713: Eighteen Lectures for Auckland second Semester 2025

Thomas Forster

January 15, 2025

Contents

6	1	Thr	ures on Ordinals and Cardinals	7				
7		1.1	Ordina	als	7			
8		1.2	Wellfo	oundedness	8			
9	2	2 Posets and Fixed-Point Theorems						
10		2.1	Lattice	es	19			
11		2.2 Separative poset – not examinable						
12		2.3	Fun with Complete Posets					
13			2.3.1	Topology	21			
14			2.3.2	Knaster-Tarski: Fixed Points, Inductively-defined Sets and Canto	r-			
15				Bernstein–like theorems	22			
16			2.3.3	Inductively defined Sets	23			
17			2.3.4	Cantor-Bernstein–like theorems	24			
18	3	Thr	ee Lect	ures on Propositional Logic	25			
19		3.1	if-ther	1	25			
20			Boole	an Algebras detect propositional Tautologies	34			
21				cations of Propositional Compactness	35			
22			3.3.1	The Interpolation Lemma	36			
23		3.4	CNF a	and DNF	38			
24			3.4.1	Resolution Theorem Proving	39			
25	4	One	ne lecture on The Axiom of Choice					
26		4.1	Weak	versions: countable choice, and a classic application thereof; DC	42			
27		4.2	**					
28	5	Six	lectures	on First order Logic and the Compactness Theorem	45			
29		5.1	•		45			
30			5.1.1	Constants and variables	45			
31			5.1.2	Predicate letters	46			
32			5.1.3	Quantifiers	46			
33			5.1.4	Function letters	47			
34			5.1.5	Quantifiers	47			
35			5.1.6	Higher-order	47			

4 CONTENTS

36			5.1.7 Signatures			
37		5.2	Axioms of LPC			
38		5.3	Semantics			
39		5.4	Completeness theorem for LPC: the set of valid sentences is semide-			
40			cidable			
41			5.4.1 ∈-terms			
42		5.5	Decidability			
43	6	Two	Lectures on Set Theory 57			
44	Ü	6.1	Transitive Closures and Transitive Sets			
45		6.2	The Cumulative Hierarchy			
46		6.3	Scott's Trick			
47		6.4	The Axiom of Foundation			
48		6.5	Mostowski Collapse			
49		6.6	Ordinals again			
50		0.0	6.6.1 Initial Ordinals			
51		6.7	$\aleph^2 = \aleph$			
52		6.8	Independence of the Axioms from each other			
53		0.0	6.8.1 Δ_0 formulae and the Lévy Hierarchy			
54			6.8.2 Some actual Independence Results			
55			6.8.3 Independence of Sumset			
56			6.8.4 Independence of the Axiom of Foundation			
57			6.8.5 Independence of the Axiom of Choice			
58		6.9	The Modern Theory of Wellfounded Sets			
59	7		Lectures on Model Theory 85			
60		7.1	The Skolem-Löwenheim Theorems 85			
61		7.2	Categoricity			
62			7.2.1 Back and forth			
63		7.3	Results related to completeness, exploiting completeness 88			
64			7.3.1 Prenex Normal Form and Quantifier-Counting 88			
65		7.4	Omitting Types			
66		7.5	Direct Products and Reduced Products			
67			7.5.1 Intersection-closed properties 91			
68			7.5.2 Reduced products			
69		7.6	Ultraproducts and Łoś's theorem			
70	8	Exai	mple Sheets 97			
71		8.1	Sheet 0: Countability			
72		8.2	Sheet 1: Mainly Ordinals			
73		8.3	Sheet 2: Posets			
74		8.4	Sheet 3: Propositional and Predicate Logic			
75		8.5	Sheet 4: More Predicate Logic and Some Set Theory 103			
76		8.6	Sheet 5			
77		We have a lot to get through in these 18 lectures; there is not only the material that				
78	is specified in the syllabus, there is also the background on which that material relie					

5 **CONTENTS**

and which may not have been earlier covered in a way that meets your current needs. Inevitably every lecturer tries to smuggle in extra material on top of what has been stipulated. I am no exception; however I have tried to flag such material for what it is: the final example sheet (which is not examinable) consists entirely of such stuff. 82

I am going to write out in full quite a lot of what you might otherwise expect to have to copy down off my slides or blackboard. I am going to make these notes available to you so that you can spend the lectures listening and thinking rather than mechanically transcribing.

I am going to assume that you have mastered all the material in Maths 315. If you are having difficulties with it, contact me, and i'll see what I can do to make your life easier.

18 lectures over six weeks – weeks 1-6.

90 92

93

97

84

86

89

The rubric says:

1. Zorn's Lemma; 2. Propositional logic and the compactness theorem (includes applications of the compactness theorem to mathematics); 3. The Axioms of set theory; 4. Ordinals; 5. Cardinals; 6. \mathbb{N} , \mathbb{Z} , \mathbb{Q} and \mathbb{R} ; 7. First order logic and the compactness theorem; 8. First order theories; 9. Cardinality of models and elementary submodels.

However I shall not follow that order; I shall do them in the order:

Ordinals (and cardinals); three lectures Propositional Logic: Boolean algebras; three lectures Zorn, Prime ideal theorem; two lectures First order logic and the compactness theorem; three lectures First order theories; two lectures Set Theory; three lectures $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ and \mathbb{R} and their implementation into Set Theory; one lecture

Some Basic Model Theory: Countable Categoricity;

hrenfeucht-Mostowski; Ło s's theorem;

one lecture

The treatment of ordinals will use ordinals forwelly.pdf;

The treatment of propositional Logic will follow logiclectures 2016.tex; 100

The treatment of set theory will use axiomsofsettheory.pdf and ACpedagogy.pdf.

101 102

103

104

105

106

107

108

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 Toad has just discovered Set Theory – i mean motor cars.

Warning! If you are reading this and your name isn't 'Thomas Forster' then you 109 are eavesdropping; these notes are my messages to myself and are made available to 110

6 CONTENTS

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. Reading these notes is not a substitute for attending the lectures: it is an adjunct to them.

Chapter 1

Three Lectures on Ordinals and Cardinals

(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.)

1.1 Ordinals

126

127

128

129

131 132

134

135

136

138

139

141

142

143

144

145

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. (There is always a *next* stage). 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.

```
\omega, \omega + n, \omega + \omega, \omega \cdot n, \omega \cdot \omega, \omega^n.
```

Ordinals are also the order types (isomorphism classes) 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 – is the same set, indeed. The ordinal we get is $\omega + 1$, 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.

149

150

151

153

155

156

157

46 Congruence relations

You will be familiar with plenty of congruence relations and of their importance. However they might never have been identified for you as an object of study, so you might not know the expression.

Congruence relations are a special kind of equivalence relation. Every equivalence relation on a set A has a quotient – the set of equivalence classes. An equivalence relation on A is said to be a *congruence relation for* for some operation $f^n: A \to A$ if that f gives rise to an operation on the quotient, often called f too.

The *locus classicus* is the equivalence relation congruence-mod-p on \mathbb{Z} , which is a congruence relation for + and \times .

Cardinals are very simple. Read my countability notes [1] Multiplication, addition and exponentiation.

1.2 Wellfoundedness

Wellfoundedness is a very important property that a binary relation might have. Ordinals are the isomorphism classes of wellorderings, and a wellordering is a wellfounded total order, so there's one reason for paying attention to wellfoundedness. Another reason is that wellfounded relations support a kind of induction, and this kind of induction is quite important generally in Pure Mathematics, and specifically in Set Theory. Let us not delay our encounter with it any longer.

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

165

166

168

170

172

174

176

177

178

179

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

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 = 00. Finally, the inference we are trying to draw is this: if x = 01 has y = 02 whenever the induction hypothesis is satisfied, then everything has y = 03. When can we do this? We must try to identify some condition on x = 03 that is equivalent to the assertion that this is a legitimate inference to draw in general (i.e., for any predicate y1).

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

183

184

185

186

187

188

189

194

195

196

197

199

200

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

$$\frac{(\forall y)(R(y,x) \to \psi(y)) \to \psi(x)}{(\forall x)(\psi(x))}$$
 R-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 1

R is **well-founded** iff for every nonempty subset A of dom(R()) we have

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

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

A wellfounded binary structure is a pair $\langle X, R \rangle$ where R is a wellfounded relation $\subseteq R^2$

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*.

A fact that is basic but worth recording is:

REMARK 1 If $\langle X, R \rangle$ is a well-founded binary structure and is a homomorphic image of $\langle Y, S \rangle$ then $\langle Y, S \rangle$, too, is a well-founded binary structure.

Perhaps leave this as an exercise? A homomorphic image of a bad subset (one without a minimal element) also lacks a minimal element.

THEOREM 1 Wellfounded induction: recursion on wellfounded relations

210

211

212

215

217

219

221

223

224

226

227

231

232

233

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

Answer:

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.

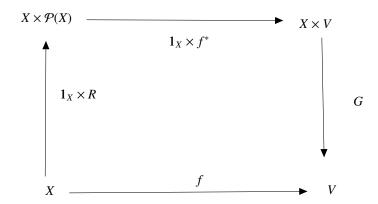
So: suppose the induction hypothesis holds for all y s.t. R(y, x).

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.



- 238 **DEFINITION 2** Wellordering a wellfounded strict total order
- "every terminal segment has a least element" is equivalent. It's the "always an immediate next stage" condition.
- COROLLARY 1 Principle of induction for wellorderings
- 42 COROLLARY 2 Definition by recursion for wellorderings
- **DEFINITION 3** Ordinals are isomorphism types of wellorderings.

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.

250 Proof:

245

246

247

248

249

- 1. The automorphism group of a total order can have no nontrivial finite cycles every nontrivial cycle looks like \mathbb{Z} and can have no least element. If τ is an automorphism of a wellordering consider $\{\tau^n(x): n \in \mathbb{Z}\}$. What is its least element?
- 255 256 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 $\langle B, <_B \rangle$.

260

261

262

264

265

266

287

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 $f(a)=\sup\{f(a'):a'<_Aa\}$ and $g:B\to A$ mutatis mutandis. We prove by wellfounded induction that $f\cdot g$ is the identity wherever 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 f the first thing not in the domain of f and f should have been in f and f should have been in f.

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

(3) tells us that the obvious order relation on ordinals is a total order. This is a nontrivial fact. The order relation on cardinals is not obviously a total order. We will talk about the order relation on cardinals later.

```
DEFINITION 4 \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 (equivalently):

"\langle Y, \leq_Y \rangle is an initial segment of \langle X, \leq_X \rangle"
```

Informally: "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.

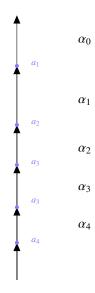
The suborder inherits totality and wellfoundedness so is a wellorder. Apply theorem 2.
You might like to try Sheet 1 q 4 at this point.

Notice that this is not true of arbitrary total orders: \mathbb{N} is a subordering of \mathbb{Z} but is not iso to an initial segment.

DEFINITION 5 $\alpha \leq_{On} \beta$ 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*) 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.

I have drawn the picture as if the index set I were \mathbb{N} and the sequence is strictly descending. We don't actually need this assumption but it does make the picture easier to draw. The assumption can be justified by appeal to a principle called DC: the principle of dependent choice of which more in section 4.1.

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. (Not even antisymmetrical – think of (0, 1) and [0, 1].

Beware! Some textbooks contain theorems with statements that sound like theorem 3 but are actually much weaker. Later we will see an implementation of ordinals into Set Theory, due to Von Neumann, and the textbooks contain proofs that the order relation on Von Neumann ordinals is wellfounded. This is true for completely banal reasons: the order relation on von Neumann ordinals is actually \in – set membership. One of the axioms of the usual set theory (ZFC) is the axiom of foundation, which says that \in is wellfounded ... so *of course* the order relation on (von Neumann) ordinals is wellfounded. (In fact if we define the class of von Neumann ordinals as the least class containing \emptyset and closed under \bigcup and $x \mapsto x \cup \{x\}$ then we can prove that the order relation on Von Neumann ordinals is wellfounded without using the axiom of foundation.)

Hoswever, a proof that the order relation on von Neumann ordinals is wellfounded is not a proof that $\langle On, <_{On} \rangle$ is 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 UBUNTU but rather a fact about the binary for my machine, and that says nothing about the binary for your machine. One does not prove facts about abstract mathematical entities by reasoning about their implementations. Proving that von Neumann ordinals are wellordered by $<_{On}$ – it doesn't prove anything about ordinals at all: what it shows is that – to that extent at least – Von Neumann ordinals behave according to spec. In that sense it's a bit like the famous demonstration of Russell-and-Whitehead that the natural numbers in their system obey 1+1=2. It's not a fact about numbers, it's a fact about their system.

Nevertheless the policy of claiming that the order relation on ordinals is well-founded by appealing to their implementation as von Neumann ordinals looks like a sensible short-cut if you are proposing to do all your mathematics by implementing it in ZF. Since this is the chief context in which people study ordinals people can be forgiven for looking for a quick hacky way of getting one of the basics sorted. *Tout comprendre, c'est tout pardonner.*

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 an artefact of set-theoretic foundationalism. If we'd decided to think of ordinals in a less set-theoretic way we would've ended up with a different theorem. Here be dragons. Look at the appendix on Burali-Forti, section ??

DEFINITION 6

54 Preorderings are transitive and reflexive;

A preorder is a set equipped with a preordering.

A Partial ordering is an antisymmetric preordering.

We assume the reader is familiar with disjoint unions, products and lexicographic products of posets.

Products not just of posets: remember we do products of groups. or of rings....

DEFINITION 7 Addition and Multiplication of ordinals defined synthetically.

[omitted; done live at the board]

359

364

365

366

368

370

371

372

374

375

384

386

388

389

Uniqueness of ordinal subtraction. What might we mean by ' $\alpha - \beta$ '? If $\beta \le \alpha$ then whenever $\langle B, <_B \rangle$ belongs to β and $\langle A, <_A \rangle$ belongs to α then there is an isomorphism Explain f "x notation $\pi : \langle B, <_B \rangle$ to a unique initial segment of $\langle A, <_A \rangle$. The truncation

$$\langle A \setminus (\pi^{"}B), <_A \upharpoonright (A \setminus (\pi^{"}B)) \rangle$$

is our wellordering of length $\alpha - \beta$. This definition ensures that $\beta + (\alpha - \beta) = \alpha$.

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 8 cofinality; regular ordinal

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

Now might be a good time to attempt sheet 1 question 6 that says that every countable limit ordinal has cofinality ω .

DEFINITION 9 Recursive definition of addition, multiplication and exponentiation of ordinals.

```
378 \alpha + 0 = \alpha;

379 \alpha + (\beta + 1) = (\alpha + \beta) + 1, \text{ and}

380 \alpha + \sup(B) = \sup(\{\alpha + \beta : \beta \in B\}).

381 \alpha \cdot 0 = 0; \alpha \cdot (\beta + 1) = \alpha \cdot \beta + \alpha, \text{ and}

382 \alpha \cdot \sup(B) = \sup(\{\alpha \cdot \beta : \beta \in B\}).

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

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

Wellorderings of length ω^{ω} : The set of polynomials in one variable with nonnegative integer coefficients ordered by dominance...

... and ϵ_0 . A bit harder!

DEFINITION 10 Countable ordinal

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

391

392

393

395

396

397

398

400

401

403

404

405

409

410

411

412

413

415

417

418

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.

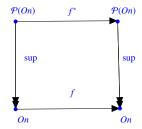
DEFINITION 11 Normal functions

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

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

The function that enumerates the members of a clubset in increasing order is normal.

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



" 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 $f: On \rightarrow On$ *is normal, and* α *is any ordinal, then there is* β *such that*

$$f(\beta) \le \alpha < f(\beta + 1).$$

407 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!

THEOREM 5 Hartogs' Lemma.

For every set X there is a wellordered set Y s.t. $Y \nleftrightarrow X$.

416 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 can wellorder it if we have AC of course, but we want to keep our assumptions to a minimum.

421

422

423

424

425

426

427

428

430

431

432

433

434

435

436

437

438

439

440

441

442

443

446

448

450

454

455

456

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.

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*. It's an expression not much used nowadays, but it's helpful and evocative, particularly when you bear in mind that \mathbb{N} , the set of finite ordinals (or cardinals, in the finite world they're the same thing) is the *first* number class¹. 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 67.

DEFINITION 12 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.

447 Proof: By coroll 1.2.

```
Now would be a good moment to attack Sheet 1 q 11
```

Hartogs' tells you ∃ enuff ordinals.

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 $\{y: (\forall x')(R(x, x') \to R(h(x'), y))\}$.

The rank function on wellfounded sets is parsimonious.

If you want to know more about ordinals read

www.dpmms.cam.ac.uk/~tf/ordinalsforwelly.pdf

¹It also sets you up to muse over what the third number class might be.

Chapter 2

Posets and Fixed-Point Theorems

- Poset you know; likewise toset; poset subsumes toset.
- Pointwise and lexicographic products of posets already done.

462 DEFINITION 13

- 463 A complete poset is one every subset of which has a least upper bound
- The expressions 'complete poset' and 'complete lattice' are used interchangeably.
- 465 A poset $\langle P, \leq \rangle$ is directed if $(\forall x, y)(\exists z)(x \leq z \land y \leq z)$.
- *A subset X of a poset* $\langle P, \leq \rangle$ *is directed if* $(\forall x, y \in X)(\exists z \in X)(x \leq z \land y \leq z)$.
- 467 A poset is

469

- directed-complete if every directed subset has a least upper bound.
 - chain-complete if every chain has a least upper bound.
- If X is a countable set then $\langle \mathcal{P}_{\aleph_1}(X), \subseteq \rangle$ is simply $\mathcal{P}(X)$, which is a complete poset. (See definition 43 for this \mathcal{P}_{κ} notation).
- Every complete poset is directed. If X is uncountable then $\langle \mathcal{P}_{\aleph_1}(X), \subseteq \rangle$ is directed but not complete.

2.1 Lattices

- We can define \leq from \wedge , \vee and =. In a distributive lattice \vee and \wedge distribute over
- each other, distributive: examples and non-examples. Subspaces of a vector space not
- distributive, nor is the lattice of partitions of a fixed set under refinement¹.

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

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. If it's injective it'll be a *scaling factor*. 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.

¹Someone 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 :

481

491

493

495

496

497

498 499

502

503

504

505

506

507

478 Complemented lattice.

Boolean Algebra

Complete lattice. Power sets and topologies.

The regular open sets form a complete poset that is actually a b.a.

482 Aside here to explain subalgebra.

2.2 Separative poset – not examinable

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 $x \le y$ then anything $\ge y$ is an upper bound for both. What we can say is that in all other cases – where x is **not** $\le y$ – there is $y' \ge y$ s.t. x and y' have no upper bound. Thus we say:

DEFINITION 14 $\langle X, \leq \rangle$ is separative iff

$$(\forall x,y \in X)(x \nleq y \to (\exists z \geq y)(\forall w)(w \ngeq z \vee w \ngeq x))$$

We will abbreviate $(\forall w)(w \not\geq z \lor w \not\geq x)$) ("z and x have no upper bouncd") as $z \perp x$. I think this notation is standard.

Actually that definition should probably be described in more detail as "upwardly separative". There is an analogous notion of *downwards separative*. Naturally they are interchangable in applications.

There are lots of separative posets and they matter: the idea of separative poset underlies the technique of *forcing* which Paul Cohen exploited to prove the independence of the Continuum Hypothesis from ZFC in 1963. All those posets are infinite.

Is the poset in this Hasse diagram separative?



2.3 Fun with Complete Posets

The Big Theorem concerning complete posets is Tarski-Knaster, theorem 6 below. It's more useful than you might at first guess, beco's complete posets are everywhere. The reals are a complete poset under $\leq_{\mathbb{R}}$; every power set is a complete poset under \subseteq ; every wellordering is a complete poset; in any topological space the open sets are a complete poset under \subseteq .

In fact there is quite a nice approach to various topological concepts (product topology, quotient topology, that sort of thing ...) through the idea of complete posets, which i will now show you.

2.3.1 Topology

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

543

544

546

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 \bigwedge operations of $\mathcal{P}(X)$ and of a topology on X are not the same (though the \bigvee operations are the same.)

We think of a topology as the set of its open sets.

The collection of *regular* open subsets (a set is regular open if it is the interior of its closure) form a boolean algebra. These two posets are important in logic.

For each set *X* the set of topologies on *X* is a complete poset.

(So each topology is a complete poset and the collection of topologies on a fixed set is itself a complete poset. Do not get confused. Hold on to your hat!)

Huge mileage can be made from the fact that the property of being a topology is intersection-closed

See section 7.5.1

For example: for any X and any $B \subseteq \mathcal{P}(X)$, the property of being a topology on X that contains all members of B is intersection-closed (obviously). The intersection of all topologies containing every element of B is clearly what we would ordinarily call the topology with basis B.

The more open sets a topology has, the finer it is;

The fewer open sets a topology has, the *coarser* it is.

The fact that the class of topologies on a fixed set is closed under arbitrary intersections underpins lots of definitions. We will consider the subspace topology, the product topology and the quotient topology. All three definitions rely on the set of topologies on a fixed set being a complete poset under inclusion, and each relies on some relevant property of topologies being intersection-closed.

DEFINITION 15 The subspace topology on $Y \subseteq X$ is the \subseteq -least topology on Y making the inclusion map continuous.

How do we know there is such a topology on Y? Any topology on Y that makes the inclusion embedding continuous must contain $(\mathrm{id}_Y)^{-1}$ "X" for any open X" $\subseteq X$. And $(\mathrm{id})_Y^{-1}$ "X" is just $Y \cap X$ ". So the subset topology must be the topology generated by all the $Y \cap X$ " with X" open.

DEFINITION 16 The Product Topology.

If $\{Y_i: i \in I\}$ is a family of topologies, X is a set, and there are maps $f_i: i \in I$ from X to Y_i , then the product topology on $\prod_{i \in I} Y_i$ is defined to be the \subseteq -least ("coarsest")

topology making all the projection maps continuous.

Recall that 'coarser' means 'fewer open sets' so the coarsest topology such that Φ should be the intersection of all topologies such that Φ (always assuming that Φ is intersection-closed). Let's see what this actually amounts to. Suppose U is an open set in Y_i , then

$$\{f \in \prod_{i \in I} Y_i : f(i) \in U\} \tag{**}$$

must be an open set. So the product topology on $\prod_{i \in I} Y_i$ is the intersection of all topolo-

gies on $\prod_{i \in I} Y_i$ that contain all the sets **. But that is simply to say that it is the topology generated by the sets **.

550 **DEFINITION 17 Quotient topology**.

The idea is this. Suppose I have a set X equipped with a topology, and a set Y, not yet so equipped. I also have a set F of functions $X \to Y$ and i'm going to call them all continuous, just beco's I happen to feel like it. There is now a finest topology on Y that makes all the functions in F continuous. It's the union of all those topologies.

'Finest' is nice, beco's the more open sets there are $\subseteq Y$ the harder it is for any one function $X \to Y$ to be cts.

With the definition of quotient topology and subset topology one is the finest and the other is the coarsest, and that's beco's they are on different sides of the arrow!

2.3.2 Knaster-Tarski: Fixed Points, Inductively-defined Sets and Cantor-Bernstein-like theorems

THEOREM 6 Tarski-Knaster

Let $\langle X, \leq \rangle$ be a complete lattice and f an order-preserving map $\langle X, \leq \rangle \rightarrow \langle X, \leq \rangle$.

Then f has a fixed point.

564 Proof:

559

560

561

565

566

567

568

572

573

574

578

579

580

Set $A = \{x : f(x) \le x\}$ 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 $f(x) \le x \to f^2(x) \le f(x)$, and so f(a) is also a lower bound for A as follows. Let $x \in A$ be arbitrary; we have $f(x) \le x$, whence $f^2(x) \le f(x)$, so $f(x) \in A$ and $a \le f(x)$.

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

- (1) holds beco's $a \le f(x)$ (as we've just showed) and f is order-preserving;
 - (2) holds beco's $f(x) \le x$ and f is order-preserving;
- (3) holds beco's $x \in A$.

... giving $f(a) \le x$ as desired. But a was the greatest lower bound, so $f(a) \le a$ and $a \in A$. But then $f(a) \in A$ since f " $A \subseteq A$, and $f(a) \ge a$ since a is the greatest lower bound.

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

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.
Notice that in theorem 6 we did not assume that f was continuous in the order topology.
With that extra assumption we can get sharper results. However there is no space for them here.

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

590 COROLLARY 4

Let $\langle X, \leq \rangle$ be a complete lattice; let a be a member of X and let f be 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 $\langle X, \leq \rangle \rightarrow \langle X, \leq \rangle$.

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 earlier 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. It's subtle but it matters. In both case the second structure (the open sets, the set of fixed points for the given increasing function) is a subposet of the first – it is a poset with the inherited structure. However (in both cases) the second structure does not have the same infinitary sup operations as the first even tho' it does have an infinitary sup operation!

The following is a perhaps more familiar-looking example. Fix an (infinite) vector space. Its power set is a complete poset under \subseteq . The collection of its *subspaces* is also a poset – a subposet of that power set indeed. It is also a complete poset. However the sup operation in the subposet of subspaces is not the same as in the power set. The sup operation in the power set is \bigcup ; the sup of a hatful H of subspaces is not $\bigcup H$, but is the space *spanned by* $\bigcup H$.

2.3.3 Inductively defined Sets

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

624

625

626

628

629

630

631

637

639

641

643 644

646

647

648

649

```
\bigwedge \{x : f(x) \le 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 **IN** 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 .

2.3.4 Cantor-Bernstein-like theorems

COROLLARY 6 Cantor-Bernstein

The preorder \leq on cardinals is antisymmetric.

638 Proof:

Suppose $f: A \hookrightarrow B$ and $g: B \hookrightarrow A$. Then the function $a \mapsto A \setminus g''(B \setminus f''a)$ is an order-preserving function $\mathcal{P}(A) \to \mathcal{P}(A)$ and will have a fixed point – call it \mathcal{A} .

Then the function

if
$$a \in \mathcal{A}$$
 then $f(a)$ else $g^{-1}(a)$

bijects A with B.

There are lower-tech proofs of corollary 6 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.

It may be worth noting that this proof of Cantor-Bernstein relies on excluded middle (" $\lambda a \in \mathcal{A}$?"). Cantor-Bernstein is not a constructive thesis.

Have a look at Sheet 3 question 1 at this point.

Other applications include Banach-Tarski.

551 Chapter 3

Three Lectures on Propositional Logic

The letters point to things that evaluate to true and false. I am going to reserve the symbols ' \top ' and – even more important – ' \bot ' for propositional constants NOT for truth-values. The symbols ' \top ' and ' \bot ' are symbols in a propositional language. They are reserved words and they evaluate to true and false.

The language. Propositional letters (aka *literals*): $p, q, r \dots$ or (better!) $p, p', p'' \dots$, 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 '\(\pera\)' as a constant symbol in propositional logic. Beware overloading.

Set of letters is a regular language.

The letters are glued together with **connectives**: \lor , \land , \rightarrow , \neg , \longleftrightarrow , XOR ...

Set of wffs is context-free

Truth-functionality. Valuations and truth-tables.

Interdefinability of connectives.

Intension and extension. Now we can talk about \rightarrow .

3.1 if-then

662

676

A conditional is a binary connective that is an attempt to formalise a relation of implication. The word 'conditional' is also used (in a second sense) to denote a formula whose principal connective is a conditional (in the first sense). Thus we say both that ' \rightarrow ' is a conditional and that ' $A \rightarrow B$ ' is a conditional. The conditional $\neg B \rightarrow \neg A$ is the **contrapositive** of the conditional $A \rightarrow B$, and the **converse** is $B \rightarrow A$. (cf., converse of a relation). A formula like $A \longleftrightarrow B$ is a **biconditional**.

The two components glued together in a conditional are the **antecedent** (from which one infers something) and the **consequent** (which is the something that one

681

683

684

685

688

689

690

691

692

693

695

696

697

698

700

701

702

704

705

infers). In *modus ponens* one *affirms* the antecedent and then *infers* the consequent, thus:

$$\frac{A \to B \quad A}{B}$$

Modus tollens is the rule:

$$\frac{A \to B \quad \neg B}{\neg A}$$

where one denies the consequent and thereby refutes the antecedent.

Affirming the consequent and inferring the antecedent:

$$\frac{A \to B \quad B}{A}$$

is the **fallacy of affirming the consequent**. A **fallacy** is a defective inference. This particular fallacy is an important fallacy, for reasons that will emerge later.

Clearly we are going to have to find a way of talking about implication, or something like it. Given that we are resolved to have a purely truth-functional logic we will need a truth-functional connective that behaves like implies. ('Necessarily' is a lost cause but we will attempt to salvage *implies*). Any candidate must at least obey *modus ponens*:

$$\frac{A \qquad A \to B}{B}$$

A useful clue to the answer comes from the combination of two thoughts:

- (i) the thought that 'A if and only if B' (you will often see 'if and only if' written as 'iff') comes out true as long as A and B agree on their truth-value; and
- (ii) the thought that A iff B must be $(A \to B) \land (B \to A)$.

So both $A \to B$ and $B \to A$ must come out true when A and B agree.

If A and B are both true or both false then $A \to B$ comes out true. We also know that $A \to B$ comes out false when A is true and B is false. This leaves to be decided only the case where A is false and B is true. A useful further thought is that $A \to B$ and $B \to A$ have to be different, but also that we can get each from the other by swapping A and B. This tells us that $A \to B$ must come out true when A is false and B is true.

Thus our resolve that 'if A then B' should be extensional has determined that $A \to B$ will be equivalent to ' $\neg (A \land \neg B)$ ' which itself is equivalent to ' $\neg A \lor B$ '. $A \to B$ evaluates to true unless A evaluates to true and B evaluates to false; \to thus defined is the **material conditional**.

This does not solve the problem of identifying the intensional conditional (it doesn't even try) but it is surprisingly useful, and we can go a long way merely with an extensional conditional. Understanding the intensional conditional is a very hard problem,

3.1. IF-THEN 27

since it involves thinking about the internal structure of intensions and nobody really
has a clue about that. (This is connected to the fact that we do not really have robust criteria of identity for intensions). The literature it has spawned is vast and inconclusive,
but we cannot avoid it altogether.

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. 714 If God exists then He is omniscient. 715 If God exists then He is benevolent. 716 If God can prevent evil then – if He knows that evil exists – then He is not 717 benevolent if He does not prevent it. 718 If God is omnipotent, then He can prevent evil. 719 If God is omniscient then He knows that evil exists if it does indeed exist. 720 Evil does not exist if God prevents it. 721 Evil exists. 722 God does not exist. 723

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

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

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

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

If God can prevent Evil then – if He

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

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

If God is omniscient then He knows that

Evil exists if it does indeed exist. $O \rightarrow (V \rightarrow K)$ (6)

Evil does not exist if God prevents it. $J \rightarrow \neg V$ (7) Evil exists. (8)We want to persuade ourselves that God does not exist¹. Well, suppose he does. Let's deduce a contradiction 726 Assume E. Then (1), (2) and (3) give us 727 P (9),728 0 (10)729 and 730 В (11)731 Now that we know O, (7) tells us that $V \to K$ (12)733 But we know V (that was (8)) so we know K (13)We know P, so (5) tells us that D(14)737 We can feed D into (4) and infer $K \to (\neg J \to \neg B)$ (15)739 But we know K (that was line 13) so we get $\neg J \rightarrow \neg B$ (16)(8) and (7) together tell us $\neg J$, so we get $\neg B$. But we got B at line 11. **DEFINITION 18** Recursive definition of satisfaction A valuation is a function from literals to truth-values. We define what it is for 744 a valuation to satisfy a compound formula by recursion on the subformula relation which (you will have noticed) is wellfounded. 746 We introduce \perp as a constant symbol in propositional logic. Beware overloading. 747 $v \, sat \, l \, [l \, a \, literal] \, iff \, v(l) = \, true;$ $v \, sat \, \phi \wedge \psi \, iff \, (v \, sat \, \phi \, and \, v \, sat \, \psi \,)$ 749 $v \ sat \ \phi \lor \psi \ iff \ (v \ sat \ \phi \ or \ v \ sat \ \psi \)$ $v \ sat \ \phi \rightarrow \psi \ iff \ (either \ not(v \ sat \ \phi) \ or \ v \ sat \ \psi \)$ 751

 $v \ sat \ \neg \phi \ iff \ not(v \ sat \ \phi).$

752

¹Purely for the sake of argument, you understand!

29 3.1. **IF-THEN**

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 756 the same valuations.

DEFINITION 19

753

754

755

757

758

759

760

761

762

763

764

765

767

769

770

771

772

773

774

775

776

777

778

779

780

782

783

784

785

786

787

788

789

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

A Logic is a theory closed under substitution.

A Logic that contains (to take a pertinant example) $A \rightarrow (B \rightarrow A)$ must contain all substitution instances of it, such as: $p \to (p \to p)$, or $(p \lor q) \to ((q \lor r) \to (p \lor q))$

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 $((p_0 \land p_8) \rightarrow \neg p_{16})$ (because an odd plus an odd is an even!).

Notice that this is a theory not a logic, co's it's not closed under substitution. It contains $(p_0 \land p_8) \rightarrow \neg p_{16}$ but not (for example) $(p_1 \land p_9) \rightarrow \neg p_{17}$ 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 20 Any set T (a theory or a Logic) of axioms-and-rules-of-inference gives rise to a deducibility relation written ' \vdash ': "T \vdash ϕ " to mean that ϕ can be deduced using the T-machinery. Sometimes people write " $\psi \vdash_T \phi$ " 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 ⊨ and ⊦ to coincide. Particular set of axioms-andrules doesn't matter; what matters is that it can be done ... proof of concept.

We'll have a brief look at some alternative rules of inference, so that we get 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.

793

794

798

799

800

801

802

803

805

806

മവമ

809

810

811

814

815

816

817

818

819

820

821

822

823

824

REMARK 2 Completeness for the Biconditional Fragment

For a formula ϕ in the language with only \longleftrightarrow and \bot the following are equivalent

- ϕ is valid (= satisfied by all valuations);
- Every literal in ϕ appears an even number of times;
- ullet ϕ is derivable from the three axiom schemes (all substitution instances of)

```
A \longleftrightarrow A;

(A \longleftrightarrow B) \longleftrightarrow (B \longleftrightarrow A) \text{ 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 of occurrences of every propositional letter. 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 $\neg (A \longleftrightarrow \neg A)$. It may even be that if we have ' \bot ' as a propositional constant in the language we don't actually need any more axioms. You may like to check.]

Remark 2 gives us a very nice illustration of a completeness theorem. It says that the *semantic* property of being valid (which involves checking a formula against valuations (which are of course external to the formula) is the same as the *syntactic* property of containing an even number of occurrences of every letter (which can be checked merely by looking inside the formula) are the same! There is another respect (which i have already mentioned) in which this completeness result is typical: the direction [syntactic property] implies [semantic property] is an easy induction; it's the other direction that is hard.

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 [4] 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 21

```
K: A \to (B \to A);
```

$$S: (A \to (B \to C)) \to ((A \to B) \to (A \to C))$$
' and

$$T: ((A \to \bot) \to \bot) \to A;$$

[Define Hilbert-style proof]

3.1. IF-THEN 31

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 22 The **Deduction Theorem** for a logic L is the assertion

if
$$L, A \vdash B$$
, then $L \vdash A \rightarrow B$.

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

Theorem 7 The deduction theorem holds for L iff L contains (all substitution instances of) K and S.

843 Proof:

844

831

835

839

 $L \rightarrow R$ 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)

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

iff (by the deduction theorem)

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

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

850 $R \to L$ 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.

859 860

862

864

865

866

867

870

871

872

873

875

876

877

878

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 8 The Adequacy Theorem

Let Γ be a set of expressions in a propositional language. If $\Gamma \models \bot$ then $\Gamma \vdash \bot$.

Proof:

"Contrapositive"? (first occuserence) 869

We prove the contrapositive. Suppose $\Gamma \nvDash \bot$. We propose to infer $\Gamma \not\models \bot$. $\Gamma \models \bot$ 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 Γ_0 , Γ_1 , Γ_2 ... 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 first-year 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 $\Gamma_i \cup \{t_{i+1}\} \vdash \bot$. If $\Gamma_i \cup \{t_{i+1}\} \not\vdash \bot$ then set $\Gamma_{i+1} = \Gamma_i \cup \{t_{i+1}\}$, else $\Gamma_{i+1} = \Gamma_i$.

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 7 *The Completeness Theorem for Propositional Logic. If* $\Gamma \models \phi$ *then* $\Gamma \vdash \phi$.

3.1. IF-THEN 33

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 8, the Adequacy Theorem, we have $\Gamma \cup \{\neg \phi\} \vdash \bot$.

Now, by the Deduction Theorem, we have $\Gamma \vdash \neg \phi \to \bot$. Finally axiom T 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 and easier to use 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 element²

Why is the compactness theorem for propositional logic like the compactness of the space of valuations? The space of valuations is compact. That is beco's it is the product of lots of copies of the two-point space (one copy for each propositional letter) and the two-point space is compact. And a product of compact spaces is compact. (That's Tikhonov – in fact a subtly weaker version of Tikhonov that sez that a product of compact *Hausdorff* spaces is compact *Hausdorff*). For any propositional formula ϕ the set $[[\phi]]$ of valuations making it true is closed (in fact clopen). Suppose now that Γ is an inconsistent set of formulæ. Then $\{[[\phi]]: \phi \in \Gamma\}$ is a family of closed sets with empty intersection. So some finite subset of it has empty intersection. So there is a finite $\Gamma' \subseteq \Gamma$ with $\Gamma' \models \bot$.

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 7 (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 8 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

²You might think that we need these finite sets of pairs to be partial functions but actually we don't: it all comes out in the wash

933

934

937

938

940

942

947

949

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.

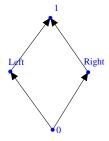
Can detect tautologousness by exhaustive search co's only finitely many cases. Spuriously easy. So consider the truth value algebra $\mathcal{P}(\mathbb{N})$.

3.2 Boolean Algebras detect propositional Tautologies

The most salient boolean algebra is the two-element boolean algebra, sometimes (as here) notated **2**.

What do we mean by this title? Suppose you try building truth tables using a boolean algebra B instead of \mathbb{Z} . 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 a Hasse diagram of the four-element Boolean Algebra.



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

p	V	(¬	p)
1	1	0	1
Left	1	Right	Left
Right	1	Left	Right
0	1	1	0

We know what \land , \lor 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 $(\neg p) \lor q$.

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 2 likewise authenticates ϕ ; in plain English, ϕ is a tautology.

So "authenticated by B" implies "authenticated by 2" (=tautology)

Now (for the other direction) suppose that B does not authenticate ϕ . Then there is $b \in B$, $b \neq \top$, and a row of the truth-table (a valuation, call it v) under which ϕ receives the value b. Consider now any maximal ideal I that contains b and the homomorphism b onto the quotient (which is b2) whose kernel is b3. This homomorphism b4 b5 sends it to b6. Then b6 b7 is a b8-valuation that makes b6 false.

So "refuted by B" implies "refuted by 2" (= not a tautology)

Conclusion:

952

953

955

957

958

959

961

962

964

966

969

970

971

972

974

976

977

978

980

We defined a tautology to be anything authenticated by the two-element boolean algebra \mathbb{Z} , 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{Z} .

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

REMARK 3 We can define a propositional tautology as

- "authenticated by all Boolean Algebras" or
- "authenticated by even one Boolean Algebra";
- 973 it makes no difference.

In fact we can characterise boolean algebras as those things with \land , \lor , 0 and 1 etc that authenticate all propositional tautologies, but there is no space to prove this fact here.

This does lead to some interesting mathematics. Think of 'boolean algebra' in the above discussion as occupying a variable slot. There may be other kinds of algebras with \vee , \wedge etc. Each such kind gives us the set of things that it authenticates, and these can be used to characterise other logics.

3.3 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

$$(\forall a, b, c)(a \le b \rightarrow a \cdot c \le b \cdot c).$$

(*Think* ... additive group of \mathbb{Q} , multiplicative group of \mathbb{R}^+ ...) They tend to be abelian so we write the group operation with a '+'.

66 REMARK 4

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

Proof:

989

990

991

992

993

995

997

1003

1004

1005

1006

1007

1009

1011

1014

1015

1016

1018

1020

1021

1022

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

Obvious tho' this is, it's a useful snotty-logician-opportunity to make the point that this happens beco's the theory of orderable groups is \forall^* , so the class of its models is closed under substructure.

For the converse we consider an arbitrary group G and 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 three axiom schemes:

```
For all a,b,c \in G
p_{a,b} \rightarrow p_{ac,bc}
p_{a,b} \rightarrow (p_{b,c} \rightarrow p_{a,c})
p_{a,b} \text{ XOR } p_{b,a} \text{ for } a \neq b
```

(thanks to Dean Miller for spotting that the third clause needs the extra condition). The first states that the order respects group multiplication, and the second and third assert that the order is total.

If we assume that every finitely generated subgroup of G is orderable then 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.

3.3.1 The Interpolation Lemma

More of a remark than a lemma but it's always called a lemma so i'll go with the flow. But first, a bit of notation. For any set Γ of formulæ, let $\mathcal{L}(\Gamma)$ be the set of propositional formulæ built up from literals in Γ . The ' \mathcal{L} ' connotes \mathcal{L} anguage.

LEMMA 4 Let P, Q, R be three pairwise-disjoint sets of literals.

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.

The formula θ is an **interpolant**.

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 satisfies 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

$$\{p: v(p) = \mathsf{true}\} \cup \{\neg p: v(p) = \mathsf{false}\} \tag{A}$$

This set entails all Q-consequences of ϕ but refutes ϕ itself. So some finite subset Δ of it refutes ϕ . Contraposing we have $\phi \to \neg \wedge \Delta$. But $\neg \wedge \Delta$ is a Q-consequence of ϕ and is therefore satisfied by $v \dots$ which is to say is implied by (A).

So there is no such v.

Observe that this proof does not tell us how to find such a θ : it merely tells us there is one. Example sheet 3 question 5 will guide you through a more effective proof that enables you to actually compute θ from ϕ and ψ . I quote:

"(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)"

There is the possibility of a nice further exercise at this point. Challenge the student: suppose the propositional letter p' has only positive occurrences in A and in B. Can you be sure of finding an interpolant in which it appears only positively? What results can you prove along these lines..?

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 of those nontrivial choice assumptions that we have to put in if we are to prove compactness for uncountable propositional languages, and use it to prove results of interest? Yes!

REMARK 5

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

Actually we don't yet know what an ultrafilter is. Look out definition 25.

Proof:

Let B be a 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} , $\neg \mathcal{U}_{\bot}$; for each $b \in B$ it contains \mathcal{U}_b XOR $\mathcal{U}_{\neg b}$; whenever $a = b \lor c$ then it contains $\mathcal{U}_a \to (\mathcal{U}_b \lor \mathcal{U}_c)$ and if $a \le b$ 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 elements³, and any principal filter generated by such an element is maximal]. So, by compactness, \mathcal{U}_B is consistent. Any valuation ν satisfying \mathcal{U}_B gives you an ultrafilter. An element b of B belongs to this ultrafilter iff $\nu(U_b) = 1$.

3.4 CNF and DNF

DEFINITION 23 A formula in a propositional language with only \land , \lor 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 ' \land ' inside a ' \lor '); 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 ' \lor ' inside a ' \land ').

REMARK 6

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 commutativity of \land *and* \lor .

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 \land over \lor and of \lor over \land , plus the de Morgan laws (ask Wikipædia) to "import" the '¬'s.

It may be worth remarking that it can take exponential time to put a formula into CNF/DNF. This is because the length of the CNF/DNF of ϕ can be exponentially longer than ϕ . Suppose our propositional language has letters $\{p_i : i \in \mathbb{N}\}$. Let Φ_0 be p_0 and let Φ_{2n+1} be $\Phi_{2n} \vee p_{2n+1}$ and let Φ_{2n+2} be $\Phi_{2n+1} \wedge p_{2n+2}$. Then the length of Φ_i is i but its CNF/DNF is of length $\sim 2^i$.

Miniexercise: What is the CNF of a tautology? What is the DNF of the negation of a tautology?

Now could be a good moment to tackle question 5 from Sheet 3:

³called **atoms**.

3.4. CNF AND DNF 39

"Establish that the class of all propositional tautologies is the maximal propositional logic in the sense that any proper 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 $(A \lor \neg A)$ and Double Negation $(\neg \neg A \to A)$.

3.4.1 Resolution Theorem Proving

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

A *clause* is 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 all 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 $(\neg A \lor B) \land (\neg B \lor A)$ giving the two clauses $\neg A \lor B$ and $\neg B \lor A$. 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. Now could be a moment to attempt Sheet 5 question 6.

Chapter 4

One lecture on The Axiom of Choice

1127 AC = axiom of Choice; 1128 ZL is Zorn's Lemma¹ 1129 WO is the Wellordering Principle: every set can be wellordered. 1130 Dean Miller makes the point that in a huge proportion of cases where we use ZL, 1131 the chain complete poset is ordered by \subseteq .

REMARK 7 WO implies AC.

1133 Proof:

1132

1134

1136

1138

1140

1141

1144

1145

1147

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.

1137 **REMARK 8** 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 4.)

You might like to look at Sheet 1 questions 8 and 12 in this connection.

REMARK 9 AC implies WO

¹What is yellow and equivalent to the axiom of choice?

1155

1156

1157

1160

1161

1163 1164

1166

1169 1170

1171

1175

1179

1180

1181

1182

1183

1184

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 1150 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 1152 has a choice function f. Then we declare $s(\alpha)$, the α th member of our sequence, to be 1153 $f(X \setminus \{s(\beta) : \beta < \alpha\}).$ 1154

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 24

A function $f: \langle X, \leq_X \rangle \to \langle X, \leq_X \rangle$ is **inflationary** if $(\forall x \in X)(x \leq_X f(x))$. 1159

Inflationary is NOT the same as increasing! Explain the difference to your friends.

For $AC \rightarrow ZL$ we need

THEOREM 9 The Bourbaki-Witt theorem

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

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 9 AC implies ZL

Proof:

Let $\langle X, \leq_X \rangle$ be a chain-complete poset. By AC we have a choice function f on 1173 $\mathcal{P}(X) \setminus \{\emptyset\}$. Then the function 1174

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

is inflationary and must have a fixed point by theorem 9. That fixed point will be a 1176 maximal element by construction.

4.1 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

4.2 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 25 Filters and ideals in boolean algebras

A filter in a boolean algebra is a subset closed under \land and \geq ; it's a collection of "big" elements. The dual concept is **Ideal**: closed under \lor and \leq . A filter is **proper** iff it does not contain \bot . \subseteq -maximal proper filters are called "ultrafilters", and an ideal dual to an ultrafilter is a **prime** ideal. (The word 'prime' comes from ring theory)

For any b.a. the set of its filters partially ordered by \subseteq is a complete poset; set of proper filters similarly ordered is chain-complete. (In fact it's directed-complete). Arbitrary intersection of a nonempty family of proper filters is a proper filter.

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 "2". yes/no. Hence the connection to logic.

Ideals are 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: $x + y = (x \wedge \overline{y}) \vee (y \wedge \overline{x})$. The boolean elements 0 and 1 become 0 and 1 of the ring.

For the other direction (turning a boolean ring into a boolean algebra) we define $x \wedge y$ to be $x \cdot y$, and $x \vee y$ is x + y + xy. How does that work?! If we are thinking of our boolean algebra as a subalgebra of a power set algebra then + is XOR and \cdot is \cap . It follows that x + y + xy is $(x \text{ XOR } y) \text{ XOR } (x \cap y)$. Now reflect that every member of $x \cup y$ belongs to precisely one of x XOR y and $x \cap y$.

You may be interested in giving a convincing mathematical account of the way in which boolean algebras and boolean rings are the same thing. There is a literature on identities of this kind. If you are interested in this, have look at sheet 4 question 13 which introduces another case of two things that are "the same".

DEFINITION 26 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.

The kernel of the homomorphism $y \mapsto y \wedge \overline{x}$ is **the ideal generated by** x. The filter dual to such an ideal is a **principal filter** (generated by \overline{x}).

A principal ideal is a boolean algebra in its own right.

DEFINITION 27 Quotient over an ideal, or filter

```
1221 x \sim_I y \text{ if } (x \text{ XOR } y) \in I;
1222 or
1223 x \sim_F y \text{ if } (x \wedge y) \vee (\overline{x} \wedge \overline{y}) \in F.
```

1225

1226

1228

1231

1232

1233

1234

1236

1238

1240

1242

1243

1244

1245

1247

1251

1252

1254

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

General patter about representation theorems.

THEOREM 10 Stone's Representation Theorem

Every Boolean algebra is isomorphic to one where

- the order relation is \subseteq set inclusion;
- \wedge is \cap ;
- 1230 \vee is \cup , and
 - complementation $(x \mapsto \overline{x})$ 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 \land , \lor 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 $b \in F$. Clearly the higher up in the algebra you are the more filters you belong to: if $a \le b$ 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 5 If $b \ngeq a$ 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 $a \wedge \overline{b}$, and partially order by \subseteq . This is a chain-complete poset and must have a maximal element by Zorn.

It's easy to check that the boolean operations are respected by the map $b \mapsto$ the set of maximal filters F such that $b \in F$. Any maximal filter must contain precisely one of b and \overline{b} so that ensures that complementation is respected. [It also explains why we

need *maximal* [proper] filters not just any-old-proper-filters... which is why we need the the prime ideal theorem]

If lemma 5 reminds you of the definition of separative poset (definition 14) the rat you have smelt is a real one. However we have no time to explain the connection.

Chapter 5

Six lectures on First order Logic and the Compactness Theorem

```
Stuff that it would be fun to fit in
            Kleene's thm on conservative extensions;
1259
           many-sorted theories;
1260
           relettering vbls not context-free;
1261
           Skolemisation;
1262
           categoricity: both 1st and 2nd order.
         Explain the syntax before anything else
1264
     DEFINITION 28
1265
     Predicate/Relation symbol
     '=' is a reserved word
     arity
1268
     function symbol
1269
     constant symbol
     atomic formula
1271
     quantifier
```

5.1 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 $A \lor B$, $A \land B$, $A \land B$, $A \land B$ 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 . . .

5.1.1 Constants and variables

1273

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.

5.1.2 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 '\'alpha' and '\'alpha' 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 ' \land ' ' \lor ' and the other connectives.

5.1.3 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:

$$(\exists x)(\ldots)$$
 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.

Perhaps one should spell this out a bit more. A constant (symbol) is a thing of the same syntactic type as a variable but which – unlike the variable – cannot be bound,

since it denotes the same single thing throughout an environment. A reserved word is a constant symbol that means the same thing in every environment. (Merely not-being-bindable doesn't make you a constant!)

For example, in a formula like

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

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

5.1.4 Function letters

1325

1326

1327

1328

1329

1331

1332

1333

1334

1335

1336

1337

1338

1340

1341

1344

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.

The other tell-tale sign of second-order logic is having function symbols that act on function symbols, or relation symbols that relate other relation symbols. For obvious reasons such gadgets are called *second order* function symbols/relation symbols. An obvious one is "apply", which takes a function symbol and a term as inputs.

5.1.5 Quantifiers

In Analysis we often use the cofinite quantifier \forall^{∞} which reads "for all but finitely many..." – think of the definition of *convergent sequence*. The sequence $f: \mathbb{N} \to \mathbb{Q}$ is convergent iff $(\forall \epsilon > 0)(\forall^{\infty} m)(\forall^{\infty} m)(|f(m) - f(n)| < \epsilon)$

You may remember the proof that there are infinitely many primes congruent to -1 mod 4^1 . It uses the quantifier "there is an odd number of ..."

Duality

The usual quantifiers \exists and \forall are *dual* in the sense of question 10 on sheet 2: $\neg \forall \neg = \exists$ and $\neg \exists \neg = \forall$

"dense" and "open" in topology are dual notions.

5.1.6 Higher-order

Difference between 1st and 2nd order theories.

 $^{^{1}}$ Suppose there are only finitely many of them. Take 4 times their product and subtract 1. It must have an odd number of factors that are congruent to $-1 \mod 4$, and they must be different from all those in what you tho'rt was the complete set.

DEFINITION 29 In second-order languages you are allowed quantifiers over function symbols and predicate letters.

Topology is a second-order concept.

1349

1350

1352

1353

1355

1356

1358

1359

1361

1363

1365

1367

1369

1370

1372

1374

1376

1377

1379

1381

1382

1383

Complete ordered fields is a 2nd order theory; Simple groups ...

Possibility of many-sorted theories – not the same! Vector spaces

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 compsci example sheet that i had to teach.)

$$(\exists x_1 \dots x_n) (\bigwedge_{1 \le i \ne j \le n} x_i \ne x_j)$$
 (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; for another the ' \wedge ' is a binder that – in some sense – is a universal quantifier 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.

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.

5.1.7 Signatures

DEFINITION 30

Signature: a structure is a set ('carrier set' better than 'domain') with knobs on. The signature is the thing that tells you what the knobs are. 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 $\mathfrak M$.

1387

1388

1389

1390

1391

1394

1396

1397

1398

1399

1400

1401

1402

1404

1406

1409

1410

1411

1412

1415

1417

1418

1419

1420

1421

1422

Typically signatures tend to be finite, altho' sometimes – for special reasons – one expands a structure to one with infinitely many constant symbols. In fact we do this in theorem 19.

I try to to use upper-case $\mathfrak{FRMRIMR}$ 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.

get these in the right order

5.2 Axioms of LPC

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 [4], and I provide them only for the sake of completeness.

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

We need the axioms:

```
\bullet ((\forall x)p) \rightarrow p[t/x]
```

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

```
\bullet (\forall x)(p \to q) \to (p \to (\forall x)q)
```

('x' not free in p)

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

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.

5.3 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 30: 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 (for example) the transitivity axiom $(\forall xyz)(x < y \land y < z. \rightarrow x < z)$ is true in a structure unless that structure has a 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 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

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

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**.² 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

$$sat(f, 'x_i = x_j') \text{ iff}_{df} f(i) = f(j).$$
 (5.1)

²It has been said that the difference between logicians and mathematicians is that logicians understand the use-mention distinction.

5.3. SEMANTICS 51

But how much like? Notice that, as it stands, 5.1 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

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 5.1 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 (5.1) 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:

$$\lceil x_i = x_j \rceil$$

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 ' \land ', ' \lor ' 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 31 First the base cases, for atomic formulæ

1495

1496

1497

1499

1500

1501

1502

1503

1516

1518

1519

1521

1523

1524

1525

1526

1527

1528

1529

1530

```
1504
        sat(f, \lceil x_i = x_j \rceil) iff f(i) = f(j);
1505
        sat(f, \lceil x_i \in x_i \rceil) \text{ iff } f(i) \in f(j).
1506
1507
        Then the inductive steps
1508
        if sat(f, \alpha) and sat(f, \beta), then sat(f, \lceil \alpha \land \beta \rceil);
1509
        if sat(f, \alpha) or sat(f, \beta), then sat(f, \lceil \alpha \vee \beta \rceil);
1510
        if for no g \supseteq f does sat(g, \alpha) hold, then sat(f, \neg \alpha \neg);
1511
        if there is some g \supseteq f such that sat(g, \lceil F(x_i) \rceil), then sat(f, \lceil (\exists x_i)(F(x_i)) \rceil);
1512
        if for every g \supseteq f with i \in dom(g), sat(g, \lceil F(x_i) \rceil), then sat(f, \lceil (\forall x_i) (F(x_i)) \rceil).
              Then we say that \phi is true in \mathfrak{M}, written \mathfrak{M} \models \phi iff sat(\bot, \phi), where \bot is the empty
1514
       partial assignment function.
```

Finally, a formula is valid iff it is true in every interpretation.

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

The full significance of this material will not become apparant until the second part of this course, where you learn to master computable function theory, since 'semidecidable' is a technical term from that area. The informal notion that 'semidecidable' captures is that a set x is semidecidable iff there is a finite engine which, on being presented with a candidate for membership of x will - sooner or later - say 'yes' if the candidate is a member of x but will dither eternally if it isn't.

The completeness theorem for LPC tells us that the set of valid sentences is semidecidable. The engine in question here is a device/gremlin/whatever that searches all possible proofs until it finds a proof of the candidate.

You might think it is obvious that there should be such an engine. That would be a big mistake: it's true but it's not obvious. If you think it is obvious then you have misunderstood something, and you need to retrace your steps.

5.4.1 ∈-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

$$(\exists x)(F(x))$$
 and $F((\epsilon x)F(x))$

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

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

THEOREM 11 Every consistent theory in a countable language has a model.

6 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;
- 1. Add axioms to T₀ to obtain a complete extension,

2. Add ϵ terms to the language.

We execute the task in (1) the way we proved theorem 8 – The Adequacy Theorem,

Notice that when we add ϵ -terms to the language we add new formula: 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

Explain Complete Extension

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}_{∞} , 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*³. 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 $\langle t_1 \dots t_n \rangle$ such that $T_{\infty} \vdash R(t_1 \dots t_n)$ and function symbols similarly.

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

LEMMA 6
$$\mathfrak{M} \models \phi(t_1, \dots t_n) \text{ iff } T_{\infty} \vdash \phi(t_1, \dots t_n)$$

Proof:

1575

1576

1577

1579

1580

1581

1582

1585

1587

1589

1590

1592

1593

1594

1596

1598

1600

1602

1603

1604

1606

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.)

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 $(L \rightarrow R)$.

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

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

So, by induction hypothesis we must have

$$T_{\infty} \vdash \phi((\epsilon x)(\neg \phi(x, t_1, \dots t_n)), t_1, \dots t_n)$$

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

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.

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.

³And we really do mean the set of epsilon terms, not the denotations of those terms...Our models really are created entirely out of syntax.

THEOREM 12 There are "nonstandard" models of arithmetic.

Proof:

What does that mean? Let T be a first-order theory of arithmetic 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 I_n of the language of arithmetic which interprets that language into the standard model and $I_n(`*") = n + 1$. I treats '0', '1' etc as usual. The sequence of I_n s "kicks the can down the road".

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 exam questions.

5.5 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 on p. 7.2.) 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. Question 9 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.

56CHAPTER 5. SIX LECTURES ON FIRST ORDER LOGIC AND THE COMPACTNESS THEOREM

Chapter 6

Two Lectures on Set Theory

We give a historically motivated introduction.

Every time Mathematics discovers it needs a new suite of entities it is liable to get into a tizz about whether or not it's safe to assume that the new entities really are around to be used as desired. Famously there is/was a problem about complex numbers. (You can see there was a problem from the use of the word 'imaginary') We usually sort things out in the end, fortunately. Now in quite a lot of cases the new entities can be introduced to the world dressed up as sets. Indeed that is why sets seem so useful. Two relevant and helpful examples are points at infinity and ideal divisors.

Why do we need points at infinity? A point at infinity is a point where two parallel lines meet. In a sense that's all we know about it, the parallel lines that meet there. So how about we just identify the desired point at infinity with – the bundle ("pencil") of parallel lines that are supposed to go through it? Points at infinity are *concretised* as pencils of lines.

There is a similar story to be told about ideal divisors in rings. One way of approaching this topic is through something which by now is familiar to you. What are integers mod 17? How do you think of them? Do you think of them as a special kind of number? Or do you think of them as equivalence classes of integers under congruence-mod-17? Both points of view are legitimate. If you think of them as equivalence classes of integers, then you will be looking for a way of feeling comfortable about the process of abstraction that will give you this new kind of number. If – on the other hand – you already think of them as this new kind of number then you might be interested in concretising them in terms of older and already familiar things – integers. . . in fact sets of integers. This is where sets come in.

Imaginary divisors get 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 $\sqrt{-5}$. Famously it does not have unique factorisation: 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 do know exactly what their nontrivial multiples are. After all, the multiples of HCF(a, b) are precisely the numbers of the form ax + by; this fact gives us a way in.

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

The set of multiples of a fixed prime is closed under addition, and also by multiplication by arbitrary ring members. And that of course is the definition of an ideal in a ring.

Integers and rationals similarly.

Also equivalence classes of Cauchy sequences as reals.

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

'extension"?

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 there is an unproblematic object-in-extension corresponding to the two intensions (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 of axioms 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 to obtain new sets, 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 *as sets* – 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. Sets are what you get if you start with a mathematical structure and erase all information about the relations that hold between its members and anything that might identify them ('anonymisation') so that the only information is presence/absence information. 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).
```

One that works is

DEFINITION 32 The Wiener-Kuratowski pair

```
pair(x, y) = \{\{x\}, \{x, y\}\}.
fst(p) = \bigcap p and
snd(p) = the unique member of \bigcup p belonging to exactly one member of p.
```

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

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 \lor w = y))

Sumset: (\forall x)(\exists y)(\forall z)(z \in y \longleftrightarrow (\exists w)(z \in w \land 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 \land \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 '∈' and '='.

Now might be a good time to tackle Sheet 4 Q (11), p. 105

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

What other axioms are we going to need...? 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 how we concretised ordered pairs.

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 $X \times Y$.

1755

1756

1757

1758

1759

1760

1763

1769

1770

Notice that we have not made any assumptions about what particular object the ordered pair $\langle x, y \rangle$ might be for $x \in X$ and $y \in Y$. However we have assumed (twice, with f_x and F_x) 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 to 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 global functions:

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

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 question 15 on sheet 4.

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.

6.1 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 further (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 collection $\{z: R^*(z,a) \land \neg F(z)\}$ which is a subset (so we use separation) of the collection $\{z: R^*(z,a)\}$ of things related to a by R^* – aka the transitive closure of R. How are we going to prove that $\{z: R^*(z,a)\}$ 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 $\mathbb N$ to the result. The trouble with this is that the 'n' is not a variable in the language. We need a relation \emptyset that relates n to $R^n\{a\}$. We do this by saying $\emptyset(n, X)$ if every set that contains (n, X) and contains (n, X) whenever it contains (n, R^*Y) also contains (0, X).

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

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

6 **DEFINITION 33** Transitive closure

The image of a set in a class sists a set

The collection

1777

1778

1779

1782

1785 1786

1787

1788

1789

1790

179

1792

1793

1794

1796

1797

1798

1799

1801

1802

1803

1805

1807

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

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! – see p. 10). Evidently TC(x) is the \subseteq -least transitive superset of x.

6.2 The Cumulative Hierarchy

DEFINITION 34 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_{α} s 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 35 $\rho(x)$ *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*. This comes from a lovely image he used

Every morning God wakes up, makes himself a nice strong cup of tea, and sets¹ to work creating sets. Each morning he creates – simultaneously – all the sets that can be made from the sets that he finds on his bench at dawn that morning. Thus each set gets created at the first opportunity, on the first day after all its members have been created. Thus the birthday of a set is the first day after the birthdays of all its members. The set of birthdays is wellordered, so each birthday has an ordinal.

The alert and suspicious reader will notice that I am using the same letter ' ρ ' here as in definition 12, and will wonder whether or not this is legitimate. It is, and I think it is safe to bounce this back to the reader.

¹Joke; JOKE!!

- (i) We prove by induction on the ordinals that $\langle V_{\alpha}, \in \upharpoonright V_{\alpha} \rangle$ 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 35.

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.

1811 **LEMMA 7** $(\forall \alpha)(V_{\alpha} \subseteq \mathcal{P}(V_{\alpha}))$

Proof:

1813

1815

1817

1819

1820

1822

1824 1825

1826

1827

1829

1830

1831

1832

1833

1835

1837

1838

1839

1840

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...

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 would give us Russell's paradox.

REMARK 10 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 11 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 . For example, if \sim is equipollence, so that $[\{x\}]_{\sim}$ is the number 1, the equivalence class of singletons, then its sumset is V. 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.

6.3 Scott's Trick

DEFINITION 36 Scott's trick

When trying to concretise/implement a mathematical entity that arises naturally from equivalence classes [of entities already implemented as sets] for an equivalence relation, then instead of using $[x]_{\sim}$ the \sim -equivalence class of a set x, we use instead the collection of things that are \sim to x 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! Observe that deployment of Scott's trick relies on everything being in the cumulative hierarchy (or at least equivalent to something in the cumulative hierarchy). If x is not in the cumulative hierarchy and nothing $\sim x$ is in the cumulative hierarchy, then Scott's definition turns up the empty set, which is not what we want.

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") 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.

6.4 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 f(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.

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

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

²A set picture is a digraph (set of ordered pairs) that looks as if it could be the graph of ∈ restricted to a transitive set

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 permute 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 is the Axiom of Infinity, and it 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 \land (\forall y)(y \in x \rightarrow 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 ...

6.5 Mostowski Collapse

```
1900 LEMMA 8 (Mostowski's collapse lemma)

1901 If \langle X, R \rangle is a well-founded structure, then there is a transitive set Y and a homomorphism \pi:\langle X, R \rangle \longrightarrow \langle Y, \in \rangle.
```

1903 Proof:

We use the theorem about wellfounded induction and recursion, theorem 1. Define $\pi(x) := {\pi(y) : R(y, x)}$. 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.

The map given by this construction is parsimonious (see p. 17).

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 59 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 . So if R is extensional then π is injective. This give us the special case:

REMARK 11 If $\langle X, R \rangle$ is a well-founded extensional structure, then there is a **unique** transitive set Y and a unique **iso**morphism between $\langle X, R \rangle$ and $\langle Y, \in \rangle$.

Why 'collapse'? A good question (and one not often asked!) It's a piece of mathematical slang, but none the worse for that. I think it's something to do with the thought that the move from the wellfounded structure to its homomorphic image under π destroys information. Any two isomorphic wellfounded structures get collapsed to the same thing.

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 ∈! And this wellordering is of course unique. [Why?] We then take this canonical representative to be our ordinal.

DEFINITION 37

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

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

The fact, noted above, that two isomorphic wellorderings get sent by Mostowski collapse to the same transitive set is why it's safe to think of these collapsed wellorderings as ordinals.

REMARK 12

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

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. 64 has a purpose. It precisely gives us a set containing 0 and closed under successor, and we can obtain the ⊆-least such set from it by separation, as follows:

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

$$\{x: x \in A \land (\forall y)(\emptyset \in y \land (\forall w)(w \in y \rightarrow w \cup \{w\} \in y) \rightarrow x \in y)\}$$

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's a very helpful exercise to crunch out the ranks of the sets that implement these various mathematical objects under the assorted possible implementations. Question 11 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 themselves³ –

³People sometimes use the phrase "essential rank" of a mathematical entity. The essential rank of a structure \mathcal{A} is the least ordinal α s.t. $|A| \leq \Xi_{\alpha}$. Alternatively it's the least ordinal α s.t. V_{α} is big enuff to contain a copy of A (the carrier set of \mathcal{A}). It's a nice phrase, and it should be standard, but it isn't.

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.

Prove that the Cauchy reals and the Dedekind reals are isomorphic. Alternatively prove that the Cauchy reals are complete. Do this without using the Axiom of Choice.

■ 6.6 Ordinals again

1986 We now pick up the thread dropped on page 17. .

DEFINITION 38

1983

1988

1990

1991

1993

1995

1996

1998

2000

2001

2006

2007

2008

2009

2010

2011

2012

2013

- (i) An aleph is the cardinality of a (usually infinite) wellordered set;
- 1989 (ii) $\aleph(\alpha)$, for α a cardinal, is the least aleph $\nleq \alpha$.

(I think we first saw this aleph-without-a-subscript notation on p 17).

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 $|Y| \leq |X|$.

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 $\aleph(|X|)$: $\aleph(\alpha) \le 2^{2^{\alpha^2}}$. By modifying the construction you can obtain better bounds (such as $\aleph(\alpha) \le 2^{\alpha^2}$ – where the asterisk means surjection) but we don't need them.

6.6.1 Initial Ordinals

DEFINITION 39 If $\langle X, <_X \rangle$ is a wellordering of order type α then $card(\alpha)$ is the cardinality |X| of the underlying ("carrier") set X.

You could also say $card(\alpha)$ is, for example, $|\{\beta : \beta <_{On} \alpha\}|$.

This 'card' notation is in the literature, but it is not in common use, and you do not need to know it once you have digested definition 40.

The chief reason why it is not in common use is the thoroughly bad, hacky, reason that: if α is a Von Neumann ordinal, then $card(\alpha)$ is in fact $|\alpha|$, the cardinality of the actual set α . So if our ordinals are Von Neumann, we can always write ' $|\alpha|$ ' instead of ' $card(\alpha)$ ' – so we don't need the latter notation!

DEFINITION 40

An ordinal α is **initial** if $(\forall \beta <_{On} \alpha)(card(\beta) <_{card} card(\alpha))$.

```
We enumerate the initial ordinals as \omega_0, \omega_1, ..., \omega_{\alpha}..., and

We define \aleph_{\alpha} to be card(\omega_{\alpha}) which of course was |\{\beta: \beta <_{On} \omega_{\alpha}\}|;

An aleph is a cardinal of a wellorderable set.
```

Observe that the third item in definition 40 is genuinely a definition not an implementation.

The following should be evident:

```
\aleph_{\alpha} is also the \alphath aleph; \aleph_{\alpha+1} is \aleph(\aleph_{\alpha});
```

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

It might do the reader no harm to explain why these things are true. Alephs are cardinals of wellorderable sets; why are they wellordered by $<_{card}$?

This notation is legitimate because, if X is wellorderable, the Y that we obtain from the construction in the proof of theorem 5 is of minimal size $\not \leq |X|$. So, if |X| is the α th aleph, |Y| is the $(\alpha + 1)$ -th aleph. Is this OK? Yes: each aleph corresponds to a unique initial ordinal, so – by theorem 3 – the alephs are wellordered by $<_{card}$, so we can enumerate them using ordinals.

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.

REMARK 13 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.

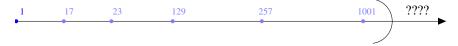
We obtain A' by recursion on $\langle A, <_A \rangle$. The first member of A' is the bottom element of $\langle A, <_A \rangle$. Thereafter the next member is always the $<_A$ -least element a of A. s.t. $f(a)>_B f(a')$ for all $a'<_A a$ that we have already put into A'. Suppose f''A' were bounded in $\langle B, <_B \rangle$. 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:

6.7. $\aleph^2 = \aleph$ 69

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 (**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 41

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 ...]

6.7
$$\aleph^2 = \aleph$$

(Using the letter 'X' as a variable to range over alephs...)

We start by noting that $\aleph = \aleph + \aleph$. (Well, all 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 its definition]. Thus – for example – $\mathfrak{S}(\omega) = 1 + 2 + 3 + 4 + \ldots = \omega$

LEMMA 9

2096

2097

2100

2102

2106

2111

2112

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

2101 Proof:

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

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

2105 ... 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_{α} .

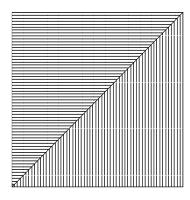
We want to show that $(\aleph_{\alpha})^2 = \aleph_{\alpha}$. \aleph_{α} is defined as $card(\omega_{\alpha})$, so it is the cardinal $|\{\beta : \beta < \omega_{\alpha}\}|$, which means that the canonical set of size $(\aleph_{\alpha})^2$ is the cartesian product

$$\{\beta:\beta<\omega_{\alpha}\}\times\{\beta:\beta<\omega_{\alpha}\}.$$

We partition this last set into three pieces:

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

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



6.7. $\aleph^2 = \aleph$ 71

To be slightly more formal about it, we partition the cartesian product

$$\{\beta: \beta < \omega_{\alpha}\} \times \{\beta: \beta < \omega_{\alpha}\}$$

into the three pieces

$$\{\langle \beta, \gamma \rangle : \beta < \gamma < \omega_{\alpha} \}, \quad \{\langle \beta, \gamma \rangle : \beta = \gamma < \omega_{\alpha} \} \quad \text{and} \quad \{\langle \beta, \gamma \rangle : \gamma < \beta < \omega_{\alpha} \}.$$

It is clear that the third piece is of order type $\mathfrak{S}(\omega_{\alpha})$ in the lexicographic order.

The idea is to show that these three pieces all have cardinality \aleph_{α} . 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_{α} .

Now we can prove

THEOREM 13 $(\forall \alpha)(\aleph_{\alpha} = (\aleph_{\alpha})^2)$.

2125 Proof:

2117

2118

2119

2120

2121

2122

2123

2126

2128

2129

2130

2132

2133

2134

2135

2136 2137

2140

By induction on α . The fact that it holds for $\alpha = 0$ you learnt in first year.

Assume true for all alephs $< \aleph_{\alpha}$. By lemma 9, 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_{α} , 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_{α} , 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_{α} . It's clearly naturally isomorphic to the upper-left triangle (as remarked earlier) so the cartesian product is now a union of three sets each of size \aleph_{α} , giving $(\aleph_{\alpha})^2 = \aleph_{\alpha} + \aleph_{\alpha} + \aleph_{\alpha} = \aleph_{\alpha}$ as desired.

Thus if the axiom of choice holds (so that every infinite cardinal is an aleph) then $\alpha = \alpha^2$ for *all* infinite⁴ cardinals α . There is a converse!

COROLLARY 12 If $\alpha = \alpha^2$ for all infinite cardinals, then AC follows.

2141 Proof:

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

$$\alpha + \aleph(\alpha) = (\alpha + \aleph(\alpha))^{2}$$

$$= \alpha^{2} + 2 \cdot \alpha \cdot \aleph(\alpha) + (\aleph(\alpha))^{2}$$

$$= \alpha + 2 \cdot \alpha \cdot \aleph(\alpha) + \aleph(\alpha)$$

$$= \alpha + \alpha \cdot \aleph(\alpha) + \aleph(\alpha)$$

$$= \alpha(1 + \aleph(\alpha)) + \aleph(\alpha)$$

$$= (\alpha \cdot \aleph(\alpha)) + \aleph(\alpha)$$

⁴In fact we can prove that $\alpha = \alpha^2$ for *all* infinite cardinals directly using Zorn's lemma. The reader might like to prove this for their own satisfaction. In fact i might set it as an exercise. But our trajectory givs more information, since we get a choice-free proof that $\alpha = \alpha^2$ for infinite alephs.

$$= (\alpha + 1) \cdot \aleph(\alpha)$$
$$= \alpha \cdot \aleph(\alpha)$$

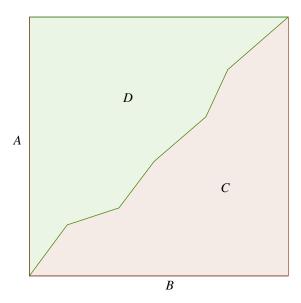
Now we use Bernstein's lemma....

REMARK 14 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 α .

Proof:



Proof:

2147

2148

2150

2152

2154 2155

2156

2157

2159

2161

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 12. There are two ways in which we could apply Bernstein's lemma here.

- (i) 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:
 - (ii) $\aleph(\alpha) \le \alpha \lor \alpha \le^* \aleph(\alpha)$.

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.

Tidy this up

Bernstein's Lemma looks like an annoying factoid of no interest. However it is an essential prop in the proof that the reals are the same size as the power set of the naturals.

We can also use theorem 13 to show that a lot of initial ordinals are regular.

2169 **THEOREM 14** (uses AC)

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

Proof.

2163

2164

2165

2166

2167

2170

2171

2172

2173

2174

2175

2176

2177

2178

2179

2181

2182

2184

2185

2186

2188

2190

2191

2192

2193

2194

2197

If $\omega_{\alpha+1}$ were the sup of fewer than $\aleph_{\alpha+1}$ (which is to say the sup of no more than \aleph_{α} smaller) ordinals – then the set of ordinals below it (which is of size $\aleph_{\alpha+1}$) would be a union of at most \aleph_{α} things each of size \aleph_{α} at most. We saw in Sheet 2 question 4 how to use AC to show that such a union is of size $(\aleph_{\alpha})^2$ at most, and theorem 13 now tells us it is of size \aleph_{α} at most, which is impossible.

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

6.8 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 Do they? 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. 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: it gives us transitive models.

6.8.1 Δ_0 formulae and the Lévy Hierarchy

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

DEFINITION 42

A restricted quantifier in the language of set theory is ' $(\forall x)(x \in y \to ...)$ ' or ' $(\exists x)(x \in y \to ...)$ '.

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.

Thereafter a Σ_{n+1} (respectively Π_{n+1}) formula is the result of binding variables in a Π_n (respectively 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_n^T and Π_n^T . 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 similarly of course).

Set the definition of 'absolute up in lights 2213

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 ' $(\exists x)(x \in M \land \ldots)$ '.

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! We consider two contrasting cases, pairing and power set.

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

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

 \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

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

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 15) and Mostowski collapse (lemma 8) 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 the power set of $\mathbb N$ has only countably many members (they're all in $\mathfrak M$ beco's $\mathfrak M$ is transitive) and it therefore cannot possibly be the true power set of the naturals of $\mathfrak M$.

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

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

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

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

6.8.2 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. However we have overstated what is needed for being a model of replacement. 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.

2279

2281

2282

2284

2286

2288

2290

2293

2301

2302

2303

2304

2305

2306

2308

2309

Put them all on the board and tick them off one by one 2275

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 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.)⁶

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

It will be a model of 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 f that sends n to $V_{\omega+n}$. You will need Quine's trick (page 60) to define it properly.

Observe that the image f \mathbb{N} of the set of naturals in this function is $\{V_{\omega+n} : n \in \mathbb{N}\}$ and the rank of this set is clearly $\omega + \omega$ so it cannot be a member of $V_{\omega+\omega}$.

Both the models we have considered so far are V_{α} s. However there are other structures we can use.

DEFINITION 43

```
2297 \mathcal{P}_{\phi}(x) = \{y \subseteq x : \phi(y)\};

2298 H_{\phi} is the least fixed point for x \mapsto \mathcal{P}_{\phi}(x);

2299 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.

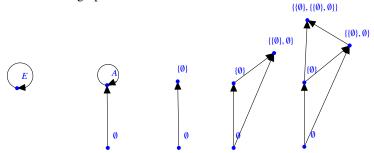
⁶There is a subtlety here, because the specially sexed-up version of the axiom of infinity $(\exists x)(\emptyset \in x \land (\forall y)(y \in x \to y \cup \{y\} \in x))$ that we saw on page 64 asserts that there is an x with a special property, and that special property is Δ_0 . The point is that you have to do a bit of work to show that if $\emptyset \in x \land (\forall y)(y \in x \to y \cup \{y\} \in x)$ then x really is Dedekind-infinite.

Set Pictures

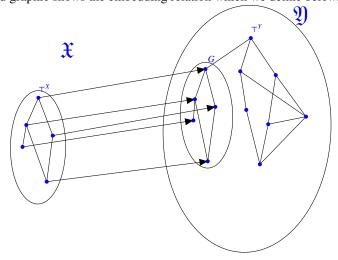
DEFINITION 44 A set picture (also known as an accessible pointed (di)graph or **APG**) is a special kind of decorated digraph, an extensional binary structure $\langle X, R, \top^R \rangle$ where \top^R is a designated ("top") element satisfying $X = (R^*)^{-1}$ " $\{\top^R\}$, a condition which says that \top^R is accessible from any vertex in X.

The idea is that it's a binary structure that looks as if it might be the graph of \in * restricted to the transitive closure of a singleton. The binary relation in an APG may be wellfounded, but much use is made of the fact that there are APGs that aren't wellfounded. Like digraphs out in the Big Wide World these digraphs may be decorated, and the decorations tend to be sets, pasted on in such a way that the membership relation between them echoes R. (Graphs can have either their edges or the vertices decorated; here we decorate only the vertices.) See pictures below.

Here are some graphics for APGs:



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



This diagram⁷ depicts the obvious embedding relation between set pictures: \mathfrak{X} =

⁷I have to confess that the binary relation in the picture isn't extensional. With any luck the reader won't

2332

2333

2334

2335

2336

2337

2338

2339

2341

2342

2343

2345

2347

2349

2351

2352

2361

```
2327 \langle X, R, \top^R \rangle embeds into \mathfrak{Y} = \langle Y, S, \top^S \rangle if there is y \in Y with S(y, \top^S) and \langle X, R, \top^R \rangle \simeq 2328 \langle S^{-1} (y), S, \top^S \rangle.
```

Clearly isomorphism is a congruence relation for this embedding relation and we write 'E' for the relation between the isomorphism classes.

If you forget this definition you can reconstruct it if you remember that it's trying to say that the set that $\langle X, R, \top^R \rangle$ is a picture of is a member of the set that $\langle Y, S, \top^S \rangle$ is a picture of.

Now to prove the existence of HC we consider $V_{\omega+1}$ and wellfounded set pictures whose carrier sets are in $V_{\omega+1}$, 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 rise to this challenge. We need countable choice to show that a union of countably many countable sets is 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.

6.8.3 Independence of Sumset

We define \beth_α to be $|V_{\omega+\alpha}|$. To prove the independence of sumset we consider H_{\beth_ω} . 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 '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_ω} .

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

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.

6.8.4 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 $\emptyset \in_{\sigma} \emptyset$, so foundation does not hold in $\langle V, \in_{\sigma} \rangle$.

notice. What matters is the isomorphism between the two smaller ellipses.

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 45 ϕ^{σ} is the result of replacing ' \in ' in ϕ throughout by ' \in_{σ} '.

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

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 variables to the left of the \in are of level 0 are of level 1 and you don't have a problem with them – they always have a ' σ ' applied to 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 ϕ where appropriate. σ "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, all occurrences of that variable have the same prefix ... the point being that – for any permutation τ whatever – $(\forall x)(\dots \tau(x)\dots)$ is equivalent to $(\forall x)(\dots x\dots)$ beco's τ 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. We now say that a formula is *stratifiable* iff 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.

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

becomes

2367

2368

2370

237

2372

2374

2376

2378

2380

2382

2383

2384

2385

2386

2387

2388

2389

239

2392

2394

2395

2396

2397

2398

2399

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

We can drop the σ s preceding 'X' and 'Y' to obtain

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

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.

6.8.5 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 (tho' there are two exercises that are sleepers for it, Sheet 5 questions 7 and 8) 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 11, 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 slight abuse of notation) 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!!)⁸

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 $\{Y \subseteq X : Y \text{ is of finite support}\}$ are distinct sets both of finite support, but they have the same members with finite support. We have to consider the class HF of elements hereditarily of finite support. Let's call it HF. This time we do get a model of ZF.

LEMMA 10 The class of sets of finite support is closed under all the definable operations that the universe is closed under.

⁸A counterexample: wellorder cofinitely many atoms. The graph of the wellorder has cofinitely many atoms in its transitive closure, but they are all inequivalent.

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

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

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

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

So the equivalence relation $\sim_{f(x_1...x_n)}$ induced on atoms by $f(x_1...x_n)$ has an equivalence class which is a superset of the intersection $A(x_1) \cap ... A(x_n)$, which is cofinite, so $f(x_1...x_n)$ 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 10) 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 is of finite support. And since we are interpreting this in HF, all members of f is a re in HF, so f 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! To show the independence of AC from ZF with foundation we would have to do more work, and it's a bit much for here.

6.9 The Modern Theory of Wellfounded Sets

Modern Set Theory, since its flowering in the 1960s, has been concerned with the fine structure of the cumulative hierarchy. That is to say, the modern tradition assumes the axiom of foundation without question. On the whole the tradition has also embraced the axiom of choice. The only area where the axiom of choice is occasionally eschewed is in the study of **The axiom of determinacy**.

To a significant extent developments in modern set theory have been driven by a set theoretic foundationalism that believes that all the problematic chickens flying around mathematics will eventually come home to roost in set theory, and that Set Theorists need to be ready with answers. To this end certain decisions have been made about how Mathematics is to be understood-as-Set-Theory: ordinals are von Neumann ordinals, reals are taken to be sets of natural numbers (aka finite (von Neumann) ordinals). By means of such identifications problems in Analysis become problems in Set Theory. In principle one might worry that the translation being used might not be faithful to the interpreted mathematics (Frege worried about this) but generally Set Theorists do not.

In 1940 Gödel exhibited a construction of a very impoverished model of ZFC which demonstrably satisfies AC and GCH. Modifications of his construction loom large in subsequent work.

The roots that Set Theory had in Analysis have had profound and lasting effects on the direction it has taken. Analysis has been a rich source of axioms, or conjectures. The axiom of determinacy is one. This is an exciting and powerful axiom that is very attractive to Set Theorists ... so attractive in fact that they are willing (part-time) to overlook the fact that it contradicts AC and pursue some of the weirdnesses that it opens up. Another is the conjecture that there is a set with a countably additive two-valued measure vanishing on singletons. Questions in (the) Descriptive Set Theory (of the reals) loom large too.

Another rich source of axioms and ideas has been generalisations of the axiom of infinity. How "long" is the cumulative hierarchy?

Roughly there are two ways of generalising the axiom of infinity.

One is to dream up ever more improbable things and postulate that the universe contains examples of them. The earliest example of this kind of move is probably the axiom of Mahlo that every normal function from ordinals to ordinals has a regular fixed point. A more striking example is the postulate that there should be a set admitting a countably additive two-valued measure vanishing on singletons. The cardinal of such a set is said to be **measurable**. It's not obvious that such a set must be *large*, but it becomes clear that such sets are very *rare*, so the smallest one is probably quite large.

The other is to spice up a theory T by adding an axiom that says, in a natural way, that T has a model. Thus we discover an axiom that says there is an ordinal κ s.t. V_{κ} is a model of ZF. We can repeat the trick, and with a little ingenuity we can devise axioms that will perform lots of such extensions with one stroke. This can require a great deal of ingenuity, and one set of ideas for how to do it grew out of the proposition that there

is a a measurable cardinal. This turns out to imply lots of things that look like strong axioms of infinity. The idea has been incredibly fertile.

Chapter 7

Two Lectures on Model Theory

Can we fit in quantifier-elimination?

All this model theory is about first-order theories No second-order stuff thank you! 2527

The Skolem-Löwenheim Theorems 7.1

Notice that the proof of theorem 11 gives us something slightly more than I have 2529 claimed. If the consistent theory T we started with was a theory in a countable lan-2530 guage then the model we obtain by the above method is also countable. It's worth 2531 recording this fact:

COROLLARY 13

2533

2536

2538

2540

2541

2542

2543

2544

2546

Every consistent theory in a countable (first-order) language has a countable model. 2534

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 15 Downward Skolem-Löwenheim

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

THEOREM 16 Upward Skolem-Löwenheim

Any theory with infinite models has arbitrarily large models.

Proof: Add lots of constants and appeal to compactness.

Actually we can do significantly better. The point is that theorem 15 tells us that if 2545 we add enough constants to ensure that the model is of size at least κ , then the model will be no bigger than κ .

2555

2559

2568

2569

2570

2571

2572

2573

2574

2576

2577

2579

2580

2581

2583

7.2 Categoricity

DEFINITION 46

2550 A theory is categorical iff it has only one model up to isomorphism;

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

2553 A structure is said to be κ-categorical if its theory is κ-categorical¹.

Duplication

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-\(\cdot\)0. Such theories are always called *countably categorical*. Here's the standard example, the theory of dense linear order without endpoints.

7.2.1 Back and forth

The theory of dense linear order has one primitive nonlogical symbol \leq and the following axioms:

```
2562 \forall xyz(x \leq y \rightarrow (y \leq z \rightarrow (x \leq z)));

2563 \forall xz(x \leq y \rightarrow y \leq x \rightarrow x = y);

2564 \forall xy\exists z(x < y \rightarrow (x < z \land z < y));

2565 \forall x\exists y(y > x);

2566 \forall x\exists y(x > y);

2567 \forall xy(x \leq y \lor y \leq x).
```

I have to confess i'm being a bit naughty here I defined only the symbol ' \leq ' but then proceeded to use '<' and '>' as well, knowing that you would know what I meant. It's naughty in the sense that one shouldn't exploit linguistic conventions without spelling them out first. Admittedly mathematicians do it all the time, and it takes a logician to make a fuss about this ... but then this *is* a logic course.

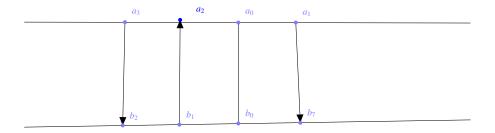
THEOREM 17 All countable dense linear orders without endpoints are isomorphic.

I shall provide a proof because it is possible to prove the theorem the wrong way. *Proof*: Suppose we have two countable dense linear orders without endpoints, $\langle \mathcal{A}, \leq_{\mathcal{A}} \rangle$ and $\langle \mathcal{B}, \leq_{\mathcal{B}} \rangle$. They are both countable, so the elements of \mathcal{A} can be enumerated as $\langle a_i : i \in \mathbb{N} \rangle$ and the elements of \mathcal{B} can be enumerated as $\langle b_i : i \in \mathbb{N} \rangle$.

We start by pairing off a_0 with b_0 . Thereafter we procede by recursion. At each stage we have paired off some things in $\langle \mathcal{A}, \leq_{\mathcal{A}} \rangle$ with some things in $\langle \mathcal{B}, \leq_{\mathcal{B}} \rangle$. Let us now consider the first thing in $\langle \mathcal{A}, \leq_{\mathcal{A}} \rangle$ not already paired off. (We mean: first in the sense of $\langle a_i : i \in \mathbb{N} \rangle$.) This lies between two things we have already paired, and we must find a mate for it in $\langle \mathcal{B}, \leq_{\mathcal{B}} \rangle$ that lies in the interval between their mates. Since the ordering is dense, this interval is nonempty, and we pick for its mate the first (in the sense of the $\langle b_i : i \in \mathbb{N} \rangle$) in it.

¹Readers should be warned that many people confusingly write ' ω -categorical' when talking about \aleph_0 -categorical structures like the countable dense linear order without endpoints.

7.3. RESULTS RELATED TO COMPLETENESS, EXPLOITING COMPLETENESS87



In the illustration we paired a_0 with b_0 . Then we sought the first b that is to the right of b_0 so it can match a_1 and (in the picture) that first b is b_7 , since $b_1 cdots b_6$ are to the left of b_0 . Next we had to find a match for b_1 , and that turned out to be a_2 . Then the a of smallest subscript that hasn't been mated so far is a_3 – and its mate is going to be b_2 .

That is the recursive process we use to build the bijection. It goes *back and forth*: $\langle \mathcal{A}, \leq_{\mathcal{A}} \rangle$ to $\langle \mathcal{B}, \leq_{\mathcal{B}} \rangle$ and then $\langle \mathcal{B}, \leq_{\mathcal{B}} \rangle$ to $\langle \mathcal{A}, \leq_{\mathcal{A}} \rangle$. That way we can be sure that by the time we have gone back and forth n times we have used up the first n things in the canonical enumeration of $\langle \mathcal{A}, \leq_{\mathcal{A}} \rangle$ and the first n things in the canonical enumeration of $\langle \mathcal{B}, \leq_{\mathcal{B}} \rangle$. We will have used n other things as well on each side, but we have no control over how late or early they are in the canonical orderings.

The union of all the finite partial bijections we thus construct is an isomorphism between $\langle \mathcal{A}, \leq_{\mathcal{A}} \rangle$ and $\langle \mathcal{B}, \leq_{\mathcal{B}} \rangle$.

Note that this construction shows that the group of order-automorphisms of the rationals acts transitively on unordered n-tuples.

Theorem 17 tells us that the theory of dense linear orders without endpoints is complete. Suppose it were not. Then there would be a formula ϕ that is undecided by it, and by the completeness theorem there would be dense linear orders without endpoints that were ϕ and dense linear orders without endpoints that were not ϕ . But then these dense linear orders would not be elementarily equivalent, and *a fortiori* not isomorphic either.

The study of countable structures that are unique up to isomorphism is a pastime widespread among logicians and has interesting ramifications.

There is a remarkable and deep theorem of Morley that says that a theory that is κ -categorical for even one uncountable κ is κ -categorical for *all* uncountable κ . It is beyond the scope of these notes. However, there are a number of natural and important countably categorical theories.

Now might be a good moment to look at question 8 on sheet 3.

2620

2621

2622

2624

2625

2627

2630

2631

2641

2642

2644

2647

2649

2650

2651

2652

7.3 Results related to completeness, exploiting completeness

7.3.1 Prenex Normal Form and Quantifier-Counting

DEFINITION 47 A formula is in prenex normal form if it is of the form string-ofquantifiers followed by stuff containing no quantifiers. All quantifiers hae been "pulled to the front" or "exported".

THEOREM 18 Every formula is equivalent to one in PNF.

Sketch of proof:

Quantifiers can be "pulled to the front". $(\forall x)(A(x)) \land (\forall y)(B(y))$ is clearly equivalent to $(\forall x)(\forall y)(A(x) \land B(y))$, 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 48

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 similarly **universal-existential** sentences, and theories² 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 49 Equational Theories and Equational Languages

A language is equational if it has no nonlogical relation symbols;

A equational theory is a theory in an equational language.

Equational theories can have *function* symbols of course, and constant symbols. Group theory is equational; ring theory is equational . . .

DEFINITION 50 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.

For a theory T, T_{\forall} is the set of universal consequences of T.

Dear Reader, in case you were thinking that T_{\forall} is the kind of thing that only a sad logician would dream up, it might be worth pointing out that the theory of integral domains is precisely the universal fragment of the theory of fields: if T is the theory of fields then T_{\forall} is the theory of integral domains. A structure is an integral domain iff it is a substructure of a field. See theorem 19 below.

²Such theories are sometimes called "inductive"; Dunno why.

LEMMA 11 For any consistent theory T and any model \mathfrak{M} of T_{\forall} , 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 19 T is universal iff every substructure of a model of T is a model of T.

Proof:

 $L \rightarrow R$ 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.

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 restrict your notion of substructure!

The next theorem after theorem 19 will say that a theory is 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, let alone prove it!

7.4 Omitting Types

You were set up for this topic by question 15 on sheet 3. That question concerned types for propositional logic. There is an analogous result for predicate logic but it is much

harder, and we are not going to prove it (most textbooks will have a proof). However we do need to engage with the ideas, so we will at least *state* it.

2695 **DEFINITION 51**

- 2696 A **type** in a first-order language \mathcal{L} is a (usually) infinite set of formulæ.
- ²⁶⁹⁷ A type Σ is an n-type if the formulæ in it all have n free variables.
- 2898 A model \mathfrak{M} realises an n-type Σ if there is a tuple $m_1 \dots m_n$) in M s.t. $\mathfrak{M} \models \sigma(m_1 \dots m_n)$ for every $\sigma \in \Sigma$. Otherwise \mathfrak{M} 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 $T \vdash (\forall \vec{x})(\neg \phi(\vec{x}))$.
- 702 **THEOREM 20** Omitting Types Theorem
- 2703 If T locally omits a type Σ , then it has a model omitting Σ .

OTT sounds a bit recherché but it does touch on some mainstream concerns: the standard model of arithmetic omits the 1-type

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

4 7.5 Direct Products and Reduced Products

If $\{\mathcal{A}_i : i \in I\}$ is a family of structures, we define the product

$$\prod_{i \in I} \mathcal{A}_i$$

to be the structure whose carrier set is the set of all functions f defined on the index set I such that $(\forall i \in I)(f(i) \in A_i)$ and the relations of the language are interpreted "pointwise": the product believes f R g iff $(\forall i \in I)(f(i) R g(i))$.

The $\{\mathcal{A}_i : i \in I\}$ are said to be the factors of the product $\prod_{i \in I} \mathcal{A}_i$.

For this operation to make sense it is of course necessary that all the \mathcal{A}_i should have the same signature!

Products are nice in various ways.

2709

2710

2711

2713

2714

2715

DEFINITION 52 Let Γ be a class of formulæ.

Products **preserve** Γ if whenever $\prod_{i \in I} \mathcal{A}_i$ is a product of a family $\{\mathcal{A}_i : i \in I\}$ and $\phi \in \Gamma$ and $(\forall i \in I)(\mathcal{A}_i \models \phi)$ then $\prod_{i \in I} \mathcal{A}_i \models \phi.$

By definition of product, products preserve atomic formulæ. Clearly they also preserve conjunctions of anything they preserve, and similarly universal quantifications over things they preserve.

What about more complex formulæ? You know that products preserve equational theories (a product of rings is a ring, after all). They also preserve Horn formulæ.

DEFINITION 53.

2719

2720

2721

2723

2724

2725

2726

2727

2728

2729

2730

2731

2732

2733

2735

2737

2738

2739

2740

274

2743

2744

2746

A Horn clause is a disjunction of atomics and negatomics of which at most one is atomic.

A Horn property is a property captured by a [closure of a] Horn expression; A Horn theory is a theory all of whose axioms are universal closures of (conjunctions of) Horn clauses.

Yes, but what is a horn formula?

7.5.1 Intersection-closed properties

A property *F* of sets is *intersection-closed* if an arbitrary intersection of sets with property *F* also has property *F*. Examples: the property of being a transitive relation (tho'rt of as a set of ordered pairs) is intersection-closed; ditto equivalence relation, symmetric relation.... An arbitrary intersection of convex figures in the plane is another such figure. An arbitrary intersection of subgroups of a group is a subgroup of that group. And many more.

The properties *symmetric reflexive*, and *transitive* are all Horn. The idea is an important one because for Horn properties one has an idea of **closure**. Take transitivity for example. *R* is transitive iff

$$(\forall xyz)((R(x,y) \land R(y,z)) \rightarrow R(x,z))$$

That is to say, the graph of R (R thought of as a set of ordered pairs) is closed under the operation that accepts the pair $\langle x, y \rangle$ and the pair $\langle y, z \rangle$ as inputs and outputs the pair $\langle x, z \rangle$.

As usual, when one has a set and an operation that can be applied to its members, one has a notion of canonical unique closure of that set under that operation. The point about Horn properties is that in the Horn clause there are lots of premisses (all positive, each saying – as it were – that the relation contains a certain tuple) and one conclusion, saying that in that case the relation contains this other tuple.

Intersection-closed. 'f" $X^n \subseteq X$ ' is Horn. The easy direction I am leaving as an exercise; it will say: if F is a Horn property, then for every x the F-closure of x exists and is well-defined and unique.

Observe that "is a total order" is not a Horn property.

REMARK 15 Products preserve Universal Horn formulæ

Proof:

Suppose every factor \mathcal{A}_n believes $(\forall \vec{x})((\bigwedge_{i < j} \phi_i(\vec{x})) \to \phi_j(\vec{x}))$, where all the ϕ are atomic. We want to show that the product believes it too. So let $\vec{f} = f_1 \dots f_k$ be a tuple of things in the product satisfying the antecedent. That is to say, for each factor

³You may recall that the collection of transitive relations on a fixed set is a complete poset. If you didn't prove it then, prove it now. Observe that the only feature of the property *transitive* that you have used in the proof is the fact that it is a Horn property.

 \mathcal{A}_n , we have $\mathcal{A}_n \models \phi_i(f_1(n), f_2(n) \dots f_k(n))$ for each i < j. But then every \mathcal{A}_n believes $\phi_i(f_1(n), f_2(n) \dots f_k(n))$ so the product believes $\phi_i(f_1, f_2 \dots f_k)$ as desired.

In particular an arbitrary product of transitive relations is a transitive relation. [This is a good point of departure]. An arbitrary product of posets is a poset (being a poset is horn) but an arbitrary product of tosets is not reliably a toset because the totality condition (trichotomy, connexity) is not Horn.

This illustrates how products do not always preserve formulæ containing \vee or \neg . This suggests that remark 15 is best possible. (We won't prove it) How so? If ϕ is preserved, then the product will fail to satisfy it if even *one* of the factors does not satisfy it even tho' all the rest do. (cf Genesis [19:23-33] where not even one righteous man is enough to save the city. The product is not righteous unless *all* its factors are). In these circumstances the product $\models \neg \phi$ but it is not the case that all the factors $\models \neg \phi$. As for \vee , if ϕ and ψ are preserved, it can happen that $\phi \vee \psi$ is not, as follows. If half the factors satisfy ϕ and half satisfy ψ , then they all satisfy $\psi \vee \phi$. Now the product will satisfy $\psi \vee \psi$ iff it satisfies one of them. But in order to satisfy one of them, that one must be true at *all* the factors, and by hypothesis it is not. Something similar happens with the existential quantifier.

7.5.2 Reduced products

We will need filters and ultrafilters from definition 25.

Given a filter F over the index set, we can define $f \sim_F g$ on elements of the product if $\{i \in I : f(i) = g(i)\} \in F$. Then we either take this \sim_F to be the interpretation of '=' in the new product we are defining, keeping the elements of the carrier set of the new product the same as the elements of the old or we take the elements of the new structure to be equivalence classes of functions under \sim . These we will write $[g]_{\sim_F}$ or $[g]_F$ or even [g] if there is no ambiguity.

This new object is denoted by the following expression:

$$(\prod_{i\in I}\mathcal{A}_i)/F$$

Similarly we have to revise our interpretation of atomic formulæ so that

$$(\prod_{i\in I}\mathcal{A}_i)/F \models \phi(f_1,\ldots f_n) \text{ iff } \{i: \phi(f_1(i),\ldots f_n(i))\} \in F.$$

REMARK 16 \sim_F is a congruence relation for all the operations that the product inherits from the factors.

Can't do any harm to write out a proof. [Not lectured but supplied for the notes]

Let H be an operation, of arity h, and let \vec{f} and \vec{g} be two h-tuples in the product, with $f_i \sim_F g_i$ for each $i \leq h$. That is to say: for each $i \leq h$, $\{n: f_i(n) = g_i(n)\} \in F$. Since h is finite, we can conclude that $\{n: \bigwedge_{i \leq h} f_i(n) = g_i(n)\} \in F$.

We want $H(\vec{f}) \sim_F H(\vec{g})$. That is to say we desire that

$${n: H(f_1(n)\cdots f_h(n)) = H(g_1(n)\cdots g_h(n))} \in F.$$

But we know (by our assumption that $f_i \sim_F g_i$ for each $i \leq h$) that $\bigwedge_{i \leq k} (f_i(n) = g_i(n))$ holds for an F-large set of n, so if H is given the same tuple of arguments it can hardly help but give back the same value.

It may be worth bearing in mind that to a certain extent the choice between thinking of elements of the carrier set of the reduced product as the \sim_F -equivalence classes of functions and thinking of them as those functions is a real one and might matter. I have proceeded here on the basis that the carrier set is the set of \sim_F -equivalence classes because that seems more natural. However, in principle there are set-existence issues involved in thinking of a product this way – how do we know that the \sim_F -equivalence classes are sets? – so we want to keep alive in our minds the possibility of doing things the second way. This will matter when we come to consider reduced products where the factor structures are proper classes (= have carrier sets that are proper classes). This happens in the extensions of ZF(C) with large cardinal axioms (specifically measurable cardinals). In practice these issues are usually swept under the carpet; this is a safe strategy only because it is in fact possible to sort things out properly! There is of course also the possibility of picking representatives from the equivalence classes, possibly by means of AC.

(For those of a philosophical turn of mind, there is an interesting contrast here with the case of quotient structures like, say, integers mod p. I have the impression that, on the whole, mathematicians do not think of integers-mod-p as sets of integers, nor as integers equipped with a nonstandard equality relation, but rather think of them as objects of a new kind. These reflections may have significance despite not really belonging to the study of *mathematics*: the study of *how we think about mathematics* is important too. And of course this also takes us back to the roots of set theory as discussed at the beginning of chapter 6.)

The reason for proceeding from products to reduced products was to complicate the construction and hope to get more things preserved. In fact nothing exciting happens (we still have the same trouble with \lor and \neg – think: *tosets*) unless the filter we use is ultra. Then everything comes right.

7.6 Ultraproducts and Łoś's theorem

At this point an excellent book to look at is [5]. It's comprehensive, clear, and probably at the right level.

THEOREM 21 (Łoś's theorem)

Let \mathcal{U} be an ultrafilter on I. For all expressions $\phi(f,g,h...)$,

$$(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U}\models\phi(f,g,h\ldots)\text{ iff }\{i:\mathcal{A}_i\models\phi(f(i),g(i),h(i)\ldots)\}\in\mathcal{U}.$$

2812 Proof:

We do this by structural induction on the rectype of formulæ. For atomic formulæ it is immediate from the definitions.

2823

2825

2826

2827

2829

2831

[wouldn't hurt to write out the details for the fainthearted!]

The proofs for \land and \forall are just as with products.

As we would expect, the only hard work comes with \neg and \lor , though \exists merits comment as well.

2819 Disjunction

Suppose we know that $(\prod_{i \in I} \mathcal{A}_i)/\mathcal{U} \models \phi$ iff $\{i : \mathcal{A}_i \models \phi\} \in \mathcal{U}$ and

$$(\prod_{i \in I} \mathcal{A}_i)/\mathcal{U} \models \psi \text{ iff } \{i : \mathcal{A}_i \models \psi\} \in \mathcal{U}.$$

We want to show

$$(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U}\models (\phi\vee\psi) \text{ iff } \{i:\mathcal{A}_i\models\phi\vee\psi\}\in\mathcal{U}.$$

The steps in the following manipulation will be reversible. Suppose

$$(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U}\models\phi\vee\psi.$$

Then

$$(\prod_{i \in I} \mathcal{A}_i)/\mathcal{U} \models \phi \text{ or } (\prod_{i \in I} \mathcal{A}_i)/\mathcal{U} \models \psi.$$

By induction hypothesis, this is equivalent to

$$\{i: \mathcal{A}_i \models \phi\} \in \mathcal{U} \text{ or } \{i: \mathcal{A}_i \models \psi\} \in \mathcal{U},$$

both of which imply

$$\{i: \mathcal{A}_i \models \phi \lor \psi\} \in \mathcal{U}.$$

 $\{i: \mathcal{A}_i \models \phi \lor \psi\}$ is $\{i: \mathcal{A}_i \models \phi\} \cup \{i: \mathcal{A}_i \models \psi\}$. Now we exploit the fact that \mathcal{U} is ultra: for all A and B it contains $A \cup B$ iff it contains at least one of A and B, which enables us to reverse the last implication.

Negation

2828 We assume

$$(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U}\models\phi \text{ iff } \{i:\mathcal{A}_i\models\phi\}\in\mathcal{U}$$

and wish to infer

$$(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U}\models\neg\phi \text{ iff } \{i:\mathcal{A}_i\models\neg\phi\}\in\mathcal{U}.$$

Suppose $(\prod_{i \in I} \mathcal{A}_i)/\mathcal{U} \models \neg \phi$. That is to say,

$$(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U}
ot\models\phi.$$

By induction hypothesis this is equivalent to

$$\{i: \mathcal{A}_i \models \phi\} \notin \mathcal{U}.$$

But, since \mathcal{U} is ultra, it must contain I' or $I \setminus I'$ for any $I' \subseteq I$, so this last line is equivalent to

$$\{i: \mathcal{A}_i \models \neg \phi\} \in \mathcal{U},$$

as desired.

2833 Existential quantifier

The step for \exists is also nontrivial:

$$(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U}\models\exists x\phi$$

$$(\exists f)(\prod_{i\in I}\mathcal{A}_i)/\mathcal{U}\models\phi(f)$$

$$(\exists f)\{i \in I : \mathcal{A}_i \models \phi(f(i))\} \in \mathcal{U},$$

and here we use the axiom of choice to pick a witness at each factor

$$\{i \in I : \mathcal{A}_i \models \exists x \phi(x)\} \in \mathcal{U}.$$

2834

2835

2836

2837

2838

2839

2840

2841

2842

2843

2844

2845

2846

2848

2849

2851

2852

2832

You will notice that in the induction step for the existential quantifier you use the axiom of choice to pick a witness from each factor, and this use of AC seems unavoidable. This might lead you to suppose that Łoś's theorem is actually equivalent to AC (after all, it implies that an ultraproduct of nonempty structures is nonempty and that sounds very like AC) but this seems not to be the case. Try it! I am endebted to Phil Freeman for drawing my attention to Paul Howard, Proc Am Math Soc Vol. 49, No. 2, Jun., 1975.

This has the incredibly useful corollary (which we shall not prove) that

COROLLARY 14 A formula is equivalent to a first-order formula iff the class of its models is closed under elementary equivalence and taking ultraproducts.

Theorem 21 enables us to show that a lot of things are not expressible in any first order language. Since, for example, an ultraproduct of finite *p*-groups (which are all simple) is not simple, it follows that the property of being a simple group is not capturable by a language in which you are allowed to quantify only over elements of the object in question.

Miniexercise: If the ultrafilter is principal ($\{J \subseteq I : i \in J\}$), then the ultraproduct is isomorphic to the *i*th factor. So principal ultrafilters are no use.

In contrast if the ultrafilter is nonprincipal you can make good use of the construction even if all the models you feed into it are the *same*.

2858

2860

2863

2864

2866

2867

2868

2870

2871

2872

2873

2876

2878

2879

2880

2881

2882

2884

2885

2886

DEFINITION 54 If all the factors are the same, the ultraproduct is called an **ultra**power, and we write ' A^{κ}/U ' for the ultraproduct where there are κ copies of A, indexed by a set K of size κ and U is an ultrafilter on K.

Not only are \mathfrak{M} and the ultrapower $\mathfrak{M}^{\kappa}/\mathcal{U}$ elementarily equivalent by Łoś's theorem, we also have the following, of which we will make frequent use.

DEFINITION 55 Elementary Embedding

An embedding $i: \mathfrak{M} \hookrightarrow \mathfrak{N}$ is **elementary** iff

$$\mathfrak{M} \models \phi(x_1, \dots x_k) \text{ iff } \mathfrak{N} \models \phi(i(x_1) \dots i(x_k))$$

for all tuples $x_1 \dots x_k$ in \mathfrak{M} and all $\phi \in \mathcal{L}(\mathfrak{M})$.

For example: the identity map is an elementary embedding from \mathbb{Q} into \mathbb{R} , where both these structures are thought of as ordered abelian groups. Not as fields! (Why not?)

LEMMA 12 The function that sends an m in \mathfrak{M} to the equivalence class (in $\mathfrak{M}^{\kappa}/\mathcal{U}$) of the function in \mathfrak{M}^{κ} that takes the value m everywhere is elementary.

This embedding *i* is just a typed version of the *K* combinator! *Proof:*

It will be sufficient to show that, for any $m \in \mathfrak{M}$, if there is an $x \in \mathfrak{M}^{\kappa}/\mathcal{U}$ such that $\mathfrak{M}^{\kappa}/\mathcal{U} \models \phi(x, i(m))$ then there is $x \in \mathfrak{M}$ s.t. $\mathfrak{M} \models \phi(x, m)$. Consider such an $x \in \mathfrak{M}^{\kappa}/\mathcal{U}$. It is the equivalence class of a family of functions which almost everywhere (in the sense of \mathcal{U}) are related to m by ϕ so – by Łoś's theorem – there must be something x in \mathfrak{M} such that $\mathfrak{M} \models \phi(x, m)$. Then $i \mapsto x$ will do.

If you are doing Set Theory you will see the utility of this later in connection with measurable cardinals.

Ultraproducts enable us to give a particularly slick proof of the compactness theorem for predicate calculus.

THEOREM 22 (Compactness theorem for predicate logic)

Every finitely satisfiable set of sentences of predicate calculus has a model.

Proof: Let Δ be a set of wffs that is finitely satisfiable. Let S be the set of finite subsets of Δ (elsewhere in these notes notated $\mathcal{PN}_0(\Delta)$), and let $X_s = \{t \in S : s \subseteq t\}$. Pick $\mathfrak{M}_s \models s$ for each $s \in S$. Notice that $\{X_s : s \in S\}$ generates a proper filter. Extend this to an ultrafilter \mathcal{U} on S. Then

$$(\prod_{s\in S}\mathfrak{M}_s)/\mathcal{U}\models\Delta.$$

This is because, for any $\phi \in \Delta$, $X_{\{\phi\}}$ is one of the sets that generated the filter that was extended to \mathcal{U} . For any $s \in X_{\{\phi\}}$, $\mathfrak{M}_s \models \phi$, so $\{s : \mathfrak{M}_s \models \phi\} \in \mathcal{U}$.

Notice we are not making any assumption that the language is countable.

Notice the relation between Arrow's paradox and the nonexistence of nonprincipal ultrafilters on finite sets. Consider an ultraproduct of finitely many linear orders: it must be isomorphic to one of the factors. This is Arrow's "dictatorship" condition.

Tidy this up!!

... Chapter 8

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
signals a particularly tasty question.
```

8.1 Sheet 0: Countability

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.

97 8.2 Sheet 1: Mainly Ordinals

- 1. Write down subsets of \mathbb{R} of order types $\omega + \omega$, ω^2 and ω^3 in the inherited order.
- 2. Which of the following are true?

2899

2900

2902

2904

2905

2906

2907

2908

- (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.
 - 3. Consider the two functions $On \to On$: $\alpha \mapsto 2^{\alpha}$ and $\alpha \mapsto \alpha^2$. Are they normal?
 - 4. Prove the converse to lemma 1.2: if $\langle X, <_X \rangle$ is a total order satisfying "every subordering is isomorphic to an initial segment" then it is a wellordering.
 - 5. What is the smallest ordinal you can not embed in the reals in the style of question (1) on this sheet?

2914

2915

2916

2917

2918

2919

2927

2928

2929

2932

2933

2934

2935

- 2910 6. Prove that every [nonzero] countable limit ordinal has cofinality ω. What about $ω_1$?
 - 7. Recall the recursive definition of ordinal exponentiation:

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

Ordinal addition corresponds to disjoint union [of wellorderings], ordinal multiplication corresponds 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: $\alpha^{\beta+\gamma} = \alpha^{\beta} \cdot \alpha^{\gamma}$.

- 8. 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, $(\forall i, j \in I)(X_i \subseteq X_j \vee X_j \subseteq X_i)$.]
- 9. Prove that every ordinal of the form ω^{α} is **indecomposible**:

$$\gamma + \beta = \omega^{\alpha} \rightarrow \gamma = \omega^{\alpha} \vee \beta = \omega^{\alpha}.$$

- 10. Show that an arbitrary intersection of transitive relations is transitive.
- 2925 11. Let $\langle X, R \rangle$ be a wellfounded binary structure, with rank function ρ . 2926 Prove that $(\forall x \in X)(\forall \alpha < \rho(x))(\exists y \in X)(\rho(y) = \alpha)$.
 - 12. 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.

(The point is that any structure whatever can be obtained as a direct limit ("colimit") of its finitely generated substructures.)

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.

14. 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?

8.3 Sheet 2: Posets

1. For $n \in \mathbb{N}$,

2937

2945

2946

2953

2956

295

2958

2960

2961

2963

2964

2966

2967

2968

2972

2973

- (a) How many antisymmetrical binary relations are there on a set of cardinality n?

 How many binary relations satisfying *trichotomy*: $(\forall xy)(R(x,y) \lor R(y,x) \lor x = y)$?

 How are your two answers related?
- (b) How many *symmetric* relations are there on a set of cardinality *n*?

 How many *antisymmetric trichotomous* relations are there on a set of cardinality *n*?

 How are your two answers related?
- 2944 (c) Contrast (a) and (b).
 - 2. Draw Hasse diagrams of all the partial orders of sets with four elements. Indicate which of them are complete posets.
- 3. Consider the set of equivalence relations on a fixed set, partially ordered by ⊆. Show that it is a lattice. Must it be distributive? Is it complete?
- 2949 4. 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.
- 5. 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 $B \hookrightarrow A$.
 - 6. 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.
 - 7. $\mathbb{N} \to \mathbb{N}$ 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?
 - 8. For each of the following functions $\Phi : (\mathbb{N} \to \mathbb{N}) \to (\mathbb{N} \to \mathbb{N})$
 - (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 determine
 - (a) whether Φ is order-preserving, and
 - (b) whether it has a fixed point.

2978

2979

2980

2982

2983

2984

2986

2987

2990

2992

2993

2998

- 9. 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.
 - 10. Let *U* be an arbitrary set and let $\mathcal{P}(U)$ be the power set of *U*. For *X* a subset of $\mathcal{P}(U)$, the **dual** X^{\vee} of *X* is the set

$$\{y \subseteq U : (\forall x \in X)(y \cap x \neq \emptyset)\}.$$

- (a) Is the function $X \mapsto X^{\vee}$ monotone? Comment.
- (b) By considering the poset *P* 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 $X = X^{\vee}$.
- (c) $X^{\vee\vee}$ is clearly a superset of X, in that it contains every superset of every member of X. Does it consist solely of supersets of members of X? (That is, do we have $Y \in X^{\vee\vee} \to (\exists Z \in X)(Z \subseteq Y)$?)
- (d) Is $A^{\vee\vee\vee}$ always equal to A^{\vee} ?
- 11. 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.
 - $A^{<\omega}$ is the set of finite sequences from A. By considering the poset of partial functions $A^{<\omega} \to \{I\}$ or otherwise prove that if X is open then one of the two players must have a winning strategy.
- 12. $\mathbb{R} = \langle 0, 1, +, \times, \leq \rangle$ 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?
- 13. (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?

8.4 Sheet 3: Propositional and Predicate Logic

- 1. (This question is a sleeper for Banach-Tarski. Look at [6].)
 - (i) State and prove the Tarski-Knaster fixed point theorem for complete lattices.
 - (ii) Let X and Y be sets and $f: X \to Y$ and $g: Y \to X$ be injections. By considering $F: \mathcal{P}(X) \to \mathcal{P}(X)$ defined by

$$F(A) = X \setminus g''(Y \setminus f''X)$$

or otherwise, show that there is a bijection $h: X \to Y$.

Suppose U is a set equipped with a group Σ of permutations. We say that a map $s: X \to Y$ is piecewise- Σ just when there is a finite partition $X = X_1 \cup \ldots \cup X_n$ and $\sigma_1 \ldots \sigma_n \in \Sigma$, so that $s(x) = \sigma_i(x)$ for $x \in X_i$. Let X and Y be subsets of U, and $f: X \to Y$ and $g: Y \to X$ be piecewise- Σ injections. Show that there is a piecewise- Σ bijection $h: X \to Y$.

If $\langle P, \leq_P \rangle$ and $\langle Q, \leq_Q \rangle$ are two posets with order-preserving injections $f: P \to Q$ and $g: Q \to P$, must there be an isomorphism? Prove or give a counterexample.

2. Show how \land , \lor and \neg can each be defined in terms of \rightarrow and \bot .

Why can you not define \land in terms of \lor ?

Can you define \vee in terms of \rightarrow ?

3004

3005

3007

3008

3009

3010

301

3015

3016

3026

3027

Can you define \land in terms of \rightarrow and \lor ?

- 3. (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.
- Explain very briefly the relation between truth-tables and Disjunctive Normal Form.
- 5. 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 proper superset of it that is a propositional logic (closed under \models and substitution) is trivial (contains all well-formed formulæ).
- 6. 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-1.
 - Show that the theory of equality plus one wellfounded relation is not axiomatisable.

- 8. 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 2⁸⁰ 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?
 - 9. Show that monadic predicate logic (one-place predicate letters only, without equality and no function symbols) is decidable.
- 3033 10.

3032

3036

3037

3038

- (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).
- 11. Why does T not follow from K and S? Show that Peirce's Law: $((A \rightarrow B) \rightarrow A) \rightarrow A$ cannot be deduced from K and S.
 - 12. 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.
 - 13. Is

$$(\forall x)(\exists y)(F(x,y)) \to (\forall x)(\exists y)(\forall x')(\exists y')[F(x,y) \land F(x',y') \land (x=x' \to y=y')]$$

valid?

- 14. (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 algebraic iff every axiom of *T* is of the form (∀x̄)Φ where φ is a conjunction of equations between *T*-terms.
 Prove that, if *T* is algebraic, 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 algebraic?

15.

3041

3042

3043

3045

3046

3047

3050

3054

3058

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 .

- 16. This question is a sleeper for **NP-completeness**.
 - 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 presumption that $\mathcal{L}(\phi') = \mathcal{L}(\phi)$.)

8.5 Sheet 4: More Predicate Logic and Some Set Theory

- 1. Let us say A@B is $\{\{a,b\}, a \in A \land b \in B\}$. If A and B are not disjoint then A@B might contain singletons.
- If I give you a set X of pairs-and-singletons that happens to be of the form A@B can you recover A and B?

3069

3070

3071

3072

3075

3077

3078

3080

3081

3082

3083

3084

3085

3086

3087

3089

3090

3091

3094

3096

3097

- 2. Show that if x is a transitive set, then so are $\bigcup x$ and $\mathcal{P}(x)$. Are the converses true?
- 3066 3. Show that the Pair-set axiom is deducible from the axioms of empty set, power set, and replacement.
 - 4. Show that $\{z : \neg(\exists u_1, \dots, u_n)((z \in u_1) \land (u_1 \in u_2) \land \dots \land (u_n \in z))\}$ is not a set for any n. What assumptions have you made?
 - 5. 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?
 - 6. (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 $On \rightarrow On$ is itself normal.
 - (c) If α is an ordinal and f is a normal function show that f has a fixed point of cofinality $cf(\alpha)$.
 - (d) Are any of your fixed points regular?
 - (e) If α is a regular ordinal and f is a normal function show that f has a fixed point of cofinality α .
 - 7. Show that the axiom of choice follows from the assumption that cardinals are totally ordered by \leq_{card} . (This is the other direction of sheet 2 question 5.)
 - 8. Explain briefly the equivalence of the four versions of the axiom of foundation given in lectures:
 - (i) The axiom scheme of ∈-induction;
 - (ii) Every set is wellfounded;
 - (iii) Axiom of Regularity;
 - (iv) Every set belongs to the cumulative hierarchy.
 - 9. f is an \in -automorphism if f is a permutation of V that preserves \in :

$$(\forall xy)(x \in y \longleftrightarrow f(x) \in f(y)).$$

Show that a model of ZF (with foundation of course) can have no nontrivial ∈-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.

- 10. Recall that ρ is set-theoretic rank. If $\langle x, y \rangle$ is the Wiener-Kuratowski ordered pair show that $\rho(\langle x, y \rangle) = \max(\rho(x), \rho(y)) + 2$.
- (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 $\rho(\langle x, y \rangle) = \max(\rho(x), \rho(y)) + 1$?

8.6. SHEET 5 105

- (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?
- 11. 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?
- 12. Let *G* be a graph where, for each vertex v, the collection N(v) of neighbours of v is a set. (v' is a neighbour of v iff there is an edge between v and v').
- Give an example to show that G might be a proper class.
- Now suppose G is connected; prove that it is a set.
- What axioms have you used?

3114

3123

3128

- 13. 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 $\langle \mathbb{N}, E \rangle$?
 - 14. Prove that, for each $n \in \mathbb{N}$, there is a set of size \aleph_n . Is there a set of size \aleph_ω ?
- Assume that the cartesian product $x \times y$ always exists however you implement ordered pairs. Infer the axiom scheme of replacement.
- 16. 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} ?
- 17. 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 $A_i \to B_i$. Show that there is no surjection $\bigcup_{i \in I} A_i \to \prod_{i \in I} B_i$.
- 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.
- 18. Let $\{A_i : i \in \mathbb{N}\}$ be a family of finite structures, and \mathcal{U} a nonprincipal ultrafilter on \mathbb{N} . Show that the ultraproduct is finite if there is a finite bound on the size of A_i and is of size $2^{\mathbb{N}_0}$ if every infinite subset of $\{A_i : i \in \mathbb{N}\}$ contains arbitrarily large elements.

8.6 Sheet 5

- This sheet is for enthusiasts who want to take this stuff further; it's a mixture of revision, consolidation-with-backfill and looking-ahead.
- 1. Explain to members of your tutorial group (or to anyone listening who might be confused) the difference between

3149

3150

3151

3152

3153

3157

3160

3161

3162

3163

- (i) Nonstandard naturals
- (ii) Countable ordinals
- (iii) Infinite Dedekind-finite cardinals
- 2. 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 $P^* \hookrightarrow P$.
- 31. 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
- How do you have to spice up the denseness condition to prove an analogous result for linear orders of size \aleph_1 ?
- 4. 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 learnt to know and love are recursive?

 Come to think of it, are *all* countable ordinals recursive?
 - 5. 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.").
 - Using propositional logic only, show that a(n undirected) graph and its complement cannot both be disconnected. Hint: propositional letters will correspond to edges.
 - 7. (This question and the next are tasters for *forcing*)
- A poset $\langle P, \leq_P \rangle$ is [upwards] separative if

$$(\forall x, y \in P)(x \nleq y \to (\exists z \ge y)(\forall w)(w \ngeq z \lor w \ngeq x)).$$

- 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 ⊆.
 - The set $\{f : f \text{ is an injection from some set of countable ordinals } \hookrightarrow \mathbb{R}\}$ ordered by \subseteq . Think of f as a set of ordered pairs.
 - 8. 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"). 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.

8.6. SHEET 5

(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 $d \le p$.
- Suppose X is a collection of subsets of P. We say that $G \subseteq P$ is X-generic if G has nonempty intersection with every downwards dense element of X.
 - We say that G is a filter if

3173

3174

3175

3176

3177

3178

3184

- 1. for any x, y in G there is z in G such that $z \le x$ and $z \le y$, and
- 2. for any x in G and $x \le y$, we have y in G.
- (c) If X is countable, show that there is an X-generic filter.
- (d) Let $\langle P, \leq \rangle$ be a downwards separative poset with no minimal elements and let X be a collection of subsets of P closed under complementation (i.e., if $X \in X$, then also $(P \setminus X) \in X$). Show that if G is an X-generic filter, then $G \notin X$.
- (e) Let $\langle P, \leq \rangle$ be $\{0, 1\}^{<\omega}$, 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 X 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.
 - 9. (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 $\sigma_{i,i} : A_i \longrightarrow A_i$ be a surjection, and let the surjections commute.
- Show that there is a set A_I with, for each $i \in I$, a surjection $\pi_i : A_I \longrightarrow A_i$.
- 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 $B \twoheadrightarrow A_I$.

10. 💂

3198

3199

3200

3201

Let G be the alternating group of permutations of V_{ω} . Any $\pi \in G$ can be extended to a permutation acting on the whole universe by fixing everything not in V_{ω} , and we still write it π .

3207

3208

3209

3210

3211

3213

3214

3215

3216

3218

3221

3223

3224

3226

3227

3228

3229

3231

3234

3235

3237

For each $n \in \mathbb{N}$ and any x whatever such a π can move x by acting "n levels 3202 down", on¹ \bigcup $^{n} x$. 3203

> 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?

11. (A taster for Large Cardinals)

Prove Łoś's theorem (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?

12.

(a) Let T be a consistent theory in a language containing a (possibly complex) expression $\phi(-,-)$ which T proves to be an infinite total order. Let $\mathcal{I} = \langle I, <_I \rangle$ be a total ordering. Show that T has a model in which \mathcal{I} is embedded as part of the graph of ϕ .

(b) $\mathcal{I} = \langle I, \leq_{\mathcal{I}} \rangle$ 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 \vec{I} taken in $\leq_{\vec{I}}$ -increasing order we have $\mathfrak{M} \models \phi(\vec{x}) \longleftrightarrow \phi(\vec{y}).$

Now let 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 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.

13. Wikipædia says:

Commutative Rings ⊇ Integral Domains ⊇ Integrally Closed Domains ⊇ GCD domains ⊇ Unique Factorization Domains ⊇ Principal Ideal Domains ⊇ Euclidean Domains ⊇ Fields

This might make a good openbook exam question!

¹Any group of permutations of a set X acts on $\mathcal{P}_{\aleph_0}(X)$ (the set of finite subsets of X) in an obvious way.

8.6. SHEET 5

All these families-of-structures can be thought of as belonging to the one signature: $0, 1, +, \cdot$ 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. Comment on whether or not each theory is finitely axiomatisable.

- 14. How many countable [linear] order types are there whose automorphism group is transitive on singletons?
- 15. (i) How many transitive subsets of V_{ω} are there?

3243

3244

3246

3247

3249

3250

3251

325

3258

3259

3261

- (ii) How many transitive sets are there all of whose members are countable?
- 16. Recall the difference between a **wellorderable** set and a **wellordered** set.
- Prove, without using AC or foundation or ordinals, that every collection 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?
- 17. Prove that a directed limit (colimit) of wellfounded structures under end-extension is wellfounded.
- 18. A **circular order** (see [3]) is a ternary relation R(x, y, z), whose typical example is the relation that holds between points x, y and z on the unit circle if, starting from x and moving clockwise, one encounters y before z.
 - (1) Write down a set of axioms for circular orders.
 - A group is **circularly-orderable** if it has a circular ordering that interacts in the obvious way with the multiplication of the group. The typical example is the additive group of integers-mod-*p*.
 - (2) Write down a set of axioms for circularly orderable groups.
- (3) Prove that a group is circularly orderable iff all its finitely generated subgroups are circularly orderable.
- (4) Is the multiplicative group of (nonzero) integers mod *p* (*p* prime) circularly ordered?

Bibliography

- 3267 [1] www.dpmms.cam.ac.uk/~tf/cam_only/countability.pdf.
- ³²⁶⁸ [2] Forster, Logic, Induction and Sets.
- ³²⁶⁹ [3] Edward V. Huntingdon 'Inter-relations among the four principle types of order' Transactions of the American Mathematical Society **38** (1935) pp 1–9.
- 3271 [4] Peter Johnstone "Notes on Set Theory and Logic" CUP
- [5] J. L. Bell and Alan Slomson "Models and Ultraproducts" North-Holland, reprinted
 by Dover.
- 3274 [6] Stan Wagon, The Banach-Tarski Paradox. Cambridge University Press.
- ³²⁷⁵ [7] Scott, D.S. "Semantical Archæology, a parable" In: Harman and Davidson eds, Semantics of Natural Languages. Reidel 1972 pp 666–674.