Logic and Set Theory ①

Chapter 1 : Posets and Zorn's Lemma

1.1 Definition

By a partial order on a set A, we mean a binary relation $\leq$ on A which is :

- reflexive $(\forall x \in A)(x \leq x)$
- transitive $(\forall x, y, z \in A)(x \leq y \text{ and } y \leq z \Rightarrow x \leq z)$
- antisymmetry $(\forall x, y \in A)(x \leq y \text{ and } y \leq x \Rightarrow x = y)$

We say that $\leq$ is a total order if it also satisfies

$(\forall x, y \in A)(x \leq y \text{ or } y \leq x)$

By a poset, we mean a partially ordered set. Similarly, a toset is a totally ordered set.

1.2 Examples

a) The usual ordering on $\mathbb{R}$, $\leq$ defined by $x \leq y \Leftrightarrow x - y$ is a square, is a total ordering.

b) For any set A, the set $\mathcal{P}A$ of all subsets of A is partially ordered by $\subseteq$.

c) For any group, G, the set $\mathrm{Sub}(G)$ of all subgroups of G is partially ordered by $\subseteq$, and similarly for a vector space V.

d) The set $\mathbb{N}_0$ of natural numbers (including 0) is partially ordered by $|$ where $m | n \Leftrightarrow \frac{n}{m} \in \mathbb{N}_0$

(Aside: If $\leq$ is reflexive and transitive on a set A, then the relation $\sim$ defined by $x \sim y \Leftrightarrow (x \leq y$ and $y \leq x)$ is an equivalence relation on A, and $\leq$ induces a partial ordering on the quotient set $A/\sim$)

e) Suppose that given a set Σ of 'letters', a <u>word</u> in Σ is a finite (possibly empty) sequence of members of Σ. The set Σ^* of all words over Σ has several partial orderings:

- the <u>prefix ordering</u> by $r \leq w \Leftrightarrow w = ru$ for some $u \in \Sigma^*$
- the <u>ifix ordering</u> by $r \leq w \Leftrightarrow w = uxv$ for some $u, v \in \Sigma^*$

f) Given sets A, B, by a <u>partial function</u> $f: A \to B$ we mean a function defined on a subset of A and taking values in B.
We write $[A \to B]$ for the set of all partial functions A to B.
This set is partially ordered by $f \leq g \Leftrightarrow g$ extends f, i.e $g(x)$ is defined and equals $f(x)$ wherever $f(x)$ is defined.

<u>1.3 Definition</u>

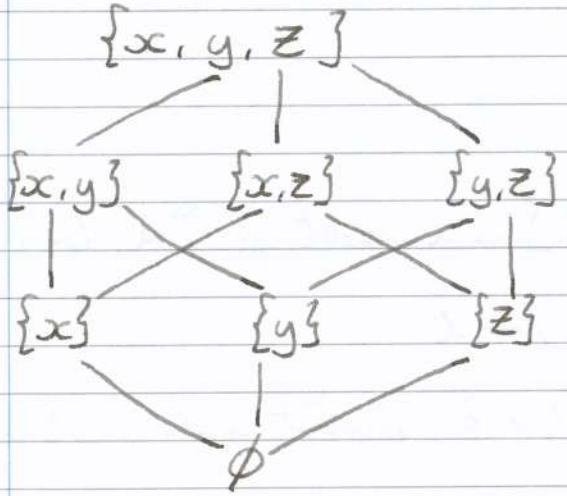
Let $(A, \leq)$ be a poset. We say that b <u>covers</u> a in A, and write $a \lhd b$ if $a \leq b$, $a \neq b$ and $(\forall c)((a \leq c$ and $c \leq b) \Rightarrow (a = c$ or $c = b$

If A is finite, then $a \leq b$ holds $\Leftrightarrow$ either $a = b$ or there exists a finite chain $a \lhd c_1 \lhd c_2 \lhd \ldots \lhd c_n \lhd b$

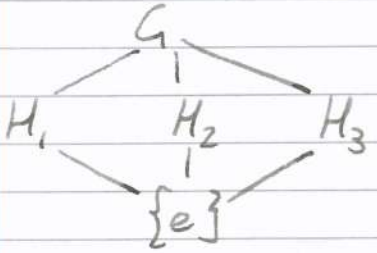
The <u>Hasse Diagram</u> of a finite poset represents the elements of P by points, with an upward line from x to y wherever $x \lhd y$.

1.4 Examples

a) $\mathcal{P}\{x, y, z\}$ has 8 elements, and its Hasse diagram is



b) Let G be the non-cyclic group of order 4. The Hasse diagram of $\mathrm{Sub}(G)$ is



1.5 Definition

Let $(P, \leq)$ be a poset, $S \subseteq P$.

a) By a <u>greatest element</u> of S, we mean an element $s \in S$ such that $t \leq s \quad \forall t \in S$

b) By an <u>upper bound</u> for S, we mean an element $p \in P$ such that $s \leq p \quad \forall s \in S$

c) By a <u>least upper bound</u> for S, we mean a least element of $\{p \in P \mid p \text{ is an upper bound for } S\}$

p is the greatest element of $S \Leftrightarrow$ (p is a least upper bound for S and $p \in S$)

We say that $(P, \leq)$ is <u>complete</u> if every $S \subseteq P$ has a least upper bound.

<u>1.6 Examples</u>

a) $\mathcal{P}A$ is complete. If $\{B_i \mid i \in I\}$ is a subset of $\mathcal{P}A$, then $\bigcup_{i \in I} B_i$ is a least upper bound for it.

b) $(\mathbb{R}, \leq)$ is not complete, since it lacks greatest and least elements. But $\mathbb{R} \cup \{\pm\infty\}$ is complete.

c) $\mathrm{Sub}(G)$ is complete. Given a set $\{H_i \mid i \in I\}$ of subgroups of G, $\bigcup_{i \in I} H_i$ needn't be a subgroup, but there is a smallest subgroup $\langle \bigcup_{i \in I} H_i \rangle$ containing it.

<u>1.7 Lemma</u>

Suppose that $(P, \leq)$ is a complete poset. Then, so is $(P, \geqslant)$.

<u>Proof</u>

Given $S \subseteq P$, we have to show that S has a greatest lower bound. Let $T = \{p \in P \mid p \text{ is a lower bound for } S\}$, and let $t = \bigvee T$, the least upper bound of T. For any $s \in S$, we have $p \leq s$ $\forall p \in T$, so s is an upper bound for T. Hence $t \leq s$. So t is a lower bound for S, i.e. $t \in T$, and so t is the greatest element of T.

08/10/12

Logic and Set Theory ②

We denote the least upper bound of a set S, if it exists, by $\bigvee S$, and call it the join of S, and the greatest lower bound $\bigwedge S$, the meet of S.

1.8 Definition

a) By a chain in a poset $(P, \leq)$, we mean a non-empty subset $C \subseteq P$ which is totally ordered by the restriction of $\leq$.

b) We say $(P, \leq)$ is chain complete if every chain $C \subseteq P$ has a join in P.

1.9 Examples

a) Let G be a group, and let P be the set of abelian subgroups of G ordered by inclusion. If $H, K \in P$ contain elements h, k such that $hk \neq kh$, then $\{H, K\}$ has no upper bound in P, so P is not complete.

But P is chain complete: If $\{H_i \mid i \in I\}$ is a chain in P, then $\bigcup_{i \in I} H_i$ is a subgroup, since if $h, k \in \bigcup_{i \in I} H_i$, then $h \in H_i$ and $k \in H_j$ for some i, j but then since either $H_i \subseteq H_j$ or $H_j \subseteq H_i$ we have either $\{h, k\} \subseteq H_j$ or $\subseteq H_i$, and hence $hk \in \bigcup_{i \in I} H_i$. Similarly, if $hk \in \bigcup_{i \in I} H_i$, then $hk = kh$. So $\bigcup_{i \in I} H_i$ is an abelian subgroup, and hence a least upper bound for $\{H_i \mid i \in I\}$ in P.

b) Let A, B be two sets, and consider the poset $[A \to B]$ of partial functions $A \to B$, ordered by extension. This is

not complete, since if f, g are both defined at $x \in A$ and $f(x) \neq g(x)$, then $\{f, g\}$ has no upper bound. But if $\{f_i \mid i \in I\}$ is a chain in $[A \to B]$ then we have $f_i(x) = f_j(x)$ whenever f_i, f_j are both defined at x, so there is a unique f with domain $\bigcup_{i \in I} \operatorname{dom} f_i$ satisfying $f(x) = f_i(x)$ $\forall x \in \operatorname{dom} f_i$, $\forall i$. So $f = \bigvee \{f_i \mid i \in I\}$.

Recursive Definitions

We define the factorial function as the function $f : \mathbb{N} \to \mathbb{N}$ satisfying $f(n) = 1$ if $n = 0$, $f(n) = n f(n-1)$ if $n > 0$.

How do we know that this defines anything? Think instead of a function $\Phi : [\mathbb{N} \to \mathbb{N}] \to [\mathbb{N} \to \mathbb{N}]$ defined by

$$\Phi(f)(n) = 1 \text{ if } n = 0$$
$$= n f(n-1) \text{ if } n > 0 \text{ and } f(n-1) \text{ is defined}$$
$$\text{undefined otherwise.}$$

Then we define the factorial function to be the unique f such that $\Phi(f) = f$, assuming that this exists.

Note that Φ is order preserving: i.e. $f \leq g \Rightarrow \Phi(f) \leq \Phi(g)$

1.10 Theorem (Knaster-Tarski)

Let $(P, \leq)$ be a complete poset and $f : P \to P$ an order preserving map. Then f has a fixed point.

Proof:

Consider the set $S = \{x \in P \mid x \leq f(x)\}$ of pre-fixed points.

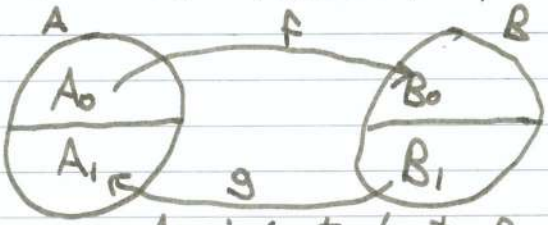
Let $y = \bigvee S$.

Thus, $\forall x \in S$, we have $x \leq y$. Hence $x \leq f(x) \leq f(y)$. So $f(y)$ is an upper bound for S. Hence $y \leq f(y)$, i.e $y \in S$. Since f is order-preserving, we also have $f(y) \leq f(f(y))$, so $f(y) \in S$. Hence $f(y) \leq y$, and $f(y) = y$ by antisymmetry.

1.11 Corollary (Cantor-Bernstein Theorem)
Suppose given sets A, B, and we have injections $f: A \to B$, $g: B \to A$. Then there exists a bijection $A \to B$.

Proof:
Consider the composite $PA \xrightarrow{f[-]} PB \xrightarrow{B\setminus(-)} PB \xrightarrow{g[-]} PA \xrightarrow{A\setminus(-)} PA$
This is order preserving since $f[-]$ and $g[-]$ are order preserving and $B\setminus(-)$, $A\setminus(-)$ are order reversing. So it has a fixed point A_0. Let $B_0 = f[A_0]$, $B_1 = B \setminus B_0$, $A_1 = g[B_1]$, then $A_1 = A \setminus$

I.e.



where f maps A_0 bijectively to B_0, and g maps B_1 bijectively to A_1. So we define $h: A \to B$ by $h(x) = f(x)$ if $x \in A_0$ and $h(x) = g^{-1}(x)$ if $x \in A_1$.

Given a poset P, we say $f: P \to P$ is inflationary if $x \leq f(x)$ for all $x \in P$.

1.12 Theorem (Bourbaki - Witt) Let P be a chain complete poset and $f: P \to P$ an inflationary map. Then for any $x \in P$, there exists $y \in P$ with $x \leq y = f(y)$.

N. Bourbaki, Sur la théorème de Zorn, Arch Math 2(1950), 434-43
E. Witt, Beweisstudien zum Satz von. M. Zorn, Math Nachrichten 4 (1951) 434-43
E. Witt, Sobre el teorema de Zorn, Rev. Mat. Hispanoamericana (4)10(1950) 82-85

Non-proof

Set $x_0 = x$. If it isn't a fixed point, set $x_1 = f(x_0)$.
If x_1 is not, set $x_2 = f(x_1)$.
If $x_0 < x_1 < x_2 < \ldots$, set $x_\omega = \bigvee \{x_n \mid n \in \mathbb{N}\}$
If x_ω isn't a fixed point, $x_{\omega+1} = f(x_\omega)$
$x_{\omega+2} = f(x_{\omega+1})$

For heaven's sake, this must stop!

10/10/12

Logic and Set Theory (3)

1.12 Theorem (Bourbaki-Witt)

Let P be a chain-complete poset and $f: P \to P$ an inflationary map. Then, for every $x \in P$, there exists $y \in P$ with $x \leq y = f(y)$

Proof

Define a subset $C \subseteq P$ to be closed if

a) For all $y \in C$, $f(y) \in C$ and

b) For all chains $S \subseteq C$, $\bigvee S \in C$

Note that arbitrary intersections of closed sets are closed.

Hence $C(x) = \bigwedge \{ D \subseteq P \mid D \text{ closed}, x \in D \}$ is the smallest closed set containing x.

(Idea: $C(x)$ is the set of all x_α's that we constructed in the non-proof)

Suppose we can show that $C(x)$ is a chain. Then $y = \bigvee C(x)$ is a member of $C(x)$ since $C(x)$ is closed, and hence we also have $f(y) \in C(x)$. So $f(y) \leq y$; but we have $y \leq f(y)$ since f is inflationary. So $y = f(y)$

(and $y \geq x$ since $x \in C(x)$).

To show that $C(x)$ is totally ordered:

Step 1

x is the least element of $C(x)$, since $\uparrow(x) = \{ y \in P \mid x \leq y \}$ is closed and $x \in \uparrow(x)$, and $C(x) \subseteq \uparrow(x)$

Now call an element y of $C(x)$ normal if
$(\forall z \in C(x))(z < y \Rightarrow f(z) \leq y)$.

Step 2

If y is normal, then for all $z \in C(x)$, we have either $z \leq y$ or $f(y) \geq z$.

To prove this, consider the set $T_y = \{ z \in C(x) \mid z \leq y \text{ or } f(y) \geq z$
$x \in T_y$ since $x \leq y$ by Step 1. Suppose $z \in T_y$: then either $z < y$, in which case $f(z) \leq y$ by normality. Or, $z = y$, in which case $f(z) \geq f(y)$. Or, $z \geq f(y)$, in which case $f(y) \leq f(z)$ since f is inflationary. So in every case $f(z) \in T_y$. Now suppose $S \subseteq T_y$ is a chain: if all $z \in S$ satisfy $z \leq y$, then $\bigvee S \leq y$. If some $z \in S$ satisfies $z \geq f(y)$, then $\bigvee S \geq f(y)$, so $\bigvee S \in T_y$.

So T_y is a closed set containing x, and hence $T_y = C(x)$.

Step 3

Every $y \in C(x)$ is normal.

Consider the set $N = \{ y \in C(x) \mid y \text{ is normal} \}$. $x \in N$ since the hypothesis $z < x$ is never satisfied, by Step 1. ($C(x) \subseteq \uparrow x$, x least element of $C(x)$.)

Suppose $y \in N$, and $z < f(y)$. Then by Step 2, we have $z \leq y$. So either $z < y$, in which case $f(z) \leq y \leq f(y)$, by normality of y, or $z = y$, so $f(z) = f(y) \leq f(y)$.

Let $S \subseteq N$ be a chain, and suppose $z < \bigvee S$. ($\in N$, want)

Logic and Set Theory (3)

Then $z \not\geq y$ for some $y \in S$, so $z \not\geq f(y)$, so $z \leq y$ (inflationary) by Step 2, so $z < y$. So $f(z) \leq y \leq VS$ by normality of y (as $y \in N$). So $VS \in N$. $\Phi \Rightarrow N$ closed, $N \supseteq C(x)$

Hence by Steps 2 and 3, $C(x)$ is a chain, so $VC(x)$ is a fixed point of f above x.

1.13 Corollary

Let P be a chain-complete poset, and $f: P \to P$ an order preserving map. Then, for any $x \in P$, with $x \leq f(x)$, there's a least $y \geq x$ with $y = f(y)$. In particular, if P has a least element 0, then f has a least fixed point.

Proof

(shows $f: P' \to P'$)

Let $P' = \{x \in P \mid x \leq f(x)\}$. Then $x \in P' \Rightarrow f(x) \in P'$ since f is order preserving. And if $S \subseteq P'$ has a join VS in P, then for all $x \in S$ we have $x \leq VS$, and so $x \leq f(x) \leq f(VS)$, so $f(VS)$ is an upper bound for S and hence $VS \leq f(VS)$. (P already chain complete; this shows that $VS \in P'$, so P' chain complete)

So by 1.12 every $x \in P'$ lies below a fixed point $y = VC(x)$

If z is any fixed point with $z \geq x$, then $\downarrow(z) = \{w \in P' \mid w \leq z\}$ is a closed set containing x, so $C(x) \subseteq \downarrow(z)$, so $VC(x) \leq z$.

Recall the recursive definition of $n!$:

$\Phi(f)(n) = 1$ if $n = 0$, $\Phi(f)(n) = n f(n-1)$ if $n > 0$, $f(n-1)$ defined; undefined otherwise.

$\mathbb{E}$ is an order preserving map $[\mathbb{N} \to \mathbb{N}] \to [\mathbb{N} \to \mathbb{N}]$
$[\mathbb{N} \to \mathbb{N}]$ is chain complete and has a least element, so there is a least f with $\mathbb{E}(f) = f$. But, for this f, we have $0 \in \mathrm{dom}\, f$ and $(\forall n)(n \in \mathrm{dom}\, f \Rightarrow n+1 \in \mathrm{dom}\, f)$. So f is a total function and hence a maximal element of $[\mathbb{N} \to \mathbb{N}]$, i.e. $f \leq g \Rightarrow f = g$.
Hence f is the unique fixed point of $\mathbb{E}$.

For the next result, we need the Axiom of Choice which asserts that if $\{A_i \mid i \in I\}$ is any set of non-empty sets, then there exists a choice function $f : I \to \bigcup_{i \in I} A_i$ i.e. a function such that $f(i) \in A_i$, for all $i \in I$.

1.14 Corollary (Zorn's Lemma)
Assume the Axiom of Choice. Let P be a chain complete poset. Then, for any $x \in P$, there exists a maximal element y of P with $x \leq y$.

Proof
For $x \in P$, let $A_x = \{y \in P \mid y > x\}$ if x is not maximal
$= \{x\}$ if x is maximal
Then A_x is non-empty for all x, so we have a choice function $f : P \to P$ with $f(x) \in A_x$ $\forall x$. Thus f is necessarily inflationary, and its fixed points are exactly the maximal elements of P.
So the result follows from 1.12.

1.15 Examples (2)

a) We show that Zorn's Lemma implies the Axiom of Choice.
Given a family $\{A_i \mid i \in I\}$ of non-empty sets, let P be the set of partial choice functions, i.e. partial functions $f : I \rightharpoonup \bigcup_{i \in I} A_i$ satisfying $f(i) \in A_i$ wherever $f(i)$ is defined.
Order P by $f \leq g \iff g$ extends f. Then P is chain-complete being closed under joins of chains in $[I \rightharpoonup \bigcup_{i \in I} A_i]$.
$P \neq \emptyset$ since the everywhere undefined function is in P, so it has a maximal element, f say. Suppose f is not total. Pick $i_0 \in I \setminus \mathrm{dom}\, f$, and pick $x_0 \in A_{i_0}$. Now define
$g(i) = f(i)$ if $i \in \mathrm{dom}\, f$, x_0 if $i = i_0$, undefined otherwise.
Then $g \in P$ and $f < g$ ⚡

b) (Hamel's Theorem) Every vector space has a basis. Let V be a vector space (over some field F). Consider the poset P of linearly independent subsets $S \subseteq V$, ordered by inclusion. P is chain complete:
If $\{S_i \mid i \in I\}$ is a chain of linearly-independent sets, then $\bigcup_{i \in I} S_i$ is linearly independent, since if we had a linear relation $\sum_{j=1}^{n} \lambda_j x_j = 0$ with $x_j \in \bigcup_{i \in I} S_i$ for all j, and then $\exists\, i \in I$ such that $x_j \in S_i$ for all j, and hence we must have $\lambda_j = 0$ for all j.
Now, let S be a maximal element of P. Suppose not.

Suppose that S is not a basis. Then we can pick $v \in V \setminus \langle S \rangle$ and consider $T = S \cup \{v\}$. Then, T is linearly independent and $S < T$ ⚔

c) (Maximal Ideal Theorem) Let R be a ring (with 1). Then, any proper ideal of R is contained in a maximal (proper) ideal. Let P be the set of proper ideals of R, ordered by inclusion. We must show that P is chain complete. Let $\{I_j \mid j \in J\}$ be a chain of proper ideals. Then $\bigcup_{j \in J} I_j$ is an ideal, since it is an additive subgroup (cf 1.9a)). But an ideal I is proper $\Leftrightarrow 1 \notin I$. So $1 \notin \bigcup_{j \in J} I_j$ and hence $\bigcup_{j \in I} I_j$ is a proper ideal.

1.16 Definition

a) A lattice is a poset L which has joins and meets for all finite subsets. In particular, L has a least element $0 = \bigvee \emptyset$ and a greatest element $1 = \bigwedge \emptyset$. Also, for any $\{x, y\} \subseteq L$, we have $x \vee y = \bigvee \{x, y\}$ and $x \wedge y = \bigwedge \{x, y\}$. These suffice, since we can construct $\bigvee \{x_1, \ldots, x_n\}$ as $((\ldots (x_1 \vee x_2) \vee x_3) \vee \ldots) \vee x_n$ (Exercise: check that $\vee$ and $\wedge$ are commutative and associative)

A mapping $f : L \to M$ between lattices is called a lattice homomorphism if $f(0_L) = f(0_M)$, $f(1_L) = f(1_M)$, and $f(x \vee y) = f(x) \vee f(y)$, $f(x \wedge y) = f(x) \wedge f(y)$.

A lattice homomorphism is order preserving since $x \leq y \Leftrightarrow x \vee y = y \Leftrightarrow x \wedge y = x$.

The converse is false: If G is a group, the inclusion $\mathrm{Sub}(G) \to \mathcal{P}G$ does not preserve $\vee$ or $\emptyset$.

b) We say that a lattice L is <u>distributive</u> if the identity $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$ holds for all $x, y, z \in L$

c) We say that y is a <u>complement</u> for $x \in L$ if $x \wedge y = 0$ and $x \vee y = 1$. By a <u>Boolean Algebra</u>, we mean a distributive lattice in which every element has a complement.

<u>1.17 Examples</u>

a) For any A, $\mathcal{P}A$ is a Boolean Algebra. It is distributive, since if $B, C, D \subseteq A$, $x \in B \cap (C \cup D) \Leftrightarrow x \in B$, and $(x \in C$ or $x \in D)$

$\Leftrightarrow (x \in B$ and $x \in C)$ or $(x \in B$ and $x \in D)$

$\Leftrightarrow x \in (B \cap C) \cup (B \cap D)$

Also, for any $B \subseteq A$, $A \setminus B$ is a complement for B in $\mathcal{P}A$.

b) Let L be a totally ordered set with greatest and least elements. Then L is a lattice, with

$x \vee y = \max\{x, y\}$ and $x \wedge y = \min\{x, y\}$.

L is distributive: e.g. if $x \leq y \leq z$, then $x \wedge (y \vee z) = x \wedge z = x$

and $(x \wedge y) \vee (x \wedge z) = x \vee x = x$

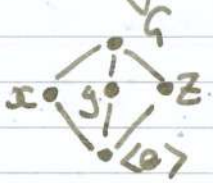
Or, if $y \leq z \leq x$, then $x \wedge (y \vee z) = x \wedge z = z$

and $(x \wedge y) \vee (x \wedge z) = y \vee z = z$

(Exercise: Check the other cases)

But if L has >2 elements, it isn't Boolean. ← Check, only top + bottom elements have complements

c) If G is the non-cyclic group of order 4, then $\mathrm{Sub}(G)$ isn't distributive.



$x \wedge (y \vee z) = x \wedge G = x$

but $(x \wedge y) \vee (x \wedge z) = \{e\} \vee \{e\} = \{e$

Note that here, x has two distinct complements, y and z.

1.18 Lemma ← reversed order relation, "opposite"

i) If L is distributive, so is L^{op}

ii) In a distributive lattice, every element has at most one complement.

Proof •

i) We have to show $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$. But

$(x \vee y) \wedge (\underbrace{x \vee z}_{\text{expand}}) = (\underbrace{(x \vee y) \wedge x}_{= x}) \vee (\underbrace{(x \vee y) \wedge z}_{\text{expand}})$

$= \underbrace{x \vee (x \wedge z)}_{= x} \vee (y \wedge z) = x \vee (y \wedge z)$

ii) Suppose y, z are both complements for x. Then:

$y \wedge (x \vee z) = y \wedge 1 = y$ but

$(y \wedge x) \vee (y \wedge z) = 0 \vee (y \wedge z) = y \wedge z$

and similarly $z = y \wedge z$, so $y = z$.

15/10/12

Logic and Set Theory (5)

We write 2 for the two element lattice $\{0, 1\}$.

1.19 Lemma

Let a, b be elements of a distributive lattice L with $a \not\leq b$. Then there exists a lattice homomorphism $f : L \to 2$ with $f(a) = 1$ and $f(b) = 0$.

Proof

Let P be the set of ordered pairs (A, B) of subsets of L satisfying:

i) $x \in A$, $x \leq y \Rightarrow y \in A$, and A is closed under finite meets

ii) $x \leq y$, $y \in B \Rightarrow x \in B$, and B is closed under finite joins.

iii) $A \cap B = \emptyset$.

We partially order P by $(A_1, B_1) \leq (A_2, B_2) \Leftrightarrow A_1 \subseteq A_2, B_1 \subseteq B_2$

P is chain complete: given a chain $\{(A_i, B_i) \mid i \in I\}$

$\bigcup_{i \in I} A_i$ satisfies the closure properties (i), and similarly

$\bigcup_{i \in I} B_i$ satisfies (ii). Because $\ldots \subseteq A_i \subseteq A_{i+1} \subseteq \ldots$

Suppose $x \in (\bigcup_{i \in I} A_i) \cap (\bigcup_{i \in I} B_i)$, then $x \in A_i$ for some i, and $x \in B_j$ for some j. But either $A_i \subseteq A_j$ or $B_j \subseteq B_i$, so either $x \in A_j \cap B_j$ or $x \in A_i \cap B_i$ ✗ since $(A_i, B_i) \leq$ or $\geq (A_j, B_j)$

So $(\bigcup_{i \in I} A_i, \bigcup_{i \in I} B_i) \in P$, and is a least upper bound for the chain.

Now $(\uparrow(a), \downarrow(b)) \in P$, since $a \not\leq b$, so we can find a maximal element (A_0, B_0) of P with $a \in A_0$, $b \in B_0$.

Suppose $A_0 \cup B_0 \neq L$; let $c \in L \setminus (A_0 \cup B_0)$

Let $A_1 = \{y \in L \mid y \geq x \wedge c \text{ for some } x \in A_0\}$ ($x \geqslant x \wedge c$)

then A_1 satisfies (i) and $A_0 \subsetneq A_1$, since $c = 1 \wedge c \in A_1$.

Hence $(A_1, B_0) \notin P$ by maximality of (A_0, B_0), so

$A_1 \wedge B_0 \neq \emptyset$, i.e. $\exists\, x \in A_0$ such that $x \wedge c \in B_0$

Similarly, considering $B_1 = \{x \in L \mid x \leq y \vee c \text{ for some } y \in B_0\}$

we deduce that there exists $y \in B_0$ such that $y \vee c \in A_0$.

Now $x \wedge (y \vee c) \in A_0$ since it is a meet of two elements of A_0, but it equals $(x \wedge y) \vee (x \wedge c)$ which is a join of two elements of B_0 ##

So $A_0 \cup B_0 = L$, and the total function $f : L \to 2$ defined by

$f(x) = \begin{cases} 1 & x \in A_0 \\ 0 & x \in B_0 \end{cases}$ is a lattice homomorphism with $f(a) = 1$, $f(b) = 0$.

1.20 Theorem (Birkhoff - Stone)

Any distributive lattice is isomorphic to a sub-lattice of some power set $\mathcal{P}A$, i.e. to a family of subsets of A closed under finite unions and intersections. In particular, any Boolean algebra is isomorphic to a sub-Boolean-algebra of $\mathcal{P}A$, i.e. a subset of $\mathcal{P}A$ closed under finite unions.

Proof

Given a distributive lattice L, let A be the set of all lattice homomorphisms $L \to 2$.

Define $\Phi : L \to \mathcal{P}A$ by $\Phi(x) = \{f \in A \mid f(x) = 1\}$

Φ is a lattice homomorphism: we have $f \in \Phi(x \wedge y) \Leftrightarrow f(x \wedge y)$

$\Leftrightarrow f(x) = f(y) = 1 \Leftrightarrow f \in \Phi(x) \cap \Phi(y)$

and similarly $f \in \Phi(x \vee y) \Leftrightarrow f(x \vee y) = 1$

$\Leftrightarrow$ either $f(x) = 1$ or $f(y) = 1 \Leftrightarrow f \in \Phi(x) \cup \Phi(y)$.

Also, Φ is injective since if $a \neq b$ then 1·19 yields $f \in A$ $\Rightarrow a \neq b$ or $b \neq a$

belonging to just one of $\Phi(a)$ and $\Phi(b)$.

So Φ is an isomorphism from L to the sub-lattice

$\{\Phi(x) \mid x \in L\}$ of $\mathcal{P}A$.

In particular, if L is Boolean, then this sub-lattice is a sub-Boolean-algebra of $\mathcal{P}A$.

Chapter II: The Propositional Calculus

2·1 Definition

A primitive proposition is an abstract symbol p which is capable of being assigned a truth value 0 (false) or 1 (true).

Given a set P of primitive propositions, a valuation of P is a function $v : P \to 2$. Given such a set P, the compound propositions (or propositional formulae) over P are the members of the set $\mathcal{L}(P)$ defined recursively by:

i) If $p \in P$ then $p \in \mathcal{L}(P)$

ii) If $s, t \in \mathcal{L}(P)$, then $(s \Rightarrow t) \in \mathcal{L}(P)$

iii) $\bot \in \mathcal{L}(P)$ ('always false')

Given a valuation $v : P \to 2$, we extend it to a function $\bar{v} : \mathcal{L}(P) \to 2$ by

i) $\bar{v}(p) = v(p)$ if $p \in P$

ii) $\bar{v}(s \Rightarrow t) = 0$ if and only if $\bar{v}(s) = 1$ and $\bar{v}(t) = 0$

iii) $\bar{v}(\bot) = 0$

The clause for $\bar{v}(s \Rightarrow t)$ can be presented as a truth-table

s	t	$s \Rightarrow t$
0	0	1
0	1	1
1	0	0
1	1	1

We can use the same idea to evaluate the truth or falsity of more complicated propositions.

We define $\neg p$ to be $(p \Rightarrow \bot)$ since this has truth table

p	$\neg p$
0	1
1	0

We now define $\top = \neg \bot$, and also $(p \vee q)$ to be $(\neg p \Rightarrow q)$, $(p \wedge q) = \neg(p \Rightarrow \neg q)$

p	q	$\neg p$	$p \vee q$
0	0	1	0
0	1	1	1
1	0	0	1
1	1	0	1

p	q	$\neg q$	$p \wedge q$
0	0	1	0
0	1	0	0
1	0	1	0
1	1	0	1

$(p \Leftrightarrow q)$ is defined to be $((p \Rightarrow q) \wedge (q \Rightarrow p))$

17/10/12

Logic and Set Theory (6)

2·2 Lemma (Functional Completeness of Propositional Calculus)
For any $n \geq 0$, each function $2^n \to 2$ occurs as the truth table of some $S \in \mathcal{L}(\{p_1, p_2, \dots, p_n\})$

Proof (By induction on n)
For $n = 0$, there are two functions $2^0 = \{*\} \to 2$, corresponding to the two elements 0 and 1, and these are the truth tables of $\bot$ and $\top = (\bot \Rightarrow \bot)$ respectively.
Assume true for n, and let $f : 2^{n+1} \to 2$. Define
$f_0, f_1 : 2^n \to 2$ by $f_i(x_1, x_2, \dots, x_n) = f(x_1, x_2, \dots, x_n$
We can find $S_0, S_1 \in \mathcal{L}(\{p_1, \dots, p_n\})$ whose truth tables are f_0, f_1. Now consider $S = ((S_0 \wedge \neg p_{n+1}) \vee (S_1 \wedge p_{n+1}))$
($x_{n+1} = 0$ for S_0, $x_{n+1} = 1$ for S_1)
It is immediately true that this has f as its truth table. □
When do two formulae have the same truth table?
Note that this happens for s, t if and only if $(s \Leftrightarrow t)$ has truth table the constant function with value 1.

2·3 Definition
Let P be a set of primitive propositions, let $S \subseteq \mathcal{L}(P)$ and let $t \in \mathcal{L}(P)$. We say that S <u>semantically entails</u> t and write $S \models t$, if any valuation $v : P \to 2$ such that $\bar{v}(s) = 1 \ \forall s \in S$ also satisfies $\bar{v}(t) = 1$.
In the particular case with $S = \emptyset$, we simply write $\models t$ and call it a <u>tautology</u>

2.4 Examples

a) $\vdash (p \Rightarrow (q \Rightarrow p))$ since this formula has truth table

p	q	$(q \Rightarrow p)$	$p \Rightarrow (q \Rightarrow p)$
0	0	1	1
0	1	0	1
1	0	1	1
1	1	1	1

b) $\vdash ((p \Rightarrow (q \Rightarrow r)) \Rightarrow ((p \Rightarrow q) \Rightarrow (p \Rightarrow r))) = S$

S	p	q	r	$(p \Rightarrow q)$	$(p \Rightarrow r)$	$(q \Rightarrow r)$	$(p \Rightarrow (q \Rightarrow r))$	$((p \Rightarrow q) \Rightarrow (p \Rightarrow r))$
1	1	1	0	1	0	0	0	0

c) $\{p, (p \Rightarrow q)\} \vDash q$

	p	q	$(p \Rightarrow q)$	
$\not\vDash$	0	0	1	p is false
	0	1	1	
	1	0	0	$p \Rightarrow q$ is false
	1	1	1	

$(MP)_1$ **2.5 Definition**

$[(MP)_1 \Rightarrow (MP)_2]$ The axioms of the propositional calculus are all formulas

$(MP)_2$ of the three forms

(K) $(S \Rightarrow (t \Rightarrow S))$, (S) $((S \Rightarrow (t \Rightarrow u)) \Rightarrow ((S \Rightarrow t) \Rightarrow (S \Rightarrow u)))$

or (T) $(((S \Rightarrow \bot) \Rightarrow \bot) \Rightarrow S)$

We have one rule of inference (Modus Ponens): from S, and $(S \Rightarrow t)$, we may infer (deduce) t.

A deduction from a set S of hypotheses is a finite list $(S_1, S_2, \ldots, S_n)$ of propositional formulae, such that for each i, we have one of:

1. $S_i \in S$
2. S_i is an axiom
3. There exist $j, k < i$ such that $S_k = (S_j \Rightarrow S_i)$

17/10/12

Logic and Set Theory ⑥

We say that S <u>syntactically entails</u> t, and write $S \vdash t$ if there exists a deduction from S whose last line is t.

We call a deduction from $\emptyset$ a <u>proof</u>, and write $\vdash t$, and say that t is a <u>theorem</u> if it is deducible from $\emptyset$.

2.6 Lemma (Examples)

a) The following is a proof of $(s \Rightarrow s)$:

1. $(s \Rightarrow (s \Rightarrow s))$ (k)
2. $(s \Rightarrow ((s \Rightarrow s) \Rightarrow s))$ (k) with s, $(s \Rightarrow s)$
3. $((s \Rightarrow ((s \Rightarrow s) \Rightarrow s)) \Rightarrow ((s \Rightarrow (s \Rightarrow s)) \Rightarrow (s \Rightarrow s)))$ (axiom S) with $s, (s \Rightarrow s), s$
4. $((s \Rightarrow (s \Rightarrow s)) \Rightarrow (s \Rightarrow s))$ (MP from lines 2 and 3)
5. $(s \Rightarrow s)$ (MP from lines 1 and 4)

b) The following is a deduction of $(s \Rightarrow u)$ from $\{(s \Rightarrow t), (t \Rightarrow u)\}$

1. $((s \Rightarrow (t \Rightarrow u)) \Rightarrow ((s \Rightarrow t) \Rightarrow (s \Rightarrow u)))$ (axiom S)
2. $(t \Rightarrow u)$ (hypothesis)
3. $((t \Rightarrow u) \Rightarrow (s \Rightarrow (t \Rightarrow u)))$ (axiom k)
4. $(s \Rightarrow (t \Rightarrow u))$ (MP from 2, 3)
5. $((s \Rightarrow t) \Rightarrow (s \Rightarrow u))$ (MP from 1, 4)
6. $(s \Rightarrow t)$ (hypothesis)
7. $(s \Rightarrow u)$ (MP from 5, 6)

2.7 Lemma (Soundness Theorem)

If $S \vdash t$, then $S \models t$. In particular, every Theorem is a tautology.

Proof

Let $(s_1, s_2, \ldots, s_n = t)$ be a deduction of t from S.

We show $S \models s_i$ for all i, by induction on i.

Clearly, $s_i \in S \Rightarrow S \models s_i$.

If s_i is an axiom, then $\emptyset \models s_i$ and so $S \models s_i$

If we have $j, k < i$ such that $s_k = (s_j \Rightarrow s_i)$, then by induction any model $\bar{v}$ for S (a valuation making all of S true) satisfies $\bar{v}(s_j) = 1$ and $\bar{v}(s_k) = 1$, so by 2.4 c) we have $\bar{v}(s_i) = 1$.

2.8 Theorem (Deduction Theorem)

Let $S \subseteq \mathcal{L}(P)$, $s, t \in \mathcal{L}(P)$. Then

$S \vdash (s \Rightarrow t)$ if and only if $S \cup \{s\} \vdash t$

Proof

($\Rightarrow$) Given a deduction $s_1, \ldots, s_n = (s \Rightarrow t)$ from S, add the two lines $s_{n+1} = s$ (hypothesis), then $s_{n+2} = t$ (MP).

($\Leftarrow$) Suppose given a deduction $(t_1, t_2, \ldots, t_n = t)$ of t from $S \cup \{s\}$ we show by induction on i that $S \vdash (s \Rightarrow t_i)$ $\forall i$.

If t_i is an axiom, or a member of S, we write down (MP)
t_i (axiom or hypothesis), $t_i \Rightarrow (s \Rightarrow t_i)$ (axiom k), $s \Rightarrow t_i$

Logic and Set Theory ⑥

If t_i is S, we write down the 5 line proof of $(S \Rightarrow S)$ in 2.6 a).

If $\exists j, k < i$ with $t_k = (t_j \Rightarrow t_i)$, we write down deductions of $(S \Rightarrow t_j)$ and of $(S \Rightarrow (t_j \Rightarrow t_i))$, then write down $((S \Rightarrow (t_j \Rightarrow t_i)) \Rightarrow ((S \Rightarrow t_j) \Rightarrow (S \Rightarrow t_i)))$ (axiom S) and apply MP twice.

$S \vdash t$: syntactic entailment, a deduction exists

$S \models t$: semantic entailment, if S true means t is true

Model (for S): a valuation $\bar{v}: \mathcal{L}(P) \to \mathbb{Z}$ on which $v(s) = 1$ for $s \in S$.

2.9 Theorem (Completeness Theorem)

Given $S \subseteq \mathcal{L}(P)$, $t \in \mathcal{L}(P)$, we have $S \vdash t \Leftrightarrow S \models t$

Proof:

($\Rightarrow$) is lemma 27.

($\Leftarrow$) Using 2.8, we may reduce this problem to the particular case $t = \bot$:

If $S \models t$, then $S \cup \{\neg t\} \models \bot$ and if $S \cup \{\neg t\} \vdash \bot$, then by 2.8 we have $S \vdash \neg\neg t$ and we can convert a deduction of $\neg\neg t$ from S into a deduction of t by adding the axiom $(\neg\neg t \Rightarrow t)$, and applying modus ponens.

In the case $t = \bot$, we prove the contrapositive implication

If S is consistent (i.e. $S \not\vdash \bot$) then S has a model (i.e. $S \not\models \bot$). Consider the set Q of all consistent subsets of $\mathcal{L}(P)$, ordered by inclusion. Q is chain-complete: if $\{T_i \mid i \in I\} \subseteq Q$ is a chain, then $\bigcup_{i \in I} T_i$ is consistent since a deduction of $\bot$ from it would involve only finitely many members of $\bigcup_{i \in I} T_i$, and we could find $i \in I$ such that all of them belong to T_i. ✗

Hence, by Zorn's Lemma there exists a maximal consistent set $\bar{S}$ containing S.

$\bar{S}$ is deductively closed i.e. $\bar{S} \vdash t$ implies $t \in \bar{S}$ since if $\bar{S} \vdash t$ then anything deducible from $\bar{S} \cup \{t\}$ is deducible from $\bar{S}$, and hence by maximality $\bar{S} \cup \{t\} = \bar{S}$.

Also, for every $t \in \mathcal{L}(P)$ we have either $t \in \bar{S}$ or $\neg t \in \bar{S}$ since if $t \notin \bar{S}$, then $\bar{S} \cup \{t\} \vdash \bot$, so by 2.8 $\bar{S} \vdash \neg t$, so $\neg t \in \bar{S}$.

$\bar{S}$ maximal consistent

Now we define $v: P \to 2$ by $v(p) = 1$ if and only if $p \in \bar{S}$.

We claim that the canonical extension $\bar{v}: \mathcal{L}(P) \to 2$ of r satisfies $\bar{v}(t) = 1$ if and only if $t \in \bar{S}$. (*)

We prove this by induction on the structure of t:

true by definition if $t \in P$, and true for $\bot$ since $\bar{S}$ is consistent. Suppose (*) holds for both s and t, and consider $(s \Rightarrow t)$.

Case 1: $t \in \bar{S}$, so $\bar{v}(t) = 1$. Then $\bar{v}(s \Rightarrow t) = 1$, but $\{t\} \vdash (s \Rightarrow t)$, so since $\bar{S}$ is deductively closed, we have $(s \Rightarrow t) \in \bar{S}$.

Case 2: $s \notin \bar{S}$, so $\bar{v}(s) = 0$, and again we have $\bar{v}(s \Rightarrow t) =$

But $\neg s \in \bar{S}$, and we have $\{\neg s\} \Rightarrow (s \Rightarrow t)$ (q.4, Sheet 2) and so $(s \Rightarrow t) \in \bar{S}$.

Case 3: $s \in \bar{S}$, $t \notin \bar{S}$, then $\bar{v}(s) = 1$, $\bar{v}(t) = 0$, so $\bar{v}(s \Rightarrow t) = 0$

But $(s \Rightarrow t) \notin \bar{S}$, since $\{s, (s \Rightarrow t)\} \vdash t$.

Hence r is a model for $\bar{S}$, and in particular a model for S

2.10 Remarks

a) Given a valuation $v: P \to 2$, the set $\{s \in \mathcal{L}(P) \mid \bar{v}(s) = 1\}$ is deductively closed, consistent, and for every s it contains either s or $\neg s$. So it is a maximal consistent set.

Hence, the maximal consistent subsets of $\mathcal{L}(P)$ are exactly the sets of formulae true under some valuation.

b) If P is countable, then there is a proof of 2·9 which does not use Zorn's Lemma. P countable $\Rightarrow \mathcal{L}(P)$ countable, so we can enumerate the members of $\mathcal{L}(P)$, as $(t_0, t_1, t_2, \dots)$

Given a consistent set S, we can enlarge it to a maximal consistent set as follows:

go through the t_i one by one, and at the n^{th} stage if t_n can be consistently added to our current set S_n we do so, and otherwise, we add $\neg t_n$. $S_{n+1} = S_n \cup \{t_n\}$ or $S_{n+1} = S_n \cup \{\neg t_n\}$ (This works since if $S_n \cup \{t_n\} \vdash \bot$, then $S_n \vdash \neg t_n$. Then $\bar{S} = \bigcup_{n \in \mathbb{N}} S_n$. $\bar{S}$ is consistent since it is the union of a chain of consistent sets, and for every $t \in \mathcal{L}(P)$, we have either $t \in \bar{S}$ or $\neg t \in \bar{S}$. So $\bar{S}$ is maximal consistent.

c) Given $S \subseteq \mathcal{L}(P)$, we define a relation $\leqslant_S$ on $\mathcal{L}(P)$ by $t \leqslant_S u$ if $S \cup [t] \vdash u$. This is reflexive and transitive, so if we define $\equiv_S$ by $t \equiv_S u$ if $t \leqslant_S u$ and $u \leqslant_S t$, then $\equiv_S$ is an equivalence relation, and $\leqslant_S$ induces a partial order on the quotient set $\mathcal{B}(S) = \mathcal{L}(P)/\equiv_S$ (the Lindenbaum algebra of a set S). We claim that $\mathcal{B}(S)$ is a Boolean algebra: the meet of $[s]$ and $[t]$ is $[s \wedge t]$, and their join is $[s \vee t]$. The complement of $[s]$ is $[\neg s]$.

S is consistent $\Leftrightarrow [T] \neq [\bot]$, i.e. $\mathcal{B}(S)$ has ≥ 2 elements.
So by 1.19 this implies that there exists a homomorphism $f : \mathcal{B}(S) \to 2$. Then the composite $\mathcal{L}(P) \to \mathcal{B}(S) \xrightarrow{f} 2$ coincides with $\bar{v}$, where v is its restriction to $P \subseteq \mathcal{L}(P)$.
Again, models for S correspond bijectively to homomorphisms $\mathcal{B}(S) \to 2$.

2.11 Corollary (Decidability Theorem)
Let $S \subseteq \mathcal{L}(P)$ be finite. Then there is an algorithm which determines, for any $t \in \mathcal{L}(P)$, whether or not $S \vdash t$

Proof
This is obvious for $\models$ (truth tables)

22/10/12 Logic and Set Theory (8)

2.12 Corollary (Compactness Theorem)

Let $S \subseteq \mathcal{L}(P)$. If every finite subset of S has a model then S has a model.

Proof

S has a model if and only if $S \not\vdash \bot$

But we know that if $S \vdash \bot$ then there is a finite $S' \subseteq S$ with $S' \vdash \bot$, the same holds for $\models$.

2.13 Remark

Let V be the set of all valuations of P (= all functions $P \to 2$). If we define U_t, for any $t \in \mathcal{L}(P)$, to be $\{v \in V \mid \bar{v}(t) = 1\}$ then the U_t form a base for a topology on V, since $U_s \cap U_t = U_{(s \wedge t)}$. The Compactness Theorem is exactly the assertion that this space is compact, since $\{U_t \mid t \in S\}$ covers $V \Leftrightarrow \{\neg t \mid t \in S\} \vdash \bot$.

(In fact, V is homeomorphic to the product of P copies of the discrete space $\{0, 1\}$, so the Compactness Theorem follows from Tychonoff's Theorem).

2.14 Examples

a) We use Compactness to show that any partial ordering on a set can be extended to a total ordering.

Given a poset $(A, \leq)$, take P to be the set of all propositions P_{xy}, $x \neq y$ in A.

Take S to consist of all propositions p_{xy} $(x < y$ in $A)$

$(p_{xy} \Leftrightarrow \neg p_{yx})$ $\quad x \neq y$ in A.

$(p_{xy} \Rightarrow (p_{yz} \Rightarrow p_{xz}))$ $\quad (x, y, z$ any 3 elements of $A)$.

Then v is a model for S if and only if

$\{(x,y) \mid$ either $x = y$ or $v(p_{xy}) = 1\}$ is a total ordering ~~extending~~ extending $\leq$.

So such an ordering exists providing that every finite subset $S' \subseteq S$ has a model. Given such an S', we can find a finite $A' \subseteq A$ such that all the p_{xy}s occuring in members of S' have $\{x, y\} \subseteq A'$. So we can construct a model for S' by constructing a total ordering of A' which extends $(\leq \cap A' \times A')$, and this can be done without Zorn's Lemma, by making a finite number of choices.

$G = (E, V)$

b) Let G be a ~~collection~~ graph. By an n-colouring of G, we mean a function $V \xrightarrow{c} \{1, 2, \ldots, n\}$ such that if $\{x, y\}$ is an edge, then $c(x) \neq c(y)$.

We claim that a graph is n-colourable if and only if all of its finite subgraphs are n-colourable.

To prove this, let $P = \{p_{xi} \mid x \in V,\ 1 \leq i \leq n\}$

and let $S = \{(p_{xi} \Rightarrow (p_{xj} \Rightarrow \bot)) \mid x \in V,\ i \neq j\}$

$\cup \{\bigwedge_{i=1}^{n} p_{xi} \mid x \in V\} \cup \{(p_{xi} \Rightarrow (p_{yi} \Rightarrow \bot)) \mid \{x, y\} \in E,\ 1 \leq i \leq n\}$

Then a model for S' is an n-colouring of (V, E).

If $S' \subseteq S$ is finite, then we can find a finite $V' \subseteq V$ such that the p_{xi} occuring in members of S' all satisfy $x \in V'$, so an n-colouring of $(V', E' = E \cap \mathcal{P}V')$ will yield a model for S'.

Chapter 3 : The Predicate Calculus

3.1 Definition

By a signature, we mean a pair $\Sigma = (\Omega, \Pi)$, where Ω is a set of operation symbols ω, each equipped with an arity $\alpha(\omega) \in \mathbb{N}$ and Π is a set of relation symbols π with a similar arity.

Given a signature Σ, a Σ-structure is a set A equipped with operations $\omega_A : A^{\alpha(\omega)} \to A$, for each $\omega \in \Omega$, and relations $[\![\pi]\!]_A \subseteq A^{\alpha(\pi)}$ for each $\pi \in \Pi$.

We could (and many logicians do) dispence with operations symbols, by replacing ~~them with~~ each n-ary operation ω by an $(n+1)$-ary relation ω^*, where $(x_1, \ldots, x_{n+1}) \in [\![\omega^*]\!]$, if an only if $x_{n+1} = \omega(x_1, \ldots, x_n)$.

3.2 Definition

Given a signature $\Sigma = (\Omega, \Pi)$, we define the set of terms over Σ as follows:

a) We have a supply of variables $x, y, z \ldots$ (or $x, x', x'', \ldots$ which are terms

b) If $t_1, \ldots, t_n$ are terms, ~~then~~ ω an n-ary operation symbol, then $\omega t_1 t_2 \ldots t_n$ is a term.

c) If A is a Σ-structure, we recursively define the <u>interpretation</u> t_A in A of each term t involving variables in the set $\{x_1, \ldots, x_n\}$ as a function $A^n \to A$:

If $t = x_i$ for some i, then $t_A(a_1, \ldots, a_n) = a_i$.

If $t = \omega u_1 u_2 \ldots u_m$ where $\alpha(\omega) = m$, then t_A is the composite

$$A^n \xrightarrow{((u_1)_A \ldots (u_m)_A)} A^m \xrightarrow{\omega_A} A$$

e.g. if we have a binary operation m, (interpreted by juxtaposition) then $mxmyz$ is interpreted in a structure A as the mapping $(a, b, c) \mapsto a(bc)$

and $mmxyz$ is interpreted by the mapping $(a, b, c) \mapsto (ab)c$.

3.3 Definition

Let $\Sigma = (\Omega, \Pi)$ be a first order signature. The set $\mathcal{L}(\Sigma)$ of (first order) formulae over Σ is defined as follows:

a) atomic formulae which are of two kinds: if $\pi \in \Pi$, $\alpha(\pi) = n$ and $t_1, \dots, t_n$ are terms then $\pi(t_1, \dots, t_n)$ is an atomic formula; and if s, t are two terms, then $(s = t)$ is an atomic formula

b) $\bot$ is a formula, and if φ, ψ are formulae, so is $(\varphi \Rightarrow \psi)$
(We introduce $\neg\varphi$, $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \Leftrightarrow \psi)$ as in Chapter 2)

c) If φ is a formula, and x is a free variable of φ, then $(\forall x)\varphi$ is a formula. (We introduce $(\exists x)\varphi$ as shorthand for $\neg(\forall x)\neg\varphi$).

The (finite) set $FV(\varphi)$ of free variables in φ is defined by:

- for an atomic formula, $FV(\varphi)$ is the set of all variables which occur in φ.
- $FV(\bot) = \emptyset$, $FV(\varphi \Rightarrow \psi) = FV(\varphi) \cup FV(\psi)$
- $FV((\forall x)\varphi) = FV(\varphi) \setminus \{x\}$

Note that a variable may appear both free and bound in the same formula: e.g. $(\pi(x, y) \wedge (\forall x)\pi(z, x))$ (*)
We say that φ, ψ are α-equivalent if ψ is obtained from φ

by renaming its bound variables.

So $(\pi(x,y) \wedge (\forall w)\, \pi(z,w))$ is α-equivalent to (*).

Given a Σ-structure A, we interpret any formula φ with $FV(\varphi) \subseteq \{x_1, x_2, \ldots, x_n\}$ as a set $[\varphi]_A \subseteq A^n$, or equivalently a function $\varphi_A : A^n \to 2$, as follows:

If φ is $\pi(t_1, \ldots, t_m)$, φ_A is the composite

$A^n \xrightarrow{((t_1)_A, \ldots, (t_m)_A)} A^m \xrightarrow{\pi_A} 2$ (where π_A is the characteristic function of $[\pi]_A \subseteq A^m$)

If φ is $(s = t)$, φ_A is the composite

$A^n \xrightarrow{(s_A, t_A)} A^2 \xrightarrow{\delta} 2$, where $\delta(x,y) = 1$ iff $x = y$

If φ is $\bot$, φ_A is the constant function with value 0.

If φ is $(\psi \Rightarrow \chi)$, φ_A is the composite

$A^n \xrightarrow{(\psi_A, \chi_A)} 2^2 \xrightarrow{\Rightarrow_2} 2$ where $\Rightarrow_2(x,y) = 1$ unless $x = 1$, $y = 0$

If φ is $(\forall x_{n+1})\, \psi$, then

$[\varphi]_A = \{(x_1, \ldots, x_n) \in A^n \mid$ for all $x_{n+1} \in A$, $(x_1, \ldots, x_n, x_{n+1}) \in [\psi]$

By a context for φ, we mean a sequence $(x_1, \ldots, x_n)$ of distinct variables including all the free variables of φ.

By the canonical context for φ, we mean a list of its free variables in the order of their first (free) occurrence in φ.

Unless otherwise stated, we interpret formulae in their canonical contexts.

3.4 Definition

We say that a formula φ is satisfied in a structure A, and write $A \models \varphi$, if $\varphi_A : A^n \to 2$ is the constant function with value 1 (equivalently, $[\varphi]_A = A^n$). Note that if $x \in FV(\varphi)$, then $A \models \varphi$ iff $A \models (\forall x)\varphi$.

By the universal closure of φ, we mean the formula $(\forall x_1)(\forall x_2)\dots(\forall x_n)\varphi$, where $(x_1, x_2, \dots, x_n)$ is the canonical context for φ. We abbreviate this to $(\forall x_1, x_2, \dots, x_n)\varphi$.

A sentence or closed formula is one with no free variables.

By a first-order theory over Σ, we mean a set T of sentences in $\mathcal{L}(\Sigma)$ - called axioms of T. We say that A is a model for T (and write $A \models T$) if $A \models \varphi$ for all $\varphi \in T$.

3.5 Examples

a) The theory of groups is written over a signature (Ω, Π) with $\Omega = \{m, i, e\}$ with $\alpha(m) = 2$, $\alpha(i) = 1$, $\alpha(e) = 0$, and $\Pi = \emptyset$. The axioms of T are:

$(\forall x, y, z)(mx\,my\,z = m\,m\,x\,c\,y\,z)$

$(\forall x)(mx\,e = x)$

$(\forall x)(mx\,ix = e)$

Then a T-model is exactly a group.

b) The theory of fields has $\Omega = \{+, \times, 0, 1, -\}$ and $\Pi = \emptyset$. We write down the axioms of commutative rings with 1 as universal closures of equations.

Add the axioms $(10 \neq 1) \Rightarrow \bot)$ and

$(\forall x)((x=0) \vee (\exists y)(x \cdot y = 1))$

c) For the theory of posets, we take $\Omega = \emptyset$, $\Pi = \{\leq\}$ (write $(s \leq t)$ for $\leq(s,t)$) and axioms

$(\forall x)(x \leq x)$

$(\forall x,y,z)((x \leq y) \Rightarrow ((y \leq z) \Rightarrow (x \leq z)))$

$(\forall x,y)((x \leq y) \Rightarrow (y \leq x) \Rightarrow (x = y)))$

For tosets, we add the axiom $(\forall x,y)((x \leq y) \vee (y \leq x))$

d) The theory of (combinatorial) projective planes has $\Omega = \emptyset$, $\Pi = \{\pi, \lambda, \in\}$ with $\alpha(\pi) = \alpha(\lambda) = 1$, $\alpha(\in) = 2$.

and axioms $(\forall x)(\pi(x) \vee \lambda(x))$

$(\forall x) \neg(\pi(x) \wedge \lambda(x))$

$(\forall x,y)((x \in y) \Rightarrow (\pi(x) \wedge \lambda(y)))$

$(\forall x,y)((\pi(x) \wedge \pi(y) \wedge \neg(x=y)) \Rightarrow (\exists z)((x \in z) \wedge (y \in z) \wedge \ldots$

$\ldots(\forall w)(((x \in w) \wedge (y \in w)) \Rightarrow (z = w))))$

and the dual of this.

Logic and Set Theory (10)

Substitution

If φ is a formula, x a variable, and t a term, we write $\varphi[t/x]$ for the formula obtained from φ by replacing each free occurrence of x in φ by a copy of t, provided that none of the variables of t occur bound in φ, otherwise, we must first rename the bound variables of φ so that they don't clash with variables in t.

Similarly, we write $\varphi[t_1, t_2/x_1, x_2]$; note that this is not the same as $\varphi[t_1/x_1][t_2/x_2]$.

3.6 Definition

Suppose that T is a set of sentences and φ a sentence. We say that T semantically entails φ, and write $T \models \varphi$, if every model for T satisfies φ.

If T and φ are free variables, then we say $T \models \varphi$ iff $T' \models \varphi'$, where φ' is obtained from φ by adding new constants (= nullary operation symbols) c_i, to our signature for each free variable x_i in φ and setting $\varphi' = \varphi[c_1, \ldots, c_n/x_1, \ldots]$ and T' is obtained from T by substituting the c_i, and then for the x_i in the members of T, and then quantifying universally over any remaining free variables.

3.7 Definition

The axioms of the predicate calculus are all formulae of the following 7 forms:

$(\varphi \Rightarrow (\psi \Rightarrow \varphi))$ (k)

$((\varphi \Rightarrow (\psi \Rightarrow \chi)) \Rightarrow ((\varphi \Rightarrow \psi) \Rightarrow (\varphi \Rightarrow \chi)))$ (s)

$(\neg\neg\varphi \Rightarrow \varphi)$ (T)

$((\forall x)\varphi \Rightarrow \varphi[t/x])$ $(x \in FV(\varphi))$ (I)

$((\forall x)(\varphi \Rightarrow \psi) \Rightarrow (\varphi \Rightarrow (\forall x)\psi))$ $(x \in FV(\psi) \setminus FV(\varphi))$ (IP)

$(\forall x)(x = x)$ (R)

$(\forall x, y)((x = y) \Rightarrow (\varphi \Rightarrow \varphi[y/x]))$ (Sub)

Our rules of inference are:

1. From φ and $(\varphi \Rightarrow \psi)$, we may infer ψ, provided either ψ has a free variable or φ doesn't. Note that if φ has a free variable, then $\emptyset \vdash \varphi$, and $\emptyset \models (\varphi \Rightarrow \bot)$, but $\emptyset \not\vdash \bot$. (MP)

2. From φ, we may infer $(\forall x)\varphi$, provided that x does not occur free in any hypothesis used in the deduction of φ. (Generalization)

By a deduction in the predicate calculus from a set of hypotheses T, we mean a finite sequence $\varphi_1, \varphi_2, \ldots, \varphi_n$ such that for each i we have either φ_i an axiom, or $\varphi_i \in T$, or φ_i obtained from earlier formulae by one of the two rules of inference.

We say that T syntactically entails φ if there is a deduction from T whose last member is φ, and write $T \vdash \varphi$.

3.8 Lemma (Soundness Theorem)

If $T \vdash \varphi$ then $T \models \varphi$.

Proof (by induction on the length of a deduction of φ from T). (essentially the same as 2.7)

3.9 Proposition (Deduction Theorem)

Suppose either that ψ has a free variable, or φ doesn't.

Then
$T \vdash (\varphi \Rightarrow \psi)$ iff $T \cup \{\varphi\} \vdash \psi$

Proof:
$\Rightarrow$ Exactly like 2.8
$\Leftarrow$ Mostly like 2.8, but we have to consider the case when $\psi = (\forall x)\chi$ was obtained by an application of (Gen).

Then, by our inductive hypothesis, we have a deduction of $(\varphi \Rightarrow \chi)$ from T, and we can add
$(\forall x)(\varphi \Rightarrow \chi)$ (Gen)
$((\forall x)(\varphi \Rightarrow \chi) \Rightarrow (\varphi \Rightarrow (\forall x)\chi))$ (IP)
$(\varphi \Rightarrow (\forall x)\chi)$ (MP)
provided $x \notin FV(\varphi)$.

But if x is free in φ, then φ wasn't used in the deduction of χ, so we actually have $T \vdash (\forall x)\chi$, from which we can obtain $T \vdash (\varphi \Rightarrow \psi)$ by (K) plus (MP).

3.10 Theorem (Completeness Theorem)
$T \vdash \varphi$ iff $T \models \varphi$.

Proof:
$\Rightarrow$ is 3.8.
$\Leftarrow$: We use 3.9 plus axiom (T) to reduce to the case $\varphi = \bot$. Then, reduce to the case when T is a set of sentences: we need to show that if T is consistent (i.e. $T \not\vdash \bot$) then T has a model.

Basic idea: suppose that T is maximal consistent (i.e. for any sentence φ we have $\varphi \in T$ or $\neg\varphi \in T$) and T has witnesses, i.e. whenever $FV(\varphi) = \{x\}$ and $T \vdash (\exists x)\varphi$ there is a closed term t (one with no variables)

such that $T \vdash \varphi[t/x]$.

Then T has a <u>term model</u>, i.e. if C is the set of closed terms, and $\equiv$ is the equivalence relation on C given by $s \equiv t$ iff $T \vdash (s = t)$. Then the set $A = C/\equiv$ can be given the structure of a model of T.

Logic and Set Theory (11)

We use two constructions:

a) enlarging T to a maximal consistent set of sentences.
(i.e. for all sentences φ we have $\varphi \in T$ or $\neg\varphi \in T$)

b) adding witnesses. We need to show that if T is consistent and $T \vdash (\exists x)\varphi$ where $FV(\varphi) = \{x\}$, then T' is consistent where T' is obtained from T by adding a new constant c_φ to Ω and adding a new-axiom $\varphi[c_\varphi/x]$.

Suppose that T' is not consistent; i.e. $T' \vdash \bot$. Then $T \vdash \varphi[c_\varphi/x]$ by 3.9. Since T does not mention c_φ, we can convert this into a deduction of $\neg\varphi$ from T, and hence $T \vdash (\forall x)\neg\varphi$ by (Gen). But $T \vdash \neg(\forall x)\neg\varphi$ by assumption, so $T \vdash \bot$.

Now, given T, we enlarge Ω by adding constants c_φ for each finite φ such that $T \vdash (\exists x)\varphi$, and enlarge T by adding the axioms $\varphi[c_\varphi/x]$ for each such φ.

This is still consistent, since a deduction of $\bot$ from it could involve only finitely many of the new axioms.

Now, starting from $T = T_0$, we construct an infinite sequence (T_n) of theories, such that for even n, T_{n+1} is a completion of T_n, and for odd n, T_{n+1} is obtained by adding witnesses to T_n.

Then $T_\omega = \bigcup_{n=0}^{\infty} T_n$ is consistent, complete, and has witnesses.
Now let C be the set of closed terms over the signature of T_ω, and define $\equiv$ on C by $s \equiv t$ iff $T_\omega \vdash (s = t)$. We can make the quotient $M = C/\equiv$ into a Σ-structure by setting $\omega_M([t_1], [t_2], \ldots, [t_n]) = [\omega t_1 t_2 \ldots t_n]$
and $([t_1], \ldots, [t_n]) \in [\pi]_M$ iff $T_\omega \vdash \pi(t_1, \ldots, t_n)$

Now, an induction over the structure of φ shows that for any φ with $FV(\varphi) = \{x_1, \ldots, x_n\}$, we have
$([t_1], \ldots, [t_n]) \in [\varphi]_M$ iff $T_\omega \vdash \varphi[t_1, \ldots, t_n / x_1, \ldots, x_n]$
Thus, for a sentence φ, we have $M \models \varphi$ iff $T_\omega \vdash \varphi$.
In particular, since $T_0 \subseteq T_\omega$, M is a model of $T_0 = T$ □

Remark

Note that if our original signature Σ_0 is countable, then so is Σ_ω and therefore the model M is countable.

3.11 Corollary (Compactness Theorem)

If T is a set of sentences such that every finite subset of T has a model, then T has a model.

Proof

Obvious for "is consistent" □

3.12 Corollary (Upward Löwenheim-Skolem Theorem)

Suppose that a first order theory T has an infinite model. Then it has models of arbitrarily large cardinality.

Proof

Given a set I, extend the signature of T by adding new constants $\{c_i : i \in I\}$, and extend T to T' by adding new axioms $\neg(c_i = c_j)$ for all $i \neq j$. Given a finite subset $T'' \subseteq T'$, we can construct a model for T'' by taking our infinite T-model M and assigning distinct values in M to all the constants which occur in members of T''. since there are only finitely many of these.

So by 3.11 T' has model M'; but this is just a T-model equipped with an injection $I \to M'$, sending i to $(c_i)_{M'}$. □

3.13 Corollary (Downward Löwenheim-Skolem Theorem)

If a theory T over a countable signature has an infinite model, then it has a countably-infinite model.

Proof

Apply the argument of 3.12 with a countably infinite family of constants. This produces a consistent theory T' over a countable signature. By the remark after 3.10, T' has a countable model, which must be countably infinite.

A theory T is called <u>categorical</u> if it has just one model, up to isomorphism.

We can have categorical first order theories with finite models, but not with infinite ones. There are categorical axiomatisations for infinite structures (e.g. Peano's Postulates for $\mathbb{N}$, 1890):

a) 0 is a natural number

b) If n is a natural number, then $S(n)$ is a natural number (→ successor)

c) $(\forall n) \neg (0 = S(n))$

d) $(\forall n, m)((S(m) = S(n) \Rightarrow (m = n))$

e) If $p(n)$ is a property of natural numbers such that $p(0)$ holds and $(\forall n)(P(n) \Rightarrow P(S(n)))$, then $(\forall n) P(n)$.

- a) and b) give $\Sigma = (\Omega, \Pi)$, with $\Omega = \{0, S\}$, $\Pi = \emptyset$, where 0 is nullary and S is unary.

Note that e) involves a quantification over subsets of the intended structure $\mathbb{N}$. (→ so not 1st order) We could replace e) by a scheme of axioms of the form

$$(\forall y_1, \ldots, y_n)((\varphi[0/x] \wedge (\forall x)(\varphi \Rightarrow \varphi[S(x)/x])) \Rightarrow (\forall x)\varphi)$$

with one for each formula φ with $FV(\varphi) = \{x, y_1, \ldots, y_n\}$

but this yields e) only for countably many subsets of $\mathbb{N}$.

3.14 Definition

First-order Peano Arithmetic is the theory over the signature with $\Omega = \{0, S, a, m\}$, $\Pi = \emptyset$, and axioms

$(\forall x) \neg (Sx = 0)$

$(\forall x, y)((Sx = Sy) \Rightarrow (x = y))$

$(\forall x)(axO = x)$

$(\forall x, y)(axSy = Saxy)$

$(\forall x)(mxO = O)$

$(\forall x, y)(mxSy = amxyx)$

$(\forall y_1, \ldots, y_n)((\varphi[0/x] \wedge (\forall x)(\varphi \Rightarrow \varphi[Sx/x])) \Rightarrow (\forall x)\varphi)$

for each formula φ with $FV(\varphi) = \{x, y_1, \ldots, y_n\}$

31/10/12 Logic and Set Theory (12)

3.14 Definition: First Order Peano Arithmetic

This is the theory over the signature with $\Omega = \{0, s, a, m\}$ and $\Pi = \emptyset$, and axioms

$(\forall x)\ \neg(sx = 0)$

$(\forall x, y)((sx = sy) \Rightarrow (x = y))$

$(\forall x)\ (ax0 = x)$

$(\forall x, y)\ (axsy = saxy)$

$(\forall x)\ (mx0 = 0)$

$(\forall xy)\ (mxsy = amxyx)$

$(\forall y_1, \ldots, y_n)((\varphi[0/x] \wedge (\forall x)(\varphi \Rightarrow \varphi[sx/x])) \Rightarrow (\forall x)\varphi)$
for each formula φ with $FV(\varphi) = \{x, y_1, \ldots, y_n\}$

3.15 Definition

a) A first order theory T is called complete iff for every sentence φ in the language of T, we have either $T \vdash \varphi$ or $T \vdash \neg\varphi$ (but not both)

b) We say that two Σ-~~models~~ structures M, N are elementarily equivalent if, for all sentences φ in $L(\Sigma)$, $M \models \varphi$ iff $N \models \varphi$.

3.16 Proposition

A first order theory is complete iff all its models are elementarily equivalent.

Proof

Suppose T is complete. Then, if M, N are T-models

$M \models \varphi$ implies $T \not\vdash \neg\varphi$, so $T \vdash \varphi$, so $N \models \varphi$.

Conversely, suppose that all T-models are elementarily equivalent. Then, if $T \not\vdash \varphi$, $T \cup \{\neg\varphi\}$ is consistent, so has a model, so all T-models satisfy $\neg\varphi$, so $T \vdash \neg\varphi$. □

We say a theory T is κ-categorical, for some cardinal κ, if any two T-models of cardinality κ are isomorphic.

For example: the theory of dense total orders without top and bottom elements is (complete and) countably categorical, since any countable model is isomorphic to $\mathbb{Q}$. However, it is not (cardinality $\mathbb{R}$)-categorical, since $\mathbb{R}$ and $\{x \in \mathbb{R} \mid (x > 0) \vee (x \in \mathbb{Q})\}$ are both models, and the second contains countable intervals but the first does not.

Another example: the theory of algebraically closed fields of characteristic 0 is complete, but not countably categorical: $\overline{\mathbb{Q}}$ (the algebraic closure of $\mathbb{Q}$) and $\overline{\mathbb{Q}(\pi)}$ are non-isomorphic countable models.

By the Löwenheim-Skolem argument, if a first order theory has finite models of arbitrarily large cardinality, then it must have infinite models.

So there is no first order theory whose models are finite groups.

We can axiomatise infinite groups: take the theory of groups

and add $(\forall x_1, \ldots, x_n)(\exists y)(\bigwedge_{i=1}^{n} \neg(x_i = y))$ for each $n \geq 1$. But there is no finite set of axioms whose models are exactly infinite groups. If there were, say $\varphi_1, \ldots, \varphi_n$, then {axioms for groups} $\cup$ $\{\neg \bigwedge_{i=1}^{n} \varphi_i\}$ would be a first order axiomatisation of finite groups.

Chapter 4 : Zermelo - Fraenkel Set Theory

We want to construct a first-order theory ZF whose intended models are "universes of sets".

Start with a signature containing a binary predicate symbol $\in$ (and nothing else apart from $=$).

The fundamental axiom is Extensionality which says:

$(\forall x, y)((\forall z)((z \in x) \Leftrightarrow (z \in y)) \Rightarrow (x = y))$

We also need the ability to 'collect' elements into sets:

Frege (1893) suggested the comprehension scheme

$(\exists y)(\forall x)(\overset{x \in y}{(y \in x)} \Leftrightarrow \varphi)$ where φ is any formula with $FV(\varphi) = \{x\}$

Russell's Paradox: consider the formula $\varphi = \neg(x \in x)$

$(\exists y)(\forall x)((x = y) \Leftrightarrow \neg(x \in x))$

Is $y \in y$? If $y \in y$ then $\varphi[y/x]$ i.e. $\neg(y \in y)$

If $\neg(y \in y)$ then $\neg\varphi[y/x]$, i.e. $(y \in y)$ ⚔

To avoid this, there are three main possibilities:

a) Russell advocated type theory, which is a many-sorted theory

in which the formula $x = y$ is only meaningful if x, y have the same sort, and $x \in y$ if $\text{sort}(x) = \text{sort}(x) + 1$ (and so on)
But this isn't set theory, since it abandons extensionality.

b) W. Quine 'New Foundations' (NF):
We work with sets, but restrict the Comprehension Scheme to stratifiable formulae φ, i.e. to formulae which could be interpreted in type theory. This seems uncomfortably strong in comparison with ZF

c) Zermelo (1904) proposed replacing Comprehension by the Separation Scheme:

$$(\forall z_1, \ldots, z_n)(\forall y_1)(\exists y_2)(\forall x)((x \in y_2) \Leftrightarrow ((x \in y_1) \wedge \varphi))$$

where φ is any formula with $FV(\varphi) = \{x, z_1, \ldots, z_n\}$

22/11/12

4.1 Definition

Zermelo set theory is the first order theory with the following axioms:

i) $(\forall x, y)\big((\forall z)((z \in x) \Leftrightarrow (z \in y)) \Rightarrow (x = y)\big)$ (Ext)

ii) $(\forall z_1, \ldots, z_n)(\forall y_1)(\exists y_2)(\forall x)\big((x \in y_2) \Leftrightarrow ((x \in y_1) \wedge \varphi)\big)$ (Sep)

where φ is any formula with $FV(\varphi) = \{x, z_1, \ldots, z_n\}$

[Note that this axiom implicitly defines an $(n+1)$ary operation on our universe V, sending $(z_1, \ldots, z_n, y_1)$ to the unique y_2 whose existence it asserts. We write $\{x \in y_1 \mid \varphi\}$ for a term denoting this y_2; note that x appears in this term as a bound variable.

iii) $(\exists x)(\forall y)\ \neg(y \in x)$ (Empty Set)

[we ~~me~~ introduce the constant symbol $\emptyset$ as a name for this x]

iv) $(\forall x, y)(\exists z)(\forall w)\big((w \in z) \Leftrightarrow ((w = x) \vee (w = y))\big)$ (Pair)

[we write $\{x, y\}$ for the set z whose existence is asserted by this, and abbreviate $\{x, x\}$ for $\{x\}$.]

v) $(\forall x)(\exists y)(\forall z)\big((z \in y) \Leftrightarrow (\exists w)((z \in w) \wedge (w \in x))\big)$ (Union)

[We write $\bigcup x$ for this y, and abbreviate $\bigcup\{x, y\}$ to $x \cup y$. Note that we don't need an extra axiom to define $\bigcap x$ if $x \neq \emptyset$, and we don't want to define $\bigcap \emptyset$. We can also form $x \cap y = \bigcap\{x, y\}$ and $x \setminus y = \{z \in x \mid \neg(z \in y)\}$

vi) $(\forall x)(\exists y)(\forall z)\big((z\in y)\Leftrightarrow(\forall w)\big((w\in z)\Rightarrow(w\in x)\big)\big)$ (Pow)

[We write $\mathcal{P}x$ for the y whose existence is asserted by this, and abbreviate $(\forall w)\big((w\in z)\Rightarrow(w\in x)\big)$ to $(z\subseteq x)$.]

We can now define the ordered pair $\langle x,y\rangle$ to be $\{\{x\},\{x,y\}\}$ (Kuratowski - Wiener ordered pair). We define

$\mathrm{First}(t) = \cup\cap t$ if $t\neq\emptyset$ and $\mathrm{First}(\emptyset)=\emptyset$, and

$\mathrm{Second}(t) = \cup(\cup t\setminus\cap t)$ if $\cup t\setminus\cap t\neq\emptyset$ and

$\mathrm{Second}(t) = \mathrm{First}(t)$ otherwise. Then we have

$\mathrm{First}(\langle x,y\rangle)=x$, $\mathrm{Second}(\langle x,y\rangle)=y$, for all x,y, so that $\langle x,y\rangle=\langle z,w\rangle$ iff $x=z$, $y=w$.

We write 'x is an ordered pair' for $x=\langle\mathrm{First}(x),\mathrm{Second}(x)\rangle$

We can form the product set $x\times y$ as

$\{z\in\mathcal{P}\mathcal{P}\cup\{x,y\}\mid(z\text{ is an ordered pair})\wedge(\mathrm{First}(z)\in x)\wedge(\mathrm{Second}(z)\in y)\}$.

We write 'x is a function' for

$(\forall y)\big((y\in x)\Rightarrow(y\text{ is an ordered pair}))\wedge(\forall u,v_1,v_2)\big((\langle u,v_1\rangle\in x)\wedge(\langle u,v_2\rangle\in x)\Rightarrow(v_1=v_2)\big)$

We write '$x:y\to z$' for $(x\text{ is a function})\wedge(\forall w)\big((w\in x)\Rightarrow\big((\mathrm{First}(w)\in y)\wedge(\mathrm{Second}(w)\in z)\big)\big)\wedge(\forall u)\big((u\in y)\Rightarrow(\exists v)(\langle u,v\rangle\in x)\big)\big)$

and we can form the set z^y of all functions y to z as

$\{x\in\mathcal{P}(y\times z)\mid x:y\to z\}$]

vii) $(\exists x)\big((\emptyset\in x)\wedge(\forall y)\big((y\in x)\Rightarrow(y\cup\{y\}\in x)\big)\big)$ (Inf)

[We abbreviate $y\cup\{y\}$ by y^+, the successor of y

02/11/12

Logic and Set Theory (13)

A set with the closure properties specified in (Inf) is called a <u>successor set</u>. By (Sep), if a successor set exists, there is a smallest one, namely

$\{z \in x \mid (\forall y)(y \text{ is a successor set}) \Rightarrow (z \in y)\}$

where x is any given successor set. We write ω for the smallest successor set]

Note that $\phi^{+\overbrace{+\cdots+}^{n \text{ times}}} = n$ is a set with n elements.

We could alternatively have used the sequence $\phi, \mathcal{P}\phi, \mathcal{P}\mathcal{P}\phi, \ldots$ or the sequence $\phi, \{\phi\}, \{\{\phi\}\}, \ldots$ to construct an infinite sequence of sets.

Note that we have a formula which is satisfied by precisely the ordered pairs $\langle \phi, \phi \rangle, \langle \phi^{+}, \mathcal{P}\phi \rangle, \langle \phi^{++}, \mathcal{P}\mathcal{P}\phi \rangle, \ldots$

This formula defines a 'function' (not yet a set), whose 'domain' is a set, but whose 'range' is not yet a set.

A. Fraenkel (1923) introduced a new axiom scheme of Replacement which solves this problem (and also makes the pair set axiom redundant)

05/11/12 Logic and Set Theory (4)

Sets and Classes

By a class, we mean an equivalence class of formulae φ with one free variable x, where we consider φ and ψ to be equivalent if $(\forall x)(\varphi \Leftrightarrow \psi)$ is deducible from the axioms of our set theory. (Equivalently, φ and ψ have the same interpretation in any model.) We say that φ and ψ are extensionally equivalent if this holds.

We write $\{x \mid \varphi\}$ for the class corresponding to φ.

We will also denote classes by letters such as M, and write $(t \in M)$ to mean that $\varphi[t/x]$ is deducible, where φ is a formula defining M.

We say the class $\{x \mid \varphi\}$ is a set if $(\exists y)(\forall x)((x \in y) \Leftrightarrow \varphi)$ is deducible. If we can deduce $\neg(\exists y)(\forall x)((x \in y) \Leftrightarrow \varphi)$, then we call $\{x \mid \varphi\}$ a proper class.

e.g. the universe V is a class, corresponding to the formula $(x = x)$ and it is a proper class, as is the Russell class $\{x \mid \neg(x \in x)\}$. Similarly, a class of pairs is an extensional-equivalence class of formulae φ with two free variable x, y denoted $\{\langle x, y\rangle \mid \varphi\}$. We say that $t \in \{\langle x, y\rangle \mid \varphi\}$ iff t is an ordered pair $\langle u, v\rangle$, and $\varphi[u, v/x, y]$ is deducible. Similarly for classes of n-tuples $(n \geq 3)$.

A <u>function class</u> is a class of pairs F for which

$(\forall x, y, y')(\langle x, y\rangle \in F \wedge \langle x, y'\rangle \in F) \Rightarrow (y = y'))$

The <u>domain</u> of the function class F is the class $\{x \mid (\exists y)(\langle x, y\rangle \in F$

<u>4.2 Definition</u> (Zermelo - Fraenkel set theory - ZF).

ZF is obtained from Zermelo set theory by adding the axiom scheme of Replacement and the axiom of Foundation, where:

- Replacement says ~~$\forall\forall\forall$~~

$(\forall z_1, \ldots, z_n)(\forall x)(\exists y)((\forall v)(v \in y) \Leftrightarrow (\exists u)(u \in x) \wedge \varphi))$

for any φ with free variables $\{u, v, z_1, \ldots, z_n\}$ satisfying

$(\forall z_1 \ldots z_n, u, v, v')((\varphi \wedge \varphi[v'/v]) \Rightarrow (v = v'))$

- Foundation says

$(\forall x)(\neg(x = \emptyset) \Rightarrow (\exists y)((y \in x) \wedge (x \cap y = \emptyset)))$

Using Replacement, we can extend the sequence of sets $\emptyset, \mathcal{P}\emptyset, \mathcal{P}\mathcal{P}\emptyset, \ldots$ (usually denoted $V_0, V_1, V_2, \ldots$) by forming the set $\{V_n \mid n \in \omega\}$, and then forming

$V_\omega = \bigcup \{V_n \mid n \in \omega\}$

Then we can form $V_{\omega+1} = \mathcal{P} V_\omega$, $V_{\omega+2} = \mathcal{P} V_{\omega+1}, \ldots$ and $V_{\omega+\omega} = \bigcup \{V_\omega, V_{\omega+1}, V_{\omega+2}, \ldots\}$

Informally, foundation says that every set is a member of some V_α (Proved later).

25/11/12 Logic and Set Theory (14)

4.3 Definition

By 'x is a transitive set' we mean

$(\forall y, z)(((y \in x) \wedge (z \in y)) \Rightarrow (z \in x))$ → $\bigcup x \subseteq \bigcup_{y \in x} y$

Equivalently, $x \subseteq \mathcal{P}x$ or $\bigcup x \subseteq x$.

Given any set x, there is a smallest transitive set $TC(x)$ with $x \subseteq TC(x)$, namely $\bigcup\{x, \bigcup x, \bigcup\bigcup x, \ldots\}$. This is transitive since $y \in TC(x)$ implies $y \in \bigcup^n x$ for some n, then any member of y is in $\bigcup^{n+1} x$.

Conversely, if $x \subseteq z$ and z is transitive, then $\bigcup x \subseteq \bigcup z \subseteq z$, $\bigcup^2 x \subseteq \bigcup^2 z^2 \subseteq \bigcup z \subseteq z$ and so on, i.e. $TC(x) \subseteq z$.

Note also that $TC(\{x\})$ is the smallest transitive set z with $x \in z$

Exercise: show that V_α is a transitive set (power set and union of transitive sets are transitive)

4.4 Proposition

In the presence of the other axioms of ZF, the axiom of Foundation is equivalent to the axiom-scheme of $\in$-Induction:

$(\forall z_1, \ldots, z_n)((\forall x)((\forall y)((y \in x) \Rightarrow \varphi[y/x]) \Rightarrow \varphi) \Rightarrow (\forall x)\varphi)$

where φ is any formula with $FV(\varphi) = \{x, z_1, \ldots, z_n\}$.

Proof

($\Leftarrow$) Define 'x is a regular set' to mean

$(\forall y)((x \in y) \Rightarrow (\exists z)((z \in y) \wedge (z \cap y = \emptyset)))$

Then (Foundation) $\Leftrightarrow$ $(\forall x)(x$ is a regular set$)$. So it suffices to prove that $((\forall y)((y \in x) \Rightarrow (y$ regular$)) \Rightarrow (x$ regular$))$

But if all members of x are regular and $x \in z$, then either $(x \cap z = \phi)$, in which case x is a member of z disjoint from z, or $(\exists y)((y \in x) \wedge (y \in z))$, in which case y is regular since $y \in x$, and so z has a member disjoint from z

($\Rightarrow$) For the converse, we do the case $n = 0$ i.e. no parameters.

Suppose $(\forall x)(\forall y)((y \in x) \Rightarrow \varphi[y/x]) \Rightarrow \varphi)$

but $(\exists x) \neg\varphi$. Now, for this x, $u = \{y \in TC(\{x\}) \mid \neg\varphi[y/x]\}$ $\neq \phi$

is non-empty since it contains x.

So by [Foundation] $\overset{(\exists z)}{}((z \in u) \wedge (z \cap u) = \phi))$

Then φ holds at all members of z, since they are members of $TC(\{x\}) \setminus u$, but $\neg\varphi[z/x]$ ⊗

07/11/12 Logic and Set Theory (15)

4.5 Definition

a) Let R be a relation class (i.e. an extensional equivalence class of formulae with two free variables). We say that R is well-founded if it satisfies

$$(\forall x)(\neg(x=\emptyset) \Rightarrow (\exists y)((y \in x) \wedge (\forall z)((\langle z,y\rangle \in R) \Rightarrow \neg(z \in x))))$$

[So the Axiom of foundation says that $\in$ is well founded]

b) More generally, let M be a class. We say that R is well-founded relative to M if it satisfies (WF)

$$(\forall x)((x \subseteq M) \wedge \neg(x=\emptyset)) \Rightarrow (\exists y)((y \in x) \wedge (\forall z)((\langle z,y\rangle \in R) \Rightarrow \neg(z \in x))))$$

c) We say R is local relative to M if the R-predecessors in M of any element of M form a set, i.e.

$$(\forall x \in M)(\exists y)(\forall z)((z \in y) \Leftrightarrow ((z \in M) \wedge (\langle z,x\rangle \in R)))$$

where $(\forall x \in M)\varphi$ means $(\forall x)((x \in M) \Rightarrow \varphi)$ and $(\exists x \in M)\varphi$ means $(\exists x)((x \in M) \wedge \varphi)$

If R is local relative to M, then for any subset x of M we can form a set $RC(x)_M$, which is the ~~subset of~~ smallest set y satisfying

$$(x \subseteq y) \wedge ((\forall u,v)(((v \in y) \wedge (\langle u,v\rangle \in R) \wedge (u \in M)) \Rightarrow (u \in y)))$$

Specifically, $RC_M(x) = \bigcup\{x, F(x), F(F(x)), \ldots\}$

where $F(x)$ denotes the set of R-predecessors of members of x which belong to M.

Hence we can prove a proposition.

4·6 Proposition

If R is well-founded and local relative to M, then we have a principle of R-induction over M:

$((\forall x \in M)((\forall y \in M)((\langle y, x\rangle \in R) \Rightarrow \varphi[y/x]) \Rightarrow \varphi) \Rightarrow (\forall x \in M)\varphi$

Proof (Just like 4·4)

Suppose that φ satisfies the inductive condition, and suppose we are given $x \in M$ for which $\neg\varphi$ holds. Form $\{y \in RC_M(\{x\}) \mid \neg\varphi[y/x]\}$. This is non-empty and $\subseteq M$, so has an R-minimal member y, say.

Then $\varphi[z/x]$ holds for all $z \in M$ with $\langle z, y\rangle \in R$, but $\varphi[y/x]$ fails] ⊗

4·7 Lemma

Suppose that R is well founded and local relative to a class M. Then there is a transitive relation $\bar{R}$ such that $R \subseteq \bar{R}$, and $\bar{R}$ is well founded and local relative to M.

Proof

We define $\bar{R}$ by $\langle y, x\rangle \in \bar{R} \Leftrightarrow y \in (RC_M(\{x\}) \setminus \{x\})$

Clearly $\bar{R} \supseteq R$, from the construction of $RC_M(\{x\})$.

R is transitive, since $\langle y, x\rangle \in \bar{R}$ implies $RC_M(\{y\}) \subseteq RC_M(\{x\})$ and hence from $\langle z, y\rangle \in \bar{R}$ we can deduce $\langle z, x\rangle \in \bar{R}$.

It is local, since the $\bar{R}$ predecessors of x are the members of the set $RC_M(\{x\}) \setminus \{x\}$

For well-foundedness, suppose we are given $\emptyset \neq x \subseteq M$ such that x has no $\bar{R}$-minimal member. Form the set

$$y = \{z \in RC_M(x) \mid (\exists u \in M)((\langle u, z\rangle \in \bar{R}) \wedge (u \in x))\}$$

Then $y \neq \emptyset$ since $x \subseteq y$, and y has no R-minimal member ⊗

4.8 Theorem (R-Recursion Theorem)

Suppose that R is well founded and local relative to a class M, and suppose that given a function class G of two variables which is defined on $M \times V$ (i.e. $G(x, y)$ is defined whenever $x \in M$). Then there is a unique function class F of one variable defined on M, and satisfying

$$(\forall x \in M)\ F(x) = G(\langle x, \{F(y) \mid ((y \in M) \wedge (\langle y, x\rangle \in R))\}\rangle) \quad (*)$$

Proof:

For uniqueness, suppose that F_1, F_2 both satisfy (*). Then we may prove that $(\forall x \in M)(F_1(x) = F_2(x))$ by R-induction over M.

For existence, we define the notion of an attempt as follows:

(f is an attempt) means (f is a function) $\wedge$ ($\mathrm{dom}\, f \subseteq M$) ●

$$\wedge\ (\mathrm{dom}\, f = RC_M(\mathrm{dom}\, f))$$

$$\wedge\ (\forall x)((x \in \mathrm{dom}\, f) \Rightarrow (f(x) = G(\langle x, \{f(y) \mid ((y \in M) \wedge \langle y, x\rangle \in R)\}\rangle)$$

Note that if f_1, f_2 are attempts, then $\mathrm{dom}\, f_1 \cap \mathrm{dom}\, f_2$ is an R-closed subset of M, and we can prove $(\forall x \in \mathrm{dom}\, f_1 \cap \mathrm{dom}\, f_2)(f_1(x) = f_2(x))$, by R-induction over this set.

Hence if we define F by

$\langle x, y\rangle \in F \Leftrightarrow (\exists f)((f \text{ is an attempt}) \wedge (f(x) = y))$ then

F is a function class, with $\operatorname{dom} F \subseteq M$, and F satisfies (*) for all $x \in \operatorname{dom} F$. So we need to show $\operatorname{dom} F = M$

i.e. $(\forall x \in M)(\exists f)((f \text{ is an attempt}) \wedge (x \in \operatorname{dom} f))$

Suppose not : given $x \in M$ not in the domain of any attempt, consider $[y \in RC_M(\{x\}) \mid y \text{ is not in the domain of any attempt}]$

This set is non-empty since it contains x, so it has an $\bar{R}$-minimal member y, say.

Now $(\forall z \in RC_M(\{y\}) \setminus \{y\})(\exists!\ w)$ ($\exists!$ a unique) $(\exists f)(f \text{ is an attempt}) \wedge \langle z, w\rangle \in f))$

So the set of all such pairs $\langle z, w\rangle$ is a function f_0 with domain $RC_M(\{y\}) \setminus \{y\}$, and f_0 is an attempt.

Now define $f_1 = f_0 \cup \{\langle y, G(\langle y, \{f_0(z) \mid (\langle z, y\rangle \in R \wedge (z \in M))\}\rangle)\}$;

Then f_1 is an attempt with domain $RC_M(\{y\})$,

contradicting the assumption that y is not in the domain of any attempt ⊗

4.9 Remark

Given a set a and a well-founded relation $r \subseteq a \times a$, we can prove the existence of functions defined by recursion over r by a simpler argument: in this case the union of all attempts is a set (by Replacement), and hence is itself an attempt. If this attempt (f, say), is not total, then an r-minimal member of $a \setminus \operatorname{dom} f$ yields a contradiction.

4.10 Example

The relation $S \subseteq \omega \times \omega$ defined by $\langle x, y \rangle \in S \Leftrightarrow y = x^+$ is well founded, as is the strict order-relation on ω (which we may take to be $\in \cap (\omega \times \omega)$).

These yield the two versions of mathematical induction:

$$(\forall x)((x \subseteq \omega) \wedge (\emptyset \in x) \wedge (\forall n \in \omega)((n \in x) \Rightarrow (n^+ \in x))) \Rightarrow (x = \omega)$$

$$(\forall x)((x \subseteq \omega) \wedge (\forall n \in \omega)((\forall m \in n)(m \in x) \Rightarrow (n \in x))) \Rightarrow (x = \omega)$$

4.11 Definition

A binary relation-class R is said to be extensional on a class M if $(\forall x, y \in M)((\forall z \in M)((\langle z, x \rangle \in R) \Leftrightarrow (\langle z, y \rangle \in R)) \Rightarrow (x = y))$

4.12 Theorem (Mostowski's Theorem)

Let a be a set and $r \subseteq a \times a$ an extensional well-founded relation. Then there exists a unique pair $\langle b, f \rangle$ such that b is a transitive set and $f : \langle a, r \rangle \to \langle b, \in \cap b \times b \rangle$ is an isomorphism of sets-with-binary-relation.

Proof

Uniqueness: suppose given $\langle b, f\rangle, \langle b', f'\rangle$ both satisfying the conditions. Then $g = f' \circ f^{-1} : b \to b'$ is an isomorphism $\langle b, \in\rangle \to \langle b', \in\rangle$ and we can prove that $(\forall x \in b)(g(x) = x)$ by $\in$-induction over b. Hence $b = b'$ and $f = f'$.

Existence: we define f by r-recursion over a:

$f(x) = \{f(y) \mid (y \in a) \wedge (\langle y, x\rangle \in r)\}$

and we define $b = \{f(x) \mid x \in a\}$ (which is a set by replacement).

Clearly f is surjective and $\langle x, y\rangle \in r$ implies $f(x) \in f(y)$.

For the converse of the latter, it suffices to show that f is injective, since we know that if $f(x) \in f(y)$ then $f(x) = f(z)$ for some z with $\langle z, y\rangle \in r$. We show that f is injective by r-induction:

suppose $(\forall y \in a)((\langle y, x\rangle \in r) \Rightarrow (\forall z \in a)(f(z) = f(y) \Rightarrow (z = y)))$

and suppose $f(x) = f(z)$.

Then $(\forall y)(\langle y, x\rangle \in r) \Rightarrow (\exists u)((\langle u, z\rangle \in r) \wedge (f(y) = f(u)))$

where $(\forall y)((\langle y, x\rangle \in r) \Rightarrow \langle y, z\rangle \in r))$ by the induction hypothesis.

Similarly if $\langle u, z\rangle \in r$ then we deduce $f(u) = f(y)$ for some y with $\langle y, x\rangle \in r$, whence $u = y$, and so $\langle u, x\rangle \in r$.

Hence x and z have the same r-predecessors, so by extensionality $x = z$.

4.13 Definition

We say that a binary relation-class R is trichotomous on a class M if $(\forall x, y \in M)(\langle x,y\rangle \in R) \vee (\langle y,x\rangle \in R) \vee (x = y))$

Note that if R is well-founded, then these three possibilities are mutually exclusive, since if two of them hold then $\{x, y\}$ has no R-minimal member.

Also, if R is well-founded, and trichotomous, then it is transitive, since if $\langle x,y\rangle \in R$ and $\langle y,z\rangle \in R$ but not $\langle x,z\rangle \in R$ then $\{x,y,z\}$ has no R-minimal member.

So if R is well-founded, and trichotomous, then $R \cup =$ is a total ordering of M.

Also, a well-founded trichotomous relation is extensional, since if x and y have the same R-predecessors, then we cannot have $\langle x,y\rangle \in R$ or $\langle y,x\rangle \in R$.

4.14 Corollary

Given a set a and a well founded trichotomous relation r on a, there is a unique pair $\langle b, F\rangle$ such that b is transitive and totally ordered by $\in$, and $f : \langle a, r\rangle \to \langle b, \in\rangle$ is an isomorphism of ordered sets.

Note: r is well-founded and trichotomous on a

$\Leftrightarrow$ every non-empty $b \subseteq a$ has an r-least upper bound.

We use the term well-ordering for a well-founded trichotomous relation.

Chapter 5 : Ordinals

5.1 Definition

An ordinal is a transitive set α which is well-ordered by $\in$.

Clearly $\emptyset$ is an ordinal, and so are $\emptyset^+$, $\emptyset^{++}$, and ω. To prove this, we need some lemmas.

5.2 Lemma

If α is an ordinal, so is $\alpha^+ = \alpha \cup \{\alpha\}$

Proof

Transitivity : if $x \in y \in \alpha^+$ then either $y \in \alpha$ or $y = \alpha$.

So in either case $x \in \alpha$ and hence $x \in \alpha^+$.

Trichotomy : if $x, y \in \alpha^+$, then we have one of

$(x \in \alpha) \wedge (y \in \alpha)$ in which case we have $(x \in y) \vee (y \in x) \vee (x = y)$

$(x \in \alpha) \wedge (y = \alpha)$ in which case $x \in y$

$(x = \alpha) \wedge (y \in \alpha)$ $y \in x$

$(x = \alpha) \wedge (y \notin \alpha)$ $x = y$

5.2 Lemma

If α is an ordinal, so is $\alpha^+ = \alpha \cup \{\alpha\}$

5.3 Lemma

Every member of an ordinal is an ordinal.

Proof

Let α be an ordinal, $x \in \alpha$. If $z \in y \in x$, then $y \in \alpha$ and $z \in \alpha$ by transitivity of α. So we have one of $(z \in x)$, $(z = x)$ or $(x \in z)$. But if either $(z = x)$ or $(x \in z)$, then $\{x, y, z\}$ has no $\in$-minimal member. So $(z \in x)$, hence x is transitive.

If y, z are both members of x, then $y, z \in \alpha$ by transitivity, so we have $(y \in z) \vee (y = z) \vee (z \in y)$.

5.4 Lemma

If α, β are ordinals, then either $(\alpha \subseteq \beta)$ or $(\beta \subseteq \alpha)$

Proof

Suppose $\alpha \not\subseteq \beta$. Then $\alpha \setminus \beta \neq \emptyset$, so it has an $\in$-least member r, say. We can show that $r = \alpha \cap \beta$:

- If $\delta \in r$, then $\delta \in \alpha$ and $\delta \notin \alpha \setminus \beta$, so $\delta \in \alpha \cap \beta$
- Conversely, if $\delta \in \alpha \cap \beta$, then r and δ are both members of α, so we have one of $(\delta \in r)$, $(\delta = r)$, $(r \in \delta)$

But either $(\delta = r)$ or $(r \in \delta)$ would imply $r \in \beta$, since $\delta \in \beta$ ✗

So by extensionality we have $r = \alpha \cap \beta$, and in particular $\alpha \cap \beta \in \alpha$.

Similarly, if $\beta \not\subseteq \alpha$, then $\alpha \cap \beta \in \beta$. So if neither inclusion holds, $(\alpha \cap \beta) \in (\alpha \cap \beta)$, contradicting Foundation ※

5.5 Corollary

i) For ordinals α and β, $(\alpha \subseteq \beta)$ is equivalent to $(\alpha \in \beta) \vee (\alpha = \beta)$

ii) For any two ordinals α and β, we have $(\alpha \in \beta) \vee (\alpha = \beta) \vee (\beta \in \alpha)$

Proof

i) $(\alpha \in \beta)$ implies $\alpha \subseteq \beta$ since β is transitive, so $(\Leftarrow)$ is obvious. But if $(\alpha \subseteq \beta)$ and $(\alpha \neq \beta)$, then $\alpha = (\alpha \cap \beta) \in \beta$ by the proof of 5.4.

ii) This is immediate from 5.4 and 5.5 i). □

Writing On for the class of ordinals, we have shown that $\alpha \in \beta \in On$ implies $\alpha \in On$ and that $\alpha, \beta \in On$ implies $(\alpha \in \beta) \vee (\alpha = \beta) \vee (\beta \in \alpha)$

5.6 Corollary (Burali-Forti Paradox)

On is a proper class.

Proof

If On were a set, we would have $On \in On$ ※

5.7 Lemma

If x is a ~~set whose~~ subset of On, then $\bigcup x \in On$.

Proof

$\bigcup x$ is transitive, since it is a union of transitive sets. The

members of Vx satisfy trichotomy by 5.5 ii)

5.8 Theorem

Let M be any class satisfying

$(\forall x)((x \in M) \Rightarrow (x^+ \in M))$ and $(\forall x)((x \subseteq M) \Rightarrow (Vx \in M))$

Then $On \subseteq M$.

Proof

Suppose not. Then there is an $\in$-least $\alpha \in On \setminus M$. Suppose that α has an $\in$-greatest member β say. Then $\beta \in M$, and we have $(\forall r)((r \in \alpha) \Leftrightarrow ((r \in \beta) \vee (r = \beta)))$ i.e. $\alpha = \beta^+$ ⚡

If α has no greatest member, then $(\forall \beta \in \alpha)(\exists r \in \alpha)(\beta \in r)$, so $\alpha = U\alpha$. But $\in$-minimality of α means that $\alpha \subseteq M$ ⚡

5.9 Proposition

$(\forall \alpha)((\alpha \in On) \Leftrightarrow (\alpha \text{ is transitive}) \wedge (\forall \beta \in \alpha)(\beta \text{ is transitive})))$

Proof

($\Rightarrow$) is immediate from 5.3.

($\Leftarrow$) Let On' denote the class of transitive sets whose members are all transitive. If $On' \neq On$, let x be an $\in$-minimal member of $On' \setminus On$. Then all members of x are in On, so they satisfy trichotomy by 5.5 ii). x is transitive by definition, so $x \in On$ ⚡

Let α be an ordinal. We say that α is a successor if it has an $\in$-greatest member, or equivalently, if $\alpha = \beta^+$ for some β. If not, we say that α is a limit; this is equivalent to saying $\alpha = \bigcup \alpha$. (Note that 0 is a limit; if we want to exclude it, we will refer to non-zero limits).

Example : $\omega = \bigcup \omega$ is a non-zero limit ordinal.

We tend to denote (non-zero) limit ordinals by λ.

5.10 Definition

a) We define a function-class rk (rank) by $\in$-recursion over V :
$rk(x) = \bigcup \{ rk(y)^+ \mid y \in x \}$. Clearly, $rk(x) \in On$ for all x.

b) We define a function-class $\alpha \mapsto V_\alpha$ by $\in$-recursion over On :
$V_\alpha = \bigcup \{ \mathcal{P} V_\beta \mid \beta \in \alpha \}$. Equivalently :
$V_0 = \emptyset$, $V_{\alpha^+} = \mathcal{P} V_\alpha$, $V_\lambda = \bigcup \{ V_\beta \mid \beta \in \lambda \}$ for λ a non-zero limit.

5.11 Theorem

For $x \in V$ and $\alpha \in On$, we have

$x \in V_\alpha \Leftrightarrow rk(x) < \alpha$ (i.e. $rk(x) \in \alpha$)

and $x \subseteq V_\alpha \Leftrightarrow rk(x) \leq \alpha$ (i.e. $rk(x) \subseteq \alpha$)

Proof

The second assertion is the special case $\alpha = \beta^+$ of the first, so we need only prove the first.

Suppose true for all $\beta < \alpha$, and suppose $x \in V_\alpha$. Then any $y \in x$ belongs to V_β for some $\beta < \alpha$. Hence ~~$(\forall y \in x)(\exists \beta < \alpha)(rk(y) < \beta)$~~. So

If $\alpha = \beta^+$ is a successor, we have

$(\forall y \in x)(rk(y) < \beta)$

so $(\forall y \in x)(rk(y)^+ \leq \beta)$, so $rk(x) \leq \beta < \alpha$.

If α is a limit, we have ~~$(\forall y \in x)(\exists \beta < \alpha)(rk(y) < \beta)$~~

so $(\exists \beta < \alpha)(rk(x) = \beta)$ and then

$(\forall y \in x)(rk(y) < \beta)$. So $(\forall y \in x)(rk(y)^+ \leq \beta)$

so $rk(x) \leq \beta^+ < \alpha$.

5.11 Proposition

i) $x \in V_\alpha \Leftrightarrow rk(x) < \alpha$

ii) $x \subseteq V_\alpha \Leftrightarrow rk(x) \leq \alpha$

Proofs

ii) Follows from i) by substituting α^+ for α.

i)($\Rightarrow$) By $\in$-induction on α:

Assume $(\forall x)(\forall \beta < \alpha)(x \in V_\beta \Rightarrow rk(x) < \beta)$

and assume that $x \in V_\alpha$. If $\alpha = \beta^+$ is a successor, then $x \subseteq V_\beta$ i.e. $(\forall y \in x)(y \in V_\beta)$. So by the induction hypothesis $(\forall y \in x)(rk(y) < \beta)$. Hence

$rk(x) = \bigcup \{ rk(y)^+ \mid y \in x \} < \beta^+ = \alpha$

If α is a limit, then since $V_\alpha = \bigcup \{ V_\beta \mid \beta < \alpha \}$ we have $x \in V_\beta$ for some $\beta < \alpha$. So by the induction hypothesis $rk(x) < \beta < \alpha$.

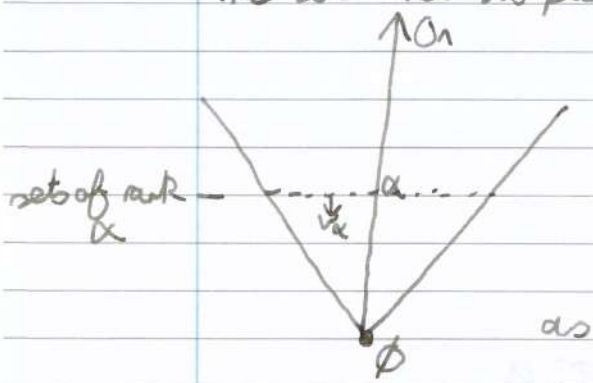
($\Leftarrow$) By $\in$-induction on x.

Suppose $(\forall y \in x)(rk(y) < \alpha \Rightarrow y \in V_\alpha)$ and suppose $rk(x) < \alpha$.

If $\alpha = \beta^+$, then $rk(y) < \beta$ for all $y \in x$, so by the induction hypothesis, $(\forall y \in x)(y \in V_\beta)$ i.e. $x \in \mathcal{P} V_\beta = V_\alpha$.

If α is a limit, then there exists β with $rk(x) < \beta < \alpha$. So $rk(y) < \beta$ for all $y \in x$, so by the induction hypothesis, $y \in V_\beta$ for all $y \in x$ i.e. $x \in \mathcal{P} V_\beta = V_{\beta^+} \subseteq V_\alpha$.

The set-theorist's picture of the universe:



If we are given a model V of $ZF \setminus \{\text{Foundation}\}$ we can still define On as the class of transitive sets well-ordered by $\in$, and we can still define the function class $\alpha \mapsto V_\alpha$

In this context, foundation is equivalent to the assertion

$(\forall x)(\exists \alpha \in On)(x \in V_\alpha)$

Indeed, in this model, $\bigcup \{V_\alpha \mid \alpha \in On\}$ is the class of regular sets as defined in 4.4 i.e. $\{x \mid (\forall y)((x \in y) \Rightarrow (\exists z \in y)(y \cap z = \emptyset))\}$

Arithmetic of Ordinals

We can define $\alpha+\beta$, $\alpha\beta$, either recursively or synthetically.

5.12 Definition

a) The recursive definitions say

$\alpha + 0 = \alpha$, $\alpha + \gamma^+ = (\alpha+\gamma)^+$

$\alpha + \lambda = \bigcup \{\alpha+\beta \mid \beta < \lambda\}$ if λ is a non-zero limit.

$\alpha \cdot 0 = 0$ $\alpha \cdot \gamma^+ = (\alpha \cdot \gamma) + \alpha$

$\alpha \cdot \lambda = \bigcup \{\alpha \cdot \beta \mid \beta < \lambda\}$ if λ is a limit.

b) The synthetic definitions say

$\alpha+\beta$ is the order-type of $\alpha \amalg \beta = (\alpha \times \{0\}) \cup (\beta \times \{1\})$ ordered so that every $\langle \gamma, 0 \rangle$ with $\gamma \in \alpha$ comes before every $\langle \delta, 1 \rangle$ with $\delta \in \beta$ (i.e. with reverse-lexicographic ordering)

$\alpha \cdot \beta$ is the order-type of the product $\alpha \times \beta$, with reverse lexicographic ordering (we must check that this is a well ordering).

5.13 Lemma

The two definitions in 5.12 coincide.

Proof

We need to show that the synthetic definition satisfies all clauses of the recursive definition. ($+_s$: $+$ via synthetic definition)

But $\alpha +_s 0 = \underbrace{otp}_{\text{order type}}(\alpha \times \{0\} \cup \emptyset) = \alpha$

$\alpha +_s (r^+) = otp(\alpha \times \{0\} \cup r \times \{1\} \cup \{\langle r, 1\rangle\})$

$= otp(\alpha \times \{0\} \cup r \times \{1\})^+ = (\alpha + r)^+$

by the induction hypothesis.

$\alpha +_s \lambda = otp(\alpha \times \{0\} \cup \lambda \times \{1\})$

$= otp(\bigcup \{\alpha \times \{0\} \cup \beta \times \{1\} \mid \beta < \lambda\})$

and each $(\alpha \times \{0\} \cup \beta \times \{1\})$ occurs as an initial segment of $\alpha \times \{0\} \cup \lambda \times \{1\}$, so this is $\bigcup \{\alpha + \beta \mid \beta < \lambda\}$

Similarly, (for $\cdot_s$ synthetic multiplication)

$\alpha \cdot_s 0 = otp(\emptyset) = 0$

$\alpha \cdot_s r^+ = otp(\alpha \times r \cup \alpha \times \{r\}) = otp(\alpha \times r) + otp(\alpha \times \{r\})$

$= \alpha \cdot r + \alpha$

$\alpha \cdot_s \lambda = otp(\bigcup \{\alpha \times \beta \mid \beta < \lambda\})$ and each $\alpha \times \beta$ is an initial segment of $\alpha \times \lambda$, so this is $\bigcup \{\alpha \cdot \beta \mid \beta < \lambda\}$

5.15 Lemma

i) If $\beta < \gamma$ then $\alpha + \beta < \alpha + \gamma$

ii) If $\beta \leq \gamma$, then $\alpha + \beta \leq \alpha + \gamma$

iii) If $\alpha \neq 0$ and $\beta < \gamma$, then $\alpha \cdot \beta < \alpha \cdot \gamma$

iv) If $\alpha \leq \beta$ then $\alpha \cdot \gamma \leq \beta \cdot \gamma$

Proof

We prove i) and iii) synthetically. If $\beta < \gamma$, then $\alpha \amalg \beta$ is a proper initial segment of $\alpha \amalg \gamma$, so $otp(\alpha \amalg \beta) < otp(\alpha \amalg \gamma)$

Similarly, if $\alpha \neq 0$, then $\alpha \times \beta$ is a proper initial segment of $\alpha \times \gamma$ so again $otp(\alpha \times \beta) < otp(\alpha \times \gamma)$

ii), iv) are proved by induction on γ:

$\alpha + 0 = \alpha \leq \beta = \beta + 0$

If $\alpha + \gamma \leq \beta + \gamma$ for all $\gamma < \lambda$, then

$\alpha + \lambda = \bigcup \{\alpha + \gamma \mid \gamma < \lambda\} \leq \bigcup \{\beta + \gamma \mid \gamma < \lambda\} = \beta + \lambda$

The proof of iv) is similar.

Note that $1 + \omega = \bigcup \{1 + n \mid n < \omega\} = \omega = 0 + \omega$

and $2 \cdot \omega = \bigcup \{2 \cdot n \mid n < \omega\} = \omega = 1 \cdot \omega$

Note also that $\square + \beta$ and $\square \cdot \beta$ are not continuous at limits i.e. $\omega + 1 \neq \bigcup \{n + 1 \mid n < \omega\}$ and

$\omega \cdot 2 = \omega + \omega \neq \bigcup \{n \cdot 2 \mid n < \omega\}$

5.15 Lemma

a) $0 + \alpha = \alpha$ for all α

b) $0 \cdot \alpha = 0$ for all α

c) $\alpha \cdot 1 = \alpha = 1 \cdot \alpha$ for all α

d) $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$ for all α, β, γ

e) $\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$ for all α, β, γ

f) $\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$ for all α, β, γ

Proof (sketch)

All this can be proved synthetically, by establishing an order-isomorphism between two well-ordered sets. For example, e) says that $\alpha \times (\beta \amalg \gamma)$ is order isomorphic to $(\alpha \times \beta) \amalg (\alpha \times \gamma)$ but the left-hand side has elements of the form $\langle \alpha', \langle \beta', 0 \rangle \rangle$ or $\langle \alpha', \langle \gamma', 1 \rangle \rangle$ where $\alpha' < \alpha$, $\beta' < \beta$, and $\gamma' < \gamma$. The right hand side has elements of the form $\langle \langle \alpha', \beta' \rangle, 0 \rangle$, $\langle \langle \alpha', \gamma' \rangle, 1 \rangle$. The obvious bijection, $\langle x, \langle y, z \rangle \rangle \mapsto \langle \langle x, y \rangle, z \rangle$ is an order isomorphism. With the exception of the first equality in c), which follows from a), plus $\alpha \cdot 0^+ = \alpha \cdot 0 + \alpha$, we can also prove by induction, over α in cases a)-c), and over γ in cases (d-f). e.g. the inductive proof of f):

$\alpha \cdot (\beta \cdot 0) = \alpha \cdot 0 = 0 = (\alpha \cdot \beta) \cdot 0$

$\alpha \cdot (\beta \cdot \gamma^+) = \alpha \cdot (\beta \cdot \gamma + \beta) = \alpha \cdot (\beta\gamma) + \alpha \cdot \beta$ by e)

$= (\alpha \cdot \beta) \cdot \gamma + \alpha \cdot \beta$ (induction hypothesis) $= (\alpha \cdot \beta) \cdot \gamma^+$

If λ is a limit, $(\alpha \cdot \beta) \cdot \lambda = \bigcup \{ (\alpha \cdot \beta) \cdot \tau \mid \tau < \lambda \}$

$= \bigcup \{ \alpha \cdot (\beta \cdot \tau) \mid \tau < \lambda \}$ by our induction hypothesis.

But $\beta \cdot \lambda = \bigcup \{ \beta \cdot \tau \mid \tau < \lambda \}$ and unless $\beta = 0$ (in which case both sides are zero) the sequence $\beta \cdot \tau$, $\tau < \lambda$ is strictly increasing

So $\beta \cdot \lambda$ is a limit and hence $\alpha \cdot (\beta \cdot \lambda) = \bigcup \{ \alpha \cdot \delta \mid \delta < \beta \cdot \lambda \}$

But $\delta < \beta \cdot \lambda$ implies that $\delta < \beta \cdot \tau$ for some $\tau < \lambda$

and hence $\alpha \cdot \delta < \alpha \cdot (\beta \cdot \tau)$ (again, unless $\alpha = 0$)

So $\bigcup \{ \alpha \cdot \delta \mid \delta < \beta \cdot \lambda \} = \bigcup \{ \alpha \cdot (\beta \cdot \tau) \mid \tau < \lambda \}$

Note that $(\alpha + \beta) \cdot \tau \neq \alpha \cdot \tau + \beta \cdot \tau$ in general.

e.g. $(1+1) \cdot \omega = 2 \cdot \omega = \omega$ but $1 \cdot \omega + 1 \cdot \omega = \omega + \omega \geq \omega$

5.16 Lemma (Division Algorithm for On)

Let $\alpha, \beta \in On$ with $\beta \neq 0$. Then there exist unique $\tau, \delta \in On$ with $\alpha = \beta \cdot \tau + \delta$ and $\delta < \beta$.

Proof

Since $\beta \geq 1$, we have $\beta \cdot \tau \geq 1 \cdot \tau = \tau$, so there exists τ' such that $\beta \cdot \tau' > \alpha$. Consider the least such τ' : it can't be a limit, since $\beta \cdot \lambda = \bigcup \{ \beta \cdot \tau \mid \tau < \lambda \}$, so $\tau' = \delta^+$, where $\beta \cdot \tau \leq \alpha < \beta \cdot \tau^+ = \beta \cdot \tau + \beta$

Now, there is a least δ' for which $\beta \cdot \tau + \delta' \geq \alpha$, and this δ' must be a successor δ^+ (moreover, $\delta' \leq \beta$, so $\delta < \beta$)

Then $\beta \cdot \tau + \delta \leq \alpha < \beta \cdot \tau + \delta^+ = (\beta \cdot \tau + \delta)^+$, so $\alpha = \beta \cdot \tau + \delta$.

Conversely, if $\alpha = \beta \cdot r + \delta$ with $\delta < \beta$, then $\beta \cdot r \leq \alpha < \beta \cdot (r^+)$ so r is uniquely determined, and $\beta \cdot r + \delta \leq \alpha < \beta \cdot r + \delta^+$, so δ is uniquely determined.

Ordinal Exponentiation

What should α^β be? Our first guess is that it should be the order-type of the set $F(\beta, \alpha)$ for all functions $\beta \to \alpha$.

This set can be totally ordered by lexicographic ordering (i.e. $f < g \Leftrightarrow f(r) < g(r)$ for the least r such that $f(r) \neq g(r)$), but this is not a well ordering if β is infinite and $\alpha \geq 2$.

Since if we define $f_i : \omega \to 2$ by $f_i(j) = 1$ if $i = j$, 0 otherwise, then the sequence (f_i) is strictly decreasing, so $\{f_i \mid i \in \omega\}$ has no least member.

To get round this, we use reverse lexicographic ordering and also cut down to the set $F_0(\beta, \alpha) \subseteq F(\beta, \alpha)$ of functions of finite support, i.e. those f such that $f(r) = 0$ for all but finitely many r.

5.17 Lemma

$F_0(\beta, \alpha)$ is well ordered by reverse lexicographic ordering.

Proof (by induction on β).

Let S be a non-empty subset of $F_0(\beta, \alpha)$. Pick $f \in S$; if f is identically zero then it is the least member of S.

Otherwise, there is a greatest $\beta' < \beta$ such that $f(\beta') \neq 0$.

Now let $S_1 = \{ g \in S \mid g(\tau) = 0 \text{ for all } \tau > \beta' \}$

Then every member of S precedes every member of $S \setminus S_1$, so we need to find the least element of S_1. Let α' be the least element of $\{ g(\beta') \mid g \in S_1 \}$ and let $S_2 = \{ g \in S_1 \mid g(\beta') = \alpha' \}$

Again, every member of S_2 precedes every member of $S_1 \setminus S_2$ (so we need to find the least element of S_2).

But S_2 is order-isomorphic to $\{ g|_{\beta'} \mid g \in S_2 \}$ which is a non-empty subset of $F_0(\beta', \alpha)$ and hence has a least element by our induction hypothesis.

We define α^β to be the order-type of $(F_0(\beta, \alpha), \text{r.l.o})$ (reverse lexicographic order)

5·18 Lemma

Ordinal exponentiation satisfies the recursive definition

$\alpha^0 = 1$, $\alpha^{\beta^+} = \alpha^\beta \cdot \alpha$

$\alpha^\lambda = \bigcup \{ \alpha^\beta \mid \beta < \lambda \}$ if λ is a non-zero limit.

Proof

$F_0(\phi, \alpha) = \{\phi\}$ which is the ordinal 1.

We have a bijection $F_0(\beta^+, \alpha) \to F_0(\beta, \alpha) \times \alpha$ given by $f \mapsto \langle f|_\beta, f(\beta) \rangle$ which is an order isomorphism when both sides are ordered by reverse lexicographic order.

If λ is a limit, $F_0(\lambda, \alpha)$ is the union of the subsets $G(\beta, \alpha)$, $\beta < \lambda$

where $G(\beta, \alpha) = \{f \in F_0(\lambda, \alpha) \mid f(\gamma) = 0 \text{ for all } \gamma \geq \beta\}$

But $G(\beta, \alpha)$ is order-isomorphic to $F_0(\beta, \alpha)$ and the

$G(\beta, \alpha)$ are initial segments of $F_0(\lambda, \alpha)$.

So the order isomorphisms $G(\beta, \alpha) \cong F_0(\beta, \alpha) \cong \alpha^\beta$

fit together to yield an order-isomorphism $F_0(\lambda, \alpha) \to \bigcup\{\alpha^\beta \mid \beta < \lambda\}$

5.19 Lemma

a) $\alpha^{(\beta+\gamma)} = \alpha^{\beta} \cdot \alpha^{\gamma}$
b) $\alpha^{(\beta \cdot \gamma)} = (\alpha^{\beta})^{\gamma}$

} for all $\alpha, \beta, \gamma \in O_n$

Proof

We prove a) synthetically: there is a bijection

$F_0(\beta \amalg \gamma, \alpha) \to F_0(\beta, \alpha) \times F_0(\gamma, \alpha)$

sending $F \mapsto \langle f_0, f_1 \rangle$ where $f_0(\beta') = F(\langle \beta', 0 \rangle)$

and $f_1(\gamma') = F(\langle \gamma', 1 \rangle)$,

which is easily seen to be an order isomorphism when both sides are ordered by reverse lexicographic ordering. We prove b) by induction on γ:

$\alpha^{(\beta \cdot 0)} = \alpha^{0} = 1 = (\alpha^{\beta})^{0}$

$\alpha^{(\beta \cdot \gamma^+)} = \alpha^{(\beta \cdot \gamma + \beta)} = \alpha^{(\beta\gamma)} \cdot \alpha^{\beta}$ by a)

$= (\alpha^{\beta})^{\gamma} \cdot \alpha^{\beta}$ by the induction hypothesis.

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

If $\gamma = \lambda$ is a limit, we deal first with the cases $\alpha = 0$ or 1, or $\beta = 0$. Otherwise, $\alpha^{\square}$ and $(\alpha^{\beta})^{\square}$ are strictly increasing functions, since we have $\alpha \geq 2$ and $\alpha^{\beta} \geq 2$.

So $\alpha^{(\beta\lambda)} = \alpha^{\bigcup\{\beta \cdot \gamma \mid \gamma < \lambda\}} = \bigcup\{\alpha^{\delta} \mid \delta < \beta \cdot \lambda\}$ since $\beta \cdot \lambda$ is a non-zero lim

$= \bigcup\{\alpha^{(\beta \cdot \gamma)} \mid \gamma < \lambda\}$ since $\{\beta \cdot \gamma \mid \gamma < \lambda\}$ is cofinal in $\beta\lambda$.

$= \bigcup\{(\alpha^{\beta})^{\gamma} \mid \gamma < \lambda\}$ by the induction hypothesis

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

We don't have $(\alpha \cdot \beta)^\gamma = \alpha^\gamma \cdot \beta^\gamma$ in general.

e.g. $(\omega \cdot 2)^2 = \omega(2 \cdot \omega) \cdot 2 = \omega \cdot \omega \cdot 2 = \omega^2 \cdot 2$

but $\omega^2 \cdot 2^2 = \omega^2 \cdot 4 > \omega^2 \cdot 2$

But it can also fail even when $\alpha \cdot \beta = \beta \cdot \alpha$, if γ is infinite:

e.g. $(2 \cdot 2)^\omega = 4^\omega = \bigcup \{4^n \mid n < \omega\} = \omega$

but $2^\omega \cdot 2^\omega = \omega \cdot \omega = \omega^2 > \omega$

Chapter 6: Choice and Well-Ordering

We've seen that $\omega < \omega+1 < \omega+2 < \ldots < \omega+\omega = \omega \cdot 2$

$\omega < \omega \cdot 2 < \omega \cdot 3 < \ldots < \omega \cdot \omega = \omega^2$

$\omega < \omega^2 < \omega^3 < \ldots < \omega^\omega$

$\omega < \omega^\omega < \omega^{\omega^\omega} < \ldots < \varepsilon_0$ = the least α such that $\alpha = \omega^\alpha$

$\varepsilon_0 < \varepsilon_1 < \ldots < \varepsilon_\omega < \ldots < \zeta_0$ = the least α such that $\alpha = \varepsilon_\alpha$

But these are all countable! Do there exist any uncountable ordinals?

Yes: Consider the set of all countable ordinals. More generally,

6.1 Lemma (Hartogs' Lemma)

For any set a, there exists an ordinal $\gamma(a)$ for which there is no injection $\gamma(a) \to a$.

Proof

Consider the set $b \subseteq \mathcal{P}(a \times a)$ of all well-orderings of subsets of a. From Mostowski's Theorem, we have a function class F assigning to each well-ordered set its order type.

So $\{F(\bigcup\{\{x \mid \langle x,x\rangle \in y\}, y\rangle) \mid y \in b\}\bigcup$ is a set by Replacement.

Clearly, this set is a transitive subset of On, and hence an ordinal $r(a)$.

But if we had an injection $f: r(a) \to a$, then $\{f(\beta) \mid \beta < r(\alpha)\}$ would be a well ordered subset of a with order type $r(a)$ ⁂

6·2 Corollary

We get a 'new' proof of the Bourbaki-Witt theorem 1·12:

Suppose, given a chain complete poset P, an inflationary map $f: P \to P$, and an element $x \in P$, we wish to find a fixed point $> x$.

We define a function class $G: On \to P$ by recursion:

$G(0) = x$

$G(\alpha^+) = f(G(\alpha))$

$G(\lambda) = \bigvee\{G(\alpha) \mid \alpha < \lambda\}$ for λ a non-zero limit.

(This works since $\alpha \leq \beta \Rightarrow G(\alpha) \leq G(\beta)$, and so $\{G(\alpha) \mid \alpha < \lambda\}$ is a chain.

The restriction of G to $r(P)$ isn't injective, i.e. $\exists \alpha < \beta < r(P)$ with $G(\alpha) = G(\beta)$.

But now $\alpha^+ \leq \beta$, so $G(\alpha^+) = f(G(\alpha)) \leq G(\beta) = G(\alpha)$ and hence $G(\alpha)$ is a fixed point of f, lying above $x = G(0)$. □

This proof requires the 'full strength' of ZF set theory, whereas the proof in chapter I works in Zermelo set theory (i.e. without Replacement)

Question: Does every set actually have a well-ordering?

Answer: (Zermelo, 1904) Yes, iff we assume the axiom of choice.

6·3 Definition

Given a set a, we write $\mathcal{P}^+a$ for $\mathcal{P}a \setminus \{\emptyset\}$

By a choice function for a, we mean a function $f: \mathcal{P}^+a \to a$ such that $f(b) \in b$ for all $b \in \mathcal{P}^+a$.

Note that, if $\{a_i \mid i \in I\}$ is a set of non-empty sets, then it is a subset of $\mathcal{P}^+ \bigcup \{a_i \mid i \in I\}$, so a choice function for $\bigcup \{a_i \mid i \in I\}$ will yield a choice function $I \to \bigcup \{a_i \mid i \in I\}$ as considered in Chapter I.

Thus, we can formulate the axiom of choice as $(\forall x)(\exists f)$ (f is a choice fn for x)

6·4 Theorem (Zermelo's Theorem)

In a model of ZF set theory, the sets which have choice functions are exactly the well orderable sets.

Proof

If $<$ is a well ordering of a, then we get a choice function $g: \mathcal{P}^+a \to a$ by setting $g(b) =$ the $<$-least member of b.

Conversely, suppose that we are given $g: \mathcal{P}^+a \to a$ (assume $a \neq \emptyset$)

We define a function class $F: On \to a$ by recursion:

$$F(\alpha) = \begin{cases} g(a \setminus \{F(\beta) \mid \beta < \alpha\}) & \text{if } \{F(\beta) \mid \beta < \alpha\} \neq a \\ g(a) & \text{otherwise} \end{cases}$$

By 6·1, $F|_{r(a)}$ is not injective. So there must be some $\alpha < r(a)$ such that $F(\alpha)$ is defined by the second clause i.e. $\{F(\beta) \mid \beta < \alpha\} = a$

Let α be the least ordinal for which this happens: then $F|_\alpha$ is injective.

Hence this is a bijection $\alpha \to a$, so $\{\langle F(\gamma), F(\beta)\rangle \mid \gamma < \beta < \alpha\}$ is a well-ordering of a.

6.5 Remarks

a) The proof we gave for 6.4 is not Zermelo's original. His proof doesn't use replacement (and is similar to the proof of Bourbaki-Witt in 1.12)

b) We can prove the Well-Ordering Theorem directly from Zorn's Lemma:
Given a set a, consider the set $P = \{\langle b, < \rangle \mid (b \subseteq a) \wedge (< \text{ is a well ordering of } b)\}$.
We order P by setting $\langle b_1, <_1 \rangle \leq \langle b_2, <_2 \rangle$ iff $b_1 \subseteq b_2$, b_1 is a $<_2$-initial segment of b_2, and $<_1 = <_2 \cap (b_1 \times b_1)$
Then P is chain-complete: given a chain $\{\langle b_i, <_i \rangle \mid i \in I\}$ the pair $\langle \bigcup\{b_i \mid i \in I\}, \bigcup\{<_i \mid i \in I\} \rangle$ is a member of P and a least upper bound for the chain.
$P \neq \emptyset$, as $\langle \emptyset, \emptyset \rangle \in P$, so it has a maximal element $\langle b_0, <_0 \rangle$ say. If $b_0 \neq a$, pick $x \in a \setminus b_0$ and set
$b_1 = b_0 \cup \{x\}$, $<_1 = <_0 \cup \{\langle y, x \rangle \mid y \in b_0\}$
Then $\langle b_0, <_0 \rangle < \langle b_1, <_1 \rangle$ ⚡
So $<_0$ is a well-ordering of a.

c) Results proved using Zorn's Lemma can also be proved using the Well Ordering Theorem.
e.g. Hamel's Theorem: given a vector space V, well order (the underlying set of) V as $\{x_\beta \mid \beta < \alpha\}$. Now define a sequence of subsets $S_\beta \subseteq V$, $\beta \leq \alpha^+$, by $S_0 = \emptyset$, and if $x_\beta \in \langle S_\beta \rangle$, set $S_{\beta^+} = S_\beta$, otherwise $S_{\beta^+} = S_\beta \cup \{x_\beta\}$

If λ is a limit ordinal, $S_\lambda = \bigcup \{ S_\beta \mid \beta < \lambda \}$

By induction, S_β is linearly independent for all β, and since $x_\beta \in \langle S_{\beta^+} \rangle$ for all β, we have $\langle S_{\alpha^+} \rangle = V$ i.e. S_{α^+} is a bas

Note also that we can prove the uncountable case of the completene theorem for propositional logic in the same style in which we did the countable case (cf. 2·16(b)), by well-ordering the set $\mathcal{L}(P)$ of a compound propositions.

6·6 Definition

We say that an ordinal is initial if there is no bijection $\alpha \to \beta$ with $\beta < \alpha$. Clearly, every $n \in \omega$ is initial, as is ω itself, and $\omega_1 = \tau(\omega)$, the set of all countable ordinals.

More generally, we define a function class $\alpha \mapsto \omega_\alpha$, from O_n to O_n, by recursion:

$\omega_0 = \omega$, $\omega_{\alpha^+} = \tau(\omega_\alpha)$

$\omega_\lambda = \bigcup \{ \omega_\alpha \mid \alpha < \lambda \}$ if λ is a non-zero limit ordinal.

6·7 Lemma

The ordinals of the form ω_α are exactly the infinite initial ordinal

Proof

We prove that every ω_α is initial by induction on α:

This is clear for $\alpha = 0$.

For a successor α^+, $\omega_{\alpha^+} = \tau(\omega_\alpha)$ does not inject into any

any $\beta < \omega_{\alpha^+}$, since every such β injects into ω_α.

For a limit λ, ω_λ is the limit of a strictly increasing sequence so if we had an injection $\omega_\lambda \to \beta$ for some $\beta < \omega_\lambda$, we would have an injection $\omega_\lambda \to \omega_\alpha$ for some $\alpha < \lambda$, and hence an injection $\omega_{\alpha^+} \to \omega_\alpha$ ✗

Now let β be an infinite initial ordinal. We have $\omega_\alpha \geq \alpha$ for all α, so there is a least α with $\omega_\alpha \geq \beta$. This α cannot be a non-zero limit, and cannot be zero since β is infinite.

So α is a successor δ^+, and then we have $\omega_\delta \leq \beta < \omega_{\delta^+} = r(\omega_\delta)$

So we have injections $\omega_\delta \to \beta$, $\beta \to \omega_\delta$, hence by Cantor-Bernstein (1.11) there is a bijection $\omega_\delta \to \beta$. Since β is initial, this implies that $\beta = \omega_\delta$ □

Informally, a <u>cardinal</u> is an equivalence class of sets under the relation of <u>equipollence</u> (i.e. that of being in bijective correspondence). But every equivalence class except $\{\emptyset\}$ is a proper class, so we need to find a way of representing this by sets.

If we assume the Axiom of Choice, then every equivalence class contains an ordinal, and hence contains a unique initial ordinal, which we can take as a canonical representative of the class i.e. we define $\mathrm{card}(x)$ to be the unique initial ordinal in bijection with x.

Without AC, we can't find a unique representative for each class, but we can find a representative subset of it, as follows

Define the essential rank of a set x as

$\bigcap \{ \beta \leq \text{rank}(x) \mid (\exists y)((\text{rk } y = \beta) \wedge (\exists \text{ a bijection } y \to x)) \}$

and then define $\text{card } x = \{ y \in V_{(\text{essential rank } x)^+} \mid (\exists \text{ a bijection } y \to x) \}$

From now on, all we assume about the function class card is

$(\forall x, y)((\text{card}(x) = \text{card}(y)) \Leftrightarrow (\exists \text{ a bijection } x \to y))$

We need a new notation for card ω_α : following Cantor, we denote it by $\aleph_\alpha$ (aleph - alpha)

Given cardinals m, n, we write $m \leq n$ to mean that there exists an injection $x \to y$ where $\text{card}(x) = m$, $\text{card}(y) = n$.

Cantor - Bernstein implies that this is a partial ordering of the class of cardinals.

The assertion that it is a total ordering is equivalent to the Axiom of Choice.

We define binary operations $+$, $\cdot$ and $\square^{\square}$ on the class of cardinals as follows :

$m + n = \text{card}(x \amalg y)$ where $\text{card } x = m$, $\text{card } y = n$,

$m \cdot n = \text{card}(x \times y)$ similarly, and

$m^n = \text{card}(x^y)$, x^y denoting the set of all functions $y \to x$

6.8 Lemma

Let m, m', n, n', p be cardinals.

a) If $m' \leq m$ and $n' \leq n$, then $m'+n' \leq m+n$ and $m' \cdot n' \leq m \cdot n$

b) If $m' \leq m$, then $m'^{n} \leq m^{n}$; and if $0 \neq n' \leq n$, then $m^{n'} \leq m^{n}$

c) $m+(n+p) = (m+n)+p$ and $m \cdot (n \cdot p) = (m \cdot n) \cdot p$

d) $m+n = n+m$ and $m \cdot n = n \cdot m$

e) $m \cdot (n+p) = m \cdot n + m \cdot p$

f) $m^{(n+p)} = m^{n} \cdot m^{p}$, $m^{n \cdot p} = (m^{n})^{p}$, $(m \cdot n)^{p} = m^{p} \cdot n^{p}$

Proof (Selected items)

Let a, b, c be sets with cardinalities m, n, p respectively.

b) Given an injection $f: a' \to a$, the map $g \mapsto f \circ g$ is an injection $a'^{b} \to a^{b}$

Given an injection $h: b' \to b$ with $b' \neq \emptyset$, pick $x_0 \in b'$ and define

$k: b \to b'$ by $k(y) = \begin{cases} x & \text{if } h(x) = y \\ x_0 & \text{if } y \notin \text{in}(h) \end{cases}$

Then k is surjective, and the map $g \mapsto g \circ k$ is an injection $a^{b'} \to a^{b}$

f) We want a bijection $a^{(b \amalg c)} \to a^{b} \times a^{c}$; take the mapping

$f \mapsto \langle f \circ i, f \circ j \rangle$ where $i: b \to b \amalg c$ is the map

$y \mapsto \langle y, 0 \rangle$, and $j: c \to b \amalg c$, $z \mapsto \langle z, 1 \rangle$

Similarly, we want a bijection $a^{b \times c} \to (a^{b})^{c}$; take the mapping

$f \mapsto \hat{f}$ where $\hat{f}(z)(y) = f(\langle y, z \rangle)$

We also want a bijection $(a \times b)^{c} \to a^{c} \times b^{c}$; take the mapping

$f \mapsto \langle p \circ f, q \circ f \rangle$ where $p(\langle x, y \rangle) = x$ and $q(\langle x, y \rangle) = y$

6.9 Proposition

For any ordinal α, $\aleph_\alpha \cdot \aleph_\alpha = \aleph_\alpha$

Proof (by induction on α)

We construct a well-ordering of $\omega_\alpha \times \omega_\alpha$, which has order-type ω_α

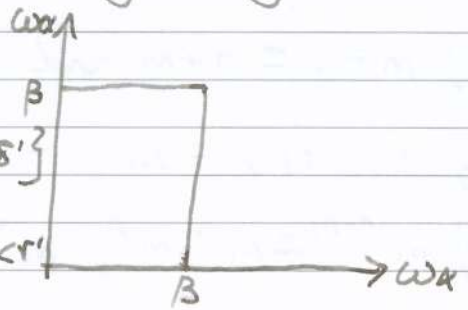
We order pairs $\langle \gamma, \delta \rangle \in \omega_\alpha \times \omega_\alpha$ by setting

$\langle \gamma, \delta \rangle < (\gamma', \delta')$

$\Leftrightarrow$ EITHER $\max\{\gamma, \delta\} < \max\{\gamma', \delta'\}$

OR $\max\{\gamma, \delta\} = \max\{\gamma', \delta'\}$ and $\gamma < \gamma'$

OR $\gamma = \max\{\gamma, \delta\} = \max\{\gamma', \delta'\} = \gamma'$ and $\delta < \delta'$



To show that this is a well-ordering, let S be a non-empty subset of $\omega_\alpha \times \omega_\alpha$.

Let $\beta = \min\{\max\{\gamma, \delta\} \mid \langle \gamma, \delta \rangle \in S\}$ and let

$S_1 = \{\langle \gamma, \delta \rangle \in S \mid \max\{\gamma, \delta\} = \beta\}$

Let $\gamma_0 = \min\{\gamma \mid \langle \gamma, \delta \rangle \in S_1\}$; if $\gamma_0 < \beta$ then $\langle \gamma_0, \beta \rangle$ is the least element of S_1 and hence of S.

If $\gamma_0 = \beta$, set $\delta_0 = \min\{\delta \mid \langle \beta, \delta \rangle \in S_1\}$, then $\langle \beta, \delta_0 \rangle$ is the least element of S_1, and hence of S.

Now any proper initial segment $\downarrow(\langle \gamma, \delta \rangle)$ of $\omega_\alpha \times \omega_\alpha$ is contained in $\beta \times \beta$, where $\beta = (\max\{\gamma, \delta\})^+$. Then either $\beta < \omega$, in which case $\mathrm{card}(\beta \times \beta) < \aleph_0 \leq \aleph_\alpha$

or $\mathrm{card}(\beta) = \aleph_{\alpha'}$, for some $\alpha' < \alpha$, in which case

card $(\beta \times \beta) = \aleph_{\alpha'} < \aleph_\alpha$ by the induction hypothesis.

So every proper initial segment of $\omega_\alpha \times \omega_\alpha$ has order type $< \omega_\alpha$ hence the order type of $\omega_\alpha \times \omega_\alpha$ is $\leq \omega_\alpha$. But ω_α injects into $\omega_\alpha \times \omega_\alpha$ (say by $\beta \mapsto \langle \beta, \alpha \rangle$) so the order type cannot be $< \omega_\alpha$. Hence card $(\omega_\alpha \times \omega_\alpha) = \aleph_\alpha$

6·8 (9)

If $n \geq 2$ and $n \geq 2$, then $m + n \leq m \cdot n$

Proof

Let x_0, x_1 be distinct elements of a, and y_0, y_1 distinct elements of b. Define $f : a \amalg b \to a \times b$ by

$f(\langle x, 0 \rangle) = \langle x, y_0 \rangle$ for all $x \in A$

$$f(\langle y, 1 \rangle) = \begin{cases} \langle x_0, y \rangle & \text{if } y \in B \setminus \{y_0\} \\ \langle x_1, y_1 \rangle & \text{if } y = y_0 \end{cases}$$

It is easy to check that f is injective.

6·10 Corollary

For all α and β, we have

$$\aleph_\alpha + \aleph_\beta = \aleph_\alpha \cdot \aleph_\beta = \aleph_{\max\{\alpha, \beta\}}$$

Proof

Assume $\alpha \leq \beta$. Then $\aleph_\beta = 0 + \aleph_\beta \underset{6\cdot8(9)}{\leq} \aleph_\alpha + \aleph_\beta \underset{(6\cdot8(4))}{\leq} \aleph_\alpha \cdot \aleph_\beta$

$\leq \aleph_\beta \cdot \aleph_\beta = \aleph_\beta$ by 6·9

By Cantor-Bernstein, all these must be equal.

Hence, if the Axiom of Choice holds, then addition and multiplication of infinite cardinals is boring. Conversely, if $m \cdot m = m$ for all infinite m, the the Axiom of Choice holds. To prove this, we need:

6.11 Lemma

If $m+n = m \cdot n$, then either $n \leq m$, or there exists a surjection from a set of cardinality n to one of cardinality m.

Proof

Let a, b be sets of cardinality m, n respectively. Consider the composite $b \xrightarrow{j} a \amalg b \xrightarrow{f} a \times b \xrightarrow{p} a$ where f is our given bijection, j is the inclusion and p is the projection. If this is surjective, we are done. Otherwise, pick $x \in a \setminus \mathrm{im}(pfj)$. Now the restriction of f^{-1} to $\{x\} \times b$ takes values in $a \times \{a\} \subseteq a \amalg b$, so it defines an injection $b \to a$.

6.12 Proposition

Suppose $m \cdot m = m$ for all $m \geq \aleph_0$. Then the Axiom of Choice holds

Proof

We show that any set a can be well ordered. If a injects into ω this is easy; if not then $\tau(a) \not\geq \omega$, so $\mathrm{card}(a \amalg \tau(a)) \geq \aleph_0$

Hence, writing m for $\mathrm{card}\, a$, n for $\mathrm{card}\, b$, we have $(m+n) \cdot (m+n) = m+n$

But $m \cdot n = (m+0) \cdot (0+n) \leq m+n$, so $m \cdot n \leq m+n$

We also have $m+n \leq m \cdot n$ by 6.8 (9)

So by Cantor Bernstein we have $m+n = m \cdot n$, but $n \not\leq m$

since $\tau(a)$ doesn't inject into a.

So we have a surjection $g : \tau(a) \to a$. Hence we can well-order a by setting

$x < y \iff$ (the least β with $g(\beta) = x$) $<$ (the least β with $g(\beta) = y$,

6.13 Lemma

For any cardinal m, we have $m < 2^m$.

Proof

For any set a, the map $x \mapsto \{x\}$ is an injection $a \to \mathcal{P}a$

But there is no injection $a \to \mathcal{P}a$, by Cantor's Diagonal Argument.

6.14 Lemma

If $\beta \leq \alpha^+$, then $\aleph_\beta^{\aleph_\alpha} = 2^{\aleph_\alpha}$ (assuming the Axiom of Choice).

Proof

We have $2^{\aleph_\alpha} \leq \aleph_\beta^{\aleph_\alpha} \leq \aleph_{\alpha^+}^{\aleph_\alpha} \leq (2^{\aleph_\alpha})^{\aleph_\alpha} = 2^{(\aleph_\alpha \cdot \aleph_\alpha)} = 2^{\aleph_\alpha}$ (by 6.5)

(since $2 \leq \aleph_\beta$; since $\beta \leq \alpha^+$; since $2^{\aleph_\alpha} > \aleph_\alpha$)

So by Cantor-Bernstein all the $\leq$'s are $=$'s.

Hence, by assuming the Axiom of Choice, we are principally interested in the function class $F : On \to On$ defined by $\aleph_{F(\alpha)} = 2^{\aleph_\alpha}$

Cantor conjectured (the Continuum Hypothesis) that $F(0) = 1$.

The Generalised Continuum Hypothesis asserts that $F(\alpha) = \alpha + 1$ for all α, or (without assuming the Axiom of Choice) that $m \lhd 2^m$ (i.e. m is covered by 2^m) for all infinite m. In fact, the latter statement implies the Axiom of Choice.

Gödel (1930) showed that if ZF is consistent, so is ZF + GCH (+AC).

Cohen (1964) showed that if ZF is consistent, then so is

ZF + AC + ¬CH

Given a family of cardinals, $\{m_i \mid i \in I\}$, we define $\sum_{i\in I} m_i$ to be the cardinality of $\coprod_{i\in I} a_i = \bigcup\{a_i \times \{i\} \mid i \in I\}$, where each a_i has cardinality m_i, and we define $\prod_{i\in I} m_i$ to be the cardinality of $\prod_{i\in I} a_i = \{f : I \to \bigcup\{a_i \mid i \in I\} \mid f(i) \in a_i \text{ for all } i \in I\}$

(Note that it requires the Axiom of Choice to show that operations on cardinals are well defined.

6·15 Lemma (König's Lemma)

If $m_i < \lambda_i$ for all $i \in I$, then $\sum_{i\in I} m_i < \prod_{i\in I} \lambda_i$

Proof

$\leqslant$ is proved very similarly to 6·8(g). To show $\neq$, suppose a_i, b_i are sets of cardinalities m_i, λ_i respectively, and consider a map $\coprod_{i\in I} a_i \xrightarrow{f} \prod b_i$. For each i, consider the composite $a_i \xrightarrow{j_i} \coprod_{i\in I} a_i \xrightarrow{f} \prod_{i\in I} b_i \xrightarrow{p_i} b_i$; this can't be surjective since $\text{card}(a_i) < \text{card}(b_i)$, so we can find $y_i \in b_i$ not in its image. But then the map $i \mapsto y_i$ is an element of $\prod_{i\in I} b_i$ not in the image of f_{j_i} for any i, and hence not in the image of f.

6·16 Corollary

Let λ be a limit ordinal with $\text{cf}(\lambda) = \omega$, i.e. such that $\lambda = \bigcup\{\alpha_i \mid i \in \omega\}$ for some increasing sequence of ordinals $\alpha_i < \lambda$. Then $2^{\aleph_0} \neq \aleph_\lambda$

Proof

Let $\alpha_0 = \omega_0$, and, for $i \geq 1$, set $a_i = \omega_{\alpha_i} \setminus \omega_{\alpha_{i-1}}$. Then $\text{card}(a_i) = \aleph_{\alpha_i}$ for each i. Set $m_i = \aleph_{\alpha_i}$ and $n_i = \aleph_\lambda$ in König's Lemma; $m_i < n_i$ for all i, so $\sum_{i \in \omega} m_i < \prod_{i \in \omega} n_i$

But $\sum_{i \in I} m_i = \aleph_\lambda$ since there is a bijection from $\omega_\lambda = \bigcup_{i \in I} a_i$ to $\coprod_{i \in I} a_i$ and $\prod_{i \in I} a_i = \aleph_\lambda^{\aleph_0}$

However, $(2^{\aleph_0})^{\aleph_0} = 2^{(\aleph_0 \cdot \aleph_0)} = 2^{\aleph_0}$, so $2^{\aleph_0} \neq \aleph_\lambda$

Chapter 7 : Problems of Consistency and Independence

Recall : since ZF is a first order theory in a countable language, it must (if consistent) have countable models (note that the function enumerating the elements of the model cannot be definable by a function class). We might still hope that ZF is complete (i.e. for every sentence φ we have either $ZF \vdash \varphi$ or $ZF \vdash \neg\varphi$), or at least that we could complete it by adding a finite ~~set~~ number of extra axioms or axiom schemes.

There are easy independence proofs of Foundation, Infinity, Power Set and Union axioms relative to the rest of ZF (see questions 3 and 9 on Example Sheet 3).

We can also prove consistency of Foundation relative to $ZF \setminus \{\text{Foundation}\}$ by cutting down to the class of regular sets (cf. Theorem 4.4).

Gödel showed that GCH and AC are consistent relative to ZF, by 'slowing down' the von Neumann hierarchy : he defined the constructable universe L by setting

$L_0 = \emptyset$, $L_{\alpha+} = Def(L_\alpha) = \{x \subseteq L_\alpha \mid x = \{y \in L_\alpha \mid \varphi\}$ where φ is a formula with parameters in $L_\alpha\}$

$L_\lambda = \bigcup \{L_\alpha \mid \alpha < \lambda\}$ for λ a limit ordinal.

$L = \bigcup \{L_\alpha \mid \alpha \in On\}$

L does satisfy all the axioms of ZF (the Power Set axiom is the difficult one), but we can show by induction that L_α has a definable well-ordering for each α, and hence that L can be well ordered.

L also satisfies GCH.

Let $\varphi(x) = \neg Thm_T(x)$

Let $c : \mathbb{N} \to \mathbb{N}$ be the coding function, i.e. $c(n) = \ulcorner n \urcorner$

and $sub_x : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ be the function such that

$sub_x(\ulcorner \varphi \urcorner, \ulcorner t \urcorner) = \ulcorner \varphi[t/x] \urcorner$

Let $\psi(y) = \varphi(sub_x(y, c(y)))$

Let $m = \ulcorner \varphi(x) \urcorner$

Now, $\psi(m) = \varphi(sub_x(m, c(m))) = \varphi(sub_x(\ulcorner \psi(x) \urcorner, \ulcorner m \urcorner)$

$= \varphi(\ulcorner \psi(m) \urcorner) = \neg Thm_T(\psi(m))$

Clearly $T \nvdash \psi(m)$ unless T is inconsistent.

If T is ω-consistent (a slight strengthening of consistency), then $T \nvdash \neg \psi(m)$ either. We can formalise the assertion that T is consistent, as the formula $Con_T = \neg Thm_T(\ulcorner \bot \urcorner)$

We claim that $T \vdash (\psi(m) \Leftrightarrow Con_T)$, and hence $T \nvdash Con_T$ unless T is inconsistent.

We have $\vdash (\bot \Rightarrow \psi(m))$ so $PA \vdash Thm_T(\ulcorner (\bot \Rightarrow \psi(m)) \urcorner)$

so $PA \vdash (Thm_T(\ulcorner \bot \urcorner) \Rightarrow Thm_T(\ulcorner \psi(m) \urcorner))$

so $PA \vdash (\psi(m) \Rightarrow Con_T)$.

Conversely, $PA \vdash (Thm_T(\ulcorner \psi(m) \urcorner) \Rightarrow Thm_T(\ulcorner Thm_T(\ulcorner \psi(m) \urcorner) \urcorner))$

i.e. $PA \vdash (Thm_T(\ulcorner \psi(m) \urcorner) \Rightarrow Thm_T(\ulcorner \neg \psi(m) \urcorner))$

and hence formalising a propositional deduction we get

$$PA \vdash \left(Thm_T(\ulcorner \Psi(m) \urcorner) \Rightarrow Thm_T(\ulcorner \bot \urcorner) \right)$$

i.e. $$PA \vdash \left(Con_T \Rightarrow \Psi(m) \right)$$

P. Cohen developed a technique called 'forcing' which can be used to adjoin a 'generic' object to a model of set theory.
He showed that if we adjoin a generic function $F: a \to \mathcal{P}\omega$, then

a) F is injective

b) 'Cardinals are preserved' i.e. if there is no injection $b \to c$ in the original model then there is not one in the extended model.

So if $a = \mathcal{P}\mathcal{P}\omega$, then the set $F(\mathcal{P}\omega)$ is a subset of $\mathcal{P}\omega$ whose cardinality is strictly between those of ω and $\mathcal{P}\omega$.
This shows that if ZFC is consistent, then so is ZFC $\cup \{\neg CH\}$
A. Fraenkel and A. Mostowski (1920s) proved the independence of AC from a theory ZFA ('ZF with atoms'), in which Extension fails.
E. Specker modified this by replacing atoms with 'autosingletons' i.e. sets x satisfying $x = \{x\}$ (so that Extension holds but Foundation fails - see question 12, sheet 3.
However, the well founded part of any such model is the original model
Cohen observed that given a Fraenkel-Mostowski-Specker model, with a set a having no choice function, if we adjoin a generic function $a \to \mathcal{P}\omega$, we get a subset of $\mathcal{P}\omega$ with no choice function.
Now by cutting down to the well-founded part of the model, we get a model of ZF $\cup \{\neg AC\}$.

Gödel's Incompleteness Theorem

Informally, no 'sufficiently complicated' theory can be both complete and consistent. Suppose, we are given a theory T (e.g ZF) in a countable language. We can encode all formulae φ in our language by natural numbers $\ulcorner\varphi\urcorner$, and also finite sequences of formulae by natural numbers.

Write $Ded_T(m,n)$ for the assertion that m encodes a deduction of the formula coded by n, from the axioms of T.

If T is an extension of ZF, we also have a model of PA (first order Peano arithmetic) inside any model of T.

Hence we have an interpretation of PA in T,

i.e. a translation from formulae of PA to formulae of T which maps the axioms of PA to deducible sentences in T.

We write $Thm_T(y)$ for the formula $(\exists x)\, Ded_T(x,y)$

Then we have $T \vdash \varphi$ implies $PA \vdash Thm_T(\ulcorner\varphi\urcorner)$ and we also have $PA \vdash (Thm_T(\ulcorner\varphi\urcorner) \Rightarrow Thm_T(\ulcorner Thm_T(\ulcorner\varphi\urcorner)\urcorner))$

We can also formalise Modus Ponens; we have a function

$inp : \mathbb{N}\times\mathbb{N} \to \mathbb{N}$ ~~such that~~ (definable in Peano Arithmetic) such that

$inp(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner) = \ulcorner(\varphi \Rightarrow \psi)\urcorner$

and then show $PA \vdash (\forall x,y)\,(Thm_T(inp(x,y)) \Rightarrow (Thm_T(x) \Rightarrow Thm_T(y)))$