Set Theory and Logic at Part II in 2016/7

Thomas Forster

June 13, 2017

Contents

The Toad never answered a word, or budged from his seat in the road; so they went to see what was the matter with him. They found him in a sort of trance, a happy smile on his face, his eyes still fixed on the dusty wake of their destroyer. At intervals he was still heard to murmer ‘Poop-poop!’

Kenneth Graham *The Wind in The Willows*, chapter 2.

The Rubric

LOGIC AND SET THEORY (D)

24 lectures, Lent term

No specific prerequisites.

Introduction

Ordinals and cardinals

Well-orderings and order-types. Examples of countable ordinals. Uncountable ordinals and Hartogs' lemma. Induction and recursion for ordinals. Ordinal arithmetic. Cardinals; the hierarchy of alephs. Cardinal arithmetic. [5]

Posets and Zorn's lemma.

Partially ordered sets; Hasse diagrams, chains, maximal elements. Lattices and Boolean algebras. Complete and chain-complete posets; fixed-point theorems. The axiom of choice and Zorn's lemma. Applications of Zorn's lemma in mathematics. The well-ordering principle. [5]

Propositional logic

The propositional calculus. Semantic and syntactic entailment. The deduction and completeness theorems. Applications: compactness and decidability.[3]

Predicate logic

The predicate calculus with equality. Examples of first-order languages and theories. Statement of the completeness theorem; *sketch of proof*. The compactness theorem and the Löwenheim-Skolem theorems. Limitations of first-order logic. Model Theory. [5]

Set theory

Consistency

Problems of consistency and independence. [1]

Appropriate books

J.L Bell and A Slomson. Models and Ultraproducts North Holland 1969 recently reissued by Dover

B.A. Davey and H.A. Priestley Lattices and Order. Cambridge University Press 2002

T. Forster Logic, Induction and Sets. Cambridge University Press

A. Hájnal and P. Hamburger Set Theory. LMS Student Texts number 48, CUP 1999

A.G. Hamilton Logic for Mathematicians. Cambridge University Press 1988

P.T. Johnstone Notes on Logic and Set Theory. Cambridge University Press 1987

D. van Dalen Logic and Structure. Springer-Verlag 1994

1 Lecture 1: First of Five Lectures on Ordinals and Cardinals

[Lecture 1: First of Five Lectures on Ordinals and Cardinals, preceded by some introductory patter]

> (The rubric says): *Well-orderings and order-types. Examples of countable ordinals. Uncountable ordinals and Hartogs' lemma. Induction and recursion for ordinals. Ordinal arithmetic. Cardinals; the hierarchy of alephs. Cardinal arithmetic.*

(Some of the arithmetic of ordinals and cardinals cannot really be done properly until we have some set theory under our belt. OTOH it's good to at least *introduce* the students to these ideas early on in the piece. The resulting exposition is inevitably slightly disjointed.)

Warning! If you are reading this and your name isn't 'Thomas Forster' then you are eavesdropping; these notes are my messages to myself and are made available to you only on the off-chance that such availability might help you in preparing your own notes for this course. This warning doesn't mean that you shouldn't be reading this document, but you should bear it in mind anyway because i do not write out here in detail things i can do off the top of my head. The things that I write out are things that i might, in the heat of the moment, get wrong, or do in the wrong order—or forget altogether.

However it is my settled intention that by fairly early in the Lent Term these notes will be in a state fit for students to use them to revise from..

1.1 Ordinals

Cantor's discovery of a new kind of number. $1_{\mathbb{R}} \neq 1_{\mathbb{N}}$ etc etc. $1_{\mathbb{R}}$ is a multiplicative unit whereas $1_{\mathbb{N}}$ is the quantum of multiplicity ("how many?"). Brief chat about datatypes.

"How many times do i have to tell you to tidy up your room?" the answer will be an ordinal (possibly finite).

Cantor's discussion of closed sets of reals.

Ordinals measure the length of **discrete deterministic monotone processes**. (synchronous/asynchronous doesn't matter)

Well, we mean something slightly more than discrete . . . the set of stages has a total order, and it's always the case that the set of unreached stages has a first element. Monotonicity ensures that it's always clear what the situation is that you are in, and determinism-and-discreteness means that there is always an immediately-next thing to do and that you know what it is.

ω, $\omega + n$, $\omega + \omega$, $\omega \cdot n$, $\omega \cdot \omega$ ω^n.

Ordinals are also the order types of special kinds of total orders. ω is the order-type of $\langle \mathbb{N}, <_{\mathbb{N}} \rangle$. (I write structures as tuples, carrier set followed by operations). Pick 0 off the front and put it on the end, get a bigger ordinal—but the underlying set is the same size. In fact we get $\omega + 1$, which illustrates how addition corresponds to concatenation.

But to understand order types we need to put the project into a more general context: *We need to do some logic.*

Definition 1 *Congruence relations.*

Congruence relations give rise to operations on the quotient.

Cardinals are very simple. Read my countability notes on `www.dpmms.cam.ac.uk/~tf/cam_only/countability.pdf`.

Multiplication, addition and exponentiation.

1.2 Wellfoundedness

Suppose we have a carrier set with a binary relation R on it, and we want to be able to infer

$$\forall x\ \psi(x)$$

from

$$(\forall x)((\forall y)(R(y,x)\ \to \psi(y)) \to \psi(x))$$

In words, we want to be able to infer that everything is ψ from the news that you are ψ as long as all your R-predecessors are ψ. y **is an R-predecessor of** x if $R(y,x)$. Notice that there is no "case $n = 0$" clause in this more general form of induction: the premiss we are going to use implies immediately that a thing with no R-predecessors must have ψ. The expression "$(\forall y)(R(y,x) \to \psi(y))$" is called the **induction hypothesis**. The first line says that if the induction hypothesis is satisfied, then x is ψ too. Finally, the inference we are trying to draw is this: **if** x has ψ whenever the induction hypothesis is satisfied, **then** everything has ψ. When can we do this? We must try to identify some condition on R that is equivalent to the assertion that this is a legitimate inference to draw in general (i.e., for any predicate ψ).

Why should anyone want to draw such an inference? The antecedent says "x is ψ as long as all the immediate R-predecessors of x are ψ", and there are plenty of situations where we wish to be able to argue in this way. Take $R(x,y)$ to be "x is a parent of y", and then the inference from "children of blue-eyed parents have blue eyes" to "everyone has blue eyes" is an instance of the rule schematised above. As it happens, this is a case where the relation R in question does *not* satisfy the necessary condition, for it is in fact the case that children of blue-eyed parents have blue eyes and yet not everyone is blue-eyed.

To find what the magic ingredient is, let us fix the relation R that we are interested in and suppose that the inference

$$\frac{(\forall y)(R(y,x) \to \psi(y)) \to \psi(x)}{(\forall x)(\psi(x))} \qquad R\text{-induction}$$

has failed for some choice ψ of predicate. Then we will see what this tells us about R. To say that R is well-founded all we have to do is stipulate that this failure (whatever it is) cannot happen for any choice of ψ.

Let ψ be some predicate for which the inference fails.

Then the top line is true and the bottom line is false. So $\{x : \neg\psi(x)\}$ is nonempty. Let us call this set A for short. Using the top line, let x be something with no R-predecessors. Then all R-predecessors of x are ψ (vacuously!) and therefore x is ψ too. This tells us that if y is something that is not ψ, *then there must be some y' such that $R(y', y)$ and y' is not ψ either.* If there were not, y would be ψ. This tells us that the collection A of things that are not ψ "has no R-least member" in the sense that everything in that collection has an R-predecessor in that collection. That is to say

$$(\forall x \in A)(\exists y \in A)(R(y, x))$$

To ensure that R-induction can be trusted it will suffice to impose on R the condition that $(\forall x \in A)(\exists y \in A)(R(y, x))$ never hold, for any nonempty $A \subseteq dom(R)$. Accordingly, we will attach great importance to the following condition on R:

Definition 2 *R is* **well-founded** *iff for every nonempty subset A of $dom(R()$ we have $(\exists x \in A)(\forall y \in A)(\neg R(y, x))$*

(x is an "R-minimal" element of A.)

This definition comes with a health warning: it is easy to misremember. The only reliable way to remember it correctly is to rerun in your mind the discussion we have gone through: well-foundedness is precisely the magic property one needs a relation R to have if one is to be able to do induction over R. No more and no less. The definition is not *memorable*, but it is *reconstructible.*

Theorem 1 *Wellfounded induction: recursion on wellfounded relations*

Induction over a wellfounded relation is immediate. Justification of recursion requires a little thought.

Let $\langle X, R\rangle$ be a binary structure, with R wellfounded. Then the recursion

$$f(x) = G(x, \{f(x') : R(x', x)\})$$

has a unique solution as long as G is everywhere defined.

> *A niggle: why does G need to look at x? Why isn't it enough for it to look merely at $\{f(x') : R(x', x)\}$?*
>
> *A: two distinct xs might have the same R-predecessors and we want to keep open the possibility of f sending them to different things.*

Fix f. We need the concept of the transitive closure of a relation. The transitive closure of R, written '$R*$' is the $\subseteq$-least transitive relation $\supseteq R$.

However the clever idea which is specific to this proof is the concept of an **attempt**. An attempt-at-x is a function f_x which is defined at x and at every y such that $R^*(y,x)$, and obeys the recursion wherever it is defined. That is to say, if f_x is defined for all z s.t. $R(z,y)$, and it is defined at y, then we must have $f_x(y) = G(y, \{f_x(z) : R(z,y)\})$.

The concept of *attempt* is the only clever part of this proof. All that remains to be done is to choose the right thing to prove by induction. We prove by R-induction on 'x' that (i) every x has an attempt-at-x and that (ii) all attempts-at-x agree at x and at all y such that $R^*(y,x)$. Everything has been set up to make that easy.

So: suppose the induction hypothesis holds for all y s.t. $R(y,x)$.

That is to say, for every y s.t. $R(y,x)$ there is f_y, an attempt-at-y, and all attempts-at-y agree on all y' s.t. $R^*(y',y)$.

Is there an attempt-at-x? Yes. We take the union of all the f_y for $R(y,x)$ and add the ordered pair that tells us to send x to $G(x, \{f_y(y) : R(y,x)\})$.

Then the function that we are declaring by this recursion is simply the function that, for each $x \in X$, sends it to whatever-it-is that all attempts-at-x want to send x to. This function is defined everywhere and it clearly obeys the recursion.

That is to say, for any set X with a wellfounded relation R on it, and every function $G : X \times V \to V$ there is a unique f making the following diagram commute.

$$\begin{tikzcd}
X \times \mathcal{P}(X) \arrow[r, "\mathbb{1}_X \times f^*"] & X \times V \arrow[d, "G"] \\
X \arrow[u, "\mathbb{1}_X \times R"] \arrow[r, "f"] & V
\end{tikzcd}$$

Definition 3 *Wellordering a wellfounded strict total order*

"every terminal seg has a least elt" is equivalent. It's the "always an immediate next stage" condition.

Corollary 1 *Principle of induction for wellorderings*

Corollary 2 *Definition by recursion for wellorderings*

Definition 4 *Ordinals are isomorphism types of wellorderings.*

2 Lecture 2

Theorem 2

1. *Every wellordering is rigid (no nonidentity automorphisms);*
2. *If there is an isomorphism between two wellorderings $\langle A, <_A\rangle$ and $\langle B, <_B\rangle$ then it is unique;*
3. *Given two wellorderings $\langle A, <_A\rangle$ and $\langle B, <_B\rangle$ one is isomorphic to a unique initial segment of the other.*

Proof:

1. The automorphism group of a total order is torsion-free—every nontrivial cycle looks like $\mathbb{Z}$ and can have no least element. If τ is an automorphism of a wellordering consider $\{\tau^n(x) : n \in \mathbb{N}\}$. What is its least element?

2. Suppose σ and τ were two distinct isomorphisms $\langle A, <_A\rangle \to \langle B, <_B\rangle$; Then $\sigma \cdot \tau^{-1}$ would be a nontrivial automorphism of $\langle B, <_B\rangle$.

3. We define an isomorphism by recursion in the obvious way. It must exhaust either $\langle A, <_A\rangle$ or $\langle B, <_B\rangle$ and, by the earlier parts, it will be unique.

   To be slightly more formal about it, define $f : A \to B$ by the recursion $f(a) =: \sup\{f(a') : a' <_A a\}$ and $g : B \to A$ *mutatis mutandis*. We prove by wellfounded induction that $f \cdot g$ is the identity where it is defined. One of f and g must be total. If not, let a be the first thing not in the domain of f and b the first thing not in the domain of g. Then $\langle a, b\rangle$ should have been in f and $\langle b, a\rangle$ should have been in g.

We will give a slightly more detailed proof of part (3) later.

Definition 5 $\langle X, \leq_X\rangle$ **is an end-extension of** $\langle Y, \leq_Y\rangle$ *iff*
(i) $Y \subseteq X$,
(ii) $\leq_Y \subseteq \leq_X$ *and*
(iii) $(\forall y \in Y)(\forall x \in X)(x \leq y \to x \in Y)$.

Alternatively "$\langle Y, \leq_Y\rangle$ **is an initial segment of** $\langle X, \leq_X\rangle$*"*

"New stuff cannot be earlier than old stuff".

For the moment we use this only where $\langle Y, \leq_Y\rangle$ and $\langle X, \leq_X\rangle$ are wellorderings, but the idea is susceptible of generalisations to arbitrary posets and even to binary structures (models of set theory) where the binary relation ($\in$) is not even transitive. But that is for later.

Lemma 1 *Every suborder of a wellorder is isomorphic to an initial segment of it.*

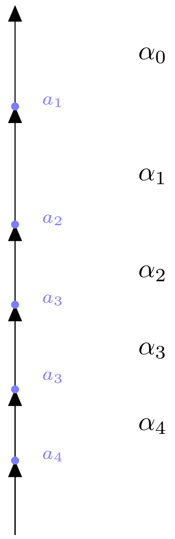
The suborder inherits totality and wellfoundedness so is a wellorder. Apply theorem 2.

Notice that this is not true of arbitrary total orders. Not every subordering of $\mathbb{Z}$ is iso to an initial segment.

Definition 6 *$\alpha \leq_{On} \beta$ if every wellordering of length β (every wellordering whose equivalence class is β) has an initial segment of length α.*

(The two ways you might define it are equivalent)

And that initial segment is unique, as we have just seen.



Theorem 3 *$<_{On}$ is wellfounded.*

Proof: Let α be an ordinal. We will show that the ordinals below α are wellfounded. The long arrow represents a wellordering $\langle A, <_A \rangle$ of length $\alpha = \alpha_0$. If (*per impossibile*[1]) there is a family $\{\alpha_i : i \in I\}$ of ordinals with no least member (and all of them $< \alpha$) then, for each $i \in I$, $\langle A, <_A \rangle$ has a (unique) proper initial segment of length α_i. For $i \in I$ let a_i be the supremum of that (unique) initial segment of $\langle A, <_A \rangle$ of length α_i. Then $\{a_i : i \in I\}$ is a subset of A with no $<_A$-least member. ■

This result is nontrivial: it's not always true that the family of isomorphism types of widgets has a widget structure. Recall linear order types without wellfoundedness; not linearly ordered.

[1] As Prof Körner points out, a well-known Swedish-Italian mathematician.

Beware! Some textbooks contain theorems with statements that sound like theorem 3 but are actually much weaker. A proof that the order relation on von Neumann ordinals is wellfounded is not a proof that $\langle On, <_{On}\rangle$ is a wellfounded any more than a check that `UBUNTU` runs properly on my laptop means that it will run safely on yours. The fact that `UBUNTU` runs safely on my laptop is not a fact about the safety of `UBUNTU` but a fact about the binary for my machine, and that says nothing about the binary for your machine.

Theorem 4 *Vital, central fact! (Cantor)*

Every ordinal is the order type of the set of ordinals below it in their natural order.

Equivalently, the order type of an initial segment of the ordinals is the least ordinal not in it.

Proof: You prove this by induction. ■

Corollary 3 *(The Burali-Forti Paradox)*☠ ☠

The collection On of all ordinals cannot be a set.

Proof:

By thm 3 $\langle On, <_{on}\rangle$ is a wellordering. Since it is downward-closed, thm 4 tells us that its order type must be the least ordinal not in it. The least ordinal that is not a ordinal? I don't need this! Beam me up, Scottie. ■

Strictly speaking we cannot correctly state and prove these last two allegations until we know what a set of ordinals is. So this is a promissory note…to be redeemed when we do some set theory. In any case one can argue that corollary 3 goes deeper than set theory. That fact that On turns out not to be a set is a consequence of the fact that we have chosen to clothe this particular mathematical spirit in set-theoretic flesh. There is something deeply weird going on, and it's not primarily a fact about set theory.

Definition 7

Preorderings are transitive and reflexive;
A preorder is a set equipped with a preordering.
A Partial ordering is an antisymmetric preordering.
Also disjoint unions, products and lexprods of posets.

Not just of posets: remember we do products of groups.

3 Lecture 3

Definition 8 *Addition and Multiplication of ordinals defined synthetically.*

Uniqueness of ordinal subtraction. What might we mean by '$\alpha - \beta$'? If $\beta \leq \alpha$ then whenever $\langle B, <_B\rangle$ belongs to β and $\langle A, <_A\rangle$ belongs to α then there *is an isomorphism $\pi : \langle B, <_B \rangle$ to a unique initial segment of $\langle A, <_A \rangle$. The truncation $\langle A \setminus \pi\text{``}B,\ <_A \restriction (A \setminus \pi\text{``}B)\rangle$ is our wellordering of length $\alpha - \beta$. This definition ensures that $\beta + (\alpha - \beta) = \alpha$.*

Explain $f\text{``}x$ notation

Part 3 of theorem 2 reassures us that ordinal subtraction is uniquely defined.

We really do need wellfoundedness here. You'd think that $\omega^* - \omega^*$ would be 0, wouldn't you? But it can be any natural number. The set of negative integers has lots of initial segments of length ω^*.

We remark without proof that it is immediate from the definitions of addition and multiplication in terms of disjoint union and lexicographic product that both operations are associative, and that multiplication distributes over addition.

We need ordinal subtraction for Cantor Normal Forms.

So now we can do induction/recursion on ordinals.

Definition 9 *cofinality; regular ordinal*

('regular' is topological jargon) You have never seen anything of cofinality $> \omega$.

Now might be a good time to attempt the example sheet exercise that says that every countable limit ordinal has cofinality ω.

Definition 10 *Recursive definition of addition, multiplication and exponentiation of ordinals.*

$\alpha + 0 = \alpha$*;* $\alpha + (\beta + 1) = (\alpha + \beta) + 1$*, and* $\alpha + sup(B) = sup(\{\alpha + \beta : \beta \in B\})$*.*

$\alpha \cdot 0 = 0$*;* $\alpha \cdot (\beta + 1) = \alpha \cdot \beta + \alpha$*, and* $\alpha \cdot sup(B) = sup(\{\alpha \cdot \beta : \beta \in B\})$*.*

$\alpha^0 = 1$*;* $\alpha^{\beta+1} = \alpha^\beta \cdot \alpha$*, and* $\alpha^{sup(B)} = sup(\{\alpha^\beta : \beta \in B\})$*.*

Remember which way round to write multiplication. Not commutative!!!

Wellorderings of length ω^ω, ϵ_0.

Definition 11 *Countable ordinal*

A countable ordinal is the order type of a wellordering of $\mathbb{N}$*.*

It's an immediate consequence of this definition, in conjunction with theorem 4, that an ordinal is countable iff there are countably many ordinals below it. This fact is too elementary to merit a label, but you need to internalise it. This absolutely must underpin your understanding of countable ordinals. Without it you would be entirely lost.

Definition 12 *Normal functions*

A total function $f : On \to On$ is **normal** *if it is total, strictly increasing and continuous.*

The range of a normal function is a **clubset** *"closed unbounded set"*

"Continuous"? It means that the following diagram commutes.

$$\begin{tikzcd}
\mathcal{P}(On) \arrow[r, "f^*"] \arrow[d, "\sup"] & \mathcal{P}(On) \arrow[d, "\sup"] \\
On \arrow[r, "f"] & On
\end{tikzcd}$$

"f^*" is a nonce notation for the function $X \mapsto f$"X. I don't expect to use it again.

Addition, multiplication and exponentiation on the Right are normal. Not on the Left!

Lemma 2 *Division Algorithm for Normal Functions.*

If $f : On \to On$ is normal, and α is any ordinal, then there is β such that $f(\beta) \leq \alpha < f(\beta + 1)$.

Proof:

The β we want is $\sup\{\beta : f(\beta) \leq \alpha\}$. What is $f(\beta)$? By normality it must be $\sup\{f(\beta) : f(\beta) \leq \alpha\}$, which is clearly $\leq \alpha$. So β is not merely the *supremum* of $\{\beta : f(\beta) \leq \alpha\}$, it is actually the *largest element* of $\{\beta : f(\beta) \leq \alpha\}$. But then $f(\beta + 1)$ must be strictly greater than α. ■

ω is a *countable ordinal.* Observe that $\omega + 1$, ω^2 and lots of other ordinals are also countable. Are *all* ordinals perhaps countable ...? No!

4 Lecture 4

Theorem 5 *Hartogs' Lemma.*

For every set X there is a wellordered set Y s.t. $Y \not\hookrightarrow X$.

Proof:

Notice that—despite Cantor's theorem—$\mathcal{P}(X)$ will not do, beco's there is no reason to suppose that it can be wellordered.

We exhibit a uniform construction of such a Y.

Consider $\mathcal{P}(X \times X)$. This is the set of all binary relations on X. We define a map $f : \mathcal{P}(X \times X) \to On$. If $R \in \mathcal{P}(X \times X)$ is a wellordering we send it to its order type, its length; if it is not a wellordering we send it to 0. The range f"$(\mathcal{P}(X \times X))$ of f is the set Y that we want.

What is the cardinality of Y? Y is naturally wellordered, so what is its order-type in this ordering? Y is *downward-closed* so, by theorem 4 its order-type is the least ordinal not in Y. The ordinals in Y are precisely the ordinals of wellorderings of subsets of X. So the order type of Y is the least ordinal not the length of a wellordering of any subset of X. So Y is not the same size as any subset of X. *It's too big.*

■

An aside about notation. Professors Leader and Johnstone write '$\gamma(X)$' for the order type of Y in the obvious ordering. This is a nice notation, but it's not standard. A standard notation is '$\aleph(|X|)$', but beware! That notation is for for the **cardinal** $|Y|$ of the Y thus obtained, *not* for the ordinal of the obvious wellordering of Y. This function is sometimes called 'Hartogs' aleph function'. Do not confuse this notation with the notation that gives subscripts to alephs: $\aleph_0$ is not $\aleph(0)$!

It's natural to ask specifically what happens if we do the construction of theorem 5 in the particular case where $X = \mathbb{N}$. The answer is that we get the set of countable ordinals, a set that Cantor called the *second number class*. We need a name for the cardinal of this set: $\aleph_1$. The supremum of the second number class is the ordinal ω_1, the least uncountable ordinal.

We take up this thread again on page **??**. .

Definition 13 *Rank functions for wellfounded (binary) structures.*

If $\langle X, R\rangle$ is a wellfounded binary structure we define:

$$\rho(x) = sup\{\rho(y) + 1 : R(y, x)\}.$$

(The intention is that $\rho(x)$ shall be the least ordinal bigger than all the $\rho(y)$ for y Related to x.)

Lemma 3 *Rank function is uniquely defined.*

Proof: By coroll 1.2.

Now would be a good moment to attack the following example sheet question:

> "Let $\langle X, R\rangle$ be a wellfounded binary structure, with rank function ρ. Prove that $(\forall x \in X)(\forall \alpha < \rho(x))(\exists y \in X)(\rho(y) = \alpha)$."

Uniqueness by recursion. Hartogs' tells you ∃ enuff ordinals.

Uniquely parsimonious. Let us say that a homomorphism $h : \langle X, R\rangle \to \langle Y, S\rangle$ between wellfounded structures is **parsimonious** if, for all $x \in X$, $h(x)$ is an S-minimal member of $\{y : (\forall x')(R(x, x') \to R(h(x'), y)\}$. But don't worry about this, co's it's not examinable.

5 Lecture 5: First of 5 lectures on Posets

(The rubric says): *Partially ordered sets; Hasse diagrams, chains, maximal elements. Lattices and Boolean algebras. Complete and chain-complete posets; fixed-point theorems. The axiom of choice and Zorn's lemma. Applications of Zorn's lemma in mathematics. The well-ordering principle.*

Loads of definitions

Definition 14 *poset you know; likewise toset; poset subsumes toset*

[pointwise and lexicographic products of posets already done]

complete poset - PTJ question about Hasse Diagrams of posets with 4 elements

directed poset. Every complete poset is directed. $\langle \mathcal{P}_{\aleph_0}(X), \subseteq \rangle$

chain-complete poset; directed-complete poset.

Lattices

distributive lattice (can define $\leq$ from $\wedge$, $\vee$ and $=$) distributive: examples and non-examples. Subspaces of a vector space[2].

Complemented lattice.
Boolean Algebra

Complete lattice. Power sets and topologies. A topology on X is a sub-poset of the complete poset $\mathcal{P}(X)$. It is also a complete poset, though it is not a sub–complete-poset, co's the $\bigvee$ and $\bigwedge$ operations are not the same in the two cases.

Aside here to explain subalgebra.

For each set X the set of topologies on X is a complete poset. This justifies such definitions as "The subset topology is the coarsest topology making the inclusion map cts"

added later

I forgot to say: The reg open sets form a complete poset that is actually a b.a.

[2] Someone in lectures asked about the meaning of distributivity. A very good question! If we have two binary operations o_1 and o_2 where o_1 distributes over o_2:

$$(\forall xyz)(o_1(x, o_2(y, z)) = o_2(o_1(x, y), o_1(x, z)))$$

what we are saying is that, for any x, the operation $y \mapsto o_1(x, y)$ is an endomorphism of the o_2 structure. Obvious when you think of it. For example, on the integers, multiplication by a fixed integer is an endomorphism of the additive structure of the integers—if it's injective it'll be a *scaling factor*.

Separative poset. A separative poset is one that is as undirected as possible. In a directed poset any two points have an upper bound. Clearly we cannot say that no two (distinct) points have an upper bound, beco's if $x \leq y$ then anything $\geq y$ is an upper bound for both. What we can say is that if x is **not** $\leq y$ *then there is $y' \geq y$ s.t. x and y' have no upper bound. Thus we say:*

$\langle X, \leq \rangle$ is separative iff

$$(\forall x, y \in X)(x \not\leq y \to (\exists z \geq y)(\forall w)(w \not\geq z \vee w \not\geq x))$$

There are lots of separative posets and they matter.

6 Lecture 6: Fixed-point theorems

THEOREM 6 *Tarski-Knaster*

Let $\langle X, \leq \rangle$ be a complete lattice and f an order-preserving map $\langle X, \leq \rangle \to \langle X, \leq \rangle$. Then f has a fixed point.

Proof: Set $A = \{x : f(x) \leq x\}$ and $a = \bigwedge A$. (A is nonempty because it must contain $\bigvee X$.)

That's the only part of the proof you need to *remember*, co's you can work the rest of it out from the definition of a.

But, for the sake of completeness, we continue ...

Since f is order-preserving, we certainly have $f(x) \leq x \to f^2(x) \leq f(x)$, and so $f(a)$ is also a lower bound for A as follows. Let $x \in A$ be arbitrary; we have $f(x) \leq x$, whence $f^2(x) \leq f(x)$, so $f(x) \in A$ and $a \leq f(x)$.

$$f(a) \leq^{(1)} f^2(x) \leq^{(2)} f(x) \leq^{(3)} x$$

(1) holds beco's $a \leq f(x)$ (as we've just showed) and f is order-preserving;

(2) holds beco's $f(x) \leq x$ and f is order-preserving;

(3) holds beco's $x \in A$.

...giving $f(a) \leq x$ as desired. But a was the *greatest* lower bound, so $f(a) \leq a$ and $a \in A$. But then $f(a) \in A$ since $f``A \subseteq A$, and $f(a) \geq a$ since a is the greatest lower bound. ■

Observe that this a is not only a fixed point, it is the *least* fixed point.

Observe further that if $\langle X, \leq \rangle$ is a complete poset then so too, for any $a \in X$, is $\{x \in X : x \geq a\}$ equipped with the restriction of $\leq$, and it has the same sup and inf operations. It is a genuine sub–complete-poset. This has the immediate consequence that

I haven't stated this correctly

COROLLARY 4 *Let $\langle X, \leq \rangle$ be a complete lattice; let a be a member of X and let f an order-preserving map $\langle X, \leq \rangle \to \langle X, \leq \rangle$. Then f has a fixed point $\geq a$.*

and this in turn has the further corollary

Corollary 5 *Let $\langle X, \leq\rangle$ be a complete lattice and f an order-preserving map $\langle X, \leq\rangle \to \langle X, \leq\rangle$. Then f has a complete poset of fixed points.*

Proof:
We need to show that every set of fixed points for f has a sup. So let A be a set of fixed points for f. Clearly A has a sup $\bigvee A$ beco's $\langle X, \leq\rangle$ is a complete lattice. Is this the thing we want? The obvious thing to do is to try to prove that it is a fixed point. You will fail! However, all is not lost, because you use corollary 4 to show that there is a least fixed point above $\bigvee A$, and that fixed point is the one we want. ■

There is an echo here of the fuss i was making last time about how the complete poset of open sets in a topology on a set X is not a sub–complete-poset of the power set of X.

This proof of theorem 6 shows not only that order-preserving functions have fixed points but that they have *least* fixed points. This gives us the existence of inductively defined sets because the operation of taking a set and adding to it the result of applying all the constructors once to all its members is order-preserving (with respect to $\subseteq$). The above definition of the element a echoes precisely the declaration of $\mathbb{N}$ as an intersection of a family of sets. Compare

$$\bigwedge\{x : f(x) \leq x\} \text{ with } \bigcap\{X : (S``X \cup \{0\}) \subseteq X\}.$$

There are three things you might worry about here:

(i) is $\{X : (S``X \cup \{0\}) \subseteq X\}$ a set? Co's, if not, it isn't there for us to take $\bigcap$ of it;

(ii) if we want to use theorem 6 to deduce the existence of $\mathbb{N}$ then we seem to be using T-K on the complete poset of the power set of the set of cardinals, and is *that* a set?;

(iii) what is $S(\alpha)$ when α is a cardinal about which we know nothing?

(i) and (ii) you are not to worry about for the moment. These are set-theoretic issues which we will sort out later.

The answer to (iii) is that actually you know this already: $S(\alpha)$ is just $\alpha+1$ which is $|x \cup \{y\}|$ whenever $|x| = \alpha$ and $y \notin x$. When we come to the axiom of choice we shall see that typically $S(\alpha) = \alpha$ for infinite α.

Corollary 6 *Cantor-Bernstein*

There are lower-tech proofs of CB that do not involve assuming that $\mathcal{P}(A)$ is a set, and you will find them in the older textbooks but they are fiddly.

Other applications include Banach-Tarski.

There are other fixed-point theorems of this flavour "a slick function from a nice poset into itself has lots of fixed points" (for example "every normal function from On to On has a fixed point") and we will deal with them as they come up, not all together.

6.1 The Axiom of Choice

AC = axiom of Choice; ZL is Zorn's Lemma[3] WO is the Wellordering Principle every set can be wellordered.

Remark 1 *WO implies AC.*

Proof: Suppose you can wellorder anything that is shown to you, and you want a choice function on a family X of nonempty sets. You wellorder $\bigcup X$ by some wellorder which you call '$<$' and then, for each $x \in X$, declare $f(x)$ to be the $<$-least element of x. ■

Remark 2 *ZL implies WO.*

Proof: You are given a set X and you want to wellorder it. Your weapon is ZL, which means that whenever you have a chain-complete poset, it will have a maximal element. How do you use ZL? Well, you seek a chain-complete poset such that a maximal element of it is a wellordering of X. How about taking your chain-complete poset to be the poset of wellorderings of subsets of X (thought of as subsets of $X \times X$) ordered by $\subseteq$? Not *quite*. The problem is that a union of a chain of wellorderings under $\subseteq$ might not be a wellordering. You need to partially order the wellorderings by **end-extension**. (Recall definition 5.) ■

7 Lecture 7

Remark 3 *AC implies WO*

Again we have a matching challenge. We want to wellorder a set X and we are told we can have a choice function on any family of nonempty sets that we like. The obvious suspect is $\mathcal{P}(X) \setminus \{\emptyset\}$. We now define, by recursion on the ordinals, a sequence s of elements of X indexed by ordinals. By AC, the family $\mathcal{P}(X) \setminus \{\emptyset\}$ of nonempty sets has a choice function f. Then we declare $s(\alpha)$, the αth member of our sequence, to be $f(X \setminus \{s(\beta) : \beta < \alpha\})$.

How can we be sure that we do not run out of ordinals? Hartogs' lemma (theorem 5) tells us that there is a wellordering too big to be embedded in X. So we must have used up all of X by the time we reach the order type of any such wellordering. ■

Definition 15
A function $f : \langle X, \leq_X \rangle \to \langle X, \leq_X \rangle$ *is* **inflationary** *if* $(\forall x \in X)(x \leq_X f(x))$.

Inflationary is NOT the same as increasing!

For AC → ZL we need

[3] What is yellow and equivalent to the axiom of choice?

Theorem 7 *The Bourbaki-Witt theorem*

Every inflationary function from a chain-complete poset into itself has a fixed point.

Proof: Let $\langle X, \leq_X \rangle$ be a chain complete poset, and let $f : X \to X$ be inflationary. The idea is to build a chain, starting at some (any) $x \in X$, extend it at successor stages by doing f to the latest element obtained, and at limit stages by taking sups—$\langle X, \leq_X \rangle$ is chain complete. If we reach a fixed point at any stage we have our hearts' desire. But Hartogs' lemma (theorem 5) tells us that we cannot run out of ordinals.

■

Corollary 7 *AC implies ZL*

Proof: Now let $\langle X, \leq_X \rangle$ be a chain-complete poset. By AC we have a choice function f on $\mathcal{P}(X) \setminus \{\emptyset\}$. The function

$$x \mapsto \texttt{if } x \texttt{ is } \leq_X\text{-maximal } \texttt{then } x \texttt{ else } f(\{x' \in X : x <_X x'\})$$

is inflationary and must have a fixed point by theorem 7. That fixed point will be maximal by construction. ■

7.0.1 Weak versions: countable choice, and a classic application thereof; DC

ctbl U of ctbls is ctbl. Do the same with DC

[not being written up for the notes: i can do this in my sleep]

A Dedekind-Infinite set is one the same size as some proper subset of itself. If countable choice fails there may be infinite sets that are not Dedekind-infinite.

König's Lemma

7.0.2 Applications of Zorn's Lemma

We look for chain-complete posets.

Comparability of cardinals

Independent sets in a vector space

Filters in a boolean algebra. Filters? Wossat?!

Definition 16 *Filters and ideals in boolean algebras*

*A filter in a boolean algebra is a subset closed under $\wedge$ and $\geq$; it's a collection of "big" elements. The dual concept is *Ideal*: closed under $\vee$ and $\leq$. A filter is proper iff it does not contain $\bot$. $\subseteq -$maximal proper filters are called "ultrafilters". and maximal ideals ("prime ideals")*

Filters in a b.a. form a poset under $\subseteq$.

8 Lecture 8: Boolean Algebras Continued

Yer typical boolean algebra is a power set algebra, which is to say a product of lots of copies of the two-element boolean algebra, commonly written '$\mathbb{2}$'. yes/no. Hence the connection to logic.

Set of proper filters chain-complete. In fact it's directed-complete.
Arbitrary intersection of a nonempty family of proper filters is a proper filter.

Dual to a filter is an ideal. Called ideals because they are ideals in boolean rings. A boolean algebra becomes a ring if we take $\times$ to be $\wedge$ and $+$ to be `XOR`: $x + y = (x \cap \overline{y}) \vee (y \cap \overline{x})$

Definition 17 *Principal and nonprincipal ideals and filters.*

Boolean homomorphism h must send $a \wedge b$ to $h(a) \wedge h(b)$ and so on for all the other operations.

Ideals are kernels of boolean algebra homomorphisms .

A principal ideal is a ba in its own right. The ideal generated by x is the kernel of the homomorphism $y \mapsto y \wedge \overline{x}$.

Definition 18 *Quotient over an ideal, or filter*
$x \sim_I y$ if $(x \ \texttt{XOR}\ y) \in I$;
or
$x \sim_F y$ if $(x \wedge y) \vee (\overline{x} \wedge \overline{y}) \in F$.

(Consider the ideal of finite sets in $\mathcal{P}(\mathbb{N})$, and the quotient algebra.

This may remind you of a puzzle about Dons and hats and only finitely many of them getting it wrong. You will need the axiom of choice.)

General patter about representation theorems.

Theorem 8 *Stone's representation theorem*
Every Boolean algebra is isomorphic to one where the order relation is $\subseteq$— set inclusion; $\wedge$ is $\cap$; $\vee$ is $\cup$, and complementation is set complementation.

Proof:
The challenge is to associate to each element of the algebra a set in such a way that elements higher in the algebra get sent to bigger sets (more elements). (Of course we also have to respect the boolean operations $\wedge$, $\vee$ and complement). In principle these sets could be anything, but in fact we don't have to look very far from home. It turns out that we can send each element b to the set of maximal filters F such that $b \in F$. The higher up in the algebra you are the more filters you belong to. Clearly if $a \leq b$ then any maximal filter containing a contains b. It remains to check that the inclusion is strict. That is ...

We need a lemma that says that

Lemma 4 *If $b \not\geq a$ then there is a maximal filter containing a but not b.*

Proof:

Consider the collection of those filters that are supersets of the principal filter generated by $a \wedge \overline{b}$. This is a chain-complete poset and must have a maximal element by ZL.

It's easy to check that the boolean operations are respected. Any maximal filter must contain either b or $\overline{b}$ so that ensures that complementation is respected. [It also explains why we need *maximal* [proper] filters not just any-old-proper-filters.]

■

8.1 Reduced products

Suppose $\{A_i : i \in I\}$ is a family of algebras. (Think *groups*, for the sake of concreteness).

We define the operations on the product $\prod_{i \in I} A_i$ in the usual way (pointwise). So that, for an operation @ and for $f, g \in \prod_{i \in I} A_i$, we declare $f@h$ to be the function $i \mapsto f(i)@g(i)$.

Definition 19
Let $F \subseteq \mathcal{P}(I)$ be a filter on I.

For f, g in $\{A_i : i \in I\}$ say $f \sim_F g$ iff $\{i \in I : f(i) = g(i)\} \in F$.

We had better check that $\sim_F$ is a congruence relation for all the operations defined pointwise-style on the product. Let @ be such a (binary) operation, written infix.

Suppose $f \sim_F f'$ and $g \sim_F g'$. Then $f@g$ is $i \mapsto f(i)@g(i)$ and $f'@g'$ is $i \mapsto f'(i)@g'(i)$. Now $\{i \in I : f'(i)@g'(i) = f(i)@g(i)\}$ is (a superset of) $\{i \in I : f(i) = f'(i)\} \cap \{i : g(i) = g'(i)\}$ and both these sets are in the filter, whence $f@g \sim_F f'@g'$ as desired.

This extends to operations of higher arity as long as the arity is finite.

The quotient is notated $\prod_{i \in I} A_i/F$.

Thus we have shown

Remark 4 *If @ is an operation defined at each factor A_i and thence defined pointwise on the product, then $\sim_F$ is a congruence relation for it and it is defined on the quotient.*

Thus if all the factors have wombat-structure then the reduced product has too.

9 Lecture 9: First of Five Lectures on Propositional Logic

The propositional calculus. Semantic and syntactic entailment. The deduction and completeness theorems. Applications: compactness and decidability.

The letters point to things that evaluate to `true` and `false`. I am going to try to remember to reserve the symbols '$\top$' and—even more important—'$\bot$' for propositional constants NOT truth-values.

The language. Propositional letters (aka *literals*): p, q, r ... or (better!) p, p', p'' ..., so that the set of literals forms a regular language. NB: the internal structure of the literals given by the prime symbol is not going to be visible to the semantics for the logic. 'p'''' is a single symbol not a string of four.

We introduce '$\bot$' as a constant symbol in propositional logic. Beware overloading.

Set of letters is a regular language.
Set of wffs is context-free

Truth-functionality. Valuations and truth-tables.
Interdefinability of connectives.
Intension and extension. Now we can talk about $\to$.

9.1 if-then

(This section was lectured in only the most cursory manner, largely because it's arguably not examinable. I have included a longer treatment here beco's it may be of interest to some students.)

Lots of students dislike the truth-functional conditional as an account of implication. The usual cause of this unease is that in some cases $p \to q$ evaluates to `true` for what seem to them to be spurious and thoroughly unsatisfactory reasons, namely: that p is false, or that q is true. How can q follow from p merely because q happens to be true? The meaning of p might have no bearing on q whatever! Standard illustrations in the literature include

> If Julius Cæsar is Emperor then sea water is salt.

These example seem odd because we feel that to decide whether or not p implies q we need to know a lot more than the truth-values of p and q.

This unease shows that we have forgotten that we were supposed to be examining a relation between *extensions*, and have carelessly returned to our original endeavour of trying to understand implication between *intensions*. $\wedge$ and $\vee$, too, are relations between intensions but they also make sense applied to extensions.

Now if p implies q, what does this tell us about what p and q evaluate to? Well, at the very least, it tells us that p cannot evaluate to `true` when q evaluates to `false`. That is to say that we require—at the very least— that the *extension* corresponding to a conditional should satisfy *modus ponens*.

How many extensions are there that satisfy *modus ponens*? For a connective C to satisfy *modus ponens* it suffices that in each of the two rows of the truth table for C where p is true, if $p\,C\,q$ is true in that row then q is true too.

p	C	q
1	?	1
0	?	1
1	0	0
0	?	0

We cannot make $p\,C\,q$ true in the third row, because that would cause C to disobey *modus ponens*, but it doesn't matter what we put in the centre column in the three other rows. This leaves eight possibilities:

$$(1):\ \frac{p \quad q}{q}$$

p	C^1	q
1	1	1
1	0	0
0	1	1
0	0	0

$$(2):\ \frac{p \quad p \longleftrightarrow q}{q}$$

p	C^2	q
1	1	1
1	0	0
0	0	1
0	1	0

$$(3):\ \frac{p \quad \neg p}{q}$$

p	C^3	q
1	0	1
1	0	0
0	1	1
0	1	0

$$(4):\ \frac{p \quad p \to q}{q}$$

p	C^4	q
1	1	1
1	0	0
0	1	1
0	1	0

$$(5):\ \frac{p \quad \bot}{q}$$

p	C^5	q
1	0	1
1	0	0
0	0	1
0	0	0

$$(6):\ \frac{p \quad p \wedge q}{q}$$

p	C^6	q
1	1	1
1	0	0
0	0	1
0	0	0

$$(7):\ \frac{p \quad \neg p \wedge q}{q}$$

p	C^7	q
1	0	1
1	0	0
0	1	1
0	0	0

$$(8):\ \frac{p \quad \neg p \wedge \neg q}{q}$$

p	C^8	q
1	0	1
1	0	0
0	0	1
0	1	0

obtained by replacing the major premiss '$p \to q$' in the rule of *modus ponens* by each of the eight extensional binary connectives that satisfy the rule.

(1) will never tell us anything we didn't know before; we can never use (5) because its major premiss is never true; (6) is a poor substitute for the rule of $\wedge$-elimination; (3), (7) and (8) we will never be able to use if our premisses are consistent.

(2), (4) and (6) are the only sensible rules left. (2) is not what we are after because it is symmetrical in p and q whereas "if p then q" is not. The advantage of (4) is that you can use it whenever you can use (2) or (6). So it's more use!

We had better check that we do not get into trouble with this policy of adopting (4), and evaluating $p \to q$ to `true` unless there is a very good reason not to. Fortunately, in cases where the conditional is evaluated to `true` *merely* for spurious reasons, then no harm can be done by accepting that evaluation. For consider: if it is evaluated to `true` *merely* because p evaluates to `false`, then we are never going to be able to invoke it (as a major premiss at least), and if it is evaluated to `true` *merely* because q evaluates to `true`, then if we invoke it as a major premiss, the only thing we can conclude—namely q—is something we knew anyway.

This last paragraph is not intended to be a *justification* of our policy of using only the material conditional: it is merely intended to make it look less unnatural than it otherwise might. The astute reader who spotted that nothing was said there about conditionals as *minor* premisses to modus ponens should not complain. They may wish to ponder the reason for this omission.

The idea is that we can use this strictly truth-functional stuff to codify arguments that only involve *and, or* and *not*, and don't involve *all* or *some.* The following example is from Kalish and Montague. It's a bit contrived but you get the idea.

> If God exists then He is omnipotent.
> If God exists then He is omniscient.
> If God exists then He is benevolent.
> If God can prevent evil then—if He knows that evil exists—then He is not benevolent if He does not prevent it.
> If God is omnipotent, then He can prevent evil.
> If God is omniscient then He knows that evil exists if it does indeed exist.
> Evil does not exist if God prevents it.
> Evil exists.
> ____________________
> *God does not exist.*

Here are the basic propositions and the letters we are going to abbreviate them to.

God exists	E
God is omnipotent	P
God is omniscient	O
God is benevolent	B
God can prevent Evil	D
God knows that Evil exists	K
God prevents Evil	J
Evil exists	V

If God exists then He is omnipotent. $E \to P$ (1)

If God exists then He is omniscient. $E \to O$ (2)

If God exists then He is benevolent. $E \to B$ (3)

If God can prevent Evil then—if He knows that Evil exists—then He is not benevolent if He does not prevent it. $D \to (K \to (\neg J \to \neg B))$ (4)

If God is omnipotent, He can prevent Evil. $P \to D$ (5)

If God is omniscient then He knows that Evil exists if it does indeed exist. $O \to (V \to K)$ (6)

Evil does not exist if God prevents it. $J \to \neg V$ (7)

Evil exists. V (8)

We want to persuade ourselves that God does not exist. Well, suppose he does. Let's deduce a contradiction

Assume E. Then (1), (2) and (3) give us

P (9),

O (10)

and

B (11)

Now that we know O, (7) tells us that

$V \to K$ (12)

But we know V (that was (8)) so we know

K (13)

We know P, so (5) tells us that

D (14)

We can feed D into (4) and infer

$K \to (\neg J \to \neg B)$ (15)

But we know K (that was line 13) so we get

$\neg J \to \neg B$ (16)

(8) and (7) together tell us $\neg J$, so we get $\neg B$. But we got B at line 11.

Definition 20 *Recursive definition of satisfaction*

A valuation is a function from literals to truth-values. We define what it is for a valuation to satisfy a compound formula by recursion on the subformula relation which (you will have noticed) is wellfounded.

We introduce $\bot$ as a constant symbol in propositional logic. Beware overloading.

v sat l [l a literal] iff $v(l) =$ `true`*;*
v sat $\phi \wedge \psi$ iff (v sat ψ and v sat ψ)
v sat $\phi \vee \psi$ iff (v sat ψ or v sat ψ)
v sat $\phi \to \psi$ iff (either not(v sat ψ) or v sat ψ)
v sat $\neg\phi$ iff not(v sat ϕ).
We say $\phi \models \psi$ iff every valuation that sat ϕ also sat ψ.

Semantic entailment and validity
"true under all valuations"; "tautology"

Logical equivalence: two formulæ are logically equivalent iff they are satisfied by the same valuations.

Definition 21
A theory is a set of sentences, closed under some notion of decidibility clear from context.

A Logic is a theory closed under substitution. A Logic that contains (to take a pertinant example) $A \to (B \to A)$ must contain all substitution instances of it, such as: $p \to (p \to p)$, or $(p \vee q) \to ((q \vee r) \to (p \vee q))$

Here is an example of a propositional theory. We might call it the theory of adding two eight-bit words (with overflow). It has 24 propositional letters, p_0 to p_7, p_8 to p_{15} and p_{16} to p_{23}, and axioms to say that p_{16} to p_{23} represent the output of an addition if p_0 to p_7 and p_8 to p_{15} represent two words of input. `true` is 1 and `false` is 0, so it contains things like $((p_0 \wedge p_8) \to \neg p_{16})$ (because an odd plus an odd is an even!).

$$\begin{array}{c|cccccccc} & p_7 & p_6 & p_5 & p_4 & p_3 & p_2 & p_1 & p_0 \\ + & p_{15} & p_{14} & p_{13} & p_{12} & p_{11} & p_{10} & p_9 & p_8 \\ \hline = & p_{23} & p_{22} & p_{21} & p_{20} & p_{19} & p_{18} & p_{17} & p_{16} \end{array}$$

Notice that this is a theory not a logic, co's it's not closed under substitution. It contains $(p_0 \wedge p_8) \to \neg p_{16}$ but not (for example) $(p_1 \wedge p_9) \to \neg p_{17}$ which is obtained from it by the substitution: $p_0 \mapsto p_1$, $p_8 \mapsto p_9$ and $p_{16} \mapsto p_{17}$.

Definition 22 *Any set T (a theory or a Logic) of axioms-and-rules-of-inference gives rise to a deducibility relation written '$\vdash$': "$T \vdash \phi$" or (sometimes) '$\psi \vdash_T \phi$" to mean that ϕ can be deduced from ψ using the T-machinery.*

Theories and Logics usually (tho' not always) arise from a set of axioms and a set of rules of inference. Thus, considered as sets of formulæ they are what we call *recursively enumerable*. We say they are *axiomatised*.

The logic consisting of all valid formulæ of propositional logic (all tautologies) does not on the face of it arise in this way. It is a nontrivial fact that by a judicious choice of theory (Logic) we can get $\models$ and $\vdash$ to coincide. Particular set of axioms-and-rules doesn't matter; what matters is that it can be done ...*proof of concept*

10 Lecture 10

We'll have a (very brief, co's it's not really examinable) look at some alternative rules of inference, so that we have some idea of the generality of a propositional theory.

Natural deduction!

Brief chat about completeness theorems. Kuratowski's theorem about planar graphs.

Can you detect semantic validity just by looking at the syntax, without looking at the models? Talk about the biconditional fragment.

Remark 5 *Completeness for the Biconditional Fragment*

A formula in the language with only $\longleftrightarrow$ and $\neg$ is valid (satisfied by all valuations) iff every literal that appears at all appears an even number of times. These two conditions are equivalent to be derivable from the three axiom schemes (all substitution instances of) $A \longleftrightarrow A$; $(A \longleftrightarrow B) \longleftrightarrow (B \longleftrightarrow A)$ and $(A \longleftrightarrow (B \longleftrightarrow C)) \longleftrightarrow ((A \longleftrightarrow B) \longleftrightarrow C)$. And your sole rule of inference is modus ponens.

Whatever your axioms and rules of inference are, it's going to be pretty easy to show that $\Gamma \vdash \phi$ implies $\Gamma \models \phi$; it's the other direction that is hard. In the biconditional logic case it's easy to see that anything deduced from the three axiom schemes by modus ponens has an even number blah. It's the other direction that is hard. [There is an extension of this to the logic with negation as well, but i can't remember what the axiom for $\neg$ is: it may be $\neg(p \longleftrightarrow \neg p)$. You may like to check.]

If we are going to prove that $\models$ and $\vdash$ coincide, we'd better have precise mathematical definitions of them. We know what $\models$ is. So we need to be clear about $\vdash$.

We also need to be crystal-clear about what a proof is.

Because we are short of time i am going to use an axiomatisation-with-rule-of-inference kit that gives a slick proof of completeness. I have in fact shamelessly lifted it from PTJ's book ch 2.

Brief chat about **Interdefinability of connectives classically** (could've been done earlier) We don't exploit interdefinability in natural deduction.

We have three axiom schemes, K, S and T plus *modus ponens*. Or three axioms plus a rule of substitution plus *modus ponens*.

Definition 23 *K, S and T; Hilbert-style proof*

A singleton list containing an axiom is a proof. What about the empty list?

T is the characteristic axiom for **classical Logic**: double negation and law of excluded middle. Not everybody likes these two axioms, so it's nice to have an axiomatisation which lists them separately so they can be dropped if we want.

K and S enable us to prove the "deduction theorem"...

First, a notational innovation you will have to get used to: people often write 'L, A' for '$L \cup \{A\}$'.

Definition 24 *The* **deduction theorem** *for a logic L is the assertion*

$$\textit{if } L, A \vdash B, \textit{ then } L \vdash A \to B.$$

The converse is trivial as long as L has *modus ponens*.

Theorem 9 *The deduction theorem holds for L iff L contains (all substitution instances of) K and S.*

Proof:

$L \to R$ The left-to-right direction is easy, for we can use the deduction theorem to construct proofs of K and S. This we do as follows:

$$L \vdash (A \to (B \to C)) \to ((A \to B) \to (A \to C))$$

(which is what we want) holds iff (by the deduction theorem)

$$L \cup \{(A \to (B \to C))\} \vdash ((A \to B) \to (A \to C))$$

iff (by the deduction theorem)

$$L \cup \{(A \to (B \to C)), (A \to B)\} \vdash (A \to C)$$

iff (by the deduction theorem)

$$L \cup \{(A \to (B \to C)), (A \to B), A\} \vdash C.$$

But this last one we can certainly do, since

$$[(A \to (B \to C));\ (A \to B);\ A;\ (B \to C);\ B;\ C]$$

is a Hilbert-proof of C from $L \cup \{(A \to (B \to C)), (A \to B), A\}$ (and we have already seen how to do this by natural deduction).

We also want $L \vdash A \to (B \to A)$. This holds (by the deduction theorem) iff $L \cup \{A\} \vdash (B \to A)$ iff (by the deduction theorem again) $L \cup \{A, B\} \vdash A$.

$R \to L$ Suppose $L, A \vdash B$. That is to say, there is a (Hilbert) proof of B in which A is allowed as an extra axiom. Let the ith member of this list be B_i. We prove by induction on i that $L \vdash A \to B_i$. $B_i \to (A \to B_i)$ is always a (substitution instance of) an axiom (because of K), so if B_i is an axiom, we have $L \vdash A \to B_i$ by *modus ponens*. If B_i is A, this follows because $L \vdash A \to A$. If B_i is obtained by *modus ponens* from two earlier things in the list, say B_j and $B_j \to B_i$ then, by induction hypothesis, we have $L \vdash A \to B_j$ and $L \vdash A \to (B_j \to B_i)$. But by S this second formula gives us $L \vdash (A \to B_j) \to (A \to B_i)$ and then $L \vdash A \to B_i$ by *modus ponens*.

■

11 Lecture 11

From now on we are going to assume that our only rules of inference are *modus ponens* and substitution. Thus when we write "$\Gamma \vdash \phi$" we mean that if we add to Γ all substitution-instances of K, S and T, and close under *modus ponens* then we can find a Hilbert-style proof of ϕ.

Theorem 10 *The Adequacy Theorem*
Let Γ be a set of expressions in a propositional language.
If $\Gamma \models \bot$ then $\Gamma \vdash \bot$.

Proof:
We prove the contrapositive. Suppose $\Gamma \not\vdash \bot$. We propose to infer $\Gamma \not\models \bot$. "Contrapositive"? $\Gamma \models \bot$ says that any valuation that satisfies Γ satisfies $\bot$, but of course no valuation satisfies $\bot$, so $S \models \bot$ says that no valuation satisfies Γ. So the challenge is to find a valuation that satisfies Γ, given that we cannot deduce $\bot$ from Γ.

The idea is to construct a sequence $\Gamma = \Gamma_0, \Gamma_1, \Gamma_2 \ldots$ s.t. $\Gamma_i \not\vdash \bot$ for each i, and such that $\Gamma_\omega = \bigcup_{i<\omega} S_i$ "decides" every formula... by which we mean that, for each formula ϕ, either $\Gamma_\omega \vdash \phi$ or $\Gamma_\omega \vdash \neg\phi$. To do this we enumerate the expressions of the language in order type ω as $\langle t_i : i \in \mathbb{N} \rangle$.

(If you have done your 1a revision exercises you will be aware that the set of expressions is countable. This is because the set of finite sequences from a countable set is countable. You can prove this using the prime powers trick.)

Given Γ_i we obtain Γ_{i+1} by asking whether or not $\Gamma_i \cup \{t_i\} \vdash \bot$. If $\Gamma_i \cup \{t_i\} \not\vdash \bot$ then set $\Gamma_{i+1} = \Gamma_i \cup \{t_{i+1}\}$, or $\Gamma_{i+1} = \Gamma_i \cup \{t_{i+1}\}$ o/w.

The valuation we want is now the valuation that sends every literal in the deductive closure of $\bigcup_{i \in \mathbb{N}} \Gamma_i$ to `true` and sends all others to `false`. ■

We can now obtain the completeness theorem as a corollary. We prove only the hard direction.

Corollary 8 *The Completeness theorem for Propositional Logic.*
If $\Gamma \models \phi$ then $\Gamma \vdash \phi$.

Proof:

Suppose $\Gamma \models \phi$. Then we must have $\Gamma \cup \{\neg\phi\} \models \bot$. (No valuation can satisfy both ϕ *and* $\neg\phi$ even if it tries with both hands). Then, by theorem 10, the Adequacy Theorem, we have $\Gamma \cup \{\neg\phi\} \vdash \bot$.

Now, by the Deduction Theorem we have $\Gamma \vdash \neg\phi \to \bot$. Axiom T now allows us to infer ϕ. ■

The combination of axioms and rule of inference used here was chosen precisely to expedite this particular proof of completeness: K and S give you the deduction theorem, and axiom T provides the final step. Other combinations will give different proofs. There are presentations of propositional logic that are more natural but they make the completeness theorem much harder to prove.

We obtain as a corollary the compactness theorem.

Consider a propositional language with a countable infinity of literals. We can topologise the set of all valuations by declaring, for each finite set x of pairs $\langle l, t\rangle$ where l is a literal and t is a truth value, that the set $\{v : x \subset v\}$ (thinking of valuations as sets of ordered pairs) is a basis element.

It turns out that this topology is compact: it's the product of copies of the two-point space (which is compact).

Clearly, proofs being finite objects, if there is a proof of ϕ from Γ, then there is a proof that uses only finitely many formulæ in Γ. But, by corollary 8 (which tells us that $\vdash$ and $\models$ are the same relation) it then follows that if $\Gamma \models \phi$ then $\Gamma' \models \phi$ for some finite subset $\Gamma' \subseteq \Gamma$. We'd better give this a name and a number:

Corollary 9 *The Compactness Theorem (for propositional Logic)*
If $\Gamma \models \phi$ then there is $\phi' \subseteq \Gamma$, Γ' finite, with $\Gamma' \models \phi$.

One consequence of the completeness theorem for propositional logic is that both "ϕ is a tautology" and "ϕ is not a tautology" become what one might call *existential* sentences. "ϕ is a tautology" becomes "there is a p s.t. p is a proof of ϕ" and "ϕ is not a tautology" becomes "there is a valuation that refutes ϕ".

This two-pronged attack looks useful if we are looking for efficient engines that answer whether or not a propositional formula is a tautology. Clearly we have a deterministic algorithm that runs in time exponential in the number of distinct propositional letters in ϕ: simply examine all valuations. Clearly any valuation that refutes ϕ can be shown to do so in time linear in the length of ϕ. Thus non-tautologousness is what they call *nondeterministic polynomial*. What about tautologousness? ϕ is tautologous iff there is a proof of it. But can a correct proof be verified in time polynomial in the length of ϕ? The question is: "is there a system of rules and axioms with the feature that there is a polynomial P in one variable such that every tautology of length n has a proof in that system of length less than $P(n)$?". Curiously this question seems to be open.

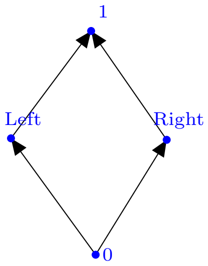
It is still not known whether or not there is a polynomial-time test for tautologousness.

Can detect tautologousness by exhaustive search co's only finitely many cases. Spuriously easy. So consider the truth value algebra $\mathcal{P}(\mathbb{N})$.

11.1 Boolean algebras detect propositional tautologies

What do we mean by this title? Suppose you try building truth tables using a boolean algebra B instead of $\mathbb{2}$. Each row of such a truth-table corresponds to what one might call a *B-valuation*—a thing like an ordinary valuation except that it takes values in B. You'll end up with $|B|^n$ rows (assuming your formula has n distinct letters in it) instead of 2^n rows, so it's not the kind of thing you would want to do unless you had a compelling reason! Let us say that a boolean algebra B **authenticates** a propositional formula ϕ if every row of this truth table puts `true` under the main connective of ϕ.

Here is the four-element boolean algebra.



and a four-valued truth-table for $p \vee (\neg p)$

p	$\vee$	$(\neg$	$p)$
1	1	0	1
Left	1	Right	Left
Right	1	Left	Right
0	1	1	0

We know what $\wedge$, $\vee$ and $\neg$ are in a boolean algebra, but we defined $\rightarrow$ purely in terms of its two-valued truth-table. So let us say that $p \rightarrow q$ is short for $(\neg p) \vee q$.

So s'pose i fill in a truth table for a formula ϕ using a boolean algebra B. I have $|B|^n$ rows. 2^n of those rows are rows in which the letters take only values $\top$ and $\bot$ (or `true` and `false`). So, if B authenticates ϕ, then $\mathbb{2}$ likewise authenticates ϕ; in plain English, ϕ is a tautology.

So "authenticated by B" implies "authenticated by $\mathbb{2}$" (=tautology)

Now suppose that B does not authenticate ϕ. Suppose there is a row of the truth-table under which ϕ receives the value b, where $b \neq \top$. Consider now any maximal ideal that contains b and the homomorphism onto the quotient, which is $\mathbb{2}$. This homomorphism *kills* b—sends it to $\perp$. This gives us a $\mathbb{2}$-valuation that makes ϕ false.

So "refuted by B" implies "refuted by $\mathbb{2}$" (= not a tautology)

Conclusion:

We defined a tautology to be anything authenticated by the two-element boolean algebra $\mathbb{2}$, but we could, for any boolean algebra B, have defined a tautology to be anything authenticated by B. Authenticated by one is the same as authenticated by all. So it suffices to check authentication by $\mathbb{2}$.

Thus all boolean algebras detect the same set of tautologies. We'd better minute this fact.

Remark 6 *We can define a propositional tautology as "authenticated by all Boolean Algebras" or "authenticated by even one Boolean Algebra"; it makes no difference.*

[in fact we can characterise boolean algebras as those things with $\wedge$, $\vee$, 0 and 1 etc that authenticate all propositional tautologies, but there is no space to prove it, and i can't really pretend it's examinaable. But it's something to guide your thoughts.]

12 Lecture 12

12.1 Applications of Propositional Compactness

Other examples of propositional theories, and applications of propositional compactness.

A group is (right)-orderable if it admits an order $\leq$ such that $(\forall a, b, c)(a \leq b \to a \cdot c \leq b \cdot c)$. (*Think* ... additive group of $\mathbb{Q}$, multiplicative group of $\mathbb{R}_+$ ...) They tend to be abelian so we write the group operation with a '+'.

Remark 7 *A group is right-orderable iff all its finitely generated subgroups are right-orderable.*

Proof:

One direction is easy: if G is right-orderable so are all its subgroups, in particular all its finitely generated subgroups.

For the converse we set up a propositional language and exploit propositional compactness. The language has, for each pair of distinct elements $a, b \in G$, a propositional letter $p_{a,b}$. (Secretly the meaning of $p_{a,b}$ is that $a < b$). The propositional theory to which we are going to apply compactness has the axiom schemes:

$p_{a,b} \to p_{ac,bc}$ for all $a, b, c \in G$
$p_{a,b} \to (p_{b,c} \to p_{a,c})$ for all $a, b, c \in G$
$p_{a,b}$ XOR $p_{b,a}$ for all $a, b \in G$

The first states that the order respects group multiplication, and the second and third assert that the order is total.

Any finite set T' of these axioms is consistent beco's each finite subset mentions only finitely many elements of G. For each such T' consider the subgroup $G^{(T')}$ of G generated by the elements mentioned in the subscripts of the propositional letters appearing in T'. This is a finitely generated subgroup of G and is accordingly orderable by hypothesis. Any ordering of $G^{(T')}$ gives a valuation which satisfies T'. ■

This is one of various standard applications of propositional compactness. Others are ...

(i) *The order extension principle*: every partial order on a set can be refined to a total order;

(ii) If every finite subgraph of a graph is n-colourable then the graph itself is n-colourable.

We will keep these up our sleeve for example-sheet questions. However there are two that we will write out in detail.

12.1.1 The Interpolation Lemma

Lemma 5 *Let P, Q, R be three pairwise-disjoint sets of literals; for any set Γ, let $\mathcal{L}(\Gamma)$ be the set of propositional formulæ built up from literals in Γ. Let $\phi \in \mathcal{L}(P \cup Q)$ and $\psi \in \mathcal{L}(Q \cup R)$ be formulæ such that $(\phi \to \psi)$ is a theorem of the propositional calculus.*

Then there is a formula $\theta \in \mathcal{L}(Q)$ such that both $(\phi \to \theta)$ and $(\theta \to \psi)$ are theorems.

More of a remark than a lemma but it's always called a lemma so i'll go with the flow.
Proof:
Consider the set $\Gamma = \{\gamma \in \mathcal{L}(Q) : \vdash (\phi \to \gamma)\}$ of Q-consequences of ϕ. The idea is to show that this set entails ψ, and that therefore (by compactness) some finite subset of it entails ψ, and the conjunction of that finite subset will be the θ we seek.

If Γ is to entail ψ we want every valuation that satisfies Γ to satisfy ψ. Now we do know that every valuation that satisfies ϕ also satisfies ψ, so it will be sufficient to show that any Q-valuation that satisfies Γ can be extended to a $(P \cup Q)$-valuation that satisifies ϕ.

We argue by contradiction. Suppose there is a Q-valuation v that satisfies Γ but cannot be extended to one that satisfies ϕ. Consider the set $\{p : v(p) = \texttt{true}\} \cup \{\neg p : v(p) = \texttt{false}\}$. This set entails all Q-consequences of ϕ but refutes ϕ itself. So some finite subset Δ of it refutes ϕ. Contraposing we have[4] $\phi \to \neg\Delta$. But $\neg\Delta$ is a Q-consequence of ϕ and is therefore satisfied by v ... which is to say is implied by $\{p : v(p) = \texttt{true}\} \cup \{\neg p : v(p) = \texttt{false}\}$. So there is no such v.

■

Observe that this proof does not tell us how to find θ: it merely tells us there is one such. The following question from example sheet 3 will guide you through a more effective proof that enables you to actually compute θ from ϕ and ψ.

> "(a)$^+$ Suppose A is a propositional formula and 'p' is a letter appearing in A. Explain how to find formulæ A_1 and A_2 not containing 'p' such that A is logically equivalent to $(A_1 \wedge p) \vee (A_2 \wedge \neg p)$.
>
> (b) Hence or otherwise establish that, for any two propositional formulæ A and B with $A \vdash B$, there is a formula C, containing only those propositional letters common to both A and B, such that $A \vdash C$ and $C \vdash B$. (Hint: for the base case of the induction on the size of the common vocabulary you will need to think about expressions over the empty vocabulary)"

Notice the way in which we wellorder the language in the proof of the completeness theorem. That's all right if we have only countably many literals, but it will require nontrivial choice assumptions if the set of literals is uncountable. Can we recover any of the extra strength we have to put in to prove compactness for uncountable propositional languages? Yes!

Remark 8 *Compactness for arbitrary propositional languages implies that every boolean algebra has an ultrafilter.*

Proof:
Let B be any Boolean Algebra. For each $b \in B$ create a propositional letter $\mathcal{U}_b$ whose meaning is secretly that b belongs to the ultrafilter whose existence we are trying to prove. We set up a propositional theory $\mathcal{U}_B$. It contains $\mathcal{U}_\top$, $\neg U_\bot$; for each $b \in B$ it contains $\mathcal{U}_b$ $\texttt{XOR}$ $\mathcal{U}_{\neg b}$; whenever $a = b \vee c$ then it contains $\mathcal{U}_a \to (\mathcal{U}_b \vee \mathcal{U}_c)$ and if $a \leq b$ it contains $\mathcal{U}_a \to \mathcal{U}_b$.

The cardinality of this theory is at least the cardinality of B. Any finite subset is consistent, since any finite set of the $\mathcal{U}_b$ can mention only finitely many elements, and every finite boolean algebra has an ultrafilter. [This is because every finite boolean algebra has minimal nonzero elements[5], and the

[4] I know i should write '$\bigwedge \Delta$' since what i mean is the conjunction of all formulæ in Δ, but ...!

[5] called **atoms**

principal filter generated by such an element is maximal]. So, by compactness, $\mathcal{U}_B$ is consistent. Any valuation v satisfying $\mathcal{U}_B$ tells you which $b \in B$ belong to the ultrafilter corresponding to v. ■

13 Lecture 13

13.1 CNF and DNF

Definition 25 *A formula in a propositional language with only $\wedge$, $\vee$ and $\neg$ is in* **conjunctive normal form** *("CNF") iff it is a conjunction of disjunctions of atomics and negatomics ('$\neg$' is attached only to propositional letters, and there is no '$\wedge$' inside a '$\vee$'); it is in* **disjunctive normal form** *("DNF") iff it is a disjunction of conjunctions of atomics and negatomics ('$\neg$' is attached only to propositional letters, and there is no '$\vee$' inside a '$\wedge$').*

Remark 9 *Every formula is logically equivalent to (is satisfied by the same valuations as) both a formula in CNF and a formula in DNF. The CNF and DNF representations are unique up to reordering of the conjunctions and disjunctions.*

(We have to be careful how we state this last observation: we don't mean identical up to permutations of literals!)

I am not proposing to provide a full proof. The manipulations needed to obtain a CNF or a DNF for a formula rely on the distributivity of $\wedge$ over $\vee$ and of $\vee$ over $\wedge$, plus the de Morgan laws (ask Wikipædia) to "import" the '$\neg$'s.

You will not be asked to prove this fact in an exam—it's not interesting enuff or difficult enuff. It *may* be worth remarking that it can take exponential time to put a formula into CNF/DNF.

Miniexercise: What is the CNF of a tautology?

Now could be a good moment to tackle the following question from Sheet 3:

> "Establish that the class of all propositional tautologies is the maximal propositional logic in the sense that any superset of it that is a propositional logic (closed under $\models$ and substitution) is trivial (contains all well-formed formulæ)."

This maximal propositional logic is always called "classical", and the salient feature that distinguishes it from most subsystems of interest is axiom T, which gives us the de Morgan laws, excluded middle ($A \vee \neg A$) and Double Negation ($\neg\neg A \to A$).

13.1.1 Resolution Theorem Proving

Worth a very brief mention: a proof system for Classical Propositional Logic.

Every formula is a *clause*: a disjunction of atomics and negatomics. The sole rule of inference is "resolution": from $\Gamma \vee p$ and $\Theta \vee \neg p$ infer $\Gamma \vee \Theta$.

The method of proof is: Take your axioms, and turn them into CNF, and thus turn each into a set of clauses. For example, if one of your axioms was $A \longleftrightarrow B$, this has CNF $(\neg A \vee B) \wedge (\neg B \vee A)$ giving the two clauses $\neg A \vee B$ and $\neg B \vee A$. Thereafter, on being challenged to prove ϕ, you turn $\neg\phi$ into CNF, which gives a set of clauses. You add these clauses to the clauses you already have, and you attempt to obtain the empty clause by using the rule of resolution. The empty clause is the `false` ...so if you obtained it you have deduced the `false` from $\neg\phi$ and thereby proved ϕ as desired.

This is the logical basis of the programming language `PROLOG`.

13.2 Predicate Logic begun

> *The predicate calculus with equality. Examples of first-order languages and theories. Statement of the completeness theorem; *sketch of proof*. The compactness theorem and the Löwenheim-Skolem theorems. Limitations of first-order logic [first-order arithmetic of reals and naturals not categorical] . Model Theory.*

No theorems in this first lecture, or two lectures. A few definitions and lots of *culture*.

Explain the syntax before anything else

Definition 26
Predicate/Relation symbol
'=' is a reserved word
arity
function symbol
constant symbol
atomic formula
quantifier

13.3 The Syntax of First-order Logic

All the apparatus for constructing formulæ in propositional logic works too in this new context: If A and B are formulæ so are $A \vee B$, $A \wedge B$, $\neg A$ and so on. However we now have new ways of creating formulæ, new gadgets which we had better spell out:

There is really an abuse of notation here: we should use quasi-quotes ...

13.3.1 Constants and variables

Constants tend to be lower-case letters at the start of the Roman alphabet ('a', 'b' ...) and variables tend to be lower-case letters at the end of the alphabet ('x', 'y', 'z' ...). Since we tend to run out of letters we often enrich them with subscripts to obtain a larger supply: 'x_1' etc.

13.3.2 Predicate letters

These are upper-case letters from the Roman alphabet, usually from the early part: 'F' 'G' .... They are called predicate letters because they arise from a programme of formalising reasoning about predicates and predication. '$F(x, y)$' could have arisen from 'x is fighting y'. Each predicate letter has a particular number of terms that it expects; this is the **arity** *of the letter.* **Unary** *predicates have one argument,* **binary** *predicates have two; n-***ary** *have n. 'loves' has arity 2 (it is binary) 'sits-on' is binary too. If we feed it the correct number of terms—so we have an expression like $F(x, y)$—we call the result an* **atomic formula.**

The equality symbol *'=' is a very special predicate letter: you are not allowed to reinterpret it the way you can reinterpret other predicate letters. The Information Technology fraternity say of strings that cannot be assigned meanings by the user that they are* **reserved***; elsewhere such strings are said to be* **part of the logical vocabulary***. The equality symbol '=' is the only relation symbol that is reserved. In this respect it behaves like '$\wedge$' and '$\forall$' and the connectives, all of which are reserved in this sense.*

Similarly arity of functions. [say a bit more about this]

Atomic formulæ can be treated the way we treated literals in propositional logic: we can combine them together by using '$\wedge$' '$\vee$' and the other connectives.

13.3.3 Quantifiers

Finally we can **bind** *variables with* **quantifiers***. There are two: $\exists$ and $\forall$. We can write things like*

$(\forall x)F(x)$:	*Everything is a frog;*
$(\forall x)(\exists y)L(x, y)$	*Everybody loves someone*

The syntax for quantifiers is variable-preceded-by quantifier enclosed in brackets, followed by stuff inside brackets:

$$(\exists x)(\ldots) \text{ and } (\forall y)(\ldots)$$

We sometimes omit the pair of brackets to the right of the quantifier when no ambiguity is caused thereby.

The difference between variables and constants is that you can bind variables with quantifiers, but you can't bind constants. The meaning of a constant is fixed. Beware! This does not mean that constants are reserved words! The constant 'a' can denote anything the user wants it to denote, it doesn't wander around like the denotation of a variable such as 'x'. Confusingly that's not to say that there are no reserved constants; there are plenty in formalised mathematics, the numerals '0', '1' ...for starters.

For example, in a formula like

$$(\forall x)(F(x) \to G(x))$$

the letter 'x' is a variable: you can tell because it is bound by the universal quantifier. The letter 'F' is not a variable, but a predicate letter. It is not bound by a quantifier, and cannot be: the syntax forbids it. In a first-order language you are not allowed to treat predicate letters as variables: you may not bind them with quantifiers. Binding predicate letters with quantifiers (treating them as variables) is the tell-tale sign of **second-order** *Logic.*

We also have

13.3.4 Function letters

These are lower-case Roman letters, typically 'f', 'g', 'h' .... We apply them to variables and constants, and this gives us **terms***: $f(x)$, $g(a, y)$ and suchlike. In fact we can even apply them to terms: $f(g(a, y))$, $g(f(g(a, y), x))$ and so on. So a term is either a variable or a constant or something built up from variables-and-constants by means of function letters.*

Quantifiers (mention cofinite quantifier for Analysis and "there is an odd number of"—proof that there is no largest prime congruent to 3 mod 4). Mention duality of $\exists$ and $\forall$ in the sense of question (v) on sheet 2.

Difference between 1st and 2nd order theories.

Definition 27 *In second-order languages you are allowed quantifiers over function symbols and predicate letters.*

Topology is a second-order concept.
Complete ordered fields is a 2nd order theory. Simple group.
Possibility of many-sorted theories—not the same!

14 Lecture 14

Bear in mind that, whatever your kit of relation symbols, function symbols etc etc is, the subformula relation between the formulæ you get is going to be wellfounded and you can perform inductions and recursions on it.

For people trying to get entirely straight in their minds what a first-order formula is, examples like the following can be quite confusing. (It's the answer to a question on a 1B compsci example sheet)

$$(\exists x_1 \ldots x_n)(\bigwedge_{1 \leq i \neq j \leq n} x_i \neq x_j) \tag{H}$$

The example sheet question asked for a first-order sentence that is true only in structures with at least n distinct inhabitants.

In a straightforward official sense this sentence (H) is not first-order, in that the recursions that generate first-order formulæ do not output it. For one thing, it exploits the fact that the variables have internal structure, and the '$\bigwedge$' is a binder that—in some sense—quantifies over the subscripts on the variables. But for all that it's not second-order either.

My take on this is that (H) is (obviously) not *literally* a sentence in a first-order language, tho' one could perhaps think of it as a program that, when provided with a numeral as input, outputs a genuine first-order sentence. Which first-order sentence you get will of course depend on the numeral you gave it. Alternatively you can think of it as a uniform parametrised description (in a metalanguage) of an infinite family of first-order sentences. That's probably the simplest way to cope with this kind of slangy mathematical shorthand. It is probably safe to think of formulæ like (H) as first-order, if only by courtesy: trying to spice up the definition of first-order formula so that (H) becomes a first-order formula would be a very messy exercise.

DEFINITION 28

Signature: a structure is a set ('carrier set' better than 'domain') with knobs on. Languages have signatures too. A structure is a structure "for" a language iff they have the same signature.

A substructure of the structure $\mathfrak{M}$ is a subset of the carrier set of $\mathfrak{M}$ equipped with the same knobs and closed under the relevant operations.

Reducts/expansions

We need the concept of signature for basic sense-making reasons. It doesn't make sense to ask whether a formula ϕ is true in a structure $\mathfrak{M}$ unless all the gadgets in ϕ appear also in $\mathfrak{M}$.

Typically signatures tend to be finite. It is true that sometimes, for special reasons, one expands a structure to one with infinitely many constant symbols. In fact we do this in theorem 16.

I try to to use upper-case $\mathfrak{FRAKTUR}$ font for variables ranging over structures, but it doesn't come out very well on a blackboard! I will write the carrier set of the structure $\mathfrak{M}$ as M, the corresponding upper-case roman letter.

$\mathcal{L}(T)$, for T a theory.

et these in the right order

14.1 Axioms

This section will be very short; since we are not going to spend much time actually deducing theorems of LPC we are not going to be very concerned about what the axioms are. In any case, the details of the proof of the completeness theorem do not seem to be very sensitive to one's choice of axioms.

I have copied the following from PTJ's book, and i provide them only for the sake of completeness.

We need the concept of a **free variable**. Brief chat.

$$((\forall x)p) \to p[t/x])$$

where p is a formula with 'x' free in it, and t is any term with no free occurrences of 'x'

$$(\forall x)(p \to q) \to (p \to (\forall x)q)$$

(‘x’ not free in p)

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

$$(\forall xy)(x = y \to p \to p[y/x])$$

p any formula with x

My proof of the completeness theorem will also use the rules:

Universal Generalisation: if we have proved $\phi(x)$ with ‘x’ free, then we have proved $\forall x \phi(x)$. (“Let x be arbitrary ...”)

and a rule that says: if we have a proof of $F(a)$ for some ‘a’ and a proof of $(\exists x)(F(x)) \to p$ then we have a proof of p.

14.2 Semantics

In this section we develop the ideas of truth and validity (which we first saw in the case of propositional logic) in the rather more complex setting of predicate logic.

What we will give is—for each language $\mathcal{L}$—a definition of what it is for a formula of $\mathcal{L}$ to be true in a structure-for-$\mathcal{L}$

The first thing we need is the concept of a signature from definition 28: for a formula ϕ to have a prayer of being true in a structure $\mathfrak{M}$, the signature of the language that ϕ belongs to must be the same as the signature of $\mathfrak{M}$. It simply does not make sense to ask whether or not the transitivity axiom $(\forall xyz)(x < y \wedge y < z. \to x < z)$ is true in a structure that has no binary relation in it.

First we need to decide what our carrier set is to be. Next we need the concept of an **interpretation**. An interpretation is the thing that married up the gadgets in the signature at the structure with the gadgets in the signature in the language. More formally it is a function $\mathcal{I}$ assigning to each predicate letter, function letter and constant in the language of ϕ a subset of M^n, or a function $M^k \to M$, or element of M *mutatis mutandis*. That is to say, to each syntactic device in the language of ϕ, the interpretation assigns a component of $\mathfrak{M}$ of the appropriate arity.

For example, one can interpret the language of arithmetic by determining that the “domain of discourse” (the carrier set) is to be $\mathbb{N}$, the set of natural numbers, and that the interpretation of the symbol ‘$\leq$’ will be the set of all pairs $\langle x, y \rangle$ of natural numbers where x is less than or equal to y, and so on

We have now equipped the language with an interpretation so we know what the symbols mean, but not what the values of the variables are. In other words, settling on an interpretation has enabled us to reach the position from which we started when doing propositional logic. It’s rather like the position we are in when contemplating a computer program but not yet running it. When we run it we have a concept of instantaneous state of the program: these states (snapshots) are allocations of values to the program variables. Let us formalise a concept of state.

A **finite assignment function** is a finite (partial) function from variables in $\mathcal{L}$ to M, the carrier set of $\mathfrak{M}$. These will play a rôle analogous to the rôle of valuations in propositional calculus. I have (see above) carefully arranged that all our variables are orthographically of the form x_i for some index i, so we can think of our assignment function f as being defined either on *variables* or on *indices*, since they are identical up to 1-1 correspondence. It is probably better practice to think of the assignment functions as assigning elements of M to the *indices* and write "$f(i) = \ldots$", since any notation that involved the actual *variables* would invite confusion with the much more familiar "$f(x_i) = \ldots$", where f would have to be a function defined on the things that the variables range over.

Next we define what it is for a partial assignment function to satisfy a sentence p (written "$sat(f, p)$"). We will do this by recursion on the set of formulæ (which comes equipped with a wellfounded subformula relation that justifies induction) so naturally we define *sat* first of all on atomic sentences.

Notice that in

$$sat(f, x_i = x_j)$$

we have a relation between a function and an expression, not a relation between f and x_i and x_j. That is to say that we wish to **mention** the variables (talk about them) rather than **use** them (to talk about what they point to). This contrast is referred to as the **use-mention distinction**.[6] This is usually made clear by putting quotation marks of some kind round the expressions to make it clear that we are mentioning them but not using them. Now precisely what kind of quotation mark is a good question. Our first clause will be something like

$$sat(f, `x_i = x_j\text{'}) \ \text{ iff}_{\text{df}} \ f(i) = f(j). \tag{1}$$

But how much like? Notice that, as it stands, 1 contains a name of the expression which follows the next colon: $x_i = x_j$. Once we have put quotation marks round this, the i and j have ceased to behave like variables (they were variables taking indices as values) because quotation is a referentially opaque context.

A context is **referentially opaque** if two names for the same thing cannot be swapped within it while preserving truth. Quotation is referentially opaque because when we substitute one of the two names for Dr. Jekyll/Mr. Hyde for the other in

'Jekyll' has six letters

we obtain the falsehood

'Hyde' has six letters

[6] It has been said that the difference between logicians and mathematicians is that logicians understand the use-mention distinction.

even though Jekyll and Hyde are the same person. The intuition behind the terminology is that one cannot "see through" the quotation marks to the thing(s) pointed to by the words 'Jekyll' and 'Hyde', so one cannot tell that they are the same. There are other important contexts that are referentially opaque: belief, for example. I might have different beliefs about a single object when it is identified by different names, and these beliefs might conflict.

But we still want the 'i' and 'j' to be variables, because we want the content of clause 1 to read, in English, something like: "for any variables i and j, we will say that f satisfies the expression whose first and fourth letters are 'x', whose third and fifth are i and j, respectively and whose middle letter is '$=$', iff $f(i) = f(j)$". It is absolutely crucial that in the piece of quoted English text 'x' and '$=$' appear with single quotation marks round them while 'i' and 'j' do not. Formula (1) does not capture this feature. To correct this Quine invented a new notational device in (1962), which he called "corners" and which are nowadays known as "Quine quotes" (or "quasi-quotes"), which operate as follows: the expression after the next colon:

$$\ulcorner x_i = x_j \urcorner$$

being an occurrence of '$x_i = x_j$' enclosed in Quine quotes is an expression that does not, as it stands, name anything. However, i and j are variables taking natural numbers as values, so that whenever we put constants (numerals) in place of i and j it turns into an expression that will name the result of deleting the quasi-quotes. This could also be put by calling it a variable name.

A good way to think of quasi-quotes is not as a funny kind of quotation mark—for quotation is referentially opaque and quasi-quotation is referentially transparent—but rather as a kind of diacritic, not unlike the LaTeX commands I am using to write these notes. Within a body of text enclosed by a pair of quasi-quotes, the symbols '$\wedge$', '$\vee$' and so on, do not have their normal function of composing *expressions* but instead compose *names of expressions*. This also means that Greek letters within the scope of quasi-quotes are not dummies for expressions or abbreviations of expressions but are variables that range over expressions (not sets, or natural numbers). Otherwise, if we think of them as a kind of funny quotation mark, it is a bit disconcerting to find that—as Quine points out—$\ulcorner \mu \urcorner$ is just μ (if μ is an expression with no internal structure). The interested reader is advised to read pages 33-37 of Quine's *Mathematical Logic*, where this device is introduced.

It might have been easier to have a new suite of operators that combine names of formulæ to get names of new formulæ so that, as it might be, putting '&' between the names of two formulæ gave us a name of the conjunction of the two formulæ named. However, that uses up a whole font of characters, and it is certainly more economical, if not actually clearer, to use corners instead.

Once we have got that straight we can declare the following recursion, where 'α' and 'β' are variables taking expressions as values.

Definition 29 *First the base cases, for atomic fomulæ*

sat$(f, \ulcorner x_i = x_j \urcorner)$ *iff* $f(i) = f(j)$*;*
sat$(f, \ulcorner x_i \in x_j \urcorner)$ *iff* $f(i) \in f(j)$*.*

Then the inductive steps
if sat(f, α) *and sat*(f, β)*, then sat*$(f, \ulcorner \alpha \wedge \beta \urcorner)$*;*
if sat(f, α) *or sat*(f, β)*, then sat*$(f, \ulcorner \alpha \vee \beta \urcorner)$*;*
if for no $g \supseteq f$ *does sat*(g, α) *hold, then sat*$(f, \ulcorner \neg \alpha \urcorner)$*;*
if there is some $g \supseteq f$ *such that sat*$(g, \ulcorner F(x_i) \urcorner)$*, then sat*$(f, \ulcorner (\exists x_i)(F(x_i)) \urcorner)$*;*
if for every $g \supseteq f$ *with* $i \in dom(g)$*, sat*$(g, \ulcorner F(x_i) \urcorner)$*, then sat*$(f, \ulcorner (\forall x_i)(F(x_i)) \urcorner)$*;*

Then we say that ϕ *is* **true in** $\mathfrak{M}$*, written* $\mathfrak{M} \models \phi$ *iff sat*$(\bot, \phi)$*, where* $\bot$ *is the empty partial assignment function. Finally, a formula is* **valid** *iff it is true in every interpretation.*

14.3 Completeness theorem for LPC: the set of valid sentences is semidecidable

14.3.1 $\in$-terms

Suppose $T \vdash (\exists x)(F(x))$. There is nothing to stop us adding to $\mathcal{L}(T)$ a new constant symbol 'a' and adding to T an axiom $F(a)$. Clearly the new theory will be consistent if T was. Why is this? Suppose it weren't, then we would have a deduction of $\bot$ from $F(a)$. But T also proves $(\exists x)(F(x))$, so we can do a $\exists$-elimination to have a proof of $\bot$ in T. But T was consistent.

Notice that nothing about the letter 'a' that we are using as this constant tells us that a is a thing which is F. We could have written the constant 'a_F' or something suggestive like that. Strictly it shouldn't matter: variables and constant symbols do not have any internal structure that is visible to the language, and the 'F' subscript provides a kind of spy-window available to anyone *mentioning* the language, but not to anyone merely *using* it. The possibility of writing out novel constants in suggestive ways like this will be useful later.

Check for yourself that $(\exists x)(\forall y)(F(y) \to F(x))$ is always true. It tells us that for any F with one free variable we can invent a constant whose job it is to denote an object which has property F as long as anything does. If there is indeed a thing which has F then this constant can denote one of them, and as long as it does we are all right. If there isn't such a thing then it doesn't matter what the constant denotes.

This constant is often written $(\epsilon x)F(x)$. Since it points to something that has F as long as there is something that has F, we can see that

$$(\exists x)(F(x)) \qquad \text{and} \qquad F((\epsilon x)F(x))$$

are logically equivalent (true in the same structures). So we have two rules

$$\frac{(\exists x)(F(x))}{F((\epsilon x)F(x))} \qquad \text{and} \qquad \frac{F((\epsilon x)F(x))}{(\exists x)(F(x))}$$

15 Lecture 15

Theorem 11 *Every consistent theory in a countable language has a model.*

Proof:

Let T_0 be a consistent theory in a countable language $\mathcal{L}(T_1)$.

We do the following things

1. Add axioms to T_0 to obtain a complete extension;
2. Add ϵ terms to the language.

Explain Complete Extension

We execute the task in (1) the way we proved theorem 10—The Adequacy Theorem,

Notice that when we add ϵ-terms to the language we add new formulæ: if '$(\epsilon x)F(x))$' is a new ϵ-term we have just added then '$G((\epsilon x)F(x)))$' is a new formula, and T_0 doesn't tell us whether it is to be true or to be false. That is to say $\mathcal{L}(T_0)$ doesn't contain '$(\epsilon x)F(x)$' or '$G((\epsilon x)F(x)))$'. Let $\mathcal{L}(T_1)$ be the language obtained by adding to $\mathcal{L}(T_1)$ the expressions like '$(\epsilon x)F(x)$' and '$G((\epsilon x)F(x)))$'.

We extend T_0 to a new theory in $\mathcal{L}(T_1)$ that decides all these new formulæ we have added. This gives us a new theory, which we will—of course—call T_1. Repeat and take the union of all the theories T_i we obtain in this way: call it T_∞. (Easy to see that all the T_i are consistent—we prove this by induction).

It's worth thinking about what sort of formulæ we generate. We added terms like $(\epsilon x)(F(x))$ to the language of T_1. Notice that if H is a two-place predicate in $\mathcal{L}(T)$ then we will find ourselves inventing the term $(\epsilon y)H(y,(\epsilon x)F(x))$ which is a term of—one might say—*depth* 2. And there will be terms of depth 3, 4 and so on as we persist with this process. All atomic questions about ϵ terms of depth n are answered in T_{n+1}.

T_∞ is a theory in a language $\mathcal{L}_\infty$, and it will be complete. The model $\mathfrak{M}$ for T_∞ will be the structure whose carrier set is the set of ϵ terms we have generated *en route*[7]. All questions about relations between the terms in the domain are answered by T_∞. The interpretation of an n-ary relation symbol 'R' from $\mathcal{L}(T)$ will be the set of all tuple $\langle t_1 \ldots t_n\rangle$ such that $T-\infty \vdash R(t_1 \ldots t_n)$ andf functions symbols similarly.

Does this make $\mathfrak{M}$ into a model of T? We will establish the following:

Lemma 6 $\mathfrak{M} \models \phi(t_1, \ldots t_n)$ *iff* $T_\infty \vdash \phi(t_1, \ldots t_n)$

Proof: We do this by induction on the complexity of ϕ. When ϕ is atomic this is achieved by stipulation. The induction step for propositional connectives is straightforward. (Tho' for one direction of the '$\vee$' case we need to exploit the fact that T_∞ is complete, so that if it proves $A \vee B$ then it proves A or it proves B.)

[7] And we really do mean the set of epsilon terms, not the denotations of those terms... Our models really are created entirely out of syntax.

The remaining step is the induction step for the quantifiers. They are dual, so we need consider only $\forall$. We consider only the hard direction ($L \to R$).

Suppose $\mathfrak{M} \models (\forall x)\phi(x, t_1, \ldots t_n)$. Then $\mathfrak{M} \models \phi(t_0, t_1, \ldots t_n)$ for all terms t_0. In particular it must satisfy it even when $t_0 = (\epsilon x)(\neg\phi(x, t_1, \ldots t_n))$, which is to say

$$\mathfrak{M} \models \phi((\epsilon x)(\neg\phi(x, t_1, \ldots t_n)), t_1, \ldots t_n)$$

So, by induction hypothesis we must have

$$T_\infty \vdash \phi((\epsilon x)(\neg\phi(x, t_1, \ldots t_n)), t_1, \ldots t_n)$$

whence of course $T_\infty \vdash (\forall x)\phi(x, t_1, \ldots t_n)$.

■

This completes the proof of theorem 11. Observe the essential rôle played by the ϵ terms.

This is a result of fundamental importance. Any theory that is not actually self-contradictory is a description of *something*. It's important that this holds only for first-order logic. It does not work for second-order logic, and this fact is often overlooked. See below, section 16.2.1.

Corollary 10 *Compactness for first-order logic.*

If T is a first-order theory all of whose finite fragments have models then T has a model.

Proof:

Such a T is obviously consistent (proves no contradictions) so, by theorem 11 it has a model. ■

This theorem looks cute and it has many, many, consequences, but most of them are unattractive, and say things like "first-order logic cannot capture this concept". The most striking of them is that there is no first-order way of saying what a natural number is.

Theorem 12 *There are "nonstandard" models of arithmetic.*

Proof:

What does that mean? Let T be a first-order theory of arithemtic of $\mathbb{N}$, with $+$, $\times$, $=$, anything you like, really. Then it has a model which is not the "standard" model. Add a constant symbol—'*' (I don't want to use anything standard and suggestive like 'ω' or '∞'.) Then we add axioms for '*' to say $* \neq 0$, $* \neq 1$, $* \neq 1+1$ .... Clearly any finite subset of T with these new axioms is consistent as long as T was, and so has a model. ■

To be specific (and this might help you get your thoughts about interpretations in order), the first n of these new axioms will tell you only that * must be at least $n+1$. That is to say, for each n, there is an interpretation $\mathcal{I}_n$ of the language of arithmetic which interprets that language *into the standard model* and $\mathcal{I}_n$('*') $= n+1$. $\mathcal{I}$ treats '0', '1' etc as usual.

This technique is used on an industrial scale to show that certain theories are not axiomatisable, by which we mean (as in this case)...

"There is no first-order theory whose models are precisely the standard model of Arithmetic."

In general we prove things like

"There is no first order theory the class of whose models is the class K"

where K is something natural like (for example) the singleton of the standard model of arithmetic, or the class of all simple groups, or the class of all fields of finite characteristic.

This train of thought is a rich source of example sheet questions and old Tripos questions.

15.1 Decidability

Propositional logic is *decidable*: there is an algorithm that tells us whether or not a candidate formula is a tautology. First-order Logic is not decidable in this sense. It's semidecidable because it is axiomatisable: every valid sentence is spat out by our axiomatisation, so if a sentence is valid we learn this fact in finite time. What about if it isn't valid? We would learn that—too—in finite time if every falsifiable first-order formula had a finite countermodel (there are only countably many possible such countermodels and we can wellorder them in order type ω and examine them one by one) but that is not true (consider the negation of the theory DLO of dense linear orders, which we encounter below. It is falsifiable, but the only structures that falsify it are infinite!) We have no time to prove that in this course, but a special case is tractable. A question on sheet 3 invites the reader to show that the monadic fragment of first-order logic (one-place predicate letters only, no function symbols) is decidable.

16 Lecture 16

16.1 The Skolem-Löwenheim Theorems

Notice that the proof of theorem 11 gives us something slightly more than I have claimed. If the consistent theory T we started with was a theory in a countable language then the model we obtain by the above method is also countable. It's worth recording this fact:

Corollary 11
Every consistent theory in a countable (first-order) language has a **countable** *model.*

We can actually prove something more general. Think about what happens to the construction in the proof of theorem 11 if our language has uncountably many constant symbols or function symbols or predicate letters. The proof will procede as before by wellordering the language, and we will build uncountably many ϵ-terms. Clearly the set of terms we generate will be no bigger than the size of the language. This is the

Theorem 13 *Downward Skolem-Löwenheim*

A consistent theory in a language $\mathcal{L}$ has a model of size $|\mathcal{L}|$ at most.

Theorem 14 *Upward Skolem-Löwenheim*

Any theory with infinite models has arbitrarily large models.

Proof: Add lots of constants and appeal to compactness. ■

Actually—tho' i didn't mention this in lectures and I should've done—we can do significantly better. The point is that theorem 13 tells us that if we add enough constants to ensure that the model is of size at least κ, then the model will be no bigger than κ.

This will be amended in next year's edition!

16.2 Categoricity

Definition 30

A theory is categorical iff it has only one model up to isomorphism;

A theory is categorical-in-κ if it has precisely one model of size κ up to isomorphism.

No interesting examples of categorical first-order theories: they all have only finite models (indeed only *one* finite model!)

Plenty of examples of first-order theories categorical-in-$\aleph_0$. Such theories are always called *countably categorical*. Here's the standard example, the theory of dense linear order without endpoints.

Insert DLO axioms here; supply back-and-forth proof of categoricity of DLO. This was done in lectures but not written up.

Now might be a good moment to look at question (vi) on sheet 3.

"Write down axioms for a first-order theory T with $=$ plus a single one-place function symbol f that says that f is bijective and that for no n and no x do we have $f^n(x) = x$.

(a) Is T finitely axiomatisable?
(b) How many countable models does T have (up to isomorphism)?
(c) How many models of cardinality $2^{\aleph_0}$ does it have (up to isomorphism)?
(You may assume that the continuum is not the union of fewer than $2^{\aleph_0}$ countable sets, a fact whose proof—were you to attempt it—would need AC.)
(d) Let κ be an uncountable aleph. How many models does T have of size κ?"

16.2.1 Failure of Completeness of Second-order Logic

It is a truth universally acknowledged among mathematicians—tho' not actually widely *articulated*—that freedom from contradiction guarantees existence. The completeness theorem ensures that for first-order theories. Sadly it is not true for second-order theories. We will exhibit a consistent second-order theory with no models.

We start with second-order arithmetic. It is a standard example of a categorical theory.

The fact that second-order arithmetic is categorical is often bandied about but rarely proved in detail, so i thought i'd do it here, so we can get our thoughts in order on this troublesome matter. In any case it is an essential lemma for the proof of remark 11.

First we need a few words about the difference between semantics for second-order theories and first-order theories. A second order theory looks syntactically exactly like a two-sorted first-order theory with a domain of (as it might be) `wombats` and another domain of `set-of-wombats`. The point about a second-order theory is that the second domain has to be exactly the power set of the first: it must contain **all** subsets of the first domain.

Remark 10 *The second-order arithmetic of* $\mathbb{N}$ *is categorical.*

Proof:

Formally we have a two-sorted language, with lower-case variables to range over natural numbers and upper-case variables to range over sets of natural numbers. It has the usual axioms about addition and multiplication being commutative associative etc, how they distribute, how everything except 0 has a unique predecessor and so on. It will have a set existence axiom saying, for any expression $\phi(\vec{Y}, \vec{y}, n)$ in this language, that $(\forall \vec{Y})(\forall \vec{y})(\exists X)(\forall n)(n \in X \longleftrightarrow \phi(\vec{Y}, \vec{y}, n))$. To capture induction we have an axiom for the least number principle: $(\forall X)(X \neq \emptyset \to (\exists n \in X)(\forall y \in X)(n \leq y))$.

That should do the trick.

This theory is two-sorted: any interpretation $\mathcal{I}$ for this theory must supply *two* domains, one for the lower case (first-order) variables to range over, and the other for the higher-order variables to range over. The fact that this is to be a *second-order* theory means that, for any interpretation $\mathcal{I}$ for this theory, the range of the upper case variables is precisely the power set of the range of the lower case variables. For a two-sorted *first-order* theory we do not have this restriction. This means that there are fewer interpretations of this theory thought of as second-order rather than two-sorted first-order and makes it less surprising that it should be categorical.

So let us place ourselves within some theory that is equipped to discuss models of this theory. The aim is to show that any two are isomorphic. Suppose we have two such models, $\mathcal{A}$ and $\mathcal{B}$ with carrier sets A-with-$\mathcal{P}(A)$ and B-with-$\mathcal{P}(B)$. We want to define a bijection between A and B. Once we have that it will lift to a bijection between $\mathcal{P}(A)$ and $\mathcal{P}(B)$ which will (in fact) be an isomorphism.

Clearly we match up the two zeros—0_A and 0_B—and we want to continue by means of a recursion in the metalanguage. The bijection we want is going to be

$$\bigcap\{R \subseteq A \times B : \langle 0_A, 0_B\rangle \in R \wedge (\forall a \in A)(\forall b \in B)(R``\{a' \in A : a' <_A a\} = \{b' \in B : b' <_B b\} \to \langle a, b\rangle \in R\}$$

We want this to be defined on the whole of A and to cover the whole of B. Suppose not, and consider the subset $X \subseteq A$ consisting of those things not covered by R. Crucially we now need the range of the upper case variables in $\mathcal{A}$ to the the set of **all** subsets of the range of the lower case variables, so this set X really is an element of the model $\mathcal{A}$ and therefore it *it must have a least element* by the least number principle. But then this least element will have been married off to a member of B by the recursion.

Similarly we argue that the whole of B must be covered.

■

We are now in a position to exhibit a consistent second-order theory with no model.

Remark 11 *There is a consistent extension of second-order arithmetic with no model.*

Proof: Add a new constant symbol—'*'—to the language for second-order arithmetic, just as we did in the proof of theorem 12. Then we add to second order arithmetic axioms for '*' to say $* \neq 0$, $* \neq 1$, $* \neq 1+1$ .... Clearly any finite subset of T with these new axioms is consistent as long as T was. However any model of this new theory must have an element denoted by '*', and it can't. ■

16.3 Results related to completeness, exploiting completeness

16.3.1 Prenex Normal Form and quantifier-counting

Definition 31 *A formula is in* **prenex normal form** *if it is of the form string-of-quantifers followed by stuff containing no quantifiers. All quantifiers hae been "pulled to the front" or "exported".*

Theorem 15 *Every formula is equivalent to one in PNF.*

Sketch of proof:

Quantifiers can be "pulled to the front". $(\forall x)(A(x)) \wedge (\forall y)(B(y))$ is clearly equivalent to $(\forall x)(\forall y)(A(x) \wedge B(y))$, and there is an analogous equation for '$\exists$'.

Less obvious that

$(\exists x)(A(x)) \to p$ is equivalent to $(\forall x)(A(x) \to p)$. ■

The significance of this is that it gives us a nice measure of the logical complexity of a formula: count the length of the quantifier prefix once it's in PNF. Better, count the number of **quantifier blocks** in the prefix. There are theorems connecting the quantifier prefixes that you find in the axioms of a theory T with the operations that the class of models of T is closed under. We shall prove the simplest of them to give a flavour.

Definition 32 *A sentence is "universal" iff it is in PNF and its quantifier prefix consists entirely of universal quantifiers. By a natural extension we say a theory is "universal" iff, once you put its axioms into PNF, their quantifier prefixes consist entirely of universal quantifiers. We define "universal-existential" sentences and theories*[8] *similarly as theories all of whose axioms, when in PNF have a block of universal quantifiers followed by a block of existential quantifiers, and so on.*

Definition 33 *The* **diagram** $D_{\mathfrak{M}}$ *of a structure* $\mathfrak{M}$ *is the theory obtained by expanding* $\mathfrak{M}$ *by giving names to every* $m \in M$*, and collecting all true atomic assertions about them.*

17 Lecture 17: The Same Continued

Lemma 7 *For any consistent theory* T *and any model* $\mathfrak{M}$ *of* $T_\forall$*, the set of universal consequences of* T*, the theory* $T \cup D_{\mathfrak{M}}$ *is consistent.*

Proof:

Let $\mathfrak{M}$ be a model of $T_\forall$, with carrier set M. Add to $\mathcal{L}(T)$ names for every member of M. Add to T all the (quantifier-free) assertions about the new constants that $\mathfrak{M}$ believes to be true. This theory is $T \cup D_{\mathfrak{M}}$. We want this theory to be consistent. How might it not be? Well, if it isn't, there must be an inconsistency to be deduced from a conjunction ψ of finitely many of the new axioms. This rogue ψ mentions finitely many of the new constants. We have a proof of $\neg\psi$ from T. T knows nothing about these new constants, so clearly we must have a UG proof of $(\forall\vec{x})\neg\psi$. But this would contradict the fact that $\mathfrak{M}$ satisfies every universal consequence of T. ■

Theorem 16

T is universal iff every substructure of a model of T is a model of T.

Proof:

L → R is easy. We prove only the hard direction.

Suppose that T is a theory such that every substructure of a model of T is also a model of T. Let $\mathfrak{M}$ be an arbitrary model of $T_\forall$. We will show that it must be a model of T. We know already from the foregoing that the theory $T \cup D_{\mathfrak{M}}$ is consistent, and so it must have a model—$\mathfrak{M}^*$, say. $\mathfrak{M}^*$ is a model of T, and $\mathfrak{M}$ is a submodel of $\mathfrak{M}^*$ and therefore (by assumption on T) a model of T—as desired.

But all we knew about $\mathfrak{M}$ was that it was a model of the universal consequences of T. So any old $\mathfrak{M}$ that was a model of the universal consequences of T is a model of T. So T is axiomatised by its universal consequences. ■

[8] PTJ calls such theories "inductive" in his lectures.

There are lots of theorems with this flavour: "The class of models of T is closed under operation burble iff T has an axiomatisation satisfying syntactic condition blah"

But wait! If we have the axiom of choice then, whenever we have an axiom that says $(\forall y)(\exists x)(F(x, y))$ then we can invent a function symbol and an axiom that says $(\forall x)(F(x, f(x)))$. In fact we don't even need the axiom of choice. [you might like to think about why, and have a look at question (xi) on Sheet 3] If we do this often enough we can invent enough function symbols to turn any theory we like into a universal theory, and then all its substructures are also models of it!

Yes you can, but when you add new function symbols you change your notion of substructure! See question (iv) on Sheet 3.

The next theorem after theorem 16 will say that a theory if universal-existential iff the class of its models, partially ordered by isomorphic embeddability, is directed complete. One direction is easy—you might even like to prove it—but the converse (The Chang-Łoś-Suszko lemma) is hard. Indeed i don't know how to pronounce it!

17.1 Omitting Types

You were set up for this topic by question (xiii) on sheet 3. That concerned types for propositional logic, which is a reasonable exercise tho' not actually examinable. There is an analogous result for predicate logic but it much harder and of course is not examinable. However we do need to engage with the ideas.

A *type* in a first-order language $\mathcal{L}$ is a (usually) infinite set of formulæ. A type Σ is an n-type if the formulæ in it all have n free variables.

A model $\mathfrak{M}$ *realises* a type Σ if $\mathfrak{M} \models$ every $\sigma \in \Sigma$. Otherwise v *omits* Σ. We say a theory T *locally omits* an n-type Σ if, whenever ϕ is a formula with n free variables such that T proves $(\forall \vec{x})(\phi(\vec{x}) \to \sigma(\vec{x}))$ for every $\sigma \in \Sigma$, then $T \vdash (\forall \vec{x})(\neg\phi(\vec{x}))$.

The omitting types theorem says that if T locally omits a type Σ, then it has a model omitting Σ. I mention this only to put the standard model of arithmetic into context. The standard model omits the type

$$\{x \neq 0;\ x \neq 1;\ x \neq 1+1;\ x \neq 1+1+1\ldots\}$$

17.2 Some Set Theory

> "*Set theory as a first-order theory; the axioms of ZF set theory. Transitive closures, epsilon-induction and epsilon-recursion. Well-founded relations. Mostowski's collapsing theorem. The rank function and the von Neumann hierarchy.*"

We give a historically motivated introduction.

Points at infinity are concretised as pencils of lines; imaginary divisors concretised as ideals (= sets). Here is the standard example: $\mathbb{Z}[\sqrt{-5}]$ is sold to us as the substructure of $\mathbb{C}$ generated by $\mathbb{Z}$ and $\sqrt{-5}$.

In $\mathbb{Z}[\sqrt{-5}]$ we can factorise 6 as $2 \cdot 3$ and also as $(1+\sqrt{-5}) \cdot (1-\sqrt{-5})$ (we can compute these products in $\mathbb{C}$) and all these four factors are irreducible in $\mathbb{Z}[\sqrt{-5}]$.

So we invent "lower" factors—four of them in fact. One to be a common factor of 2 and $1+\sqrt{-5}$, a second to be a common factor of 2 and $1-\sqrt{-5}$, the third to be a common factor of 3 and $1+\sqrt{-5}$, and finally the fourth to be a common factor of 3 and $1-\sqrt{-5}$. How are we to concretise these ficticious factors? The key observation is that, although we (think) we do not know what these new roots are, we know exactly what their nontrivial multiples are, and that gives us a way in.

Different ideal divisors will correspond to different sets, so we concretise the ideal divisor of 3 and $1+\sqrt{-5}$ *as that set*: $\{a \cdot 3 + b \cdot (1+\sqrt{-5}) : a, b \in \mathbb{Z}\}$.

Integers and rationals similarly.

Also equivalence classes of Cauchy sequences as reals.

We need this beco's—as TWK sez—*prima facie* there seems to be a question about whether or not it is consistent to assume that natural numbers have additive inverses, that integers have multiplicative inverses and that *all* holes in the rationals can be simultaneously filled.

OK, so the set of multiples of the ideal divisor exists as a comprehended object, some suitably concrete object-in-extension. Ditto the pencil-of-lines. So ‘extension”? there is an unproblematic object-in-extension corresponding to the two intension (ideal divisor, point at infinity). Does this always work? Does every set-in-intension have a corresponding set-in-extension? No! Russell was able to show this, using very old ideas going back at least to the Greeks. Russell's paradox. It's an interesting object proof-theoretically but for us it's just a pain. We are going to have to come up with some subset of the set of axioms of naïve set theory plus a good story.

There are various subsets one can use, but—altho' I am an expert on one particular one, due to Quine—i am not going to tell you about that subset, but talk only about the mainstream version which everyone uses. It's known as **Zermelo-Fraenkel Set theory** or ‘ZF’ for short.

A guiding principle in trying to suss out the most suitable subset to use is the recurring thought that set theory started off (as outlined above) as a way of concretising abstract mathematical objects. Thus the axioms of set theory arose largely out of a desire to manipulate sets and prove the existence of such sets as might serve as *simulacra* for mathematical objects. Thus the axioms largely consist of assertions that sets can be manipulated in certain ways, and that the world of sets is closed under certain operations.

Most set theories do not have axioms giving us sets that are interesting in their own right—such as the set of all sets, or the set of all cardinals... largely because the existence of such sets is not compatible with axioms saying that sets can be manipulated freely. In particular they tend not to be compatible with separation...

But first we deal with the most fundamental axiom: extensionality.
$(\forall x, y)(x = y \longleftrightarrow (\forall z)(z \in x \longleftrightarrow z \in y))$.

It's called 'extensionality' because a binary relation R is called 'extensional' as long as $(\forall xy)(x = y \longleftrightarrow (\forall z)(R(z, y) \longleftrightarrow R(z, x)))$.

(Do not confuse this use of 'extensional' with 'extensional' meaning 'truth-functional', contrasted with *intensional*.) The thought behind the axiom of extensionality is that sets are the datatype with absolutely minimal internal structure: sets without knobs on. You don't do anything to their members so the only way of telling two sets apart is by seeing if they have different members.

Let pursue this idea of concretisation and see what axioms it leads us to. We concretise functions as sets of ordered pairs so let's concretise ordered pairs. We want a total [binary] function `pair` and two [unary] partial functions `fst` and `snd` (or π_1 and π_2 if you prefer) s.t.
$(\forall xy)(\texttt{fst}(\texttt{pair}(x, y)) = x)$ and
$(\forall xy)(\texttt{snd}(\texttt{pair}(x, y)) = y)$.

18 Lecture 18

One that works is

Definition 34 *The* **Wiener-Kuratowski** *pair*
$\texttt{pair}(x, y) = \{\{x\}, \{x, y\}\}$.
$\texttt{fst}(p) = \bigcap\bigcap p$ *and*
$\texttt{snd}(p) =$ *the unique member of* $\bigcup p$ *belonging to exactly one member of* p.

$$x = \texttt{snd}(p) \longleftrightarrow (\exists! z)(z \in p \wedge x \in z).$$

If ordered pairs are concretised as above, what axioms do we need if we are to construct and deconstruct them?

Pairing: $(\forall xy)(\exists z)(\forall w)(w \in z \longleftrightarrow (w = x \vee w = y))$

Sumset: $(\forall x)(\exists y)(\forall z)(z \in y \longleftrightarrow (\exists w)(z \in w \wedge w \in x))$

Power set: $(\forall x)(\exists y)(\forall z)(z \in y \longleftrightarrow (\forall w)(w \in z \to w \in x))$

Separation: $(\forall \vec{y})(\forall x)(\exists z)(\forall w)(w \in z \longleftrightarrow w \in x \wedge \phi(w, \vec{y}))$

I have written these out in primitive notation as far as possible. Set theory is a first-order theory in the language with just '$\in$' and '$=$'.

(Think a bit here about our recent proof—remark 10—that second-order arithmetic is categorical. Go back and check: which axioms did we use?)

Separation implies that there is no universal set, lest we get Russell's paradox.

What other axioms are there...? Well, it shouldn't matter how we concretise ordered pairs. Let's try to prove the existence of $X \times Y$ (which is a set, after all, even if it's not a set *of sets*) without knowing what an ordered pair is.

For any $x \in X$ we consider the function $f_x : y \mapsto \langle x, y\rangle$. Then f_x"Y is just $\{x\} \times Y$. Consider now the function $F_x : x \mapsto \{x\} \times Y$. Then F_x"X is $\{\{x\} \times Y : x \in X\}$ and $\bigcup$ of this is just $X \times Y$. ■

Notice that we have not made any assumptions about what particular object $\langle x, y\rangle$ might be for $x \in X$ and $y \in Y$. However we have assumed (twice) that *the image of a set in a function is a set.* This assumption is the **axiom scheme of replacement**.

(If you want to prove the existence of $X \times Y$ in the special case where your ordered pairs are Wiener-Kuratowski you don't need replacement, tho' you do need power set. This is an old example sheet question, and you might like the try it—it'll help you to get a feel for set-theoretic manipulation.)

The formulation of the replacement scheme in the language of set theory is slightly fiddly, because we do not want variables ranging over functions:

$$(\forall y)(\exists! x)(\phi(y, x)) \to (\forall Y)(\exists X)(\forall x)(x \in X \longleftrightarrow (\exists y \in Y)(\phi(y, x)))$$

Of course this can be done with parameters, but stating that makes it even harder to read.

The upper case 'X' and 'Y' are not second-order variables; i'm using upper case to make it easier to read. [There is actually a converse to this: if $X \times Y$ always exists however you implement pairing and unpairing then the axiom scheme of replacement follows. It's an exercise on Sheet 4.]

The image of a set in a class is a set

Now that we have replacement various things become possible. We can give a proper definition of transitive closure and we can construct the cumulative hierarchy.

Let us take these in turn.

18.1 Transitive Closures and Transitive Sets

The justification I gave of R-induction on the assumption that R is wellfounded was an informal one. Now that we are doing set theory formally the time has come to formally deduce R-induction from the assumption that R is wellfounded.

Suppose $(\forall x)(((\forall y)(R(y, x) \to F(y)) \to F(x))$. Suppose (with a view to obtaining a contradiction) that $\neg F(a)$ for some a. Naturally we want a to give rise to a set with no R-minimal element, thereby contradicting wellfoundedness of R. The obvious candidate is the set $\{z : R^*(z, a) \wedge \neg F(z)\}$ which is a subset (so we use separation) of the set $\{z : R^*(z, a)\}$ of things related to a by R^* the transitive closure of R. How are we going to prove that this is a set? If we are to do it with the axioms we have seen so far we are clearly going to have to use replacement, as follows. Use the function $n \mapsto R^n$"$\{a\}$ and take the image of $\mathbb{N}$ in it; then do $\bigcup$ to the result. The trouble with this is that the 'n' is not a variable in the language. We need a relation ϕ that relates n to R^n"$\{a\}$. We do this by saying $\phi(n, X)$ if every set that contains $\langle n, X\rangle$ and contains $\langle m-1, Y\rangle$ whenever it contains $\langle m, R``Y\rangle$ also contains $\langle 0, X\rangle$.

$$\Phi(n, X) \longleftrightarrow (\forall A)((\langle n, X\rangle \in A \wedge (\forall m, Y)((\langle m, R``Y\rangle \in A \to \langle m-1, Y\rangle \in A) \to \langle 0, a\rangle \in A)$$

It takes a while to get your head round this definition[9]! This is *Quine's trick*.

Definition 35 *Transitive closure*

The collection

$$TC(x) = \bigcup_{n\in\mathbb{N}} (\bigcup{}^n x)$$

is a set. This set is the **Transitive closure** *of x.*

We also say: x is **transitive** *if $x \subseteq \mathcal{P}(x)$.*

By Quine's trick we prove that $TC(x)$ is always a set.

Beware overloading of this terminology! (We have "transitive closures" of relations too!) Evidently $TC(x)$ is the $\subseteq$-least transitive superset of x.

18.2 The Cumulative Hierarchy

Definition 36 *The* **Cumulative Hierarchy**

is defined by recursion on the ordinals:

$$V_\alpha =: \bigcup_{\beta<\alpha} \mathcal{P}(V_\beta).$$

We need, perhaps, to say a little bit about why this definition is legitimate. *Prima facie* there is a worry because we are doing a recursion over all the ordinals (which is not a set—see corollary 3) rather than merely over an initial segment of it (which is). It's OK because the functions we define by recursion on those initial segments all agree.

We want to be sure that V_α exists for all α. Obviously we want to do an induction over the ordinals. No problem at successor ordinals, co's we use Power set. The justification at limit ordinals needs replacement. We need to take the image of the collection of ordinals below α in the function $\gamma \mapsto V_\gamma$.

Can do no harm to take some time out to think about what the various V_αs look like. V_0 is empty; $V_1 = \{\emptyset\}$; $V_2 = \{\emptyset, \{\emptyset\}\}$ .... (How big is V_n?). What does V_ω consist of?

We now define a rank function on members of the cumulative hierarchy:

Definition 37 *$\rho(x)$ is the least ordinal α such that $x \subseteq V_\alpha$.*

Conway used to say of the rank of a set that it was the set's *birthday*.

The alert and suspicious reader will notice that i am using the same letter 'ρ' here as in definition 13, and will wonder whether or not this is legitimate. It is, and i think it is safe to bounce this back to the reader.

[9] And thanks to Mr Irving of Churchill for spotting the typo.

(i) We prove by induction on the ordinals that $\langle V_\alpha, \in\restriction V_\alpha\rangle$ is a wellfounded binary structure, so it has a rank function.
(ii) Then we prove that all the rank functions we obtain, for all α, agree.
(iii) Finally we prove that they agree with the *other* (novel) rank function that we have just defined in definition 37.

Now is as good a place as any to record the fact that every V_α is transitive. So every set in the cumulative hierarchy has a transitive closure in the cumulative hierarchy.

LEMMA 8 $(\forall\alpha)(V_\alpha \subseteq \mathcal{P}(V_\alpha))$

Proof:
Of course you do this by induction. I think i can safely leave the details to the reader. ■

There is an intimate connection between the cumulative hierarchy and the concretisation project, to which we now return ...

We cannot straightforwardly concretise/implement cardinals as equivalence classes if we have separation beco's $\bigcup\alpha = V$ whenever α is an equipollence class, so separation will give us Russell's paradox.

REMARK 12 *If α is an equipollence class (other than $\{\emptyset\}$) then $\bigcup\alpha = V$.*

Proof:
Suppose not. Then there is b s.t $(\forall A \in \alpha)(b \notin A)$. Let A be a member of α (any will do). For any $a \in A$, the set $(A \cup \{b\}) \setminus \{a\}$ is in bijection with A and is therefore in α. But then $b \in \bigcup\alpha$ after all.
■

COROLLARY 12 *The only equipollence class that is a set is $\{\emptyset\}$.*

In fact something analogous happens for any any equivalence relation $\sim$ with a natural global definition: if $[X]_\sim$ is an equivalence class then $\bigcup^n [X]_\sim = V$ for some small concrete n depending only on $\sim$. However the details of the proof depend very sensitively on the definition of the equivalence relation, so we don't bother with the details, but just draw the moral: equivalence classes are not the way to concretise mathematical objects arising from equivalence relations.

However, now that we have the cumulative hierarchy, we are in a position to solve the problem of implementing objects that arise from equivalence relations.

18.3 Scott's Trick

DEFINITION 38 Scott's trick
When trying to concretise/implement a mathematical entity that arises naturally from equivalence classes for an equivalence relation, then instead of $[x]_\sim$ the $\sim$-equivalence class of a set x, we use the collection of things $\sim$ to x that are in the cumulative hierarchy and are of minimal rank with that property.

Thus, if $\sim$ is an equivalence relation we instantiate the (as it might be, cardinal) not as the true equivalence class—which might not be available—but instead as $[x]_\sim \cap V_\alpha$ where α is the least ordinal α s.t. $[x]_\sim \cap V_\alpha$ is nonempty. Observe that x might not be a member of its (as-it-might-be) cardinal thus construed!

For this to work we need to be sure that, for all x and all equivalence relations $\sim$, there is some y in the cumulative hierarchy with $y \sim x$. There are various axioms that deliver this (one of them is the antifoundation axiom of Forti and Honsell, which you may have heard of: "every set picture is a picture of a unique set"[10]) but the simplest way to ensure it is to brutally assume that every set is in the cumulative hierarchy.

This is one of the various forms of the axiom of foundation.

19 Lecture 19

19.1 The Axiom of Foundation

This axiom takes various forms, and it's worth taking some time to straighten them out.

One form is the assertion that every set is wellfounded. What do we mean by a wellfounded set? We know what a wellfounded *relation* is, but a wellfounded *set*? The most intuitively appealing way to characterise wellfounded sets is to say that x is a wellfounded set iff there is no ω-sequence $\langle x_i : i \in \mathbb{N}\rangle$ with $x = x_0$ and $(\forall n \in \mathbb{N})(x_{i+1} \in x_i)$, but this is equivalent to the correct definition only if we have dependent choice. The correct definition is that x is a wellfounded set iff $\in\upharpoonright TC(\{x\})$ is a wellfounded relation.

So we want an axiom that says that all sets are wellfounded. We can do this by saying that $\in$ is a wellfounded relation, but that's a bit suspect because the universe is not a set if foundation holds, so we are cutting off the branch we are sitting on.

> We can adopt a scheme of $\in$-induction.
>
> We can say that every set is wellfounded, as above.
>
> There is also the axiom of restriction ...

The axiom of restriction says $(\forall x)(\forall y)(x \in y \to (\exists z \in y)(z \cap y = \emptyset))$. "$(\exists z \in y)(z \cap y = \emptyset)$" sounds a bit more like foundation if you read it as "y has an $\in$-minimal element". But what about the "$(\forall x)(\forall y)(x \in y \to$" bit? This harks back to the proof by mathematical induction that every nonempty set of natural numbers has a $<_\mathbb{N}$-least element. You prove by induction on n that every subset of $\mathbb{N}$ containing n has a $<_\mathbb{N}$-minimal element.

[10] A set picture is a digraph (set of ordered pairs) that looks as if it could be the graph of $\in$ restricted to a transitive set.

The axiom of restriction is an attempt to say that every nonempty set has an $\in$-minimal element:

$$(\forall y)(y \neq \emptyset \to (\exists z \in y)(z \cap y = \emptyset)).$$

$$(\forall y)((\exists x)(x \in y) \to (\exists z \in y)(z \cap y = \emptyset)).$$

By standard manipulation of first-order formulæ this becomes

$$(\forall y)(\forall x)(x \in y \to (\exists z \in y)(z \cap y = \emptyset)).$$

and then you permuste the quantifiers

$$(\forall x)(\forall y)(x \in y \to (\exists z \in y)(z \cap y = \emptyset)).$$

and then it appears to say that every x has a certain property, which we call 'regular'.

This is why the axiom of restriction/foundation is important: by unleashing Scott's trick it enables us to always concretise any mathematical entity arising from an equivalence relation. If we do *not* have the axiom of foundation then models can be found in which there are illfounded sets that are not the same size as any wellfounded set. That would mean, at the very least, that we cannot use Scott's trick to implement cardinals.

Thus Scott's trick in conjunction with the axiom of foundation has solved the concretisation problem for objects arising from equivalence relations.

There are still two axioms we haven't mentioned, at least not in this connection. One is the axiom of choice, which we saw earlier. The other arises from the need to implement $\mathbb{N}$ and $\mathbb{R}$. It's clear than any set that implements $\mathbb{N}$ must be infinite, and we have not so far had an axiom that tells us there are infinite sets and we can no longer postpone postulating them. The axiom of infinity will tell us that there is an infinite set. It comes in various forms, and if we have the axiom scheme of replacement and foundation and AC then all the forms you might think of turn out to be equivalent. One specially fiddly version that is often seen in the literature is

Axiom of Infinity: $(\exists x)(\emptyset \in x \wedge (\forall y)(y \in x \to y \cup \{y\} \in x))$

Quite why it should take this form has something to do with the implementation of ordinals, to which we now turn.

We can of course use Scott's trick to implement ordinals but with ordinals we have an extra trick up our sleeve. Every equivalence class (= abstract ordinal) contains a wellordering whose order relation is set membership, and this wellordering is unique. We prove this using ...

19.2 Mostowski Collapse

Lemma 9 *(Mostowski's collapse lemma)*

If $\langle X, R\rangle$ is a well-founded structure, then there is a transitive set Y and a homomorphism $f{:}\langle X, R\rangle \to \langle Y, \in\rangle$.

Proof: We use the theorem about wellfounded induction and recursion, theorem 1. Set $\pi(x) := \{\pi(y) : R(y, x)\}$. The *definiens* (The RHS) is a set by replacement.

The desired Y is simply the range of π. Y is transitive because nothing ever gets put into Y unless all its members have been put in first. ■

Mostowski collapse shows that every wellfounded structure $\langle X, R\rangle$ has a homomorphism π onto a structure $\langle \pi``X, \in\rangle$ where $\pi``X$ is a transitive set.

[Explain *homomorphism*?]

In general there is no reason to expect that the homomorphism π is injective. It's simple to give illustrations where it is and also illustrations where it isn't. If $\{y : R(y, x_1)\} = \{y : R(y, x_2)\}$ then clearly $\pi(x_1) = \pi(x_2)$. Clearly if there is no such pair x_1 and x_2 then π will be injective. Recall from page 52 that in these circumstances we say that R is **extensional**. Reflect that the axiom of extensionality says that $\in$ is extensional.

If R is extensional, then no two things in X have the same set of R-predecessors and so no two things ever get sent to the same thing by π. This give us the special case:

If $\langle X, R\rangle$ is a well-founded extensional structure, then there is a **unique** transitive set Y and a unique isomorphism between $\langle X, R\rangle$ and $\langle Y, \in\rangle$.

Mostowski collapse is a crucial lemma in the study of wellfounded sets, and it gets used all the time, but we mustn't lose track of the fact that we are encountering it in the context of a story about how to implement ordinals. So we ask: What happens in the cases where $\langle X, R\rangle$ is a wellordering? Wellorders are total orders so distinct things have distinct predecessors so the homomorphism is an isomorphism.

Thus every wellordering is isomorphic to a wellordering whose order relation is $\in$! And this wellordering is of course unique. [Why?] We then take this canonical representative to be our ordinal.

Definition 39

Every wellordering is isomorphic to a unique wellordering $\langle X, \in\rangle$ where X is a transitive set. Such a wellordering is a **von Neumann ordinal**.

(often just called plain 'ordinals' [which is naughty]).

Remark 13

The order relation $<_{On}$ on von Neumann ordinals is $\in$;
Each ordinal is identical to the set of its predecessors;
$\alpha + 1 = \alpha \cup \{\alpha\}$.

Notice that in showing that every wellordering is isomorphic to a unique von Neumann ordinal we have used replacement but have not used foundation.

The fact that every von Neumann ordinal coincides with the set of the ordinals below it fits very cutely with theorem 4 that says that every ordinal counts the set of ordinals below it in their natural order.

Notice also that although we have shown that every wellordering is isomorphic to a special one (which we can use as its ordinal) namely the wellordering whose order relation is $\in$, there doesn't seem to be a similar move available for cardinals. Given a set x is there an obvious special set in bijection with x, something that we can use as its cardinal? Not clear. We will return to this later.

Now is the moment to observe that the peculiarly specific form of the axiom of infinity we saw on p. 58 has a purpose. It precisely gives us a set containing 0 and closed under successor, and we can obtain the $\subseteq$-least such set from it by separation, as follows:

Let A be a set given by the fancy version of the axiom of infinity. Then the set we want is

$$\{x : x \in A \wedge (\forall y)(\emptyset \in y \wedge (\forall w)(w \in y \to w \cup \{w\} \in y) \to x \in y)\}$$

which is a set by separation. That set is of course the set of finite von Neumann ordinals, which will do for our implementation of $\mathbb{N}$.

Once we've implemented ordinals we can implement integers, rationals, reals and complexes. In lots of different ways, in fact.

> Naturals can be von Neumann naturals or Zermelo naturals or Scott's trick naturals;
>
> Integers can be signed naturals or equivalence classes of ordered pairs of naturals;
>
> Rationals can be signed ordered pairs of naturals or equivalence classes of ordered pairs of integers;
>
> Reals can be Dedekind cuts in rationals or equivalence classes of Cauchy sequences of rationals;
>
> Complex numbers typically are thought of as ordered pairs of reals.

And in every case where you are using equivalence classes to implement something there is the possibility of using Scott's trick to cut the class down to something smaller.

It would be a very helpful exercise to crunch out the ranks of the sets that implement these various mathematical objects under the assorted possible implementations. A question on sheet 4 invites you to do that...

The answers themselves do not matter in the slightest—the ordinals obtained are properties of the implementing sets, not of the mathematical entities themselves[11]—but the exercise will give you experience in manipulating some purely set theoretic quantities, and prepare you for doing some more idiomatic set theory in the days to come—something you will not have done before.

19.3 Ordinals again

We now pick up the thread dropped on page 4.

DEFINITION 40
(i) An aleph is the cardinality of a (usually infinite) wellordered set;
(ii) $\aleph(\alpha)$, for α a cardinal, is the least aleph $\not\leq \alpha$.

(I think we first saw this aleph-without-a-subscript notation in lecture 4 on p 14).

Look again at the proof of lemma 5, Hartogs' lemma, which told us that $\aleph(\alpha)$ is always defined. The proof i gave there used replacement (tho' we didn't bring out the use of replacement!) It is possible to give a proof without replacement (as Hartogs in fact originally did, the axiom scheme of replacement not having been formulated at that stage) as follows.

Given X we seek a wellordered set Y with $|Y| \not\leq |X|$.

Consider $\mathcal{P}(X \times X)$ (use Wiener-Kuratowski ordered pairs if you want to be specific); throw away every subset that isn't a wellordering; quotient out what's left under isomorphism. The result is (a concretisation) of the set of ordinals of wellorderings of subsets of X—as it were equivalence-classes-local-to-X—and is the Y we desire.

This argument gives us an upper bound for $\aleph(|X|)$: $\aleph(\alpha) \leq 2^{2^{\alpha^2}}$. By modifying the construction you can obtain better bounds (such as $\aleph(\alpha) \leq^* 2^{\alpha^2}$—where the asterisk means surjection) but we don't need them.

19.3.1 Initial ordinals

For the moment write 'card(α)' for $|\{\beta : \beta <_{On} \alpha\}|$. (This 'card' notation is in the literature, but it is not in common use, and you do not need to know it). Then

DEFINITION 41
An ordinal α is **initial** *if $(\forall \beta <_{On} \alpha)(card(\beta) <_{card} card(\alpha))$.*
We enumerate the initial ordinals as ω_0, ω_1, $\ldots \omega_\alpha \ldots$, and
We define $\aleph_\alpha$ to be $card(\omega_\alpha)$ which of course was $|\{\beta : \beta <_{On} \omega_\alpha\}|$.

The following should be evident:

$\aleph_\alpha$ is also the αth aleph;

$\aleph_{\alpha+1}$ is $\aleph(\aleph_\alpha)$;

The alephs are wellordered by $<_{card}$.

[11] In PTJ's book p.87 he uses the phrase "essential rank" of a mathematical entity. It's a nice phrase, and it should be standard, but it isn't.

We can use initial ordinals to implement alephs as sets. Every aleph corresponds to a unique initial ordinal, so we can implement an aleph as the corresponding (von Neumann) initial ordinal. If we are willing to adopt AC then every cardinal is an aleph, and we have in fact implemented all cardinals. Could we not have implemented cardinals by Scott's trick? Yes, if we have foundation, or even if we have the (weaker) assertion that every set is the same size as a wellfounded set. This route *via* von Neumann initial ordinals doesn't need either of these assumptions, but it does use AC.

However it is blindingly cute, and has become the industry standard.

20 Lecture 20

Remark 14 *Every regular ordinal is initial.*

Proof:

It's not a particularly deep or important fact but it's basic and will help you orient yourself. And the proof is idiomatic. Actually we prove the contrapositive.

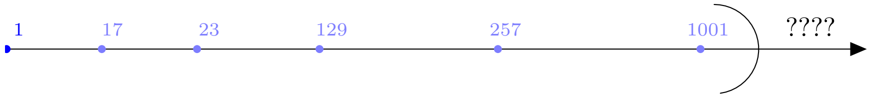
We need a factoid. Suppose $\langle A, <_A\rangle$ and $\langle B, <_B\rangle$ are (strict) total orders, with $<_A$ a wellorder, and there is a bijection $f : A \twoheadrightarrow B$. (We probably want B to not have a last element; must check what else we might need). We are *not* assuming that f is order-preserving! Nevertheless f does have a maximal order-preserving restriction, a rather special one: there is $A' \subseteq A$ s.t $f \restriction A'$ is order-preserving, and $f``A'$ is cofinal (unbounded) in $\langle B, <_B\rangle$.

We obtain A' by recursion on $\langle A, <_A\rangle$. The first member of A' is the bottom element of $\langle A, <_A\rangle$. Thereafter the next member is always the $<_A$-least element a of A. s.t. $f(a) >_B f(a')$ for all $a' <_A a$ that we have already put into A'. Suppose $f``A'$ were bounded in $\langle B, <_B\rangle$. Consider the subset $B' \subseteq B$ consisting of things not dominated by any $f(a)$ for $a \in A'$, and consider the $b \in B'$ s.t. $f^{-1}(b)$ is $<_A$-minimal. $f^{-1}(b)$ should have been put into A'.

End of factoid: ■

Now suppose β is not an initial ordinal. (As I said, we are proving the contrapositive). Then there is $\alpha < \beta$ s.t. α has as many predecessors as β. Let $\langle A, <_A\rangle$ and $\langle B, <_B\rangle$ (as in the factoid) be the ordinals below α and the ordinals below β respectively. The factoid gives us a set of ordinals cofinal in β whose order type $\leq \alpha < \beta$. So β is not regular. ■

Here is an illustration of a particular case.



The picture shows why every countable limit ordinal has cofinality ω. The long right-pointing arrow represents a countable ordinal manifested as a wellordering of naturals ($\mathbb{N}$ in a funny order). The (unbounded!) increasing sequence of natural numbers reading from the left are the numbers chosen as in the recursion ... 1001 is the least natural number > 257 that is above 257 in both orders. The semicircle represesents where this increasing sequence of naturals comes to a halt, closes off. Are there any natural numbers in the region flagged by the question marks? Suppose there were—347, say. OK, so what were doing declaring 1001 to be the 6th member of the sequence? We should have used 347!

Thus every countable limit ordinal λ is the sup of an ω-sequence $\langle \lambda_i : i < \omega \rangle$ of smaller ordinals.

Definition 42
Such a sequence of smaller ordinals is a **fundamental sequence** *for λ.*

Fundamental sequences give you a way of using ordinals to measure how rapidly growing a function $f : \mathbb{N} \to \mathbb{N}$ is. One can define a sequence f_α over countable ordinals α by something like $f_0(n) = n+1$; $f_{\alpha+1}(n) = (f_\alpha)^n(n)$ and (and this is the clever bit) if λ is the sup of $\langle \lambda_n : n < \omega \rangle$ set $f_\lambda(n) = f_{\lambda_n}(n)$.

[Something to think about ... every regular ordinal is initial ... is every initial ordinal regular... ? ω is initial and is regular; you saw in an example sheet question that ω_1 (which is obviously initial) is regular ...]

20.1 $\aleph^2 = \aleph$

(Using the letter '$\aleph$' as a variable to range over alephs...)

We start by noting that $\aleph = \aleph + \aleph$. (Well, what we will *actually* need is $\aleph + \aleph + \aleph = \aleph$, but never mind). Beginners might like to have this spelled out, and it holds because $2 \cdot \omega_\alpha = \omega_\alpha$. How so? Any order of limit order-type consists of lots of concatenated copies of $\mathbb{N}$, each of length ω. You can interleave two (or indeed three) worders of length ω to get a worder of length ω so you can do this for all the copies simultaneously.

We start by defining a function $\mathfrak{S} : On \to On$. Given an ordinal α, take a wellordering $\langle A, <_A \rangle$ of order type α, make disjoint copies of all its proper initial segments, and then concatenate the copies ... with longer things appended after shorter things.

The result is a wellordering and its order type is defined to be $\mathfrak{S}(\alpha)$. [This notation is not standard, and I am not going to use it outside this proof so i'm not numbering it]. Thus—for example—$\mathfrak{S}(\omega) = 1 + 2 + 3 + 4 + \ldots = \omega$

Lemma 10

(i) $\mathfrak{S} : On \to On$ is a normal function;
(ii) Every initial ordinal is a value of $\mathfrak{S}$.

Proof:
(i) $\mathfrak{S} : On \to On$ evidently also has a recursive definition:

$\mathfrak{S}(\alpha+1) = \mathfrak{S}(\alpha) + \alpha$ and
$\mathfrak{S}(\lambda) = \mathrm{Sup}\{\mathfrak{S}(\alpha) : \alpha < \lambda\}$ for λ limit.

...from which it is clear that $\mathfrak{S}$ is a normal function.

(ii)
Use the division algorithm for normal functions to show that there is a β s.t $\mathfrak{S}(\beta) \leq \omega_\alpha < \mathfrak{S}(\beta+1)$. If $\mathfrak{S}(\beta) < \omega_\alpha$ then we have $\omega_\alpha \leq \mathfrak{S}(\beta+1) = \mathfrak{S}(\beta) + \beta$ which is impossible, since $\mathfrak{S}(\beta)$ and β both have cardinality below $\aleph_\alpha$. ■

We want to show that $(\aleph_\alpha)^2 = \aleph_\alpha$. $\aleph_\alpha$ is defined as the cardinal $\{\beta : \beta < \omega_\alpha\}$, which means that the canonical set of size $(\aleph_\alpha)^2$ is the cartesian product $\{\beta : \beta < \omega_\alpha\} \times \{\beta : \beta < \omega_\alpha\}$. We partition this last set into three pieces:

(i) the [graph of] the identity relation restricted to $\{\beta : \beta < \alpha\}$, and

(ii), (iii)

the two triangles above-and-to-the-left, and below-and-to-the-right of the diagonal.

To be slightly more formal about it, we partition the cartesian product $\{\beta : \beta < \alpha\} \times \{\beta : \beta < \alpha\}$ into the three pieces $\{\langle\beta,\gamma\rangle : \beta < \gamma < \alpha\}$, $\{\langle\beta,\gamma\rangle : \beta = \gamma < \alpha\}$ and $\{\langle\beta,\gamma\rangle : \gamma < \beta < \alpha\}$.
It is clear that the third piece is of order type $\mathfrak{S}(\alpha)$ in the lexicographic order.
The idea is to show that these three pieces all have cardinality $\aleph_\alpha$. That's obvious for the second piece, the identity relation. Also there is an obvious bijection between the first and third piece ("flip your ordered pairs") so it will suffice to prove that the third piece ("the bottom-right triangle") has cardinality $\aleph_\alpha$.

Now we can prove

Theorem 17 $(\forall\alpha)(\aleph_\alpha = (\aleph_\alpha)^2)$.

Proof:
By induction on α. The fact that it holds for $\alpha = 0$ you learnt in 1a.[12]
Assume true for all alephs $< \aleph_\alpha$. By lemma 10, ω_α is a value of $\mathfrak{S}$; we want to show that it is actually a fixed point. Now ω_α is an initial ordinal, which is to say that for any $\beta < \omega_\alpha$, the cardinal $|\{\gamma : \gamma < \beta\}|$ is less than $\aleph_\alpha$, and (by induction hypthesis) is equal to its own square. Suppose ω_α were $\mathfrak{S}(\beta)$ for some $\beta < \omega_\alpha$. This would entail that the size of the cartesian product $\{\gamma : \gamma < \beta\} \times \{\gamma : \gamma < \beta\}$ is at least $\aleph_\alpha$, contradicting the induction. So ω_α is a fixed point of $\mathfrak{S}$. This means that the lower-right triangle of the cartesian product $\{\gamma : \gamma < \omega_\alpha\} \times \{\gamma : \gamma < \omega_\alpha\}$—which can be wellordered to length $\mathfrak{S}(\omega_\alpha) = \omega_\alpha$—is of cardinality $\aleph_\alpha$. It's clearly naturally isomorphic to the

[12] Thank you Mr Rogers, for pointing out that i had not supplied the base case!!

upper-left triangle (as remarked earlier) so the cartesian product is now a union of three sets each of size $\aleph_\alpha$, giving $(\aleph_\alpha)^2 = \aleph_\alpha + \aleph_\alpha + \aleph_\alpha = \aleph_\alpha$ as desired.

■

Thus if the axiom of choice holds (so every infinite cardinal is an aleph) then $\alpha = \alpha^2$ for all infinite cardinals. There is a converse!

Corollary 13 *If $\alpha = \alpha^2$ for all infinite cardinals, then AC follows.*

Proof: Let α be an arbitrary infinite cardinal, and suppose $\beta^2 = \beta$ for all infinite cardinals β. Then we have

$$\begin{aligned} \alpha + \aleph(\alpha) &= (\alpha + \aleph(\alpha))^2 \\ &= \alpha^2 + 2 \cdot \alpha \cdot \aleph(\alpha) + (\aleph(\alpha))^2 \\ &= \alpha + 2 \cdot \alpha \cdot \aleph(\alpha) + \aleph(\alpha) \\ &= \alpha + \alpha \cdot \aleph(\alpha) + \aleph(\alpha) \\ &= \alpha(1 + \aleph(\alpha)) + \aleph(\alpha) \\ &= (\alpha \cdot \aleph(\alpha)) + \aleph(\alpha) \\ &= (\alpha + 1) \cdot \aleph(\alpha) \\ &= \alpha \cdot \aleph(\alpha) \end{aligned}$$

Now we use Bernstein's lemma which i put in my initial handout but which of course you haven't read, so i'll include it here.

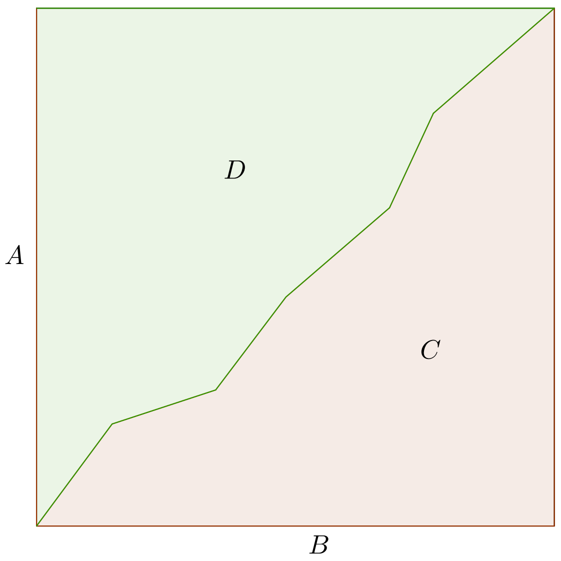
21 Lecture 21

Remark 15 *Bernstein's lemma*

$$\gamma + \delta = \alpha \cdot \beta \ \rightarrow \ \alpha \leq^* \gamma \vee \beta \leq \delta$$

Here '$\alpha \leq^* \gamma$' means that there is a surjection from a set of size γ to a set of size[13] α.
Proof:

[13] Or $\alpha = 0$, yes.



Proof: Suppose A and B are two sets (of size α and β). Suppose further that we have split $A \times B$ (represented by the square figure above) into two pieces, C and D (of size γ and δ), so that $C \cap D = \emptyset$ and $C \cup D = A \times B$. Now project the C region onto the A axis. Does it cover the whole of the A-axis? (I've tried to draw the picture so that it's not clear whether it does or not!) If it does, then $|A| \leq^* |C|$. If it doesn't, then there is a line through D parallel to the B axis, whence $|B| \leq |D|$.

■

Returning to the proof of corollary 13. We can apply Bernstein's lemma in two ways. We can infer $\aleph(\alpha) \leq^* \alpha \vee \alpha \leq \aleph(\alpha)$. The second disjunct is the one we want so we would like to exclude the first disjunct: $\aleph(\alpha) \leq^* \alpha$. For all we know this could happen if α is not an aleph, so we have to use Bernstein the other way round:

$$\aleph(\alpha) \leq \alpha \vee \alpha \leq^* \aleph(\alpha)$$

The first disjunct is of course impossible—by definition of $\aleph(\alpha)$—so we infer the second, which tells us that any set of size α is a surjective image of a wellordered set. But any such surjective image can be wellordered, and this gives us our result. ■

We can also use theorem 17 to show that a lot of initial ordinals are regular.

THEOREM 18 *(uses AC)*

Every ordinal $\omega_{\alpha+1}$ is regular.

Proof:

If $\omega_{\alpha+1}$ is the sup of fewer than $\aleph_{\alpha+1}$—which is to say the sup of no more than $\aleph_\alpha$ smaller ordinals—then the set of ordinals below it (which is of size

$\aleph_{\alpha+1}$) is a union of at most $\aleph\alpha$ things each of size $\aleph_\alpha$ at most. We saw in an example sheet question how to use AC to show that such a union is of size $(\aleph_\alpha)^2$ at most, and theorem 17 now tells us it is of size $\aleph_\alpha$ at most, which is impossible. ■

The obvious follow-up question is: if λ is limit can ω_λ be regular? It is if $\lambda = 0 \ldots$. The conttext in which to consider this question is the context of independence proofs, to which we now turn.

21.1 Independence of the Axioms from each other

We've spent quite a lot of time and energy rolling out set theory as a platform on which to do mathematics; it can do no harm to do something a bit more idiomatic; set theory does, after all, have a life of its own. The schedules require me to cover problems of consistency and independence of the axioms, so let's do that.

We prove independence results by exhibiting models. We emphasise that for philosophical reasons we are interested only in transitive models. The idea is that if i give you a set x i must also give you all its members—since a set, after all, is nothing more than the set of all its members. So any sensible model with an element x must contain everything in the transitive closure of x as well. Hence our restriction to transitive models only.

That is one reason why Mostowski collapse is so important.

21.1.1 Δ_0 formulae and the Lévy Hierarchy

First we define Δ_0 formulæ and a quantifier hierarchy associated with them

Definition 43 *A Δ_0-formula in the language of set theory is a formula built up from atomics by means of boolean connectives and restricted quantifiers. A restricted quantifier in the language of set theory is '$(\forall x)(x \in y \to \ldots)$' or '$(\exists x)(x \in y \wedge \ldots)$'. Thereafter a Σ_{n+1} (respectively Π_{n+1}) formula is the result of binding variables in a Π_n (repectively Σ_n) formula with existential (respectively universal) quantifiers.*

We immediately extend the Σ_n and Π_n classes by closing them under interdeducibility-in-a-theory-T, and signal this by having 'T' as a superscript so our classes are Σ^T_n and Π^T_n. As usual, we omit the superscripts when they are clear from context.

We find that Δ_0 formulæ behave in many ways as if they contained no quantifiers at all. An unrestricted quantifier is an injunction to scour the whole universe in a search for a witness or a counterexample; a restricted quantifier invites us only to scour that part of the universe that lies in some sense "inside" something already given. The search is therefore "local" and should behave quite differently: that is to say, restricted universal quantification ought to behave like a finite conjunction and ought to distribute over disjunction in the approved de Morgan way. (And restricted existential quantification too, of course).

One effect of this is that Δ_0 predicates are **absolute** between transitive models. This merits a short discussion. If $\phi(x)$ is a formula with one free variable and no quantifiers, and $\mathfrak{M}$ believes there is an x such that $\phi(x)$, then any $\mathfrak{M}' \supseteq \mathfrak{M}$ will believe the same. This much is obvious. The dual of this is similarly obvious: If $\phi(x)$ is a formula with one free variable and no quantifiers, and $\mathfrak{M}$ believes that $\phi(x)$ holds for every x, then any $\mathfrak{M}' \subseteq \mathfrak{M}$ will believe the same. We say that existential formulæ **generalise upwards** and universal formulæ **generalise downwards**. Something analogous holds for Σ_1 formulæ and Π_1 formulæ. They generalise upwards and downwards in the same way *as long as $\mathfrak{M}$ and $\mathfrak{M}'$ are both transitive models.* Δ_0 formulæ of course generalise both upward and downward and are therefore **absolute**.

We need this gadgetry if we are to cope with what is usually the first problem students have with finding models for fragments of ZF. The first thing to note is that if $\mathfrak{M} = \langle M, \in \rangle$ is a model of set theory then '$\mathfrak{M} \models \phi$' is actually a formula of set theory. Which formula? The formula we obtain from 'ϕ' by replacing every quantifier '$(\forall x)(\ldots)$' by '$(\forall x)(x \in M \to \ldots)$' and replacing every quantifier '$(\exists x)(\ldots)$' by '$(\exists x)(x \in M \wedge \ldots)$'.

The problem i have just spoken of is this: most of the axioms of ZF take the form of an assertion that the universe is closed under some operation or other. If we are to get straight which sets (or classes) are models of which axioms we will need to be absolutely clear about the difference between being closed under an operation and being a model for the axiom that says you are closed under that operation. You might think that for a set to be a model of the axiom that says the world of sets is closed under operation blah it is necessary and sufficient for that set to be closed under operation blah. But you'd be wrong!

$\mathfrak{M} \models$ the axiom of pairing iff

$$(\forall x \in M)(\forall y \in M)(\exists z \in M)(\forall w \in M)(w \in z \longleftrightarrow w \in x \vee w \in y)$$

$\mathfrak{M}$ is closed under the pair set operation iff $(\forall x, y \in M)(\{x, y\} \in M)$.

In contrast $\mathfrak{M} \models$ the axiom of power set iff

$$(\forall x \in M)(\exists y \in M)(\forall z \in M)(z \in y \longleftrightarrow (\forall w \in M)(w \in z \to w \in x))$$

Now, since $\mathfrak{M}$ is transitive, the last bit—$(\forall w \in M)(w \in z \to w \in x)$—is equivalent to $z \subseteq x$, so the displayed formula simplifies slightly to

$$(\forall x \in M)(\exists y \in M)(\forall z \in M)(z \in y \longleftrightarrow z \subseteq x)$$

$\mathfrak{M}$ is closed under the power set operation iff

$$(\forall x \in M)(\mathcal{P}(x) \in M)$$

Are these two equivalent? Clearly not. Reflect that, by Downward Skolem-Löwenheim (theorem 13) and Mostowski collapse (lemma 9) ZF has a countable transitive model $\mathfrak{M}$. In a countable transitive model every set must be countable.

So the thing in $\mathfrak{M}$ that $\mathfrak{M}$ believes to be power set of $\mathbb{N}$ will be a countable set and cannot possibly be the true power set of the naturals.

The point is that "$x = \{y, z\}$" is just "$y \in x \ \wedge \ z \in x \ \wedge \ (\forall w \in x)(w = y \vee w = z)$" which is Δ_0 and is absolute;

In contrast $x = \mathcal{P}(y)$ is

$(\forall w \in x)(\forall z \in w)(z \in y) \wedge (\forall w)((\forall u)(u \in w \to u \in y) \to y \in x)$ which is not Δ_0!

We are now in a position to look at some actual independence results.

21.1.2 Some actual independence results

Let's start with the simplest possible example. It exploits V_ω, a set i talked about earlier, and whose existence i proved in lectures. For which axioms ϕ can we establish that $\langle V_\omega, \in\rangle \models \phi$?

Well, it's transitive so it's a model for extensionality. It's a model for pairing and power set, and is actually closed under pairing and under power set. It's a model of separation because any subset of a member of V_ω is also a member of V_ω. What about replacement? Is the image of a set in V_ω in some function also a set in V_ω? Well, obviously not, beco's such a function could send its arguments from V_ω into the wide blue yonder, but it doesn't have to! For V_ω to be a model of replacement all that is necessary is that if we have a function from V_ω to V_ω *which is definable with all its parameters in V_ω and all its bound variables constrained to range over things in V_ω* then the image of an element of V_ω in such a function is also in V_ω. And *that* is clearly true—we don't even need the italicised condition.

Reflect that $\mathfrak{M} \not\models \bot$, for all $\mathfrak{M}$, so no inconsistent theory can have a model. Therefore the fact that V_ω is a set means that we have proved the consistency of *something*, that something being whatever the set of things is that are all true in $\langle V_\omega, \in\rangle$.

To cut a long story short it's pretty clear that it is a model of all the axioms except infinity: V_ω not only does not contain any infinite set, it doesn't even contain any set that it mistakenly believes to be infinite. However it satisfies all the other axioms. In fact it's even a model of the Axiom of choice, and it's a model of the axiom of choice even if the theory in which we are conducting this discussion does not assume AC.

This shows that the axiom of infinity does not follow from the other axioms of ZFC

Another structure to consider is $V_{\omega+\omega}$. This is transitive, so it's a model of extensionality. It's obviously a model of pairing, sumset and power set. Also separation (Put them all on the board and tick them off one by one). This time it's clearly a model of infinity. Not only does it contain an infinite set, it contains an infinite set which is infinite in the sense of the model. ("x is infinite" is not Δ_0 so we have to be careful.)[14]

It's going to be a model of sumset because something gets into $V_{\omega+\omega}$ as long as its rank is less than $\omega+\omega$ ... and $\bigcup$ decreases rank. (and $y = \bigcup x$ is Δ_0).

It will be a model of the AC as long as the theory in which we are conducting the analysis has AC as an axiom. As long as our ordered pairs are Wiener-Kuratowski any wellordering of a member of $V_{\omega+\omega}$ will also be a member of $V_{\omega+\omega}$, a couple of layers higher up. (W-K pairs increase rank by 2).

So: which axiom or axiom scheme is left? Replacement!

You want to say ... "it can't be a model of replacement, beco's—if it were—it would then be a model of the whole of ZF, and so we would have proved the consistency of ZF inside ZF" and you read somewhere about the Incompleteness theorem of Gödel that says that can't happen. And you'd be right of course. However it would be nice to have an actual instance of replacement that fails. Ideally i'd let you think about it but time is short. Consider the function that sends n to $V_{\omega+n}$. You will need Quine's trick to define it properly

Both the models we have considered so far are V_αs. However there are other structures we can use.

22 Lecture 22 (Independence results continued)

Definition 44

$\mathcal{P}_\phi(x) = \{y \subseteq x : \phi(y)\}$;
H_ϕ is the least fixed point for $x \mapsto \mathcal{P}_\phi(x)$;
Alternatively $H_\phi = \{x : (\forall y \in TC(\{x\}))\phi(y)\}$.

We also write '$\mathcal{P}_\kappa(x)$' (where κ is a cardinal) for $\{y \subseteq x : |y| < \kappa\}$, and H_κ for the least fixed point of this function, so that $H_\kappa = \{x : (\forall y \in TC(\{x\}))(|y| < \kappa)\}$

The 'H' means 'hereditarily'.

Observe that $V_\omega = H_{\aleph_0}$. $V_{\omega+\omega}$ is not H_α for any cardinal α.

The next H after $H_{\aleph_0}$ is $H_{\aleph_1}$, the set of hereditarily countable sets, commonly notated 'HC'.

It's perhaps not blindingly obvious that HC is a set. However if you have countable choice (so that ω_1 is regular) then every hereditarily countable set is in V_{ω_1} and then HC is a set by separation.

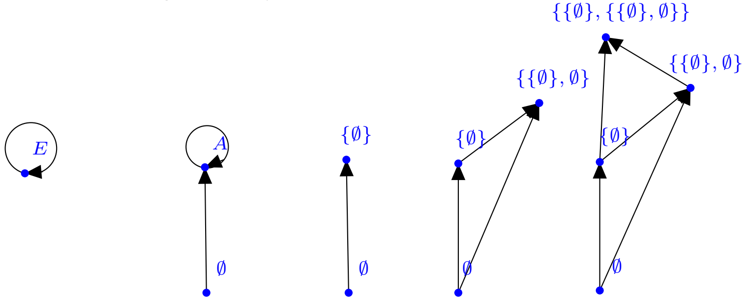
However we can prove the existence by the natural device of set pictures

[14] There is a subtlety here, because the specially sexed-up version of the axiom of infinity $(\exists x)(\emptyset \in x \wedge (\forall y)(y \in x \to y \cup \{y\} \in x))$ that we saw on page 58 asserts that there is an x with a special property, and that special property is Δ_0. The point is that you have to do a bit of work to show that if $\emptyset \in x \wedge (\forall y)(y \in x \to y \cup \{y\} \in x)$ then x really is Dedekind-infinite.

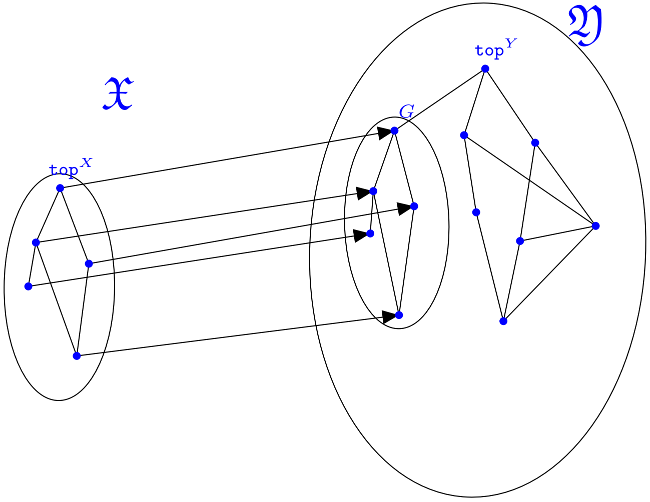
Definition 45 *Set Pictures*

A (wellfounded) **set picture** *(also known as an* **accessible pointed (di)graph** *or* **APG***) is a special kind of decorated digraph, a (wellfounded) extensional binary structure* $\langle X, R, \texttt{top}^R\rangle$ *with a designated ("top") element* $\texttt{top}^R$ *s.t.* $X = (R^*)^{-1}$“$\{\texttt{top}^R\}$. *I.e., it's a binary structure that looks as if it might be the graph of* $\in^*$ *restricted to the transitive closure of a singleton.*

Here are some graphics for APGs:


$\{\{\emptyset\}, \{\{\emptyset\}, \emptyset\}\}$
$\{\{\emptyset\}, \emptyset\}$
$\{\{\emptyset\}, \emptyset\}$
E
A
$\{\emptyset\}$
$\{\emptyset\}$
$\{\emptyset\}$
$\emptyset$
$\emptyset$
$\emptyset$
$\emptyset$

This second graphic shows the embedding relation which we define below.



This diagram[15] *depicts the obvious embedding relation between set pictures:* $\mathfrak{X} = \langle X, R, \texttt{top}^R\rangle$ *embeds into* $\mathfrak{Y} = \langle Y, S, \texttt{top}^S\rangle$ *if there is* $y \in Y$ *with* $S(y, \texttt{top}^S)$ *and* $\langle X, R, \texttt{top}^R\rangle \simeq \langle S^{-1}$“$\{y\}, S, \texttt{top}^S\rangle$.

Clearly isomorphism is a congruence relation for this embedding relation and we write '$\mathcal{E}$*' for the relation between the isomorphism classes.*

If you forget this definition you can reconstruct it if you remember that it's

[15] I have to confess that the binary relation in the picture isn't extensional. With any luck the reader won't notice. What matters is the isomorphism between the two smaller ellipses.

trying to say that the set that $\langle X, R, \texttt{top}^R\rangle$ is a picture of is a member of the set that $\langle Y, S, \texttt{top}^S\rangle$ is a picture of.

Now to prove the existence of HC we consider $V_{\omega+1}$ and wellfounded set pictures, and Scott's-trick isomorphism classes thereof. Evidently the family of (Scott's-Trick) equivalence classes is a wellfounded binary structure, so we can take the Mostowski collapse. The Mostowski collapse is HC.

Which axioms are true in HC?

I'm hoping that by now you can be trusted, Dear Reader, to calculate which axioms are true in HC. We need countable choice to show that a union of countably many countable sets iss countable (and we'll need that if we are to verify sumset). We can verify all of them except power set. Why is power set not true in HC? Well, everything in HC is countable, and the power set of a countably infinite set is uncountable. But life is not that simple. Remember that by downward Skolem-Löwenheim ZF must have a countable model and indeed (by Mostowski collapse) a countable *transitive* model. In any such model every set is countable! However not all the inhabitants of such a model are countable *in the sense of the model*: the model contains some (externally) countable sets for which it does not supply a bijection to the naturals of the model. In HC, in contrast, every set is internally countable, so the axiom of power set really does fail.

We still have to prove the independence of extensionality, AC, pairing, sumset and foundation. Let's press on.

22.0.3 Independence of Sumset

Let $\beth_\omega$ be the cardinal $\sup\{\aleph_0, 2^{\aleph_0}, 2^{2^{\aleph_0}}, 2^{2^{2^{\aleph_0}}} \ldots\}$. (We can define $\beth$ numbers with arbitrary subscripts but we don't need to)

To prove the independence of sumset we consider $H_{\beth_\omega}$. This is a set for the same reason that HC is. This time we consider the set of all wellfounded set pictures in $V_{\omega+\omega}$ and consider the set of Scott's-trick equivalence classes of them. The embedding just described is inherited by the quotient, and we write the inherited embedding as '$\mathcal{E}$'. Evidently the family of (Scott's-Trick) equivalence classes is a wellfounded binary structure, so we can take the Mostowski collapse. The Mostowski collapse is $H_{\beth_\omega}$.

$H_{\beth_\omega} \not\models$ sumset because $\{V_{\omega+n} : n < \omega\}$ is in $H_{\beth_\omega}$ but $\bigcup\{V_{\omega+n} : n < \omega\} = V_{\omega+\omega}$ is not.

It is a model of all the other axioms for familiar reasons. It satisfies AC as long as the universe within which we constructed it satisfies AC.

22.1 Independence of the Axiom of Foundation

Let σ be the transposition $(\emptyset, \{\emptyset\})$. Equip the universe with a *new* membership relation $x \in_\sigma (y)$ defined as $x \in \sigma(y)$. Observe that $\emptyset \in_\pi \emptyset$, so foundation does not hold in $\langle V, \in_\sigma\rangle$. What about the other axioms? The first thing to note is that all the axioms of ZFC (except foundation) are preserved whatever permutation you use.

Definition 46 *ϕ^σ is the result of replacing '$\in$' in ϕ throughout by '$\in_\sigma$'.*

Then $(\langle V, \in_\sigma\rangle \models \phi) \longleftrightarrow \phi^\sigma$. You want to prove $\vdash (\forall\sigma)(\phi \longleftrightarrow \phi^\sigma)$ where ϕ is an axiom other than foundation.

OK, so you look at ϕ^σ, and you notice that *prima facie* distinct occurrences of a given variable have different prefixes. Variables that never appear to the right of an '$\in$' you say are of level 0, and you don't have a problem with them. Variables that appear to the right of an '$\in$' only when the variable to the left of the $\in$ are of level 1 and you don't have a problem with them—unless they also appear to the left of an $\in$. Let 'y' be such a variable. Then we have subformulæ like $x \in \sigma(y)$ and $y \in \sigma(z)$.

We make the elementary observation that '$x \in \sigma(y)$' is equivalent to '$\sigma(x) \in \sigma``(\sigma(y))$' and so can be replaced by it in $\phi$ where appropriate. $\sigma``z$ is $\{\sigma(w) : w \in z\}$ and the function $z \mapsto \sigma``z$ is of course just yet another permutation. We might find that we have to "lift" σ in this way more than once .... So the notation '$j(\sigma)$' for this new permutation might come in handy.

The key is to manipulate the formulæ you are dealing with so as to ensure that, for every variable, every occurrence of that variable has the same prefix ...the point being that $(\forall x)(\ldots\sigma(x)\ldots)$ is equivalent to $(\forall x)(\ldots x\ldots)$ beco's σ is a permutation.

This is a description of the recursive step in an algorithm for rewriting atomic formulæ in such a way that, for each variable, all its occurrences end up with the same prefix, so we can reletter. The second step? We now find that some occurrences of 'z' have no prefix, whereas some have.

The definition of *stratifiable* for a formula is simply that this algorithm succeeds.

It's now simple to verify that ϕ^σ is equivalent to ϕ as long as ϕ is stratifiable. Not all instances of replacement are stratifiable but it turns out not to matter.

$$(\forall x\exists! y)\phi(x,y) \ \rightarrow\ (\forall X)(\exists y)(\forall z)(z \in Y \longleftrightarrow (\exists w)(w \in X) \wedge \phi(w,z))$$

becomes

$$(\forall x\exists! y)\phi^\sigma(x,y) \ \rightarrow\ (\forall X)(\exists y)(\forall z)(z \in \sigma(Y) \longleftrightarrow (\exists w)(w \in \sigma(X)) \wedge \phi^\sigma(w,z))$$

We can drop the σs preceding 'X' and 'Y' to obtain

$$(\forall x\exists! y)\phi^\sigma(x,y) \ \rightarrow\ (\forall X)(\exists y)(\forall z)(z \in Y \longleftrightarrow (\exists w)(w \in X) \wedge \phi^\sigma(w,z))$$

which is merely another instance of replacement (as long as σ is a function class). Thus the map (on the syntax) sending each ϕ to ϕ^σ sends every stratifiable formula ϕ to (something logically equivalent to) ϕ, and sends every instance of replacement to something logically equivalent to another instance.

We now check that every axiom other than foundation is either stratifiable or interdeducible with a stratifiable formula, and accordingly remains true in the new model. Observe that in the new model $\emptyset$ has become an object equal to its own singleton. Such objects are called **Quine atoms**. We added only one Quine atom, but if (say) we had swapped every natural number with its singleton we would have added countably many. We will need this when we come to prove the independence of AC.

22.2 Independence of the Axiom of Choice

Proving the independence of the axiom of choice from ZF is hard work, and was finally cracked by Cohen in 1963 with the advent of *forcing*. Forcing is too demanding for a course like this, but there are other ideas that go into the independence proof, and some of them can be profitably covered here.

One useful thought is that the axiom of choice says that the universe contains some highly asymmetrical objects. After all, as we saw in theorem 2 on page 9, any wellordering is rigid. If we can arrange matters so that everything in the universe has some symmetries then we will break AC. I've made it sound easier than it is, but that's the idea.

We start with a model of ZF + foundation, and use the permutation methods seen above to obtain a permutation model with a countable set A of Quine atoms. The permutation we use to achieve this is the product of all transpositions $(n, \{n\})$ for $n \in \mathbb{N}^+$.

A will be a **basis** for the illfounded sets in the sense that any class X lacking an $\in$-minimal element contains a member of A. Since the elements of A are Quine atoms every permutation of A is an $\in$-automorphism of A, and since they form a basis we can extend any permutation σ of A to a unique $\in$-automorphism of V in the obvious way: declare $\sigma(x) := \sigma``x$. Notice that the collection of sets that this definition does not reach has no $\in$-minimal member if nonempty, and so it must contain a Quine atom. But σ by hypothesis is defined on Quine atoms.

Any permutation σ of the atoms can be extended to an $\in$-automorphism of the universe (also written σ) by declaring $\sigma(x) = \sigma``x$. Now (a, b) is of course the transposition swapping a and b, and we will write '(a, b)' also for the unique automorphism to which the transposition (a, b) extends. Every set x gives rise to an equivalence relation on atoms. Say $a \sim_x b$ if (a, b) fixes x. We say x is of (or has) **finite support** if $\sim_x$ has a cofinite equivalence class. (At most one equivalence class can be cofinite).

The union of the (finitely many) remaining (finite) equivalence classes is the **support** of x. Does that mean that x is of finite support iff the transitive closure $TC(x)$ contains finitely many atoms? Well, if $TC(x)$ contains only finitely many atoms then x is of finite support (x clearly can't tell apart the cofinitely many atoms not in $TC(x)$) but the converse is not true: x can be of finite support if $TC(x)$ contains cofinitely many atoms. (Though that isn't a sufficient condition for x to be of finite support!!)[16]

[16] A counterexample: wellorder cofinitely many atoms. The graph of the wellorder has cofinitely many atoms in its transitive closure, but they are all inequivalent.

It would be nice if the class of sets of finite support gave us a model of something sensible, but extensionality fails: if X is of finite support then $\mathcal{P}(X)$ and the set $\{Y \subseteq X : Y$ is of finite support$\}$ are both of finite support and have the same members with finite support. We have to consider the class of elements hereditarily of finite support. Let's call it HF. This time we *do* get a model of ZF.

Lemma 11 *The class of sets of finite support is closed under all the definable operations that the universe is closed under.*

Proof:

When x is of finite support let us write '$A(x)$' for the cofinite equivalence class of atoms under $\sim_x$. For any two atoms a and b the transposition (a, b) induces an $\in$-automorphism which for the moment we will write (a, b), too.

Now suppose that $x_1 \ldots x_n$ are all of finite support, and that f is a definable function of n arguments. $x_1 \ldots x_n$ are of finite support, and any intersection of finitely many cofinite sets is cofinite, so the intersection $A(x_1) \cap \ldots A(x_n)$ is cofinite. For any a, b we have

$$(a, b)(f(x_1 \ldots x_n)) = f((a, b)(x_1) \ldots (a, b)(x_n))$$

since (a, b) is an automorphism. In particular, if $a, b \in A(x_1) \cap \ldots A(x_n)$ we know in addition that (a, b) fixes all the $x_1 \ldots x_n$ so

$$(a, b)(f(x_1 \ldots x_n)) = f(x_1 \ldots x_n).$$

So the equivalence relation $\sim_{f(x_1 \ldots x_n)}$ induced on atoms by $f(x_1 \ldots x_n)$ has an equivalence class which is a superset of the intersection $A(x_1) \cap \ldots A(x_n)$, which is cofinite, so $f(x_1 \ldots x_n)$ is of finite support. ■

This takes care of the axioms of empty set, pairing, sumset and power set. To verify the axiom scheme of replacement we have to check that the image of a set hereditarily of finite support in a definable function (with parameters among the sets hereditarily of finite support and all its internal variables restricted to sets hereditarily of finite support) is hereditarily of finite support too. The operation of translating a set under a definable function (with parameters among the sets hereditarily of finite support and all its internal variables restricted to sets hereditarily of finite support) is definable and will (by lemma 11) take sets of finite support to sets of finite support.

So if X is in HF and f is a definable operation as above, $f``X$ is of finite support. And since we are interpreting this in $HF$, all members of $f``X$ are in HF, so $f``X$ is in HF too, as desired.

To verify the axiom of infinity we reason as follows. Every wellfounded set x is fixed under all automorphisms, and is therefore of finite support. Since all members of x are wellfounded they will all be of finite support as well, so x is hereditarily of finite support. So HF will contain all wellfounded sets that were present in the model we started with. In particular it will contain the von Neumann ω.

It remains only to show that AC fails in HF. Consider the set of (unordered) pairs of atoms. This set is in HF. However no selection function for it can be. Suppose f is a selection function. It picks a (say) from $\{a, b\}$. Then f is not fixed by (a, b). Since f picks one element from every pair $\{a, b\}$ of atoms, it must be able to tell all atoms apart; so the equivalence classes of $\sim_f$ are going to be singletons, $\sim_f$ is going to be of infinite index, and f is not of finite support.

So the axiom of choice for countable sets of pairs fails. Since this axiom is about the weakest version of AC known to man, this is pretty good. The slight drawback is that we have had to drop foundation to achieve it. On the other hand the failure of foundation is not terribly grave: the only illfounded sets are those with a Quine atom in their transitive closures, so there are no sets that are gratuitously illfounded: there is a basis of countably many Quine atoms. On the other hand it is only the illfounded sets that violate choice!

23 Lecture 23

24 Lecture 24: Independence

Internalising: Deduction theorem. Gödelisation. There are at least four clever ideas in Gödel's paper:
(i) p.r. functions;
(ii) arithmetisation of syntax;
(iii) the β-function and
(iv) diagonalisation.
(iv) was not his, and (iii) is not essential (tho' it is clever). The key is the combination of (ii) and (iv).

25 Example Sheets

Questions marked with a '+' are brief reality-checks; questions marked with a '*' are for enthusiasts/masochists only; ☠ means what you think it means, and particularly tasty questions are decorated with a pink marzipan pig: 🐖

Sheet 0: Numbers and Sets Revision

1A Numbers and Sets is the only prerequisite for this course, and it can do you no harm to give a quick going-over to your notes for that course. You might like to have a quick glance at my supervision/lecture notes for Discrete maths for Computer Scientists, linked from my 1a teaching page. It's *Sets* rather than *Numbers* but that's OK beco's there is no number theory in Part II ST&L.

Countability

"uncountably many" wasn't ever a complete answer to the question "How many wombats are there?" It *just may* (sometimes) still be an *adequate* answer but—now that you are doing Part II you should always be prepared to give more detail. Read `www.dpmms.cam.ac.uk/~tf/countability.pdf` and do the exercises therein; it won't take you long.

(i)

Explain briefly why the diagonal argument that shows that $\mathcal{P}(\mathbb{N})$ is uncountable doesn't show that there are uncountably many finite sets of naturals.

Set Theory and Logic, Michaelmas 2016, Sheet 1: Ordinals and Induction

Questions marked with a '*' may be skipped by the nervous.

(i)

Write down subsets of $\mathbb{R}$ of order types $\omega + \omega$, ω^2 and ω^3 in the inherited order.

(ii)

Which of the following are true?

(a) $\alpha + \beta$ is a limit ordinal iff β is a limit ordinal;
(b) $\alpha \cdot \beta$ is a limit ordinal iff α or β is a limit ordinal;
(c) Every limit ordinal is of the form $\alpha \cdot \omega$;
(d) Every limit ordinal is of the form $\omega \cdot \alpha$.

For these purposes 0 is a limit ordinal.

(iii)

Consider the two functions $On \to On$: $\alpha \mapsto 2^\alpha$ and $\alpha \mapsto \alpha^2$. Are they normal?

(iv)

Prove the converse to lemma 2: if $\langle X, <_X \rangle$ is a total order satisfying "every subordering is isomorphic to an initial segment" then it is a wellordering.

(v)

What is the smallest ordinal you can not embed in the reals in the style of question (i)?

(vi)

Prove that every [nonzero] countable limit ordinal has cofinality ω. What about ω_1?

(vii)*

Recall the recursive definition of ordinal exponentiation:

$$\alpha^0 = 1;\ \alpha^{\beta+1} = \alpha^\beta \cdot \alpha, \text{ and } \alpha^{sup(B)} = \sup(\{\alpha^\beta : \beta \in B\}).$$

Ordinal addition corresponds to disjoint union [of wellorderings], ordinal multiplication correponds to lexicographic product, and ordinal exponentiation corresponds to ...? Give a definition of a suitable operation on wellorderings and show that your definition conforms to the spec: $\alpha^{\beta+\gamma} = \alpha^\beta \cdot \alpha^\gamma$.

(viii)

Let $\{X_i : i \in I\}$ be a family of sets, and Y a set. For each $i \in I$ there is an injection $X_i \hookrightarrow Y$. Give an example to show that there need not be an injection $(\bigcup_{i \in I} X_i) \hookrightarrow Y$. But what if the X_i are nested? [That is, $(\forall i, j \in I)(X_i \subseteq X_j \vee X_j \subseteq X_i)$.]

(ix)

Prove that every ordinal of the form ω^α is **indecomposible**: $\gamma + \beta = \omega^\alpha \;\; \to \;\; \gamma = \omega^\alpha \;\vee\; \beta = \omega^\alpha$.

(x)

Show that an arbitrary intersection of transitive relations is transitive. The **transitive closure** R^* (sometimes written '$tr(R)$') is the $\subseteq$-least transitive relation $\supseteq R$.

Let $\langle X, R \rangle$ be a wellfounded binary structure, with rank function ρ. Prove that $(\forall x \in X)(\forall \alpha < \rho(x))(\exists y)(\rho(y) = \alpha)$.

[A later—perhaps preferable—version of this question...

Let $\langle X, R \rangle$ be a wellfounded binary structure, with rank function ρ. Prove that $(\forall x \in X)(\forall \alpha < \rho(x))(\exists y \in X)(\rho(y) = \alpha)$.]

(xi)

Let $\{X_i : i \in \mathbb{N}\}$ be a nested family of sets of ordinals.

(a) Give an example to show that the order type of $\bigcup_{i \in \mathbb{N}} X_i$ need not be the sup of the order types of the X_i.

(b) What condition do you need to put on the inclusion relation between the X_i to ensure that the order type of $\bigcup_{i \in \mathbb{N}} X_i$ is the sup of the order types of the X_i?

(c) Show that the ordered set of the rationals can be obtained as the union of a suitably chosen ω-chain of some of its finite subsets.

(xii)

Using the uniqueness of subtraction for ordinals, and the division algorithm for normal functions, show that every ordinal can be expressed uniquely as a sum

$$\omega^{\alpha_1} \cdot a_1 + \omega^{\alpha_2} \cdot a_2 + \cdots \omega^{\alpha_n} \cdot a_n$$

where all the a_i are finite, and where the α_i are strictly decreasing.

(xiii)

Let f be a function from countable [nonzero] limit ordinals to countable ordinals satisfying $f(\alpha) < \alpha$ for all (countable limit) α. (f is "*pressing-down*".) Can f be injective?

Set Theory and Logic, Michaelmas 2016, Sheet 2: Posets

‘+’ signifies a question you shouldn’t have trouble with; ‘☠’ means what you think it means.

(i)

(a) For $n \in \mathbb{N}$, how many antisymmetrical binary relations are there on a set of cardinality n? How many binary relations satisfying *trichotomy*: $(\forall xy)(R(x,y) \vee R(y,x) \vee x = y)$? How are your two answers related?

(b) How many *symmetric* relations and how many *antisymmetric trichotomous* relations are there on a set of cardinality n? How are your two answers related?

(c) Contrast (a) and (b)

(ii)

Consider the set of equivalence relations on a fixed set, partially ordered by $\subseteq$. Show that it is a lattice. Must it be distributive? Is it complete?

(iii)

Cardinals: Recall that $\alpha \cdot \beta$ is $|A \times B|$ where $|A| = \alpha$ and $|B| = \beta$. Show that a union of α disjoint sets each of size β has size $\alpha \cdot \beta$. Explain your use of AC.

(iv)

Let $\langle A, \leq\rangle$ and $\langle B, \leq\rangle$ be total orderings with $\langle A, \leq\rangle$ isomorphic to an initial segment of $\langle B, \leq\rangle$ and $\langle B, \leq\rangle$ isomorphic to a terminal segment of $\langle A, \leq\rangle$. Show that $\langle A, \leq\rangle$ and $\langle B, \leq\rangle$ are isomorphic.

(v)

(Mathematics Tripos Part II 2001:B2:11b, modified).

Let U be an arbitrary set and $\mathcal{P}(U)$ be the power set of U. For X a subset of $\mathcal{P}(U)$, the **dual** $X^\vee$ of X is the set $\{y \subseteq U : (\forall x \in X)(y \cap x \neq \emptyset)\}$.

1. Is the function $X \mapsto X^\vee$ monotone? Comment.
2. By considering the poset of those subsets of $\mathcal{P}(U)$ that are subsets of their duals, or otherwise, show that there are sets $X \subseteq \mathcal{P}(U)$ with $X = X^\vee$.
3. $X^{\vee\vee}$ is clearly a superset of X, in that it contains every superset of every member of X. What about the reverse inclusion? That is, do we have $Y \in X^{\vee\vee} \to (\exists Z \in X)(Z \subseteq Y)$?
4. Is $A^{\vee\vee\vee}$ always equal to $A^\vee$?

(vi)

Use Zorn’s Lemma to prove that

(i) every partial ordering on a set X can be extended to a total ordering of X;

(ii) for any two sets A and B, there exists either an injection $A \hookrightarrow B$ or an injection $B \hookrightarrow A$.

(vii)

(Tripos IIA 1998 p 10 q 7)

Let $\langle P, \leq_P \rangle$ be a chain-complete poset with a least element, and $f : P \to P$ an order-preserving map. Show that the set of fixed points of f has a least element and is chain-complete in the ordering it inherits from P. Deduce that if $f_1, f_2, \ldots, f_n$ are order-preserving maps $P \to P$ which commute with each other (i.e. $f_i \circ f_j = f_j \circ f_i$ for all i, j), then they have a common fixed point. Show by an example that two order-preserving maps $P \to P$ which do not commute with each other need not have a common fixed point.

(viii)

$\mathbb{N} \rightharpoonup \mathbb{N}$ is the set of partial functions from $\mathbb{N}$ to $\mathbb{N}$, thought of as sets of ordered pairs and partially ordered by $\subseteq$.

Is it complete? Directed-complete? Separative? Which fixed point theorems are applicable?

For each of the following functions $\Phi : (\mathbb{N} \rightharpoonup \mathbb{N}) \to (\mathbb{N} \rightharpoonup \mathbb{N})$, determine (a) whether Φ is order-preserving, and (b) whether it has a fixed point:

(i) $\Phi(f)(n) = f(n) + 1$ if $f(n)$ is defined, undefined otherwise.

(ii) $\Phi(f)(n) = f(n) + 1$ if $f(n)$ is defined, $\Phi(f)(n) = 0$ otherwise.

(iii) $\Phi(f)(n) = f(n-1) + 1$ if $f(n-1)$ is defined, $\Phi(f)(n) = 0$ otherwise.

(ix)

Players `I` and `II` alternately pick elements (`I` plays first) from a set A (repetitions allowed: A does not get used up) thereby jointly constructing an element s of A^ω, the set of ω-sequences from A. Every subset $X \subseteq A^\omega$ defines a game $G(X)$ which is won by player `I` if $s \in X$ and by `II` otherwise. Give A the discrete topology and A^ω the product topology.

By considering the poset of partial functions $A^{<\omega} \to \{\mathtt{I}\}$ ($A^{<\omega}$ is the set of finite sequences from A) or otherwise prove that if X is open then one of the two players must have a winning strategy.

(x) 🐖

$\mathbb{R} = \langle 0, 1, +\times, \leq \rangle$ is a field. Consider the product $\mathbb{R}^{\mathbb{N}}$ of countably many copies thereof, with operations defined pointwise. Let $\mathcal{U}$ be an ultrafilter $\subseteq \mathcal{P}(\mathbb{N})$ and consider $\mathbb{R}^{\mathbb{N}}/\mathcal{U}$. Prove that it is a field. Is it archimedean?

(xi)

(i)$^+$ How many order-preserving injections $\mathbb{R} \to \mathbb{R}$ are there?

(ii)☠ Let $\langle X, \leq_X \rangle$ be a total order with no nontrivial order-preserving injection $X \to X$. Must X be finite?

Set Theory and Logic, Michaelmas 2016, Sheet 3: Propositional and Predicate Logic

(i)

Show how $\wedge$, $\vee$ and $\neg$ can each be defined in terms of $\to$ and $\bot$. Why can you not define $\wedge$ in terms of $\vee$? Can you define $\vee$ in terms of $\to$? Can you define $\wedge$ in terms of $\to$ and $\vee$?

(ii)

(a) Show that for every countable set A of propositions there is an independent set B of propositions with the same deductive consequences.

(b) If A is finite show that we can find such a B with $B \subseteq A$.

(c) Give an example to show that we should not expect $B \subseteq A$ if A is infinite.

(d) Show that if A is an infinite independent set of propositions then there is no finite set with the same deductive consequences.

(iii)

Explain briefly the relation between truth-tables and Disjunctive Normal Form.

Explain briefly why every propositional formula is equivalent both to a formula in CNF and to a formula in DNF.

Establish that the class of all propositional tautologies is the maximal propositional logic in the sense that any superset of it that is a propositional logic (closed under $\models$ and substitution) is trivial (contains all well-formed formulæ).

(iv)

A formula (of first-order Logic) is in **Prenex Normal Form** if the quantifiers have been “pulled to the front”—every propositional connective and every atomic subformula is within the scope of every quantifier.

> Explain briefly why every first-order formula is equivalent to one in PNF.
>
> Axiomatise the theory of groups in a signature with ‘$=$’ and a single three-place relation “x times y is z”. Put your axioms into PNF. What are the quantifier prefixes?
>
> Find a signature for Group Theory which ensures that every substructure of a group is a semigroup-with-$\mathbb{1}$.

(v)

Show that the theory of equality plus one wellfounded relation is not axiomatisable.

(vi)

Write down axioms for a first-order theory T with equality plus a single one-place function symbol f that says that f is bijective and that for no n and no x do we have $f^n(x) = x$.

(a) Is T finitely axiomatisable?

(b) How many countable models does T have (up to isomorphism)?

(c) How many models of cardinality of the continuum does it have (up to isomorphism)? (You may assume that the continuum is not the union of fewer than $2^{\aleph_0}$ countable sets, a fact whose proof—were you to attempt it—would need AC.)

(d) Let κ be an uncountable aleph. How many models does T have of size κ?

(e) Is T complete?

(vii)

Show that monadic predicate logic (one place predicate letters only, without equality and no function symbols) is decidable.

(viii) 🐖

(a)$^+$ Suppose A is a propositional formula and 'p' is a letter appearing in A. Explain how to find formulæ A_1 and A_2 not containing 'p' such that A is logically equivalent to $(A_1 \wedge p) \vee (A_2 \wedge \neg p)$.

(b) Hence or otherwise establish that, for any two propositional formulæ A and B with $A \models B$, there is a formula C, containing only those propositional letters common to both A and B, such that $A \models C$ and $C \models B$. (Hint: for the base case of the induction on the size of the common vocabulary you will need to think about expressions over the empty vocabulary).

(ix)

Why does T not follow from K and S?

Show that Peirce's Law: $((A \to B) \to A) \to A$ cannot be deduced from K and S.

(x$^+$)

Look up *monophyletic*. Using only the auxiliary relation "is descended from" give a definition in first-order logic of what it is for a set of lifeforms to be monophyletic.

(xi)

Is

$$(\forall x)(\exists y)(F(x, y)) \to (\forall x)(\exists y)(\forall x')(\exists y')[F(x, y) \wedge F(x', y') \wedge (x = x' \to y = y')]$$

valid?

(xii)

(a) Show that the theory of fields of characteristic zero is (first-order) axiomatisable but not finitely axiomatisable. Show that the theory of fields of finite characteristic is not first-order axiomatisable.

(b) Recall that a simple group is one with no nontrivial normal subgroup. Is the theory of simple groups first order?

(c) A local ring is a ring with a unique maximal ideal. Is the theory of local rings first-order? [Hint: what might the unique maximal ideal be?]

(d) Is the theory of posets in which every element belongs to a unique maximal antichain first-order?

(e) A theory T is **equational** iff every axiom of T is of the form $(\forall\vec{x})\Phi$ where ϕ is a conjunction of equations between T-terms.

Prove that, if T is equational, then a pointwise product of models of T is another model of T, and substructures and homomorphic images of models of T are models of T.

Which of the theories in (a)–(d) are equational?

(xiii)

A *type* in a propositional language $\mathcal{L}$ is a countably infinite set of formulæ.

For T an $\mathcal{L}$-theory a *T-valuation* is an $\mathcal{L}$-valuation that satisfies T. A valuation v *realises* a type Σ if v satisfies every $\sigma \in \Sigma$. Otherwise v *omits* Σ. We say a theory T *locally omits* a type Σ if, whenever ϕ is a formula such that T proves $\phi \to \sigma$ for every $\sigma \in \Sigma$, then $T \vdash \neg\phi$.

(a) Prove the following:

Let T be a propositional theory, and $\Sigma \subseteq \mathcal{L}(T)$ a type. If T locally omits Σ then there is a T-valuation omitting Σ.

(b) Prove the following

Let T be a propositional theory and, for each $i \in \mathbb{N}$, let $\Sigma_i \subseteq \mathcal{L}(T)$ be a type. If T locally omits every Σ_i then there is a T-valuation omitting all of the Σ_i.

(xiv)

Prove that, for every formula ϕ in CNF, there is a formula ϕ' which

(i) is satisfiable iff ϕ is;

(ii) is in CNF where every conjunct contains at most three disjuncts.

(Hint: there is no assumption that $\mathcal{L}(\phi') = \mathcal{L}(\phi)$.)

Set Theory and Logic, Michaelmas 2016, Sheet 4: More Predicate Logic and Some Set Theory

(i)$^+$

Show that if x is a transitive set, then so are $\bigcup x$ and $\mathcal{P}(x)$. Are the converses true?

(ii)$^+$

Show that the Pair-set axiom is deducible from the axioms of empty set, power set, and replacement.

(iii)$^+$

Show that $\{z : \neg(\exists u_1, \ldots, u_n)((z \in u_1) \wedge (u_1 \in u_2) \wedge \cdots \wedge (u_n \in z))\}$ is not a set for any n. What assumptions have you made?

(iv)

Write down sentences in the language of set theory to express the assertions that, for any two sets x and y, the product $x \times y$ and the set y^x of all functions from x to y exist. You may assume that your pairs are Wiener-Kuratowski.

Which axioms of set theory are you going to have to assume if these assertions are to be provable?

(v)

(a) Prove that every normal function $On \to On$ has a fixed point.
(b) Prove that the function enumerating the fixed points of a normal function $On \to On$ is itself normal.
(c) If α is a regular ordinal and f is a normal function show that f has a fixed point of cofinality α.
(d) Are any of your fixed points regular?

(vi)

Show that the axiom of choice follows from the assumption that cardinals are totally ordered by $\leq_{card}$.

(vii)

Explain briefly the equivalence of the four versions of the axiom of foundation given in lectures: (i) The axiom scheme of $\in$-induction; (ii) The assertion that every set is wellfounded; (iii) Axiom of Regularity; (iv) Every set belongs to the cumulative hierarchy.

(viii)

f is an $\in$-automorphism if f is a permutation of V that preserves $\in$: $x \in y \longleftrightarrow f(x) \in f(y)$. Show that a model of ZF (with foundation of course) can have no nontrivial $\in$-automorphisms.

Give an example to show that the surjectivity condition on f is necessary; that is to say, there are non-trivial injective $\in$-homomorphisms.

(ix)

For the Wiener-Kuratowski ordered pair $\rho(\langle x, y\rangle) = \max(\rho(x), \rho(y)) + 2$. ($\rho$ is set-theoretic rank.)

(a) Can you define a ordered pair such that $\rho(\langle x, y\rangle) = \max(\rho(x), \rho(y)) - 1$?
(b) Can you define a ordered pair such that $\rho(\langle x, y\rangle) = \max(\rho(x), \rho(y)) + 1$?
(c)* Can you define a ordered pair such that $\rho(\langle x, y\rangle) = \max(\rho(x), \rho(y))$ for all but finitely many x and y?

(x)

There are various ways of constructing implementations (as sets) of $\mathbb{Q}$, $\mathbb{Z}$, $\mathbb{R}$ and $\mathbb{C}$ from an implementation (as sets) of the naturals. For one of these constructions compute the ranks of the sets that have the rôles of $\mathbb{Q}$, $\mathbb{Z}$, $\mathbb{R}$ and $\mathbb{C}$.

Different implementations will almost certainly give you different answers. Are there any lower or upper bounds on the answers you might get?

(xi)

Consider the binary relation E on $\mathbb{N}$ defined by: $n \, E \, m$ iff the nth bit (counting from the right, starting at 0) in the binary expansion of m is 1. What can you say about the structure $\langle \mathbb{N}, E\rangle$?

(xii)

Prove that, for each $n \in \mathbb{N}$, there is a set of size $\aleph_n$. Is there a set of size $\aleph_\omega$?

(xiii)

Assume that the cartesian product $x \times y$ always exists however you implement ordered pairs. Infer the axiom scheme of replacement.

(xiv)

Assume that every normal function $On \to On$ has a regular fixed point. Consider the function that enumerates the initial ordinals and deduce that there is a "weak inaccessible" κ. Which axioms of ZF hold in V_κ?

(xv)

Suppose $\{A_i : i \in I\}$ and $\{B_i : i \in I\}$ are families of sets such that for no $i \in I$ is there is a surjection $A_i \twoheadrightarrow B_i$. Show that there is no surjection $\bigcup_{i \in I} A_i \twoheadrightarrow \prod_{i \in I} B_i$.

You will need the axiom of choice. Is there a converse?

Using these ideas you can show that $\aleph_\omega \neq 2^{\aleph_0}$ *without* using AC.

Sheets from here on are still under construction!!

Set Theory and Logic, Michaelmas 2016, Sheet 5

Sheet 5 is for Part II enthusiasts who want to take this stuff further; it's a mixture of revision, consolidation and looking-ahead. It is also for Part III students who want something to get them used to what they are going to be doing later in the year. (Part III Logic is lectured in Lent.)

(i)

Explain to your supervision partner (or to anyone listening who might be confused) the difference between

(i) Nonstandard naturals
(ii) Countable ordinals
(iii) Infinite Dedekind-finite cardinals

(ii)

For P a poset, let P^* be the poset of chains-in-P partially ordered by end-extension. (Chains are allowed to be empty). Show that there is no injective homomorphism $P^* \hookrightarrow P$.

(iii)

Any two countable dense linear orders without endpoints are isomorphic. Give an illustration to show how your back-and-forth construction might not work for dense linear orders of size $\aleph_1$. How do you have to spice up the denseness condition to prove an analogous result for linear orders of size $\aleph_1$?

(iv)

(For those of you who did Languages and Automata)

A wellordering of $\mathbb{N}$ is *recursive* iff its graph (subset of $\mathbb{N} \times \mathbb{N}$) is decidable ("recursive"); an ordinal is *recursive* iff it is the order type of a decidable ("recursive") wellordering of $\mathbb{N}$. Which of the countable ordinals you have learned to know and love are recursive? Come to think of it, are *all* countable ordinals recursive?

(v)*

(For those of you who did Languages and Automata)

Prove Trakhtenbrot's theorem that if S is a signature with equality and at least one binary relation symbol then the set of S-sentences true in all finite structures is not semidecidable ("r.e.").

(vi)

(A taster for *forcing*)

A poset $\langle P, \leq_P \rangle$ is [upwards] separative if $(\forall x, y \in P)(x \not\leq y \to (\exists z \geq y)(\forall w)(w \not\geq z \vee w \not\geq x))$

For each of the following posets say whether or not it is (i) separative (ii) directed (iii) chain-complete.

The set of finite sequences of countable ordinals (thought of as sets of ordered pairs) partially ordered by $\subseteq$.

The set $\{f : f$ is an injection from some set of countable ordinals $\hookrightarrow \mathbb{R}\}$ ordered by $\subseteq$. Think of f as a set of ordered pairs.

(vii)

(For those of you who did some graph theory in Lent term)

Using propositional logic only, show that a(n undirected) graph and its complement cannot both be disconnected. (Hint: propositional letters will correspond to edges)

(viii)

A poset $\langle P, \leq \rangle$ is called *downwards separative* if for all $x \not\leq y$ there is $z \leq x$ with z incompatible with y. ("incompatible" means "have no lower bound"). We say that a poset is *downwards splitting* if for every x there are y and z such that $y, z \leq x$, and y and z are incompatible.

(a) Show that not every downwards separative poset is downwards splitting.

(b) Show that if a poset has no minimal elements and is downwards separative, then it is downwards splitting.

A set $D \subseteq P$ is called *downwards dense* if for every p in P there is a d in D such that $d \leq p$.

Suppose XX is a collection of subsets of P. We say that $G \subseteq P$ is XX-*generic* if G has nonempty intersection with every downwards dense element of XX.

We say that G is a *filter* if

1. for any x, y in G there is z in G such that $z \leq x$ and $z \leq y$, and
2. for any x in G and $x \leq y$, we have y in G.

> (c) If XX is countable, show that there is an XX-generic filter.
>
> (d) Let $\langle P, \leq \rangle$ be a downwards separative poset with no minimal elements and let XX be a collection of subsets of P closed under complementation (i.e., if $X \in XX$, then also $P \setminus X \in XX$). Show that if G is an XX-generic filter, then $G \notin XX$.

(e) Let $\langle P, \leq \rangle$ be the set of finite sequences of zeros and ones, ordered by reverse inclusion. Show that this is a downwards separative poset without minimal elements.

(f) Let XX be the collection of recursive sets of finite sequences of zeros and ones. Show, using (c), (d), and (e), that there is a non-recursive such set.

(ix)

(Concrete constructions of limits in ZF)

Let $\langle I, \leq_I \rangle$ be a directed poset and, for each $i \in I$, let A_i be a set and, for all $i \leq_I j$, let $\sigma_{i,j} : A_i \hookrightarrow A_j$ be an injection, and let the injections commute.

Show that there is a set A_I with, for each $i \in I$, an injection $\sigma_i : A_i \hookrightarrow A_I$ and the $\sigma_{i,j}$ commute with the σ_i.

Show also that A_I is minimal in the sense that if B is any set such that for each $i \in I$ there is an injection $\tau_i : A_i \hookrightarrow B$ and the τ_i commute with the $\sigma_{i,j}$, then there is a map $A_I \hookrightarrow B$.

Let $\langle I, \leq_I \rangle$ be a directed poset and, for each $i \in I$, let A_i be a set and, for all $i \leq_I j$, let $\sigma_{j,i} : A_j \twoheadrightarrow A_i$ be a surjection, and let the surjections commute.

Show that there is a set A_I with, for each $i \in I$, a surjection $\pi_i : A_I \twoheadrightarrow A_i$.

Show also that A_I is minimal in the sense that, if B is any set such that for each $i \in I$ there is a surjection $\tau_i : B \twoheadrightarrow A_i$ and the τ_i commute with the $\sigma_{i,j}$, then there is a map $B \twoheadrightarrow A_I$.

(x)☠

Let G be the alternating group of permutations of V_ω. For each $n \in \mathbb{N}$ its members can move x by permuting those elements of $\bigcup^n x$ that are of finite rank and fixing the remainder. A set that is fixed by everything in G under the nth action of G is said to be n-**symmetric**; if it is n-symmetric for all sufficiently large n it is just plain **symmetric**.

Consider the collection of sets that are hereditarily symmetric. Which axioms of ZFC are true in this structure?

(xi)* (A taster for Large Cardinals)

Prove Łoś's theorem :

Theorem 19 *Let $\mathcal{U}$ be an ultrafilter $\subseteq \mathcal{P}(I)$. For all first-order expressions ϕ,*

$$(\prod_{i \in I} \mathcal{A}_i)/\mathcal{U} \models \phi \text{ iff } \{i : \mathcal{A}_i \models \phi\} \in \mathcal{U}.$$

(You may assume AC)

Suppose there is a set K with a nonprincipal ultrafilter $\mathcal{U} \subseteq \mathcal{P}(K)$ that is closed under countable intersections. By using Scott's trick concretise the elements of the ultrapower $V^K/\mathcal{U}$. Prove that it is wellfounded. What can you say about the Mostowski collapse?

(xii)* 🐖

$\mathcal{I} = \langle I, \leq_{\mathcal{I}} \rangle$ is a **set of indiscernibles** for a model $\mathfrak{M}$ for a language $\mathcal{L}$ iff $\leq_{\mathcal{I}}$ is a total order, and for all $\phi \in \mathcal{L}$, if ϕ is a formula with n free variables in it then for all distinct n-tuples $\vec{x}$ and $\vec{y}$ from $\mathcal{I}$ **taken in $\leq_{\mathcal{I}}$-increasing order** we have $\mathfrak{M} \models \phi(\vec{x}) \longleftrightarrow \phi(\vec{y})$.

Now let $\mathcal{I}$ be a total order, T a theory with infinite models and a formula $P()$ with one free variable s.t. T thinks that the extension of P is an infinite total order. Then T has a model $\mathfrak{M}$ in which $\mathcal{I}$ is embedded in (the interpretation of) P as a set of indiscernibles.

(Notice that there is no suggestion that the copy of $\mathcal{I}$ in $\mathfrak{M}$ is a set of $\mathfrak{M}$, or is in any way definable.)

It is comparatively straightforward, given $\mathcal{I}$ and T and $P()$, to find $\mathfrak{M}$ as in the theorem if we do not ask that I should be embedded as a set of indiscernibles: compactness does the trick. To get the set of indiscernibles you need to use Ramsey's theorem from Graph theory.

(xiii)

(GRM revision from a logical point of view). Wikipædia says:

> Commutative Rings $\supseteq$ Integral Domains $\supseteq$ Integrally Closed Domains $\supseteq$ GCD domains $\supseteq$ Unique Factorization Domains $\supseteq$ Principal Ideal Domains $\supseteq$ Euclidean Domains $\supseteq$ Fields

All these families-of-structures can be thought of as belonging to the one signature: 0, $\mathbb{1}$, $+$, $\cdot$ and $-$. Which of them are first-order axiomatisable? In each case provide axiomatisations or explain why there are none. Identify the quantifier complexity of the axiomatisations you find.

(xiv)(☠)

How many countable [linear] order types are there whose automorphism group is transitive on singletons?

(xv)

How many transitive subsets of V_ω are there?

How many transitive sets are there all of whose members are countable?

(xvi)

Recall the difference between a **wellorderable** set and a **wellordered** set.

Prove without using AC or foundation or ordinals that every set of wellorderable sets has a member that injects into all the others.

Is this the same as saying that the collection of alephs is wellordered by the order relation on cardinals?

(xvii)

A directed limit of wellfounded structures under end-extension is wellfounded.