)

"Just as ordinal products were defined in terms of repeated addition, so ordinal exponentiation can be defined in terms of repeated multiplication." It is desirable that the following properties should be true by definition

$0^\alpha = 0$ for $\alpha \geq 1$

$1^\alpha = 1$

$\alpha^{\beta+\gamma} = \alpha^\beta \alpha^\gamma$ and

$\alpha^{\beta\gamma} = (\alpha^\beta)^\gamma$

However not all the laws of exponents hold true. In general

$(\alpha\beta)^\gamma \neq \alpha^\gamma \beta^\gamma$

e.g. $(2 \cdot 2)^\omega = 4^\omega$ but $2^\omega 2^\omega = \omega\omega = \omega^2$ (p. 207)

Copi adds a few more observations before concluding this discussion. If m is finite and > 1 then m^ω is the limit of the expanding product $m \cdot m \cdot m$ … "which is an infinite sequence of m-tuples and thus, ω. Alternatively, "m^ω is the limit of the set of all finite ordinals m^n for $n < \omega$, so again $m^\omega = \omega$. Similarly, ω^ω is the limit of the set of all powers ω^n for $n < \omega$. Since 1 + ω + ω^2 + … + ω^n = ω^n, we may also write ω^ω = 1 + ω + ω^2 + … + ω^k… ." (p. 207)

Next Copi addresses the question of whether all the ordinal numbers mentioned above exist within the ZF system. "Since ω is a set, we can use the Axiom of Pairing and the Union Axiom to form successor sets on the 'far' side of ω." Thus the successor of ω or ω^+ is $\omega \cup \{\omega\}$, which using the notation for ordinal addition already employed, is written as ω + 1. Next the successor to the successor of ω or $(\omega^+)^+$ is $\omega^+ \cup \{\omega^+\}$ which is written ω + 2, and so on. Obviously each of these exists as a set and as an ordinal number. However is there a larger set containing them all? Just as we needed to introduce the Axiom of Infinity to prove that there was a set of all natural numbers, so we now need to introduce another axiom, or rather axiom schema, in order to prove that there are sets of ordinal numbers beyond ω. Thus

ZF-9 **Axiom Schema of Replacement:** If $\phi(x; y)$ is a formula such that for each member x of the set X, $\phi(x; y)$ and $\phi(x; z)$ imply that $y = z$, then there exists a set S such that $y \in S$ if and only if there is an x in X such that $\phi(x; y)$. (*l.c.*)

The reasons for the somewhat convoluted statement of this axiom schema will become apparent below. However for now, if we let X be ω and let $\phi(x; y)$ be $y = \omega + x$, then according to the Axiom of Replacement there exists a set containing

ω; ω + 1; ω +2; …

The union of this set with ω is written $\omega 2$ which is also an ordinal number. After $\omega 2$ come the ordinal numbers $\omega 2$ + 1, $\omega 2$ + 2, $\omega 2$ + 3, and so on. After all of these we can use the Axiom of Replacement again to produce $\omega 3$, followed by $\omega 3$ + 1, $\omega 3$ + 2, $\omega 3$ …, all of which are ordinal numbers. By the same process again we can generate $\omega 4$, $\omega 5$, $\omega 6$ and so on. We can then use the Axiom of Replacement on the sequence ω, $\omega 1$, $\omega 2$… to obtain the ordinal number ω^2, which is related to the sequence in the same way that ω is related to the natural numbers. And so we can go on to produce further transfinite numbers, however some new notation is required. The next ordinal number after the sequence

ω^ω; $\omega^{(\omega^\omega)}$; $\omega^{\omega^{(\omega^\omega)}}$; … is written ε_0. (p. 207 - 208)

At this point Copi presents three examples of transfinite well-ordered sets that comprise of familiar elements.

- The set of all fractions $\frac{2^n-1}{2^n}$ and 1 ordered by the relation '$<$', which looks like

  0; 1/2; 3/4; 7/8; … ; 1

  the initial segment $s(1)$ of which is similar to ω and has an ordinal number is ω + 1.

- The set of all odd numbers in order of magnitude followed by the set of all even numbers in order of magnitude, which looks like

  1; 3; 5; … ; 2; 4; 6; …

  has an ordinal number twice as large as ω *i.e.* $\omega + \omega = \omega \cdot 2$.

- The set of all positive integer powers of prime numbers p^n ordered by the relation '<' defined as:

  $p_m^i < p_n^j$ if either $p_m < p_n$ or both $p_m = p_n$ and $i < j$.

  This set looks like

  2; 4; 8; 16; … ; 3; 9; 27; 81; … ; 5; 25; 125; 625; … ; …

  and has an even larger ordinal number *i.e.* $\omega \cdot \omega$ or ω^2. (p. 208)

Copi offers some clarifications and remarks about the Axiom of Replacement. Firstly, the condition that $\phi(x; y)$ and $\phi(x; z)$ imply that $y = z$ is to prevent the specification of $\phi(x; y)$ as $x \subset y$. If that were allowed then, since the empty set is a member of every set, taking X as $\{\emptyset\}$ would make the guaranteed set S contain all sets, which, as we have seen, is impossible. Secondly, invoking an expression like $\phi(x; y)$ might seem superfluous when we could just as well express the specifying function as f such that $f(x) = y$. Yet a function is well defined only if its domain is a set and a set can be specified to contain its range. However the purpose of the Axiom of Replacement, remember, is precisely to produce a set that *is* the range of f. Therefore, if we could simply take the latter for granted we would have no need of the axiom. (*l.c.*)

According to Copi, the Axiom of Replacement is "enormously powerful". In fact, the Axiom of Separation can be derived from it. Given any set A and any condition $\phi(x)$, if we let the X in the Axiom of Replacement be A and specify the condition $\phi(x; y)$ to be ($x = y$; $\phi(x)$) then the '$\phi(x; y)$ and $\phi(x; z)$ imply that $y = z$' part of the hypothesis of the Axiom of Replacement is satisfied. Apart from the trivially true '$x = x$', the conclusion then asserts

$$(\exists S)(\forall x)(x \in S \equiv x \in A \bullet \varphi x)$$

which is the Axiom of Separation encountered earlier. (*l.c.*)

The Axiom of Pairing can also be derived from the Axiom of Replacement together with the Power Set Axiom. Firstly, let set A above be the power set $\wp\wp\{\emptyset\}$, *i.e.* $\{\emptyset; \{\emptyset\}\}$. Then, if a and b are two objects whose pair set we desire, we specify the condition

$\phi(x; y)$ to be ($x = \emptyset \bullet y = a$) v ($x = \{\emptyset\} \bullet y = b$).

So for each x in $\wp\wp\{\emptyset\}$ there is exactly one y such that $\phi(x; y)$

i.e. for $x = \emptyset$, $y = a$ and for $x = \{\emptyset\}$, $y = b$.

These satisfy the hypothesis of the Axiom of Replacement so that its conclusion asserts

$$(\exists S)(\forall x)(x \in S \equiv x = a \vee x = b)$$

which is the symbolic statement of the Axiom of Pairing. (p. 208- 209)

We have already listed parts of the ascending series of ordinal numbers in the discussions above; however it is helpful to record them in one place as Copi does. Thus we have:

$0; 1; 2; \ldots ; \omega; \omega + 1; \omega +2; \ldots ; \omega 2 ; \omega 2 + 1; \omega 2 + 2; \ldots ; \omega 3; \ldots ; \omega 4; \ldots ; \omega^2; \omega^2 + 1; \omega^2 + 2; \ldots ; \omega^2 + \omega; \omega^2 + \omega + 1; \omega^2 + \omega + 2; \ldots ; \omega^2 + \omega 2; \omega^2 + \omega 2 + 1; \omega^2 + \omega 2 + 2; \ldots ; \omega^2 + \omega 3; \ldots ; \omega^2 + \omega 4; \ldots ; \omega^2 2; \ldots ; \omega^2 3; \ldots ; \omega^3; \ldots ; \omega^4; \ldots \omega^\omega; \ldots ; \omega^{(\omega^\omega)}; \ldots; \omega^{\omega^{(\omega^\omega)}}; \ldots ; \varepsilon_0; \varepsilon_0 + 1; \varepsilon_0 + 2; \ldots ; \varepsilon_0 + \omega; \ldots ; \varepsilon_0 + \omega 2; \ldots \varepsilon_0 + \omega^2; \ldots \varepsilon_0 + \omega^\omega; \ldots ; \varepsilon_0 2; \ldots ; \varepsilon_0 \omega; \ldots ; \varepsilon_0 \omega^\omega; \ldots ; {\varepsilon_0}^2; \ldots$

According to Copi, "Up to any point in this ascension, the set of all ordinal numbers, to that point is a well ordered set, W, which has an ordinal number greater than any member of W. If there *were* a set of *all* ordinal numbers, Ω, then it would have an ordinal number greater than any ordinal number *in* Ω, that is an ordinal number greater than any ordinal". This is known as the **Burali-Forti Paradox** which proves that there can be no set of *all* cardinal numbers. (p. 209)

Under the section on Cardinal Numbers and the Axiom of Choice we observed that, just as we can identify finite cardinal numbers with special *representative* sets that the **ZF** axioms guarantee to exist, from among equivalent sets, the same can be done for infinite cardinal numbers. For good reason Copi postponed their definition to the present section; however we are now in a position to specify just *which sets* are to serve as their special representatives. According to a theorem proved by Zermelo, using the Axiom of Choice, every set can be well ordered. Using the Power Set Axiom repeatedly we can form an endless series of sets of ever larger cardinality. So by Zermelo's theorem every set is equivalent to some ordinal number. (*l.c.*)

Copi points out that an infinite set can be equivalent to a great many different ordinal numbers. (In fact, all the transfinite sets above have a cardinality of 'only' $\aleph_0$.) However, of all the ordinal numbers equivalent to a given set S, themselves form a set. This can be shown as follows: The power set of S, *i.e.* $\wp S$, obviously has a greater cardinality than that of S. Therefore, any ordinal number σ that is equivalent to $\wp S$ must also have a greater cardinality than that of S as well as all the ordinal numbers equivalent to S. Now for every ordinal number β that is less than σ is also a member of σ, therefore σ contains every ordinal number equivalent to S. (p. 209 -210)

Using the Axiom of Separation we can obtain a subset of σ, call it σ' that contains all and only those ordinal numbers equivalent to S. Because σ' is a well-ordered set, it must contain a smallest, or least element, call it α, which can serve as *the* representative set equivalent to S and as the cardinal number of S. Copi therefore defines a *cardinal number* as "an ordinal number α such that if β is another ordinal number equivalent to α, then $\alpha < \beta$…. [T]his definition accords with the essential characteristic of cardinal numbers, that card A = card B if and only if $A \sim B$." (p. 210)

However, Copi advises that powers and products of ordinals must not be interpreted cardinally. "If A and B are well ordered sets, then in general,

$$\text{card}\left((\text{ord}B)^{\text{ord}\,A}\right) < (\text{card}\,B)^{\text{card}\,A}$$

The following examples are provided: For 2^ω, we have

$$\text{card}\,((\text{ord}\,2)^\omega) = \text{card}\,\omega = \aleph_0, \text{ but } (\text{card}\,2)^{\text{card}\,\omega} = 2^{\aleph_0} = \aleph, \text{ where } \aleph_0 < \aleph$$

And for ω^{ω}, we have

$$(\text{card}\ \omega)^{\text{card}\ \omega} = \aleph_0{}^{\aleph_0} = \aleph, \text{ but } \text{card}\big((\text{ord}\ \omega)^{\text{ord}\ \omega}\big) = \aleph_0$$

The reason for this being that all positive integers can be arranged into a well-ordered set whose ordinal number is ω^{ω}, and whose cardinality is $\aleph_0$. (p. 210)

The non-denumerable set of real numbers $\mathbb{R}$ however has a cardinality of $\aleph$ and an ordinality beyond anything we have so far encountered, including all the ordinal numbers on the previous page. According to Copi, there are ordinal numbers with non-denumerably many members. The set of all ordinal numbers up to a point beyond the ordinal number of the real numbers is one such well-ordered set. As such it has a smallest, or least element, symbolised as ω_1, which is the least ordinal number with non-denumerable cardinality, the cardinal number of which is $\aleph_1$. Then there is a whole sequence of alephs with ever increasing cardinality, with cardinal numbers

$$\aleph_0; \aleph_1; \dots ; \aleph_n; \dots ; \aleph_\omega; \aleph_{\omega+1} ; \dots ; \aleph_{\omega 2}; \dots \qquad (\textit{l.c.})$$

From the discussion above we know that $2^{\aleph_0} = \aleph$ and that $\aleph_1$ is the least cardinal greater than $\aleph_0$. From this we can deduce that $\aleph_1 \leq \aleph$. According to the **continuum hypothesis** proposed by Cantor in 1878

> There is no set whose cardinality is strictly between that of the integers and the real numbers.

In other words, there is no transfinite cardinal between the cardinal of the set of positive integers and that of the set of real numbers. Therefore either $\aleph_1 = \aleph$ or $\aleph_1 = 2^{\aleph_0}$. According to the **generalised continuum hypothesis**, for every ordinal number α

$$2^{\aleph_\alpha} = \aleph_{\alpha+1}$$

In other words, the cardinality of the power set of any infinite set is the smallest cardinality greater than that of the set. According to Copi, various mathematicians have attempted to prove the continuum hypothesis and in 1939 Gödel proved that, if the axioms of set theory are consistent, then the continuum hypothesis cannot be *disproved*. However in 1963 Paul Cohen proved that the continuum hypothesis cannot be *proved* from the axioms of **ZF** set theory either. Therefore the answer to this problem is independent of **ZF** set theory. (Wikipedia: Continuum hypothesis; Copi, *l.c.*)

It is believed that the nine axioms so far listed are sufficient to construct all of mathematics; therefore in a sense they are all that we require to that end. However it may be desirable to add, not so much another axiom, rather than a restriction stating that there are no sets other than those required or derivable from those axioms already listed. This is partially motivated by Mirimanoff's paradox which involves the concept of a **grounded set**, defined such that for any set x, for which there is no sequence of (not necessarily distinct) sets y_1, y_2, y_3, … such that … $\in y_3 \in y_2 \in y_1 \in x$. If we let W be the set of all grounded sets, then if W is grounded then $W \in W$ and $W \in W \in W$, and so on, then W is not grounded. If however W *is not* grounded then there is a sequence of sets y_1, y_2, y_3, … such that … $\in y_3 \in y_2 \in y_1 \in W$, hence y_1 is not grounded and cannot be a member of W. Therefore in either case W leads to a contradiction. According to Copi, the limited comprehension built into the Axiom of Separation should prevent the formation of sets such as W; however the idea of a nonempty set with no ultimate members other than an infinitely descending sequence of sets that terminates nowhere is repugnant to reason. Although we do need to distinguish between a singleton $\{x\}$ and its only member, x, the idea of a set being a member of itself is dubious. Recall that the limitation of comprehension imposed by the Axiom of Separation was to prevent Russell's paradox; however without it every $x \in x$ would lead to Russell's paradox. According to Copi, "This paradox can be generalized into infinitely many others involving what might be called '∈-cycles': $x \in y \bullet y \in x$, $x \in y \bullet y \in z \bullet z \in x$, and so on. Such ungrounded sets, self-membered sets, or ∈-cycle sets, all seem both repugnant to our thinking and of no possible utility in mathematics or logic". All of these "undesirables" are ruled out *a fiat* by the final **ZF** axiom. (p. 211)

ZF-10 Axiom of Regularity: Every nonempty set A contains an element b such that $A \cap b = \emptyset$.

How this axiom rules out these undesirable sets can be readily shown.

1. If A is a set, then it is not a member of itself, $A \notin A$. If A is an empty set then it has no members, therefore $A \notin A$. If A is not an empty set, then by the Axiom of Regularity, it contains an element x such that $\{A\} \cap x = \emptyset$. Since $\{A\}$ is a singleton, only $A \in \{A\}$ which means that $x = A$ and $\{A\} \cap A = \emptyset$. However since $A \in \{A\}$ it follows that $A \notin A$. (*l.c.*)

2. There are no sets A and B such that $A \in B \bullet B \in A$. Suppose by *reductio ad absurdum* that there're were such sets, then we would have

   $A \in \{A; B\} \cap B$ and $B \in \{A; B\} \cap A$ …①

   By the Axiom of Regularity there must be an x in $\{A; B\}$, such that $\{A; B\} \cap x = \emptyset$. But since $\{A; B\}$ is a doubleton of A and B, either $x = A$ or $x = B$. Therefore, either

   $\{A; B\} \cap A = \emptyset$ or $\{A; B\} \cap B = \emptyset$

   which contradicts …①. Therefore our supposition must be false. (*l.c.*)

3. There is no ungrounded set y_0 such that there exists a sequence of sets y_1, y_2, y_3, … such that … $\in y_3 \in y_2 \in y_1 \in y_0$. Suppose again by *reductio ad absurdum* that if there were such a set y_0 and a sequence of sets y_1, y_2, y_3, … then by the Axiom of Replacement, we could let ω be our set X and $y_{x^+} \in y_x$ be the formula $\phi(x; y)$. Then the condition that for each member x of the set ω, $\phi(x; y)$ and $\phi(x; z)$ implies that $y = z$ must be the case because we have already shown that there can be no ∈-cycles. Therefore the sets y_1, y_2, y_3, … must all be distinct. Recall that if the Axiom of Replacement hypothesis is true, then it would guarantee the existence of a set $R = \{x | x = y_n \bullet n \in \omega\}$. Hence, there would be a function f with ω as its domain and R as its range, where $f(n^+) \in f(n)$ for each $n \in \omega$. Thus any member r of R would have to equal $f(n)$ for some n so that $f(n^+) \in R \cap r$. This implies that for every r in R, $R \cap r \neq \emptyset$, which is a direct contradiction of the Axiom of Regularity. (p. 211 - 212)

This concludes Copi's very succinct yet rigorous chapter on set theory from its logical beginnings to its outline of the ZF system still in use today.

Tasks

Unlike most other chapters Copi did not provide any exercises for this one. He did however invite the interested reader to attempt to deduce some of the twenty theorems on pages 175 to 176 from the ten axioms of Boolean algebra listed above on page 175; however not all of them are as "fairly easy to derive" as Copi maintained, especially given some of the ontological assumptions involved, such as the non-existence of the universal set, discussed afterwards.

We have not set any proofs involving the **ZF** system of axioms because our task is not to teach mathematics. However we do encourage the reader to think about the relation between set theory, logic and mathematics and philosophy in general. It has sometimes been claimed that mathematicians can do without logic once the foundations of set theory have been formally established. Is this true, tenable or desirable? What is the ontological status of a set? What about the empty set?

Feedback

It is true that once a sufficiently expressive theory of sets has been established, sets can then be used to define or represent the basic logical connectives in terms of which all logical operations can be derived. However, every axiom of set theory, including the **ZF** system, is stated as a logical expression regarded as self-evidently true or true by definition. When theorems of set theory are proved on the basis these axioms they are done so by proofs of logic. Indeed some axioms of set theory follow logically from others, though they need not so long as they are consistent with each other.

As we have seen, according to the Axiom of Extensionality, a set is defined or determined by its members; however the empty set is guaranteed to exist by the Empty Set Axiom. That the empty set has no members does not mean that it does not exist as a concept. Just as the number zero counts no objects, does not mean that it does not exist. Indeed the discovery of the number zero must count as one of humanities greatest intellectual achievements. Therefore when (some) mathematicians claim that their subject requires no justification because it assumes only the empty set, we should be weary. The empty set itself is a concept that requires logic to define and further axioms to use iteratively. Therefore, we believe that any effort to divorce logic from set theory or subsume one by the other is ill-conceived and undesirable.

References

CANTOR, G. (1874) Ueber eine Eigenschaft des Inbegriffes aller reellen algebraischen Zahlen. *Journal für die reine und angewandte Mathematik* **77** pp**.** 258 - 262

CANTOR, G. (1892) Ueber eine elementare Frage der Mannigfaltigkeitslehre. *Jahresbericht der Deutschen Mathematiker-Vereinigung* (1890 - 1891) **1** pp. 75 - 78

COPI, I.M. (1979) Symbolic Logic, 5th Edition. Macmillan: New York