tures through equivalence classes. In Chapter Five, these structures are applied to problems in the semantics of expressions indicating change and becoming.

Chapters Six and Seven discuss lattices. Chapter Six deals with the properties of lattices, join semilattices, and Boolean algebras; it further extends the discussion of homomorphisms and introduces generated and free lattices; and it discusses filters, ideals and set theoretic representation of distributive lattices, join semilattices and Boolean algebras. Chapter Seven discusses the relation between Boolean algebras and the types for noun phrases and verbs; and it discusses the relation between join semilattices and the semantics of plurals and mass nouns.

I am greatly indebted to my students at Cornell who have been exposed to previous versions of this book and helped improve it enormously. Manfred Krifka's very extensive, detailed and helpful comments were of invaluable use in shaping the final version. I thank my friends Nirit Kadmon and Craige Roberts for telling me and reminding me, from the time we first met in 1984 onwards, that I should write this book and what it should be like. Veneeta Srivastav carries the most direct responsibility for my having written this book; she forced me to teach, in the spring of 1988, the course out of which this book grew. I am very grateful for that.

I want to express a great intellectual debt to Hans Kamp. In reading this book (and the topics that didn't make it into this book), it should be obvious how much Hans' thinking has shaped semantics into what it is today.

Finally, I want to thank my friends and colleagues, Gennaro Chierchia, Nirit Kadmon and Sally McConnell-Ginet for their continuous support, moral and otherwise, while writing this book.

CHAPTER ONE

LOGIC AND SET THEORY

Let me first state what the present chapter does not do. It is not intended to give an introduction to logic or set theory. As mentioned in the Preface, such an introduction is presupposed. Secondly, it is not meant to give to any satisfactory degree an overview of the basic results and techniques in logic (see for this Hodges, 1983).

With respect to logic and set theory, semanticists are on the consumers' side. The fruitful application of logical techniques to semantic problems does not require one to be a logician. Yet, something can be said for the view that it is a good thing if semanticists have some idea of what it is that logicians do with these techniques. For one thing, it may be of use if you want to ask your logician-friends for help: you don't have to switch to the automatic pilot at the first mention of, say, completeness or comprehension. Secondly, it may help to get a better grip on what the use of these techniques is committing you to. Thirdly, it is, of course, so inherently rewarding if some such knowledge is part of your intellectual baggage.

More importantly, certain logical concepts are used and mentioned (and, sometimes, mis-used and mis-mentioned) so frequently in the literature, that their proper understanding may make the literature more accessible. The most central of these, and the one that I will focus on, is the notion of 'first order'.

Probably most people who have gone through introductory semantics courses know that 'first order' has to do with quantification over individuals, that many semanticists prefer 'higher order' while many logicians prefer 'first order', and that certain things 'are not first order'. I think that, for the reasons mentioned, some deeper understanding of the notion of 'first order' is useful.

The basic aim of this chapter, then, is to shed some light on notions like first order logic, first order theories and first order definability, in short, on what it means to be first order.

1.1. FIRST ORDER LOGIC

1.1.1. *Basic Concepts*

First order predicate logic

To set the stage I will start with a definition of the syntax, semantics and logic of elementary predicate logic or first order predicate logic (with identity).

Syntax

We will define the notion of a *first order language*. We start by specifying the notion of a *first order lexicon*.

A *first order lexicon* L has the following ingredients:

L has basic symbols: $\neg \wedge \exists = (\ ),$

These symbols are the same for every first order lexicon.
L contains a set of individual constants:

$$\mathrm{CON}_L = \{c_1, c_2, \ldots\}$$

This set is particular to L and may be empty, finite or countably infinite.
L contains a set VAR of countably many individual variables:

$$VAR = \{x_1, x_2, \ldots\}$$

This set is the same for every first order lexicon.
For every $n > 0$, L contains a set of n-place predicate constants:

$$\text{for every } n > 0\text{: } PRED^n_L = \{P^n_1, P^n_2, \ldots\}$$

Also these sets $PRED^n_L$ are particular to L and may be empty, finite or countable.

Every first order language is based on a first order lexicon.

All these languages will have the same syntactic rules, they only differ in their lexicon, and hence they only differ in what their individual constants and predicate constants are. Given the rules to follow, once we specify the lexicon of a language, we know exactly which first order language we have. Given this, we will use the same name L both for the lexicon and the language and we understand that first order language L is the unique first order language based on lexicon L.

Given a lexicon L, the language L is defined by the following definition of the terms, atomic formulas and formulas of L.

Let L be a first order lexicon.

The set of terms of L, $TERM_L$, is given by:

$$TERM_L = CON_L \cup VAR$$

The set of atomic formulas of L, $ATFORM_L$, is given by:

$ATFORM_L$ is the smallest set such that:

1. If $P \in PRED^n_L$ and $t_1, \ldots, t_n \in TERM_L$ (for any n) then

   $P(t_1, \ldots, t_n) \in ATFORM_L$

2. If $t_1, t_2 \in TERM_L$ then $t_1 = t_2 \in ATFORM_L$

The set of formulas of L, $FORM_L$, is the smallest set such that:

1. $ATFORM_L \subseteq FORM_L$
2. If $\varphi \in FORM_L$ then $\neg\varphi \in FORM_L$
3. If $\varphi, \psi \in FORM_L$ then $(\varphi \wedge \psi) \in FORM_L$
4. If $x \in VAR$ and $\varphi \in FORM_L$ then $\exists x\varphi \in FORM_L$

As said before, every first order language is based on the same set of basic symbols: $\neg \ \wedge \ \exists =) \ (\ ,$. We call these the *logical constants*. Not only are they the same in every language, but they also have the same interpretation in all models (given by the truth definition), as we will see. Other logical constants can be defined:

$$(\varphi \vee \psi) := \neg(\neg\varphi \wedge \neg\psi)$$
$$(\varphi \rightarrow \psi) := (\neg\varphi \vee \psi)$$
$$\forall x\varphi := \neg\exists x \neg \varphi$$

I will not be very consistent in the choice of primitives.

First order languages differ in their choice of *non-logical individual and predicate constants* (CON_L and $PRED^n_L$), and, as we will see, they are non-logical precisely in that their interpretation can vary in different models.

Predicates can only apply to individual level terms, and there are only individual level variables. So we can only predicate things of individuals and we can only quantify over individuals. The first order is the order of individuals. Although we can express that a certain individual has a certain property or that certain individuals stand in a certain relation, we cannot quantify over such properties of individuals or relations between them. The second order is the order of such properties and relations. Generally, a first order language is a language in which we can only quantify over first order entities, i.e. individuals.

First order languages are unambiguous: every formula is built in a unique way by the above rules. If φ and ψ are formulas of L then ψ is a subformula of φ if ψ is used in the rules for building up φ. ψ can be used more than once in building up φ. In that case we say that there are different occurrences of subformula ψ in φ.

If φ is a formula of L and there is an occurrence of subformula $\exists x_n\ \psi$ or $\forall x_n\ \psi$ (i.e. $\neg\exists x_n \neg\ \psi$) in φ, then we say that the *scope* of that occurrence of quantifier $\exists x_n$ or $\forall x_n$ is ψ.

An occurrence of a variable x_n in φ is *bound* iff it is in the scope of some occurrence of a quantifier $\exists x_n$ or $\forall x_n$ in φ, otherwise it is *free*.

A variable x_n is *bound* in φ iff every occurrence of x_n in φ is bound, otherwise x_n is free.

A *sentence* is a formula without free variables.

Semantics

Let L be a first order language as specified above.

A *model for* L is a pair $M = \langle D, i\rangle$, where:

> D is a nonempty set (of individuals)
> i is an interpretation function for the non-logical constants of L: i is a function such that:
> for every $c \in CON_L$: $i(c) \in D$
> for every $P \in PRED^n_L$ (for any n): $i(P) \subseteq D^n$

So every individual constant is interpreted as an individual and every n-place predicate is interpreted as a set of n-tuples of individuals.

An *assignment function* is a function g from VAR to D, i.e. g assigns every variable a value in D.

A model gives us an interpretation for the non-logical constants, an assignment function gives us an interpretation for the variables. Let us give a notation for the interpretation of terms and predicates in a model M ($= \langle D, i\rangle$) relative to an assignment function g:

– terms $\quad [\![c]\!]_{M,g} = i(c) \quad$ if $\quad c \in CON_L$
$\quad\quad\quad\ \ [\![x]\!]_{M,g} = g(x) \quad$ if $\quad x \in VAR$

– predicates: $[\![P]\!]_{M,g} = i(P) \quad$ if $\quad P \in PRED^n_L$

The *truth definition* gives us an interpretation for the whole language ($FORM_L$), based on the above interpretation of the constants and variables.

There are two standard ways of giving the truth definition.

Either you define $[\![\varphi]\!]_{M,g}$, the truth value of formula φ in model M relative to assignment g.

Or you define truth as a relation between a formula φ, a model M and an assignment function g:

$$M \models \varphi[g] \qquad \varphi \text{ is true in } M \text{ relative to } g$$

(φ is false in M relative to g iff φ is not true in M relative to g. We use notation $M \not\models \varphi[g]$ to indicate: it is not the case that $M \models \varphi[g]$.)

Here I will use the second one.

TRUTH DEFINITION

Let $M = \langle D, i\rangle$ be a model for L and g an assignment function: we specify $M \models \varphi[g]$ for any formula of L.

Atomic formulas:

1. $M \models P(t_1, \ldots, t_n)[g]$ iff $\langle [\![t_1]\!]_{M,g}, \ldots, [\![t_n]\!]_{M,g}\rangle \in i(P)$
2. $M \models t_1 = t_2[g]$ iff $[\![t_1]\!]_{M,g} = [\![t_2]\!]_{M,g}$

Formulas:

3. $M \models \neg\varphi[g]$ iff $M \not\models \varphi[g]$
4. $M \models (\varphi \wedge \psi)[g]$ iff $M \models \varphi[g]$ and $M \models \psi[g]$

Given assignment g, g^d_x is that assignment that at most differs from g in that g^d_x assigns d to variable x.

5. $M \models \exists x\varphi[g]$ iff for some $d \in D$: $M \models \varphi[g^d_x]$

Now that we have defined for every formula when it is true or false in a model relative to an assignment function, we can define when a formula is true or false in a model (independently of an assignment function). These notions will not always be defined for formulas with free variables, but they will be for sentences:

Let M be a model and φ a formula.

Truth:

$$\varphi \text{ is } \textit{true} \text{ in } M,\ M \models \varphi \text{ iff for all } g\text{: } M \models \varphi[g]$$

φ is *false* in M, $M \not\models \varphi$ iff for all g: $M \not\models \varphi[g]$

Exercise 1. The following lemma holds:

LEMMA. *If φ is a sentence then:*

for all g: $M \models \varphi[g]$ iff for some g: $M \models \varphi[g]$
for all g: $M \not\models \varphi[g]$ iff some g: $M \not\models \varphi[g]$

(a) Show that if φ is a sentence and M a model then φ is true in M or φ is false in M
(b) Let P be a predicate and M a model such that $i(P) \neq \emptyset$ and $i(P) \neq D$ (i.e. there are Ps, but not everything is a P). Show that $P(x)$ and $\neg P(x)$ are neither true nor false in M.
(c) The same predicate and model. Show whether $(P(x) \vee \neg P(x))$ and $(P(x) \wedge \neg P(x))$ are true, false, or undefined in M.

In general we are not very much interested in formulas, so we will will often restrict ourselves to sentences.

Now that we have given the truth definition, we can define the notions of validity and entailment.

An *argument* Δ/φ consists of a (empty, finite or infinite) set of sentences Δ, the premises of the argument, and a sentence φ, the conclusion of the argument.

Let Δ be a set of sentences and φ a sentence.

We say that argument Δ/φ is valid iff Δ *entails* φ, where entailment is defined as follows:

Δ *entails* φ, $\Delta \models \varphi$ iff for every model M: if for every $\delta \in \Delta$: $M \models \delta$ then $M \models \varphi$

Note that we use the same symbol for entailment and truth. No confusion should arise though, because when we use $\models$ for entailment it is a relation between a set of sentences and a sentence, and when we use $\models$ for truth it is a relation between a model and a sentence.

So an argument from premises Δ to conclusion φ is valid if every model in which the premises are true is also a model in which the conclusion is true.

We write $\models \varphi$ for $\emptyset \models \varphi$. In other words,

φ *is logically valid*, $\models \varphi$ iff for every M: $M \models \varphi$

So validity means truth in all models.

φ and ψ are logically equivalent iff $\varphi \models \psi$ and $\psi \models \varphi$.

Exercise 2. This exercise continues the last one. We have defined entailment as a relation between a set of sentences and a sentence. Let us assume that we give exactly the same definition of entailment for Δ and φ where Δ is a set of formulas and φ is a formula.

Let $\varphi[x]$ be any formula with free variable x.

Prove that $\varphi[x]$ and $\forall x \varphi[x]$ are logically equivalent.

Logic

There are various ways of giving a proof system.

If you want to learn how to prove things in predicate logic, *natural deduction systems* are usually the best. A natural deduction system is a system of rules that tells you for each logical constant how you can legitimately introduce that constant in a proof and, if it already occurs in a proof, how you can legitimately use (exploit, eliminate) it in a proof. I will give a natural deduction system for predicate logic in Chapter Three.

In talking and proving things *about* predicate logic an *axiomatic theory* is often the easiest. An axiomatic theory consists of a list of *axioms* (formulas) and *inference rules* to derive new formulas from axioms and already derived formulas. It is very common to formulate the axiomatic theory in the form of *axiom schemata* in the metalanguage. That is, instead of listing one particular formula of L as an axiom and having a substitution rule deriving all L-formulas with the same form, you give the schema, form, of that formula and say: all L-formulas of this form are axioms.

I will present an axiom system for first order predicate logic with identity.

First two more notions.

$\varphi[t/x]$ is the result of substituting term t for every free occurrence of variable x in φ.

t is *substitutable for* x in φ if no occurrence of t in $\varphi[t/x]$ is bound where x was free in φ.

Logic L0

Definitions of non-primitive logical constants: these are the definitions of $\vee$, $\rightarrow$, $\forall x$ that were given above.

Propositional axioms:

A1: $\varphi \to (\psi \to \varphi)$
A2: $(\varphi \to (\psi \to \chi) \to ((\varphi \to \psi) \to (\varphi \to \chi))$
A3: $(\neg \psi \to \neg \varphi) \to (\varphi \to \psi)$

Quantifier axioms:

A4: $\forall x(\varphi \to \psi) \to (\varphi \to \forall x\psi)$ if x is not free in φ
A5: $\forall x\varphi \to \varphi[t/x]$ if t is substitutable for x in φ

Identity axioms:

A6: $x = x$
A7: $x = y \to (\varphi \to \varphi[y/x])$ if y is substitutable for x in φ

Inference rules:

MP: From φ and $(\varphi \to \psi)$ infer ψ *Modus Ponens*
GEN: From φ infer $\forall x\varphi$ *Generalization*

Let us turn to the notion of proof.

A *proof* of φ from premises $\varphi_1, \ldots, \varphi_n$ is a *finite* sequence of formulas $\langle\psi_1, \ldots, \psi_m\rangle$ such that:

1. $\psi_m = \varphi$
2. *each* ψ_i is either (a) an axiom or (b) a premise or (c) inferred with MP and GEN from earlier formulas in the sequence.

Furthermore we assume that we can freely use definitions and that once we've proved a formula, we can use it as an axiom schema: we've proved every formula with that same form.

Let Δ be a set of formulas and φ a formula.

We write $\Delta \vdash \varphi$ (or more precisely $\Delta \vdash_{L0} \varphi$), φ is *provable from* Δ, φ is *derivable from* Δ iff there is an $L0$-proof of φ from premises $\delta_1, \ldots, \delta_n \in \Delta$.

We write $\vdash \varphi$, φ is *provable*, φ is a *tautology* for $\emptyset \vdash \varphi$, that is, φ is provable from the axioms and inference rules without further assumptions.

φ is a *contradiction* iff $\vdash \neg \varphi$.

It is common to use a special symbol $\perp$ to stand for an arbitrary contradiction.

Note that if α is an axiom, then $\vdash \alpha$.

Let me as an example of an axiomatic proof derive the law of excluded middle, $\neg p \vee p$, i.e. we want to prove $\vdash \neg p \vee p$. Here is the proof:

(1) $p \to ((q \to p) \to p)$

This formula is an instance of A1.

(2) $(p \to ((q \to p) \to p)) \to ((p \to (q \to p)) \to (p \to p))$

This is an instance of A2.

(3) $(p \to (q \to p)) \to (p \to p)$

This is derived from (1) and (2) with MP.

(4) $(p \to (q \to p)$

Another instance of A1.

(5) $p \to p$

This is derived from (3) and (4) with MP

(7) $\neg p \vee p$

This we get by applying the definition of $\to$.

We have discussed now how we can prove things in predicate logic (which is $L0$). Logicians, as we will see, are generally more interested in proving things about predicate logic. One very frequently used method of proving things about predicate logic is the *proof with induction to the complexity of formulas*, formula induction.

You want to prove that every formula of L has a certain property P. In an induction proof you do that by showing that the atomic formulas have P and that the rules for building complex formulas preserve P. A proof with formula induction always has the following form:

1. *Base step*.
   a. prove that $P(t_1, \ldots, t_n)$ has P
   b. prove that $t_1 = t_2$ has P
2. *Induction step*.
   a. assume that φ has P. This is called the *induction hypothesis*. Prove that $\neg\varphi$ has P.
   b. assume that φ has P and that ψ has P (two induction hypotheses). Prove that $(\varphi \wedge \psi)$ has P.

c. assume that φ has P. Prove that $\exists x\varphi$ has P.

If you can prove all of that, you know that every formula has P, because every formula is constructed from atomic formulas (that have P) by the construction rules (that preserve P).

Exercise 3. Prove with induction that every formula has an equal number of left and right brackets.

Theories

We have now introduced two fundamental relations between sets of formulas and formulas:

$\Delta \vdash_{L0} \varphi$, φ is derivable, provable from Δ in $L0$
$\Delta \models \varphi$, φ follows logically from Δ, Δ entails φ
with their borderline cases:
$\vdash_{L0} \varphi$, φ is a tautology (φ is provable)
$\models \varphi$, φ is logically valid (φ is true in all models)

A lot of classical results in logic concern the relation between these two notions. In order to talk about those and their consequences in the next section, it is convenient to broaden the perspective somewhat.

Let L be a first order language.
A *theory* in L is a set of sentences of L.
So a first order theory is a set of first order sentences.

Let Δ be a theory.
Δ is *inconsistent* iff $\Delta \vdash \bot$; consistent if $\Delta \nvdash \bot$.
A theory is inconsistent if you can derive a contradiction from it.
Δ is *deductively closed* iff if $\Delta \vdash \varphi$ then $\varphi \in \Delta$ (for all φ).
A theory is deductively closed if everything you can derive from it is already in it.
Δ is *maximally consistent* iff Δ is consistent and there is no Δ' such that $\Delta \subseteq \Delta'$ and $\Delta \neq \Delta'$ and Δ' is consistent.
A theory is maximally consistent if it is consistent and everything you add to it would make it inconsistent.
Δ^c, the *deductive closure* of Δ is the set $\{\varphi : \Delta \vdash \varphi\}$. So Δ^c is the result of adding to Δ every sentence that can be derived from Δ.

Δ is *complete* iff Δ^c, the deductive closure of Δ, is maximally consistent.
A theory is complete if adding any sentence to it that cannot be derived from it would make it inconsistent.

A remark on terminology here. We also say that a logic like $L0$ is complete in the sense that there is a completeness proof for it. This does not mean the same. Be sure to keep 'theory Δ is complete' and 'logic L is complete' apart. I will come back to it later.

Δ is a *set of axioms* for Γ iff $\Delta^c = \Gamma^c$.
Γ is *finitely axiomatizable* iff there is a finite set of axioms for Γ.

Some facts.

FACT 1: If Δ is consistent then Δ^c is consistent.

FACT 2: If Δ is maximally consistent then Δ is deductively closed.

FACT 3: (Deduction theorem) If $\Delta \cup \{\varphi\} \vdash \psi$ then $\Delta \vdash \varphi \to \psi$.

FACT 4: $\Delta \cup \{\varphi\}$ is inconsistent iff $\Delta \vdash \neg\varphi$.

FACT 5: (Lindenbaum's Lemma) Any consistent theory can be extended to a maximally consistent theory.

Exercise 4
(a) Prove fact 2.
(b) Prove that if Δ is maximally consistent then: $\Delta \vdash \varphi$ or $\Delta \vdash \neg\varphi$, for any sentence φ.

Having introduced some syntactic concepts (concerning $\vdash$), let us now define some similar semantic concepts.
Δ is *satisfiable*, Δ *has a model* iff there is a model M such that for all $\delta \in \Delta$: $M \models \delta$.
A theory Δ has a model if there is a model in which all the sentences in Δ are true.
Δ is *closed under logical consequence* iff if $\Delta \models \varphi$ then $\varphi \in \Delta$.

Some more notions

If φ is a sentence and $M \models \varphi$ we say that φ holds on M, also that M is a model for φ.

Similarly, we say that M is a model for theory Δ, and we write $M \models \Delta$, if M is a model for every $\delta \in \Delta$.

MOD(φ) is the class of all models for φ, i.e. the class of all models on which φ holds.

MOD(Δ) is the class of all models for Δ.

Let $\mathbf{M}$ be a class of models. We say that φ holds in the class $\mathbf{M}$ if φ holds in every model $M \in \mathbf{M}$ Similarly for Δ.

Let Δ be a theory and $\mathbf{M}$ a class of models.

Δ *axiomatizes* $\mathbf{M}$ iff $\mathbf{M} = \text{MOD}(\Delta)$.

The sentences in Δ are called the axioms for $\mathbf{M}$.

We call a class of models $\mathbf{M}$ *first order definable* iff

$$\mathbf{M} = \text{MOD}(\varphi)$$

for some first order sentence φ, that is, if there is a first order sentence φ, such that $\mathbf{M}$ is the class of all models where φ holds.

We are often interested in finding out whether a certain property of models is first order definable (like the property that a model has iff its domain has exactly three elements). What this means is that we want to find out whether the class of all models that have that property (like the class of all models whose domain has exactly three elements) is first order definable. The above definition tells us that this class is first order definable if there is a first order sentence that holds on all models that have that property and doesn't hold on all models that don't have that property (and in fact there is a first order sentence that holds on all models whose domain has exactly three elements and that does not hold on any other model, namely:

$$\exists x \exists y \exists z[\neg(x = y) \wedge \neg(x = z) \wedge \neg(y = z)$$
$$\wedge\ \forall u[(u = x) \vee (u = y) \vee (u = z)]])$$

First order definable means definable by a first order sentence. We have a similar notion for definability by a first order theory (the two notions are the same if the theory is finite, because then we can replace it by the conjunction of the sentences in it):

$\mathbf{M}$ is *generalized first order definable* iff $\mathbf{M} = \text{MOD}(\Delta)$, for some first order theory (set of first order sentences) Δ.

With these notions we have already advanced a lot in the field of *model theory*, in which the relation between classes of models and sentences, theories, is a primary object of study. In the next sub-section I will discuss the central theorem linking the notions of derivability and entailment and then discuss some facts concerning first order definability.

1.1.2. *Metalogic*

When we talk about first order logic we mean the logic $L0$ in relation to the semantics we have given. First order logic is characterized by three central theorems: the completeness theorem, the compactness theorem and the Löwenheim–Skolem theorem(s).

The completeness theorem

The (generalized) completeness theorem says the following: Let Δ be a theory and φ a sentence.

COMPLETENESS: $\Delta \vdash_{L0} \varphi$ *iff* $\Delta \models \varphi$

In other words, an argument is derivable iff it is valid.

This theorem implies the so called weak completeness theorem:

WEAK COMPLETENESS: $\vdash_{L0} \varphi$ *iff* $\models \varphi$

i.e. φ is a tautology iff φ is logically valid.

It is common to split the completeness theorem into its two parts:

Part I: SOUNDNESS: *if* $\Delta \vdash \varphi$ *then* $\Delta \models \varphi$

This theorem tells us that $L0$ is correct with respect to the semantics given: the derivations that you can make with the $L0$-rules are all semantically valid.

Part II: COMPLETENESS: *if* $\Delta \models \varphi$ *then* $\Delta \vdash \varphi$

This theorem tells us that $L0$ is complete with respect to the semantics given: every argument that is semantically valid can be derived with the $L0$-rules.

There is another formulation of the completeness theorem:

COMPLETENESS (formulation 2): Let Δ be any theory in L. Δ *is consistent iff* Δ *has a model.*

Again we can split this formulation into two parts:

Part I: SOUNDNESS: *if* Δ *has a model then* Δ *is consistent.*

Part II: COMPLETENESS: *if* Δ *is consistent then* Δ *has a model.*

Exercise 5. Show that the two formulations of soundness and of completeness are equivalent.

The soundness theorem is the easy (but tedious) part. It is also called the correctness theorem because, as mentioned above, it says that the axioms and inference rules that we have chosen are correct.

In order to prove this you primarily check that all the axioms are logically valid and that the inference rules preserve validity. If you have checked that, you are basically done, because, assume that you have a proof $\langle\psi_1, \ldots, \psi_n\rangle$ from Δ to φ (that is, $\Delta \vdash \varphi$) and assume that in some model M all the premises in Δ are true. The formulas occurring in our proof are either axioms, true in all models (we have just checked that), hence true in M, or premises in Δ and by assumption true in M, or derived from earlier formulas with the inference rules. Those rules, we have checked, preserve validity, so those formulas are also true in M. This means that φ has to be true in M, so indeed $\Delta \models \varphi$. I will not give the proof that this holds for $L0$ here.

The completeness part is the hard part. It says that our proof theory is rich enough to prove everything there intuitively is to prove, everything that is valid.

Standardly, one proves that if a theory (set of sentences) is consistent it has a model, by taking an arbitrary consistent theory Δ and constructing a model for it.

The method most commonly used nowadays was developed by Leon Henkin, and is called a Henkin proof. We have to prove that for any consistent theory Γ there is a model for Γ. In a Henkin proof we build this model for Γ by enriching the language that Γ is formulated in with enough new individual constants and build a model for Γ out of that enriched language.

I will sketch here some of the steps (for ease for a first order language L without identity).

First we define:

Let L be any first order language.

A *Henkin theory* is a set of sentences of L, Δ, such that:

1. Δ is maximally consistent.
2. if $\exists x\varphi \in \Delta$ then for some individual constant $c \in CON_L$: $\varphi[c/x] \in \Delta$

The point of this definition is that, by the special nature of a Henkin theory, we can build a model out of it in the following way:

Let Δ be a Henkin theory. We define a model M in the following way:

The domain D of M is the set of individual constants of the language in which Δ is formulated: $D = CON_L$.

The interpretation function i of M is given as follows:

$$i(c) = c$$
$$i(P^n) = \{\langle c_1, \ldots, c_n\rangle : P^n(c_1, \ldots, c_n) \in \Delta\}$$

THEOREM. *This model* $M = \langle D, i\rangle$ *is a model for Henkin theory* Δ, *i.e. for all* $\delta \in \Delta$: $M \models \delta$

In fact, what you can prove is:

For every sentence φ: $\varphi \in \Delta$ *iff* $M \models \varphi$

The proof goes with induction to sentences rather than with the more standard induction to formulas. Induction to sentences is usually impossible and is only possible here because of the special condition on existential sentences in a Henkin theory.

Proof. Base step: $P(c_1, \ldots, c_n) \in \Delta$ iff [by definition of M]

$\langle c_1, \ldots, c_n\rangle \in i(P)$ iff [by the semantics] $M \models P(c_1, \ldots, c_n)$

Induction step: (a) *Assume* $\varphi \in \Delta$ iff $M \models \varphi$. Note that, since φ is a sentence, this is equivalent to:

$\varphi \notin \Delta$ iff $M \not\models \varphi$.

We prove $\neg\varphi \in \Delta$ iff $M \models \neg\varphi$:
$\neg\varphi \in \Delta$ iff

[because Δ is maximally consistent]

$\varphi \notin \Delta$ iff

[by the induction hypothesis]

$$M \not\models \varphi \text{ iff } \quad [\text{semantics}] \quad M \models \neg\varphi$$

(b) Assume $\varphi \in \Delta$ iff $M \models \varphi$ and $\psi \in \Delta$ iff $M \models \psi$. We prove $(\varphi \wedge \psi) \in \Delta$ iff $M \models (\varphi \wedge \psi)$:

$$(\varphi \wedge \psi) \in \Delta \text{ iff}$$

[because Δ is maximally consistent]

$$\varphi \in \Delta \text{ and } \psi \in \Delta \text{ iff}$$

[induction hypothesis]

$$M \models \varphi \text{ and } M \models \psi \text{ iff}$$

[semantics]

$$M \models (\varphi \wedge \psi)$$

(c) Let $\psi[t]$ stand for a sentence ψ with term t occurring in it. Assume that for any constant $c \in CON_L$: $\varphi[c] \in \Delta$ iff $M \models \varphi[c]$. We prove

$$\exists x\varphi[x/c] \in \Delta \text{ iff } M \models \exists x\varphi[x/c]$$

If $\exists x\varphi[x/c] \in \Delta$ then, by Condition 2 of the definition of Henkin set, for some c: $\varphi[c] \in \Delta$.

If for some c: $\varphi[c] \in \Delta$ then, by the fact that Δ is maximally consistent, $\exists x\varphi[x/c] \in \Delta$.

So we know: (1) $\exists x\varphi[x/c] \in \Delta$ iff for some c: $\varphi[c] \in \Delta$.

The induction hypothesis says that $\varphi[c] \in \Delta$ iff $M \models \varphi[c]$ (for any c).

So we know: (2) for some c: $\varphi[c] \in \Delta$ iff for some c: $M \models \varphi[c]$.

Since the objects in the domain of M are the individual constants, every object in the domain trivially has a name. But that means that the following equivalence holds:

(3) for some c: $M \models \varphi[c]$ iff $M \models \exists x\varphi[x/c]$.

Combining (1), (2) and (3) we have proved that

$$\exists x\varphi[x/c] \in \Delta \text{ iff } M \models \exists x\varphi[x/c]$$

This completes the proof.

If we have identity in the language, the model corresponding to the Henkin theory Δ is a little bit more complicated: the individuals in the model will not be the constants but the equivalence classes of the constants under the relation:

$$c_1 \approx c_2 \text{ iff } c_1 = c_2 \in \Delta.$$

What we have proved now is that every Henkin theory has a model. What we have to prove is that any consistent theory has a model. We do that by proving the following:

Let L be a first order language.

LEMMA. *Any consistent set of sentences* Γ *in* L *can be extended to a Henkin theory* Δ (*in a language* L' *of which* L *is a sublanguage*).

If we can prove this, we have indeed proved that any consistent Γ has a model. The reason is this: we know that Henkin theory Δ has a model M' in L' (we have just proved that). If we take this model and ignore the interpretation of everything that is not in our original language L, we get a model M for L: M is the restriction of M' to L. Since M' is a model for Δ and $\Gamma \subseteq \Delta$, it follows that M' is also a model for Γ. But M is just the restriction of M' to L, so it does exactly the same with the L-sentences as M' does. Consequently M is a model for Γ.

The hard part then is to prove this lemma.

This can be done with the following construction.

We start with theory Γ in language L.

(a) Extend L to L' by adding countably many new individual constants $c_1, c_2, \ldots$ These are called *witnesses* for the truth of existential sentences $\exists x\varphi$.

The point is this: in our original model it is not true that every object has to have a name. So $\exists x\varphi$ can be true, without there being any constant c such that $\varphi[c/x]$ is true. Since in the end in Δ we are going to construct our model out of the individual constants (as we have seen before), i.e. the individual constants will be the objects in our model, and since the truth of $\exists x\varphi$ in *that* model requires there to be an object in D that satisfies φ, there has to be an individual constant c in *that* language such that $\varphi[c/x]$ is true (that constant then can be taken as the required object).

(b) We put all the sentences of L' in a list (this can be done because there are at most countably many) $\varphi_1, \varphi_2, \ldots$

(c) Now we define a sequence of theories $\Delta_0, \Delta_1, \ldots$ in L' in the following way:

1. $\Delta_0 = \Gamma$ (our starting consistent theory)
2. every Δ_i in the sequence is consistent

3. for every i: $\Delta_i \subseteq \Delta_{i+1}$
4. for every i: finitely many witnesses appear in the sentences in Δ_i
5. Look at the ith sentence φ_i in our list. If $\Delta_i \cup \{\varphi_i\}$ is consistent then $\varphi_i \in \Delta_{i+1}$
6. if $\varphi_i \in \Delta_{i+1}$ and φ_i is of the form $\exists x\psi$ then there is a witness c which does not occur in Δ_i or in φ_i such that $\psi[c/x] \in \Delta_{i+1}$

(d) Now define:

$$\Delta = \cup \Delta_i$$

i.e. Δ is the union of all the theories that occur in this sequence, hence it is the set of all the sentences that occur in the theories in this construction.

We can prove that:

(a) Δ is *consistent*. This follows because, if Δ is inconsistent then some finite subset of Δ is inconsistent. Since every formula in Δ comes from some Δ_i in the construction, some Δ_i will contain all the sentences in this finite inconsistent subset, so this Δ_i is inconsistent. But all the Δ_i are consistent by definition, hence Δ is consistent.

(b) Δ is *maximally consistent*. This means, beyond the fact that Δ is consistent that: for every L'-sentence φ: if $\Delta \cup \{\varphi\}$ is consistent, then $\varphi \in \Delta$. This follows from condition (5): if φ is consistent with Δ, then it is consistent with any subtheory of Δ in the sequence. Since φ occurs in our list of sentences, say at place j, it was added to Δ_j, and hence it is in Δ. (So what (a) and (b) prove is Lindenbaum's Lemma: every consistent theory can be extended to a maximally consistent theory.)

So we have proved that Δ is maximally consistent, this means that Δ satisfies the first condition for being a Henkin theory. Further, condition (2) of the definition of Henkin theory is satisfied by condition (6) of the construction. So we have proved:

Δ is a Henkin theory.

Since by definition of the construction $\Gamma \subseteq \Delta$, we have now proved our lemma: we started with an arbitrary consistent theory Γ and extended it to a Henkin theory Δ. Since the latter has a model, so does Γ. So indeed every consistent theory has a model. This ends the proof of the completeness theorem.

The Compactness Theorem

The notion of inconsistency is, by the nature of what a proof is, a finitary notion: a theory Δ is inconsistent if a contradiction can be proved from it. Since a proof is a finite sequence of formulas, only finitely many premises of Δ are needed to show the inconsistency. In other words: if Δ is inconsistent, then there is a finite subset of Δ that is already inconsistent. Or, to formulate it the other way round: if every finite subtheory of Δ is consistent, then Δ is consistent. For consistency, this follows from the definition. What happens at the semantic side? Suppose a theory does not have a model. Is it true that then already some finite subtheory does not have a model? The compactness theorem tells us that this is indeed the case. The compactness theorem is a corrollary of the completeness theorem:

COMPACTNESS: Let Δ be a theory.

If every finite subset of Δ has a model then Δ has a model.

Exercise 6. Show that the completeness theorem implies the compactness theorem.

The Löwenheim–Skolem Theorem

Let $M = \langle D, i \rangle$ be a model. The cardinality of M is the cardinality of its domain D. So M is finite if D is, countable if D is, etc. . .

If you inspect the proof of the completeness theorem and in particular the construction of the Henkin theory out of our consistent starting theory Γ, then you observe that we add to the (maximally countable) set of individual constants of our starting language L countably many new individual constants to get L'. This means that the set of constants of L' is also countable and since this set of constants forms the domain of the model we construct out of the Henkin theory we have in fact shown that every consistent theory has a model that is at most countable (some consistent theories have only finite models). This is the downwards Löwenheim–Skolem theorem:

DOWNWARDS LÖWENHEIM–SKOLEM THEOREM. *If a first order theory has infinite models it also has a* countable infinite *model.*

The general (downwards and upwards) Löwenheim–Skolem(–Tarski) theorem says:

If a first order theory has an infinite model it has models of every infinite cardinality.

The compactness theorem and the Löwenheim–Skolem theorem tell us a lot about the (lack of) expressive power of first order logic. To show this, let me give a consequence of the compactness theorem:

THEOREM. *If a theory Δ has arbitrary large finite models then Δ has an infinite model.*

Proof. Let Δ have arbitrary large finite models. Consider the following sentences:

φ_1: $\exists x\ x = x$ ('the domain has at least one object')
φ_2: $\exists x \exists y\ x \neq y$ ('the domain has at least two objects')
etc. . . .

Consider the theory $\Delta \cup \{\varphi_1, \varphi_2, \ldots\}$.

Take any finite subset of this theory. It consists of a finite set Δ_0 of sentences from Δ and a finite set of sentences of the form φ_i. In fact we only have to consider the largest φ_i in this finite subset, because it implies the ones with a smaller number (if the domain has at least 20 elements it has at least 10). So we can write this finite subset as $\Delta_0 \cup \{\varphi_i\}$. Since Δ has arbitrary large finite models it has a model where φ_i is true, hence $\Delta_0 \cup \{\varphi_i\}$ has a model. From this we can conclude that every finite subset of $\Delta \cup \{\varphi_1, \varphi_2, \ldots\}$ has a model. Then it follows with compactness that $\Delta \cup \{\varphi_1, \varphi_2, \ldots\}$ has a model. Because all the sentences $\varphi_1, \varphi_2, \ldots$ are true on this model it is infinite (for every n, its domain has at least n elements). But of course this model is a model for Δ as well, so Δ has an infinite model.

We can use this to give some examples of the lack of expressibility of first order logic.

We know that the property 'having exactly two elements' is first order definable.

$$\varphi = \exists x \exists y [x \neq y \wedge \forall z [z = x \vee z = y]]$$

MOD(φ) is the class of all models with exactly two elements. So this sentence defines that class. Similarly, for every n, the properties of having at least, at most, exactly n elements are first order definable.

But:

FACT 1. The property *finiteness* is not first order definable (there is no first order formula that is true in all and only all finite models).

FACT 2. *Finiteness* is not generalized first order definable (definable by a set of sentences).

FACT 3. *Infinity* is not first order definable.

FACT 4. *Infinity* is generalized first order definable.

Proofs.

(1) Finiteness is not generalized first order definable. This means that there is no set of sentences Δ such that MOD(Δ) is the class of all finite models. This follows directly from the previous theorem (and hence from compactness). If MOD(Δ) were the class of all and only finite models, then Δ would have arbitrary large finite models and hence it would have an infinite model M. So $M \in MOD(\Delta)$ which contradicts the assumption.

(2) If finiteness can't be defined by a set of sentences it certainly cannot be defined by a sentence, so finiteness is not first order definable.

Exercise 7. Prove fact 3.
Exercise 8. Prove fact 4.

We can formulate these results about definability in a somewhat different way. First order logic is good in telling models of one finite cardinality apart from models of another finite cardinality (there will be a first order sentence holding in the one but not in the other). First order logic is not capable of carving out the class of all finite models, although it can pick out (with an infinite set of sentences, though) the class of all infinite models. Beyond that, the Löwenheim-Skolem theorem tells us, first order logic is completely blind. First order logic is totally incapable of distinguishing models of one infinite cardinality from models of some other cardinality, because the theorem says: if MOD(Δ) contains an infinite model, it contains models of all infinite cardinalities. As we will see this has some interesting consequences.

Another interesting fact that is proved in very much the same way is:

THEOREM. Most *is not first order definable, that is, there is no first order sentence* $\varphi[A, B]$ *that expresses: most* As *are* B.

Proof. First note that most As are B means that there are more As that are B than As that are not B, i.e. the cardinality of $A \cap B$ is bigger than that of $A - B$.

Assume that there is a first order sentence $\varphi[A, B]$ expressing this. We consider again lists of first order sentences, two this time:

$\varphi_1 :=$ at least 1 A is B $\quad\quad$ $\psi_1 :=$ at least 1 A is not B
$\varphi_2 :=$ at least 2 As are B $\quad\quad$ $\psi_2 :=$ at least 2 As are not B
... $\quad\quad$ ...

Let $\Delta = \{\varphi(A, B), \varphi_1, \varphi_2, \ldots, \psi_1, \psi_2, \ldots\}$

Take any finite subset of Δ. It consists maybe of $\varphi[A, B]$ and of some sentences saying that at least n As are B and at least m As are not B.

Any such subset has a model where more As are B than As are not B. For instance, a model where $|A - B| = m$ and $|A \cap B| = k$, where $k \geqslant n$ and $k > m$ (where $|X|$ stands for the cardinality of X).

But now we know with compactness that our whole theory Δ has a model. Because Δ contains all these sentences $\varphi_1, \varphi_2, \ldots$ and $\psi_1, \psi_2, \ldots$, we know that Δ has only infinite models, even stronger, we know that Δ has only models where $A \cap B$ and $A - B$ are infinite, because all the φ-sentences and all the ψ-sentences are true on it.

Now we apply the Löwenheim–Skolem theorem and we find that Δ has a countable model, say M. M is a model for Δ, so $\varphi[A, B]$ is true on M.

On M, $A \cap B$ and $A - B$ are infinite, and since M is countable, this means that both $A \cap B$ and $A - B$ are countable. But then $|A \cap B| = |A - B|$ in M. But that means that *most As are B* is false on M, because exactly as many As are B as there are As that are not B. Since $\varphi[A, B]$ is true on M, we have proved that $\varphi[A, B]$ does not define *most As are B*, because there is a model where the first is true, but the second is false.

Now you might say: Well, that is because we stretch our intuitions about *most*, that are derived from finite sets, to infinite sets, and so we shouldn't be surprised if we get strange results there. Leaving aside the question whether this is justified criticism, it leads to a new question: Suppose we only consider finite models. Is *most* definable with respect to the class of finite models? This means the following. Let *FIN* be the class of all finite models. Is there a first order sentence φ such that for all $M \in FIN$: $M \models \varphi$ iff $M \models$ *most As are B*?

The answer to this question is a bit harder to prove. I won't give the proof, but indicate the reason for it:

THEOREM. Most *is not first order definable with respect to the class of finite models*.

The reason is the following. You can prove that first order formulas have a certain insensitivity to size that *most* lacks. Basically what you can show is that for every first order sentence, there is a certain number n, related to the complexity of that formula (roughly, the number of embedded quantifiers), such that the size of the relevant sets in a model can make a difference to the truth value of the sentence if their size is smaller than n, but not if it is bigger than n.

To give an example, take the sentence: *either there are at most two men or there are at least four men*. The crucial set here is the set of men and the crucial number here is four. If you vary the size of the set of men below five, then first the sentence is true (4), then it is false (3), then it is true again (2). But on models with a set of men bigger than 4 it is invariably true. This is the intended insensitivity: just changing the size of the models higher than n is not going to affect the truth value: it is invariably true there, or invariably false there.

So, what we can prove is that we can find such a number for every first order sentence.

The reason then that *most* is not definable on the finite models, is that for *most As are B* we cannot find a number n, such that it is insensitive to size higher than n.

The point is, *most* essentially compares the sizes of two sets $A \cap B$ and $A - B$. Whatever finite size you give those two sets, you can change the truth value of *most As are B* by changing the sizes of those two sets. I.e. say $|A \cap B| = n$ and $|A - B| = m$ and $n > m$, so the sentence is true. Change $|A - B|$ to k, $k > n$ and the sentence is false. Change $|A \cap B|$ to l, $l > k$ and the sentence is true again, and of course there is no number after which this switching of truth values becomes impossible. For that reason *most* is not definable even on finite models.

To end this section, let me mention in how far first order logic is characterized by the above theorems:

LINDSTRÖM'S THEOREM. *Every logic (with semantics) that extends* *first order logic and satisfies the compactness theorem and the Löwenheim–Skolem theorem coincides with first order logic. Similarly if it satisfies the completeness theorem and the Löwenheim–Skolem theorem.*

1.2. SECOND ORDER LOGIC

1.2.1. *Basic Concepts*

To get a better view on what it means to be first order and to be complete, I will briefly discuss some aspects of higher order logic, in particular second order logic.

In a second order language we can not only talk about and quantify over individuals (John is ill, Every man walks), but also about and over properties of individuals:

Being ill is terrible.
Every property John has, Mary has too.

We can represent these examples as:

Terrible (Ill)
$\forall X[X(j) \rightarrow X(m)]$.

We will base second order languages on the same logical constants as first order languages (except that we need no longer take = as a primitive) and on the following sets (I suppress index L):

$CON = \{c_1, c_2, \ldots\}$	individual constants
$PRED_1^n = \{P_1^n, P_2^n, \ldots\}$	n-place predicate constants of individuals
$PRED_2^n = \{\mathbf{P}_1^n, \mathbf{P}_2^n, \ldots\}$	predicate constants of 1-place place properties of individuals
$VAR_1 = \{x_1, x_2, \ldots\}$	individual variables
$VAR_2^n = \{X_1^n, X_2^n, \ldots\}$	n-place predicate variables

The second order languages that I define here are not as general as could be. We see that, as before, we have individual constants and variables, and n-place predicates of individuals. The latter will, as before, denote sets of n-tuples of individuals.

Besides that we have n-place predicate variables, which will range over sets of n-tuples of individuals. Let's call a set of n-tuples of individuals an (extensional) n-place property of individuals. We will be able, then, to quantify over n-place properties of individuals.

The restriction of the present languages lies in the second order predicates. For simplicity I will assume that there are only (n-place) predicates of 1-place properties. If we want to be more general, we have to associate with every predicate a type, indicating what kind of entity the predicate takes at what argument place (i.e. that a predicate is for instance a three place relation between individuals, one place properties of individuals and two place relations between individuals). I won't go into the trouble of doing this, because it can be done more elegantly in type logic, which we will see below.

I-TERM, the set of individual level terms:

$$I\text{-}TERM = CON \cup VAR_1$$

P-TERM, the set of predicate level terms:

$$P\text{-}TERM = VAR_2^1 \cup PRED_1^1$$

ATFORM is the smallest set such that:

1. if $P \in PRED_1^n$ and $t_1, \ldots, t_n \in I\text{-}TERM$ then $P(t_1, \ldots, t_n) \in ATFORM$
2. if $X \in VAR_2^n$ and $t_1, \ldots, t_n \in I\text{-}TERM$ then $X(t_1, \ldots, t_n) \in ATFORM$
3. if $\mathbf{P} \in PRED_2^n$ and $T_1, \ldots, T_n \in P\text{-}TERM$ then $\mathbf{P}(T_1, \ldots, T_n) \in ATFORM$

FORM is the smallest set such that:

1. $ATFORM \subseteq FORM$
2. If $\varphi \in FORM$ then $\neg\varphi \in FORM$
3. If $\varphi, \psi \in FORM$ then $(\varphi \wedge \psi) \in FORM$
4. If $x \in VAR_1$ and $\varphi \in FORM$ then $\exists x\varphi \in FORM$
5. If $X \in VAR_2$ and $\varphi \in FORM$ then $\exists X\varphi \in FORM$

We see that the only interesting new clauses are *ATFORM* 2 and 3 and *FORM* 5.

Let L be a second order language.

A *model for L* is a pair $\langle D, i\rangle$ where (as before) D is a non-empty set of individuals and i is given as follows:

if $c \in CON$ then $i(c) \in D$

if $P \in PRED_1^n$ then $i(P) \subseteq D^n$

We use pow(A) for the power set of A, the set of all As subsets. Then this clause can be written as:

If $P \in PRED_1^n$ then $i(P) \in \text{pow}(D^n)$

And we add:

If $\mathbf{P} \in PRED_2^n$ then $i(\mathbf{P}) \subseteq (\text{pow}(D))^n$

i.e. we assume that second order properties are n-place relations between one-place first order properties. The latter are sets of individuals, so second order properties denote n-tuples of sets of individuals.

An *assignment* function is a function g that assigns to every $x \in VAR_1$ an object $g(x) \in D$ and to every $X \in VAR_2^n$ a set $g(X) \subseteq D^n$.

Now we can give the truth definition. All clauses stay the same, except that the following are added:

$[\![\mathbf{P}]\!]_{M,g} = i(\mathbf{P}) \qquad [\![X]\!]_{M,g} = g(X)$

(a) $M \models X(t_1, \ldots, t_n)[g]$ iff $\langle [\![t_1]\!]_{M,g}, \ldots, [\![t_n]\!]_{M,g} \rangle \in g(X)$

(b) $M \models \mathbf{P}(T_1, \ldots, T_n)[g]$ iff $\langle [\![T_1]\!]_{M,g}, \ldots, [\![T_n]\!]_{M,g} \rangle \in i(\mathbf{P})$

(c) $M \models \exists X^n \varphi[g]$ if for some set $A \subseteq D^n$: $M \models \varphi[g_X^A]$

We see that these definitions are totally along the lines of those for first order logic.

We get second order logic, $L1$, by adding to our first order logic $L0$ the second order versions of A4,A5 and GEN:

A4′: $\forall X(\varphi \to \psi) \to (\varphi \to \forall X\psi)$ if X is not free in φ

A5′: $\forall X\varphi \to \varphi[T/X]$ if T is substitutable for X in φ

GEN: if $\vdash \varphi$ then $\vdash \forall X\varphi$

and the following comprehension axiom schema:

A8: Let $\varphi[x]$ be a formula with free variable x.
$\exists X \forall x[X(x) \leftrightarrow \varphi[x]]$

The latter postulate tells you that sentences with a free variable define one place properties.

We no longer need the identity axioms in our logic, because identity can be *defined* in second order logic as follows:

$t_1 = t_2 := \forall X[X(t_1) \leftrightarrow X(t_2)]$

This definition says that two individuals are identical iff they have the same properties. Two individuals are distinct iff there is a property that the one has but the other doesn't. This principle is called Leibniz' Law. Note that the definition works, because there are enough properties around to do the job (i.e. if d and d' are distinct then d has the property $\lambda x.x = d$, but d' doesn't).

After second order logic comes third order logic. In third order logic we can quantify over second order properties. Usually we jump from second order logic directly to higher order logic in general. A particular simple format for higher order logic is *functional type* logic.

Functional type logic differs in three respects from the logical systems we have seen up to now. In the first place we have as much higher order quantification as we want. Secondly, we don't interpret predicates as sets, but as their characteristic functions, all predicates are interpreted as functions. Thirdly, we add to the language a λ-operator with which we can define complex functions in the language.

Though I assume familiarity with type logic, I present its syntax and semantics here for completeness' sake.

A *type logical language* has logical constants $\neg$ $\wedge$ $\exists$) (, = λ and is based on a set of types (the same for every type logical language):

TYPE is the smallest set such that:

1. $e, t \in TYPE$
2. if $a, b \in TYPE$ then $\langle a, b \rangle \in TYPE$

e is the type of individuals, t the type of formulas. For every type a we specify:

CON_a, the set of constants of type a
VAR_a, the set of variables of type a (countably many)

We define for every a: EXP_a, the expressions of type a:

EXP_a is the smallest set such that:

1. $CON_a \cup VAR_a \subseteq EXP_a$
2. if $\alpha \in EXP_{\langle a,b \rangle}$ and $\beta \in EXP_a$ then $\alpha(\beta) \in EXP_b$
3. if $x \in VAR_a$ and $\beta \in EXP_b$ then $\lambda x \beta \in EXP_{\langle a,b \rangle}$
4. if $\varphi, \psi \in EXP_t$ and $x \in VAR_a$ and $\alpha, \beta \in EXP_b$ then $\neg\varphi$, $(\varphi \wedge \psi)$, $\exists x\varphi$, $\alpha = \beta \in EXP_t$

A *model* for type logical language L is a pair $M = \langle D, i \rangle$.
D is set of individuals.

For every type a, we define the domain of a, D_a by:

$$D_e = D \qquad \text{individuals}$$
$$D_t = \{0, 1\} \qquad \text{truth values}$$
$$D_{\langle a,b\rangle} = D_b^{Da} \qquad \text{the set of all functions from } D_a \text{ into } D_b$$

The interpretation function i assigns to every $c \in CON_a$ an interpretation $i(c) \in D_a$.

For every type a and $x \in VAR_a$, an assignment function g assigns to x a value $g(x) \in D_a$.

This time we define for every type a and $\alpha \in EXP_a$: $[\![\alpha]\!]_{M,g}$, the interpretation of α in M relative to g.

1. $[\![c]\!]_{M,g} = i(c)$ if $c \in CON_a$
   $[\![x]\!]_{M,g} = g(x)$ if $x \in VAR_a$
2. $[\![\alpha(\beta)]\!]_{M,g} = [\![\alpha]\!]_{M,g}([\![\beta]\!]_{M,g})$
3. $[\![\lambda x\beta]\!]_{M,g} = \lambda d \in D_a \,.\, [\![\beta]\!]_{M,gx^d}$

Note that I use here λ-abstraction in the meta language to denote the function that assigns to every d: $[\![\beta]\!]_{M,gx^d}$

4. $[\![\neg\varphi]\!]_{M,g} = 1$ iff $[\![\varphi]\!]_{M,g} = 0$ (0 otherwise)
   $[\![(\varphi \wedge \psi)]\!]_{M,g} = 1$ iff $[\![\varphi]\!]_{M,g} = [\![\psi]\!]_{M,g} = 1$
   $[\![\exists x\varphi]\!]_{M,g} = 1$ iff for some $d \in D_a$: $[\![\varphi]\!]_{M,gx^d} = 1$
   $[\![\alpha = \beta]\!]_{M,g} = 1$ iff $[\![\alpha]\!]_{M,g} = [\![\beta]\!]_{M,g}$

I won't give the axiom system for type logic (see Gallin, 1975), except for mentioning the central axiom:

$$\lambda x\alpha(\beta) = \alpha[\beta/x]$$

if no variable that is free in β becomes bound in $\alpha[\beta/x]$.

This is the principle of λ-conversion.

1.2.2. *The Expressive Power of Second Order Logic*

Second order logic has a much greater expressive power than first order logic. For instance, properties like finiteness and infinity are second order definable. Another example concerns sentences like *most As are Bs*. We have seen that *most* is not first order definable, but *most* is second order definable. The reason is the following. Most As are Bs means that the cardinality of the set $\lambda x[A(x) \wedge B(x)]$ is greater than the cardinality of $\lambda x[A(x) \wedge \neg B(x)]$. Since in second order logic we can quantify over properties and relations, we can express that there is a relation between two properties X and Y: $\exists R[R(X, Y)]$; we can express that this relation is a function with X as its domain and Y as its range; and we can express that this relation is an injection, a one–one function between X and a proper subset of Y. Given this, we can express that such a function exists between $\lambda x[A(x) \wedge \neg B(x)]$ and $\lambda x[A(x) \wedge B(x)]$. The formula expressing the latter expresses that most As are Bs.

We see that in second order logic we can express facts about the cardinality of sets. In fact, we can define the natural number series and their natural relations and properties (= Peano Arithmetics) in second order logic.

Since second order logic is an extension of first order logic and it is clearly more expressive than the latter, we know by Lindström's theorem that it will lose some of the characteristics of first order logic: completeness, compactness and Löwenheim–Skolem. In fact, it loses all of them.

It should be clear that the Löwenheim-Skolem theorem cannot hold, because we can now express whether a model is countable or has a higher cardinality. If a theory contains a sentence expressing the latter, it cannot have a countable model, contradicting Löwenheim–Skolem.

Neither generalized completeness nor weak completeness hold. The correctness theorem does hold, but that only tells us that the axioms we have chosen do not lead us into problems. But there are sentences that are logically valid (true in all models) that cannot be proved by the proof theory for second order logic.

Nor does it help to think that maybe we haven't enough proof rules in our axiom system. Gödel's incompleteness theorem (to which I will come back) applies to second order logic. This theorem tells us that even if you were to add to the proof system all the true sentences you can't prove in the proof system (which is not very intuitive as a system of proofrules, for sure), you will get true sentences in the new proof system you cannot prove.

Ironically, it is precisely the strength of second order logic, its expressibility, that turns into its weakness here. When theories or logics get as strong as this, it becomes possible to encode facts *about* their own proof theories in their sentences. In an indirect way, it becomes possible to express in the theory facts like: φ is an axiom of the theory, φ is *provable* in the theory.

Godel proved (for Peano Arithmetics, but it holds for second order logic as well) that it becomes possible to define a sentence which expresses: 'I am not provable in this proof theory'. This sentence is true, because if it were provable, the theory would be inconsistent, and because of what it says it is not provable. As said before, adding the sentence brings no relief. It gives you a richer theory in which you can prove the fact about the first theory, but there will be a similar sentence saying the same about the new theory, which is not provable in the new theory.

It is important to realize that when we say that a logic is incomplete, this means that it is incomplete *with respect to the* intended semantics. The axiom system of second order logic is intended as a theory to prove sentences that quantify over extensional properties (sets) of individuals. The intended semantics is reflected in the definition of a second order model and the truth definition. In particular, the intended semantics assumes that applying a predicate to an argument means that the denotation of the argument is *a set theoretic element of the* denotation of the predicate, as is shown in clauses like:

$$M \models P(t)[g] \text{ iff } [\![t]\!]_{M,g} \in [\![P]\!]_{M,g}$$

What the incompleteness theorem says is that the logic is not capable of capturing all valid inferences about these intended structures.

This is very important to keep in mind for the following reason: we can give a *first order* formulation of second order logic (as we will see, more correctly: a first order approximation).

Up to now we have only seen first order models where the variables range over the whole domain D. A *sorted* first order logic has different sorts of variables that range only over a subset of the domain (an example of semantic use of sorted quantification is given in Cooper 1983). One can prove that this sorting, if done properly, does not add anything to the unsorted first order logic.

Basically, if you have variable v of sort P and variable α of sort Q, then in a sorted logic $\forall v\varphi$ and $\forall \alpha\varphi$ need not have the same truth value, because the variables quantify over different domains. But you can add predicate constants P and Q to the language that denote the sorts, and then you can translate these expressions into unsorted ones:

$$\forall x[P(x) \rightarrow \varphi[x/v]] \text{ and } \forall x[Q(x) \rightarrow \varphi[x/\alpha]],$$

equivalent to the sorted ones.

Completeness, compactness and Löwenheim–Skolem do apply to this sorted logic.

Now suppose we split the domain of a model D into two sorts, and we have two corresponding sorts of variables. The first sort we call D_I and the corresponding variables are $x_1, x_2, \ldots$. We think of the elements of this sort as individuals.

The second sort we call D_p with variables $X_1, X_2, \ldots$ and we think of these as properties.

Now we add a special relation A (for application), a relation between elements of sort P and elements of sort I. So, for $A(c_p, c_I)$ we can read: property c_p applies to individual c_I.

In this way we can also mimic second order quantification with first order quantification. For instance, we define:

$$t_1 = t_2 := \forall X[A(X, t_1) \leftrightarrow A(X, t_2)]$$

We can rewrite the axiom system of second order logic in this new language. In fact, we don't have to do much at all, from which we can see that an axiom system is only second order in relation to a second order interpretation, its intended semantics, not all by itself. In other words, it is not something about the formulas itself that makes the logic second order, except in a trivial way.

What we do then is put conditions on the special relation A to make it as much as possible like the *intended* application ($\in$) of second order logic, and to make properties as much like 'real' properties. But it is clear that there is one thing that A cannot be, and that is $\in$ itself (then the models become second order again).

The completeness theorem for sorted first order logic can be carried over to this 'second order logic', but, of course, this logic is not, and cannot be complete with respect to its *intended* semantics (which is the second order one), but only with respect to that part that can be expressed with A.

The models for this theory will have a relation $[\![A]\!]_{M,g}$ that captures some, but not all aspects of the intuitive meaning of application, but there are parts of the meaning of application that cannot be captured. So we can mimic second order logic in first order logic, we can not reduce it to it.

It is an open (re-opened) question whether first order property theories capture enough of the meaning of application to be useful for natural language semantics (see for instance Bealer, 1982; Bealer and Mönnich, 1989; Chierchia 1984; Chierchia and Turner, 1988).

In such a theory we would have a first order sentence expressing that most *A*s are *B*s, but that sentence can only approximate its standard, set theoretic meaning. It is not clear that such an approximation conforms to our intuitions (i.e. our intuitions may very well be truly second order).

When we compare first order logic and second order logic, we see that the second is attractive for its expressive capacity (something that semanticists are especially interested in), but that this attraction is also its great weakness: you lose the possibility of connecting semantically valid inference patterns with inference rules that manipulate formulas (or representations, for that matter). Whether this is bad or unproblematic depends on your occupation and your views about it.

From a semantic point of view, the expressibility of higher order logic is highly useful: what you're interested in is precisely expressing semantic properties and relations, and making fine semantic distinctions. Whether first order approximations give you a workable or natural framework remains to be seen.

On the other hand, a semantic theory may have more aims than just descriptive adequacy in characterizing entailments. For instance, if you think that the business of semantics is not just describing entailments, but giving inference rules that the meanings of expressions arise out of (some sort of 'meaning is use' theory) then you'd probably better stay first order. The same applies if you think that all semantic phenomena should be completely dealt with at a level of logical form. If your theory is second order, the incompleteness theorem tells you that this is impossible. Finally, if your concerns are computational, i.e. concern computing inferences, then first order logic (or in fact some subsystem of that) is to be preferred.

In short, the advantage of higher order logic is that it is a very powerful tool, which is very useful where we need powerful tools. The disadvantage is that we know precious little about the logical properties of this powerful tool. There is no comparison between the two for logicians. The main reason is that the weakness in expressibility of first order logic really turns out to be a magnificent logical strength. The Löwenheim–Skolem theorem plays a crucial role in a host of very deep results and developments in logic and set theory.

For instance, suppose you have a first order theory T (like set theory) and you want to prove that a certain first order sentence (say, the axiom of choice) is independent of it ($T \cup \{\varphi\}$) and $T \cup \{\neg\varphi\}$ are both consistent). To show that, you have to construct a model both for $T \cup \{\varphi\}$ and for $T \cup \{\neg\varphi\}$. Given Löwenheim–Skolem, you know that if these theories have models, they have countable models. It is usually much easier to construct a countable model than an arbitrary model, because there are a lot of by now well known techniques that you can use to construct countable models. And indeed, this is the way such proofs usually proceed (I will come back to this later in the part about set theory).

Another advantage is the use of so called transfer arguments. First order logic is bad in telling infinite structures apart. Suppose you have structures that you cannot tell apart in a first order way. If you then can prove that a first order property holds of the one structure, you have automatically proved that it holds of the other as well (else that property would tell the structures apart after all).

These are strong reasons for logicians to be interested in first order theories: we know a lot about them and there are a lot of techniques available for them, that we lose in higher order logic.

1.3. FIRST ORDER THEORIES

1.3.1. *Some Examples of First Order Theories*

A first order theory is a set of first order sentences. I will give some examples of first order theories and the different ways in which they are interpreted.

One type of first order theories is the type that defines a mathematically interesting class of structures, that is, a theory such that MOD(Δ) is exactly this class.

The language that such first order theories are written in consists of (of course) the logical constants and as non-logical constants the individual and predicate constants that the sentences (axioms) of the theory make statements about. These constants we can call the *mathematical constants*: their meaning is constrained by the axioms of the theory (you can think of those axioms as meaning postulates; in fact, this is what meaning postulates are).

If we want to prove something about just this theory, we assume that there are no other non-logical, non-mathematical constants. That is, if we want to prove something about, say, the theory of groups, we are only interested in sentences about the special relations and individuals that are constrained by the axioms.

The first theories we will consider are theories about *ordering*. The language they are formulated in consists of a single two place relation constant that we will write as $\leqslant$. So

$$CON = \emptyset,$$
$$PRED = \{\leqslant\}.$$

Again, what this means is that things like:

$$\exists x \exists y [\neg (x \leqslant y) \vee \neg (y \leqslant x)]$$

are sentences of our language, but $c_1 \leqslant c_2$ is not. The reason is that we are interested in the *inherent* properties of $\leqslant$, and not in contingent properties of arbitrary elements. If our purpose is just to study orderings, then it is unwanted to have such sentences. If we want to look at orderings with a special *element* (say a minimal element), we add a new mathematical constant c denoting that element, with axioms specifying its properties. So we only allow *as primitives* constants and predicates that the theory constrains, not arbitrary ones (note the as 'primitives': we are interested in relations we can define).

Of course, if our purpose is to *use* such theories in natural language semantics, say, as an ordered set of moments of time, then we can be interested (and normally are) in a language that has other non-logical constants as well (the interpretations of the basic lexical items in Montague Semantics, for instance).

But for the present discussion we are interested in language $\{\leqslant\}$.

The ordering theories that I will mention here will be discussed more systematically later in Chapter Two.

The first order theory consisting of sentences (1)–(3) defines the class of all *partial orders*:

Theory of partial orders, PO

(1)	$\forall x [x \leqslant x]$	reflexivity
(2)	$\forall x \forall y \forall z [(x \leqslant y \wedge y \leqslant z) \rightarrow x \leqslant z]$	transitivity
(3)	$\forall x \forall y [(x \leqslant y \wedge y \leqslant x) \rightarrow x = y]$	antisymmetry

Let us define:

$$x < y := x \leqslant y \wedge x \neq y$$

$<$ is called a strict partial order. We can define the theory of partial orders also with a strict partial order. We will see later that this is essentially the same theory:

Theory of strict partial orders

(1′)	$\forall x \neg (x < x)$	irreflexivity
(2′)	$\forall x \forall y \forall z [(x < y \wedge y < z) \rightarrow x < z]$	transitivity
(3′)	$\forall x \forall y [x < y \rightarrow \neg (y < x)]$	asymmetry

$PO \cup \{4\}$ defines the class of all linear orders:

Theory of linear orders

(1) (2) (3)

(4) $\forall x \forall y [x \leqslant y \vee y \leqslant x]$ connectedness

The corresponding postulate for strict partial orders is:

(4′) $\forall x \forall y [x < y \vee y < x \vee x = y]$ connectedness

An example of a linear order is **N**, the set of all natural numbers, ordered by smaller than (or smaller than or is equal to). If we add to the theory of linear orders postulates (6) and (7), we get:

Theory of linear orders without endpoints

(1) (2) (3) (4)

(5)	$\forall x \exists y \; x < y$	no greatest element
(6)	$\forall x \exists y \; y < x$	no smallest element

An example of a linear order without endpoints is **Z**, the set of all integers.

If we add to the theory of linear orders postulates (7) and (8) we get:

Theory of dense linear orders

(7)	$\exists x \exists y \; x \neq y$	at least two elements
(8)	$\forall x \forall y [x < y \rightarrow \exists z [x < z \wedge z < y]]$	density

The last principle says that between every two elements there is a third element.

Putting all these together gives us:

Theory of dense linear orders without endpoints

(I will call this theory D here): (1) (2) (3) (4) (5) (6) (7) (8).

Examples of such structures are **Q**, the set of rational numbers, and **R**, the set of real numbers.

The theory of dense linear orders without endpoints is a very interesting theory.

Remember that a theory Δ is *complete* iff its deductive closure is maximally consistent, iff for every φ (in the language that Δ is formulated in): $\Delta \vdash \varphi$ or $\Delta \vdash \neg\varphi$ (not both).

THEOREM. *The theory of dense linear orders without endpoints is complete.*

This theorem tells you that theory D says everything there is to be said about dense linear orders without endpoints in the language: every statement you can formulate follows from the theory or is inconsistent with it.

None of the other theories has this property. For instance, take the theory of partial orders, a sentence like $\exists x \exists y\ x \neq y$ is independent of this theory: there are models where it holds and models where it doesn't hold. This is not the case for D: every sentence either holds in all models for D or in none.

The proof of this theorem is very instructive because it uses almost everything that we have talked about, and techniques varying from simple observation, elementary logical proof, some logical adoptation of mathematical notions, the completeness and the Löwenheim–Skolem theorem and a very interesting mathematical construction (due to Georg Cantor). I give the proof in an appendix.

We have now seen examples of first order theories that are meant to characterize a class of mathematical structures. This is not the only way first order theories can be understood.

Sometimes a theory is written down with one particular model in mind, the *intended model* or *standard model*, and the axioms are meant to allow us to derive facts about this model.

Examples of this are Arithmetics or Number Theory, where we have the natural numbers in mind, and Set Theory where we think about the set theoretic universe.

Peano Arithmetics is the natural numbers series with its standard ordering and operations.

When we write down axioms for arithmetics we are not, as before, interested in the class of all models for that theory, but we are interested in how well that theory describe things that are arithmetically true, true in the standard model of arithmetics.

The standard model of arithmetics is $\langle \mathbf{N}, S, 0, +, . \rangle$, where **N** is the set of natural numbers, 0 is zero, S is the direct successor function, $+$ is addition and . is multiplication.

Note that our aim here is much more ambitious than in the case of partial orders. There we were interested in what is true for all partial orders. Here we are interested in what is true in one particular model. That is, ideally we would want the axioms of the theory to be such that we can derive everything that is true for the natural numbers, everything that is true in the standard model. This means that, ideally we would want our theory to be complete. I will discuss Peano Arithmetics in the next subsection.

1.3.2. *Peano Arithmetics (PA)*

First we have to extend the notion of first order theory slightly. Besides individual constants and predicate constants we introduce a new set of constants, *function constants.*

$$FUN^n = \{f_1, f_2, \ldots\}$$

n-place function constant f^n will be interpreted by the interpretation function i of the model as a function from D^n into D, a function that maps n-tuples of elements of D onto elements of D.

The syntax is extended with the rule:

If $t_1, \ldots, t_n \in TERM$ and $f \in FUN^n$ then
$f(t_1, \ldots, t_n) \in TERM$

and this rule has the obvious interpretation:

$$[\![f(t_1, \ldots, t_n)]\!]_{M,g} = i(f)(\langle [\![t_1]\!]_{M,g}, \ldots, [\![t_n]\!]_{M,g} \rangle$$

All theorems go through for first order logic with function constants.

The language for PA has the following mathematical constants:

$CON = \{\mathbf{0}\}$, where $\mathbf{0}$ is a constant denoting the number 0.
$FUN^1 = \{S\}$, $S(x)$ is the direct successor of x.
$FUN^2 = \{+, .\}$.

We write **n** for $S\ldots S0$ (with n-times S). These expressions are called *numerals*, they denote *numbers*. From now on I shall no longer write the numerals in bold type.

Peano Arithmetics, PA

(1) $\forall x[\neg(Sx = 0)]$

0 is not the successor of anything.

(2) $\forall x \forall y[Sx = Sy \rightarrow x = y]$

Numbers that have the same direct successor are identical.

(3) $\forall x[x + 0 = x]$
(4) $\forall x \forall y[x + Sy = S(x + y)]$

This justifies us in writing: $Sy = y + 1$

Proof. $1 = S0$

$$y + 1 = y + S0 = \text{(by 4) } S(y + 0) = \text{(by 3) } Sy$$

Written in that form, (4) is obvious: $x + (y + 1) = (x + y) + 1$.

(5) $\forall x[x.0 = 0]$
(6) $\forall x \forall y[x.Sy = (x.y) + x$

Again, this is clear in the form: $x.(y + 1) = (x.y) + x$.

(7) The last axiom is again an axiom schema: the induction schema: Let φ be a formula with free variable x.

$$(\varphi[0/x] \wedge \forall x[\varphi \rightarrow \varphi[Sx/x]) \rightarrow \forall x \varphi$$

This means the following: suppose 0 has property P and you can prove that if n has P then $n + 1$ has P, then every natural number has P:

$$(P(0) \wedge \forall n[P(n) \rightarrow P(n + 1)) \rightarrow \forall n\ P(n)$$

Let us prove with induction that every number differs from its successor: $\forall x[\neg(x = Sx)]$.

In other words, we choose for φ: $\neg(x = Sx)$.

(1) Base step: $\neg(0 = S0)$. This follows from axiom 1.

(2) Suppose (induction hypothesis) that for some number n: $\neg(n = Sn)$ (that is, $\neg(x = Sx)$ holds of n). We have to show that $\neg(Sn = SSn)$ (that is, $\neg(x = Sx)$ holds of Sn).

Assume that $Sn = SSn$, then with Axiom 2 we know that $n = Sn$, contrary to the assumption, so indeed $\neg(Sn = SSn)$.

This means that $\forall n[\neg(n = Sn) \rightarrow \neg(Sn = SSn)]$.

Given (1) and (2) the induction axiom with $\neg(x = Sx)$ chosen for φ tells us that indeed: $\forall x[\neg(x = Sx)]$.

PA is a very powerful theory (innocent though it looks). We can derive the typical kinds of things in it that we learn at school (like $x + y = y + x$ or $2 + 2 = 4$) and much more, because we can define natural classes of functions like the primitive recursive functions (like $\lambda x.2x^2 + x + 1$) in PA.

Exercise 9. Prove in PA: $2 + 2 = 4$

In fact, Peano does a very good job in capturing our intuitions about the natural numbers; so good that before Gödel it was widely believed that it did capture all our intuitions about them (as it was widely believed before Skolem that it actually uniquely defines the natural numbers). But we know that strong theories can be weak.

Gödel showed how you can encode in PA, that is, in statements about numbers, facts about the proof theory of PA, so that you can prove a fact about PA in PA itself by deriving from PA the number theoretic statement encoding that fact. Gödel showed that the encoding power of PA is very large and he was able to encode as a number theoretic statement the sentence expressing: 'I am not provable in PA'. The encoding is chosen in such a way that number theoretic encoding of this sentence is true on the standard model. By the way the coding is done, it follows that the encoded sentence is indeed a fact about PA.

It follows that this number theoretic statement cannot be derived from PA, although it is true on the standard model. Because, if it could be derived, then it would be provable in PA after all, but it expresses the fact about PA that it isn't provable in PA. Neither can the negation of this number theoretic statement be derived from PA. So: PA is not complete, there are sentences φ (in the language that PA is formulated in) such that PA $\nvdash \varphi$ and PA $\nvdash \neg\varphi$.

We have seen examples of incomplete theories before (partial orders), so we could ask: so what? It is important to see the difference. PA, as said before, was intended to prove facts about the natural numbers, and not to describe a class of structures.

We already knew, by the way, that Peano does not uniquely characterize the standard model. The reason is that PA is a first order theory and hence, by the Löwenheim–Skolem theorem, it has, besides the standard model all sorts of non-standard models.

Even given the existence of such non-standard models we could hope that at least PA could prove anything that is true in the standard model. Gödel showed that we can't: his sentence is true in the standard model, but not provable in PA.

But, we could say, suppose we take the following theory: $\Delta = \{\varphi\colon \mathbf{N} \models \varphi\}$, the set of all sentences true on the standard model. That theory is complete and does trivially capture what we want (it is called complete PA). Sure, but the problem is that we have no idea what that theory looks like, it is not axiomatizable (not with a finite set of axioms (but neither is PA), nor with the help of a finite set of axiom schemata). So that theory can never serve as a proof theory in which you can actually prove things (like PA).

A further consequence of the incompleteness theorem is that we cannot prove in PA that PA is consistent! This may seem strange. Isn't **N** a model for PA? And since PA is a first order theory, isn't PA consistent then? Yes. But we can't prove that in PA (while we can prove other things about the proof theory), rather we prove that in *set theory*, a richer theory (so we prove that PA is consistent relative to set theory). And that doesn't tell us much, because we cannot prove in set theory that set theory is consistent, so we have only shifted the problem. It is more of a problem there, however, because set theory is regarded as a foundation for mathematics, the stuff that mathematics takes place in. So we can't prove that that foundation is consistent. (Again, we can prove that it is consistent in some richer theory, but then, of course, we can't prove that that theory is consistent.)

Note that in second order logic, with its strong definability capacity, we can define the natural numbers and with them arithmetics up to isomorphism. So all the Löwenheim–Skolem non-standard models can be eliminated there. But, of course, that is not a good help either. First order Peano consists of a semantically complete set of inference rules and an incomplete mathematical theory that, hence, describes the natural numbers badly. In second order logic we can carve out the natural numbers completely, but now our inference rules are incomplete, so the result is as incomplete as before.

The different notions of completeness that we have discussed may be a bit confusing. Let me make a few remarks on that. Note first that there is a difference between a logic being complete and a theory being

complete. Take first order logic L. We have said that L is strongly complete and this means:

$$\Delta \vdash_L \varphi \text{ iff } \Delta \models \varphi$$

(weakly means: $\vdash_L \varphi$ iff $\models \varphi$).

On the other hand, we have said that theory Δ is complete iff

$$\Delta \vdash \varphi \text{ or } \Delta \vdash \neg\varphi$$

What is the relation between these two notions?

L is complete as a logic if it can prove the semantically valid inferences. This does not mean that L as a theory is complete. That would mean that for all sentences φ of first order logic

$$\text{for every } \varphi\colon \vdash_L \varphi \text{ or } \vdash_L \neg\varphi.$$

That only holds for the sentences that are either necessarily true or necessarily false, the non-contingencies.

Let $N(L)$ be the set of all non-contingencies of logic L. What completeness of a logic says is that the fact that logic L is complete means that:

$$\text{for every } \varphi \in N(L) \vdash_L \varphi \text{ or } \vdash_L \neg\varphi$$

Let Δ be a theory. Let $N(\Delta)$, the set of all non-contingencies relative to Δ, be the set of all sentences true in all models of Δ or false in all models for Δ. Now take a consistent theory Δ. Given that our logic L is complete it follows that Δ is complete as a logic even if it is not complete as a theory:

$$\text{for every } \varphi \in N(\Delta)\ \Delta \vdash_L \varphi \text{ or } \Delta \vdash_L \neg\varphi$$

The reason is that: $\varphi \in N(\Delta)$ iff $\Delta \models \varphi$ or $\Delta \models \neg\varphi$. The completeness of L gives us then the above statement.

So, in this sense even first order Peano Arithmetics is complete as a logic, in the sense that you can prove everything that holds on every first order model for Peano.

As we have seen earlier, neither second order logic nor second order Peano Arithmetics is complete as a logic.

Let us introduce a vague notion: $I(\Delta)$ is the class of intended models for Δ (where Δ can be a theory or a logic). Logic L is complete with respect to $I(L)$ iff:

$$\Delta \vdash_L \varphi \text{ iff for every } M \in I(L)\text{: } M \models \varphi$$

We can define the non-contingencies with respect to the intended models, $NI(L)$ to be the set of all sentences true on all intended models, or false on all intended models.

I said above that we can give a first order translation of second order logic that is complete with respect to a certain type of first order models. However, its intended models are second order models, and it is not complete with respect its intended models.

If $I(\Delta)$ consists of only one model, as in the case of Peano, and Δ is complete with respect to $I(\Delta)$ then Δ is complete *as a theory*, i.e. $\Delta \vdash \varphi$ or $\Delta \vdash \neg\varphi$. The reason is that if M is a model then for every sentence φ $M \models \varphi$ or $M \models \neg\varphi$. Δ proves every sentence true on the intended model and disproves every sentence false there, so every sentence is either proved or disproved. In this sense, the fact that Peano is incomplete as a theory shows that it is incomplete with respect to its intended model. This also means that there are sentences true on the standard model that are not proved by Peano.

By now one should have an idea why the incompleteness of PA was such an earth shaking result that destroyed the most influential doctrines about what mathematics is.

Nobody believes that there is 'the partial order' that we want to prove things about. There are partial orders. But mathematicians do believe that there is the set of natural numbers. If mathematics is the art of proving things about mathematical entities, then we now know that there are essential limitations on this, limitations at the very heart of mathematics, the natural numbers.

Let me briefly make a few remarks about the impact of Gödel's Theorem on the major doctrines about the nature of mathematics. A warning: when I say that a theory was refuted, I mean to say that it was refuted in the form in which it was popular. No foundational theory is ever refuted for good, they tend to reappear in more subtle forms.

Frege believed that all of mathematics could be reduced to (second order) logic (so: no mathematical axioms, only logical axioms). This is called Logicism. This paradigm was no longer current at the time of Gödel's paper, because Russell had shown that Frege's reconstruction crucially led to inconsistency (the Russell Paradox). However, the Incompleteness Theorem killed its dead body once more: you can't get a logic in which you can derive all of mathematics.

Probably the most prominent doctrine before Gödel was Hilbert's Formalism: the task of mathematics is to give mathematical axiom systems and prove all mathematical statements with them. The underlying idea is that the meaning of mathematical statements is completely determined by the proof theoretic inference rules that manipulate mathematical expressions. This doctrine, too, was refuted by Gödel. No mathematical theory is powerful enough to prove every true mathematical statement.

The foundational theory that, since Gödel, has gained most popularity among mathematicians is Realism. Realism assumes that there is a mathematical world (independent of us) out there, containing mathematical entities. Mathematics is discovering the properties of those mathematical entities. This doctrine is not refuted by Godel's theorem (on the contrary): if there is a realm of mathematical entities independent of us, containing, for instance, the natural numbers, there is no principled reason to expect that our finitary means of proving statements about them can tell us all there is to know about them.

Finally, Intuitionism, like Realism, assumes that there is a domain of mathematical entities, but that it isn't independent of us: mathematical entities are mental constructions that have their justification in the mathematical (structural) intuition that we human beings are equipped with.

One could think that at first sight this is incompatible with Gödel's result (What is a mental construction? Isn't it a proof?), but this is not correct. Like all mathematical entities, we should be able to construct the natural numbers in order to accept their existence, but this is not the same as proving their existence in a formal proof system. A better way to characterize the difference between a realist and an intuitionist is by saying that for a realist the models of a theory are things existing in the world, while for an intuitionist they are things constructed by us. Given this, there is also for an intuitionist no principled reason to expect that everything that is true in a constructed structure is provable in the proof theory (on the contrary).

What we can say is that both realism and intuitionism are in a way strongly semantically inspired theories, and that may very well be the reason that they have survived Gödel's theorem.

Whether it is independently existing structures or structures constructed on the basis of our intuition that form the basis for mathematical truth, in both cases these structures are non-syntactic entities, that serve as the interpretation for syntactic structures and as justification for syntactic procedures (like proof theories). If we want to come to grips with the meaning of mathematical statements it is at the level of those structures that we have to look. Gödel tells us that we won't find it anywhere else.

Beyond that, the roads part: some will try to justify mathematical truth in terms of arguments on what the mathematical world happens to be like, others will try to justify it in terms of the methods with which we construct the mathematical world. In this sense, the problems in the foundations of mathematics and of (model theoretic) semantics for natural language are strongly related and we find the same distribution of opinions about them in both fields.

1.4. ZERMELO–FRAENKEL SET THEORY

1.4.1. *Basic Set Theory*

The last first order theory discussed here is Zermelo–Fraenkel set theory or ZF. I will give the axioms and indicate to a certain extent what they do for you.

ZF differs from other set theories in that it is a theory of pure sets and sets only. That is, all entities postulated by ZF are sets. If you think that there are, say, individuals that are not sets, you can add a set of so called ur-elements to ZF (with appropriate axioms). Technically, the only result of that is that the theory becomes more cumbersome. In our semantic modeling we usually start with a domain of individuals D. This means that the elements of D are primitives *in our model*. That ZF, the theory in which we define our models, requires them to be sets is not incompatible with that: as long as *our models* regard them as primitives, we wouldn't care even if ZF were to tell us that in fact they are fried bananas (I owe this apt expression to Gennaro Chierchia).

Take the natural numbers as an example. The standard model for Peano consists of the set of natural numbers. Now, if the theory in which we describe this model is ZF, this requires natural numbers to be sets. We can define the natural numbers as sets in ZF, and in this way reduce Peano to ZF, but this does not mean that we show in that way that the natural numbers *are* sets, only that we can model them with sets. (To stretch our metaphor, if we can define the natural

numbers in set theory FB, where all things are fried bananas, we haven't shown that numbers *are* fried bananas, for instance, if FB is a different theory from ZF, then fried bananas are not sets, and natural numbers would be both.)

So it is important to distinguish between what individuals are according to the theory that they are individuals of and according to the theory in which that theory is modeled. ZF is a very powerful theory, because we can model practically everything in it we want to; in that sense it is a great help for all those for whom modeling is what they are interested in (like linguists). To be sure, there are things where one can seriously ask whether we are not losing too much by modeling them in ZF (these things mostly have to do with the extensionality of ZF and with its strict laws against self application) and where alternatives to ZF are proposed, but I think it is better to first understand how splendid and strong the modeling power of ZF really is, before one allows oneself to be dissatisfied with it and looks at alternatives. One fact is, for instance, that usually real alternatives to set theory can nevertheless be modeled in set theory. This certainly should make one really cautious.

Let's get to the theory. The language of ZF has one mathematical constant, the two-place relational constant $\in$.

The intuition behind sets is that they are objects, taken together, grabbed together, collected together, often under a certain perspective (the set of all chairs grabs objects together under the perspective of being a chair). Our first attempt at set theory is the following principle of *comprehension*:

Comprehension. Let $\varphi(x)$ be a first order sentence with free variable x, then the set of all objects that have φ exists:

$$\exists B \forall x [x \in B \leftrightarrow \varphi(x)]$$

We write this set as $\{x: \varphi(x)\}$.

This principle was assumed in the early age of set theory, when Cantor invented it, but it was soon shown to lead to contradictions, the most famous one being the Russell Paradox. Assume comprehension. Define the set of all sets that are not an element of themselves: $\{x: \neg(x \in x)\}$. Let us call this set A. $\neg(x \in x)$ is a first order formula, so comprehension says that this set exists. But then we can ask: is *this* set an element of itself or not? i.e.

$A \in A$?

Suppose $A \in A$. Then A has the property $\lambda x.\neg(x \in x)$, so $\neg(A \in A)$. Suppose $\neg(A \in A)$. But A is by definition the set of all sets that have the property $\lambda x.\neg(x \in x)$. Because $\neg(A \in A)$, A has that property, so $A \in A$. So indeed, comprehension leads to a contradiction.

Zermelo's diagnosis of what goes wrong in comprehension was that it quantifies over all sets that have a certain property, without guaranteeing that there is a universe that all those sets are taken out of. This universe would be $\{x: x = x\}$, the set of everything. The assumption that such a set exists leads to contradiction (as we have seen). The basic rule of set theory that avoids the paradoxes is that you cannot arbitrarily grab things together and form a set out of them, but only if you can prove that the things that you want to grab together formed part of a universe (set) to start with.

Instead of the comprehension axiom ZF has an axiom of *separation*. Suppose that you know that there is a set A, then you can use properties to separate the elements of A that have a property from those that don't. In other words, *given set* A, and first order formula $\varphi(x)$, then the separation axiom says that the set of all elements of A that have $\lambda x.\varphi(x)$ exists: i.e. if A exists, then $\{x \in A: \varphi(x)\}$ exists: Separation is again an axiom schema:

1. *Separation*: Given set A and formula $\varphi(x)$ (where A does not occur in $\varphi(x)$):

$$\exists B \forall x[x \in B \leftrightarrow x \in A \wedge \varphi(x)]$$

Note that with the separation axiom (unlike comprehension) you cannot prove that any set exists. It builds new sets out of other sets. Once there are some sets, it becomes very powerful, but it doesn't do anything, as long as there are no other axioms postulating the existence of some set.

Note that separation does not allow for the formation of the Russell set $\{x: \neg(x \in x)\}$, but at most for a set $\{x \in A: \neg(x \in x)\}$. It doesn't lead to the Russell Paradox (and neither will the paradox be brought in by the other axioms).

We said that the intuition behind set theory is: grabbing objects together under a certain aspect. We have seen that we cannot accept the first formulation of this leading to $\{x: \varphi(x)\}$, but only a restricted form: grabbing objects from a certain universe together under some aspect: $\{x \in A: \varphi(x)\}$.

A second way in which we modify the original intuition is by assuming the following: to build $\{x \in A: \varphi(x)\}$ we rely on this aspect $\varphi(x)$, but once the set exists, it is *only determined by the elements it has* and not by the aspect under which we brought those elements together. This is the axiom of *extensionality* saying that two sets are identical iff they have the same elements:

2. *Extensionality:* Given sets A and B:

$$\forall x[x \in A \leftrightarrow x \in B] \rightarrow A = B$$

Sets that have the same elements are identical. Note again that extensionality does not postulate the existence of any sets. It does not build new sets either, it only puts a strong condition on the identity of sets. Extensionality forms the basis of our use of the notation $\{a, b, c\}$ or $\{x: \varphi(x)\}$. Separation only says that there is *some* set B such that the elements of B are exactly the elements of A that have $\varphi(x)$. But it doesn't say that there is only one such set, which is needed, to motivate the notation $\{x \in A: \varphi(x)\}$ for *the* set of all elements of A that have $\varphi(x)$. Extensionality tells you that if there are two such sets, they have the same elements, and hence those two sets are identical.

We can define the relation of subset on sets (if they exist):

$$A \subseteq B := \forall x[x \in A \rightarrow x \in B]$$

Exercise 10. Prove that $\subseteq$ is a partial order, i.e. that $\subseteq$ is reflexive, transitive and antisymmetric.

Now that we have extensionality, we can show the defining power of separation. Assume that A and B are sets. Let us define:

The *intersection* of A and B,

$$A \cap B := \{x \in A: x \in B\}$$

Normally we write this as $\{x: x \in A \wedge x \in B\}$, but separation only allows a definition like the first. ($\{x \in B: x \in A\}$ is equally good.)

Those readers that feel they could use some exercise in basic set theory could prove that all these are equivalent and that $A \cap B \subseteq A$, $A \cap B \subseteq B$. And that $A \cap A = A$; $A \cap B = B \cap A$; $A \cap (B \cap C) = (A \cap B) \cap C$. And that $A \subseteq B$ iff $A \cap B = A$.

The *complement* of B in A,

$$A - B := \{x \in A : x \notin B\}$$

(we write $x \notin B$ for $\neg(x \in B)$).

So, with only extensionality and separation we can already define intersection (conjunction) and complementation (negation). We cannot define union (disjunction), though. We need some other things for that.

Up to now, sets don't exist (at least, nothing requires them to exist). We will now give an axiom that tells us that at least one set exists, *the empty set*:

3. *Empty set:* $\exists x[\forall y: \neg(y \in x)]$

There is a set without elements, we call it *the empty set* $\emptyset$. Those that want some practice can prove that there is only one such set and can prove some of its properties:

For every set A: $\emptyset \subseteq A$. But note $\emptyset \notin \emptyset$.

$$(A - B) \cap B = \emptyset$$
$$A \cap \emptyset = \emptyset$$

The empty set axiom is only a temporary one, because it becomes superfluous as soon as there is any other axiom postulating the existence of some set:

Exercise 11. Prove that if we add to extensionality and comprehension any axiom postulating that some set A exists, we can prove that $\emptyset$ exists as well.

With axiom 3 we have now at least one set postulated (the empty one), and separation applies to it, but that doesn't give us any other sets (it is hard to carve out elements of a set that doesn't have any elements). It is time now to change that. The next axiom, the pair axiom, is an axiom that builds a new set out of two given sets (like the empty set and the empty set). This axiom cannot be derived from separation, because separation can only select within one given set, but Pair can glue sets together to give a new set (i.e. separation only gives you sets within a universe, but pair can make new universes):

4. *Pair:* Given two sets A and B.

$$\exists C \forall x[x \in C \leftrightarrow x = A \vee x = B]$$

This means that given sets A and B, there is a set that contains exactly A and B as elements, i.e. given A and B, $\{A, B\}$ exists. Note that indeed this axiom builds new sets and can't be derived from separation. For instance, our universe up to now consisted of the empty set $\emptyset$ only. With pair, we know that the set $\{\emptyset, \emptyset\}$ exists, because of extensionality, this is $\{\emptyset\}$, the *singleton set* (one-element set) containing $\emptyset$ as its only element. Note that $\{\emptyset\} \neq \emptyset$ (why?).

And now that we have started, in fact we already get infinitely many sets: $\emptyset$, $\{\emptyset\}$, $\{\{\emptyset\}\}$, $\{\{\{\emptyset\}\}\}, \ldots$ And not only that, but now that we know that both $\emptyset$ and $\{\emptyset\}$ exist, the pair axiom tells us that $\{\emptyset, \{\emptyset\}\}$ exists as well.

Note that the pair axiom, though it is a building axiom, does not have existential import itself. Without the empty set axiom, Pair doesn't give you anything. We can't get sets with more than two elements with Pair. That is what the next axiom, the sum axiom, is going to add, because it is going to allow us to form unions.

First, one more notion: $\{x, y\}$ is called the unordered pair of x and y. Intuitively the *ordered pair* of x and y, $\langle x, y\rangle$ is like $\{x, y\}$, except that it takes x and y in the given order. This means that $\langle x, y\rangle$ satisfies the following conditions:

$$\langle a, a'\rangle = \langle b, b'\rangle \text{ iff } a = b \text{ and } a' = b'$$

It is common not to introduce $\langle x, y\rangle$ separately, but to find sets that satisfy this condition and then define ordered pairs as those sets (it doesn't mean that ordered pairs as we intuitively think of them *are* those sets, but that as far as set theory is concerned, everything works perfectly with those sets as ordered pairs). One definition is: $\langle x, y\rangle = \{\{x\}, \{x, y\}\}$. (Those who want practice can prove that the above condition holds under this definition.)

Given this, the pair axiom tells us that ordered pairs exist. (But not yet that the cartesian product of A and B, the set of ordered pairs $\langle a, b\rangle$ such that $a \in A$, $b \in B$ exists. This means that we cannot yet do relations and functions.)

The sum axiom builds a new set out of a set A that has elements by gluing the elements of the elements of A together (even if those elements are not part of the same universe). Together, the two axioms Pair and Sum will justify the operation of union.

5. *Sum:* Let A be a set. There is a set B that has as elements exactly those things that are element of at least one element of A:

$$\exists B \forall x[x \in B \leftrightarrow \exists y[x \in y \wedge y \in A]]$$

This set B is called the *sum of* A and is written as: $\cup A$. Let us define:

The *union of* A and B, $A \cup B := \{x: x \in A \vee x \in B\}$

This notation, of course, has to be justified, because we can't just quantify over any x. We can define, with pair and sum, the union of A and B formally as:

$$A \cup B := \cup\{A, B\}$$

You can, of course, prove that this definition gives you what you want, but it is probably easier to see it with an example. Let $A = \{1, 2\}$ and $B = \{3, 4\}$. Intuitively, the union of A and B should be $\{1, 2, 3, 4\}$. First form the pair $\{A, B\}$, that is $\{\{1, 2\}, \{3, 4\}\}$. The sum of this set (and here we use the sum axiom) is the set of all things that occur in at least one element of this set. The elements of this set are A and B, so the sum is the set of all things that are element of either A or B: $\{1, 2, 3, 4\}$. Note that now indeed we can form sets with more elements than two.

I said earlier that we are not yet able to get relations and functions, the next axiom, the power set axiom is going to provide them for us.

6. *Power sets:* If A is a set then there is a set that consists of all subsets of A:

$$\exists B \forall x[x \in B \leftrightarrow x \subseteq A]$$

We write the power set of A as pow(A). So pow(A) $= \{x: x \subseteq A\}$. Again, those that want some exercise can show that:

for every A: $\emptyset \in \text{pow}(A)$
for every A: $A \in \text{pow}(A)$

and can calculate: pow($\emptyset$), pow(pow($\emptyset$)), pow(pow(pow($\emptyset$))).

We now define:

The *Cartesian product of* A and B,

$$A \times B = \{\langle a, b\rangle: a \in A \wedge b \in B\}$$

We can prove that $A \times B$ exists if A and B do. For this we note that on the definition of $\langle a,b\rangle$ as $\{\{a\}, \{a, b\}\}$ it holds that if $a \in A$ and $b \in B$ then $\langle a, b\rangle \in \text{pow}(\text{pow}(A \cup B))$, and this set exists on the basis of Sum and Power. In fact, it is not hard to see that:

$$A \times B = \{x \in \text{pow}(\text{pow}(A \cup B)): \\ \exists a \in A \wedge \exists b \in B: x = \langle a, b\rangle\}$$

and given the existence of pow(pow($A \cup B$)), the Cartesian product of A and B is defined with separation.

Now we can define relations and functions. I will introduce here only the notions that are crucial in this section. In Chapter Two, I will develop the theory of relations and functions more extensively.

A *relation* between A and B is a subset of $A \times B$. If R is a relation between A and B then:

the *domain* of R, $\text{dom}(R) = \{a \in A: \exists b \in B[\langle a, b\rangle \in R]\}$
the *range* of R, $\text{ran}(R) = \{b \in B: \exists a \in A[\langle a, b\rangle \in R]\}$

Relation f between A and B is a *function* from A into B, $f: A \rightarrow B$, iff: for every $a \in \text{dom}(f)$ there is exactly one $b \in \text{ran}(f)$ such that $\langle a, b\rangle \in f$.

Function $f: A \rightarrow B$ is an *injection*, or *one–one*, iff

$$\forall a \in A\ \forall a' \in A: \text{if } f(a) = f(a') \text{ then } a = a'$$

(No two elements of the domain have the same value.)

Function f: $A \rightarrow B$ is a *surjection*, or *onto*, iff

$$\forall b \in B\ \exists a \in A: b = f(a)$$

(Every element in B is the value of some element in A.)

f is a *bijection* iff f is a surjection and f is an injection.

The *function space* of $A \times B$, B^A, is the set of all functions from A into B.

We see that by now our axioms are getting very powerful.

The next axiom (which is Fraenkel's contribution to ZF) is related to the separation axiom (with Empty set it implies separation). It says that if φ is a functional expression, that means: $\forall x \exists! y\ \varphi$ (for every x there is exactly one y such that φ) and set A exists, then the set of values of this function for domain A exists:

7. *Substitution schema:* Given set A and formula φ (such that B is not free in φ):

$$\forall x \exists! y\ \varphi \rightarrow \exists B \forall y[y \in B \leftrightarrow \exists x[x \in A \wedge \varphi]]$$

We now have a rich, infinite universe of sets. We don't yet have infinite sets though.

For this, we give von Neumann's set theoretic definition of the natural numbers:

1. $0 := \emptyset$
2. $\forall x\ S(x) := x \cup \{x\}$

On this definition:

$$0 = \emptyset$$
$$1 = \emptyset \cup \{\emptyset\} = \{\emptyset\}$$
$$2 = \{\emptyset\} \cup \{\{\emptyset\}\} = \{\emptyset, \{\emptyset\}\}$$
$$3 = \{\emptyset, \{\emptyset\}\} \cup \{\{\emptyset, \{\emptyset\}\}\} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$$

Or: $n = \{0, 1, \ldots, n-1\}$. (Note the correspondence between the number of elements of n and n.)

A is a *successor set* iff $0 \in A \wedge \forall x[x \in A \to S(x) \in A]$

8. *Infinity:* there is a successor set

We will show later that the infinity axiom does not follow from the other axioms.

Note that Infinity is the first existential axiom after Empty set. This time extensionality does not tell us that there is a unique successor set (If A is a sucessor set, then so is $S(A)$). The smallest successor set is that successor set that is a subset of every other successor set. This successor set is called ω, and it is the set of natural numbers.

We can easily define + and . on the natural numbers defined above and we can prove that all the Peano axioms, including the induction axiom hold for ω. So, indeed ω is a model for Peano in ZF.

Once we have the infinity axiom, the number of sets and their size really explodes. Note, furthermore, that once we have the infinity axiom, we no longer need the empty set axiom.

1.4.2. *The Set Theoretic Universe*

The last axiom of ZF is more like the extensionality axiom than like the others. It does not build sets, nor postulates the existence of sets, but it puts a condition on what sets we allow. The idea is the following. We would like to be able to picture the set theoretic universe in terms of the complexity of the sets, with the simplest set ($\emptyset$) at the bottom, and the other sets ranked according to their complexity (both the set $\{\omega\}$ and the set $\{\emptyset\}$ are singleton sets, still the first is more complex than the other). Intuitively, the way to define the complexity of a set is in terms of its element structure (that is, the structure consisting of the set, its elements, their elements, the elements of those, etc.) So we want to define the complexity of a set in terms of that structure. Given the huge size of sets and hence of those element structures, a definition of the complexity of a set is only possible if it is inductively defined on the element structure. An inductive definition like that is only possible if it is guaranteed that if you go down in the element structure (from a set to one of its elements, to one of the elements of that set and so on) that after having done that a finite number of times, you can't go down any more. Now the only sets for which the latter is not possible are sets that we cannot build on the basis of our axioms anyway, but cannot (yet) exclude either. Those are typically sets like sets that contain themselves as an element: $x \in x$ (in other words, our axioms have not yet excluded the possibility of the existence of a set that contains itself as an element). (If $x \in x$ you can go infinitely down in the element structure, because you can get in a loop: from x you go to x and then down to $x, \ldots$)

The *foundation axiom* tells us that such sets do not exist and with that it guarantees that we can define the complexity of every set.

9. *Foundation:* Every non-empty set A has an element b that has no element in common with A:

$$A \neq \emptyset \to \exists b \in A[b \cap A = \emptyset]$$

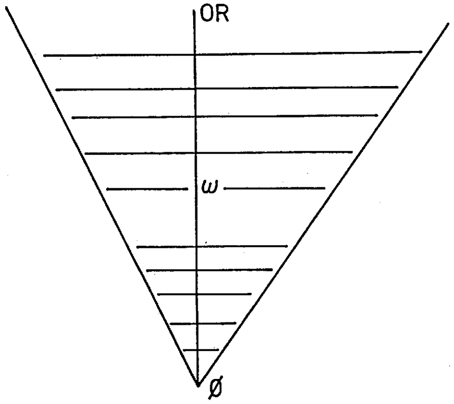
You can see that our funny set x such that $x \in x$ is excluded now. Suppose $a \in a$. Consider $\{a\}$. This is a non-empty set, according to foundation it has an element that has no element in common with a. But its only element is a, and according to the assumption $a \in a$, so a and $\{a\}$ do have an element in common, contradicting foundation.

Let's say a bit more about the set theoretic universe. The elements of ω are also called the finite *ordinal numbers*. ω is the first infinite ordinal number. The class of all ordinal numbers (note, class and not set, we use *class* when something is too big to be a set) OR is given by the following principles (it can be defined in other ways):

1. $0 \in OR$
2. if $x \in OR$ then $S(x) \in OR$

3. if $x \subseteq OR$ then $\cup x \in OR$

You can prove that the ordinal numbers are linearly ordered. Since $\omega = \cup\omega$, indeed ω is the smallest infinite ordinal number. Note that $S(\omega)$, $\omega + 1$, is also an ordinal number. An ordinal number α is called a successor ordinal if for some ordinal number β: $\alpha = \beta + 1$. Ordinal number α is called a limit ordinal if $\alpha = \cup\alpha$ (and $\alpha \neq 0$).



The axiom of foundation tells us that we can use the ordinal numbers as the complexity ranking of the sets in our set theoretic universe. So we can see the set theoretic universe as an expanding universe ranked along the ever ongoing line of ordinal numbers (this is called the cumulative hierarchy).

Besides ordinal numbers, we have cardinal numbers. Cardinal numbers are numbers not indicating the complexity but the size of a set (i.e. indicating that ω is infinite, but $\{\omega\}$ has one element). The definition of cardinal number is rather complex. It is easy to define when two sets have the same cardinality:

a and b have the same *cardinality* iff there is a *bijection* between a and b.

With this definition we can show that if we have a set with 6 elements and a set with 7 elements, then the second is bigger. **N** ($= \omega$) has the same cardinality as **Z** and as **Q**, but **R**, the set of real numbers is bigger (i.e. you can make a bijection between **N** and **Z** and between **N** and **Q**, but not between **N** and **R**, there are more real numbers than natural numbers. Cantor gave a famous proof of the latter theorem).

Also, for any set A, the cardinality of the power set of A is greater than the cardinality of A (this should give you a good idea of the explosion in size of the set theoretic universe, if we start with ω and keep applying the power set axiom).

When in the 19th century set theory and the notion of cardinality were developed, the cardinality of set A, $|A|$, was defined as the set of all sets with the same cardinality as A: $\{x: \exists f[f$ is a bijection between x and $A]\}$. However, that definition leads to the set theoretic paradoxes. The class of all such sets cannot itself be a set.

However, within ZF it is possible to approximate this class with a set. Let $K(A)$ be the above class of sets with the same cardinality as A. The *cardinal number* of set A, $|A|$, is the set of all sets in $K(A)$ whose complexity is smaller or equal than that of any other set in $K(A)$, that is, the set of all sets in $K(A)$ of minimal complexity.

Within the class of ordinal numbers we define an *initial number* as an ordinal number that is the smallest ordinal number of a certain cardinality. On this definition the natural numbers are initial numbers, so is ω, but not $\omega + 1$, $\omega + \omega$, $\omega.\omega$, since all those have the same cardinality as ω (countably infinite). The first initial number after ω we call ω_1 (and ω we call ω_0). Now we define: a cardinal number is an *aleph*, $\aleph$, iff it is the cardinality of an initial ordinal number. So, the cardinality of all countable sets (sets with the same cardinality as ω_0) is $\aleph_0$, and similarly, $\aleph_1 = |\omega_1|$, etc.

Two questions can now be asked: (1) Are the alephs the only cardinal numbers there are, or are there other cardinalities? (2) On the assumption that the alephs are the only cardinal numbers, given that we have sets of cardinality $\aleph_n$, do we have a method to construct sets of cardinality $\aleph_{n+1}$? Related to this is the question: what is the cardinality of **R**? We know, by Cantor's proof, that the cardinality of **R** is $2^{\aleph_0}$. But which aleph is that?

ZF doesn't answer these questions, in fact it has been proved that the statements answering them one way or another are independent of ZF.

Two axioms that are often used in mathematics, but that are not part of ZF proper take a stand on these issues.

The first is the *axiom of choice*. Many versions of this principle exist (we will see another one, Zorn's Lemma, later). Let A be a set of sets. A *choice function* for A is a function that chooses out of every non-empty element X of A some element x of X.

Axiom of choice (AC): every set has a choice function.

This principle is much debated in the foundations of mathematics. Nobody has problems with it for finite sets (that follows already in ZF), but the problem is that the principle tells you that for infinite sets, of which we have no idea of what they look like, such a function exists, but it doesn't tell you what that function looks like. It is a highly non-constructive principle: it says, even if you can't construct such a function, it still exists. Constructivists regard such a principle as nonsense, but even non-constructivists will always tell you whether a proof makes use of the principle or not.

If we assume the Axiom of Choice, then we can prove that any two sets can be compared in cardinality:

$$\text{for all } A, B\text{: } |A| \leqslant |B| \text{ or } |B| \leqslant |A|.$$

Given this, it can easily be seen that if we assume AC, we can prove that the alephs are the only cardinal numbers.

So the Axiom of Choice gives an answer to the first question. However, adding AC to ZF does not help us in answering the second questions: answers to this question are still independent from ZF + AC. Concerning this question: by Cantor's Theorem we know that $\aleph_n < 2^{\aleph_n}$. With AC we know that $\aleph_{n+1}$ is the direct successor (within the class of cardinal numbers) of $\aleph_n$. From this it follows that: $\aleph_{n+1} \leqslant 2^{\aleph_n}$. The second question, then, reduces to: are there cardinal numbers in between $\aleph_n$ and $2^{\aleph_n}$. As said, this question is independent of ZF + AC.

An axiom (which implies the axiom of choice) that answers this is the *generalized continuum hypothesis*:

$$\text{GCH: } \aleph_{\alpha+1} = 2^{\aleph_\alpha}$$

This principle tells you that the alephs (and hence the initial numbers) are the only cardinal numbers and it tells you that the cardinality of the power set of ω (which is 2^ω) is $\aleph_1$ (and not something bigger). The real numbers can be constructed with infinite subsets of the set of natural numbers. Given that, the cardinality of **R** (which is also called the continuum) is the same as the cardinality of pow(ω). The generalized continuum hypothesis hence implies that the cardinality of **R** is $\aleph_1$. (This is the *continuum hypothesis*.)

This principle has the advantage that it makes the notion of size of a set simple and elegant (in fact, we no longer have to have the cumbersome definition of cardinal numbers, we can simply take initial numbers to be cardinal numbers). However, the problem with it, and the reason why it is not part of ZF strictly (apart from the fact that it implies the axiom of choice) is that it is rather arbitrary.

The other principles of ZF all had their crucial role in building sets and constraining the notion of set, and all of them were motivated by what we think sets are, what we can do with them, and how the universe is structured. The situation with GCH is different, it has nice consequences, and that may be a reason for accepting it, but it is not our idea of what set theory is that forces us to adopt it.

I will finally make some remarks on models for ZF and models in ZF. Let me first remind you that the incompleteness theorem applies to ZF, which means that ZF is incomplete and which means that we can't prove the consistency of ZF in ZF. This means that we cannot build a model for ZF in ZF. But ZF is a very rich theory. We can build models for parts of ZF in ZF, we can build models for other set theories in ZF, and we can even say what a model for ZF in ZF should look like, although we can't prove that it exists.

Of course, given the fact that we can't prove ZF consistent, this is not going to give us any consistency proofs for such other theories, but we may learn interesting things from such models anyway. So, let us assume, take for granted that ZF is consistent.

One thing we are interested in is whether we have given the most concise axiomatization of ZF, that is, whether the ZF-axioms are independent of each other. We have already seen that that is not the case, the empty set axiom follows from the infinity axiom and separation follows from empty set and substitution. We can show in ZF that all the other axioms are independent, though.

We do that by constructing a model inside ZF for all those axioms except one. If we can do that, then it and its negation are consistent with the other axioms (its negation is, because of this model, and it is, because of the assumption that ZF is consistent).

A model for ZF-axioms consists of a universe of sets on which all the ZF axioms in question are true. The claim that there is an internal model for those axioms means that we can find an ordinal number such that if we take all the sets that have a lower complexity than this number, the structure of sets we get is closed under all the set theoretic operations that the axioms allow us (like separation, substitution, pair, union and power set).

A model for ZF would be a model that is so big that it can be closed under all operations.

Apart from independence, such models are also interesting, because they tell you what modeltheoretic property corresponds with, say, the substitution axiomschema.

Let me give one (easy) example. Take the infinity axiom, and let us call ZF – FIN the theory that we get by leaving the infinity axiom out of ZF. We can show (on the assumption that ZF is consistent) that ω_0, or rather the universe of sets of smaller complexity than ω_0 (the universe of finite sets), is a model for this theory. With that we have shown that the infinity axiom is independent of all other axioms. We have a model of the other axioms on which infinity fails (namely ω_0), and we have models on which it holds together with the other axioms (assuming ZF to be consistent). So both it and its negation are consistent with the other axioms.

Once we assume the infinity axiom, universes that satisfy other axioms as well tend to be extremely big (think again of repeatedly applying pow to ω, all those sets have to be in the universe).

We know what property a universe has to have for it to be a model of all of ZF (we assume the axiom of choice in its characterization). Look at ω_0. We know that for all $n < \omega_0$: $2^n < \omega_0$ (i.e. for all X, if $|X| < \omega$ then $|\text{pow}(X)| < \omega$). And we know that ω_0 satisfies all ZF-axioms except infinity. Now suppose we can find ordinal numbers bigger than ω (so infinity holds), but with this same property: Let α be such an ordinal number: $\forall\beta < \alpha$: $2^\beta < \alpha$.

We can prove that the universe corresponding to such an ordinal number is a model for ZF. Such an ordinal number is called a strongly inaccessible cardinal. It doesn't seem implausible to set theorists that such inaccessible cardinals in fact exist, but you can't prove it in ZF (nor can you prove in ZF that the statement that at least one of them exists is consistent with ZF). So the funny thing is that we can't prove in ZF that ZF has a model, but we can describe exactly what that model would have to look like.

At this moment I would like to make a little side step.

If you remember the discussion of the foundation axiom earlier then two things may have stuck in your memory. In the first place, that it allowed us to define a complexity ranking and in the second place that

it forbids sets of the form $x \in x$. At the very beginning of this section, I talked about the Russell Paradox and you probably remember that that precisely had to do with such weird sets. A lot of people tend to come out of an acquaintance with set theory with the idea that it is the axiom of foundation that blocks the paradox in ZF. We can have some understanding for this confusion, because in Type theory (one version of which is used in Montague's work) it is the type structure that forbids the paradox and what the foundation axiom does is very much imposing a type (= complexity) structure on set theory. Still, this illusion is wrong for ZF. As I have argued in the beginning of this section, it is the adoption of the restricted separation principle, rather than the unrestricted comprehension principle that avoids the paradox. Apart from that, you are able to argue yourself that it can't be the foundation axiom. That is the next exercise:

Exercise 12. Assume that ZF is consistent. Argue that ZF – {Foundation} is consistent.

The foundation axiom tells us that the set theoretic universe has a certain structure, it excludes models that don't have that structure, but it doesn't block the paradox, because that was already blocked. Let me add a question to this:

Exercise 13. So it's the separation axiom and not foundation. Now then, what happens to ZF if we leave out the separation axiom?

Another use of internal models is to prove the 'consistency' of other set theories or property theories. Suppose we develop a new set theory (or property theory). Then we would want to know its strength and adequacy. We usually can't prove that it is consistent (if it is not very weak), but we often can give a model for it inside ZF. We then have proved that the theory is consistent, if ZF is, this is called a relative consistency proof (relative to ZF).

Such a relative consistency proof can show us a lot. In the first place, it is nice to know that this new theory is at least not worse off than ZF. Secondly, the ZF-model for this theory may tell us a lot about this theory.

All in all this may give you the impression that, even with our hands tied behind our back by the incompleteness proofs, we can get a lot of information about consistency out of ZF.

Let me finally come back to models for ZF itself. I argued earlier that those models have to be incredibly big (inaccessible cardinals). Here then is a brain twister. Although ZF is a theory with infinitely many axioms (because of replacement and substitution), it still is a first order theory. Now we assume again that ZF is consistent, that means that it has a model, an infinite model, which – as I just said – is one of these huge ordinal numbers. But then we know, with the Löwenheim–Skolem theorem, that ZF has a *countable model*. That is, there is a model on which all the axioms of ZF are true, but which has only countably many elements. This is known as Skolem's paradox.

It is good to think for a while about how ridiculous this fact really is: here we are talking about those huge infinite ordinal numbers (like pow(pow(pow(ω_ω)), and everything we are saying is at the same time true of this tiny countable model.

The paradox is only a seeming paradox. If we recall the discussion of second order logic and the first order imitation of it, we may get a clearer view on what is going on. The fact that ZF has countable models tells us that our first order theory is not rich enough to carve out the 'real' set theoretic universe and the 'real' relation of element of, $\in$. There are countable models in which we can interpret our variables and the *symbol* $\in$ of our language in such a way that all the axioms are true. But in those models this symbol $\in$ does not have its standard interpretation of 'element of' in the set theoretic sense. The infinity axiom tells us that there is a successor set in that model, say ω. This is an element of the domain of that model. Sentences like $1 \in \omega$, $2 \in \omega, \ldots, \omega \in \text{pow}(\omega)$ are all true, but the interpretation of ω in this model is just as an arbitrary individual, and $\in$ is interpreted as *some* relation relating $1, 2, 3 \ldots$ with ω and ω with pow(ω), *not* as set membership. We cannot force $\in$ to be set membership, it is only a first order approximation.

So we have to live in ZF with Skolem's Paradox. And again, for a logician that is made easy, because these countable models for ZF are an enormous tool. There are techniques for constructing countable models for ZF. For instance, Cohen developed a technique called forcing, with which he could construct a model for ZF on which the axiom of choice is false, showing the axiom of choice to be independent of ZF, and the same for the generalized continuum hypothesis.

So again we see that the first order nature of ZF and the fact that the Löwenheim–Skolem theorem applies to it can be used as a logical/mathematical strength, rather than a weakness.

APPENDIX

The theory of dense linear orders without endpoints, which I will call D has the following axioms:

(1) $\forall x\, x \leqslant x$	reflexivity
(2) $\forall x \forall y \forall z[(x \leqslant y \wedge y \leqslant z) \rightarrow x \leqslant z]$	transitivity
(3) $\forall x \forall y[(x \leqslant y \wedge y \leqslant x) \rightarrow x = y]$	antisymmetry
(4) $\forall x \forall y[x \leqslant y \vee y \leqslant x]$	connectedness
(5) $\forall x \exists y\, x < y$	no greatest element
(6) $\forall x \exists y\, y < x$	no smallest element
(7) $\exists x \exists y\, x \neq y$	at least two elements
(8) $\forall x \forall y[x < y \rightarrow \exists z[x < z \wedge z < y]]$	density

What we are going to prove is: D is complete, i.e. for all φ: $D \vdash \varphi$ or $D \vdash \neg\varphi$.

Let me start out by borrowing some notions from Chapter Two.

Let $A = \langle A, \leqslant_A \rangle$ and $B = \langle B, \leqslant_B \rangle$ be two ordered sets.

An *isomorphism* from A to B is a one–one function f from A to B that preserves the structure, i.e. such that for all a and a' in A: $a \leqslant_A a'$ iff $f(a) \leqslant_B f(a')$.

A and B are isomorphic iff there is an isomorphism from A to B.

Let me note here an obvious fact that if f is an isomorphism from A to B then the inverse function of f, f^{-1} $(= \{\langle b, a \rangle : \langle a, b \rangle \in f\})$ is an isomorphism from B to A. Since f and f^{-1} only differ in their direction (i.e. from A to B or from B to A) we do not hesitate to call them the same isomorphism.

So, isomorphism is a mathematical relation between structures.

In our logical setting we have to adapt this notion for models. We have only one non-logical constant $\leqslant$.

Let $A = \langle A, i_A \rangle$ and $B = \langle B, i_B \rangle$ be two models.

f is an isomorphism from model A to model B iff f is an isomorphism from structure $\mathbf{A} = \langle A, i_A(\leqslant) \rangle$ to structure $\mathbf{B} = \langle B, i_B(\leqslant) \rangle$.

So this means that for all a and a' in A:

$$\langle a, a' \rangle \in i_A(\leqslant) \text{ iff } \langle f(a), f(a') \rangle \in i_B(\leqslant)$$

So the only thing we have done is rewrite the notion of isomorphism for models, but it is the same notion.

One more notion. Let f be an isomorphism between A and B. Let g be an assignment function from VAR into A. The image of g in B under f, g^* is that assignment function from VAR into B such that for all $x \in VAR$: $g^*(x) = f(g(x))$.

Here is our first theorem:

THEOREM 1. *If models A and B are isomorphic, then for every first order sentence* φ: $A \models \varphi$ *iff* $B \models \varphi$.

So, if A and B are isomorphic then the same first order sentences are true on them.

In fact, we will prove Theorem 1 indirectly, by proving something stronger:

THEOREM 2. *If A and B are isomorphic and* φ *is a first order formula, then for all* g: $A \models \varphi[g]$ *iff* $B \models \varphi[g^*]$.

Before we prove Theorem 2, let me argue that Theorem 2 implies Theorem 1. The reason is the following:

LEMMA 1. *If* φ *is a sentence and for all* g: $A \models \varphi[g]$ *iff* $B \models \varphi[g^*]$ *then* $A \models \varphi$ *iff* $B \models \varphi$.

Proof of Lemma 1. Let φ be a sentence. Remember that if φ is a sentence, $A \models \varphi$ holds iff for some g: $A \models \varphi[g]$. We assume that $A \models \varphi[g]$ iff $B \models \varphi[g^*]$.

If $A \models \varphi$ then for some g: $A \models \varphi[g]$. Then by assumption: $B \models \varphi[g^*]$. Then $B \models \varphi$.

If $B \models \varphi$ then for some h: $B \models \varphi[h]$. Now there is a unique assignment function on A, g, such that $h = g^*$ $(g = \lambda x.f^{-1}(g(x)))$, so again by the assumption: $A \models \varphi[g]$. Then $A \models \varphi$. So indeed we have proved that $A \models \varphi$ iff $B \models \varphi$. This completes the proof of the lemma.

It should be clear that Theorem 2 and Lemma 1 together imply Theorem 1. So, the only thing left to prove is Theorem 2.

Proof of Theorem 2. The proof goes with formula induction. Assume that f is an isomorphism from A to B (hence A and B are isomorphic).

1. *Base step*

Atomic formulas have the following form:

either: $x \leqslant y$ (x and y variables)
or: $x = y$

We have to prove for all g:

(a) $A \models x \leqslant y[g]$ iff $B \models x \leqslant y[g^*]$

that is, $\langle g(x), g(y)\rangle\rangle \in i_A(\leqslant)$ iff $\langle g^*(x), g^*(y)\rangle \in i_B(\leqslant)$.
$g^*(x) = f(g(x))$; $g^*(y) = f(g(y))$, so this follows from the fact that A and B are isomorphic (in other words, this is a place where you crucially use that assumption).

(b) $A \models x = y[g]$ iff $B \models x = y[g^*]$

that is, $g(x) = g(y)$ iff $g^*(x) = g^*(y)$. This follows from the fact that our isomorphism f is a one–one function, i.e.:

$$a = a' \text{ iff } f(a) = f(a')$$

Here again we use the fact that A and B are isomorphic. This completes the base step. We will see that the induction steps follow mainly from the meaning of the logical constants:

2. *Induction steps*

(a) Assume (induction hypothesis (IH)):

for all g: $A \models [g]$ iff $B \models \varphi[g^*]$

prove:

for all g: $A \models \neg\varphi[g]$ iff $B \models \neg\varphi[g^*]$

This follows from the following list of equivalences: $A \models \neg\varphi[g]$ iff [semantics $\neg$] $A \not\models \varphi[g]$ iff [IH] $B \not\models \varphi[g^*]$ iff [semantics $\neg$] $B \models \neg\varphi[g^*]$.

(b) Assume: (IH1) for all g: $A \models \varphi[g]$ iff $B \models \varphi[g^*]$ and assume: (IH2) for all g: $A \models \psi[g]$ iff $B \models \psi[g^*]$
Prove:

for all g: $A \models (\varphi \wedge \psi)[g]$ iff $B \models (\varphi \wedge \psi)[g^*]$

Again a list of equivalences:

$A \models (\varphi \wedge \psi)[g]$ iff $A \models \varphi[g]$ and $A \models \psi[g]$ iff (IH1 and IH2) $B \models \varphi[g^*]$ and $B \models \psi[g^*]$ iff $B \models (\varphi \wedge \psi)[g^*]$

(c) Assume: (IH) for all g: $A \models \varphi[g]$ iff $B \models \varphi[g^*]$
Prove:

for all g: $A \models \exists x\varphi[g]$ iff $B \models \exists x\varphi[g^*]$

Again: assume $A \models \exists x\varphi[g]$. Then for some $a \in A$; $A \models \varphi[g_x^a]$. Now the induction hypothesis is that for all g: $A \models \varphi[g]$ iff $B \models \varphi[g^*]$. An instance of that is: $A \models \varphi[g_x^a]$ iff $B \models \varphi[g^{*f(a)}_x]$. So we can conclude $B \models \varphi[g^{*f(a)}_x]$. and hence $B \models \exists x\varphi[g^*]$.

Assume $B \models \exists x\varphi[h]$. Then for some $b \in B$: $B \models \varphi[h_x^b]$. Again, we know that there is a unique assignment g on A such that $h = g^*$, and hence also $h_x^b = g^{*f-1(b)}_x$. Then we know with the induction hypothesis that $A \models \varphi[g^{*f-1(b)}_x]$. And then we know that $A \models \exists x\varphi[g]$.

We have gone through all the cases now, so we have completed our proof of Theorem 2. With that we have also proved Theorem 1, the theorem we wanted to prove, and the theorem that we will use later.

Let us now make an observation:

LEMMA 2. *D does not have any finite models.*

Proof. The main reason is density. The theorem holds already for dense linear orders with at least two elements. A model for D has at least two elements (Axiom 8). It is a simple observation that density tells you that between any two elements there are infinitely many distinct other elements.

Suppose that only finitely many elements lie between a and a', say n: $a < a_1 < \cdots < a_n < a'$. Density tells you that, for instance, between a and a_1 there is a third element, so the assumption that n elements lie between a and a' is false. This holds for any n, so indeed D has only infinite models.

Let me now state the main theorem, but leave the proof to the end. The main theorem is a famous mathematical theorem, first proved by Georg Cantor:

THEOREM 3. *Any two countable dense linear orders without end points are isomorphic.*

It is clear that given the notion of isomorphism for models that we have defined, this means that any two countable models for D are isomorphic.

Assuming all the theorems we have proved now plus Theorem 3, let's put them together to prove the starting theorem:

THEOREM 4. *D is complete.*

Proof. Suppose there is a first order sentence φ such that $D \nvdash \varphi$ and $D \nvdash \neg\varphi$ (i.e. D is incomplete). Then both $D \cup \{\varphi\}$ and $D \cup \{\neg\varphi\}$ are consistent. Then by *completeness* both $D \cup \{\varphi\}$ and $D \cup \{\neg\varphi\}$ have a model, say, $M \models D \cup \{\varphi\}$ and $M' \models D \cup \{\neg\varphi\}$.

Since all models for D are infinite, so are all models for $D \cup \{\varphi\}$ and $D \cup \{\neg\varphi\}$. But then it follows with Löwenheim–Skolem that both $D \cup \{\varphi\}$ and $D \cup \{\neg\varphi\}$ have a *countable model*, say $A \models D \cup \{\varphi\}$ and $B \models D \cup \{\neg\varphi\}$.

A is a model for $D \cup \{\varphi\}$, so it is also a model for D, and similarly B is also a model for D.

But, by Cantor's theorem, we know that all countable models for D are isomorphic. Hence A and B are isomorphic.

But then we know, by Theorem 1, that the same first order sentences are true on A and B. Since A is a model for $D \cup \{\varphi\}$, φ is true on A. φ is a first order sentence, so φ is true on B, and similarly $\neg\varphi$ is true on A. That means that both $A \models (\varphi \wedge \neg\varphi)$ and $B \models (\varphi \wedge \neg\varphi)$, which is a contradiction.

So indeed there is no sentence φ such that $D \nvdash \varphi$ and $D \nvdash \neg\varphi$, so for all φ : $D \vdash \varphi$ or $D \vdash \neg\varphi$, hence D is complete.

The only thing left to prove, then, is Cantor's theorem:

THEOREM 3. *Any two countable dense linear orders without end points are isomorphic.*

The proof is, as I said, very famous, and it goes with a so-called zigzag or back and forth construction (a device very commonly used since).

We will prove the theorem with $<$ rather than $\leqslant$. This doesn't make a difference.

Let me define a few notions.

A *partial* isomorphism from $\langle A, <\rangle$ to $\langle B, <\rangle$ is an isomorphism between a finite subset of A (ordered by $<$ as in A) and a finite subset of B (ordered by $<$ as in B).

Let p be a partial isomorphism from $\langle A, <\rangle$ to $\langle B, <\rangle$.

We call an element from A or B *new with respect to p* iff that element is neither in the domain, nor in the range of p. If an element is not new with respect to p it is *old with respect to p*. The same for q where q is a partial isomorphism from $\langle B, <\rangle$ to $\langle A, <\rangle$.

What I said before for isomorphisms holds as well for partial isomorphisms: a partial isomorphism and its inverse differ only with respect to their direction. I will here use p if the function goes from A to B and q if it goes from B to A.

It is time for a lemma:

LEMMA 3. *If p is a partial isomorphism from $\langle A, <\rangle$ to $\langle B, <\rangle$ and $a \in A$ is new wrt. p then p can be extended to a partial isomorphism p' from $\langle A, <\rangle$ to $\langle B, <\rangle$ with domain* $\mathrm{dom}(p) \cup \{a\}$ *(the same, of course, for q).*

Proof. There are three cases.

(a) *Either:* for every $a' \in \mathrm{dom}(p)$: $a < a'$. Because $<$ has no begin point in B, we can always find an element $b \in B$ which is new wrt. p such that for all $b' \in \mathrm{ran}(p)$: $b < b'$. For any such b, $p \cup \{\langle a, b\rangle\}$ is a partial isomorphism.

(b) *Or:* for every $a' \in \mathrm{dom}(p)$: $a' < a$. Exactly the same argument holds, because $<$ has no end point in B.

(c) *Or:* a lies between some a' and $a'' \in \mathrm{dom}(p)$: $a' < a < a''$. Since B is densely ordered by $<$, we can always find a new $b \in B$ such that: $p(a') < b < p(a'')$ and the same argument holds: for any such b, $p \cup \{\langle a, b\rangle\}$ is a partial isomorphism. This concludes the proof of the lemma.

Now both $\langle A, <\rangle$ and $\langle B, <\rangle$ are countable. This means that for each one of them there is a way of counting its elements. Such a way of counting is called an enumeration. So let α be an enumeration of $\langle A, <\rangle$ and β an enumeration of $\langle B, <\rangle$. α is one way of putting all elements of A in an infinite list. α_1 is the first element of that list, α_2 the second, etc. So we have the following listings of A and B:

$$\alpha_1, \alpha_2, \alpha_3, \ldots.$$
$$\beta_1, \beta_2, \beta_3, \ldots.$$

We are going to construct, using these enumerations, a sequence of partial isomorphisms between A and B. The idea is the following: we are going to make a one–one function between A and B and we are going to do that in stages, where in every stage we make sure that we preserve the order. We use the enumerations to make sure that in the construction we hit every element of A and every element of B and hit them only once. Else we won't have a one-one function.

Again, I write p (or p_n) if the function goes from A to B and q if the function goes from B to A, but remember that $q = p^{-1}$ and $p = q^{-1}$.

We start in A and set: $p_0(\alpha_1) = \beta_1$.

Now we take α_2.

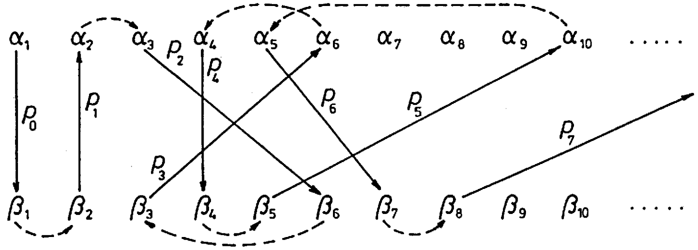
If $\alpha_2 < \alpha_1$ we let $p_1(\alpha_2)$ be the first element b in enumeration β of B such that $b < \beta_1$. So $p_1 = \{\langle \alpha_1, \beta_1\rangle, \langle \alpha_2, b\rangle\}$. p_1 is a partial isomorphism from A to B. If $\alpha_1 < \alpha_2$ we let $p_1(\alpha_2)$ be the first element b in β such that $\beta_1 < b$, and also then p_1 is a partial isomorphism. So is q_1, because it is the same isomorphism from the other direction. We now continue in β. We look for the first element in the list β that is new with respect to p_1. (That is β_2 if we haven't already mapped α_2 on it, else it is β_3.) Let this be β_n.

- If $\beta_n < p_1(\alpha_1)$ and $\beta_n < p_1(\alpha_2)$ we let $q_2(\beta_n)$ be the first new element a of α such that $a < \alpha_1$ and $a < \alpha_2$.
- If $p_1(\alpha_1) < \beta_n$ and $p_1(\alpha_2) < \beta_n$ we let $q_2(\beta_n)$ be the first new element a of α such that $\alpha_1 < a$ and $\alpha_2 < a$.
- If $p_1(\alpha_1) < \beta_n < p_1(\alpha_2)$ we let $q_2(\beta_n)$ be the first new element a of α such that $\alpha_1 < a < \alpha_2$.

Now we have partial isomorphism with three elements.

Now we look in α again and we look at the first element of α that is new with respect to q_2 and we repeat the procedure.

In general, we extend partial isomorphism $p_n : A \to B$ by choosing the first element of α which is new wrt. p_n: either this one is before all elements in $\mathrm{dom}(p_n)$ and we choose the first new element of β before all p_n-values, or it is after all element in $\mathrm{dom}(p_n)$ and we choose the first new element in β after all p_n-values, or it lies somewhere in the middle and we look at the two old elements, which, of the elements of $\mathrm{dom}(p_n)$, are the direct neighbors of that new element, and we choose the first new element in β in-between the values of these neighbors. This gives us p_{n+1}. Then we choose the smallest new element in β and we start the same procedure from B, giving us q_{n+2} ($= p_{n+2}$). We get a picture like the following:



The crucial point, and the reason for looking all the time at the first new element satisfying a condition and for the zigzagging between α and β is that in this way we work ourselves completely through both enumeration α and β: every element comes in a partial isomorphism at some stage (either because it is chosen as the value of some β_n, or, if it is not, because then at some stage it will be the first new element in α) and the same holds for β. The zigzag construction guarantees that. The requirement that we always choose new elements furthermore guarantees that the elements in α and β are chosen only once. In the limit this process gives us a one-one function between A and B and the construction guarantees that it is an isomorphism. More precisely, given the sequence of functions $p_1, p_2, p_3, p_4, \ldots$ we can define the following function f from A to B: f is that function from A to B such that:

$$\text{for every } \alpha_n \in A: f(\alpha_n) = p_n(\alpha_n)$$

We should check that this is well defined. It is, if for every n, $p_n(\alpha_n)$ is defined (has a value in B). This is the case, because either at stage n α_n was chosen earlier as the value of some β_m, and then $p_m(\alpha_n) = \beta_m$, or at stage n α_n is the first element in α that is new wrt. p_{n-1} and p_n gives it a value.

f clearly is a one–one function (we have argued that above) and it preserves the structure of $\langle A, <\rangle$:

$$\text{for all } a, a' \in A\text{: if } a < a' \text{ then } f(a) < f(a').$$

Let $n < m$ and let $a = \alpha_n$ and $a' = \alpha_m$.

If $a < a'$ then by the construction $p_m(\alpha_n) < p_m(\alpha_m)$; in other words, $p_m(a) < p_m(a')$. Since $f(\alpha_n) = p_m(\alpha_n)$ and $f(\alpha_m) = p_m(\alpha_m)$, that is, $f(a) = p_m(a)$ and $f(a') = p_m(a')$, it indeed follows that $f(a) < f(a')$.

Concluding: by using the fact that A and B are countable, and hence have enumerations, we have defined a zigzag construction of partial isomorphisms (using the fact that A and B are dense linear orders without endpoints) on these enumerations, and in terms of that we were able to define an isomorphism between $\langle A, <\rangle$ and $\langle B, <\rangle$.