A **character** is one of the 52 upper or lower case letters of the English alphabet.

Our set theoretic universe is populated by **objects**. An object is either a **class** or an **atom** but not both.

2. Propositional calculus.

Suppose $\mathcal{L}$ is the collection of the symbols

$$(\quad) \quad \neg \quad \vee \quad \wedge \quad \to \leftrightarrow$$

$\mathcal{P}$ is a collection of symbols none which occurs in $\mathcal{L}$; a symbol in $\mathcal{P}$ is called a **propositional variable**. Suppose **s** is a symbol which does not occur in either $\mathcal{L}$ or $\mathcal{P}$. Suppose S is a finite sequence of symbols each of which occurs in either $\mathcal{L}$ or $\mathcal{P}$. Let $\mathbf{T}(S)$ be the finite sequence of symbols obtained by replacing the leftmost occurrence of a propositional variable p preceeded by $\neq\ p$ in S by **s**; note that such a substitution decreases the length of S by one. Let $\mathbf{U}(S)$ the finite sequence of symbols obtained by replacing leftmost occurrence of $(\,p\,\mathbf{b}\,q\,)$, where p, q are propositional variables and **b** is one of $\vee, \wedge, \to, \leftrightarrow$; note that such a substitution decreases the length of S by four. Let $\mathbf{V}(S) = \mathbf{U}(\mathbf{T}(S))$. We say S is **propositional form** if applying **V** the length of S times results in the sequence of one statement variable.

A **statement** is an expression which either true or false. We will use $\top$ to stand for "true" and $\bot$ to stand for "false". If p is a statement then $\neg\, p$ is the statement with truth values given by the table

p	$\neg\, p$
$\top$	$\bot$
$\bot$	$\top$

;

thus $\neg\, p$ is the **negation of** p. If p and q are statements then

$$(p \vee q) \quad (p \wedge q) \quad (p \to q) \quad (p \leftrightarrow q)$$

are the statements whose truth values given by the table

p	q	$(p \vee q)$	$(p \wedge q)$	$(p \to q)$	$(p \leftrightarrow q)$
$\top$	$\bot$	$\top$	$\top$	$\top$	$\top$
$\top$	$\bot$	$\top$	$\bot$	$\bot$	$\bot$
$\bot$	$\top$	$\top$	$\bot$	$\top$	$\bot$
$\bot$	$\bot$	$\bot$	$\bot$	$\top$	$\bot$

;

thus $(p \vee q)$ means p **or** q; $p \wedge q$ means p **and** q; $p \to q$ means **if** p **then** q; $p \leftrightarrow q$

means p **iff** q; here **iff** means $(p\ \to\ q)$ and $(q\ \to\ p)$.

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

We will use very elementary properties of the collection

$$\mathbb{N}$$

of **natural numbers**

$$0, 1, 2, \ldots, \mathrm{N}, \ldots.$$

Whenever i, j, k are natural numbers

$$j|(i,k), \quad j|(i,k], \quad j|[i,k), \quad j|[i,k]$$

means

$$i < j < k, \quad i < j \le k, \quad i \le j < k, \quad i \le j \le k,$$

respectively.

For each positive natural number N an **N-tuple** is a sequence

$$S_1\, S_2\, \ldots\, S_N$$

where, if $i|(1,N)$, S_i is some entity; we let

$$|S| = N$$

and call N the **length** of S.

We say S is a **tuple** if S is an N-tuple for some natural number N.

Suppose S and T are tuples. The **concatenation of S and T**, written

$$S\,T$$

is the $(|S| + |T|)$-tuple such that

$$(S\,T)_i = \begin{cases} S_i & \text{if } i = 1, \ldots, |S|, \\ T_{i-|S|} & \text{if } i = |S| + 1, \ldots, |S| + |T|; \end{cases}$$

note that concatenation is associative. If $J|[1, |T|]$ we say S is a **subtuple of T starting at J** if $J + |S| - 1 \le |T|$ and

$$S_j = T_{J+j} \quad \text{whenever } j|[J, J + |S| - 1].$$

3. Propositional logic.

We begin by defining the notion of **statement**. To do this we will need the following **logical symbols**: the symbols

$$(\qquad)$$

are the **punctuation characters;** the symbol

$$\neg$$

is the **negation operator**; the symbols

$$\vee \qquad \wedge \qquad \rightarrow \qquad \leftrightarrow$$

are the **logical connectives**; $\vee$ means "...or..."; $\wedge$ means "...and..."; $\rightarrow$ means "if...then..."; $\leftrightarrow$ means "...if and only if...". We also have a collection of **propositional variables** none of which is a logical symbol.

For each natural number L we define the **statements of level** L by induction on L as follows. The statements of level 0 are the 1-tuples S such that S_1 is a propositional variable. If $L > 0$ the statements of level L are the concatenations

$$\neg\, A$$

where A is a statement of level $L - 1$ as well as the concatenations

$$(\, A\, \mathbf{b}\, B\,)$$

where, for some natural numbers J and K with $\max\{J, K\} = L - 1$, A is a statement of level J and $\mathbf{b}$ is a logical connective and B is a statement of level K.

C is a **statement** if C is a statement of level L for some natural number L. The tuple B is a **subproposition of the statement** C if B is a proposition and B is a subtuple of C.

Theorem 3.1. Suppose C is a statement. There is a unique natural number L such that C is a statement of level L and exactly one of the following holds:

(I) $L = 0$ then $|C| = 1$ and C_1 is a propositional variable;
(II) $L > 0$ and there is a unique statement A of level $L-1$ such that $C = \neg A$;
(III) $L > 0$ and there are unique $J, K, A, B, \mathbf{b}$ such that J and K are natural numbers such that $\max\{J, K\} = L - 1$, A is a statement of level J, B is a statement of level K and $\mathbf{b}$ is a logical connective.

Theorem 3.2. Suppose S is a statement and $I|[1, |S|]$. There are unique R and s such that

(i) R is a tuple of statements;
(ii) $R_1 = S$;
(iii) s is an $|R|$-tuple of natural numbers such that
$$1 = s_1 \leq \cdots \leq s_{|R|} < |S|;$$
(iv) $s_i + R_i \subset S$ whenever $i|[1, |R|]$;
(vi) if $S_I =$ (then $I = s_M$; if $S_I =$) then $I = s_M + |R_M| - 1$; if $S_I = \neg$ then $s_M = I$; if S_I is a propositional variable p then $R_M = p$.

Proof. Induct on the level L of S. If $L = 0$ the Theorem holds trivially so suppose $L > 0$. If $I = 1$ we are done so suppose $I > 1$.

Suppose $S = \neg A$ where A is a proposition of level $L - 1$. If $I = 1$ we are done and if $I > 1$ we apply the inductive hypothesis to A.

Suppose $S = (\, A\, \mathbf{b}\, B\,)$

□

Theorem 3.3. Suppose a truth value $\top$ for "true" and $\bot$ for "false" has been assigned to each propositional variable. There is a unique assignment of a truth value $\mathcal{T}(S) = \top$ or $\mathcal{T}(S) = \bot$ to each statement S in such a way that

$$\mathcal{T}(\neg A) = \neg\mathcal{T}(A)$$

and

$$\mathcal{T}((\, A\, \mathbf{b}\, B\,)) = \begin{cases} \mathcal{T}(A) \vee \mathcal{T}(B) & \text{if } \mathbf{b} = \vee; \\ \mathcal{T}(A) \wedge \mathcal{T}(B) & \text{if } \mathbf{b} = \wedge; \\ \mathcal{T}(A) \rightarrow \mathcal{T}(B) & \text{if } \mathbf{b} = \rightarrow; \\ \mathcal{T}(A) \leftrightarrow \mathcal{T}(B) & \text{if } \mathbf{b} = \leftrightarrow \end{cases}$$

whenever A and B are statements and the above values are given as in the following table:

$\mathcal{T}(A)$	$\mathcal{T}(B)$	$\mathcal{T}(A) \vee \mathcal{T}(B)$	$\mathcal{T}(A) \wedge \mathcal{T}(B)$	$\mathcal{T}(A) \rightarrow \mathcal{T}(B)$	$\mathcal{T}(A) \leftrightarrow \mathcal{T}(B)$
$\top$	$\bot$	$\top$	$\top$	$\top$	$\top$
$\top$	$\bot$	$\top$	$\bot$	$\bot$	$\bot$
$\bot$	$\top$	$\top$	$\bot$	$\top$	$\bot$
$\bot$	$\bot$	$\bot$	$\bot$	$\top$	$\bot$

Definition 3.1. Suppose C is a statement. For each $i = 1, \dots, |C|$ let $\mathbf{L}(i, C)$ be the number of $j = 1, \dots, i$ such that $C_i =$ (and let $\mathbf{R}(i, C)$ be the number of $j = 1, \dots, i$ such that $C_j =$).

Lemma 3.1. If C is a statement then $\mathbf{L}(|C|, C) = \mathbf{R}(|C|, C)$.

Proof. If C is a statement of level 0 then $\mathbf{L}(|C|, C) = 0 = \mathbf{R}(|C|, C)$ Suppose L is a positive natural number the Lemma holds C is of level $L - 1$. If $C = \neg A$ where A is a statement of level $L - 1$ then

$$\mathbf{L}(|C|, C) = \mathbf{L}(|A|, A) = \mathbf{R}(|A|, A) = \mathbf{R}(|C|, C).$$

If $C = (\, A \, \mathbf{b} \, B \,)$ where, for some natural numbers J and K with $\max\{J, K\} = L - 1$, A is a statement of level J and $\mathbf{b}$ is a logical connective and B is a statement of level K then

$$\mathbf{L}(|C|, C) = 1 + \mathbf{L}(|A|, A) + \mathbf{L}(B) = \mathbf{R}(|A|, A) + \mathbf{R}(|B|, B) + 1 = \mathbf{R}(|C|, C).$$

□

Lemma 3.2. Suppose C is a statement. If $\mathbf{L}(|C|, C) > 0$ then

(i) there is a unique I among $1, \dots, |C|$ such that C_I is a logical connective and $\mathbf{L}(I, C) = 1 + \mathbf{R}(I, C)$;
(ii) if i is among $1, \dots, |C|$, $i \neq I$ and C_i is a logical connective then $\mathbf{L}(i, C) > 1 + \mathbf{R}(i, C)$.

Proof. The Lemma holds trivially if C is of level 0. Suppose L is a positive natural number and the Lemma holds whenever C is of level K where $K < L$.

Suppose C is of level L. If $C = \neg A$ where A is of level $L - 1$ we infer inductively that the Lemma holds. So suppose $C = (\, A \, \mathbf{b} \, B \,)$ where, for some natural numbers J and K with $\max\{J, K\} = L - 1$, A is a statement of level J and $\mathbf{b}$ is a logical connective and B is a statement of level K. Let $I = 1 + |A| + 1$ so $C_I = \mathbf{b}$. By ?? we have

$$\mathbf{L}(I, C) = 1 + \mathbf{L}(|A|, A) = 1 + \mathbf{R}(|A|, A) = 1 + \mathbf{R}(I, C).$$

Suppose i is among $1, \dots, |C|$, $i \neq I$ and C_i is a logical connective. Then *either* (I) $1 < i \leq 1 + |A|$ and $C_i = A_{i-1}$ *or* (II) $1 + |A| + 1 < i < 1 + |A| + 1 + |B|$ and $C_i = B_{i-(1+|A|+1)}$. In case (I) holds we have $\mathbf{L}(|A|, A) > 0$ so, as A is of level J and $J < L$, we have $\mathbf{L}(i, C) = 1 + \mathbf{L}(i - 1, A) > 1 + \mathbf{R}(i - 1, A) = 1 + \mathbf{R}(i, C)$. In case (II) holds we have $\mathbf{L}(|B|, B) > 0$ so, as B is of level K and $K < L$, we have

$$\begin{aligned}\mathbf{L}(i, C) &= 1 + \mathbf{L}(|A|, A) + 1 + \mathbf{L}(i - (1 + |A| + 1), B) \\ &> 1 + \mathbf{R}(|A|, A) + 1 + \mathbf{R}(i - (1 + |A| + 1), B) \\ &> 1 + \mathbf{R}(i, C).\end{aligned}$$

□

Theorem 3.4. Suppose L is a natural number. If C is a statement of level $L + 1$ then *either* $C_1 = \neg$ in which case there is a unique statement A of level L such that $C = \neg A$ *or* $C_1 =$ (and there is a unique statement A of level L, one and only one logical connective $\mathbf{l}$ and a unique statement B of level L such that $C = (\, A \, \mathbf{l} \, B \,)$.

4. Sets, relations and functions.

We hope the reader is familiar with elementary set theory as it is used in mathematics today. Nonetheless, we shall now give a careful treatment of set theory if only to to allow the reader to become conversant with our notation. Although our treatment will be "naive" or "informal" and not axiomatic it will be very nearly axiomatic.

We begin by defining the notion of **well formed formula**, or **wff** for short.[1] To do this we will need the following symbols: the punctuation characters

$$(\qquad)$$

the negation operator

$$\neg$$

the logical connectives

$$\vee \qquad \wedge \qquad \rightarrow \qquad \leftrightarrow$$

from propositional logic as well as the **quantifiers**

$$\forall \qquad \exists$$

special infix binary predicates

$$= \qquad \neq \qquad \in \qquad \notin$$

the **special unary predicates**

Class Set Atom

Our set theoretic universe is populated by **objects**.

We will also use **variables** which are typically letters at the end of the Roman alphabet. *No object is a variable and no variable is an object.* A **term** is either an object or a variable.

For each natural number L we define the **wffs of level** L by induction on L as follows. By a **wff of level zero** we mean the concatenations

$$\mathbf{u}\,(\,\mathbf{s}\,)$$

where **u** is a special prefix unary predicate and **s** is a term as well as the concatenations

$$(\,\mathbf{s}\,\mathbf{i}\,\mathbf{t}\,)$$

where **s** and **t** are terms and **i** is a special infix binary predicate. ZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZ
If $L > 0$ the wffs of level L are the concatenations

$$\neg\, A \qquad \mathbf{q}\, x\, A$$

where A is a statement of level $L-1$, x is a variable and **q** is a quantifier as well as the concatenations

$$(\,A\,\mathbf{b}\,B\,)$$

where, for some natural numbers J and K with $\max\{J,K\} = L-1$, A is a wff of level J and **b** is a logical connective and B is a wff of level K.

C is a **wff** if C is a wff of level L for some natural number L.

ZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZ

---

[1]Our **wff**s are close to but *not* the **wff**s of first order logic. What we give below is basically a model of the Morse-Kelley set theory.

Theorem 4.1. Suppose F is a wff of level 0. Then *either* F_1 is a special prefix unary predicate $\mathbf{u}$ and there is a unique term $\mathbf{s}$ such that

$$F = \mathbf{u}\,(\,\mathbf{s}\,)$$

or $F_1 =$ (and there a unique term $\mathbf{s}$, a unique special infix binary predicate $\mathbf{i}$ and a unique term $\mathbf{t}$ such that

$$F = (\,\mathbf{s}\,\mathbf{i}\,\mathbf{t}\,).$$

Suppose L is a natural number and F is a wff of level $L+1$. Then *either* F_1 is a quantifier $\mathbf{q}$, F_2 is a variable x and there is a unique wff A of level L such that

$$F = \mathbf{q}\,x\,A$$

or $F_1 = \neg$ and there is a unique wff A of level L such that

$$F = \neg\,A$$

or $F_1 =$) and there are a unique wff A, a unique logical connective $\mathbf{l}$ and a unique wff B such that

$$F = (\,A\,\mathbf{l}\,B\,).$$

Corollary 4.1. Suppose $A, \tilde{A}, B, \tilde{B}$ are wffs, $\mathbf{i}, \tilde{\mathbf{i}}$ are logical connectives and

$$(\,A\,\mathbf{i}\,B\,) = (\,\tilde{A}\,\tilde{\mathbf{i}}\,\tilde{B}\,).$$

Then $A = \tilde{A}$, $\mathbf{i} = \tilde{\mathbf{i}}$ and $B = \tilde{B}$.

To every variable free wff of level 0 a is assigned a truth value $\top$ *or* $\bot$ such that the following hold: and spell out some basic relations among such wffs. Suppose o is an object; we assume that each of

$$\mathbf{Class}(o) \quad \text{and} \qquad \mathbf{Atom}(o)$$

has a truth value; we say o **is a class** if $\mathbf{Class}(o)$ is true and we say o **is an atom** if $\mathbf{Atom}(o)$ is true; we assume that

$$(\mathbf{Class}(o) \vee \mathbf{Atom}(o)) \text{ is true and } (\mathbf{Class}(o) \wedge \mathbf{Atom}(o)) \text{ is false}$$

which is to say that every object is either an **atom** or a **class** but not both.

Intuitively, a class is a collection of objects. In particular,

$o \in A$ means o is an object, A is a class and o is a member of the collection A

so

$$(o \in A) \rightarrow \mathbf{Class}(A).$$

If a is an atom then a has no members; thus

$$\mathbf{Atom}(a) \rightarrow \neg(o \in a) \text{ for all objects } o.$$

If o is an object we declare that

$$\mathbf{Set}(o) \text{ iff } \mathbf{Class}(o) \text{ and, for some class } A,\ (o \in A).$$

This singling out of some classes is done to avoid the Russell Paradox which we spell out shortly.

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

We now present the all important **Axiom of Comprehension**. Suppose x is a variable and *either* (I) P is a quantifier free wff whose only variable is x *or* (II)

$P = \mathbf{q}XQ$ where $\mathbf{q}$ is a quantifier, X is a variable different from x and Q is a quantifier free wff whose only variables are x and X. Then

$$P_{x\to a} \text{ has a truth value for any object } a.$$

Axiom of Comprehension. There is one and only one class

$$\{x : P\}$$

such that such that, for any object a,

$$a \in \{x : P\} \text{ iff } a \text{ is a set and } P_{x\to a} \text{ is true.}$$

The uniqueness of $\{x : P\}$ is redundant since it is implied by the Axiom of Extent.

The condition "a is a set" to avoid the Russell Paradox which we now present. Let

$$R = \{x : \neg(x \in x)\}.$$

Theorem 4.2. R is not a set.

Proof. Suppose, to the contrary, R were a set. If $R \in R$ we would have $\neg(R \in R)$ which is a contradiction. On the other hand if $\neg(R \in R)$ then $\neg\neg(R \in R)$ which is equivalent to $R \in R$. □

Remark 4.1. Suppose P is a wff with exactly one "free" variable x. For any object a let $P_{x\to a}$ the wff obtained by replacing each "free occurrence of x in P" by a. Then $P_{x\to a}$ has a truth value. All this requires more syntax analysis that I am inclined to make at this point. See the Appendix for a thorough treatment of this neglected subject.

Let

$$0 = \{x : (x \neq x)\} \quad \text{and let} \quad \mathcal{U} = \{x : (x = x)\}.$$

Then for any object a we have

$$\neg(a \in 0) \quad \text{and} \quad a \in \mathcal{U}.$$

The **Axiom of the Empty Set** says that 0 is a set. We will see shortly that $\mathcal{U}$ is not a set. We will for the most part avoid classes which are not sets. Classes which are not sets, for some reason that escapes me, are called **proper classes**.

For any object a we let

$$\{a\} = \begin{cases} \{x : x = a\} & \text{if } a \text{ is an atom or a set,} \\ \mathcal{U} & \text{else.} \end{cases}$$

For classes A and B we let

$$A \cup B = \{x : (x \in A) \vee (x \in B)\} \quad \text{and} \quad A \cap B = \{x : (x \in A) \wedge (x \in B)\}$$

and call these classes the **union** and **intersection** of A and B.

For any class A we let

$$\sim A = \{x : x \notin A\}.$$

For classes A and B we let

$$A \sim B = A \cap \sim B = \{x : (x \in A) \vee (x \notin B)\}.$$

For any class $\mathcal{A}$ we let

$$\cap\mathcal{A} = \{x : \forall X((X \in \mathcal{A}) \to (x \in X))\} \quad \text{and we let} \quad \cup\mathcal{A} = \{x : \exists X((X \in \mathcal{A}) \vee (x \in X))\}.$$

We say $\mathcal{A}$ is **disjointed** if

$$A \cap B = \emptyset \quad \text{whenever } A, B \in \mathcal{A} \text{ and } A \neq B.$$

Note that $\mathcal{A}$ is disjointed if and only if

$$A, B \in \mathcal{A} \text{ and } A \cap B \neq \emptyset \;\Rightarrow\; A = B.$$

(Parentheses?) Note that

$$\bigcup \emptyset = \emptyset \quad \text{and that} \quad \bigcap \emptyset = \mathcal{U}.$$

Let X be a set and $\mathcal{A}$ is a family of subsets of X. We leave to the reader as an exercise the verification of the **DeMorgan Laws**:

$$X \sim \cup\mathcal{A} = \cap\{X \sim A : A \in \mathcal{A}\}, \tag{1}$$

$$X \sim \cap\mathcal{A} = \cup\{X \sim A : A \in \mathcal{A}\}. \tag{2}$$

Wait a minute! We have not said what we mean by

$$\{X \sim A : A \in \mathcal{A}\};$$

it should really be

$$\{B : \text{for some } A,\ A \in \mathcal{A} \text{ and } B = X \sim A\}$$

we shall abuse notation in this manner unless it might cause confusion. We also point out that the proof of the DeMorgan Laws reduces immediately to rules of elementary logic. *We assume the reader is quite proficient at elementary logic.* Heh, heh.

We let

$$2^X = \{A : A \subset X\}$$

and call this set the **power set of** X.

We say $\mathcal{A}$ is a **partition** of X if $\mathcal{A}$ is a disjointed family of sets such that

$$X = \bigcup \mathcal{A}.$$

Note that if $\mathcal{A}$ is a partition of X then so is $\mathcal{A} \cup \{\emptyset\}$. A family $\mathcal{C}$ of sets is **nested** if either $C \subset D$ or $D \subset C$ whenever $C, D \in \mathcal{C}$.

4.1. **Ordered pairs and relations.** Suppose a and b are objects. We let

$$\{a\} = \{x : x = a\};$$

$$\{a, b\} = \{a\} \cup \{b\};$$

and we let

$$(a, b) = \{\{a\}\} \cup \{\{a, b\}\}$$

and note that

$$\cup(a, b) = \{a\} \cup \{b\} \quad \text{and} \quad \cap(a, b) = \{a\}. \tag{3}$$

Proposition 4.1. Suppose a, b, c and d are objects. Then

$$(a, b) = (c, d) \;\leftrightarrow\; a = c \text{ and } b = d.$$

Proof. Use (3). □

We say p is an **ordered pair** if there exist objects a and b such that

$$p = (a, b).$$

It follows from the preceding Proposition that if a and b are uniquely determined by p so we may define the **first coordinate** of p to be a and the **second coordinate** of (a, b) to be b.

A **relation** is a set whose members are ordered pairs. Whenever r is a relation we let

$$\mathbf{dmn}\, r = \{x : \text{ for some } y,\ (x, y) \in r\}$$

and we let

$$\mathbf{rng}\, r = \{y : \text{ for some } x,\ (x, y) \in r\};$$

we call these sets the **domain** and **range** of r, respectively; we let

$$r^{-1} = \{(x, y) : (y, x) \in r\}$$

and call this relation the **inverse** of r. If r and s are relations we let

$$r \circ s = \{(x, z) : \text{ for some } y,\ (x, y) \in s \text{ and } (y, z) \in r\};$$

we call this relation the **composition** of the relations s and r. More formally,

$$r \circ s = \{(x, z) : \exists y\big((x, y) \in s\big) \wedge \big((y, z) \in t\big)\}.$$

Theorem 4.3. Suppose r, s, t are relations. Then

$$(r \circ s) \circ t = r \circ (s \circ t).$$

That is, composition of relations is associative.

Proof. The following statements are equivalent:

$$(w, z) \in \big((r \circ s) \circ t\big)$$

$$\exists x\Big(\big((w, x) \in t\big) \wedge \big((x, z) \in (r \circ s)\big)\Big)$$

$$\exists x\Bigg(\big((w, x) \in t\big) \wedge \exists y\Big(\big((x, y) \in s\big) \wedge \big((y, z) \in r\big)\Big)\Bigg)$$

$$\exists x \exists y\Bigg(\big((w, x) \in t\big) \wedge \Big(\big((x, y) \in s\big) \wedge \big((y, z) \in r\big)\Big)\Bigg)$$

$$\exists x \exists y\Bigg(\Big(\big((w, x) \in t\big) \wedge \big((x, y) \in s\big)\Big) \wedge \big((y, z) \in r\big)\Bigg)$$

$$\exists y \exists x\Bigg(\Big(\big((w, x) \in t\big) \wedge \big((x, y) \in s\big)\Big) \wedge \big((y, z) \in r\big)\Bigg)$$

$$\exists y\Bigg(\exists x\Big(\big((w, x) \in t\big) \wedge \big((x, y) \in s\big)\Big) \wedge \big((y, z) \in r\big)\Bigg)$$

$$\exists y\Big(\big((w, y) \in (s \circ t)\big) \wedge \big((y, z) \in r\big)\Big)$$

$$(w, z) \in \big(r \circ (s \circ t)\big)$$

□

Remark 4.2. It boils down to the associativity of $\wedge$:

$$\big(A \wedge (B \wedge C)\big) \leftrightarrow \big((A \wedge B) \wedge C\big)$$

for statements A, B, C as well as

$$\exists y \exists z A \leftrightarrow \exists z \exists y A$$

for any statement A.

Exercise 4.1. Show that $(r \circ s)^{-1} = s^{-1} \circ r^{-1}$ whenever r and s are relations.

Definition 4.1. Suppose r is a relation and A is a set. We let

$$r|A = \{(x, y) : (x, y) \in r \text{ and } x \in A\}$$

and call this relation r **restricted to** A and we let

$$r[A] = \{y : \text{ for some } x,\, x \in A \text{ and } (x, y) \in r\}.$$

Remark 4.3. Many, if not most, authors write $r(A)$ for $r[A]$; this is problematic because if r is a function and A is a member of the domain of r then $r(A)$ can be different from $r[A]$. See the simple example in (4.2) below.

Exercise 4.2. Show that

$$r[s[A]] = (r \circ s)[A]$$

whenever r and s are relations and A is a set.

Proposition 4.2. Suppose r is a relation and $\mathcal{A}$ is a family of sets. Then

$$r[\bigcup \mathcal{A}] = \bigcup\{r[A] : A \in \mathcal{A}\}.$$

Proof. Suppose $y \in r[\bigcup \mathcal{A}]$. Then there is $x \in \bigcup \mathcal{A}$ such that $(x, y) \in r$. Since $x \in \bigcup \mathcal{A}$ there is $A \in \mathcal{A}$ such that $x \in A$ so $y \in r[A]$ so $y \in \bigcup\{r[A] : A \in \mathcal{A}\}$.

On the other hand, suppose $y \in \bigcup\{r[A] : A \in \mathcal{A}\}$. There is $A \in \mathcal{A}$ such that $y \in r[A]$. In particular, there is $x \in A$ such that $(x, y) \in r$ so, as $x \in \bigcup \mathcal{A}$, $y \in r[\bigcup \mathcal{A}]$. □

Remark 4.4. We will give an example shortly of a relation r and a nonempty family of sets $\mathcal{A}$ such that $r[\bigcap \mathcal{A}] \neq \bigcap\{r[A] : A \in \mathcal{A}\}$.

Definition 4.2. Whenever A and B are sets we let

$$A \times B = \{(a, b) : a \in A \text{ and } b \in B\}$$

and call this set the **Cartesian product of** A **and** B. If A is a set we say r is a **relation on** A if $r \subset A \times A$ and we frequently write

$$x\, r\, y \quad \text{instead of} \quad (x, y) \in r.$$

4.2. **Functions.** We say a relation f is a **function** if

$$(x, y_1) \in f \text{ and } (x, y_2) \in f \Rightarrow\ y_1 = y_2.$$

In other words, $f[\{x\}]$ has at most one member for any object x. Whenever f is a function and $x \in \mathbf{dmn}\, f$ we let

$$f(x) \quad \text{or} \quad f_x$$

be the unique member of $f[\{x\}]$ and call this object the **image of** x **under** f. We write

$$f : X \to Y$$

and say f **is a function from** X **to** Y if

(i) f is a function;
(ii) $X = \mathbf{dmn}\, f$; and
(iii) $\mathbf{rng}\, f \subset Y$.

Note that if $\mathbf{rng}\, f \subset Y_i$, $i = 1, 2$, then

$$f : X \to Y_i, \quad i = 1, 2.$$

We say the function f is **univalent** if

$$(x_1, y) \in f \text{ and } (x_2, y) \in f \Rightarrow\ x_1 = x_2;$$

this amounts to saying that f^{-1} is a function which we call the **inverse function to** f. Note that if f and g are functions then $g \circ f$ is a function whose domain is $f^{-1}[\mathbf{dmn}\, g]$ and whose range is $g[\mathbf{rng}\, f]$.

Proposition 4.3. Suppose f is a function and B is a set. Then

$$x \in f^{-1}[B] \ \Leftrightarrow\ f(x) \in B.$$

Proof. We have $x \in f^{-1}[B]$ iff for some y, $y \in B$ and $(y, x) \in f^{-1}$ iff for some y, $y \in B$ and $(x, y) \in f$ iff for some y, $y \in B$ and $y = f(x)$ iff $f(x) \in B$. □

Here is an extremely useful fact about functions.

Proposition 4.4. Suppose $\mathcal{A}$ is a family of sets. Then

$$f^{-1}[\bigcap \mathcal{A}] = \bigcap \{f^{-1}[A] : A \in \mathcal{A}\} \quad \text{and} \quad f^{-1}[\bigcup \mathcal{A}] = \bigcup \{f^{-1}[A] : A \in \mathcal{A}\}.$$

Suppose A and B are sets. Then

$$f^{-1}[A \sim B] = f^{-1}[A] \sim f^{-1}[B].$$

Proof. Exercise for the reader. Note that

$$f^{-1}[\bigcup \mathcal{A}] = \bigcup \{f^{-1}[A] : A \in \mathcal{A}\}$$

is a special case of a previous Proposition. □

Remark 4.5. Keeping in mind that forward images of unions are preserved by relations we see that this Proposition says that all the set theoretic operations are preserved by taking the counterimage under a function.

That forward images of unions are preserved under functions follows from earlier work. This is *not* true for forward images of intersections as the following simple example illustrates.

Example 4.1. Let $f = \{(0, 0), (1, 0)\}$, let $A = \{0\}$ and let $B = \{1\}$. Then

$$f[A \cap B] = f[\emptyset] = \emptyset \neq \{0\} = f[A] \cap f[B].$$

Exercise 4.3. Prove or disprove:

$$f[A \sim B] = f[A] \sim f[B]$$

whenever f is a function and A and B are sets.

Example 4.2. Let

$$r = \{(a, a), (\{a\}, a)\}.$$

Then r is a function, $r(a) = a$ and $r[\{a\}] = \{a\}$.

One often uses the following elementary Proposition.

Proposition 4.5. Suppose f is a function and B is a set. Then

$$x \in f^{-1}[B] \;\Leftrightarrow\; x \in \mathbf{dmn}\, f \text{ and } f(x) \in B.$$

Proof.

$$\begin{aligned} x \in f^{-1}[B] &\Leftrightarrow \text{ for some } y,\ (y,x) \in f^{-1} \text{ and } y \in B \\ &\Leftrightarrow \text{ for some } y,\ (x,y) \in f \text{ and } y \in B \\ &\Leftrightarrow \text{ for some } y,\ x \in \mathbf{dmn}\, f,\ y = f(x) \text{ and } y \in B \\ &\Leftrightarrow x \in \mathbf{dmn}\, f \text{ and } f(x) \in B. \end{aligned}$$

□

4.3. **More on relations.**

Definition 4.3. Suppose r is a relation on the set X.

r is **reflexive** if $(x,x) \in r$ for all $x \in X$;

r is **irreflexive** if $(x,x) \notin r$ for all $x \in X$;

r is **symmetric** if $(x,y) \in r \;\Rightarrow\; (y,x) \in r$;

r is **anstisymmetric** if $(x,y) \in r \;\Rightarrow\; (y,x) \notin r$;

r is **transitive** if $(x,y) \in r$ and $(y,z) \in r \;\Rightarrow\; (x,z) \in r$;

r is **trichotomous** if for all $(x,y) \in r$ exactly one of the following holds:

$$(x,y) \in r; \quad x = y; \quad (y,x) \in r.$$

4.4. **Partial orderings.** Let X be a set.

Definition 4.4. r is a **partial ordering of** X if r is a relation on X which is irreflexive and transitive.

r is a **linear ordering of** X if r is a partial ordering of X and r is trichotomous.

r is a **well ordering of** X if r is a linear ordering of X and for each nonempty subset of A of X there is $a \in A$ such that $(a,x) \in r$ whenever $x \in A$.

Remark 4.6. Some authors define a partial ordering on X to be a transitive relation on X.

Theorem 4.4. Suppose r is a relation on X, $W \subset X$ and $q = r \cap (W \times W)$.

If r is partial ordering of X then q is a partial ordering of W.

If r is linear ordering of X then q is a linear ordering of W.

If r is well ordering of X then q is a well ordering of W.

Proof. Exercise. □

For the remainder of this subsection we fix a partial ordering $<$ *of* X. We will write

$$x \le y \text{ if } x < y \text{ or } x = y;$$

Proposition 4.6. Suppose $(x,y) \in X \times X$. Then at most one of the following holds:

$$x < y; \quad x = y; \quad y < x.$$

In particular, $<$ is antisymmetric.

Proof. Were it the case that $x < y$ and $y < x$ we would have $x < x$ by transitivity which is impossible since $<$ is irreflexive. □

Proposition 4.7. Suppose

$$\prec = \{(U,V) \in 2^X \times 2^X : U \subset V \text{ and } U \neq V\}.$$

Then $\prec$ is a partial ordering of 2^X.

Moreover, if $f : X \to 2^X$ is such that

$$f(x) = \{w \in X : w \leq x\} \quad \text{for } x \in X.$$

Then f is univalent and

$$x < y \ \Rightarrow f(x) \prec f(y).$$

Proof. We leave as an exercise for the reader to prove that $\prec$ is a partial ordering of 2^X.

Suppose $x, y \in X$, $x \neq y$ and $f(x) = f(y)$. Were it the case $f(x) = f(y)$ we would have $x \in f(x) = f(y)$ so $x < y$ and $y \in f(y) = f(x)$ so $y < x$ which is incompatible with ??. Thus f is univalent.

Suppose $x < y$. Then $f(x) \subset f(y)$ by the transitivity of $<$. Also, $y \in f(y) \sim f(x)$ by ?? so $f(x) \neq f(y)$. Thus $f(x) \prec f(y)$. □

Remark 4.7. The Proposition says that $\prec$ is the mother of all partial orderings on X.

Definition 4.5. Suppose $A \subset X$.

We say a u is an **upper bound for** A if $u \in X$ and

$$a \in A \ \Rightarrow \ a \leq u.$$

We say g is a **greatest member of** A if $g \in A$ and g is an upper bound for A.

We say M is a **maximal member of** A if $M \in A$ and

$$M < a \text{ for no } a \in A.$$

We say l is an **lower bound for** A if $l \in X$ and

$$a \in A \ \Rightarrow \ l \leq a.$$

We say l is a **least member of** A if $l \in A$ and l is a lower bound for A.

We say m is a **minimal member of** A if $m \in A$ and

$$a < m \text{ for no } a \in A.$$

Remark 4.8. If $A = \emptyset$ and $x \in X$ then x is both upper bound for A and a lower bound for A.

Proposition 4.8. Suppose $A \subset X$. Then

(i) a greatest member of A is a maximal member of A;
(ii) A has at most one greatest member;
(iii) a least member of A is a minimal member of A;
(iv) A has at most one least member.

Proof. Straightforward exercise for the reader. □

Definition 4.6. Suppose $A \subset X$.

A least member of the set of upper bounds for A is called a **least upper bound for** A. By Proposition 4.8 any least upper bound for A is unique; it is called **the supremum of** A and it is denoted

$$\sup A \quad \text{or} \quad \mathbf{l.u.b.}\, A.$$

A greatest member of the set of lower bounds for A is called a **greatest lower bound for** A**.** By Proposition 4.8 any greatest lower bound for A is unique; it is called **the infimum of** A and is denoted

$$\inf A \quad \text{or} \quad \mathbf{g.l.b.}\, A.$$

Proposition 4.9. The following two conditions are equivalent.

(i) If A is a nonempty subset of X and there is an upper bound for A then there is a least upper bound for A.

(ii) If A is a nonempty subset of X and there is a lower bound for A then there is a greatest lower bound for A.

Proof. For any subset A of X let $U(A)$ be the set of upper bounds for A and let $L(A)$ be the set of lower bounds for A.

Suppose (i) holds, $A \subset X$ and $A \neq \emptyset$ and $L(A) \neq \emptyset$. Note that $A \subset U(L(A))$. Thus $L(A)$ and $U(L(A))$ are nonempty so $U(L(A))$ has a least member b. I claim that b is a greatest lower bound for A.

Since $A \subset U(L(A))$ we have $b \leq a$ for $a \in A$ so $b \in L(A)$. Suppose $c \in L(A)$. Since $b \in U(L(A))$ we have $c \leq b$. That is, b is a greatest lower bound for A.

In a similar fashion one shows that (ii) implies (i). □

Definition 4.7. We say $<$ is **complete** if either of the equivalent conditions in the above Proposition holds.

Example 4.3. Suppose $\prec$ is the partial order on 2^X in ??. Let $\mathcal{A}$ be a nonempty family of subsets of X. That is, $\mathcal{A} \subset 2^X$ and $\mathcal{A} \neq \emptyset$. I claim that $\cup\mathcal{A}$ is a least upper bound for $\mathcal{A}$ (with respect to $\prec$).

Suppose $A \in \mathcal{A}$. If $x \in A$ then $x \in \cup\mathcal{A}$ so $A \subset \cup\mathcal{A}$. Thus $\cup\mathcal{A}$ is an upper bound for $\mathcal{A}$.

Suppose the subset B of X is an upper bound for $\mathcal{A}$. If $x \in \cup\mathcal{A}$ then there is $A \in \mathcal{A}$ such that $x \in A$. Since B is an upper bound for $\mathcal{A}$ we have $A \subset B$. Thus $x \in B$ and we have show that $\cup\mathcal{A} \subset B$.

This verifies my claim.

Exercise 4.4. Let $\mathcal{A}$ be a nonempty family of subsets of X. Show that $\cap\mathcal{A}$ is a greatest lower bound for $\mathcal{A}$.

4.4.1. *Linear orderings. We fix a linear ordering $<$ of X.*

Proposition 4.10. Suppose $A \subset X$. Then

(i) the set of maximal members of A has at most one member;

(ii) the set of least upper bounds for A has at most one member;

(iii) the set of maximal members of A has at most one member;

(iv) the set of greatest lower bounds for A has at most one member;

Proof. Straightforward exercise for the reader. □

Definition 4.8. Suppose $<$ linearly orders X and $I \subset X$. Then I is an **interval** if

$$x, z \in I,\ y \in X \text{ and } x < y < z \ \Rightarrow\ y \in I.$$

Definition 4.9. A subset I of X is an **initial segment** if

$$x \in X,\ y \in I \text{ and } x < y \ \Rightarrow\ x \in I.$$

Trivially, $\emptyset$ and X are initial segments.

Proposition 4.11. An initial segment is an interval. The union of a family of initial segments is an initial segment. The intersection of a nonempty family of initial segments is an initial segment.

Exercise 4.5. Prove this Proposition.

Proposition 4.12. Suppose I and J are initial segments. Then either $I \subset J$ or $J \subset I$.

Proof. It will suffice to show that

$$J \sim I \neq \emptyset \;\Rightarrow\; I \subset J. \tag{4}$$

So suppose $y \in J \sim I$ and, contrary to (4), there is $x \in I \sim J$. Since $x \neq y$ we have *either* (i) $x < y$ *or* $y < x$ by trichotomy. If $x < y$ we have $x \in J$ since J is an initial segment and if $y < x$ we have $y \in I$ since I is an initial segment, neither of which is possible since $x \notin J$ and $y \notin I$. Thus (4) holds. □

Remark 4.9. let $\mathcal{I}$ be the family of initial segments. For $I, J \in \mathcal{I}$ we let

$$I \prec J \quad \text{if } I \subset J \text{ and } I \neq J.$$

So $\prec$ is the intersection of $\prec$ in ?? with $\mathcal{I} \times \mathcal{I}$.

From Proposition 4.12 we see that $\prec$ is linear.

Now suppose $\mathcal{A}$ is a nonempty subfamily of $\mathcal{I}$. By Proposition 4.11 we find that $J = \cup\mathcal{A} \in \mathcal{I}$. It is easy to see that J is a least upper bound for $\mathcal{A}$ with respect to $\prec$.

In ?? we will construct the real numbers $\mathbb{R}$ from the rational numbers $\mathbb{Q}$ by using a slight variant of this construction.

Definition 4.10. For each $x \in X$ we let

$$\mathbf{i}(x) = \{w \in X : w < x\}.$$

Proposition 4.13. $\mathbf{i}(x)$ is an initial segment for any $x \in X$.

Proof. This follows directly from the transitivity of $<$. □

Theorem 4.5. The linear ordering $<$ is complete if and only if the only initial segments are $\emptyset$, X and the sets

$$\mathbf{i}(x) \text{ and } \mathbf{i}(x) \cup \{x\}, \quad x \in X.$$

Exercise 4.6. Prove this Theorem.

4.5. **Well ordering.** Note that $\emptyset$ is a well ordering of $\emptyset$.

We now suppose $X \neq \emptyset$ and we fix a well ordering $<$ of X.

For each nonempty subclass A of X we let

$$\mathbf{l}(A)$$

be the $<$-least member of A. We let

$$\mathbf{o} = \mathbf{l}(X).$$

Proposition 4.14. Suppose I is an initial segment of X and $I \neq X$. Then $I = \mathbf{i}(\mathbf{l}(X \sim I))$.

Proof. Let $x = \mathbf{l}(X \sim I)$. If $w \in \mathbf{i}(x)$ then $w \in I$ since $w < x$ and $x = \mathbf{l}(X \sim I)$; thus $\mathbf{i}(x) \subset I$. Were there $y \in I \sim \mathbf{i}(x)$ we would have $x < y$ since $x \notin I$ and that would imply $x \in I$ since I is an initial segment; thus $I \subset \mathbf{i}(x)$. □

Definition 4.11. We say $x \in X$ is a **limit point** if $\{w \in X : w < x\}$ has no largest member.

Proposition 4.15. Suppose $x \in X$ and x is a limit point. Then $X = \bigcup\{\mathbf{i}(w) : w \in X \text{ and } w < x\}$.

Proposition 4.16. Suppose I and J are $<$-initial segments of X. Then *either* $I \subset J$ *or* $J \subset I$.

Proof. If either I or J equals X the Proposition holds trivially. So, by virtue of the preceding Proposition, there are x and y in X such that $I = \mathbf{i}(x)$ and $J = \mathbf{i}(y)$. By trichotomy, $x < y$, $x = y$ or $y < x$. If $x < y$ and $w \in I$ then $w < x$ so $w < y$ and $w \in J$ by transitivity of $<$ so $I \subset J$; if $x = y$ then $I = J$; if $y < x$ and $w \in J$ then $w < x$ and $w \in I$ by transitivity of $<$ so $J \subset I$. □

Definition 4.12. Let

$$\tilde{X} = \{x \in X : x < y \text{ for some } y \in X\}.$$

(So $\tilde{X} = X$ if X has no largest member and $\tilde{X} = X \sim \{b\}$ if X has a largest member b.)

Let

$$S : \tilde{X} \to X$$

be such that

$$S(x) = \mathbf{l}(\{y \in X : x < y\}) \quad \text{whenever } x \in \tilde{X}.$$

We call S the **successor function**.

Proposition 4.17. $X \sim \mathbf{rng}\, S = \{\mathbf{o}\} \cup \{x \in X : x \text{ is a limit point}\}$.

Proof. Exercise for the reader. □

Definition 4.13. We say a subset A of X is **inductive** if $\mathbf{o} \in A$ and

$$\mathbf{i}(x) \subset A \;\Rightarrow\; x \in A \quad \text{whenever } x \in X.$$

Theorem 4.6. The principle of transfinite induction. Suppose A is an inductive subset of X. Then $A = X$.

Proof. Suppose, to the contrary, that $X \sim A$ is nonempty and let $x = \mathbf{l}(X \sim A)$. Then $\mathbf{i}(x) \subset A$ by trichotomy which implies $x \in A$ since A inductive. This is a contradiction. □

Theorem 4.7. Defining a Function by Transfinite Induction. Suppose

(i) Y is a set;
(ii) $\mathcal{G} = \bigcup\{Y^{\mathbf{i}(x)} : x \in X\}$;
(iii) $G : \mathcal{G} \to Y$.

There is one any only one f such

$$f : X \to Y$$

and such that

$$f(x) = G(f|\mathbf{i}(x)) \quad \text{whenever } x \text{ in } X.$$

Proof. Let $\mathcal{H}$ be the family of functions h mapping initial segments J into Y such that

$$h(x) = G(h|\mathbf{i}(x)) \quad \text{whenever } x \in J.$$

Let $f = \cup\mathcal{H}$.

Lemma 4.1. If $h_1, h_2 \in \mathcal{H}$ then either $h_1 \subset h_2$ or $h_2 \subset h_1$.

Proof. Let J_1, J_2 be the domains of h_1, h_2, respectively, and note that, by a previous Proposition, $J_i \subset J_j$ where $\{i, j\} = \{1, 2\}$. Let $K = \{x \in J_i : h_i(x) = h_j(x)\}$. Then

$$h_i(\mathbf{o}_i) = G(h_i|\mathbf{i}(\mathbf{o}_i)) = G(\emptyset) = G(h_j|\mathbf{i}(\mathbf{o}_j)) = h_j(\mathbf{o}_j)$$

so $\mathbf{o}_i \in K$. Suppose $w \in K$ and $\mathbf{i}(w) \subset K$. Then

$$h_i(w) = G(h_i|\mathbf{i}(w)) = G(h_j|\mathbf{i}(w)) = h_j(w)$$

so $w \in K$. Thus K is an inductive subset of J_i so $K = J_i$. This implies $h_i \subset h_j$. $\square$

It follows from the Lemma that f is a function. It is a simple matter to show that its domain is inductive as is the set of $x \in X$ such that $f(x) = G(f|\mathbf{i}(x))$. $\square$

Theorem 4.8. Suppose X_i is well ordered by $<_i$, $i = 1, 2$. Then there is one and only one function f such that if $W_1 = \mathbf{dmn}\, f$ and $W_2 = \mathbf{rng}\, f$ then

(i) W_i is a $<_i$-initial segment for $i = 1, 2$;
(ii) if $w, x \in W_1$ and $w <_1 x$ then $f(w) <_2 f(x)$;
(iii) $f[\{w \in X_1 : w <_1 x\}] = \{y \in X_2 : y < f(x)\}$ for each $x \in X_1$;
(iv) exactly one of the following holds:
  (A) $W_1 = X_1$ and $W_2 \neq X_2$;
  (B) $W_1 = X_1$ and $W_2 = X_2$;
  (C) $W_1 \neq X_1$ and $W_2 = X_2$.

Proof. For each $i = 1, 2$ let $\mathbf{l}_i(A_i)$ be the $<_i$-least member of A_i whenever A_i is a nonempty subset of X_i and let $\mathbf{i}_i(x) = \{w \in X_i : w <_i x\}$ whenever $x \in X_i$.

Let $\mathcal{G} = \bigcup\{X_2^{\mathbf{i}_1(x)} : x \in X_1\}$ and define $G : \mathcal{G} \to X_2 \cup \{\emptyset\}$ by setting

$$G(g) = \begin{cases} \mathbf{l}_2(X_2 \sim \mathbf{rng}\, g) & \text{if } \mathbf{rng}\, g \neq X_2, \\ X_2 & \text{if } \mathbf{rng}\, g = X_2 \end{cases} \quad \text{for } g \in \mathcal{G}.$$

By the preceding Theorem there is one and only one $F : X_1 \to X_2 \cup \{X_2\}$ be such that

$$F(x) = G(F|\mathbf{i}_1(x)) = \begin{cases} \mathbf{l}_2(X_2 \sim F[\mathbf{i}_1(x)]) & \text{if } F[\mathbf{i}_1(x)] \neq X_2, \\ X_2 & \text{if } F[\mathbf{i}_1(x)] = X_2 \end{cases} \quad \text{for } x \in X_1.$$

Let I be the set of $x \in X_1$ such that

$$F[\mathbf{i}_1(x)] \in \{\mathbf{i}_2(F(x))\}, X_2\}.$$

I claim that I is inductive. Clearly, $\mathbf{l}_1(X_1) \in I$. Suppose $x \in X_1$, $\mathbf{i}_1(x) \subset I$ and $F[\mathbf{i}_1(x) \neq X_2$.

Case One. $x = \mathbf{l}_1(X_1 \sim \mathbf{i}_1(w))$ **for some** $w \in \mathbf{i}_1(x)$**.** Since $\mathbf{i}_1(x) = \mathbf{i}_1(w) \cup \{w\}$ we have that $J = \mathbf{i}_2(F(w)) + \{F(w)\}$ is a $<_2$ initial segment

$$\begin{aligned} F(x) &= \mathbf{l}_2(X_2 \sim F[\mathbf{i}_1(x)]) \\ &= \mathbf{l}_2(X_2 \sim (F[\mathbf{i}_1(w)] \cup \{F(w)\})) \\ &= \mathbf{l}_2(X_2 \sim (\mathbf{i}_2(F(w)) \cup \{F(w)\})) \\ &= \mathbf{l}_2(X_2 \sim J); \end{aligned}$$

by ?? we find that $J = \mathbf{i}_2(F(x))$. Thus $x \in I$.

Case Two. x is a $<_1$ limit point. Since $\mathbf{i}_1(x) = \bigcup\{\mathbf{i}_1(w) : w \in \mathbf{i}_1(x)\}$ we have

$$\begin{aligned} F[\mathbf{i}_1(x)] &= F\left[\bigcup\{\mathbf{i}_1(w) : w \in X_1 \text{ and } w <_1 x\}\right] \\ &= \bigcup\{F[\mathbf{i}_1(w)] : w \in X_1 \text{ and } w <_1 x\} \\ &= \bigcup\{\mathbf{i}_2(F(w)) : w \in X_1 \text{ and } w <_1 x\} \\ &= \mathbf{i}_2(y) \end{aligned}$$

for some $y \in X_2$ Now

$$F(x) = \mathbf{l}_2(X_2 \sim F[\mathbf{i}_1(x)]) = \mathbf{l}_2(X_2 \sim \mathbf{i}_2(y));$$

by ?? we find that $\mathbf{i}_2(y) = \mathbf{i}_2(F(x))$. Thus $x \in I$.

Let $f = F|\{x \in X_1 : F[\mathbf{i}_1(x)] \neq X_2\}$. □

4.6. **Ordinals.** We let

$$0 = \emptyset.$$

Definition 4.14. For a set x we let

$$\mathbf{w}(x) = \{(v, w) \in x \times x : v \in w\}.$$

A set x is an ordinal if $\mathbf{w}(x)$ is a well ordering of x and every member of x is a subset of x.

$R = \{x : x \text{ is an ordinal}\}$.

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

Proposition 4.18. Suppose x is an ordinal. Then I is a $\mathbf{w}(x)$-initial segment if and only if either $I = x$ or $I = w$ for some $w \in X$.

Proof. Suppose I is a $\mathbf{w}(x)$-initial segment and $I \neq x$. By ?? there is $w \in x$ such that $I = \{v : v \in x \text{ and } v \in w\}$. Since $w \subset x$ we find that $\{v : v \in x \text{ and } v \in w\} = \{v : v \in w\} = w$ for some $w \in x$ so $I = w$.

On the other hand, if $w \in x$, $v \in w$ and $u \in v$ then then w is a $\mathbf{w}(x)$ initial segment since $w \subset x$. □

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

Theorem 4.9. 0 is an ordinal.

Theorem 4.10. Suppose x is an ordinal. Then $x \notin x$

Proof. Were it the case that $x \in x$ we would have $(x, x) \in \mathbf{w}(x)$ which is impossible since $\mathbf{w}(x)$ is irreflexive. □

Theorem 4.11. Suppose x is an ordinal and x is a set. Then $x \cup \{x\}$ is an ordinal.

Theorem 4.12. Suppose x is an ordinal and $x \neq 0$. The 0 is the $\mathbf{w}(x)$-least member of x.

Proof. Let w be the $\mathbf{w}(x)$-least member of x. Were $v \in w$ for some $w \in x$ we would have $v \in x$ so that $(v, w) \in \mathbf{w}(x)$ which is a contradiction. Thus $w = 0$. □

Theorem 4.13. Suppose x and y are ordinals, $x \subset y$ and $x \neq y$. Then $x \in y$.

Proof. Suppose $v \in x$ and $u \in v$. Then $u \in x$ because x is an ordinal. Thus x is a $\mathbf{w}(y)$-initial segment of y. Since $x \neq y$ there is by Theorem ?? a member w of y such that $x = \{t \in y : t \in w\}$. Since every member of w is a member of y, $x = \{t : t \in w\} = w \in y$. □

Theorem 4.14. Suppose x is an ordinal and $w \in x$. Then w is an ordinal.

Proof. That $\mathbf{w}(w)$ is a well ordering of w follows from ??. Suppose $v \in w$ and $u \in v$. Since $w \subset x$ we have $v \in x$ and this implies $v \subset x$. Thus $u \in x$. Since $\{u, v, w\} \subset x$, $u \in v$ and $v \in w$ by the transitivity of $\mathbf{w}(x)$ we find that $u \in w$. Thus $v \subset w$. □

Theorem 4.15. Suppose x and y are ordinals. Then either $x \subset y$ or $y \subset x$.

Proof.

Lemma 4.2. Suppose x and y are ordinals. Then $x \cap y$ is an ordinal.

Proof. Clearly, $\mathbf{w}(x \cap y)$ is a well ordering of $x \cap y$. Suppose $w \in x \cap y$. Then $w \subset x$ since x is an ordinal and $w \subset y$ since y is an ordinal so $w \subset x \cap y$. Thus $x \cap y$ is an ordinal. □

Suppose, to the contrary, $x \sim y \neq \emptyset$ and $y \sim x \neq \emptyset$. Since $x \cap y$ is an ordinal we infer from Theorem ?? that $x \cap y \in x$ and $x \cap y \in y$. Thus $x \cap y \in x \cap y$ and this contradicts Theorem ??. □

Theorem 4.16. Suppose x and y are ordinals. The exactly one of the following holds:

$$x \in y; \quad x = y; \quad y \in x.$$

Proof. Suppose $x \subset y$ and $x \neq y$. Let w be the $\mathbf{w}(y)$-first member of $y \sim x$. Then $x = w \in y$. Were it the case that $y \in x$ we would have $x \in x$. □

Theorem 4.17. R is an ordinal and R is not a set.

Proof. That $\mathbf{w}(R)$ is a well ordering of R follows directly from the preceding Theorem. Suppose $x \in R$ and $w \in x$. Then w is a set and w is an ordinal by Theorem ??; so $w \in R$ and thus $x \subset R$. □

4.7. **Ordinals.** We let

$$0 = \emptyset.$$

Definition 4.15. We let

$$E = \{(x, y) : x \in y\}.$$

Definition 4.16. For any set x we let $\mathbf{S}(x) = x \cup \{x\}$.

Definition 4.17. We let

$$\mathbf{s} = \{y : \exists x\, ((x \text{ is a set}) \wedge (y = (x, x \cup \{x\})))\}.$$

Definition 4.18. Whenever x is a set we let

$$\mathbf{w}(x) = \{(v, w) \in X \times X : v \in w\}$$

and we say x is an **ordinal** if

(i) $\mathbf{w}(x)$ is a well ordering of x;
(ii) every member of x is a subset of x.

Definition 4.19. We say x is an **ordinal** if

(i) $E|(x \times x)$ is a well ordering of x;
(ii) every member of x is a subset of x.

Theorem 4.18. 0 is an ordinal. Moreover, if x is an ordinal and $x \neq 0$ then 0 is the $\mathbf{w}(x)$-first member of x.

Proof. Since $\mathbf{w}(0) = 0$ we find that $\mathbf{w}(0)$ is, trivially, a well ordering of 0. Since 0 has no members it is trivial that every member of 0 is a subset of 0.

Suppose x is an ordinal and $x \neq 0$. Let w be the $\mathbf{w}(x)$-least member of x. Were it the case that $v \in w$ for some v we would have $v \in x$ since $w \subset x$. That implies that $(v, w) \in \mathbf{w}(x)$ which is impossible since w is the $\mathbf{w}(x)$-first member of x. Thus $v \in w$ for no v so $w = 0$. □

Theorem 4.19. Suppose x is an ordinal and $y \in x$. Then $y = \{w \in x : w \in y\}$ and y is an ordinal.

Proof. Let $Y = \{w \in x : w \in y\}$. That $Y \subset y$ is immediate. Since $y \subset x$ we find that $y \subset Y$. Thus $y = Y$.

It follows from ?? that $\mathbf{w}(y)$ is a well ordering of y.

Let $I = \{w \in y : v \in w \;\; \Rightarrow \;\; v \subset w\}$. Then $0 \in I$. Suppose $w \in x$ and $\{v \in x : v \in w\} \subset I$.

□

Theorem 4.20. Suppose x is an ordinal, I is a $\mathbf{w}(x)$-initial segment of x and $I \neq x$. Then $I = y$ for some $y \in x$.

Proof. Let y be the $\mathbf{w}(x)$-least member of $x \sim I$. By ?? $I = \{w \in x : w \in y\}$. So $I = y$ by the preceding Theorem. □

Theorem 4.21. Suppose x and y are ordinals. Then exactly one of the following holds:

$$x \in y; \quad x = y; \quad y \in x.$$

Proof. By Theorem ?? there is a function f such that if $W_x = \mathbf{dmn}\, f$ and $W_y = \mathbf{rng}\, f$ then

(i) W_x is a $\mathbf{w}(x)$-initial segment and W_y is a $\mathbf{w}(y)$-intial segment;
(ii) if $u, v \in W_x$ and $(u, v) \in \mathbf{w}(x)$ then $(f(u), f(v)) \in \mathbf{w}(y)$;
(iii) exactly one of the following holds:
  (A) $W_x = x$ and $W_y \neq y$;
  (B) $W_x = x$ and $W_y = y$;
  (C) $W_x \neq x$ and $W_y = y$.

I claim that $f(w) = w$ for $w \in W_x$. If this were not true we could let w be the $\mathbf{w}(x)$-least member of W_x such that $f(w) \neq w$. Then

$$\begin{aligned} f(w) &= f[\{v \in W_x : (v, w) \in \mathbf{w}(x)\}] \\ &= \{f(v) : v \in W_x \text{ and } (v, w) \in \mathbf{w}(x)\} \\ &= \{v \in W_x : (v, w) \in \mathbf{w}(x)\} \\ &= w \end{aligned}$$

which is impossible. The Theorem follows.

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

$I = \{w \in W_x : f(w) = w\}$. $0 \in I$. Suppose $v \in W_x$ and $\{u \in W_x : u \in v\} \subset I$. Since $v = \{u \in W_x : u \in v\}$ we find that

$$\begin{aligned} f(v) &= f[\{u \in W_x : u \in v\}] \\ &= \{f(u) : u \in W_x \text{ and } u \in v\} \\ &= \{u \in W_x : u \in v\} \\ &= v \end{aligned}$$

so I is and inductive subset of W_x and, therefore, our claim holds. □

Proof.

Lemma 4.3. $x \cap y$ is an ordinal.

Proof. It is obvious that $\mathbf{w}(x \cap y)$ is a linear ordering of $x \cap y$. Suppose C is a nonempty subset of $x \cap y$. Let a equal the $\mathbf{w}(x)$-least member of A. If $b \in C$ then, as $C \subset X$, we have $a \in b$ so a is the $\mathbf{w}(x \cap y)$-least member of $x \cap y$. Thus $\mathbf{w}(x \cap y)$ is a well ordering of $x \cap y$. Suppose $w \in x \cap y$. Since x is an ordinal $w \subset x$ and since y is an ordinal $w \subset y$. Thus $w \subset x \cap y$. □

Suppose neither $x \sim (x \cap y)$ nor $y \sim (x \cap y)$ are empty. Let v equal the $\mathbf{w}(x)$-first member of $x \sim (x \cap y)$ and let w be the $\mathbf{w}(y)$-first member of $y \sim (x \cap y)$. Since $x \cap y$ is an ordinal we have $v = (x \cap y) \cup \{x \cap y\} = w$. Thus *either* $x \subset y$ *or* $y \subset x$.

Lemma 4.4. Suppose x and y are ordinals, $x \subset y$ and $x \neq y$. Then $x \in y$.

Proof. Let w be the $\mathbf{w}(y)$-least member of $y \sim x$. Let $z = \{v \in y : (v, w) \in \mathbf{w}(y)\}$; then $z \in y$ and $x \subset z$. I claim that $x = z$. If this were not so there would be a $\mathbf{w}(y)$-least member u of $z \sim x$. □

□

Theorem 4.22. If x, y and z are ordinals, $x \in y$, and $y \in z$, then $x \in z$.

Proof. We have $x \in z$ or $x = z$ or $z \in x$. By the preceding Theorem we cannot have $y \in x$. Were it the case that $x = z$ we would have $y \in z = x$. Were it the case that $z \in x$ we would have $y \in x$ since $y \subset z$. □

Theorem 4.23. Suppose C is a nonempty set of ordinals. There is $x \in C$ such that $x \in y$ for all $y \in C \sim \{x\}$.

Proof. Suppose $z \in C$.

Suppose $z \cap C = 0$. Then $\{w \in C : w \in z\} = 0$. If $y \in C \sim \{z\}$ then $z \in y$ since otherwise $z = y$ or $y \in z$ neither of which is possible. So we may let $x = z$.

Suppose $z \cap C \neq 0$. The x equal the $\mathbf{w}(z)$-least member of $z \cap C$. Suppose $y \in C \sim \{x\}$. Then either $x \in y$ or $y \in x$ by ??. Were $y \in x$ we would have $y \in z \cap C$ which would imply that $x \in y$. □

4.8. **The natural numbers. Axiom.** There is an ordinal $\mathbb{N}$ such that

(i) $0 \in \mathbb{N}$;
(ii) the domain of the successor function of $\mathbb{N}$ equals $\mathbb{N}$;
(iii) the range of the successor function of $\mathbb{N}$ equals $\mathbb{N} \sim \{0\}$.

4.9. **Choice functions and the axiom of choice.**

Definition 4.20. Suppose $\mathcal{C}$ is a family of nonempty sets. We say c is a **choice function for** $\mathcal{C}$ if c is a function, $\mathbf{dmn}\, c = \mathcal{C}$ and

$$c(C) \in C \quad \text{whenever } C \in \mathcal{C}.$$

The Axiom of Choice (AC). Suppose $\mathcal{C}$ is a family of nonempty sets. Then there is a choice function for $\mathcal{C}$.

The Well Ordering Axiom (WO). Every set can be well ordered.

Theorem 4.24. (AC)$\Leftrightarrow$(WO).

Proof. Suppose (AC). Let X be a set and let ξ be a choice function on $2^X \sim \{\emptyset\}$. We will show that there is one and only one well ordering of X such that for any initial segment J of X not equal to X the least member of $X \sim J$ is $\xi(X \sim J)$. We prove this as follows. We let $\mathcal{W}$ be the set of ordered pairs(w, W) such that W is a subset of X, w is a well ordering of W and such that if J is a w-initial segment of W and $J \neq W$ then $\xi(X \sim J)$ is the w-least least $W \sim J$. Then

$$\bigcup \{w : \text{ for some } W,\ (w, W) \in \mathcal{W}\}.$$

is the desired well ordering of X. We leave the details to the reader; use the ideas in our previous results on well ordering.

Now suppose (WO) and let $\mathcal{C}$ be a family of nonempty sets. Let $X = \{(C, c) : C \in \mathcal{C} \text{ and } c \in C\}$ and let w be a well ordering of X. Then

$$\{(C, c) : C \in \mathcal{C} \text{ and } (C, c) \text{ is the } w\text{-least member of } \{(C, b) : b \in C\}\}$$

is a choice function for $\mathcal{C}$. □

Henceforth we go with the crowd and assume (AC).

4.10. **Equipotence and cardinal numbers.** Suppose X and Y are sets. We say that X is **equipotent with** Y and write

$$X \approx Y$$

if there exists a relation f such that $f : X \to Y$ and $f^{-1} : Y \to X$. It is obvious that

$$X \approx X$$

whenever X is a set;

$$X \approx Y \ \Rightarrow Y \approx X$$

whenever X and Y are sets and

$$X \approx Y \text{ and } Y \approx Z \ \Rightarrow \ X \approx Z$$

whenever X, Y and Z are sets. Thus we have introduced an equivalence relation on the set of all sets the equivalence classes corresponding to which are called **cardinal numbers**. (Forming the set of all sets was never allowed in public, was allowed in secret until about forty years ago and is forbidden under any circumstances today!)

Theorem 4.25. Suppose X is nonempty set and

$$f : X \to 2^X.$$

Then

$$\mathbf{rng}\, f \neq 2^X.$$

Remark 4.10. This simple but fundamental Theorem says that 2^X is larger than X when X is nonempty. The proof is an abstraction (the right one, in my opinion) of what is called Cantor's diagonal argument.

Proof. Let $A = \{x \in X : x \notin f(x)\}$. Were it the case that $A \in \mathbf{rng}\, f$ there would be $a \in X$ such that $f(a) = A$. But then

$$a \in A \;\Rightarrow\; a \notin f(a) \;\Rightarrow a \notin A$$

and

$$a \notin A \;\Rightarrow\; a \in f(a) \;\Rightarrow a \in A$$

neither of which is possible. Thus $A \notin \mathbf{rng}\, f$. □

Lemma 4.5. Suppose x is an ordinal and y is an ordinal such that $2^x \approx y$. Then $x \in y$.

Proof. Were it the case that $x = y$ we would have $x \approx 2^x$ which contradicts the preceding Theorem. Were it the case that $y \in x$ we would have $x \approx 2^x$ by the Schroeder-Bernstein Theorem. So, by ??, $x \in y$. □

Corollary 4.2. For each ordinal x there is one and only one ordinal $\mathbf{c}(x)$ such that if w is an ordinal, $w \neq x$ and $w \approx x$ then $\mathbf{c}(x) \in w$.

5. More on relations.

5.1. **Equivalence relations.**

Definition 5.1. We say r is an **equivalence relation on** X if it is reflexive, symmetric and transitive in which case we let

$$X/r = \{r[\{x\}] : x \in X\}.$$

Here is a basic theorem about equivalence relations that is used throughout pure mathematics in building new mathematical objects out of old ones.

Theorem 5.1. Suppose X is a nonempty set.

If r is an equivalence relation on X. Then X/r is a partition of X each of whose members is nonempty.

On the other hand, if $\mathcal{A}$ is a partition of X and no member of $\mathcal{A}$ is empty then

$$r = \cup\{A \times A : A \in \mathcal{A}\}$$

is an equivalence relation on X and $X/r = \mathcal{A}$.

Exercise 5.1. We leave the proof of this Theorem as an exercise for the reader.

5.2. **Integral powers of relations.** For any relation r and any $n \in \mathbb{Z}$ we define the relation

$$r^n$$

as follows. We let $r^0 = \{(x, x) : x \in \mathbf{dmn}\, r\}$. We require that $r^{n+1} = r^n \circ r$ for $n \in \mathbb{N}$; we require that $r^{-n} = (r^n)^{-1}$ for $n \in \mathbb{N}^+$.

Proposition 5.1. Suppose r is a relation. Then $r^{m+n} = r^m \circ r^n$ for $m, n \in \mathbb{Z}$.

Proof. Let

$$\mathcal{N} = \{(m, n) \in \mathbb{Z} \times \mathbb{Z} : r^{m+n} = r^m \circ r^n\}$$

and let $i = \{(x, x) : x \in \mathbf{dmn}\, x\}$.

Suppose $m \in \mathbb{Z}$. Then $r^{m+0} = r^m = r^m \circ i = r^m \circ r^0$ so $(m, 0) \in \mathcal{N}$. Suppose $n \in \mathbb{N}$ and $(m, n) \in \mathcal{N}$. Then

$$r^{m+(n+1)} = r^{(m+n)+1} = r^{m+n} \circ r = (r^m \circ r^n) \circ r = r^m \circ (r^n \circ r) = r^m \circ r^{n+1}.$$

Inducting on n we find that $\mathbb{Z} \times \mathbb{N} \subset \mathcal{N}$.

Suppose $(m, n) \in \mathbb{Z} \times \mathbb{Z}^-$. Then

$$r^{m+n} = (r^{-m-n})^{-1} = (r^{-m} \circ r^{-n})^{-1} \circ r^n \circ r^m = r^m \circ r^n.$$

Thus $\mathbb{Z} \times \mathbb{Z}^- \subset \mathcal{N}$. □

5.3. **Incompleteness of the rational numbers.**

Exercise 5.2. Let $<$ be the standard linear ordering on the set $\mathbb{Q}$ of rational numbers. (We will construct $\mathbb{Q}$ from the natural numbers $\mathbb{N}$ shortly.) Let

$$I = \{x \in \mathbb{Q} : x < 0 \text{ or } x^2 < 2\}.$$

Show that I is an initial segment. (This is easy.) Show that

$$q \in \mathbb{Q} \text{ and } I = \mathbf{i}(q) \ \Rightarrow q^2 = 2.$$

(This gives many students fits.) We will show shortly that $q^2 = 2$ for no $q \in \mathbb{Q}$.