1. The Zermelo Fraenkel Axioms of Set Theory

The naive definition of a set as a collection of objects is unsatisfactory: The objects within a set may themselves be sets, whose elements are also sets, etc. This leads to an infinite regression. Should this be allowed? If the answer is "yes", then such a set certainly would not meet our intuitive expectations of a set. In particular, a set for which $A \in A$ holds contradicts our intuition about a set. This could be formulated as a first axiom towards a formal approach towards sets. It will be a later, and not very essential axiom. Not essential because $A \in A$ will not lead to a logical contradiction. This is different from Russel's paradox:

Let us assume that there is something like a universe of all sets. Given the property $A \notin A$ it should define a set R of all sets for which this property holds. Thus $R = \{A|A \notin A\}$. Is R such a set A? Of course, we must be able to ask this question. Notice, if R is one of the $A's$ it must satisfy the condition $R \notin R$. But by the very definition of R, namely $R = \{A|A \notin A\}$ we get $R \in R$ iff $R \notin R$. Of course, this is a logical contradiction. But how can we resolve it? The answer will be given by an axiom that restricts the definition of sets by properties. The scope of a property $P(A)$, like $P(A) \equiv A \notin A$, must itself be a set, say B. Thus, instead of $R = \{A|A \notin A\}$ we define for any set B the set $R(B) = \{A|A \in B, A \notin A\}$. This resolves the Russel paradox to the statement that for any set B, one has that $R(B) \notin B$.

In axiomatic set theory we will formalize relationships between abstract objects, called sets. The only relation we are dealing with is the membership relation. For specified sets x and y, we either have $x \in y$ or $x \notin y$. Our axioms will also guarantee the existence of certain sets, like the empty set $\emptyset$ (which will be defined as the number zero), as well as the existence of the set ω of all natural numbers. We may think that all sets are within a universe, a mathematical structure $(\mathcal{U}, \in)$ for which the axiomas apply. Any axiomatic approach requires specification of a language. The axioms then will be certain expressions within that language. We will define a "first-order" language for set theory. Axiomatic set theory then is a collection of axioms on which the rules of logic are applied in order to draw further conclusions.

Notice that we use lower case letters for sets. We do this because the elements of a set will also be sets. Thus a distinction of a set A and its elements a no longer applies. We will use upper case letters, like R and P for properties that define sets.

We will now present the axioms and derive the most basic elements of set theory.

> **The Axiom of Extensionality:** If every element of the set a is an element of the set b and every element of the set b is an element of the set a, then $a = b$.

In other words, two sets are equal iff they contain the same elements. This should not be considered as a definition of equality of sets. Equality is an undefined, primitive relation and clearly, equal sets have the same elements. The axiom of extensionality merely states a condition on the relation $\in$. We may formalize extensionality:

$$\forall x \forall y \big[\forall z \big((z \in x) \leftrightarrow (z \in y)\big) \rightarrow (x = y)\big]$$

The elements of the universe $(\mathcal{U}, \in)$ are in the first place just objects without any structure. What matters is their relationship to other elements with respect to $\in$. We may think of $\mathcal{U}$ as a directed graph where the sets in $\mathcal{U}$ are nodes and $a \in b$

corresponds to an edge $a \leftarrow b$. Part of the universe may have nodes called 0, 1, 2, $\{1\}$ and edges $0 \leftarrow 1$, $0 \leftarrow 2$, $1 \leftarrow 2$, $1 \leftarrow \{1\}$:

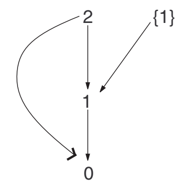

FIGURE 1. Snapshot of the Universe

An edge $0 \leftarrow \{1\}$ would violate the axiom of extensionality, because then 2 and $\{1\}$ would have the same elements.

The Null Set Axiom: There is a set with no elements:

$$\exists x \forall y \, \neg(y \in x)$$

By extensionality, there is only one such set. It is denoted by $\emptyset$ and called the *empty set*. It is a *constant* within the universe $\mathcal{U}$, i.e., a unique element defined by a formula.

The Pairing Axiom: For any sets a and b there is a set c whose only elements are a and b:

$$\forall x \forall y \exists z \forall t \big[(t \in z) \leftrightarrow \big((t = x) \vee (t = y)\big)\big]$$

By extensionality again, there is for given a and b only one such set c. We write $c = \{a, b\}$ for the set whose only elements are a and b. If a and b are the same set, then c has only one element, namely a. That is, for any set a of the universe $\mathcal{U}$ there is a set whose only element is a. This set is called the *singleton* $\{a\}$; $\{a, b\}$ is called a *pair* if a is different from b. Three applications of the pairing axiom lead to the existence of the set $\{\{a\}, \{a, b\}\}$. This is *Kuratowski's* definition of the *ordered pair* (a, b) of a and b. One easily proves the

Theorem 1. *One has that* $(a, b) = (a', b')$ *if and only if* $a = a'$ *and* $b = b'$.

The Union Axiom: For any set a there is a set b whose members are precisely the members of members of a:

$$\forall x \exists y \forall z \big[(z \in y) \leftrightarrow \exists t \big((t \in x) \wedge (z \in t)\big)\big]$$

The set b is called the *union* of a and denoted by $\bigcup a$ or $\bigcup \{x | x \in a\}$. We mention some consequences:

- For any sets a, b, c there is a set d whose elements are a, b and c: $d = \bigcup \{\{a, b\}, \{c\}\}$
- The union of $c = \{a, b\}$ is denoted by $a \cup b$. It is easy to see that $a \cup b = \{x | x \in a \text{ or } x \in b\}$.

Let a and b be sets. We say that a is a subset of b if every element of a is also an element of b:

$$(x \subseteq y) \equiv \forall z \big[(z \in x) \rightarrow (z \in y)\big]$$

The left-hand side is not a formula, because $\in$ is the only relation of our universe; $(x \subseteq y)$ is only an abbreviation of the formula in the variables x and y on the right hand side. In particular we have by extensionality that

$$\forall x \forall y \big[(x = y) \leftrightarrow \big((x \subseteq y) \wedge (y \subseteq x)\big)\big]$$

The Power Set Axiom: Let a be a set of the universe $\mathcal{U}$. Then there is a set b whose elements are precisely the subsets of a:

$$\forall x \exists y \forall z \big[(z \in y) \leftrightarrow (z \subseteq x)\big]$$

The set b is called the power set of a and we use the notation $b = \mathcal{P}(a)$. We have $\mathcal{P}(\emptyset) = \{\emptyset\}$, $\mathcal{P}(\{\emptyset\}) = \{\emptyset, \{\emptyset\}\}$, $\mathcal{P}(\{\emptyset, \{\emptyset\}\}) = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}$.

If a is any set of our universe, any $c \in \mathcal{P}(a)$ corresponds to an intuitive subset of a, namely $\{d | d \leftarrow c\}$ where for each such d, $d \leftarrow a$ holds. However, not every proper collection of edges $d \leftarrow a$ will lend itself to a set c of the universe. For example, if $\mathcal{U}$ happens to be countable then any infinite set a in $\mathcal{U}$ will have "subsets" which don't correspond to sets in $\mathcal{U}$. What kind of properties now lead to subsets? We have reached the point where we have to talk a bit about mathematical logic.

The Language of Axiomatic Set Theory. We are going to describe a *formal language* that has the following ingredients.

(1) Symbols
  (a) An **unlimited** supply of *variables* x_0, x_1, $x_2 \ldots$.
  (b) The elements of the universe $\mathcal{U}$ are the *constants* of the language.
  (c) The *membership* symbol $\in$ and the *equality* symbol $=$.
  (d) The symbols for the *propositional connectives*: $\wedge$ which stands for **and**, $\vee$ which stands for **or**, $\neg$ which stands for **not**, $\rightarrow$ which stands for **if, then**, $\leftrightarrow$ which stands for **if and only if**.
  (e) For each variable x_n one has the *universal quantifier* $\forall x_n$ which stands for **for all** x_n and the *existential quantifier* $\exists x_n$ which stands for **there exists some** x_n.

(2) Formation Rules for Formulas
  (a) Let u and v stand for any variable or constant. Then $(u \in v)$ and $(u = v)$ are formulas. These are the *atomic* formulas.
  (b) If P and Q are formulas then $(P \wedge Q)$, $(P \vee Q)$, $\neg P$, $(P \rightarrow Q)$, $(P \leftrightarrow Q)$ are formulas.
  (c) If P is a formula then $\forall x_n P$ and $\exists x_n P$ are formulas.

Only expressions that can be constructed by finitely many applications of these rules are formulas.
For better readability, different kinds of parentheses will be used, and letters, like x, y, $z, \ldots$ will stand for variables. There are standard conventions concerning the priorities of the binary propositional connectives in order to avoid an excessive accumulation of parentheses.

The axioms of set theory as stated so far are all formulas, actually *sentences*, that is, all occurrences of variables are *bound*. If Q is a formula then every occurrence of x_n within P of a subformula $\forall x_n P$ or $\exists x_n P$ of Q is said to be bound. Variables x_n which are not bound, i.e., which are not within the *scope* of a quantifier $\forall x_n$ or $\exists x_n$ of Q, are said to be *free*. If we underline in a formula a variable then this variable is meant to occur only bound.

Formulas can be represented by certain labelled, directed trees. An atomic formula is just a node, e.g.,

$$(x \in a)$$

which is a tree. If Γ_1 is the tree for P_1 and if Γ_2 is the tree for P_2 , then the tree for $(P_1 \wedge P_2)$ is the graph:

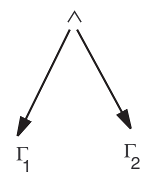

FIGURE 2. The Graph of a Conjunction

Any node of the tree Γ for the formula Q determines a subformula P of Q. For example, a node labelled $\wedge$ determines a conjunction $P \equiv (P_1 \wedge P_2)$ as a subformula of Q, where P_1 and P_2 are subformulas of P; P_1 and P_2 are the scope of the node $\wedge$. Similarly, a node $\forall x$ determines a subformula $P \equiv \forall x_n P_1$, where the subformula P_1 of P is the scope of the node $\forall x_n$ within Q.

Whenever we indicate a formula P as $P(x_0, x_1, \ldots, x_{n-1})$, it is understood that the free variables of P, if there are any, are are among $x_0, x_1, \ldots, x_n$. The constants within a formula are often called *parameters.* So we write $P(x_0, \ldots, x_{n-1}, a_0, \ldots, a_{m-1})$ to indicate the free variables and parameters of a formula. A sentence P is either true or false in the universe $\mathcal{U}$. More generally, if $P(x_0, \ldots, x_{n-1})$ is a formula with free variables $x_0, \ldots, x_{n-1}$ and if $a_0, \ldots, a_{n-1}$ belong to $\mathcal{U}$, then a simultaneous substitution of the x_i by the a_i makes $P(a_0, \ldots a_{n-1})$ either true or false. When we say that a formula $P(x_0, \ldots x_{n-1})$ holds on $\mathcal{U}$, it is meant that its *closure*, i.e.,

$$\forall x_0 \ldots \forall x_{n-1} P(x_0, \ldots, x_{n-1})$$

holds on $\mathcal{U}$. Because we have used the equality sign $=$ as a symbol within the language, equality of formulas, or more generally their equivalence, is denoted by $\equiv$, e.g., $x = y \equiv y = x$. That is, we write $P \equiv Q$ if and only if $P \leftrightarrow Q$ is a theorem of logic. Formulas without parameters are called *pure* formulas of set theory.

A formula in one free variable, or *argument*, is called a class.

$$S(x, a) \equiv (x \in a) \text{ and } R(x) \equiv \neg(x \in x)$$

are examples of classes. The first class defines a set, namely a, while the second one does not define a set: b satisfies $S(x, a)$ iff $b \in a$; there is no set r such that b satisfies $R(x)$ iff $b \in r$.

Formulas $P(x_0, \ldots, x_{n-1})$ are called *n-ary relations.* Formulas in two arguments are called *binary* relations. We also use the terms *predicates*, *properties* and *expressions* for formulas. Let $E(x)$ be a class. We say that $R(x, y)$ is a *relation on* $E(x)$ if

$$\forall x \forall y \big[R(x, y) \to \big(E(x) \wedge E(y)\big)\big]$$

holds on $\mathcal{U}$.

Let $R(x,y)$ be a binary relation. We define *domain* and *range* as the classes

$$\text{dom\_ of\_}R(x,\underline{y}) \equiv \exists y R(x,y) \text{ and } \text{ran\_of\_}R(\underline{x},y) \equiv \exists x R(x,y)$$

Then $R(x,y)$ is a relation on

$$E(z) \equiv \text{dom\_ of\_}R(z,\underline{y}) \vee \text{ran\_of\_}R(\underline{x},z)$$

which we call the *extent* of $R(x,y)$.

A binary relation $R(x,y)$ is called *reflexive* if

$$\forall x \forall y \big[R(x,y) \to \big((R(x,x) \wedge R(y,y) \big) \big]$$

holds on $\mathcal{U}$.

The relation $R(x,y)$ is *symmetric* if

$$\forall x \forall y \big[R(x,y) \to R(y,x) \big]$$

holds on $\mathcal{U}$.

The relation $R(x,y)$ is *transitive* if

$$\forall x \forall y \forall z \big[\big(R(x,y) \wedge R(y,z) \big) \to R(x,z) \big]$$

holds on $\mathcal{U}$.

The binary relation $E(x,y)$ is called an *equivalence* if it is reflexive, symmetric and transitive. It is easy to see that for any reflexive relation, e.g., an equivalence $E(x,y)$ one has that,

$$\text{dom\_of\_}E(x,\underline{y}) \equiv E(x,x) \text{ and } \text{ran\_of\_}E(\underline{x},y) \equiv E(y,y)$$

and therefore $E(x,y)$ is a relation on its domain $D(x)$.

The binary relation $R(x,y)$ is called *anti-symmetric* if

$$\forall x \forall y \big[\big(R(x,y) \wedge R(y,x) \big) \to (x = y) \big]$$

holds on $\mathcal{U}$.

A binary relation $PO(x,y)$ which is reflexive, transitive and anti-symmetric is called a *partial order*. Again we have by reflexivity that domain and range define the same class and that $PO(x,y)$ is a relation on its domain. A partial order $L(x,y)$ is called *linear* or *total*, if

$$\forall x \forall y \big[\big(D(x) \wedge D(y) \big) \to \big(L(x,y) \vee L(y,x) \big) \big]$$

holds on $\mathcal{U}$. $D(x)$ denotes the domain of $L(x,y)$.

An $(n+1)$-ary relation $F(x_0,\ldots,x_{n-1},y)$ is called *functional* if

$$\forall x \ldots \forall x_{n-1} \forall y_1 \forall y_2 \big[\big(F(x_0,\ldots,x_{n-1},y_1) \wedge F(x_0,\ldots,x_{n-1},y_2) \big) \to (y_1 = y_2) \big]$$

holds on $\mathcal{U}$.

We define for any relation $P(x_0,\ldots,x_{n-1},y)$ domain and range

$$\begin{aligned} \text{dom\_of\_}P(x_0,\ldots,x_{n-1},\underline{y}) &\equiv \exists y P(x_o,\ldots,x_{n-1},y) \\ \text{ran\_of\_}P(\underline{x_0},\ldots,\underline{x_{n-1}},y) &\equiv \exists x_0 \ldots \exists x_{n-1} P(x_0,\ldots,x_{n-1},y) \end{aligned}$$

The domain is an $n-ary$ relation $D(x_0,\ldots,x_{n-1})$ while the range is a class $R(y)$.

The binary relation

$$P(x,y) \equiv \forall z \big[(z \in y) \leftrightarrow (z \subseteq x) \big]$$

is functional in the variable x. It assigns to a set a the power set $b = \mathcal{P}(a)$. We have that

$$\text{dom\_of\_}P(x,\underline{y}) \equiv (x = x) \ \text{ and } \text{ran\_of\_}P(x,y) \equiv \exists x \forall z \big[(z \in y) \leftrightarrow (z \subseteq x) \big]$$

Instead of $P(x,y)$ we will often use the more suggestive notation $y = \mathcal{P}(x)$. We similarly write $z = (x,y)$, $z = \{x,y\}$ and $z = x \cup y$ for the corresponding functional predicates.

We define for a formula $F(x,y,z)$ the expression

$$fun\big(F(\underline{x},\underline{y},z)\big) \equiv \forall x \forall y_1 \forall y_2 \big[F(x,y_1,z) \wedge F(x,y_2,z) \rightarrow y_1 = y_2\big]$$

which holds for a set a in $\mathcal{U}$ if and only if $F(x,y,a)$ is functional.

The Schemes of Replacement and Comprehension. In the previous section we didn't stipulate the existence of sets. For example, domain and range of a binary relation were defined as classes, i.e., as formulas in one variable. Of course, given a binary relation on a given set a, domain and range should be subsets of a. The existence of sets according to standard constructions in mathematics is guaranteed by

The Axiom Scheme of Replacement: Let $F(x,y,x_0,\ldots,x_{n-1})$ be a pure formula of axiomatic set theory such that for sets $a_0,\ldots,a_{n-1}$ the binary relation $F(x,y,a_0,a_1,\ldots,a_{n-1})$ is functional. Let a be any set. Then there is a set b such that $d \in b$ holds if and only if there is some $c \in a$ such that $F(c,d,a_0,\ldots,a_{n-1})$ holds on $\mathcal{U}$:

$$\forall x_0 \ldots \forall x_{n-1} \big(fun\big(F(\underline{x},\underline{y},x_0,\ldots,x_{n-1})\big) \rightarrow$$
$$\forall x \exists y \forall v [v \in y \leftrightarrow \exists u [u \in x \wedge F(u,v,x_0,\ldots,x_{n-1})]]\big)$$

Because this is supposed to hold for **every** pure formula $F(x,y,x_0,\ldots,x_{n-1})$, where at least x and y are free, this list of axioms is called a *scheme*. It is called replacement because it allows us to replace some of the elements c of the set a simultaneously by sets d in order to create a set b. As a first application of replacement we will deduce its weaker cousin

The Scheme of Comprehension: Let $A(x,x_0,\ldots,x_{n-1})$ be a pure formula of axiomatic set theory and let $a_0,\ldots,a_{n-1}$ be sets. Then for any set a there is a set b which consists exactly of those elements c of a for which $A(c,a_0,\ldots,a_{n-1})$ holds on $\mathcal{U}$:

$$\forall x_0 \ldots \forall x_{n-1} \forall x \exists y \forall z \big[z \in y \leftrightarrow \big(z \in x \wedge A(z,x_0,\ldots,x_{n-1})\big)\big]$$

In order to deduce this from replacement, we only have to note that

$$F(x,y,x_0,\ldots,x_{n-1}) \equiv \big(y = x \wedge A(x,x_0,\ldots,x_{n-1})\big)$$

is functional.

The standard notation for the subset b of a, which is defined by the property $A(x,a,..,a)$, is

$$b = \{x | x \in a \wedge A(x,a_0,\ldots,a_{n-1})\}$$

Constructions within the Universe. The existence of the union of a set a was stipulated as an axiom. We don't need a further axiom for the intersection.

The Intersection of a Set: Let a be **non-empty** set. Then there is a set b whose members are precisely the members of all members of a.

$$\forall x \big[\neg(x = \emptyset) \rightarrow \exists y \forall z \big[(z \in y) \leftrightarrow \forall t \big(t \in x \rightarrow z \in t\big)\big]\big]$$

This follows at once from comprehension. Note that the intersection of the set a is contained in any of its members c. The standard notation for the intersection of a set a is $\bigcap a$ or $\bigcap\{x|x \in a\}$. Why is it important to assume that the set a is non-empty?

The Cartesian product of Two Sets: Let a and b be sets. Then there is a set c such that $e \in c$ if, and only if, $e = (f, g)$ where $f \in a$ and $g \in b$:

$$\forall x \forall y \exists z \forall t \big[(t \in z) \leftrightarrow \exists u \exists v \big(t = (u, v) \wedge (u \in x) \wedge (v \in y)\big)\big]$$

The equation $z = (x, y)$ is shorthand for the functional relation $Q(x, y, z)$ which says that z is the ordered pair (x, y), which according to Kuratowski's definition is the set $\{\{x\}, \{x, y\}\}$. Thus:

$$Q(x, y, z) \equiv \forall t \Big[(t \in z) \leftrightarrow$$

$$\exists u \exists v \Big\{\big(t = u \vee t = v\big) \wedge \forall s \big[s \in u \leftrightarrow s = x\big] \wedge \forall s' \big[s' \in v \leftrightarrow \big(s' = x \vee s' = y\big)\big]\Big\}\Big]$$

If $= (f, g) = \{\{f\}, \{f, g\}\}$ where $f \in a$ and $g \in b$ then $\{f\} \in \mathcal{P}(a)$ and $\{f, g\} \in \mathcal{P}(a \cup b)$. Hence $(f, g) \in \mathcal{P}(\mathcal{P}(a \cup b))$. We now apply comprehension to

$$P(z, a, b) \equiv \exists u \exists v \big[Q(u, v, z) \wedge u \in a \wedge v \in b\big]$$

which says that "z is an ordered pair whose two components belong to a and b", respectively and get the desired result as

$$c = \{e|e \in \mathcal{P}(\mathcal{P}(a \cup b)) \wedge P(e, a, b)\}$$

The set c is called the *cartesian product* $a \times b$ of a and b. The cartesian product of finitely many sets is similarly defined. The formula

$$C(x, y, z) \equiv \forall t \big[t \in z \leftrightarrow \exists u \exists v \big[Q(u, v, t) \wedge (u \in x) \wedge (v \in y)\big]$$

is functional and says that z is the cartesian product of x and y.

We remark that a binary relation $R(x, y)$ may be perceived as a unary relation $R^*(z)$:

$$R^*(z) \equiv \exists x \exists y \big(Q(x, y, z) \wedge R(x, y)\big)$$

That is, $R^*(e)$ holds if and only if $e = (c, d)$ and $R(c, d)$ holds.

Relations as Sets: Let $R(x, y)$ be a binary relation. Assume that domain and range of $R(x, y)$ are sets a and b, respectively. Then define the set

$$r = \{e|e = (c, d) \in a \times b, R(c, d)\}$$

We now have $e = (c, d) \in r$ if and only if $R^*(e)$ holds. In this sense we may identify a binary relation, for which the extent is a set, by a set of ordered pairs.

Graphs of Functions: If the binary relation $F(x, y)$ is functional and the domain of $F(x, y)$ is a set a, then, according to replacement, the range is also a set. Let b be any set containing the range of $F(x, y)$. The set

$$\{e|e = (c, d) \in a \times b, F(c, d)\}$$

is called the graph of the function $f : a \to b$.

The *projections* are important examples of functional relations:

$$\begin{aligned} F_1(x, y, z, p) &\equiv Q(x, y, z) \wedge p = x \\ F_2(x, y, z, q) &\equiv Q(x, y, z) \wedge q = y \end{aligned}$$

We have that $F_1(a,b,c,d)$ holds on $\mathcal{U}$ iff $c=(a,b)$ and $d=a$, i.e., d is the first component of the ordered pair c. The predicate

$$P_1(t,x) \equiv \exists u \exists v \big[F_1(u,v,t,x)\big]$$

then holds if "x is the first component of the ordered pair t". That f is a function from a to b is expressed by $F(a,b,f)$ where

$$F(x,y,z) \equiv \exists p\big[C(x,y,p) \wedge z \subseteq p\big] \wedge \forall u\big[u \in x \rightarrow \exists t\big(t \in z \wedge P_1(t,u)\big)\big] \wedge$$
$$\forall t \forall t' \forall s \forall u \forall u'\big[t \in z \wedge t' \in z \wedge P_1(t,s) \wedge P_1(t',s) \wedge P_2(t,u) \wedge P_2(t',u') \rightarrow u = u'\big]$$

The Exponentiation of Sets: Let a and b be sets. Then there is a set c whose elements are given by the functions $f : a \rightarrow b$.

The function $f : a \rightarrow b$ is an element of $\mathcal{P}(a \times b)$. Hence c is defined by comprehension:

$$c = \{f | f \in \mathcal{P}(a \times b), F(a,b,f)\}$$

For the set c one uses *exponential* notation $c = b^a$.

Union and Intersection of a Family of Sets: A function s with domain i is sometimes called a *family of sets* $a_j, j \in i$, where, of course $a_j = s(j)$. The union of the family s is the union of the range r of s, which is, according to the replacement axiom, a set u. We write $u = \bigcup\{a_j | j \in i\} = \bigcup s$. The intersection of a non-empty family s is defined similarly.

The Cartesian Product of a Family of Sets: Let s be a family of sets, indexed by the set i. A function $f : i \rightarrow u$ from i into the union u of the range r of s is called a *choice function* if for every $j \in i$ one has that $f(j) \in a_j$. Then there is a set c whose members are all the choice functions for s. This set is called the *cartesian product of the family* s and is denoted by $c = \prod\{a_j | j \in i\}$.

This follows from comprehension: We will use the expression $(x,y) \in z$ as shorthand for $\exists p(Q(x,y,p) \wedge (p \in z))$. Then $c = \{f | f \in u^i \wedge \forall x \forall y \forall z\big(\big((x,y) \in f \wedge (x,z) \in s\big) \rightarrow (y \in z)\big)\}$

The Remaining Axioms of ZF. Within the universe $\mathcal{U}$ we certainly can find the sets $\underline{0} = \emptyset$, $\underline{1} = \{\underline{0}\}$, $\underline{2} = \{\underline{0}, \underline{1}\}$,..., $\underline{n} = \{\underline{0}, \underline{1}, \ldots, \underline{n-1}\}$. Note that $\underline{n+1} = \underline{n} \cup \{\underline{n}\}$ where $\underline{n}$ is not a member of $\underline{n}$. Hence $\underline{n}$ has exactly n elements and $n \mapsto \underline{n}$ is an injective map from the "set" $\mathbb{N}$ of natural numbers into the universe $\mathcal{U}$. The sets $\underline{n}$ are called the *natural number objects* of $\mathcal{U}$. Notice that we have $n < m$ if and only if $\underline{n} \in \underline{m}$, and $\underline{n} \in \underline{m}$ is the same as $n \subset m$, $\subset$ standing for strict inclusion. On the basis of the axioms stated so far we have no way of telling whether there is a set whose elements are exactly the sets $\underline{n}$. An axiom stating that there is a set ω consisting exactly of all natural number sets $\underline{n}$ is objectionable: Such a definition of ω would not be given by a first order sentence of our language of set theory. This deficiency was resolved by *Dedekind:*

For any set x, $x \cup \{x\}$ is called the *successor* x^+ of x. For example, $\underline{1}$ is the successor of $\underline{0}$.

A set i is called *inductive* if we have that $\emptyset \in i$ and $x \in i$ implies that x^+ is in i.

The Axiom of Infinity: There is an inductive set i_0.

It is obvious that any inductive set i_0 must contain all $\underline{n}$. In this sense we have stipulated the existence of some infinite set, namely of i_0. Notice that the intersection of any inductive set i with i_0 is inductive. We therefore have a smallest inductive set, called ω, namely the intersection of all inductive sets within i_0. There is no danger to think that ω consists exactly of all finite numbers $\underline{n}$. And the vast majority of mathematicians feels that way. On the other hand, any axiomatic definition of ω allows for elements ν which are *nonstandard*, i.e., different from any ordinary number $\underline{n}$. However, whether one realizes this possibility or not seems to be irrelevant for the formal development of mathematics.

There are two more axioms most mathematicians consider as"true", the *Axiom of Choice* and the *Axiom of Foundation.* These axioms are listed separately, mainly because because a great deal of set theory can be developed without them.

The Axiom of Choice (AC): The Cartesian product of a family of non-empty sets is itself non-empty.

The Axiom of Foundation (AF): Every non-empty set a contains a set b which is disjoint to a.

Both axioms are independent of ZF.

The Axiom of Choice has many other equivalent formulations. For example, it says that given any set p of non-empty, and pairwise disjoint sets, then there is a set r such that for every $c \in p$ one has that $r \cap c$ is a one-element set. We may think that p is the partition according to an equivalence relation e on the union of p. Then the AC says that there is a set r of *representatives* which picks from every equivalence class c exactly one element, a representative for the whole class c. In this version, AC looks rather obvious and this is what most mathematicians think. But how can we find such a set of representativs? If it is the case that somehow every class c is ordered in a way that it has a minimum then just take for every c the minimum. But what if c is not ordered? Can we always order any non-empty set c? The answer is yes, but only if we assume AC. This is actually a big theorem. Recall that a total order is called a *well-order* if every non-empty subset has a minimum. Then one has

Well-Ordering Theorem. *The Axiom of Choice holds if and only if every non-empty set admits a well ordering.*

We have developed set theory in this chapter to the extent that all elementary mathematics can be carried out in a rigorous, axiomatic fashion. Sets that should exist intuitively, have been be shown to exist according to ZF axioms. Higher set theory deals primarily with the theory of infinite sets, in particular with well-ordered sets for which the set of natural numbers, ω, is the smallest infinite example.

We conclude with a few problems.

Problem 1. *Is it true that that $\mathcal{P}(a) = \mathcal{P}(b)$ only if $a = b$?*

Problem 2. *Show that for any set a one has that $r(a) \notin a$ where*

$$r(a) = \{c | c \in a, c \notin c\}.$$

Conclude that $R(x) \equiv (x = x)$ defines a proper class.

Problem 3. *Show that* **AF** *implies $a \notin a$ for every set a. Hint: Apply* **AF** *to $c = \{a\}$.*

This consequence of AF is probably the most intuitive statement about sets. A set consists of items and these items make up the set. A set is of higher *rank* than any of its members. In naive set theory one writes $S = \{a, b, \ldots\}$ and uses capital letters, like S, to denote a set, and lower-case letters to denote its elements.

The following problems show that our set ω satisfies the *Peano Axioms* for natural numbers:

P1: 0 is not a successor, that is $0 \neq n^+$ for every $n \in \omega$.
P2: $n^+ = m^+$ only if $n = m$.
P3=Complete Induction: If S is a set of natural numbers, that is of elements of ω such that $0 \in \omega$ and that if $n \in \omega$ then $n^+ \in \omega$, one has that $S = \omega$.

Recall that we have defined 0 as the empty set $\emptyset$, and ω as the smallest inductive set. Of course $\emptyset$ cannot be a successor, a successor is of the form $x \cup \{x\}$, and therefore contains at least one element, namely x. Thus P1 is obvious for ω. Because the successor of n is $n^+ = n \cup \{n\}$, complete induction is also trivial for ω. However, the second Peano Axioms requires a proof. In order for doing this, one best proceeds by proving some preliminary statements. You may have some fun proving these basic facts on ω.

Lemma. *If $x \in y$ where $y \in \omega$ then $x \subseteq y$.*

This lemma says that $\in$ is a transitive relation on ω: If x, y, z are elements of ω where $z \in x$ and $x \in y$ then $z \in y$. Or, if y is an element of ω then every element x of y is a subset of y. That it is a proper subset is the

Proposition. *For every $y \in \omega$ one has that an element $x \in y$ is a proper subset.*

Corollary. *$x \notin x$ for every $x \in \omega$.*

The second Peano Axiom is now the

Proposition. *Let $x, y \in \omega$. Then*

$$x^+ = y^+ \Longrightarrow x = y$$

Quite useful is

Problem 4. *Let $x \neq 0$ be an element of ω. Then $x = y^+$ for some unique $y \in \omega$.*

We have defined ω as the smallest set in the universe of sets which contains $0 = \emptyset$ and which contains with any x the successor of x^+. Can we conclude that $\omega = \{0, 1, 2, \ldots\}$? The answer is no! One can think that ω contains elements which are different from the "ordinary" natural numbers, which are "infinite". On the other hand, one can also assume that ω conains just these ordinary numbers which we can construct from $\emptyset$ by applying "finitely" many times the successor operation. If this makes you wonder, our short introduction to axiomatic set theory has served its purpose.

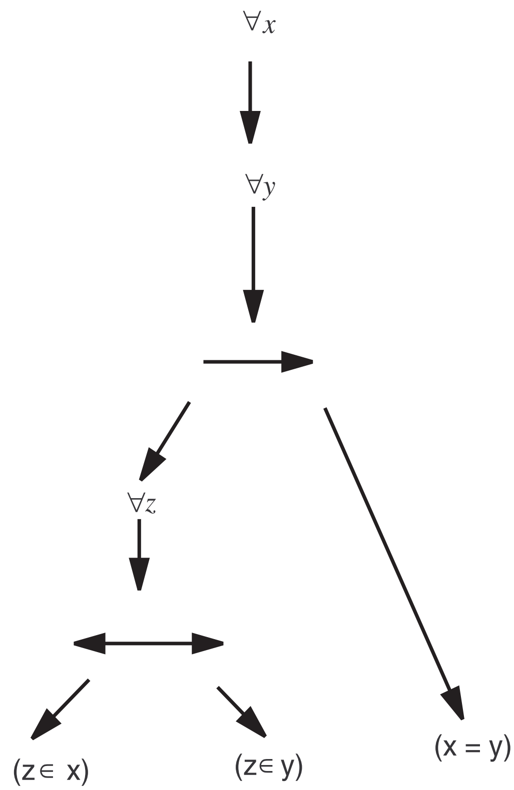

Figure 3. The Graph of the Extensionality Axiom