1. Foreword

In model theory as in no other branch of mathematics (maybe with the exception of logic, out of which model theory grew anyway) we mind our language. Swearing is frowned upon, to be sure, but also, and more importantly, we distinguish clearly between what we talk about (*semantics*) and the language we use to talk about it (*syntax*). To see what is involved, consider the sentence

This sentence is false.

If the sentence is true, then what it says is the case, so it is false. If it is false, then what it says is not the case, so it is not false; hence true. Contradiction! This is the famous *liar paradox*, known to the ancient Greeks, and often attributed to Epimenides the Cretan (about 600BC) who is reported to have said that all Cretans are liars. It has always been clear that the problem with liar paradox is that the sentence displayed above speaks about itself: *it* says of *it*self that *it* is false. Many clever and not-so-clever ways of dealing with the paradox have been tried in the course of about 26 centuries, before Alfred Tarski in his article *The concept of truth in formalized languages*[1] offered a solution, which became widely accepted[2] and, one can say, started off model theory. We will see what his solution is when we meet the notion of *satisfaction*.

In usual mathematical contexts such paradoxes do not occur, for the most part, so good old ways of not distinguishing clearly between language and its object are perfectly acceptable. But occasionally even there we can be led astray. Consider Euclid's *Five Postulates* of what is now known as Euclidean geometry.

(1) A straight line segment can be drawn joining any two points.
(2) Any straight line segment can be extended indefinitely in a straight line.
(3) Given any straight line segment, a circle can be drawn having the segment as radius and one endpoint as center.
(4) All right angles are congruent.
(5) Given any straight line and a point not on it, there exists one and only one straight line which passes through that point and never intersects the first line, no matter how far they are extended.

[1]Published in 1933 in Polish as *Pojęcie prawdy w językach nauk dedukcyjnych*. English translation appeared in 1956.

[2]Except by (some) philosophers.

For about 23 centuries many mathematicians tried to prove that the fifth postulate follows from the first four. But, it is rather easy to see that interpreting “line segment” as “segment of a great circle of some fixed ball” we have postulates (1) through (4) come out true, but postulate (5) false. So it cannot follow from the other four. Now, could otherwise brainy people really have overlooked such a simple proof? Most probably not. What is much more likely to have happened is they did not think of words “line”, “point”, “circle” etc., as *defined by the postulates alone*, but as something that is *intuitively clear* and also happens to be described by the postulates. Completely described, they believed, as it turned out, wrongly. Although of course we engage here in a bit of historical fiction, we can say that their semantics creeped into their syntax and clouded their vision.

CHAPTER 1

Syntax: first order languages and first order logic

The mysterious *first order* means, intuitively, that our variables range over "basic objects" (whatever they are), but not over sets of objects, or relations (including functions) between objects. For example, in arithmetic, the sentence "for every number x there is a number y such that $x < y$", is first order, whereas "if K is a set of numbers such that: (i) 0 is in K, and (ii) for every number n, if n is in K, then the successor of n is in K, then K contains every number", is not. Now, let us make this precise.

A *first order language* consists of:

- A countable set $\mathsf{Var} = \{v_n : n \in \omega\}$ of *variables*.
- A possibly empty set $\mathsf{Const} = \{c_k : k \in K\}$ of *constant symbols*,
- A possibly empty set $\mathsf{Func} = \{f_j^n : j \in J\}$ of *function symbols* of *arity* n, for each $0 < n$.
- A set $\mathsf{Rel} = \{R_i^n : i \in I\}$ of *relation symbols*, (also called *predicates*) of arity n, for each $0 < n$. The set Rel is nonempty: it contains as least one member, R_0^2, called the equality symbol, and written $=$.
- *Quantifiers*: $\forall$, $\exists$
- Logical *connectives*: $\wedge, \vee, \rightarrow, \leftrightarrow, \neg$.
- parentheses: (,).

The sets Const and Func can be of any cardinality (empty, finite, countably infinite, uncountably infinite), and Rel can be of any cardinality greater than 1. In practice however, all these sets will be at most countable, and very often finite.

A *signature* (or *type*) of a language L is an enumeration of relation, function and constant symbols present in L, together with their respective arities. In this enumeration, identity symbol is normally not mentioned, unless it is the only relation symbol of L (in which case L is called the *pure identity language*). If a signature has no relation symbols (except $=$) it is called *algebraic*. Languages L_1 and L_2 are said to be *similar*, or of the same type, if for any arity n there are bijective mappings between relation (function, constant) symbols of arity n in L_1 and relation (function, constant) symbols of arity n in L_2.

Example 1.1

Three examples of typical first order languages:

(1) The language L_{group} of groups is of the type $\langle f_0^2, f_0^1, c_0 \rangle$. In particular, $\mathsf{Rel}_{\text{groups}}$ contains only the identity symbol, so it is an algebraic language. Usually, the function symbol f_0^2 is written as $+$ or $\cdot$, the function symbol f_0^1 is written as $-$ or ${}^{-1}$, and the constant symbol c_0 is written as 0 or e.

(2) The language L_{graph} of graphs is of the type $\langle R_1^2 \rangle$. Usually R_1^2 is written as E. Notice that R_0^2 was not mentioned, as it is reserved for the identity symbol.

(3) The language L_{ord} of ordered sets is of the type $\langle R_1^2 \rangle$. Usually R_1^2 is written as $\leq$.

Let L be a language. Any finite string of symbols of L is an *expression*, certain expressions are *well-formed* others are not. Well-formed expressions are either *terms* or *formulae*. We define them inductively. A *term* of L is defined inductively as follows:

- every variable is a term,
- every constant symbol is a term,
- if f is an n-ary function symbol and $t_1, \ldots, t_n$ are terms, then $f(t_1, \ldots, t_n)$ is a term,
- nothing else is a term.

We will write Term_L (or just Term, if L is clear from context) for the set of terms of L.

Example 1.2

Some examples of terms:

- If L is L_{group} of Example 1.1, then Term_L contains strings c_0, $f_0^2(v_1, v_7)$, $f_0^2(v_8, v_8)$, $f_0^1(v_3)$, $f_0^2(c_0, v_7)$, etc. Normally, these would be written either multiplicatively, as e, xy, zz, u^{-1}, ey, or additively, as 0, $x + y$, $z + z$, $-u$, $0 + y$. Variables are typically written not as $v_1, v_2, \ldots$ but using a selection of letters from somewhere close to the end of the alphabet.
- If L is L_{graph} or L_{ord} of Example 1.1, then the only terms of L are variables, that is, $\mathsf{Term}_L = \mathsf{Var}_L$.

Next, we define formulae. An *atomic formula* of L is a string of the form $R(t_1, \ldots, t_n)$, where R is an n-ary relation symbol, and $t_1, \ldots, t_n$

are terms. This includes strings of the form $R_0^2(s,t)$ for terms s and t; these are normally written as $s = t$. Also, if R is of arity 2, it can be written in the infix form, that is as vRw rather than $R(v,w)$. A *formula* of L is defined inductively as follows:

- every atomic formula is a formula,
- if φ is a formula, and v a variable, then $(\forall v)\ \varphi$, $(\exists v)\ \varphi$ and $\neg\varphi$ are formulae,
- if φ, ψ are formulae, and $\star \in \{\wedge, \vee, \rightarrow, \leftrightarrow\}$, then $\varphi \star \psi$ is a formula
- nothing else is a formula

We will write For_L (or just For, if L is clear from context) for the set of formulae of L.

Example 1.3

Some examples of formulae:

- If L is L_{group}, then For_L contains strings $c_0 = f_0^2(v_1, v_7)$, $\neg f_0^2(v_8, v_8) = f_0^1(v_3)$, $(\forall v_7)\ \neg f_0^2(c_0, v_7) = v_3$, etc. Again, they would normally be written as $e = xy$, $zz \neq u^{-1}$, $(\forall y)\ ey \neq u$. So, the axioms of group theory are usually written as
  - $xe = x$, $x = ex$,
  - $xx^{-1} = e$, $x^{-1}x = e$,
  - $x(yz) = (xy)z$.
- If L is L_{ord}, then formulae of L include $(\forall x)(\exists y)\ x \leq y \vee y \leq x$, $(\forall x)\ x \leq y$, $(\exists x)(\forall z)\ \neg(x \leq u) \vee (\exists u)\ u \leq x$. Let us recall some formulae one often sees in connection with ordering relations:
  - $(\forall x)\ x \leq x$ (reflexivity)
  - $(\forall x)(\forall y)(\forall z)\ x \leq y \wedge y \leq z \rightarrow x \leq z$ (transitivity)
  - $(\forall x)(\forall y)\ x \leq y \wedge y \leq z \rightarrow x = y$ (antisymmetry)
  - $(\forall x)(\forall y)(\exists z)\ x \leq y \wedge x \neq y \rightarrow$
    $x \leq z \wedge z \leq y \wedge x \neq z \wedge z \neq y$ (density)
  - $(\forall x)(\forall y)\ x \leq y \vee y \leq x$ (linearity)
  - $(\exists x)(\forall y)\ x \leq y$ (there is a smallest element)
  - $(\forall x)(\exists y)\ x \leq y \wedge x \neq y$ (no largest element)

For a term t (or a formula φ) we write $\mathsf{Var}(t)$ (or $\mathsf{Var}(\varphi)$) for the set of variables occurring in t (or φ). Occurrences of variables in a formula can be *free* or *bound*. An occurrence of v in φ is bound if it is within the scope of $\forall v$ or $\exists v$; otherwise it is free. A formula without free variables is called a *closed* formula or a *sentence*.

Free variables are free to substitute terms for them. The result of substituting a term t for a variable v in a formula φ, written (typically) as $\varphi(v/t)$, is the formula arising out of φ by replacing uniformly all free occurrences of v in φ by (occurrences of) t. We will often write $\varphi(v_1, \ldots, v_n)$ with $\{v_1, \ldots, v_n\} \subseteq \mathsf{Var}(\varphi)$ being the set of variables occurring freely in φ. Using this notation, we can write $\varphi(v_1, \ldots, v_{i-1}, t, v_{i+1}, \ldots, v_n)$ for $\varphi(v_i/t)$.

More formally, a *substitution* is any function $\sigma\colon \mathsf{Var} \to \mathsf{Term}$, extended to terms by putting $\sigma(t(v_1, \ldots, v_n)) = t(\sigma(v_1), \ldots, \sigma(v_n))$, and to formulae by putting $\sigma(\varphi(v_1, \ldots, v_n)) = \varphi(v_1/\sigma(v_1), \ldots, v_n/\sigma(v_n))$, or in our shorthand notation $\sigma(\varphi(v_1, \ldots, v_n)) = \varphi(\sigma(v_1), \ldots, \sigma(v_n))$. Using our shorthand for a particular substitution σ_0 such that $\sigma_0(v) = t$ and $\sigma_0(u) = u$ for any variable u different from v, we have $\sigma_0(\varphi) = \varphi(v/t)$.

Example 1.4

Let $\mathcal{L}$ be the language with two binary operations: $\cdot$ and $+$. Let φ be $(\exists u)\ x + u = y$. Consider two substitutions:

- $\varphi(x/(x \cdot y), y/z)$ that is $(\exists u)\ x/(x \cdot y) + u = y/z$ that is $(\exists u)\ (x \cdot y) + u = z$
- $\varphi(x/(x \cdot u), y/z)$ that is $(\exists u)\ x/(x \cdot u) + u = y/z$ that is $(\exists u)\ (x \cdot u) + u = z$

If we interpret φ in $\mathbb{N}_0$ with the usual arithmetical operations, and evaluate the variables x, y, z all to 1. Then, $(\exists u)\ x + u = y$ gives $(\exists u)\ 1 + u = 1$, satisfied by $u = 1$. The first substitution gives $(\exists u)\ (1 \cdot 1) + u = 1$, again satisfied by $u = 1$. The second gives $(\exists u)\ (1 \cdot u) + u = 1$, which cannot be satisfied. Such substitutions will not be *allowed*.

To avoid problems such as the one of Example 1.4, we define a term t with free variables $v_1, \ldots, v_n$ to be *allowed for* a variable v *in* a formula φ, if v is free in φ, and none of $v_1, \ldots, v_n$ becomes bound in $\varphi(v/t)$. Fortunately, in practice people have the natural good sense to avoid such substitutions, but to state the axioms of first order logic we will need the notion.

A word on terminology: here we say a term t is *allowed for* a variable v in a formula φ. In many books (including stadard textbooks) such a t is called *free for* v in φ instead. This has nothing to do with 'free' in the sense of 'free variable'. It can be confusing, but unfortunately it is standard terminology.

1. Logic in an abstract setting

Languages are only useful if they can express something interesting. In a mathematical context, we would like to state assumptions (axioms among them) and draw conclusions (prove theorems). Enter logic, which deals with what follows from what. Let L be a first order language. To begin in a rather abstract setting, we define a *consequence operator* to be a map $\mathsf{Cn}\colon \wp(\mathsf{For}) \to \wp(\mathsf{For})$ such that:

- $X \subseteq \mathsf{Cn}(X)$,
- $X \subseteq Y \Rightarrow \mathsf{Cn}(X) \subseteq \mathsf{Cn}(Y)$,
- $\mathsf{Cn}(\mathsf{Cn}(X)) = \mathsf{Cn}(X)$.

If Cn is moreover *invariant under substitution*, that is, if

- $\varphi \in \mathsf{Cn}(X)$ implies $\sigma(\varphi) \in \mathsf{Cn}(\sigma(X))$, for any *substitution* σ,

then Cn is *structural*. Further, if Cn also satisfies

- $\mathsf{Cn}(X) = \bigcup\{\mathsf{Cn}(Y)\colon Y \subseteq_{fin} X\}$

then it is *finitary* (or *compact*, or *algebraic*). A *logic* then, is the set $\mathsf{Cn}(\varnothing)$ for a structural (but not necessarily finitary) consequence operator. Once a consequence operator has been defined and is clear from context, we write $X \vdash \varphi$ instead of $\varphi \in \mathsf{Cn}(X)$.

2. First order logic as a Hilbert system

By far the most common way of specifying a Cn is via *axioms* and *rules of inference*. Such a presentation of a logic is called a *Hilbert system*. The role of axioms is just what one expects: they are basic assumptions. Rules of inference are basic recipes for constructing proofs: they say, if you have such and such assumptions, you can infer a conclusion of such and such form. One possible Hilbert-style presentation of first order logic goes as follows.

2.1. Axioms.

The *pure axiom schemes* are:

A1. $\varphi \to (\psi \to \varphi)$.
A2. $(\varphi \to (\psi \to \chi)) \to ((\varphi \to \psi) \to (\varphi \to \chi))$.
A3. $(\neg\varphi \to \neg\psi) \to (\psi \to \varphi)$.
A4. $(\forall v\ \varphi) \to \varphi(v/t)$, where t is allowed for v in φ.
A5. $\varphi \to \forall v\ \varphi$, where v is not free in φ.
A6. $\big(\forall v\ (\varphi \to \psi)\big) \to (\varphi \to \forall v\ \psi)$, where v is not free in φ.

The next group of pure axiom schemes deals with the equality relation:

E1. $v = v$.
E2. $v = w \to w = v$.
E3. $v = w \wedge w = z \to v = z$.

E4. $v_1 = w_1 \wedge \cdots \wedge v_n = w_n \to f(v_1, \ldots, v_n) = f(w_1, \ldots, w_n)$, for any function symbol f.

E5. $v_1 = w_1 \wedge \cdots \wedge v_n = w_n \to R(v_1, \ldots, v_n) \leftrightarrow R(w_1, \ldots, w_n)$, for any relation symbol R.

An *axiom* is an instantiation of any of the schemes, prefixed by a (possibly empty) string of universal quantifiers. This way, each schema above represents infinitely many axioms: one for a particular choice of the formulae φ, ψ, χ, the variables, terms, etc., and a prefix of universal quantifiers.

Example 1.5

Consider the language with a unary relation symbol P, a binary relation symbol R, and a unary function symbol t. Then each of the following is an axiom:

- $\forall x\ \forall y\ R(x, z) \to \big(\neg P(v) \to R(x, z)\big)$
  / A1 preceded by universal quantifiers /
- $(\forall x\ x = x) \to y = y$
  / A4: because y is free for x in y /
- $\forall x\ x = f(x) \to f(x) = f(f(x))$
  / E4 preceded by a universal quantifier /

Logical connectives other than implication ($\to$), negation ($\neg$) and the universal quantifier are "officially" viewed as abbreviations.

- $\varphi \vee \psi$ abbreviates $\neg\varphi \to \psi$,
- $\varphi \wedge \psi$ abbreviates $\neg(\varphi \to \neg\psi)$,
- $\varphi \leftrightarrow \psi$ abbreviates $(\varphi \to \psi) \wedge (\psi \to \varphi)$,
- $(\exists v)\ \varphi$ abbreviates $\neg(\forall v)\ \neg\varphi$.

However, in practice, we will use them just as if they were part and parcel of the language.

There is only one rule of inference, *modus ponens*

R1. From φ and $\varphi \to \psi$, infer ψ.

2.2. Proofs. Let $T \cup \{\varphi\}$ be a set of formulae ($\varphi \in T$ is allowed). We say that a sequence $\alpha_1, \ldots, \alpha_n$ of formulae is a *proof* of φ from (assumptions in) T if

- α_n is φ
- each α_k for $k \in \{1, \ldots, n\}$ is a member of T, or an axiom, or for some $i, j < k$ we have $\alpha_j = \alpha_i \to \alpha_k$ (that is, α_k is a result of applying *modus ponens* to some earlier steps in the proof).

Example 1.6

Here is what a proof of $\varphi \to \varphi$ from an empty set of assumptions looks like:

(1) $\big(\varphi \to \big((\varphi \to \varphi) \to \varphi\big)\big) \to \big(\big(\varphi \to (\varphi \to \varphi)\big) \to \big(\varphi \to \varphi\big)\big)$
/ A2 with χ replaced by φ and ψ by $\varphi \to \varphi$ /

(2) $\varphi \to \big((\varphi \to \varphi) \to \varphi\big)$
/ A1 with ψ replaced by $\varphi \to \varphi$ /

(3) $\varphi \to (\varphi \to \varphi)$
/ A1 with ψ replaced by φ /

(4) $\big(\varphi \to (\varphi \to \varphi)\big) \to \big(\varphi \to \varphi\big)$
/ modus ponens applied to (2) and (1) /

(5) $\varphi \to \varphi$
/ modus ponens applied to (3) and (4) /

So it takes five lines to prove a statement as ponderous as "if 3 is less than 2, then 3 is less than 2". Ah, the beauty of logic! But, just to be fair, this is not how the proof systems are used in practice. A proof systems is not so much a tool to prove theorems *in* the system, but to prove theorems *about* the system (sometimes referred to as *meta-theorems*). We will see some examples shortly, but before we do it, let us make a rather obvious observation. Once a formula has been shown to be a theorem, it can be used in other proofs: you can always precede it by its proof. To this end, the following cheat list of theorems may be useful.

C1. $(\varphi \wedge \psi) \to \varphi$ (laws of conjunction)
C2. $(\varphi \wedge \psi) \to \psi$
C3. $(\chi \to \varphi) \to ((\chi \to \psi) \to (\chi \to (\varphi \wedge \psi)))$
D1. $\varphi \to (\varphi \vee \psi)$ (laws of disjunction)
D2. $\psi \to (\varphi \vee \psi)$
D3. $(\varphi \to \chi) \to ((\psi \to \chi) \to ((\varphi \vee \psi) \to \chi))$
N1. $(\varphi \to \psi) \to ((\varphi \to \neg\psi) \to \neg\varphi)$ (laws of negation)
N2. $\varphi \to (\neg\varphi \to \psi)$
N3. $\varphi \leftrightarrow \neg\neg\varphi$
N4. $(\varphi \to \neg\varphi) \to \neg\varphi$
EQ1. $(\varphi \leftrightarrow \psi) \to (\varphi \to \psi)$ (laws of logical equivalence)
EQ2. $(\varphi \leftrightarrow \psi) \to (\psi \to \varphi)$
EQ3. $(\varphi \to \psi) \to ((\psi \to \varphi) \to (\varphi \leftrightarrow \psi))$
CO1. $(\varphi \to \psi) \to (\neg\psi \to \neg\varphi)$ (contraposition laws)
CO2. $(\varphi \to \neg\psi) \to (\psi \to \neg\varphi)$
CO3. $(\neg\varphi \to \psi) \to (\neg\psi \to \varphi)$
DM1. $\neg(\varphi \wedge \psi) \leftrightarrow (\neg\varphi \vee \psi)$ (DeMorgan laws)

DM2. $\neg(\varphi \lor \psi) \leftrightarrow (\neg\varphi \land \psi)$
EM. $\varphi \lor \neg\varphi$ (excluded middle)
NC. $\neg(\varphi \land \neg\varphi)$ (non-contradiction)
EX. $(\varphi \land \neg\varphi) \to \psi$ (explosion: contradiction implies everything)

You can try your hand at proving some of them from the axioms: it can be a challenging game. It was a favourite pastime of logicians in the early 20th Century, when they tried to find the smallest possible axiomatic base for logic. Our axiom system here is essentially due to Łukasiewicz.

2.3. Provability. We begin with a definition.

Definition 1.1

Let $T \cup \{\varphi\}$ be a set of formulae. We say that φ is *provable from* T (written $T \vdash \varphi$) if there is a proof of φ from T. If $\varnothing \vdash \varphi$, we call φ a logical theorem (and write $\vdash \varphi$).

We can now state what the consequence operator for first order logic is. If X is a set of formulae, then $\mathsf{Cn}(X) = \{\varphi\colon X \vdash \varphi\}$. The following important result is an immediate corollary.

Theorem 1.1: Syntactic compactness

The consequence operator of first order logic is compact. That is, if for some set T of formulae and a formula φ, we have $T \vdash \varphi$, then for some finite $T_0 \subseteq T$ we have $T_0 \vdash \varphi$

Proof. Let $\alpha_1, \ldots, \alpha_n$ be a proof of φ from T. Then, putting $T_0 = \{\alpha_1, \ldots, \alpha_n\} \cap T$ we get a finite subset T_0 of T such that $\alpha_1, \ldots, \alpha_n$ is a proof of φ from T_0. $\square$

Interestingly, a semantic version of compactness is much more roundabout. One can use *completeness* of first-oder logic (to come in Chapter 3), or *ultraproducts* (to come somewhere***).

On the other hand, the following three theorems are easier to prove semantically than syntactically. It fact, their model-theoretic proofs will be given in Chapter 3. However, for these proofs we need completeness theorem, and the standard proof of completeness needs them in turn, so we state and prove them here. They formally reflect three, or in fact four, familiar proof techniques: (i) proof of an implication by assuming the antecedent, (ii) two forms of proof by contradiction, (iii) proof by cases, and (iv) proof by an arbitrary-but-fixed instance.

Theorem 1.2: Deduction Theorem

Let $T \cup \{\varphi, \psi\}$ be a set of formulae. The following equivalences hold:

(1) $T \cup \{\varphi\} \vdash \psi$ iff $T \vdash \varphi \to \psi$
(2) $T \cup \{\varphi\} \vdash \chi \wedge \neg\chi$ for some formula χ iff $T \vdash \neg\varphi$
(3) $T \cup \{\neg\varphi\} \vdash \chi \wedge \neg\chi$ for some formula χ iff $T \vdash \varphi$

Proof. For the backward direction of (1), let $\alpha_1, \ldots, \alpha_n$ be a proof of $\varphi \to \psi$ from T (so, α_n is $\varphi \to \psi$). Observe that it is also a proof of $\varphi \to \psi$ from $T \cup \{\varphi\}$. Then the sequence $\alpha_1, \ldots, \alpha_n, \varphi, \psi$ is a proof of ψ from $T \cup \{\varphi\}$. For the converse, let $\alpha_1, \ldots, \alpha_n$ be a proof of ψ from $T \cup \{\varphi\}$. Transform this sequence as follows. First replace each α_i by $\varphi \to \alpha_i$. Then, suppose inductively that the steps up to $\varphi \to \alpha_i$ have been justified, and consider $\varphi \to \alpha_{i+1}$. If α_{i+1} is an axiom or a member of T, then since $\alpha_{i+1} \to (\varphi \to \alpha_{i+1})$ is an axiom, we can add it to the sequence, and then applying modus ponens we get $\varphi \to \alpha_{i+1}$ as needed. If α_{i+1} is φ, then $\varphi \to \alpha_{i+1}$ is $\varphi \to \varphi$, so we replace it by its proof from no assumptions (see Example 1.6). If finally α_{i+1} is a result of applying modus ponens to previous steps, then for some $\ell, k < i+1$ we have $\alpha_k = \alpha_\ell \to \alpha_{i+1}$. Applying A2 we get $(\varphi \to (\alpha_\ell \to \alpha_{i+1})) \to ((\varphi \to \alpha_\ell) \to (\varphi \to \alpha_{i+1}))$, and by the inductive assumption $\varphi \to (\alpha_\ell \to \alpha_{i+1})$, which is $\varphi \to (\alpha_k$, and $\varphi \to \alpha_\ell$ have already been justified. Then, applying modus ponens twice, we get the required $\varphi \to \alpha_{i+1}$. This procedure produces sequence which satisfies all the conditions of being a proof from T, and whose last term is $\varphi \to \psi$.

For (2), we will use some theorems on the cheat list. This makes life much easier. For the forward direction, suppose $\chi \wedge \neg\chi$ is provable from $T \cup \{\varphi\}$. Since $(\chi \wedge \neg\chi) \to \neg\varphi$ is a theorem (an instance of EX), we get that $\neg\varphi$ is provable from $T \cup \{\varphi\}$. Then, by (1) we get that $T \vdash \varphi \to \neg\varphi$ and applying N4 we obtain $T \vdash \neg\varphi$ as required. For the backward direction, using N2 we get that $T \vdash \varphi \to (\chi \wedge \neg\chi)$ which by (1) gives $T \cup \{\varphi\} \vdash \chi \wedge \neg\chi$.

The argument for (3) is similar. We leave it as an exercise. $\square$

Theorem 1.3: Proof by cases

Let $T \cup \{\varphi, \psi, \chi\}$ be a set of formulae. The following are equivalent:

(1) $T \cup \{\varphi\} \vdash \chi$ and $T \cup \{\psi\} \vdash \chi$
(2) $T \cup \{\varphi \lor \psi\} \vdash \chi$.

Proof. We will use Theorem 1.2(1) implicitly. For the downward direction, we have $T \vdash \varphi \to \chi$ and $T \vdash \psi \to \chi$. Using D3, we get $T \vdash (\varphi \lor \psi) \to \chi$, so we have $T \cup \{\varphi \lor \psi\} \vdash \chi$ as needed.

For the upward direction, assuming $T \vdash (\varphi \lor \psi) \to \chi$ we get $T \vdash (\varphi \lor \psi) \to \chi$. Now, using D1 and D2, we obtain that $T \cup \{\varphi\} \vdash \varphi \lor \psi$ and $T \cup \{\psi\} \vdash \varphi \lor \psi$. Therefore $T \cup \{\varphi\} \vdash \chi$ and $T \cup \{\psi\} \vdash \chi$ as needed. $\square$

Theorem 1.4: Proof by arbitrary constant

Let $\varphi(v)$ be a formula with exactly one free variable v, in some first-order language $\mathcal{L}$, and T be a set of formulae of $\mathcal{L}$. Let $\mathcal{L}^c$ be the expansion of $\mathcal{L}$ by a new constant symbol c. The following are equivalent:

(1) $T \vdash \varphi(c)$
(2) $T \vdash \forall v \colon \varphi(v)$.

Proof. For the upward direction, observe that c is free for v in φ, so by A3 we get $T \vdash \varphi(c)$. For the downward direction, let P be a proof of $\varphi(c)$ from T. We construct P' as follows. For any formula $\psi(c)$, in which c occurs, we replace $\psi(c)$ by $\forall x \colon \psi(x)$, where x is a 'fresh' variable that does not occur anywhere previously in P'. We claim that P' is still a proof. We check P' inductively, step by step. Since $\psi(c)$ cannot be a member of T, we have two possible cases.

- If $\psi(c)$ is an axiom, then so is $\psi(x)$ and therefore, so is $\forall x \colon \psi(x)$; thus axioms remain axioms.
- If $\psi(c)$ arises an an application of modus ponens, then $\chi \to \psi(c)$ and χ occur in P before $\psi(c)$. By inductive hypothesis, $\forall x \colon \chi \to \psi(x)$ is justified. Since x is fresh, it does not appear free in χ. Thus, we can apply A6 to obtain $\chi \to \forall x \colon \psi(x)$.

Now, P' is a proof whose last formula is is $\forall x \colon \varphi(x)$. Renaming the bound variables throughout we get $\forall v \colon \varphi(v)$, as required. $\square$

3. Notation

Having completed a very formal description of first order syntax, we can now relax our strict rules for notation to bring it more into line with everyday mathematical practice. In particular, we let:

- $\forall v_1, \ldots, v_n$ abbreviate $(\forall v_1) \ldots (\forall v_n)$
- $\exists v_1, \ldots, v_n$ abbreviate $(\exists v_1) \ldots (\exists v_n)$
- $\&$ be used for logical 'and', especially if $\wedge$ is used for lattice meet in the same context.
- $|$ be used for logical 'or', especially if $\vee$ is used for lattice join in the same context.
- $\Rightarrow$ be used for logical implication, especially if $\rightarrow$ is used for an algebraic operation of residuation (as in Heyting or Boolean algebras) in the same context.
- $\sim$ be used for logical negation, especially if $\neg$ is used for a Boolean complement operation in the same context.
- colons, square brackets, braces and similar devices will be used to enhance readability of formulae.

Example 1.7: Fields

Here are the axioms of field theory, in our shorthand:

- $\forall x\colon x + 0 = x \;\&\; x = 0 + x$
- $\forall x\colon x + (-x) = 0 \;\&\; (-x) + x = 0$
- $\forall x, y, z\colon x + (y + z) = (x + y) + z$
- $\forall x, y\colon x + y = y + x$
- $\forall x\colon x \cdot 0 = 0 \;\&\; 0 = 0 \cdot x$
- $\forall x\colon x \neq 0 \Rightarrow (x \cdot 1 = x \;\&\; x = 1 \cdot x)$
- $\forall x\; \exists y\colon x \neq 0 \Rightarrow (x \cdot y = 1 \;\&\; y \cdot x = 1)$
- $\forall x, y, z\colon x \cdot (y \cdot z) = (x \cdot y) \cdot z$
- $\forall x, y\colon x \cdot y = y \cdot x$
- $\forall x, y, z\colon x \cdot (y + z) = (x \cdot y) + (x \cdot z)$
- $0 \neq 1$

CHAPTER 2

Semantics: models, interpretations, satisfaction

First order languages are interpreted in (speak about) *models*, also called *structures* or *relational structures*. A structure $\mathcal{M}$ is:

$$\mathcal{M} = \langle M, (R_i)_{i \in I}, (f_j)_{j \in J}, (c_k)_{k \in K} \rangle$$

where

(1) M is a nonempty set (the *universe* of the model).
(2) $R_i \subseteq M^n$ is an n-ary relation, for each $i \in I$.
(3) $f_j \colon M^m \to M$ is an m-ary function (an *operation*), for each $j \in J$.
(4) c_k is an element of M (a *constant*), for each $k \in K$.

It is assumed that among the relations there is a distinguished binary relation $=$ of equality. A *signature* of a structure $\mathcal{M}$ is an enumeration of its relations, functions and constants, together with their respective arities, just as in the case of languages. Two structures of the same type are called *similar*, again as in the case of languages. A structure without relations is an *algebra*, in the general algebra sense.

1. Interpretations and valuations

Languages and formulae are interpreted in structures. Which is why structures are called models: they model formulae. Some formulae turn out to be satified, some not. Intuitively we all know what that means: something like what we mean when we say that Abelian groups satisfy the equation $x + y = y + x$. We will make it precise.

Let $\mathcal{M}$ be a structure, and I a function (called *interpretation*), which assigns:

- to each constant symbol c, an element of M,
- to each function symbol f of arity n, a total function from M^n into M,
- to each relation symbol R of arity n, a subset of M^n (that is, an n-ary relation over M).

We commonly leave I out of the notation, because in an overwhelming majority of cases, I is clear from context. So, we write $c^{\mathcal{M}}$ instead of

$I(c)$ for a constant symbol c, and $f^{\mathcal{M}}$ instead of $I(f)$ for a function symbol f, and finally $R^{\mathcal{M}}$ instead of $I(R)$, for a relation symbol R.

A *valuation* in $\mathcal{M}$ is a function ν assigning to each variable x a member of M, and then extended to compound terms by: $\nu(f(t_1 \dots t_n)) = f^{\mathcal{M}}(\nu(t_1) \dots \nu(t_n))$. We will typically use the Greek ν for a valuation, just because we often use v for a variable.

We will illustrate the roles of interpretation and valuation in the example below. The example is admittedly silly, but it brings out some important points. In particular, it emphasises the fact that uninterpreted language is just a game of symbols.

Example 2.1: Interpretation and valuation

An anthropologist studies a newly discovered tribe. She focuses on their understanding of numbers and the language of arithmetic. To her surprise, she finds out that the tribe uses a kind of symbolic notation. Probing harder, she learns that:

- $\star$ is used as a unary function,
- $\times$ and $\diamond$ are used as binary functions,
- $/$ is used as a binary relation.

She finally finds and old sage who speaks English and she learns further that they write $\star$ in front of a number to indicate its negative, $\times$ in front of two numbers for addition, and $\diamond$ in front of two numbers for multiplication. Also, $/$ written between two numbers means that the first is smaller than the second. Here is an arithmetical inscription in their language:

$$\star\mathrm{I} \mathbin{/} \times\mathrm{II} \diamond \star\mathrm{III}\ \mathrm{I}$$

Puzzled by I, II and III, she asks, and hears: "Oh, these are just placeholders for numbers. We use them so that we knew which must be the same and which needn't be." To discover what the inscription means, our anthropologist must first *interpret* the symbols in the structure $(\mathbb{Z}; \leq, +, \cdot, -)$ of the integers with the usual operations and order. Her interpretation function is: $\star^{\mathbb{Z}} = -$, $\cdot^{\mathbb{Z}} = +$, $\diamond^{\mathbb{Z}} = \times$, $/^{\mathbb{Z}} = \leq$. Next, to make better sense of the inscription, she *evaluates* the placeholdes. Attracted by their uncanny resemblance to Roman numerals, she takes her valuation function to be: $\nu(\mathrm{I}) = 1$, $\nu(\mathrm{II}) = 2$, $\nu(\mathrm{III}) = 3$. Next she calculates the right-hand side of /:

$$\begin{aligned}\nu(\times\mathrm{II}\diamond\star\mathrm{III}\ \mathrm{I}) &= +(\nu(\mathrm{II}),\nu(\diamond\star\mathrm{III}\ \mathrm{I}))\\ &= +(2,\cdot(\nu(\star\mathrm{III}),\nu(\mathrm{I})))\\ &= +(2,\cdot(-\nu(\mathrm{III}),\nu(\mathrm{I})))\\ &= +(2,\cdot(-3,1))\\ &= 2+(-3\cdot 1)\\ &= -1\end{aligned}$$

writing $+$ and $\cdot$ in prefix notation for the first few lines, just to conform to the local usage. Then, the left-hand side:

$$\begin{aligned}\nu(\star(\mathrm{I})) &= -\nu(\mathrm{I})\\ &= -1\end{aligned}$$

To proceed, she needs to learn about satisfaction and truth, so we leave her for the time being...

2. Satisfaction and truth

We say that a structure $\mathcal{M}$ *satisfies* a formula $\varphi(x_1,\ldots,x_m)$ under a valuation ν, written $\mathcal{M}\models_\nu \varphi(x_1,\ldots,x_m)$, or that a valuation ν satisfies $\varphi(x_1,\ldots,x_m)$ in $\mathcal{M}$. Satisfaction if defined by induction on complexity of $\varphi(x_1,\ldots,x_m)$.

Definition 2.1: Satisfaction

Base case.
If $\varphi(x_1,\ldots,x_m)$ is atomic, that is, of the form $R(t_1\ldots t_n)$ where $t_1,\ldots,t_n$ are terms in some variables from $\{x_1,\ldots,x_m\}$, then $\mathcal{M}\models_\nu R(t_1\ldots t_n)$ if $R^{\mathcal{M}}(\nu(t_1^{\mathcal{M}})\ldots\nu(t_n^{\mathcal{M}}))$ holds in $\mathcal{M}$.
Inductive step.

- $\mathcal{M}\models_\nu \chi\to\psi$ if either $\mathcal{M}\not\models_\nu \chi$ or $\mathcal{M}\models_\nu \psi$.
- $\mathcal{M}\models_\nu \neg\psi$ if $\mathcal{M}\not\models_\nu \psi$,
- $\mathcal{M}\models_\nu \forall x\ \psi$ iff for all valuations μ that differ from ν only at x, we have $\mathcal{M}\models_\mu \psi$.

Although our official logical vocabulary contains only implication, negation and universal quantifier, it is convenient to have satisfaction defined for the other connectives as well.

Lemma 2.1

Let $\mathcal{M}$ be a structure and ν a valuation in $\mathcal{M}$. Then,

(1) $\mathcal{M} \models_\nu \varphi \wedge \psi$ iff $\mathcal{M} \models_\nu \varphi$ and $\mathcal{M} \models_\nu \psi$,
(2) $\mathcal{M} \models_\nu \varphi \vee \psi$ iff $\mathcal{M} \models_\nu \varphi$ or $\mathcal{M} \models_\nu \psi$,
(3) $\mathcal{M} \models_\nu \exists v\ \varphi$ iff for some valuation μ that differs from ν only at v, we have $\mathcal{M} \models_\mu \varphi$.

Proof. For (1), the offical version of $\varphi \wedge \psi$ is $\neg(\varphi \to \neg\psi)$. We have

$$\begin{aligned} \mathcal{M} \models_\nu \neg(\varphi \to \neg\psi) &\text{ iff } \mathcal{M} \not\models_\nu \varphi \to \neg\psi \\ &\text{ iff } \mathcal{M} \models_\nu \varphi \text{ and } \mathcal{M} \not\models_\nu \neg\psi \\ &\text{ iff } \mathcal{M} \models_\nu \varphi \text{ and } \mathcal{M} \models_\nu \psi \end{aligned}$$

We leave (2) and (3) as an exercise. □

A common notational variant, which we will be using quite often, makes use of the fact that valuations in a structure $\mathcal{M}$ can be identified with sequences $(\nu(v_n) \colon n \in \omega)$ of elements of M, and further, given any formula φ, only a finite number of variables, say, $x_1, \ldots, x_n$ occur in φ. So, instead of specifying ν and writing $\mathcal{M} \models_\nu \varphi$, we can write $\mathcal{M} \models \varphi[a_1, \ldots, a_n]$, with the implied understanding that φ is $\varphi(x_1, \ldots, x_n)$ and $\nu(x_1) = a_1, \ldots, \nu(x_n) = a_n$. Alternatively, we can also write $\varphi^{\mathcal{M}}(a_1, \ldots, a_n)$; this notation resembles the one used for *polynomials* in general algebra. Or indeed, the usual polynomials, for if $\mathcal{M}$ is a field and φ is $\exists x \colon y_n x^n + \cdots + y_1 x + y_0 = 0$, then $\mathcal{M} \models \varphi[a_0, \ldots, a_n]$ is commonly written as $\exists x \colon a_n x^n + \cdots + a_1 x + a_0 = 0$.

Finally, we come to the definition of truth.

Definition 2.2: Truth

Let $\mathcal{M}$ be a structure, φ a formula and T a set of formulae. Then,

- φ is *true* in $\mathcal{M}$ or *verified* by $\mathcal{M}$ iff it is satisfied by all valuations. Notation: $\mathcal{M} \models \varphi$ means 'φ is true in $\mathcal{M}$', and $\mathcal{M} \models T$ means 'τ is true in $\mathcal{M}$ for every $\tau \in T$'.
- $T \models \varphi$ iff every structure verifying every $\tau \in T$ also verifies φ, that is, if $\mathcal{M} \models T$ implies $\mathcal{M} \models \varphi$, for every structure $\mathcal{M}$.

Example 2.2: Satisfaction and truth

The anthropologist story from Example 2.1 continues. Having evaluated the terms on the sides of

$$\star\mathrm{I} \,/\, \times\mathrm{II} \diamond \star\mathrm{III}\ \mathrm{I}$$

under v, she now asks herself if the formula inscribed above is satisfied in $\mathbb{Z}$ under v. This is an atomic formula (a binary relation between two terms), so by definition $\mathbb{Z} \models_v \star\mathrm{I} \,/\, \times\mathrm{II} \diamond \star\mathrm{III}\ \mathrm{I}$ if and only if $v(\star\mathrm{I}) /^{\mathbb{Z}} v(\times\mathrm{II} \diamond \star\mathrm{III}\ \mathrm{I})$ holds in $\mathbb{Z}$. She already knows (see Example 2.1) that $v(\star\mathrm{I}) = -1$ and $v(\times\mathrm{II} \diamond \star\mathrm{III}\ \mathrm{I}) = -3$. She also knows that $/^{\mathbb{Z}}$ is $\leq$. So $\mathbb{Z} \models_v \star\mathrm{I} \,/\, \times\mathrm{II} \diamond \star\mathrm{III}\ \mathrm{I}$ if and only if $-1 \leq -3$. Thus, our formula is not satified by v in $\mathbb{Z}$.

Example 2.3: Satisfaction, truth and quantifiers

The anthropologist's saga continues. Now she overhears the old sage say: *Aye, for all numbers II and III, one can find a number I such that*

$$\star I \,/\, \times II \diamond \star III\ I$$

holds true. She is now a little tired of the alien notation, so she quickly translates it into:

$$\mathbb{Z} \models \forall x, y\ \exists z\colon -z \leq x + (-y \cdot z)$$

Now, 'true' means 'satisfied under every valuation' so in particular the valuation $v(y) = 1$, $v(x) = -1$, $v(z) = 0$ should do. But for this valuation, sadly, she is unable to find any v' differing from v only on z and such that $-v'(z) \leq -1 + (-1 \cdot v'(z))$. Even old sages are wrong sometimes.

Example 2.4: Some familiar structures

- $\mathcal{L} = (L; \wedge, \vee)$, where $\wedge$ and $\vee$ are binary operations, such that the following sentences are true in $\mathcal{L}$:
  - $\forall x\colon x \wedge x = x$
  - $\forall x\colon x \vee x = x$
  - $\forall x, y\colon x \wedge y = y \wedge x$
  - $\forall x, y\colon x \vee y = y \vee x$
  - $\forall x, y, z\colon x \wedge (y \wedge z) = (x \wedge y) \wedge z$
  - $\forall x, y, z\colon x \vee (y \vee z) = (x \vee y) \vee z$
  - $\forall x, y\colon x \wedge (x \vee y) = x$

  - $\forall x, y \colon x \vee (x \wedge y) = x$

  That is, $\mathcal{L}$ is any lattice. All these sentences are universally quantified equations; in such situations, we typically write them more succintly, as in the next example.
- $\mathcal{G} = \langle G, \cdot, ^{-1}, e\rangle$, where $\cdot$, $^{-1}$ and e are, respectively, a binary operation, a unary operation and a constant, such that the following hold
  - $(x \cdot y) \cdot z = x \cdot (y \cdot z)$
  - $x \cdot x^{-1} = e = x^{-1} \cdot x$
  - $x \cdot e = x = e \cdot x$

  for all $x, y, z \in G$. That is, $\mathcal{G}$ is any group.
- $\mathcal{G}_2 = \langle G, \cdot, ^{-1}, e\rangle$, satisfying all properties above, together with: (a) $\exists x \colon x \neq e$ and (b) $\forall x, y \colon x \neq y \Rightarrow x = e$ or $y = e$. That is, $\mathcal{G}_2$ is any two-element group.
- $\mathcal{B}_2 = \langle \{0, 1\}, \cdot, ^{-1}, e\rangle$, where $\cdot$ is addition mod 2, $^{-1}$ is the identity function, and e is 0. That is, $\mathcal{B}_2$ is a two-element group $\mathbb{Z}_2$.
- $\mathcal{D} = \langle D, \leq\rangle$, such that $\leq$ is a dense linear ordering without endpoints.
- $\mathcal{Q} = \langle \mathbb{Q}, \leq\rangle$, the rationals with their natural ordering.
- $\mathcal{N} = \langle \mathbb{N}_0, \leq, +, \times, 0, 1\rangle$, the standard model of arithmetic.

Notice that the equality relation is present in all these structures. Following custom, we do not mention it explicitly. Structures $\mathcal{G}_2$ and $\mathcal{B}_2$ illustrate two common ways of specifying models: one by listing some formulae that we want our model to satisfy, the other by giving the model explicitly. The same goes for $\mathcal{Q}$ and $\mathcal{D}$. Refer to Example 1.3 for formulae defining $\mathcal{D}$.

Example 2.5: Duality principle for lattices

Consider the language with two binary functions: $\times$ and $+$. Let $\mathcal{L} = (L; \wedge, \vee)$ be any lattice. Consider two interpretations: $I_1(\times) = \wedge$, $I_1(+) = \vee$, and $I_2(\times) = \vee$, $I_2(+) = \wedge$. Then, all axioms of lattices are true in $(L; I_1(\times), I_1(+))$ *and* in $(L; I_2(\times), I_2(+))$. In plain words: swapping meet and join in a lattice gives another lattice. This is known as *duality principle* for lattices.

CHAPTER 3

Completeness of first order logic

Let $T \cup \{\varphi\}$ be a set of formulae. If $T \vdash \varphi$ (there is a proof of φ from T), we call φ a *syntactic* consequence of T. If $T \models \varphi$ (every model of T is a model of φ), we call φ a *semantic* consequence of T. Are the two the same?

Theorem 3.1: Completeness theorem, Gödel-Maltsev

$T \vdash \varphi$ iff $T \models \varphi$

Before we sketch a proof of this theorem, we reformulate it in a way suitable to the method of proof we intend to use. We call a set T of formulae *satisfiable* if for some structure $\mathcal{M}$ and some valuation ν in $\mathcal{M}$ we have $\mathcal{M} \models_\nu T$. So, T is satisfiable if it has a model. If T is not satisfiable, it is called *unsatisfiable*. We call T *consistent* if there is no formula α such that $T \vdash \alpha \wedge \neg\alpha$. So, T is consistent if there is no proof of contradiction from T. If T is not consistent, it is called *inconsistent*.

Lemma 3.1

For any set of formulae T and a formula φ the following hold:

(1) $T \vdash \varphi$ iff $T \cup \{\neg\varphi\}$ is inconsistent.
(2) $T \models \varphi$ iff $T \cup \{\neg\varphi\}$ is unsatisfiable.

Proof. We will sketch a proof for the first equivalence. Suppose $T \vdash \varphi$. Let $\alpha_1, \dots, \alpha_n$ be a proof of φ from T. Then, $\alpha_1, \dots, \alpha_n$ is also a proof of φ from $T \cup \{\neg\varphi\}$, so $T \cup \{\neg\varphi\} \vdash \varphi$ holds. Trivially, $T \cup \{\neg\varphi\} \vdash \neg\varphi$ holds. It is quite a nontrivial exercise (given our parsimonious axiom system) to show that in this case $T \cup \{\neg\varphi\} \vdash \varphi \wedge \neg\varphi$ holds as well. Conversely, suppose $T \cup \{\neg\varphi\} \vdash \alpha \wedge \neg\alpha$. Then, we can show that $T \cup \{\neg\varphi\} \vdash \varphi$, using (a) the fact that $\alpha \wedge \neg\alpha \vdash \beta$ for any β and (b) the fact that $\vdash$ is transitive. We can show further that then $T \vdash \neg\varphi \rightarrow \varphi$, and from this, applying some more formal proving, we get $T \vdash \varphi$ (the latter is not too surprising given that shorthand for $\neg\varphi \rightarrow \varphi$ is $\varphi \vee \varphi$).

For the second equivalence we will give a more complete proof. Assume $T \models \varphi$. Suppose for contradiction that $T \cup \{\neg\varphi\}$ has a model, that is, for some satructure $\mathcal{M}$ and some valuation ν in $\mathcal{M}$ we have $\mathcal{M} \models_\nu T$ and $\mathcal{M} \models_\nu \neg\varphi$. By assumption then it follows that $\mathcal{M} \models_\nu \varphi$. So, by definition of satisfaction, we have $\mathcal{M} \models_\nu \varphi$ and $\mathcal{M} \not\models_\nu \varphi$. A contradiction. Conversely, suppose $T \cup \{\neg\varphi\}$ is unsatisfiable. Then, for every structure $\mathcal{M}$ and every valuation ν in $\mathcal{M}$ we have either $\mathcal{M} \not\models_\nu \tau$ for some $\tau \in T$, or $\mathcal{M} \not\models_\nu \neg\varphi$ (that is $\mathcal{M} \models_\nu \varphi$). This is equivalent to the following. For every structure $\mathcal{M}$ and every valuation ν in $\mathcal{M}$, if $\mathcal{M} \models_\nu \tau$ for every $\tau \in T$, then $\mathcal{M} \models_\nu \varphi$. Since this holds for every valution ν, we get that $\mathcal{M} \models T$ implies $\mathcal{M} \models \varphi$, that is, $T \models \varphi$ as required. $\square$

Theorem 3.2: Completeness theorem, equivalent version

T is satisfiable iff T is consistent.

1. Henkin method

The difficult part in proving completeness is to show that if T is consistent, then it has a model. We describe a method of proof due in this form to Leon Henkin. We call a set S of formulae a *Henkin set*, if

- S is maximally consistent (consistent and such that for every formula φ, either $\varphi \in S$ or $\neg\varphi \in S$)
- $(\exists v)\ \varphi \in S$ implies $\varphi(c) \in S$, for some constant c.

So, the constant c serves as a witness for the existential quantifier in $(\exists v)\ \varphi$. It is called a *witnessing constant* or just a *witness*.

Lemma 3.2

Let S be a Henkin set. Then, S induces an equivalence relation on the set Term of all terms, by $t \sim s$ iff $t = s \in S$. Moreover, $\sim$ has the following compatibility property[a]:

- if $t_1 \sim s_1, \ldots, t_n \sim s_n$, then $f(t_1, \ldots t_n) \sim f(s_1, \ldots s_n)$, for any n-ary function f.
- if $t_1 \sim s_1, \ldots, t_n \sim s_n$, then $R(t_1, \ldots, t_n) \in S$ iff $R(s_1, \ldots, s_n) \in S$, for any n-ary relation R.

[a] An equivalence relation with the compatibility property is called a *congruence* in algebra.

PROOF. Since the pure axioms of equality (E1–E5) are quantifier free, any term t is free for any variable occurring in these axioms. Then, by several applications of A4 we get that

- $t = t$
- $t = s \to s = t$
- $t = s \wedge s = r \to t = r$
- $t_1 = s_1 \wedge \cdots \wedge t_n = s_n \to f(t_1, \ldots, t_n) = f(s_1, \ldots, s_n)$ for any function symbol f
- $t_1 = s_1 \wedge \cdots \wedge t_n = s_n \to R(t_1, \ldots, t_n) \leftrightarrow R(s_1, \ldots, s_n)$ for any relation symbol R.

are logical theorems, for any terms $t, t_1, \ldots, t_n, s, s_1, \ldots, s_n r$. It follows that $t = t \in S$ for any term t, because S is consistent, so $\sim$ is reflexive. If $t = s \in S$, then since $t = s \to s = t$ is a logical theorem, and so belongs to S, we have $S \vdash s = t$. If $s = t \notin S$, then by maximality $s \neq t \in S$, and so S would be inconsistent. Therefore, $s = t \in S$, showing that $\sim$ is symmetric.

By similar reasoning, we can show that $\sim$ is transitive and has the compatibility property. □

Now, define the universe, functions and relations:

- $H = \mathsf{Term}/{\sim}$.
- $f^U([t_1], \ldots, [t_n]) = [f(t_1, \ldots, t_n)]$.
- $([t_1], \ldots, [t_n]) \in R^U$ iff $R(t_1, \ldots, t_n) \in S$.

The definitions above are correct: they do not depend of the choice of representatives of equivalence classes, by Lemma 3.2. Next, define a valuation ν:

- $\nu(v) = [v]$.

Lemma 3.3

Let $\mathcal{H}$ and ν be the structure and the valuation defined above. Then $\mathcal{H} \models_\nu \varphi$ if and only if $\varphi \in S$.

PROOF. Induction on complexity of φ. If φ is atomic, then the claim follows immediately from the definitions. If φ is $\neg\psi$, then $\varphi \in S$ iff $\psi \notin S$, by maximal consistency of S. Then, by inductive hypothesis $\mathcal{H} \not\models_\nu \psi$ iff $\psi \in S$. Therefore, $\mathcal{H} \models_\nu \neg\psi$ iff $\neg\psi \in S$, as claimed. If φ is $\alpha \to \beta$, a similar argument works.

It remains to deal with the quantifier. We will use the 'unofficial' existential quantifier in the proof; the proof for the 'official' universal quantifier can be obtained using the logical theorem $(\forall v)\ \varphi \leftrightarrow \neg(\exists v)\ \neg\varphi$. Suppose φ is $(\exists v)\ \psi$. If $\varphi \in S$, then $\psi(c) \in S$ for some constant c, because S is a Henkin set. The inductive hypothesis applies to $\psi(c)$, so we get that $\mathcal{H} \models_\nu \psi(c)$. By definition of satisfaction, then $\mathcal{H} \models_\nu (\exists v)\ \psi$. Conversely, if $\mathcal{H} \models_\nu (\exists v)\ \psi$, then by definition of satsfaction again, $\mathcal{H} \models_\mu \psi(v)$, for some valuation μ differing from ν at most on v. Thus, $\mu(v) = [t]$ for some term t. Now, consider $\psi(t)$. By inductive hypothesis, $\psi(t) \in S$ iff $\mathcal{H} \models_\nu \psi(t)$, and notice that $\nu(t) = [t] = \mu(v)$. It follows that $\mathcal{H} \models_\nu \psi(t)$, so $\psi(t) \in S$. But $\psi(t) \to (\exists v)\ \psi$ is a logical theorem (again, it is a tedious exercise in using our axiom system: the relevant axiom is A4 and we can always choose[1] a variable v so that t is free for v in ψ), so $(\exists v)\ \psi \in S$ as required. □

Lemma 3.4

Every consistent set of formulae T of $\mathcal{L}$ can be extended to a Henkin set T^* in a language $\mathcal{L}^* = \mathcal{L} \cup C$, where C a countable set of new constants.

PROOF. An inductive construction iterated inductively. We will present a reasonably detailed proof sketch. Assume for simplicity that the language is countable. Also, pretend (contrary to our official setup) that all quantifiers are existential (this is harmless, in view of $\exists = \neg\forall\neg$). Now, order all formulae of $\mathcal{L}$ into a sequence, and consider them one by one. Put $T_0^0 = T$.

(1) If $T_n^0 \cup \{\varphi_{n+1}\}$ is inconsistent, then $T_{n+1}^0 = T_n^0$.
(2) If $T_n^0 \cup \{\varphi_{n+1}\}$ is consistent, and φ_{n+1} is not of the form $(\exists v)\ \psi$, then $T_{n+1}^0 = T_n^0 \cup \{\varphi_{n+1}\}$.
(3) If $T_n^0 \cup \{\varphi_{n+1}\}$ is consistent, and φ_{n+1} is of the form $(\exists v)\ \psi$, then $T_{n+1}^0 = T_n^0 \cup \{\varphi_{n+1}, \psi(v/c)\}$, for some new constant c.
(4) $T^1 = \bigcup_{n\in\omega} T_n^0$.

It is clear that steps (1) and (2) produce a T_{n+1}^0 that is consistent. For step (3) observe that since T_n^0 is finite at each stage and there are infinitely many constants, a suitable $\psi(v/c)$ with a fresh constant c always exists. Now, suppose $T_n^0 \cup \{\varphi_{n+1}, \psi(v/c)\}$ is inconsistent, so there is a proof of contradiction from it. Thus, by Theorem 1.2(2), $T_n^0 \cup \{\varphi_{n+1}\} \vdash \neg\psi(v/c)$ must hold. Observing that c, being a constant, is free for v in ψ and applying A4, we get that $\neg(\exists v)\ \psi \to \neg\psi(v/c)$

[1] In fact, to be completely justified in doing so, we would need to show that the formulae $(\exists v)\ \psi$ and $(\exists v')\ \psi'$ are logically equivalent, where ψ' results from ψ by replacing all occurrences of v by occurrences of v'. This procedure is called 'renaming bound variables', and we do it nearly subconsciously all the time. Just think how many times you have heard or said "there is an x, ooops, I've used x already, let's call it y".

is a logical theorem. Thus, since $\neg(\exists v)\ \psi$ is $\neg\varphi_{n+1}$, we have $T_n^0 \cup \{\neg\varphi_{n+1}\} \vdash \neg\psi(v/c)$ as well. Applying Theorem 1.3, we obtain $T_n^0 \cup \{\varphi_{n+1} \vee \neg\varphi_{n+1}\} \vdash \neg\psi(v/c)$. But $\varphi_{n+1} \vee \neg\varphi_{n+1}$ is a logical theorem, so any proof from $T_n^0 \cup \{\varphi_{n+1} \vee \neg\varphi_{n+1}\}$ is also a proof from T_n^0 alone. Therefore, $T_n^0 \vdash \neg\psi(v/c)$. Since the constant c was not in the original language, Theorem 1.4 applies, giving $T_n^0 \vdash (\forall v)\ \neg\psi(v)$. But then, $T_n^0 \vdash \neg(\exists v)\ \psi(v)$ and so $T_n^0 \cup \{(\exists v)\ \psi(v)\}$, that is, $T_n^0 \cup \{\varphi_{n+1}\}$ is inconsistent, contradicting the assumption.

It follows (by compactness) that T^1 is consistent. But notice that we have only dealt with the existential formulae in the original language $\mathcal{L}$. By adding witnessing constants and new formulae, we have created new existential formulae to take care of. For example, if $(\exists x)(\exists y)\ \varphi(x,y)$ was in $\mathcal{L}$, our "first pass" would add the formula $(\exists y)\ \varphi(c,y)$ which is existential, but does not belong to $\mathcal{L}$. Let $\mathcal{L}^1$ be the expanded language. Thus, T^0 is a consistent theory in $\mathcal{L}^1$. Then, repeat the previous procedure with respect to $\mathcal{L}^1$: list all the formulae, put $T_0^1 = T^1$, and expand adding witnessing constants where necessary. The result of the "second pass" is a consistent theory T^2. We have again created new existential formulae, so we repeat the procedure.

Finally, we put $T^* = \bigcup_{n\in\omega} T^n$. This is a consistent theory (again, by compactness) in the language $\mathcal{L}^* = \bigcup_{n\in\omega} \mathcal{L}^n$, and obviously extends T. Notice that $\mathcal{L}^*$ is a countable union of countable sets, hence countable. Moreover, T^* is now maximal, since every formula of $\mathcal{L}^*$ belongs to some $\mathcal{L}^n$ and therefore has been considered. The existence of witnessing constants in T^* is immediate by the construction. $\square$

Henkin's construction is by far the most standard method of proving completeness theorem for first-order logic. It can be adapted for uncountable languages: there are uncountably many formulae then, so one needs transfinite induction to proceed but essentially the method is unchanged (except that it needs the axiom of choice, whereas the countable version needs something slightly weaker).

We are now ready to prove the completeness theorem.

Proof of Theorem 3.1. For the easy direction, suppose that T is satisfiable, that is $\mathcal{M} \models_\nu T$, for some structure $\mathcal{M}$ and valuation ν. It is easy to verify that all logical axioms are true in every structure, and applications of *modus ponens* preserve satisfaction, that is, if $\mathcal{M} \models_\nu \varphi \to \psi$ and $\mathcal{M} \models_\nu \varphi$, then $\mathcal{M} \models_\nu \psi$. Thus, if there were a proof of $\alpha \wedge \neg\alpha$ from T, then we would have $\mathcal{M} \models_\nu \alpha \wedge \neg\alpha$, which quickly yields a contradiction with the definition of satisfaction.

For the hard direction, suppose T is consistent. Then, by Lemma 3.4, T can be extended to a Henkin set T^*. By Lemmas 3.2 and 3.3, there is a structure $\mathcal{H}$ such that $\mathcal{H} \models T^*$, so $\mathcal{H} \models T$ as well. $\square$

2. Corollaries of completeness

2.1. Compactness. The following theorem is the semantic counterpart of Theorem 1.1. It is the commonest form of *Compactness Theorem* for first order logic.

Theorem 3.3: Semantic compactness

Let T be a set of formulae. If every finite subset of T has a model, then T has a model.

Proof. Suppose every finite subset of T has a model, but T has no model. By Theorem 3.2, T is inconsistent, so $T \vdash \chi \wedge \neg\chi$ for some χ. Then, by Theorem 1.1 we have $T_0 \vdash \chi \wedge \neg\chi$ for some finite $T_0 \subseteq T$. Then, T_0 is inconsistent, so it has no model, contradicting the assumption. $\square$

Definition 3.1: Stricly elementary class

Let T be a theory. We write $\mathrm{Mod}(T)$ for the class of structures $\mathcal{C}$ such that $\mathcal{C} \models T$. If a class $\mathbb{C}$ is equal to $\mathrm{Mod}(\{\varphi\})$ for a single formula φ, then $\mathbb{C}$ is called a *strictly elementary class*.

Lemma 3.5

Let ψ be a first-order sentence in pure equality language. If ψ holds in all infinite pure equality structures (sets), then there is an $n \in \mathbb{N}$ such that ψ holds in all sets S with $|S| > n$.

Proof. Suppose ψ holds in no finite sets, that is, $\neg\psi$ holds in all finite sets. Let $\Sigma_n = \{\neg\psi, \neg\varphi_1, \ldots, \neg\varphi_n\}$, where φ_n is the sentence stating that there are at most n elements (exercise: write φ_n explicitly). Suppose each Σ_n has a model. This model is a set with at least $n+1$ elements. Put $\Sigma = \bigcup_{n\in\mathbb{N}} \Sigma_n$. By compactness, Σ has a model, say, $\mathcal{S}$. Then, for each $n \in \mathbb{N}$ we have that $\mathcal{S}$ has strictly more elements than n, so $\mathcal{S}$ is infinite. But, $\mathcal{S} \models \neg\psi$. Contradiction. $\square$

Corollary 3.1: Infinity is not strictly elementary

The class $\mathbb{INF}$ of infinite pure equality structures is not strictly elementary.

Lemma 3.5 and Corollary 3.1 may seem to be just curiosities. The next theorem s below

Theorem 3.4: Robinson's principle

Let φ be a first-order sentence in the language of fields. If φ holds in all fields of characteristic 0, then there is a prime p such that φ holds in all fields of characteristic $\geq p$.

Proof. Analoguous to proof of Lemma 3.5. Let Φ be the set of sentences from Example 1.7 (a first-order rendering of field axioms), and suppose there is an infinite sequence of primes $(p_i)_{i\in I}$, such that each set $\Sigma_i = \{\neg\varphi, \underbrace{1+1+\cdots+1 \neq 0}_{p_i \text{ times}}\} \cup \Phi$, has a model. By compactness, $\Sigma = \bigcup_{i\in I}\Sigma_i$ has a model. This model is a field of characteristic 0 satisfying $\neg\varphi$, so not satisfying φ. Contradiction. $\square$

A group $\mathbf{G}$ is a *torsion* group, if for every $g \in G$ we have $g^n = e$ for some n. A group $\mathbf{G}$ is a *torsion free* group, if for every $g \in G \setminus \{e\}$ we have $g^n \neq e$ for every n.

Theorem 3.5

The class $\mathbb{TF}$ of all torsion-free groups is not finitely axiomatisable. That is, there is no finite set S of first-order formulae, such that $\mathbf{G} \in \mathbb{TF}$ iff $\mathbf{G} \models S$.

Proof. Observe that $\mathbf{G}$ is torsion-free iff $\mathbf{G} \models \Sigma$, where $\Sigma = \{x \neq e \to x^n \neq e \colon n \in \mathbb{N}\}$. Suppose, towards a contradiction, that S is a finite set axiomatising $\mathbb{TF}$. Let $\sigma = \bigwedge S$. Then, $\Sigma \models \sigma$, that is, $\Sigma \cup \{\neg\sigma\}$ has no model. Therefore, $\Sigma_0 \cup \{\neg\sigma\}$ has no model, for some finite $\Sigma_0 \subseteq \Sigma$. Take a group $\mathbf{H}$ such that all elements of H are of order big enough to satisfy Σ_0. Contradiction. $\square$

2.2. Löwenheim-Skolem Theorems. The next theorem is known as *Lower Löwenheim-Skolem Theorem*, or *Downward Löwenheim-Skolem Theorem*. It is also sometimes called *Löwenheim-Skolem Paradox*, because it states that if a first order theory (in a countable language), has an infinite model, then it has a countable one. To see where the apparent paradoxicality lies apply this to Zermelo-Fraenkel set theory (ZF): Cantor theorem, provable in ZF implies that there are sets of arbitrarily large cardinalities, yet ZF, if it consistent, has a countable model. To see why it is not a real paradox, observe that cardinalities are measured by bijections and bijections are functions, that is, sets of pairs; hence, sets. Viewed 'from outside' they show that all sets in the model are countable. But, some of these bijections may not be elements of the model, and so 'from the point of view of the model' some of its elements may not be in a bijective correspondence with the set of natural numbers.

Theorem 3.6: Lower Löwenheim-Skolem Theorem

Let L be a first order language, and T a set of formulae of L. If T has an infinite model, then T has a model of cardinality $\max\{\omega, |L|\}$.

Proof. Extend T to a Henkin set T^*. If necessary, first expand L to L_C by adding a suitable set C of constant symbols to L. Since we only need as many constants as there are formulae of L, we have $|C| \leq \max\{\omega, |L|\}$. Perform Henkin construction on T^*. We obtain a model of T^* whose cardinality is infinite but no greater than the cardinality of the language. $\square$

A companion result is *Upper Löwenheim-Skolem Theorem*.

Theorem 3.7: Upper Löwenheim-Skolem Theorem

Let L be a first order language, and T a set of formulae of L. If T has an infinite model, then T has a model of any cardinality $\kappa \geq \max\{\omega, |L|\}$.

Proof. If T has an infinite model, then T has a model of cardinality $\max\{\omega, |L|\}$, by Theorem 3.6. For a cardinal $\kappa \geq \max\{\omega, |L|\}$, expand the language by adding κ new constants. Then, expand T to T^+ by adding the sentence $c \neq d$ for each pair of distinct new constants. Carry out Henkin construction and get a model of cardinality κ. $\square$

2.3. Three theorems on provability. Compactness and Löwenheim-Skolem theorems are examples of syntax helping in establishing an essentially semantic result. We will now go the other way, and give semantic proofs of Theorems 1.2, 1.3 and 1.4.

Theorem 3.8: Deduction Theorem

Let $T \cup \{\varphi, \psi\}$ be a set of formulae. The following equivalences hold:

(1) $T \cup \{\varphi\} \models \psi$ iff $T \models \varphi \to \psi$
(2) $T \cup \{\varphi\} \models \chi \wedge \neg\chi$ for some formula χ iff $T \models \neg\varphi$
(3) $T \cup \{\neg\varphi\} \models \chi \wedge \neg\chi$ for some formula χ iff $T \models \varphi$

Proof. For (1), assume $\mathcal{M} \models T \cup \{\varphi\}$ implies $\mathcal{M} \models \psi$, for every structure $\mathcal{M}$. Let $\mathcal{N}$ be a model of T. We need to show that $\mathcal{N} \models_\nu \varphi \to \psi$, for every valuation ν. Now, if $\mathcal{N} \models_\nu \psi$, then $\mathcal{N} \models_\nu \varphi \to \psi$ by definition of satisfaction, so suppose $\mathcal{N} \not\models_\nu \psi$. We will show that $\mathcal{N} \not\models_\nu \varphi$. Contraposing the assumption we get that $\mathcal{N} \not\models_\nu T \cup \{\varphi\}$. But, $\mathcal{N}$ is a model of T, so, in particular, $\mathcal{N} \models_\nu T$. Therefore, $\mathcal{N} \not\models_\nu \varphi$, as claimed. Now applying the definition of satisfaction, we get that $\mathcal{N} \models_\nu \varphi \to \psi$ as required. The converse is proved similarly.

For (2) we prove the backward direction. Assume every model of T is a model of $\neg\varphi$. We want to show that $T \cup \{\varphi\}$ is unsatisfiable (has no models). Suppose for contradiction that $\mathcal{M} \models T \cup \{\varphi\}$. Then, $\mathcal{M} \models T$ and so $\mathcal{M} \models \neg\varphi$ by assumption. On the other hand, $\mathcal{M} \models \varphi$ as well. Take any valuation ν. We then have $\mathcal{M} \models_\nu \varphi$ and $\mathcal{M} \models_\nu \neg\varphi$, so by definition of satisfaction $\mathcal{M} \not\models_\nu \varphi$. This is a contradiction, showing that $T \cup \{\varphi\}$ is unsatisfiable.

For (3) we prove the contrapositive of the forward direction. Assume $T \not\models \varphi$. Then, there is a structure $\mathcal{M}$ and a valuation ν such that $\mathcal{M} \models_\nu T$ and $\mathcal{M} \not\models_\nu \varphi$. Therefore $\mathcal{M} \models_\nu \neg\varphi$ and so $T \cup \{\neg\varphi\}$ is satisfiable. $\square$

Theorem 3.9: Proof by cases

Let $T \cup \{\varphi, \psi, \chi\}$ be a set of formulae. The following are equivalent:

(1) $T \cup \{\varphi\} \models \chi$ and $T \cup \{\psi\} \models \chi$
(2) $T \cup \{\varphi \vee \psi\} \models \chi$.

Proof. By Deduction Theorem, (1) is equivalent to $T \models \varphi \to \chi$ and $T \models \psi \to \chi$, and (2) is equivalent to $T \models (\varphi \vee \psi) \to \chi$. Assume (1), and take a model $\mathcal{M}$ of T and an arbitrary valuation ν. If $\mathcal{M} \models_\nu \chi$, then $\mathcal{M} \models_\nu (\varphi \vee \psi) \to \chi$ as well, so assume $\mathcal{M} \not\models_\nu \chi$. Then, $\mathcal{M} \not\models_\nu \varphi$ and $\mathcal{M} \not\models_\nu \psi$, so $\mathcal{M} \not\models_\nu \varphi \vee \psi$. Thus, $\mathcal{M} \models_\nu (\varphi \vee \psi) \to \chi$ for any valuation ν, that is, $\mathcal{M} \models (\varphi \vee \psi) \to \chi$, as required.

The converse is left as an exercise. □

Theorem 3.10: Proof by arbitrary constant

Let $\varphi(v)$ be a formula with exactly one free variable v, in some first order language L, and T be a set of formulae of L. Let L^c be the expansion of L by a new constant symbol c. The following are equivalent:

(1) $T \models \varphi(c)$
(2) $T \models \forall v\colon \varphi(v)$.

Proof. Let $\mathcal{M}$ be a model of T (in the signature of L). Let $(\mathcal{M}_k)_{k \in M}$ be an enumeration of the set of all expansions of $\mathcal{M}$ by adding a new constant c to the language and interpreting it as $c^{\mathcal{M}_k} = k \in M$. Each $\mathcal{M}_k$ is still a model of T. Now, assume (1). Then, $\mathcal{M}_k \models \varphi(c)$, for each $k \in K$. Consider a valuation ν in $\mathcal{M}$. For any μ differing from ν at most on v, we have $\mu(v) = c^{\mathcal{M}_k}$ for some $k \in K$, so because $\mathcal{M}_k \models \varphi(c)$ holds by assumption, we conclude that $\mathcal{M} \models_\mu \varphi(v)$. Therefore, $\mathcal{M} \models_\nu \forall v\colon \varphi(v)$ for any ν, so $\mathcal{M} \models \forall v\colon \varphi(v)$, as needed.

The converse is left as an exercise. □

CHAPTER 4

Homomorphisms, substructures and direct products

*** write it up ****

CHAPTER 5

Games and equivalence

Techniques based around games are absolutely fundamental in modern logic and model theory. The basic seed for these games are the back and forth "Ehrenfeucht-Fraïssé games": formulated by Fraïssé in the 1950s, and rephrased in the context of games by Ehrenfeucht in the 1960s. These games involve two players, and are played on two structures $\mathcal{A}$ and $\mathcal{B}$ of the same signature. (In more general instances not considered in these notes, we may want a family of pairs of structures $\{(\mathcal{A}_n, \mathcal{B}_n) \mid n \in \omega\}$. The basic case we consider may be thought of as the situation where each of the pairs is identical.)

The game assumes that both players have complete knowledge of the two structures. The first (male) player—$\forall$, or *spoiler*, or $\forall$belard—is determined to show that $\mathcal{A}$ is different to $\mathcal{B}$. He always makes the first move in each round of play. The second (female) player—$\exists$, or *duplicator*, or $\exists$loise—is determined to try to show that $\mathcal{A}$ and $\mathcal{B}$ are the same.

Definition 5.1: the back and forth game

Round 0. If the signature L contains constants, then $\exists$ loses the game immediately (after 0 rounds) if the substructure $\langle\varnothing\rangle_{\mathcal{A}}$ of $\mathcal{A}$ generated by constants is not isomorphic to the corresponding constant-generated substructure $\langle\varnothing\rangle_{\mathcal{B}}$ of $\mathcal{B}$. Otherwise $\exists$ wins the 0th round of play.

Now assume that $j \geq 0$ rounds have been played, and that sequences $a_1, \ldots, a_k \in A$ and $b_1, \ldots, b_k \in B$ have been selected in these first j rounds.

Round $\mathbf{j+1}$.

Part 1. $\forall$ selects some $a_{j+1} \in A$ (or $b_{j+1} \in B$),

Part 2. $\exists$ selects some $b_{j+1} \in B$ (or $a_{j+1} \in A$, respectively).

If the map ϕ^-_{j+1} defined by $\phi^-_{j+1} : a_i \mapsto b_i$ extends an isomorphism ϕ_{j+1} from $\langle a_1, \ldots, a_{j+1}\rangle$ to $\langle b_1, \ldots, b_{j+1}\rangle$, then $\exists$ wins this play of the $(j+1)$-round back and forth game. Otherwise, $\forall$ wins this play of the $(j+1)$-round back-and-forth game.

FIGURE 1. The 11th century philosophers, collaborators and tragic lovers, Abelard and Heloïse as depicted in a 14th century manuscript *Roman de la Rose*. From Wikipedia commons.

Note that when $j = 0$, the sequences $a_1, \ldots, a_j$ and $b_1, \ldots, b_j$ are empty, as is consistent with the description of round 0. Also note that if ϕ^-_{j+1} extends to an isomorphism, then it (that is, ϕ_{j+1}) is unique. Thus if ∃loise loses some round k, then she loses any further rounds of play.

We play two main kinds of game. The ω back-and-forth game is when a round is played for each $j \in \omega$. In other cases we play k rounds of the game, where k is specified in advance. ∃loise has a *winning strategy* for the ω game (or the k-round game) if by suitable play, she can avoid losing the ω (k-round, respectively) back-and-forth game, regardless of how ∀belard plays. Obviously, if ∃loise has a strategy for the k-round game, then the same strategy will win her any shorter game as well. However in general her strategy (if it exists) may depend on k, so even if for every $k \in \omega$ she has a strategy, she may not have a strategy to win the ω game.

Definition 5.2

Write $\mathcal{A} \sim_\omega \mathcal{B}$ if ∃loise has a winning strategy for the ω back-and-forth game. We say that $\mathcal{A}$ is *back-and-forth equivalent* to $\mathcal{B}$. Write $\mathcal{A} \sim_k \mathcal{B}$ if ∃loise has a winning strategy for the k-round back-and-forth game.

Note that $\sim_\omega$ and $\sim_k$ really are equivalence relations. Reflexivity and symmetry are obvious. For transitivity, assume that $\mathcal{A} \sim_\omega \mathcal{B}$ (using strategy 1) and $\mathcal{B} \sim_\omega \mathcal{C}$ (using strategy 2). Then ∃loise can win the ω-round game played on the pair $\mathcal{A}$ and $\mathcal{C}$ as follows: when ∀belard selects a_{j+1} from $\mathcal{A}$ then, using some spare working paper,

∃loise uses strategy 1 to select b_{j+1} from $\mathcal{B}$, then she uses strategy 2, to make a selection of c_{j+1} from $\mathcal{C}$, as if ∀belard had made the b_{j+1} selection. Similarly, if ∀belard selects from $\mathcal{C}$, she uses strategy 2 to make a selection from $\mathcal{B}$, then selects from $\mathcal{A}$ (with strategy 1) as if this selection from $\mathcal{B}$ has been made by ∀belard. We leave the details as an exercise. The $\sim_\omega$ case is identical.

In the following example, ∃loise has a strategy to win the ω-back and forth game.

Example 5.1

Let $\mathcal{A}$ and $\mathcal{B}$ be any two dense linear orders without endpoints. ∃loise does not lose the 0 round game, as there are no constants. Also, it is impossible for her to lose round 1, because after any selection of $a_1 \in A$ and $b_1 \in B$, the structures $\langle a_1 \rangle$ and $\langle b_2 \rangle$ are simply one-element ordered sets, and these are always isomorphic. Now assume that ∃loise has avoided defeat after k rounds. We have sequences $a_1, \ldots, a_k$ from A and $b_1, \ldots, b_k$ from B. Assume ∀belard selects a_{k+1} from $\mathcal{A}$. One of the following is true, and ∃loise's play will maintain isomorphism, to win the $(k+1)$th round:

(1) a_{k+1} is equal to a_i for some $i \leq k$: ∃loise should choose $b_{k+1} := b_i$.
(2) a_{k+1} is strictly greater than a_i, for all $i \leq k$: ∃loise should choose b_{k+1} to be any element of $\mathcal{B}$ strictly greater than b_i for all $i \leq k$.
(3) a_{k+1} is strictly smaller than a_i, for all $i \leq k$: ∃loise should choose b_{k+1} to be any element of $\mathcal{B}$ strictly smaller than b_i for all $i \leq k$.
(4) there are $i, j \leq k$ with a_j covering a_i in $\langle a_1, \ldots, a_k \rangle$ and a_{k+1} has been selected between a_i and a_j. Then b_i is covered by b_j in $\langle b_1, \ldots, b_k \rangle$ (as ϕ_k is an isomorphism), and the theory of dense linear orders means that there exists a choice of b_{k+1} strictly between b_i and b_j: ∃loise should make such a choice.

There is an obvious dual strategy for when ∀belard selects from $\mathcal{B}$ instead of $\mathcal{A}$. So ∃loise has a uniform strategy for winning the ω-back and forth game: *dense linear orders without endpoints are back-and-forth equivalent.*

Notice that we implicitly used the axioms for dense linear orders to inform ∃loise's strategy. In items 2 and 3, we used the "without endpoints" property to be confident that a choice of b_{k+1} exists. In item 4, we used denseness to ensure that the interval $[b_i, b_j]$ contained some b_{k+1} with $b_i < b_{k+1} < b_j$. Notice that the strategy is easily adapted to allow for dense linear orders *with* endpoints: if ∀belard chooses the (unique) top element, then ∃loise also chooses the top element of the other structure, and so on.

In the next example, ∃loise has no winning strategy for the ω back-and-forth game.

Example 5.2

Let $\mathcal{A}$ be the ω chain $0 \leq 1 \leq 2 \leq \ldots$ and $\mathcal{B}$ be the linear sum of ω with a $\mathbb{Z}$-chain. See Figure 2. ∀belard can always win the ω back and forth game by playing as follows. First he should select b_1 of infinite height in $\mathcal{B}$. There are no elements of infinite height in $\mathcal{A}$, so ∃loise is forced to choose some element $n \in \omega$. Now ∀belard simply chooses the unique element b_2 covered by b_1 in $\mathcal{B}$ (this is not the labellig shown in Figure 2 however). To avoid defeat, ∃loise must choose some element a_2 less than $a_1 = n$ in $\mathcal{A}$. Then ∀belard chooses b_3 covered by b_2, and ∃loise must choose some a_3 strictly less than $a_2 \leq n - 1$ in $\mathcal{A}$. After at most $n + 1$ rounds of these moves by ∀belard, ∃loise has either already lost, or she has been forced to chose $a_{n+1} = 0$ (while b_{n+1} still has infinite height in $\mathcal{B}$). ∀belard now chooses an element strictly less than b_{n+1}, and ∃loise has no response that maintains isomorphism. She loses the ω back-and-forth game.

∃loise's failure in Example 5.2 suggests that she can survive for any fixed finite number of rounds by choosing a_1 sufficiently high in $\mathcal{A}$. This is true, but it is slightly less obvious than it looks: the play we described for ∀belard is not his only trick! Imagine for example that ∃loise chose $a_1 = 30$. It might appear that she will survive another 30 rounds. However now Abelard could choose b_2 to be 30 positions lower than b_1. ∃loise now must choose some element a_2 beneath 30. If she chooses $a_2 < 15$, then ∀belard can win with *at most* 15 moves simply by using the strategy of Example 5.2. If she chooses $a_2 \geq 15$, then he can again win in at most 15 moves using the existing strategy, but this time, sandwiching between a_2 and a_1. So contrary to initial impressions, ∀belard can win in only 16 rounds. In fact, he can continue

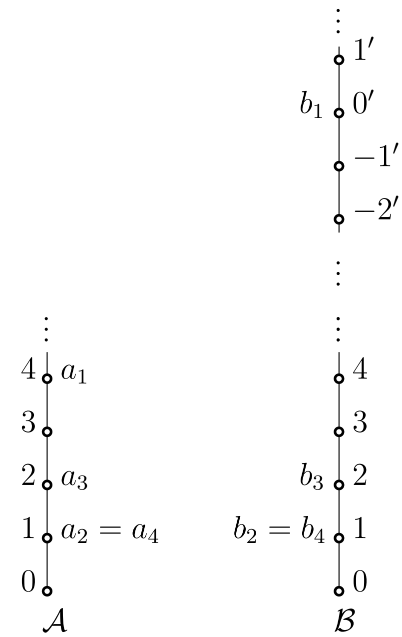

FIGURE 2. The structures $\mathcal{A}$ and $\mathcal{B}$ from Example 5.2 and 5.3. Here the labelling $a_1, \ldots, a_4$ and $b_1, \ldots, b_4$ show four rounds having been played, and ∃loise has managed to maintain a winning position. In a further two rounds of play however, ∀belard could win, because he can select b_5 and b_6 in the interval $[b_3, b_1]$, but there is only one remaining selection for ∃loise in the corresponding interval $[a_3, a_1]$ in $\mathcal{A}$.

to adjust his strategy to win the game in around $\log_2(n) + 1$ moves, where n is the value of the ∃loise's initial selection.

Let's make this more formal.

Example 5.3

For any $k \in \omega$, ∃loise has a strategy to win the k-round back-and-forth game when played in the structures $\mathcal{A}$ and $\mathcal{B}$ of Example 5.2 (and Figure 2).

PROOF. For the sake of streamlined notation, let a_0 denote the bottom element of $\mathcal{A}$ and b_0 denote the bottom element of $\mathcal{B}$.

For each $i = 1, \ldots, k$, we will say that the interval between two points $c < d$ in $\mathcal{A}$ or $\mathcal{B}$ is "large$_i$" if there is a sequence of elements $c < c_1 < \cdots < c_\ell < d$, where $\ell \geq 2^{k-i} - 1$. Now ∃loise plays in a very similar fashion to her strategy for dense linear orders, as described in Example 5.1, however now the goal is to not only maintain isomorphism at each of the rounds $i = 1, \dots, k$, but to ensure that if a_{i_1} is covered by a_{i_2} in $\langle a_0, \dots, a_i \rangle$, then

- if the interval $[a_{i_1}, a_{i_2}]$ is large$_i$, then the interval $[b_{i_1}, b_{i_2}]$ is also large, and
- if the interval $[a_{i_1}, a_{i_2}]$ is not large$_i$ (say, it is $\ell < 2^{k-i} - 1$), then height of the interval $[b_{i_1}, b_{i_2}]$ is identical to $[a_{i_1}, a_{i_2}]$

(and similarly for selections from $\mathcal{B}$). We will show, inductively, that this hypothesis on interval sizes can be maintained. The crucial observation is that if $[c, d]$ is a large$_i$ interval (as witnessed by $c_1 < \cdots < c_{2^{k-i}-1}$), then both the intervals $[c, c_{2^{k-(i+1)}}]$ and $[c_{2^{k-(i+1)}}, d]$ are large$_{i+1}$, as is routinely verified. Note also that a large$_k$ gap is simply an instance of covering: crucially, $\exists$loise is never forced to choose a repeat when $\forall$belard has not.

So suppose, by induction, that $\exists$loise has managed to maintain the the condition of interval gaps to the completion of round i. If $i = k$ we are done. Otherwise, consider a selection of a_{i+1} by $\forall$belard from $\mathcal{A}$. If a_{i+1} is equal to some earlier point, then clearly $\exists$loise can make the corresponding selection from $\mathcal{B}$. If a_{i+1} strictly greater than $a_0, a_1, \dots, a_i$, then consider the interval between the highest point a_j amongst $a_0, a_1, \dots, a_i$ and a_j. If this is large$_{i+1}$, then $\exists$loise should select a point b_{i+1} from $\mathcal{B}$ that is above all of $b_0, \dots, b_i$ such that the interval from b_j to b_{i+1} is also large$_{i+1}$. If a_{i+1} lies within some interval $a_{i_1} < a_{i_2}$ (where a_{i_2} covers a_{i_1} in $\langle a_0, \dots, a_i \rangle$), then there are four possibilities, which we consider as follows.

- If $[a_{i_1}, a_{i_2}]$ is not large$_i$, then (by the induction hypothesis) the gap from a_{i_1} to a_{i_2} is identical to that between b_{i_1} and b_{i_2}, so that b_{i+1} may be selected by $\exists$loise in $\mathcal{B}$ to lie in an exact corresponding position to a_{i+1}. The intervals $[a_{i_1}, a_{i+1}]$ and $[b_{i_1}, b_{i+1}]$ have identical size, so the large$_{i+1}$ property is carried through, and similarly for $[a_{i+1}, a_{i_2}]$ and $[b_{i+1}, b_{i_2}]$.
- If $[a_{i_1}, a_{i_2}]$ is large$_i$, but $[a_{i_1}, a_{i+1}]$ is not large$_{i+1}$, then the definition of large$_i$ ensures that $[a_{i+1}, a_{i_2}]$ *is* large$_{i+1}$. This also ensures that $\exists$loise may select b_{i+1} such that from $[b_{i_1}, b_{i+1}]$ is of identical size to $[a_{i_1}, a_{i+1}]$ and then (necessarily) $[b_{i+1}, b_{i_2}]$ is large$_{i+1}$. So the induction hypothesis on interval size is maintained to the $(i+1)$th round.
- The case where $[a_{i_1}, a_{i_2}]$ is large$_i$, but $[a_{i+1}, a_{i_2}]$ is not large$_{i+1}$ is dual to the previous case.

- If $[a_{i_1}, a_{i_2}]$ is large$_i$, and both $[a_{i_1}, a_{i+1}]$ and $[a_{i+1}, a_{i_2}]$ are large$_{i+1}$. By the inductive hypothesis, the interval $[b_{i_1}, b_{i_2}]$ in $\mathcal{B}$ is also large$_i$, so that there is a selection of b_{i+1} between b_{i_1} and b_{i_2} such that the intervals $[b_{i_1}, b_{i+1}]$ and $[b_{i+1}, b_{i_2}]$ are both large$_{i+1}$. Thus again, the inductive hypothesis on interval gaps is maintained.

The argument for selections by ∀belard in $\mathcal{B}$ instead of $\mathcal{A}$ are basically identical: because the definition of large$_i$ is identical on $\mathcal{A}$ and $\mathcal{B}$. After k rounds, ∃loise has not lost, so that $\mathcal{A} \sim_k \mathcal{B}$ as claimed. □

We now consider the back-and forth equivalence.

Theorem 5.1

For countable structures $\mathcal{A}, \mathcal{B}$, we have $\mathcal{A} \cong \mathcal{B}$ if and only if $\mathcal{A} \sim_\omega \mathcal{B}$.

Proof. (⇒) If $\iota : \mathcal{A} \to \mathcal{B}$ is an isomorphism, then ∃loise's strategy is simply to use this isomorphism to determine her play. Clearly she wins the 0th round of play. If at the ith round, ∀belard chooses $a_i \in \mathcal{A}$, then ∃loise should choose $b_i := \iota(a_i)$. If ∀belard chooses $b_i \in B$, then ∃loise chooses $a_i := \iota^{-1}(b_i)$. In this way, the map ϕ_i^- is always just the restriction of the isomorphism ι to the sets $\{a_1, \ldots, a_i\}$, with $b_1 = \iota(a_1)$, $b_2 = \iota(b_2)$ and so on. So ϕ_i always exists: it is the restriction of ι to $\langle a_1, \ldots, a_i \rangle_{\mathcal{A}}$. Thus ∃loise wins using this strategy.

(⇐) Assume $\mathcal{A} \sim_\omega \mathcal{B}$. Now $\mathcal{A}$ and $\mathcal{B}$ are both countable. If one is finite, with say n elements, then ∀belard can easily play in such a way that after n rounds there is an isomorphism from $\mathcal{A}$ to $\mathcal{B}$. So we may assume that both $\mathcal{A}$ and $\mathcal{B}$ are countably infinite. Thus we may list their elements as sequences $a_1, a_2 \ldots$ and $b_1, b_2, \ldots$ with $A = \{a_i \mid i \in \mathbb{N}\}$ and $B = \{b_i \mid i \in \mathbb{N}\}$. We now inductively detail a sequence of moves $a'_1, b'_2, a'_3, b'_4, \ldots$ by ∀belard in the ω back-and-forth game, such that if ∃loise plays according to her strategy (selecting $b'_1, a'_2, b'_3, a'_4, \ldots$), a full isomorphism is obtained.

∀belard begins by selecting $a'_1 := a_1$. We let b'_1 denote the the element selected by ∃loise. Next, let ∀belard select the earliest element (in the list $b_1, b_2, \ldots$) in $B \backslash \{b'_1\}$. Let a'_2 denote ∃loise's selection. Because ∃loise is playing with her winning strategy, the map $\phi_2 : \langle a'_1, a'_2 \rangle_{\mathcal{A}} \to \langle b'_1, b'_2 \rangle_{\mathcal{B}}$ is an isomorphism.

Now assume that after k of these two-part steps, we have selected $a'_1, a'_2, \ldots, a'_{2k}$ and $b'_1, b'_2, \ldots, b'_{2k}$ such that ϕ_{2k} (described in Definition 5.1) is an isomorphism from $\langle a'_1, a'_2, \ldots, a'_{2k} \rangle$ to $\langle b'_1, b'_2, \ldots, b'_{2k} \rangle$ *and* such that $\{a_1, \ldots, a_k\} \subseteq \{a'_1, \ldots, a'_{2k}\}$ and $\{b_1, \ldots, b_k\} \subseteq \{b'_1 \ldots, b'_{2k}\}$.

We show how to extend this to the $(k+1)$th step. $\forall$belard should select a'_{2k+1} as the earliest element of $A \backslash \{a'_1, \ldots, a'_{2k}\}$, while b'_{2k+1} denotes $\exists$loise's selection. Then $\forall$belard selects b'_{2k+2} as the earliest element of $B \backslash \{b'_1 \ldots, b'_{2k}\}$, and a'_{2k+2} denotes $\exists$loise's selection from $\mathcal{A}$. As $\exists$loise plays her winning strategy, ϕ_{2k+2} is an isomorphism and both

$$\{a_1, \ldots, a_k, a_{k+1}, a_{2k+2}\} \subseteq \{a'_1, \ldots, a'_{2k}, a'_{2k+1}, a'_{2k+2}\}$$

and

$$\{b_1, \ldots, b_k, b_{k+1}, b_{k+2}\} \subseteq \{b'_1 \ldots, b'_{2k}, b'_{2k+1}, b'_{2k+2}\}.$$

In this way $\phi_1 \subseteq \phi_2 \subseteq \phi_3 \subseteq \ldots$, with each $a_i \in A$ in the domain of cofinitely many ϕ_j, and each b_i in the range of cofinitely many ϕ_j. So $\bigcup_{i \in \mathbb{N}} \phi_i$ is an isomorphism from $\mathcal{A}$ to $\mathcal{B}$, as required. □

Corollary 5.1

If all models of some consistent theory T in a countable language are back-and-forth equivalent, then T is complete and ω-categorical.

Proof. In a countable signature, completeness follows from ω-categoricity. But by Theorem 5.1, two countable models of T will be isomorphic because they are back-and-forth equivalent. So T is ω-categorical. □

Example 5.4

Corollary 5.1 shows that the theory of dense linear orders without endpoints is complete (and ω-categorical), because Example 5.1 shows that all dense linear orders without endpoints are back-and-forth equivalent. The same is true for dense linear orders with endpoints, or with bottom but no top, or with top but no bottom.

Example 5.5

The theory of the random graph is ω-categorical. This is axiomatised (within the family of simple graphs) by the family of sentences $\phi_{n,m}$ (for $n, m \in \omega$) stating: "for all disjoint pairs of sets of vertices $\{x_1, \ldots, x_n\}$ and $\{y_1, \ldots, y_m\}$, there is a vertex z such that z is adjacent to each of $x_1, \ldots, x_n$ and none of $y_1, \ldots, y_m$. So the theory of the random graph is ω-categorical.

Proof. Let $\mathcal{A}$ and $\mathcal{B}$ be two models of these sentences. ∃loise's strategy is as follows. Assume that after k rounds elements $a_1, \ldots, a_2$ from $\mathcal{A}$ and elements $b_1, \ldots, b_k$ from $\mathcal{B}$ have been chosen, with the induced subgraph of $\mathcal{A}$ on $\{a_1, \ldots, a_k\}$ isomorphic to that of $\mathcal{B}$ on $\{b_1, \ldots, b_k\}$ (under the map $a_i \mapsto b_i$). Without loss of generality, we may assume that ∀belard selects some a_{k+1} from $\mathcal{A}$. If a_{k+1} is already amongst the $\{a_1, \ldots, a_k\}$, then ∃loise should choose b_{k+1} from $\{b_1, \ldots, b_k\}$ in the same way. Otherwise, let A_k denote the subset of $\{a_1, \ldots, a_k\}$ that are adjacent to a_{k+1}, and let A'_k denote $\{a_1, \ldots, a_k\} \backslash A_k$. Let B_k denote $\{b_i \mid a_i \in A_k\}$ and B'_k denote $\{b_1, \ldots, b_k\}$. As $\mathcal{B}$ satisfies the sentence $\phi_{|A_k|,|A'_k|}$, it follows that there exists a vertex b_{k+1} that is adjacent to all of B_k and none of B'_k. Isomorphism is maintained provided ∃loise chooses this vertex. □

Back-and-forth equivalence is clearly too strong in general to determined elementary equivalence: for example, not every theory is ω-categorical! Now we turn to a finer analysis of the k-round back and forth game.

Definition 5.3: Quantifier rank

- Atomic formulæ Φ have quantifier rank $\mathrm{qr}(\Phi) = 0$,
- $\mathrm{qr}(\Phi_1 \wedge \Phi_2) = \mathrm{qr}(\Phi_1 \vee \Phi_2) := \max\{\mathrm{qr}(\Phi_1), \mathrm{qr}(\Phi_2)\}$,
- $\mathrm{qr}(\neg\Phi) := \mathrm{qr}(\Phi)$,
- $\mathrm{qr}(\forall x \Phi(x, \ldots)) = \mathrm{qr}(\exists x \Phi(x, \ldots)) := \mathrm{qr}(\Phi(x, \ldots)) + 1$.

The following definition is a refinement of elementary equivalence to sentences of bounded quantifier rank: you should observe that $A \equiv B$ is the same as $A \equiv_k B$ for all $k \in \omega$. We'll only use this notion for finite relational signatures, but the definition makes sense more generally.

Definition 5.4

For structures $\mathcal{A}$ and $\mathcal{B}$ of the same signature, we will write $A \equiv_k B$ if for every sentence Φ of quantifier rank at most k we have $A \models \Phi$ if and only if $B \models \Phi$.

Theorem 5.2: Revisiting Theorem 5.1

Let $\mathcal{A}$ and $\mathcal{B}$ be structures in the same *finite relational signature* L.

(1) $\mathcal{A} \equiv_k \mathcal{B}$ if and only if $\mathcal{A} \sim_k \mathcal{B}$.
(2) $\mathcal{A} \equiv \mathcal{B}$ if and only if $\mathcal{A} \sim_k \mathcal{B}$ for every $k \in \omega$.

Before we prove Theorem 5.2 we observe a corollary, which is a powerful generalisation of the rather restrictive Corollary 5.1

Corollary 5.2

A theory T in a finite relational signature is complete if and only if every pair of countable models of T are $\sim_k$ equivalent for all k.

Proof. If T is not complete, then there is a sentence ϕ such that both $T \cup \{\phi\}$ and $T \cup \{\neg\phi\}$ are consistent. Let $\mathcal{A}$ be a model of $T \cup \{\phi\}$ and $\mathcal{B}$ be a model of $T \cup \{\neg\phi\}$. Using the downward Lowenheim-Skolem Theorem, we may assume that $\mathcal{A}$ and $\mathcal{B}$ are countable. Then if k denotes the quantifier rank of ϕ, we have two countable models of T that are not $\sim_k$ equivalent, by Theorem 5.2.

Conversely, assume that T is complete. So every pair of models $\mathcal{A}$ and $\mathcal{B}$ satisfy the same sentences, so have $\mathcal{A} \equiv \mathcal{B}$. That is, $\mathcal{A} \sim_k \mathcal{B}$ for all $k \in \omega$, as required. □

Example 5.6

The following set Σ of sentences in the signature $<$ define a complete theory:

(1) $<$ is a strict total order: antireflexive, transitive and $\forall x \forall y\ (x < y \vee y < x)$.
(2) there is a bottom element: $\exists \bot \forall x\ (x = \bot \vee \bot < x)$.
(3) every element has a cover: $\forall x \exists y \forall z\ (x < y \wedge \not(x < z < y))$.
(4) every element covers an element: $\forall x \exists y \forall z\ (y < x\ \ \&\ \neg(y < z < x))$.

Proof. The idea is to show that two countable models of Σ are $\sim_k$ equivalent, for every $k \in \omega$ (countability is not really used in this argument). We use the idea of Example 5.3. A typical model of Σ looks a little like $\mathcal{A}$ or $\mathcal{B}$, but possibly there are further copies of $\mathbb{Z}$-chains stacked higher and higher. Indeed, let $\mathcal{C}$ be a model. By the second kind of sentence (item 2), there is a bottom element to $\mathcal{C}$, call it 0. By item 3, there is a cover of 0, call it 1. Because $\mathcal{C}$ is a linear order (item 1), this cover is unique. Similarly, there is a cover 2 to 1, a cover 3 to 2 and so on. In general an element $c \in \mathcal{C}$ is either one of these elements of finite height above 0, or we can use item 4 to find an infinite descending chain of elements, $c > c_{-1} > c_{-2} > \dots$ each covering the next, as well as an infinite ascending chain $c < c_1 < c_2 < \ldots$, each covering the previous one. In Example 5.3, the structure $\mathcal{A}$ had all elements of finite height above 0, while in $\mathcal{B}$, there was one family of elements of infinite height (but only finite distance between them). The strategy we described for ∃loise now carries across to general models of Σ with essentially no change: use exactly the same definition of large$_i$: all discussion in the proof of Example 5.3 carries across without change. Then Corollary 5.2 shows that Σ is complete. □

The proof of Theorem 5.2 is by induction on k, and is a little more technical than Theorem 5.1. Some new definitions are needed.

For a structure $\mathcal{C}$ and elements $c_1, \ldots, c_n \in C$ we write $(\mathcal{C}, c_1, \ldots, c_n)$ to be the structure in language $L \cup \{c_1, \ldots, c_n\}$ that agrees with $\mathcal{C}$ in L and with the obvious interpretation of $c_1, \ldots, c_n$.

We need the following fundamental model theoretic notion.

Definition 5.5

Let $\mathcal{A}$ be a structure and $\bar{a} \in A^n$ for some n. The *n-type* $\operatorname{tp}(\mathcal{A}, \bar{a})$ *of $\bar{a}$ in $\mathcal{A}$* is the set of formulas $\{\phi(\bar{x}) \mid \mathcal{A} \models \phi(\bar{a})\}$. When n is clear this is just called the *type* of $\bar{a}$ in $\mathcal{A}$.

- The *rank-k* type of $\bar{a}$ in $\mathcal{A}$ is the restriction of $\operatorname{tp}(\mathcal{A}, \bar{a})$ to formulæ of rank k. It can be denoted $\operatorname{tp}_k(\mathcal{A}, \bar{a})$.

The type of $\bar{a}$ is essentially the set of all possible first order properties that hold at $\bar{a}$. We will need only rank-k 1-types to complete our proof, but it should not be a surprise that the notion of an n-type might arise on many model theoretic situations.

An important observation is that $(\mathcal{A}, a) \equiv_k (\mathcal{B}, b)$ if and only if $\operatorname{tp}_k(\mathcal{A}, a) = \operatorname{tp}_k(\mathcal{B}, b)$, because the rank k sentences satisfied by $(\mathcal{A}, a)$ just the evaluation of rank k formulæ in $\operatorname{tp}_k(\mathcal{A}, a)$ at the point a (and because b is just the name of the interpretation of the constant a in $\mathcal{B}$). For example, on the usual ordered set $\mathbf{2} := (\{0,1\}, \leq)$, the formula $\exists x_1(0 < x_1)$ is a rank-1 sentence if 0 is added to the signature, while while the formula $\exists x_1(x < x_1)$ is a formula in $tp_1(\mathbf{2}, 0)$.

The main reason we restrict to finite relational signatures is so that the following lemma is true! With adjustments to the notion of "rank", or with restrictions on the applicability of Theorem 5.2, it is possible to relax the restriction to relational signatures, but we do not pursue this here.

Lemma 5.1

Fix a finite relational signature L.

(1) Up to logical equivalence, there are only finitely many rank-k formulæ in m free variables $x_1, \ldots, x_m$.
(2) There are only finitely many rank-k 1-types in the language L; say $T_1(x), \ldots, T_n(x)$.
(3) For each rank-k 1-type $T_i(x)$, there is a rank k formula $\alpha_i(x)$ such that $\forall x\,(T_i(x) \leftrightarrow \alpha_i(x))$.

Proof. Only part 1 takes effort. The proof is by induction on k. It is trivial for $k = 0$ as the signature is finite and relational. For $k+1$ (assuming truth at k), note that each formula of rank $k+1$ is a Boolean combination of formulæ of the form $\exists x_{m+1}\phi(x_1, \ldots, x_m, x_{m+1})$ where $\phi(x_1, \ldots, x_{m+1})$ is rank k; there are (by hypothesis) only finitely many such formulæ, so any finitely many different Boolean combinations.

(2) follows from (1) and (3) follows from (2) by taking the conjunction of a representative of each different formula in the rank-k type, up to logical equivalence. $\square$

Now we prove Theorem 5.2.

Proof of Theorem 5.2. The proof is by induction on k. The base case is trivial, because it just concerns whether or not the substructure generated by any constants in the signature L are isomorphic, while the quantifier-free sentences are just those that describe the diagram of these constant-generated substructures.

So now assume that we have proved Theorem 5.2 up to some value k. Assume that $\mathcal{A}$ and $\mathcal{B}$ are $\sim_{k+1}$ equivalent, and consider a rank-$(k+1)$ sentence σ. Using $\forall = \neg\exists\neg$, we may assume that σ is a Boolean combination of rank-$(k+1)$ sentences of the form $\exists x\ \phi(x)$. So it suffices to show that $\mathcal{A}$ and $\mathcal{B}$ satisfy the same sentences of this form. Let $\exists x\ \phi(x)$ hold on $\mathcal{A}$. So there is $a_1 \in A$ with $\phi(a_1)$ true in $\mathcal{A}$. Now let $b_1 \in B$ be the move played (using her winning strategy) by $\exists$loise in response to $\forall$belard's selection of a_1. Now $\mathcal{A} \sim_{k+1} \mathcal{B}$, so $\exists$loise can use her strategy to win a further k-rounds of the back-and-forth game, and from this it follows that $(\mathcal{A}, a_1) \sim_k (\mathcal{B}, b_1)$. Hence, by the induction hypothesis (which held for all structures and up to rank k), we have $(\mathcal{A}, a_1) \equiv_k (\mathcal{B}, b_1)$. As $\phi(a_1)$ is a sentence of rank k satisfied by $\mathcal{A}$ (and because b_1 is the interpretation of the constant a_1 in $\mathcal{B}$), we have that $\phi(b_1)$ holds on $\mathcal{B}$, giving $\exists x\phi(x)$ holding on $\mathcal{B}$, as required.

For the converse direction, assume that $\mathcal{A} \equiv_{k+1} \mathcal{B}$. Consider the first move of play in the a $(k+1)$-round back and forth game. Assume $\forall$belard selects a from $\mathcal{A}$, and consider the rank-k 1-type $\mathrm{tp}_k(\mathcal{A}, a)$ of a. By Lemma 5.1, there is a single formula $\alpha(x)$ of rank k such that $\alpha(x)$ defines $\mathrm{tp}_k(\mathcal{A}, a)$. Thus $\mathcal{A} \models \exists x \alpha(x)$, a sentence of rank $k+1$. Hence, as $\mathcal{A} \equiv_{k+1} \mathcal{B}$, we have that $\mathcal{B} \models \exists x \alpha(x)$. Let b be a witness. So $(\mathcal{A}, a) \equiv_k (\mathcal{B}, b)$ (in particular, so far, $\exists$loise has not lost). Then by induction hypothesis, we have $(\mathcal{A}, a) \sim_k (\mathcal{B}, b)$, so that $\exists$loise may survive a further k rounds. Thus $\mathcal{A} \sim_{k+1} \mathcal{B}$ as required. $\square$

Example 5.7

This example will concern simple graphs. Let C_n denote the graph on $0, 1, \ldots, n-1$ with edges $i \sim i+1$, where addition is taken modulo n (so that $n-1 \sim 0$). Let P_n denote the graph on $0, 1, \ldots, n-1$ with the same edges as C_n, except with $n-1 \sim 0$ removed.

Then $\exists$loise has a winning strategy for the 2-round back-and-forth game played on C_n and P_n, but not the 3-round back-and-forth game. A quantifier rank 3 sentence that holds on C_n but fails on P_n is $\forall x \exists y \exists z (y \sim x \sim z \wedge y \neq z)$.

Proof. There are no constants, so $\exists$loise cannot lose the 0 round game. Similarly, there is only one simple graph on one point, so she cannot lose the 1-round game. After $\forall$belard makes his second move, he has identified either a two point graph with no edges, or a two point graph with one edge. (If $n \leq 3$ the first of these options are impossible.) Now $\exists$loise can find a suitable choice in the other structure to maintain the isomorphism. For the 3-round game however, $\forall$belard can play to win as follows. He should select one of the endpoints of P_n as his first selection b_1. $\exists$loise must choose some point a_1 in C_n. As his second move, $\forall$belard should choose one of the neighbours, a_2, to a_1 in C_n. If she is to win this second round, $\exists$loise must choose a neighbour b_2 of b_1 in P_n: there is only one choice because b_1 was an endpoint. Finally, $\forall$belard chooses a_3 to be the other choice of a neighbour to a_1 in C_n: $\exists$loise has no response, because the only neighbour to b_1 in P_n has already been chosen as b_2. If she choose $b_3 = b_2$ she loses because ϕ_3 is not a bijection. If she chooses $b_3 \neq b_2$, then she loses because ϕ_3 is not a graph homomorphism: $a_1 \sim a_3$, but $\phi_3(a_1) = b_1 \not\sim b_3 = \phi_3(a_3)$. $\square$

We'll finish this with a few more examples.

Example 5.8

Consider the signature $\{\}$ of sets. Two sets are back and forth equivalent if and only if they are both infinite. Two sets are $\sim_k$ equivalent if they have at least k elements each.

Proof. There is only $=$ to be preserved. If $\forall$belard selects b_{i+1} from amongst $b_1, \ldots, b_i$, then $\exists$loise makes the corresponding selection from amongst $a_1, \ldots, a_i$. If he chooses b_{i+1} to be a fresh element, then $\exists$loise also chooses a_{i+1} to be a fresh element. This strategy wins $\exists$loise the k-round game, provided that there are at least k elements (so that fresh elements are available). So the strategy wins her the ω back-and-forth game too, provided that there are infinitely many elements to play with. $\square$

Example 5.9

Half example, half exercise. Consider the cyclic graph C_n of Example 5.7. It is clear that for $n \neq m$ the graphs C_n and C_m are not elementarily equivalent. Find a k such that if $k < n, m$, then $C_n \sim_k C_m$.

Proof. The approach is reminiscent of that taken in Example 5.3. Without loss of generality, assume $n < m$. To start with, observe that any two vertices from C_n or C_m have a shortest distance: adjacent vertices have distance 1, and if two vertices share some adjacent vertex, they have distance 2, and so on. We let $d(u, v)$ denote this shortest distance. In general however, there are two intervals described by the vertices u, v, depending on whether we travel clockwise or anticlockwise (as the vertices are an initial segment of $0, 1, \ldots,$ we interpret "clockwise" to mean "increasing", modulo n or m). When we make a selection of some new vertex, it lies in one and only one of these intervals, and hence it is convenient to refer to the "length" of this interval as the distance from one end to the other, *as travelled through the interval*. This length will either be $d(u, v)$ or $n - d(u, v)$ (or $m - d(u, v)$ if we are in C_m). For example, if there is only one vertex selected (from C_n, say), then the distance from it to itself is 0, however this vertex also determines an interval of length n (go all the round the cycle).

Our goal is to define a notion of "large$_i$" distance to mean something along the lines of "big enough to never completely tied by a path after the remaining $k - i$ rounds". Notice that the fundamental relation (the edge relation) is distance 1, while equality (the relation that comes for free) is distance 0: thus distance 0 and 1 must be precisely the same after the kth round if ∃loise is to win. In order to find out the appropriate definition of length$_i$, let us work backwards. So, after completion of the kth round, in order that ϕ_k be an isomorphism from $\langle a_1, \ldots, a_k \rangle$ to $\langle b_1, \ldots, b_k \rangle$ it is necessary and sufficient that for each $i, j \leq k$ we have $d(a_i, a_j) = 0$ if and only if $d(b_i, b_j) = 0$ (this ensures ϕ_k is a bijection), and $d(a_i, a_j) = 1$ if and only if $d(b_i, b_j) = 1$ (this ensures that ϕ_k preserves the edge relation). So the appropriate definition of "large$_k$" should be 2.

Working backwards: if ∃loise has survived to the $(k-1)$th round and then ∀belard sees an interval from a_i clockwise through to a_j with $d(a_i, a_j) = 2$ but $d(b_i, b_j) > 2$, then in the kth round he can win by selecting a_k to be the unique common adjacent vertex to a_i and a_j: no such vertex exists between b_i and b_j, so ∃loise loses. (The same would be true if $d(a_i, a_j) < 2$ and $d(b_i, b_j) > 2$, but then ∃loise would already have lost the $(k-1)$th round.) Also, if ∀belard sees $d(a_i, a_j) = 3$, but $d(b_i, b_j) > 3$, then he can select b_k within the interval between b_i and b_j, but not adjacent to b_i, b_j. Now ∃loise cannot make a selection of a_k in the shortest interval connecting a_i to a_j, because such choices will be adjacent to a_i or to a_k. Now, it is possible that ∃loise can play a_k in some other large interval that remains elsewhere. However, to simplify her strategy, we will fix large$_{k-1}$ to be big enough to avoid such searches: we're going to describe a strategy for ∃loise where she always attempts to choose within the same interval. We've eliminated the possibility of 3 as large$_{k-1}$, however it appears that setting large$_{k-1}$ to be 4 suffices. For then

(1) if ∀belard chooses a_k adjacent to a_i (so a_k will not be adjacent to a_j), then ∃loise can choose b_k adjacent to b_i but not b_j.
(2) if he chooses a_k adjacent to a_j (so a_k will not be adjacent to a_i), then she can choose b_k adjacent to b_j but not b_i.
(3) if he chooses a_k within the interval between a_i and a_j but adjacent to neither, then as the corresponding interval between b_i and b_k is also at least large$_{k-1} = 4$, she can make a corresponding selection of b_k within the interval from b_i to b_k.

From here there is a pattern: it will suffices to set "large$_{k-2}$" to be big enough that we can always select from the middle of such interval to create two large$_{k-1}$ gaps (otherwise ∀belard may be able to force ∃loise to create a less-than-large$_{k-1}$ interval, when he created only large$_{k-1}$ intervals). So "large$_{k-2}$" is 8 (a selection as close as possible to the middle creates at least one interval that is large$_{k-1}$, that is, 4).

Similarly, large_{k-3} can be 16, and so on. Thus in general, large_i can be equal to 2^{k-i+1}. Then, during the ith round, if ∀belard selects a_i from between two vertices a_{j_1} and a_{j_2}, if the length of this interval was not large_{i-1}, then the corresponding gap between b_{j_1} and b_{j_2} is identical, so ∃loise has a matching move. If $d(a_{j_1}, a_{j_2})$ was large_{i-1}, then we get 3 possibilities (like cases 1,2,3 above), and the definition of large_{i-1} enables ∃loise to maintain the pattern of large_i gaps to the completion of round i.

So how big need n be for this strategy? After round 1, it is necessary that the distance from the single point chosen be large_1, as otherwise our strategy for ∃loise requires that this interval be of identical size in C_n and C_m (which it won't be, as $n \neq m$). So we need $n, m \geq 2^k$. □

This bound is not quite tight, but finding the tight bound is tricky. It is easily seen that the bound is tight for $k = 1$ and $k = 2$. For the 3-round game, we have shown that ∃loise can win when played on cycles of length at least 8, but in fact a more refined strategy enables her to survive provided that $n, m \geq 7$ (but not $n = 6$ and $m = 7$). For 4 steps, $n, m \geq 11$ is enough. It was shown in [J. Brown and R. Hoshino, The Ehrenfeucht-Fraïssé game for paths and cycles, *Ars Combinatoria* 83 (2007), 193–212] that for general $k > 2$, ∃loise has a winning strategy if and only if $n = m$ or $n, m \geq 2^{k-1} + 3$. The minimal values for $\sim_k$ equivalent of undirected paths (the next example) are also classified in this article.

Example 5.10

A very similar approach may be taken to compare the undirected paths P_n and P_m. A useful technique here is to first assume that the endpoints have been successfully selected and played by ∀belard and ∃loise (denote these by a_{-1} and a_0, and b_{-1}, b_0 respectively). This simplifies notions of distance and intervals for the remainder of the play. If ∃loise can survive k rounds with these two elements pre-selected, she can survive at least this many rounds without these.

Recall that a field with characteristic 0, is a field (Example 1.7) satisfying the sentence $\neg(\overbrace{1 + 1 + \cdots + 1}^{n} = 0)$ for each n, which is a sentence because 1 is in the signature.. Note that because the constant 1 is in the signature of fields, there is an empty-generated structure and it is isomorphic to the field of rationals $\mathbb{Q}$: we get the integers using addition (and characteristic 0 ensures they are all distinct), then multiplicative inverses of nonzero elements force general elements of $\mathbb{Q}$. This is actually a field, so no further elements are generated.

A field is *algebraically closed* if it satisfies the following sentence for all n: $\forall x_0 \dots x_n \exists y \ (x_n \neq 0 \to x_n y^n + x_{n-1} y^{n-1} + \cdots + x_1 y^1 + x_0 = 0)$. In other words, all polynomial equations have a solution. The field of complex numbers is the most well known instance of an algebraically closed field, and the fact that $\mathbb{C}$ has this property is what is usually known as the *Fundamental Theorem of Algebra*. So it's pretty important!!

An element of a field F is *algebraic* over a subfield $G \leq F$ if it is the solution to a polynomial equation over F. For example, i and $\sqrt{2}$ are algebraic over $\mathbb{Q}$, as they are solutions to $x^2 + 1 = 0$ and $x^2 - 1 = 0$ respectively. An element of $F \backslash G$ that is not algebraic over G is said to be *transcendental*. The number π is transcendental over $\mathbb{Q}$: this looks "obvious" enough, but is a little tricky to prove. The *transcendence degree* of F is the cardinality of the largest subset T of F such that each $t \in T$ is a transcendental over subfield generated by $T \backslash \{t\}$. If the transcendence degree is only countably infinite, then the field is only countable, because there are are only countably many elements that can be generated by countably many transcendentals and algebraic numbers. So the complex numbers have infinite transcendence degree. This next example shows that all algebraically closed fields are very similar.

Example 5.11

Two algebraically closed fields of characteristic 0 are back-and-forth equivalent provided they have either the same finite transcendence degree, or both have infinite transcendence degree. As a consequence the theory of algebraically closed fields of characteristic 0 is complete and there is an algorithm to decide the truth of first order sentences on $\mathbb{C}$ (in the signature $\{+, -, \cdot, 0, 1\}$ of fields).

PROOF. (Very much a sketch.) The empty-generated structure is always isomorphic to $\mathbb{Q}$, so $\exists$loise survives at least to play the first round. Her strategy is simply to copy a selection of the kind of element that $\forall$belard chooses. The key property is that the algebraic behaviour of an element is basically determined by whether it is algebraic (just look at a minimal degree polynomial equation it satisfies) or transcendental. Let the two fields by $\mathcal{F}$ and $\mathcal{G}$. After k rounds there are finitely generated subfields $\mathcal{F}_k$ of $\mathcal{F}$ and $\mathcal{G}_k$ of $\mathcal{G}$, isomorphic under

ϕ_k. Assume without loss of generality that $\forall$belard selects a_{k+1} from $\mathcal{F}$. $\exists$loise simply copies $\forall$belard's selection: if he chooses an existing element in $\mathcal{F}_k$, she selects $\phi_k(a_{k+1})$ from $\mathcal{G}_k$. If a_{k+1} is transcendental over $\mathcal{F}_k$, she chooses an element transcendental over $\mathcal{G}_k$ (so provided the transcendence degrees of $\mathcal{F}$ and $\mathcal{G}$ are either identical, or both are infinite, she can win this round), and if he chooses a_{k+1} algebraic over $\mathcal{F}_k$, she can look at the smallest degree polynomial equation $p(x) = 0$ that $\forall$belard's selection a_{k+1} satisfies. The isomorphism ϕ_k translates $p(x)$ to a polynomial over $\mathcal{G}_k$, and because $\mathcal{G}$ is algebraically closed, $\exists$loise can find in $\mathcal{G}$ a solution to it, and this is her choice of b_{k+1}.

Completeness follows from adjustment to the proof of Corollary 5.1: instead of aiming for ω-categoricity, aim for showing that two models of the same uncountable cardinality are elementarily equivalent. The main technicality is that we have not proved that back-and-forth equivalence implies elementary equivalence (except for countable structures where it even implies isomorphism). As back-and-forth equivalence implies $\sim_k$ equivalence for every $k \in \omega$, we could instead use Corollary 5.2, but that assumes a finite relational signature, and algebraically closed fields involve operations. However, we may replace operations by their graphs, and while the translation from operations to graphs of operations can increase the quantifier rank, it remains true that after a sentence transformed to graphs-of-operations form still has *some* quantifier rank, and thus is equivalently satisfied or not satisfied on back-and-forth equivalent algebraically closed fields.

Finally, we describe the algorithm to decide which sentences are true on $\mathbb{C}$. Let Σ denote the sentences defining "algebraically closed field of characteristic 0": we have just shown that these are complete. As $\mathbb{C}$ is a model of Σ we have that the theory of $\mathbb{C}$ coincides with the consequences of Σ. So to prove that a sentence ϕ is true on $\mathbb{C}$, it suffices to simultaneously search for proofs of both ϕ and $\neg\phi$ from Σ. Because Σ is complete, there is a proof of one (and only one) of these, so a systematic search eventually discovers this proof. If ϕ is proved, then this is true of $\mathbb{C}$. If $\neg\phi$ is proved then ϕ is not true of $\mathbb{C}$. $\square$

This algorithm described in the proof of Example 5.8 is extremely inefficient, but it nevertheless is in stark contrast to the theory of $(\omega, +, \cdot, 0, 1)$, for which it is possible to prove that there is *no algorithm at all* for deciding truth of sentences (this is basically what Gödel's First Incompleteness Theorem states). There's also nothing special about characteristic 0 in Example 5.8: all the arguments work with minor adjustment for any fixed characteristic.

CHAPTER 6

Skolemisation

Thoralf Skolem (1887–1963) was a Norwegian mathematician who made several deep contributions to the early development of model theory. The process of Skolemisation, is a method of eliminating quantifiers by the addition of new operations. He used this to prove the downward Löwenheim-Skolem Theorem. The technique also plays a role in computer science, particularly in logic programming.

The idea is to introduce operations selecting solutions to existential statements. The essence of the idea was well demonstrated by the role of the inverse operation $^{-1}$ in the language of groups. In the language $\{\cdot, 1\}$, the property of having an inverse is usually given by an axiom such as

$$\forall x \exists y\ (x \cdot y = 1 \wedge y \cdot x = 1).$$

In the language $\{\cdot, ^{-1}, 1\}$, the existential quantifier is eliminated, with the law becoming

$$\forall x\ (x \cdot x^{-1} = 1 \wedge x^{-1} \cdot x = 1).$$

The value of x^{-1} picks out a solution y to $x \cdot y = 1 \wedge y \cdot x = 1$. Group inverses are unique, so this is a rather atypical instance. In general operations introduced by Skolemisation are simply designed to pick out one of possibly many solutions. In model theory, Skolemisation is mostly a theoretical tool, but aspects of the idea are commonly encountered.

Definition 6.1

Let T be a theory in a language L. A *Skolemisation of T* is a theory T^+ in a language $L^+ \supseteq L$ such that

(1) Every L-structure that is a model of T can be expanded to be a model of T^+

(2) Every formula $\phi(\bar{x}, y)$ with $\bar{x} = x_1, \ldots, x_n$ nonempty there is a term t in variables $\bar{x}$ such that

$$T^+ \quad \vdash \qquad \forall \bar{x}\left((\exists y\ \phi(\bar{x}, y)) \rightarrow \phi(\bar{x}, t(\bar{x}))\right).$$

The terms t in item 2 of this definition are called *Skolem functions*. The theory T is said to *have Skolem functions* if it is a Skolemisation of itself (so, $T^+ = T$ and $L^+ = L$). In this case, T is said to be a *Skolem theory*. Note that implication $\to$ described in item 2 of this definition, could be replaced by $\leftrightarrow$, as if $\phi(\bar{x}, t(\bar{x}))$ is true under some interpretation, then certainly $\exists y\ \phi(\bar{x}, y)$ is true (with y chosen as $t(\bar{x})$).

A Skolemisation T^+ (in signature L^+) of a theory T (in signature $L \subseteq L^+$) is an example of what is known as a *conservative extension*:

> if ϕ is a sentence in the language L then $T \vdash \phi$ if and only if $T^+ \vdash \phi$.

Indeed if $T \vdash \phi$ then $T^* \vdash \phi$ because $T \subseteq T^*$. Conversely, if $T \not\vdash \phi$, then there is a model $\mathcal{A}$ of T that fails ϕ (see Theorem 3.1). As T^* is a Skolemisation of T, it can be expanded to a model $\mathcal{A}^*$ of T^*: and ϕ still fails on this expansion!! So $T^* \not\vdash \phi$.

Theorem 6.1

Let T be a Skolem theory: that is, a theory containing its own Skolem functions. Then every formula $\phi(\bar{x})$ (with $\bar{x}$ nonempty) is equivalent modulo T to a quantifier-free formula $\phi^*(\bar{x})$.

Proof. This is by induction on the complexity of $\phi(\bar{x})$. If ϕ is quantifier free, then there is nothing to prove. Now assume we have proved this up to some level of complexity, and let $\phi(\bar{x})$ be a formula of one higher complexity. Now $\phi(\bar{x})$ is a Boolean combination of subformulæ, and it suffices to remove quantifiers in each of these. So we may assume without loss of generality that $\phi(\bar{x})$ is itself of the form $\forall y\ \psi(\bar{x}, y)$ or $\exists y\ \psi(\bar{x}, y)$ for some formula $\psi(\bar{x}, y)$. Moreover, if we can complete the second case, then the first can be treated by performing our technique on the negation $\exists y\ \neg\psi(\bar{x}, y)$ and then negating the outcome. Now, because T is a Skolem theory, thee is a term t such that $\exists y\ \psi(\bar{x}, y)$ is equivalent to $\psi(\bar{x}, t(\bar{x}))$. The formula $\psi(\bar{x}, t(\bar{x}))$ is of smaller complexity (it is one quantifier smaller). So by the induction assumption, there is $\psi'(\bar{x})$ with no quantifiers such that $\forall \bar{x}\ \big((\exists y\ \neg\psi(\bar{x}, y)) \leftrightarrow \psi(\bar{x}, t(\bar{x})) \leftrightarrow \psi'(\bar{x})\big)$, as required. $\square$

Example 6.1

Consider some Skolemisation T' of the theory of ordered sets (in the language of a single binary relation $\leq$). How might the elimination of quantifiers proceed for a formula such as: $\forall y\ x \leq y$ (that in the theory of ordered sets asserts that x is the bottom element), or for $\exists y\ x \not\leq y$ (which asserts that x is *not* the bottom)?

DISCUSSION: First note that we will shortly prove that every theory has a Skolemisation, so the assumptions of the example are satisfiable! We do not attempt to describe what T' looks like at this point however.

Let us consider $\exists y\ x \not\leq y$ first. In this case, there must be some unary term $t(x)$ in the language of $T' \supseteq \{\leq\}$ such that $\exists y\ x \not\leq y$ is equivalent to $x \not\leq t(x)$. Note that every ordered set can be enriched in its signature to include operations such as t. For example, it suffices to ensure that $t(a)$ fixes any bottom element (if there is one), and otherwise selects some element that is either incomparable with a or is strictly less than a. Typically there will be many ways to do this, and any one of them would provide an example of a Skolem function for $\exists y\ x \not\leq y$.

Now let us consider the formula $\forall y\ x \leq y$. To find an equivalent quantifier-free expression, we need to rewrite in the form $\neg\exists y\ x \not\leq y$. We already know how to deal with $\exists y\ x \not\leq y$: this property for x is equivalent to $x \not\leq t(x)$. So $\neg\exists y\ x \not\leq y$ is equivalent to $\neg x \not\leq t(x)$, or in other words $x \leq t(x)$. □

Observe that there is a fundamental asymmetry to how we "see" these equivalences. It was easy to see why the property $x \not\leq t(x)$ implies $\exists y\ x \not\leq y$, and it was easy to see how some t could in principle be introduced to satisfy. However in the case of the universal quantifier we cannot "see" that $x \leq t(x)$ implies $\forall y\ x \leq y$! Why would the fact that $x \leq t(x)$ somehow guarantee that *all* y have $x \leq y$? The answer is that we can't directly see this: it crucially depends on first knowing that $x \not\leq t(x)$ was equivalent to $\exists y\ x \not\leq y$, and *only then* noticing that if $x \not\leq t(x)$ fails (that is $x \leq t(x)$) then $\exists y\ x \not\leq y$ fails (that is, $\forall y\ x \leq y$).

Here is a slightly more involved example.

Example 6.2

Let us consider some Skolem theory T in a language containing some binary operation $\{\cdot\}$, and imagine how the formula

$$\exists y \forall z \big(x \cdot y = z \cdot z\big)$$

might reduce to an equivalent quantifier-free formula by the inductive process in the proof of Theorem 6.1.

Discussion! First note that there is one free variable, x. As T is assumed to be a Skolem theory, there must be some term $t_1(x)$ built from operations in the signature of T such that

$$\exists y \forall z \big(x \cdot y = z \cdot z\big) \longleftrightarrow \forall z \big(x \cdot t_1(x) = z \cdot z\big)$$

Now we consider the formula $\forall z \ \big(x \cdot t_1(x) = z \cdot z\big)$, which we write in the equivalent form $\neg \exists z \ \big(x \cdot t_1(x) \neq u \cdot v\big)$. As T is a Skolem theory, there must be some unary term t_2 such that

$$\exists z \ \big(x \cdot t_1(x) \neq z \cdot z\big) \longleftrightarrow \big(x \cdot t_1(x) \neq t_2(x) \cdot t_2(x)\big)$$

So, $\forall z \big(x \cdot t_1(x) = z \cdot z\big)$ is equivalent to $x \cdot t_1(x) = t_2(x) \cdot t_2(x)$. $\square$

The comments on asymmetry made after Example 6.1 are even more evident here: it is completely unobvious why $x \cdot t_1(x) = t_2(x) \cdot t_2(x)$ should be equivalent to $\exists y \forall z \big(x \cdot y = z \cdot z\big)$. It isn't even clear how the order $(\exists \forall)$ of the quantifiers in the original formula are recorded in the quantifier-free equivalent. We can only work such facts out by *already knowing* that $t_2(x)$ was a test for the existence of an element z satisfying $x \cdot t_1(x) \neq z \cdot z$, and that $t_1(x)$ was a test for the existence of some y satisfying $\forall z \ \big(x \cdot y = z \cdot z\big)$.

Most uses of the notion of "Skolemisation" do not involve elimination of quantifiers in the style of Examples 6.1 and 6.2. In fact, there is usually no treatment of universal quantifiers at all, and the process does not involve the same inductive principle used in the proof of Theorem 6.1. The following example presents a typical treatment.

Example 6.3

Consider the signature $\{<\}$ consisting of a single binary relation. A typical Skolemisation of the sentence

$$(1) \qquad \forall u \exists v \forall x \exists y \; \big(v < u \wedge (x < v \to y < x)\big)$$

reduces it to

$$\forall u \forall x \; \big(f(u) < u \wedge (x < f(u) \to g(u,x) < x)\big).$$

where f is a new unary operation symbol and g is a new binary operation symbol.

Discussion. The idea is to work from the left, removing the leftmost existential quantifier. The operation f is added to witness the existence of a value of v such that $\forall x \exists y \; \big(v < u \wedge (x < v \to y < x)\big)$ holds: this value might depend on v, so f is a unary operation. The sentence now becomes

$$(2) \qquad \forall u \forall x \exists y \; f(u) < u \wedge (x < f(u) \to y < x).$$

A model of the sentence (1) becomes a model of the sentence (2) by letting $f(u)$ take any value of v for which $\forall x \exists y \; v < u \wedge (x < v \to y < x)$ holds (and at least one exists if sentence (1) is true, while if the sentence is false, then it does not matter what value we choose for $f(v)$, because all will correctly lead to a false universal sentence).

Next the choice of y in sentence (2) depends on *both* the value of u and v, so a binary operation is needed: g, say. (Note that in a sense it also depends on the choice we previously made for $f(v)$. However f is now in our signature, and its value is determined by u.) Sentence (2) becomes $\forall u \forall x \; f(u) < u \wedge (x < f(u) \to g(u,x) < x)$. □

Notice how we can now “see” backwards: the value $f(u)$ depends on u only, while the value $g(u,x)$ depends on both u and x, so that it can be “seen” that $g(u,x)$ could replace an existentially quantified variable after the $\forall u \forall x$, and that then $f(u)$ could replace an existential quantified variable immediately following $\forall u$.

The following is another important fact about Skolem theories.

Theorem 6.2

Let T be a Skolem theory in some signature L. If $\mathcal{A}$ is an L-model of T and $\varnothing \subsetneq X \subseteq A$, then $\langle X \rangle$ is an elementary substructure of $\mathcal{A}$.

Proof. Let $\phi(\bar{x}, y)$ be a formula in L, and $\bar{a}$ be a tuple in X. We need to show that if $\exists y\ \phi(\bar{a}, y)$ holds in $\mathcal{A}$, then it holds already in $\langle X \rangle$. As T is a Skolem theory, there is a term t such that $T \vdash \forall \bar{x}\ \big(\exists y\ \phi(\bar{x}, y) \to \phi(\bar{x}, f(\bar{x}))\big)$. So $\mathcal{A} \models \exists y\ \phi(\bar{a}, y)$ implies $\mathcal{A} \models \phi(\bar{a}, t(\bar{a}))$. However $\langle X \rangle$ is a substructure containing the elements in $\bar{a}$, so it also contains the value $t(\bar{a})$. That is, there is an element $b \in \langle X \rangle$ (namely, $b := t(\bar{a})$) such that $\langle X \rangle \models \phi(\bar{a}, b)$. So $\langle X \rangle \models \exists y\ \phi(\bar{a}, y)$ as required. □

Theorem 6.1 and 6.2 show that Skolem theories have very strong properties. Unfortunately, not many natural theories are Skolem theories. Rather, the value of the the idea lies in the fact that arbitrary theories can be enriched to become Skolem theories, and the strong properties of these more expressive languages can be used to prove properties of the original. The idea is basically to follow the technique in the proof of Theorem 6.1, but instead of assuming that Skolem functions exist and using them to remove existential quantifiers, instead we introduce new operations to play the role of the Skolem functions. A new operation is introduced for each formula, and then the whole process is repeated ω times.

In the following theorem, the set Ξ might be empty, or it might be $\mathrm{Th}(\mathcal{A})$ for some structure of interest $\mathcal{A}$.

Theorem 6.3

For any language L and any set Ξ of L-sentences, there is a signature $L^\Sigma \supseteq L$ and a set $\Sigma \supseteq \Xi$ of sentences in L^Σ such that

(1) every L-structure $\mathcal{A}$ satisfying T expands to a model of Σ in the language L^Σ.
(2) Σ is a Skolem theory in L^Σ.
(3) $|L^\Sigma| = |L| + \omega$.

Proof. Put $L_0 := L$ and $\Sigma_0 = \Xi$. The set Ξ will actually play no role in the argument. We begin by expanding the language to patch skolemisation for L, however this introduces new formulæs in need of Skolemisation, so the process is repeated, and so on (ω times).

For every formula $\phi(\bar{x}, y)$ where $\bar{x}$ is nonempty, introduce a new operation symbol $f_{\phi,\bar{x}}$ of the same arity n (where $\bar{x} = x_1, \ldots, x_n$). We let L_1 denote the language L combined with all these new operation symbols. We define the sentences Σ_1 to consist of Σ_0 along with all sentences (across all L_0-formulæ $\phi(\bar{x}, y)$):

$$\forall \bar{x}\ \big(\phi(\bar{x}, y) \to \phi(\bar{x}, f_{\phi,\bar{x}}(\bar{x}))\big). \tag{3}$$

We want to show that Σ_1 is a Skolemisation of the empty theory in language L. These sentences are chosen precisely so that the property (2) in Definition 6.1 holds. It remains to show that every L-model $\mathcal{A}$ of T can be expanded to an L_1-model of Σ_1. For each $\bar{a}$ from $\mathcal{A} \models T$ we need to define $f_{\phi,\bar{x}}(\bar{a})$. Now, if there is a solution b such that $\mathcal{A} \models \phi(\bar{a}, b)$, then choose $f_{\phi,\bar{x}}(\bar{a}) = b$. There may be many such selections, but it does not matter which we select, because we just need to ensure that $\phi(\bar{a}, f_{\phi,\bar{x}}(\bar{a}))$ is true. If there is no such solution, choose $f_{\phi,\bar{x}}(\bar{a}) = a_1$ (an essentially arbitrary choice). The implication in Equation (3) holds vacuously in this case, because the premise of the implication is false. This completes the proof that Σ_1 is a Skolemisation of the empty theory in L. Note that $|L_1| = |L| + \omega$.

Unfortunately Σ_1 is not necessarily a Skolem theory itself, because there are *new* formulæ that can be built in all these new operations we introduced. So now apply the same construction again to the language L_1 (to produce L_2) and let Σ_2 denote Σ_1 along with the new sentences created in Equation 3 (applied to L_1 now). Continuing inductively, define L_{i+1} from L_i and Σ_{i+1} from Σ_i by repeating this technique. Each is a Skolemisation of the previous case, and each has $|L_i| = \omega_0 + |L|$. Finally, define L^Σ to be the union $\bigcup_{i\in\omega} L_i$ and $\Sigma(L) := \bigcup_{i\in\omega} \Sigma_i$. We can inductively expand each L-model $\mathcal{A}$ of T through the L_i, satisfying Σ_i for each i. As $L_0 \subseteq L_1 \subseteq L_2 \subseteq \ldots$ and $\Sigma_0 \subseteq \Sigma_1 \subseteq \Sigma_2 \subseteq \ldots$, each expansion of $\mathcal{A}$ at the $(i+1)$th stage, is an expansion of the expansion at the ith stage, and so in the ω limit, one obtains an expansion of $\mathcal{A}$ to the language L^Σ. The sentences in $\Sigma(L)$ hold, because each such sentence lies in Σ_i for some i. For the same reason, $\Sigma(L)$ is a Skolem theory: each formula $\exists y\ \phi(\bar{x}, y)$ in L^Σ is a formula in L_i for some i, so was Skolemised in moving to L_{i+1} and $\Sigma_{i+1} \subseteq \Sigma(L)$. Finally, an ω union of sets of size $|L| + \omega$ is itself of size $|L| + \omega$, which was the last detail required to prove. □

A nice application of the Skolemisation Theorem (and the original motivation) is the Downward Löwenheim Skolem Theorem 3.6.

Corollary 6.1: The Downward Löwenheim-Skolem Theorem. Again!

Let $\mathcal{A}$ be a model in a signature L. Let $X \subseteq A$ and λ be a cardinal with $|X|+|L|+\omega_0 \leq \lambda \leq |A|$. Then $\mathcal{A}$ has an elementary substructure B with $X \subseteq B$ and $|B| = \lambda$.

Proof. We assume that X is nonempty, or there are constants (if X were empty and there were no constants, then as $\omega_0 \leq \lambda \leq |A|$, we may add an arbitrary element of A to X without changing the assumptions of the statement). Now use Theorem 6.3 to expand $\mathcal{A}$ to a structure $\mathcal{A}'$ on A satisfying $\Sigma(L)$. Let $B' \subseteq A$ be some set of size λ and let $\mathcal{B}$ be the substructure of $\mathcal{A}'$ generated by $X \cup B'$. Note that $|B| = |\langle X \cup B' \rangle| = |X| + |B'| + |L| \leq \lambda + \lambda + |L| = \lambda$ (as $\omega_0 + |L| \leq \lambda$) as required. And $\mathcal{B}$ is an elementary substructure of $\mathcal{A}'$ by Theorem 6.2. Hence the reduct to the language L is also an elementary substructure (if $\exists y\ \phi(\bar{b}, y)$ in the language L has a solution $y = a$ in $\mathcal{A}$, then the same value a witnessed a solution in $\mathcal{A}'$, and hence there was a solution in $\mathcal{B}'$. But $\mathcal{B}'$ and $\mathcal{B}$ agree on the language L, so $\exists y\ \phi(\bar{b}, y)$ holds in $\mathcal{B}$). ☐

The following two corollaries have only half been observed.

Corollary 6.2

The following are equivalent for a consistent theory T in a countable signature:

(i) T is complete;
(ii) for some cardinal $\kappa \geq \omega$, all cardinality κ models of T are elementarily equivalent;
(iii) countable models are of T are elementarily equivalent;
(iv) (for relational signatures...) countable models of T are $\sim_k$ equivalent for all $k \in \omega$

Proof. These are all almost trivial if all models of T are finite, so we now assume that T has infinite models.

(iii) and (iv) were shown equivalent in Corollary 5.2. Also, (iii) implies (ii) trivially. Now, if (i) holds (that is, T is complete), then all models of T are elementarily equivalent, so (ii) and (iii) hold. Now we show (ii) implies (i) in the contrapositive.

Assume that T is not complete. So there is some sentence ϕ such that both $T \cup \{\phi\}$ and $T \cup \{\neg\phi\}$ are consistent. By the Completeness Theorem, there is a model $\mathcal{A}_1$ of $T \cup \{\phi\}$ and $\mathcal{A}_2$ of $T \cup \{\neg\phi\}$. By the Upward Lowenheim-Skolem Theorem, we may assume that $\mathcal{A}_1$ and $\mathcal{A}_2$ have cardinality at least κ. By the Downward Lowenheim Skolem Theorem, there are elementary substructures $\mathcal{B}_1 \preceq \mathcal{A}_1$ and $\mathcal{B}_2 \preceq \mathcal{A}_2$ with $|B_1| = |B_2| = \kappa$. But $\mathcal{B}_1$ and $\mathcal{B}_2$ are not elementarily equivalent because they disagree on ϕ (it holds on $\mathcal{B}_1$ but fails on $\mathcal{B}_2$). So (ii) fails, as required. ☐

Corollary 6.3

The following are equivalent for a consistent countable theory T:

(i') T is ω-categorical;
(ii') countable models of T are back-and-forth equivalent.

Proof. This is just Theorem 5.1 combined with Corollary 6.2. $\square$

1. Appendix: Sets and ordinals

Many of the statements in this section assume the axiom of choice (the statement that a Cartesian product of nonempty sets is nonempty).

Definition .2: Well ordered sets

A *linearly ordered set* $(S; \leq)$ is an ordered set such that $(\forall a, b \in S)\ a \leq b$ or $b \leq a$.
A well ordered set $(S; \leq)$ is a linearly ordered set where every non-empty subset $A \subseteq S$ has a least element.

If $(S; \leq)$ is a well-ordered set (with $S \neq \varnothing$) then as S is a non-empty subset of itself, the definition of being well-ordered shows that S has a least element. There are stronger consequences too: every element of a well ordered set is either a maximum or has a successor (that is a cover). To see this, observe that if a is not the maximum, then $\{b \in S \mid a < b\}$ is non-empty so has a minimum. This is the successor of a. Another nice property is that any subset of a well ordered set is itself well ordered. It is not hard to prove (using some amount of the axiom of choice) that being well ordered is equivalent to the property that there is no infinite descending chain $x_1 > x_2 > x_3 > \dots$. This is known as the *descending chain condition* (DCC).

Example .4

(1) Every finite chain is well ordered.
(2) $(\mathbb{N}, \leq)$ (the ω-chain) is well ordered.
(3) $(\mathbb{Z}, \leq)$ is not well ordered.
(4) $(\mathbb{Q}_0^+, \leq)$ (where $\mathbb{Q}_0^+$ denotes $\{x \in \mathbb{Q} \mid x \geq 0\}$) is not well ordered.

Proof. (1) is trivial. (2) is almost trivial: it is certainly very easy to show that $\mathbb{N}$ satisfies the DCC, which is equivalent (for a linearly ordered set) to being well ordered. For $(\mathbb{Z}, \leq)$: it has no minimum, so is not a well-ordering. For $(\mathbb{Q}_0^+, \leq)$ use the subset $\{x \in \mathbb{Q} \mid x > 0\}$, which has no minimum. □

A well known equivalent of the axiom of choice is that statement that "every set can be well ordered". The following example gives easy well-orderings for the sets encountered in Example .4, but you will find it quite impossible to explicitly describe a well-ordering for $\mathbb{R}$.

Example .5

$0, 1, -1, 2, -2, 3, -3, \ldots$ is a well ordering of the integers $\mathbb{Z}$. $1/1, 2/1, 3/1, \ldots, 1/2, 3/2, 5/2, \ldots, 1/3, 2/3, 4/3, \ldots$ is a well ordering of the non-negative rationals $\mathbb{Q}_0^+$. Another is $0, 1, 2, 1/2, 3, 4, 1/3, 3/2, 5, 2/3, \ldots$ where p/q (a fraction with $\gcd(p, q) = 1$) appears before p'/q' if $2^p 3^q < 2^{p'} 3^{q'}$.

The isomorphism classes of well orderings are called *ordinals*; we actually identify an ordinal with a member of its class rather than use the entire class. In the case of the smallest ordinal: there is only one, $\varnothing$. In general, ordinals are themselves well-ordered (or at least, any set of ordinals is well-ordered): for ordinals α, β we define $\alpha \leq \beta$ if α is (isomorphic to) an initial segment of β, that is, there is an order isomorphism of α onto a downset of β.

The finite ordinals are simply the finite chains. The first infinite ordinal ω is order isomorphic to $\mathbb{N}_0$ with its usual order. The second infinite ordinal is $\omega + 1$. The ω^{th} infinite ordinal is $\omega + \omega$.

Definition .3: Cardinality

Two sets A and B are said to have the same *cardinality* if there exists a bijection $\varphi : A \to B$.

The ordinals ω, $\omega + 1$ and $\omega + \omega$ all have the same cardinality. The ordinals can be partitioned into classes of the same cardinality, that is the relation $\equiv$ given by $\alpha \equiv \beta$ if α and β have the same cardinality, is an equivalence relation. The least ordinal in an $\equiv$ class is called a *cardinal number* (or simply a *cardinal*). Note that ω is an ordinal (hence also an ordered set) as well as a cardinal. It is a consequence of the axiom of choice that every set is in bijective correspondence with some cardinal (that this cardinal is unique is a consequence of the definition). The cardinal bijectively equivalent to a set A is said to be the *cardinality* of A, and is denoted by $|A|$.

The cardinality of a structure $\mathcal{A} = (A, (f_i)_{i \in I}, (r_j)_{j \in J}, (c_k)_{k \in K})$ is the cardinality of the set A.

Example .6

The finite ordinals are cardinal numbers.
The smallest infinite cardinal is ω (often written $\aleph_0$).
The reals $\mathbb{R}$ have the same cardinality as the set of all subsets of $\mathbb{N}$ (denoted by $2^{\aleph_0}$).

A set is *countable* if it has cardinality at most ω (some texts will make this "equal to ω"). A set is *uncountable* if it has cardinality $\lambda > \omega$.

Example .7

- The sets $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$ are all countably infinite sets: they are all of cardinality ω.
- The sets $\wp(\mathbb{N})$, $\mathbb{R}$ and $\wp(\mathbb{R})$ are all uncountable. $\wp(\mathbb{N})$ and $\mathbb{R}$ are of the same cardinality (the continuum, often denoted c), while $|\wp(\mathbb{R})|$ is of larger cardinality still: 2^c.

The next cardinal larger than $\omega = \aleph_0$ is known as $\aleph_1$ (you can use well ordering of the ordinals to show that $\aleph_1$ exists). The assertion that c is $\aleph_1$ is known as the *continuum hypothesis*. The truth or falsity of the continuum hypothesis is known to be independent of the axioms of set theory with axiom of choice. It is not common for mathematicians to ask for the continuum hypothesis to be true.

2. Appendix: ultrafilters and ultraproducts

Here we list some useful properties of ultrafilters.

Recall that a *filter* F in a Boolean algebra $B = \langle B, \vee, \wedge, ', 0, 1\rangle$ is a nonempty subset of B with

(1) $x \in F$ and $x \leq y$ implies $y \in F$ (an upset of B)
(2) if $x \in F$ and $y \in F$ then $x \wedge y \in F$.

If $0 \notin F$ then F is a proper filter. A useful exercise encountered in MAT3DS is that the smallest filter containing F and some $a \in B$ is the upset of $\{a \wedge b \mid b \in F\}$ (this is true even for a lattice, but certainly for a Boolean algebra).

An *ultrafilter* is a maximal proper filter. The following lemma uses Zorn's Lemma to show that there is always a plentiful supply of ultrafilters.

Lemma .1

Every proper filter extends to an ultrafilter.

PROOF. Let F be a proper filter of the Boolean algebra B. We use Zorn's Lemma. Let $\mathcal{F}$ denote the family of all proper filters extending F. Then $\mathcal{F}$ is nonempty (it contains F) and is ordered by $\subseteq$. A union of a chain of filters in $\mathcal{F}$ is a proper filter in $\mathcal{F}$ (another standard exercise in MAT3DS), so by Zorn's Lemma $\mathcal{F}$ contains a maximal element U. We show that U is an ultrafilter: certainly we have chosen it to be maximal with respect to containing F and not containing 0. Consider any $a \in B\backslash U$, which exists as U is proper. Now, the smallest filter containing a and extending U is the upset V generated by $\{a \wedge b \mid b \in U\}$. But this filter extends both F and U, so the maximality of U ensures that V is *not* a proper filter. Thus U is a maximal with respect to being proper—it is an ultrafilter. $\square$

For us, most Boolean algebras will take the form $\wp(I)$ for some set I. In this instance there are certain general properties that may be observed, which are useful in constructing ultraproducts.

Let I be an index set. An ultrafilter on I is an ultrafilter of the Boolean algebra $\wp(I)$ of all subsets of I. The ultrafilter U is principal if it is the upset of some singleton subset of I: that is, if there is $i \in I$ such that $U = \{A \subseteq I \mid i \in A\}$. These ultrafilters are not very interesting to us.

Lemma .2

If I is an infinite set, then every ultrafilter on I is either principal or contains the filter all cofinite subsets of I.

PROOF. Assume that U is an ultrafilter on I and that A is a cofinite subset of I that is not contained in U. Let B denote the complement of A in I. So $B \in U$. Let $C \subseteq B$ be the smallest subset of B such that $C \in U$ (which exists as B is finite): we show that C is a singleton subset of I. Now as U is a filter, it is closed under pairwise intersections and $\varnothing \notin U$ it follows that every element of U contains C as a subset. Select any $a \in C$. Then the principal upset of $\{a\}$ in $\wp(I)$ is a proper filter containing U. As U is maximal, it follows that U is this filter. $\square$

Because the set of all cofinite subsets is itself a filter of $\wp(I)$, we may use Lemma .1 to find nonprincipal ultrafilters. In fact there are at least $2^{\aleph_0}$ (more if $|I| > \aleph_0$) different ways to extend the filter of cofinite subsets of I.

2.1. Two kinds of ultraproduct. In constructing ultraproducts, there are two every commonly encountered kinds of nonprincipal ultrafilter.

(1) Any nonprincipal ultrafilter works. this is commonly encountered when one has a family of structures, and the properties one wants hold on cofinitely many of them. For instance, the property of “having more than n-elements”, will hold cofinitely on any family of finite structures $\{A_i \mid i \in I\}$ of unbounded size. Hence an ultraproduct of the A_i with respect to any nonprincipal ultrafilter on I will itself “have more than n-elements” because all cofinite subsets of I are contained in the nonprincipal ultrafilter (by Lemma .2). As $n \in \omega$ was arbitrary, the ultraproduct will be infinite.

(2) Ultrafilters over the set of all finite subsets of some set. This is used where some desired property can be seen to hold “locally” (in “finite regions”). Typically we consider the set I consisting of all finite subsets of some set S. For any $i \in I$ (a finite subset of S), let J_s denote the set of all elements of I that extend i. So $J_i = \{j \in I \mid i \subseteq j\}$. The J_i are subsets of I, and their upset in $\wp(I)$ is a (proper) filter because $J_i \cap J_j = J_{i \cup j}$. By Lemma .1, this filter can be extended to an ultrafilter on I.

As an example of the second kind of ultrafilter in action, we use ultraproducts to prove the Compactness Theorem. Assume that Σ is a set of sentences such that each finite subset i of Σ has a model A_i. Let I denote the set of finite subsets of Σ, and construct an ultrafilter $\mathcal{U}$ as described in method 2. We claim that $\prod_{I/\mathcal{U}} A_i \models \Sigma$. Consider any $\phi \in \Sigma$. Then $J_{\{\phi\}} \in \mathcal{U}$ and $\{i \in I \mid A_i \models \phi\} \supseteq J_{\{\phi\}} \in \mathcal{U}$. Thus ϕ holds “almost everywhere” (as measured by membership in $\mathcal{U}$) and hence holds in $\prod_{I/\mathcal{U}} A_i$. As $\phi \in \Sigma$ was arbitrary it follows that $\prod_{I/\mathcal{U}} A_i \models \Sigma$.

Index