

1 Completeness

In this chapter, we're going to meet some important results about the powers and deficiencies of first-order logic. Our starting point is the completeness theorem, which shows that there is a mechanical way to check for first-order entailment. We'll see that this positive result is tightly connected to some negative results: the compactness theorem and the Löwenheim-Skolem theorems. These results concern the *size* of models, by which we mean the size of their domain. To fully appreciate their implications, I'll begin with some background about the sizes of sets, which will be needed in later chapters anyway.

1.1 Cardinalities

$\{\text{Athens}\}$ is a set with a single member: the city Athens. We say that $\{\text{Athens}\}$ has size 1. $\{\text{Athens}, \text{Berlin}\}$ has size 2. $\{\text{Athens}, \text{Berlin}, \text{Cairo}\}$ has size 3. And so on. It seems straightforward. But now consider the set \mathbb{N} of natural numbers 0, 1, 2, 3, What is its size?

The official set-theoretic term for the size of a set is *cardinality*. So $\{\text{Athens}, \text{Berlin}, \text{Cairo}\}$ has cardinality 3. Finite sets have a finite cardinalities, which are natural numbers. But infinite sets also have a cardinality. These aren't natural numbers, but numbers of a more general sort, called *cardinals*. The infinite cardinals are also known as the *alephs*, because they are conventionally written using the Hebrew letter 'ℵ' ("aleph"). For example, the cardinality of \mathbb{N} is called 'ℵ₀'. This is the smallest infinite cardinal. The next one is ℵ₁, followed by ℵ₂, and so on.

How do we determine the cardinality of an infinite set? The crucial idea goes back to Galileo and Hume, although it was only fully developed by Georg Cantor in the 19th century. Cantor, following Galileo and Hume, suggested that two sets have the same cardinality iff there is a one-to-one correspondence between their members.

To make this precise, we define the notion of a one-to-one correspondence or bijection.

Definition 1.1

A function f from a set A to a set B is a *bijection* if it satisfies the following two conditions.

- (i) For every $x, y \in A$, if $f(x) = f(y)$ then $x = y$. (Injectivity)
- (ii) For every $b \in B$ there is some $a \in A$ such that $f(a) = b$. (Surjectivity)

Intuitively, a bijection pairs up each element of A with exactly one element of B , and vice versa, so that every element of either set gets a unique partner. As a shorthand, let's say that sets A and B are *equinumerous* if there is a bijection from A to B . (In this case, there is always also a bijection from B to A .)

The Galileo-Hume-Cantor principle now says that *two sets have the same cardinality iff they are equinumerous*. Obviously, no finite set of numbers is equinumerous with \mathbb{N} . So \mathbb{N} has an infinite cardinality. Let's introduce ' \aleph_0 ' as a name for this cardinality. Using the Galileo-Hume-Cantor principle, we can determine, for any other set, whether it also has cardinality \aleph_0 .

Consider, for example, the set of odd numbers $1, 3, 5, 7, \dots$. The following function f is a bijection from \mathbb{N} to the set of odd numbers:

$$f(n) = 2n + 1.$$

The function maps 0 to 1, 1 to 3, 2 to 5, and so on. Every natural number is mapped to a unique odd number, and no odd number is left unmapped. So the set of odd numbers also has cardinality \aleph_0 .

Galileo found this paradoxical: how can there be as many odd numbers as natural numbers, given that the odd numbers are a proper subset of the natural numbers? Never mind, said Cantor: the resulting theory of cardinalities is consistent and mathematically fruitful, even if it may initially seem strange.

Sets that are equinumerous with \mathbb{N} are also called *countably infinite* or *denumerable*. A set is *countable* if it is either finite or countably infinite. The label alludes to the fact that a bijection between \mathbb{N} and another set effectively assigns a "counter" to each member of the set. For example, the above bijection between \mathbb{N} and the odd numbers

can be written as follows:

0.1
1.3
2.5
3.7
⋮

Exercise 1.1 Show that (a) the set of even natural numbers and (b) the set of integers $\dots, -2, -1, 0, 1, 2, \dots$ are both countably infinite.

Are there any *uncountable* sets? We might suspect that the set of *pairs* of natural numbers is uncountable. But not so. Cantor's *zig-zag method* shows that there is a bijection between \mathbb{N} and the set of pairs of natural numbers. We begin by arranging all pairs of natural numbers in a two-dimensional grid.

	0	1	2	3	4	...
0	(0,0)	(0,1)	(0,2)	(0,3)	(0,4)	...
1	(1,0)	(1,1)	(1,2)	(1,3)	(1,4)	...
2	(2,0)	(2,1)	(2,2)	(2,3)	(2,4)	...
3	(3,0)	(3,1)	(3,2)	(3,3)	(3,4)	...
4	(4,0)	(4,1)	(4,2)	(4,3)	(4,4)	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮

We then define a path through this grid that visits each cell exactly once. The orange arrows indicate that path. It effectively enumerates all pairs (x, y) , starting with $(0, 0)$, followed by $(0, 1)$, $(1, 0)$, $(2, 0)$, $(1, 1)$, $(0, 2)$, and so on. The enumeration amounts to a lists of all the pairs. We get a bijection to \mathbb{N} by assigning to each pair its position in the list (starting with position 0). Thus $(0, 0)$ is mapped to 0, $(0, 1)$ to 1, $(1, 0)$ to 2, and so forth.

(We can find a formula for this bijection. As we follow the arrow, the pairs *before* any given pair (x, y) comprise all the pairs on the left-to-top diagonals before (x, y) , which

have 1, 2, 3, etc. pairs, plus the y pairs on the diagonal containing (x, y) itself. The total number of pairs preceding (x, y) is therefore $(1 + 2 + \dots + (x + y)) + y$. In exercise ??, you showed that $1 + 2 + \dots + (x + y) = (x + y)(x + y + 1)/2$. So the position of any pair (x, y) in the enumeration is $(x + y)(x + y + 1)/2 + y$.)

Exercise 1.2 Show that the set of ordered triples of natural numbers is countably infinite.

But it's true that not all sets are countable. Cantor established this with another powerful technique: *diagonalization*.

Theorem 1.1: Cantor's Theorem

The set of all sets of natural numbers not countable.

Assume for reductio that the set of all sets of natural numbers is countable. This means that we can list them as S_0, S_1, S_2, \dots . Now consider the set $D = \{n \in \mathbb{N} : n \notin S_n\}$. This is a set of natural numbers, so it must be somewhere in the list. That is, there is some S_k such that $D = S_k$. Now the number k is either in D or not. If k is in D then by definition of D , k is not in S_k . This is impossible, as $D = S_k$. If k is not in D , then by definition of D , k is in S_k . Again, this is impossible. \square

We can again picture this method with a two-dimensional grid.

	0	1	2	3	4	...
S_0	✓	✗	✗	✓	✓	...
S_1	✗	✗	✗	✓	✓	...
S_2	✗	✗	✓	✗	✗	...
S_3	✓	✗	✗	✓	✗	...
S_4	✓	✗	✗	✗	✗	...
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\ddots

Each row represents a set of natural numbers. Each column stands for a number. A

checkmark indicates that the number is in the set, a cross that it is not. (So 0 is in S_0 , 1 is not in S_0 , 2 is not in S_0 , and so on.) Cantor's method now looks at the diagonal, in bold. It defines the new set D by *inverting* the diagonal, swapping crosses and checkmarks. In the picture, the inverted diagonal would begin with **$\times\checkmark\times\times\checkmark\dots$** , meaning that D does not contain 0, does contain 1, does not contain 2 and 3, does contain 4, and so on. The so-defined set D is called the *antidiagonal* of the grid. It can't be anywhere in the list. And so there can be no list of all sets of natural numbers.

Theorem 1.2 can be strengthened:

Theorem 1.2: Cantor's Theorem (general version)

For any set A , the set of subsets of A has strictly greater cardinality than A .

Proof. The proof proceeds essentially as before. Suppose for reductio that there is a bijection $f : A \rightarrow \mathcal{P}(A)$, where $\mathcal{P}(A)$ (called the *power set* of A) is the set of all subsets of A . Define the antidiagonal set $D = \{X \in A : X \notin f(X)\}$. Since D is a subset of A , it is in $\mathcal{P}(A)$. But it can't be in an output of f . For suppose it is. That is, suppose $D = f(X)$ for some $X \in A$. Either X is in D or not. If X is in D , then by definition of D , X is not in $f(X)$. But $f(X) = D$, so this is impossible. If X is not in D , then by definition of D , X is in $f(X)$. Again, this is impossible. \square

Cantor's theorem reveals a hierarchy of infinities. Starting with \mathbb{N} , we can produce larger and larger infinities by taking power sets: $\mathcal{P}(\mathbb{N})$ is larger than \mathbb{N} , $\mathcal{P}(\mathcal{P}(\mathbb{N}))$ is larger than $\mathcal{P}(\mathbb{N})$, and so on, without end. There are infinitely many infinite cardinals.

Exercise 1.3 Show that if A and B are countably infinite sets, then their union $A \cup B$ is countably infinite.

Exercise 1.4 Show that the set of finite subsets of \mathbb{N} is countably infinite.

1.2 Planning the completeness proof

In section ??, we showed that all valid sentences in propositional logic are derivable from the axiom schemata A1–A3 by Modus Ponens. This particular result is wasn't easy to foresee, but it wasn't a surprise that there is *some* mechanical way of checking if a sentence of propositional logic is valid. Models of propositional logic are essentially

finitary. A model is a truth-value assignment to the sentence letters. Since each sentence of propositional logic is composed of finitely many proposition letters, and there are only finitely many ways of assigning truth-values to these sentence letters, it is to be expected that there is a finitary algorithm for deciding whether all of these assignments make the sentence true.

In first-order logic, the situation very different. A first-order model consists of an arbitrary set D together with an interpretation of the non-logical vocabulary, where, for example, every subset of D is a possible interpretation of any monadic predicate. Even if D itself is countable, this means that there are usually uncountably many models with that domain. And D doesn't have to be countable. It can have any cardinality whatsoever. The space of first-order models is enormous. Their number is vastly greater than \aleph_0 . There is, in fact, no aleph big enough to measure the number of first-order models. It is therefore astonishing that there is a finitary, mechanical method – a proof system – by which one can determine precisely the sentences that are true across the realm of all first-order models.

The completeness theorem for first-order logic is noteworthy for other reasons as well. For example, it supports the conjecture that all intuitively valid mathematical arguments can be formalized as deductions in the first-order calculus. Let's take for granted, for the sake of the argument, that every mathematical statement can be expressed in a first-order language. Now suppose some mathematical argument *can't* be replicated in the first-order calculus. By the completeness theorem, it follows that there is a model in which the premises are true but the conclusion false. This strongly suggests that the argument wasn't valid. We'll meet further applications of the completeness theorem later. First we need to prove it.

The first completeness proof for first-order Logic was developed by Kurt Gödel in 1929. We're going to follow the simpler strategy due to Henkin (1949) that we've already used in chapter ???. Let's recall the key steps.

We want to show that whenever a sentence A is entailed by a set of sentences Γ (meaning that every model that makes all members of Γ true also makes A true), then there is a deduction of A from Γ . For short: if $\Gamma \models A$ then $\Gamma \vdash A$. The proof is by contraposition. We assume $\Gamma \not\models A$ and derive $\Gamma \not\vdash A$, as follows.

1. Assume $\Gamma \not\models A$.
2. Infer that $\Gamma \cup \{\neg A\}$ is consistent.
3. Show that $\Gamma \cup \{\neg A\}$ can be extended to a maximal consistent set Γ^+ .
4. Construct a model \mathcal{M} based on Γ^+ in which all members of Γ^+ are true.
5. Infer that $\Gamma \not\vdash A$.

A set of sentences is *consistent* if no contradiction can be derived from it. That is, Γ is consistent iff there is no sentence A such that $\Gamma \vdash A$ and $\Gamma \vdash \neg A$. Equivalently (as you've shown in exercise ??): the proof carries over from propositional logic), Γ is consistent iff $\Gamma \not\vdash \perp$.

Steps 2 and 5 are easy. The following lemma establishes step 2, and is proved exactly like its propositional counterpart, lemma ??.

Lemma 1.1

$\Gamma \not\vdash A$ iff $\Gamma \cup \{\neg A\}$ is consistent.

Proof. Suppose $\Gamma \cup \{\neg A\}$ is inconsistent. Then $\Gamma \vdash \neg\neg A$ by RAA and so $\Gamma \vdash A$ by DNE. Contraposing, this means that if $\Gamma \not\vdash A$ then $\Gamma \cup \{\neg A\}$ is consistent. Conversely, suppose $\Gamma \vdash A$. Then $\Gamma, \neg A \vdash A$ by Mon and $\Gamma, \neg A \vdash \neg A$ by Mon and Id. So $\Gamma \cup \{\neg A\}$ is inconsistent. \square

It remains to fill in steps 3 and 4: we need to show that *every consistent set of sentences is satisfiable*, meaning that there is a model in which all its members are true.

Here is the plan. Let Γ be a consistent set of sentences in some first-order language \mathcal{L} (with identity and function symbols). We'll show how one can construct a model \mathcal{M}_Γ in which all members of Γ are true.

As in chapter ??, it helps to extend Γ to a maximal consistent set Γ^+ . For example, if Γ contains $Fa \vee Gb$, it's not obvious if we our model should satisfy Fa or Gb or both. Γ^+ will contain Fa or Gb or both, so it answer this question. The atomic sentences in Γ^+ will instruct us how to build the model.

Suppose we have Fa in Γ^+ . Our model \mathcal{M} must then contain an object $\llbracket a \rrbracket^{\mathcal{M}}$ denoted by a that falls in the extension $\llbracket F \rrbracket^{\mathcal{M}}$ of F . The nature of the object $\llbracket a \rrbracket^{\mathcal{M}}$ is irrelevant. For convenience, we might choose *the individual constant* a as the object denoted by a . Then we can say that the extension of F comprises all individual constants c for which $Fc \in \Gamma^+$. (It might seem odd to have a model whose domain consists of individual constants, and in which each constant denotes itself. But noting in the definition of a first-order model prevents us from doing this.)

Unfortunately, there are some complications. Suppose Γ contains $\exists xFx$. (Which is short for $\neg\forall x\neg Fx$.) We want this sentence to be true in our model \mathcal{M} . So there must be some object in $\llbracket F \rrbracket^{\mathcal{M}}$. But if $\llbracket F \rrbracket^{\mathcal{M}}$ is defined as the set of individual constants c for which $Fc \in \Gamma^+$, there might be no such object, as there might be no constant c for which $Fc \in \Gamma^+$. We need to ensure that this doesn't happen. That is, we need to ensure that for each existential sentence $\exists xFx$ in Γ^+ , there is a "witnessing" sentence Fc in Γ^+ . The

problem is that Γ may already contain $\neg Fc$ for every individual constant c . Then we can't add the required witness without making Γ inconsistent.

To get around this problem, we'll construct Γ^+ in an extended language \mathcal{L}^+ that adds new individual constants to \mathcal{L} . We can then use these constants to ensure that every existential sentence in Γ has a witness.

Another complication arises from the presence of the identity symbol. If we let each individual constant denote itself, any identity statement involving different individual constants will be false. After all, no constant is identical to any other constant. But $a = b$ is consistent, and might be in Γ and therefore in Γ^+ . So we'll actually let each constant denote a *set* of terms: the constant a will denote the set $\{c \mid a = c \in \Gamma^+\}$.

Let's fill in the details.

1.3 The completeness proof

We'll show that every consistent set of first-order sentences is satisfiable. As we've seen above (and as we'll spell out again below), from this it is only a small step to the completeness theorem.

So let Γ be a consistent set of sentences in a first-order language \mathcal{L} . We'll extend Γ to a maximal consistent set in a language \mathcal{L}^+ that has infinitely many further individual constants. We first need to confirm that Γ is still consistent in the extended language \mathcal{L}^+ . Consistency is language-relative because the language determines which instances of the axioms are available. As we switch from \mathcal{L} to \mathcal{L}^+ , new axioms become available. We need to confirm that these new axioms don't allow deriving a contradiction from Γ .

Lemma 1.2

If Γ is a set of \mathcal{L} -sentences that is consistent within \mathcal{L} , and \mathcal{L}^+ extends \mathcal{L} by a set of new individual constants, then Γ is consistent within \mathcal{L}^+ .

Proof. Assume for contraposition that Γ is inconsistent within \mathcal{L}^+ . Then there is a deduction A_1, \dots, A_n of \perp from Γ , where each A_i is an \mathcal{L}^+ -sentence. Being finite, the deduction only uses finitely many of the new constants: call them c_1, \dots, c_k . The deduction also can't use more than finitely many of the old constants in \mathcal{L} . Since first-order languages have infinitely many constants, we can choose k distinct constants d_1, \dots, d_k from \mathcal{L} that don't occur in the deduction. Now consider the sequence of sentences A'_1, \dots, A'_n that results from A_1, \dots, A_n by replacing each c_i by d_i . It is easy to see that

- (i) if A_i is an axiom then so is A'_i ;
- (ii) if $A_i \in \Gamma$ then $A'_i \in \Gamma$ (sentences in Γ don't contain any new constants);
- (iii) if A_i follows from A_1, \dots, A_{i-1} by MP or Gen, then A'_i follows from A'_1, \dots, A'_{i-1} by MP or Gen.

So A'_1, \dots, A'_n is a deduction of \perp from Γ , showing that Γ is inconsistent within \mathcal{L} . \square

Next, we show that Γ can be extended to a maximal consistent set Γ^+ in which every existential sentence $\exists xA$ has a witness $A(x/c)$. Such sets are called *Henkin sets*.

Definition 1.2

A set of sentences Γ in a first-order language \mathcal{L} is a *Henkin set in \mathcal{L}* if it satisfies the following conditions.

- (i) Γ is consistent (within \mathcal{L}).
- (ii) For every \mathcal{L} -sentence A , either $A \in \Gamma$ or $\neg A \in \Gamma$. (Maximality)
- (iii) Whenever Γ contains a sentence of the form $\neg \forall xA$ then it also contains $\neg A(x/c)$ for some individual constant c . (Witnessing)

In the presence of (i) and (ii), (iii) is equivalent to the requirement that if Γ contains $\neg \forall x \neg A$ then it contains $A(x/c)$ for some c .

Exercise 1.5 Show that if Γ is a Henkin set and $\Gamma \vdash A$, then $A \in \Gamma$.

Exercise 1.6 Show that if Γ is a Henkin set (in first-order language with identity) then for every closed term t of the language there is an individual constant c such that $t = c$ is in Γ .

Exercise 1.7 Show that if Γ is a Henkin set then $\forall xA \in \Gamma$ iff $A(x/c) \in \Gamma$ for every individual constant c .

Lemma 1.3

Every consistent set of \mathcal{L} -sentences Γ can be extended to a Henkin set Γ^+ in any language \mathcal{L}^+ that adds infinitely many individual constants to \mathcal{L} .

Let Γ be a consistent set of \mathcal{L} -sentences. Let A_1, A_2, \dots be a list of all \mathcal{L}^+ -formulas with exactly one free variable. We define a sequence of sets $\Gamma_0, \Gamma_1, \dots$ as follows:

$$\begin{aligned}\Gamma_0 &:= \Gamma \\ \Gamma_{n+1} &:= \Gamma_n \cup \{\neg\forall x A_n \rightarrow \neg A_n(x/c_n)\},\end{aligned}$$

where x is the free variable in A_n and c_n is a new \mathcal{L}^+ -constant that does not occur in Γ_n . (There must be some such constant because \mathcal{L}^+ contains infinitely many constants that don't occur in Γ .) Let Γ' be the union $\bigcup_n \Gamma_n$ of all sets in this sequence. (That is, a sentence A is in Γ' iff it is in some Γ_n .)

We show that Γ' is consistent. Suppose not. Then there is a derivation of \perp from Γ and the “Henkin sentences” $\neg\forall x A_n \rightarrow \neg A_n(x/c_n)$. This derivation can use only finitely many of the Henkin sentences. So one of the Γ_n must be inconsistent. But we can show by induction on n that each Γ_n is consistent.

The base case, for $n = 0$, hold by assumption: Γ is consistent.

For the inductive step, assume Γ_n is consistent and suppose for reductio that Γ_{n+1} is inconsistent. By RAA, we then have

$$\Gamma_n \vdash \neg(\neg\forall x A_n \rightarrow \neg A_n(x/c_n)). \quad (1)$$

It's easy to show that if $\Gamma_n \vdash \neg(A \rightarrow B)$ then $\Gamma_n \vdash A$ and $\Gamma_n \vdash B$. (Both $\neg(A \rightarrow B) \rightarrow A$ and $\neg(A \rightarrow B) \rightarrow B$ are truth-functional tautologies.) So (1) implies

$$\Gamma_n \vdash \neg\forall x A_n, \text{ and} \quad (2)$$

$$\Gamma_n \vdash \neg\neg A_n(x/c_n). \quad (3)$$

From (3), we get $\Gamma_n \vdash A_n(x/c_n)$ by DNE. As c_n does not occur in Γ_n , Gen yields

$$\Gamma_n \vdash \forall x A_n. \quad (4)$$

(2) and (4) show that Γ_n is inconsistent, which contradicts the induction hypothesis.

Next, we extend Γ' to a maximal consistent set Γ^+ . The construction follows the proof of Lindenbaum's Lemma (lemma ??). Let S_1, S_2, \dots be a list of all \mathcal{L}^+ -sentences. Start-

ing with Γ' , We define another sequence of sets $\Gamma'_0, \Gamma'_1, \dots$:

$$\begin{aligned}\Gamma'_0 &:= \Gamma' \\ \Gamma'_{n+1} &:= \begin{cases} \Gamma'_n \cup \{S_n\} & \text{if } \Gamma'_n \cup \{S_n\} \text{ is consistent,} \\ \Gamma'_n \cup \{\neg S_n\} & \text{otherwise.} \end{cases}\end{aligned}$$

We show by induction that each Γ'_n is consistent. The *base case*, for $n = 0$, holds because Γ'_0 is Γ' , which we've just shown to be consistent. For the *inductive step*, assume that a set Γ'_n in the list is consistent. By lemma ??, either $\Gamma'_n \cup \{S_n\}$ or $\Gamma'_n \cup \{\neg S_n\}$ is consistent. If $\Gamma'_n \cup \{S_n\}$ is consistent, then Γ'_{n+1} is $\Gamma'_n \cup \{S_n\}$ (by construction), so Γ'_{n+1} is consistent. If $\Gamma'_n \cup \{S_n\}$ is not consistent, then Γ'_{n+1} is $\Gamma'_n \cup \{\neg S_n\}$, so again Γ'_{n+1} is consistent.

Let Γ^+ be the union $\bigcup_n \Gamma'_n$ of $\Gamma'_0, \Gamma'_1, \dots$. Evidently, Γ is a subset of Γ^+ . We show that Γ^+ is a Henkin set.

Maximality holds because for each sentence S_n , one of S_n or $\neg S_n$ is in Γ'_{n+1} and therefore in Γ^+ .

To show consistency, suppose that Γ^+ is inconsistent. Then there are sentences A_1, \dots, A_n in Γ^+ from which \perp is deducible. All of these sentences have to occur somewhere on the list S_1, S_2, \dots . Let S_j be the first sentence from S_1, S_2, \dots that occurs after all the A_1, \dots, A_n . Then all A_1, \dots, A_n are in Γ'_j . So Γ'_j is inconsistent. But we've seen that all of the Γ'_n are consistent.

It remains to show that Γ^+ has the witnessing property: whenever $\neg \forall x A$ is in Γ^+ then Γ^+ contains a corresponding sentence $\neg A(x/c)$. Let A be any formula in which x is the only free variable. By construction, Γ' contains $\neg \forall x A \rightarrow \neg A(x/c)$, for some constant c . Since Γ^+ extends Γ' , it also contains this sentence. So if Γ^+ contains $\neg \forall x A$ then it contains $\neg A(x/c)$, by lemma ??. \square

Next, we show how to read off a model from a Henkin set. The model's domain will consist of sets of closed terms, so that we can stipulate that each constant c denotes the set of all terms t for which $c = t$ is in the Henkin set. Let's have a closer look at these sets.

Definition 1.3

A binary relation R on some domain D is an *equivalence relation* if it is

- (i) *reflexive*: for every $x \in D$, xRx ;
- (ii) *symmetric*: for every $x, y \in D$, if xRy then yRx ; and
- (iii) *transitive*: for every $x, y, z \in D$, if xRy and yRz then xRz .

Lemma 1.4

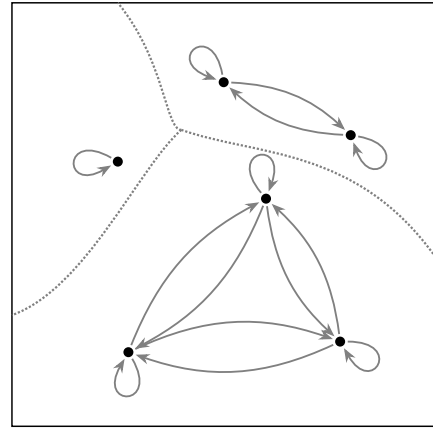
If Γ is a Henkin set then the relation R that holds between \mathcal{L}^+ -terms t, s iff $t = s \in \Gamma$ is an equivalence relation.

Proof. By A6, $\vdash t = t$. So $t = t \in \Gamma$ by exercise 1.5. Symmetry and transitivity follow similarly from exercise ??.

□

An equivalence relation *partitions* the domain over which it is defined into distinct cells so that within each cell, all objects stand in the relation to one another. These cells are called *equivalence classes*. If R is an equivalence relation and x an object in the domain, we write ' $[x]_R$ ' for the equivalence class of R that contains x . That is, $[x]_R = \{y \in D \mid xRy\}$. If the relation R is clear from the context, we may simply write ' $[x]$ '.

When we're talking about a Henkin set Γ , the relevant equivalence relation is the one defined in lemma 1.4. So ' $[t]$ ' denotes the set of all terms s for which $t = s \in \Gamma$.



If Γ is a Henkin set, the model \mathcal{M} that we'll construct from Γ will have as its domain the set of all equivalence classes $[t]$, where t is a closed term in the language of Γ . By exercise 1.6, $[t]$ always contains an individual constant. So we can also say that $D = \{[c] \mid c \text{ is an individual constant}\}$.

We'll stipulate that the interpretation function I of \mathcal{M} assigns $[c]$ to each individual constant c . So we'll have $\llbracket c \rrbracket^{\mathcal{M}} = [c]$. We'll extend this to function terms, so that $\llbracket f(c) \rrbracket^{\mathcal{M}} = [f(c)]$. But we can't *directly* stipulate this, because interpretation functions don't assign denotations to complex terms. We have to interpret f as denoting a function on D . The denotation of $f(c)$ is then the denotation of f applied to the denotation of

c . Since the denotation of c is $[c]$, we want the denotation of f to be a function that returns $[f(c)]$ for input $[c]$. So we'll stipulate that for any one-place function symbol f and constant c , $I(f)$ is the function that maps $[c]$ to $[f(c)]$.

This kind of stipulation can go wrong. Suppose $[c]$ contains two constants c and d , and $[f(c)] \neq [f(d)]$. Our stipulation entails that $I(f)$ returns $[f(c)]$ for input $[c]$ and $[f(d)]$ for input $[d]$. But a function can't return two different values for the same input. we must show that this problem can never arise.

A similar issue arises for the interpretation of predicates. To ensure that Ft is in Γ iff $\mathcal{M} \models Ft$, we'll stipulate that $I(F)$ is the set of all $[t]$ for which $Ft \in \Gamma$. This set isn't well-defined if there are cases where $[t]$ contains two terms t and s for which $Ft \in \Gamma$ but $Fs \notin \Gamma$.

The following lemma shows that neither problem can arise.

Lemma 1.5

If f is an n -ary function symbol, P an n -ary predicate symbol, and $s_1, \dots, s_n, t_1, \dots, t_n$ are terms such that $[s_1] = [t_1], \dots, [s_n] = [t_n]$, then

- (i) $[f(s_1, \dots, s_n)] = [f(t_1, \dots, t_n)]$;
- (ii) $P(s_1, \dots, s_n) \in \Gamma$ iff $P(t_1, \dots, t_n) \in \Gamma$.

Proof. Assume $[s_i] = [t_i]$, for $i = 1, 2, \dots, n$. So Γ contains $s_i = t_i$, for each $i = 1, 2, \dots, n$.

(i). By A6 (and exercise 1.5), $f(s_1, \dots, s_n) = f(s_1, \dots, s_n)$ is in Γ . By n instances of A7 and MP, it follows that $f(s_1, \dots, s_n) = f(t_1, \dots, t_n)$ is in Γ , which entails that $[f(s_1, \dots, s_n)] = [f(t_1, \dots, t_n)]$.

(ii). Assume $P(s_1, \dots, s_n)$ is in Γ . By n instances of A7 and MP, $P(t_1, \dots, t_n)$ is in Γ as well. The converse holds by the same reasoning. \square

Lemma 1.5 ensures that the following definition is legitimate.

Definition 1.4

Let Γ be a Henkin set in a language \mathcal{L} . The *Henkin model* \mathcal{M}_Γ of Γ is defined as follows.

The domain D of \mathcal{M}_Γ is the set $\{[c] \mid c \text{ is an individual constant of } \mathcal{L}\}$, where $[c]$ is the set of all closed \mathcal{L} -terms t for which $c = t \in \Gamma$.

The interpretation function I of \mathcal{M}_Γ maps

- (i) each individual constant c to $[c]$,
- (ii) each n -ary function symbol f to the function on D that maps any n -tuple $[t_1], \dots, [t_n]$ to $[f(t_1, \dots, t_n)]$,
- (iii) each (non-logical) n -ary predicate symbol P to the set of all n -tuples $([t_1], \dots, [t_n])$ such that $P(t_1, \dots, t_n) \in \Gamma$.

Now remember what we're trying to achieve. We want to show that every consistent set of sentences is satisfiable. We've shown in lemma 1.3 that every such set can be extended to a Henkin set. Definition 1.4 tells us how to construct a model from this Henkin set. It remains to show that all members of the Henkin set (and therefore all members of the original set) are true in this model.

We'll need the following two facts.

Lemma 1.6

If \mathcal{M} is a Henkin model and t a closed term then $\llbracket t \rrbracket^{\mathcal{M}} = [t]$.

Proof. The proof is by induction on the complexity of t . The base case is covered by clause (i) in definition 1.4. So let t be $f(t_1, \dots, t_n)$. Then

$$\begin{aligned}
 \llbracket f(t_1, \dots, t_n) \rrbracket^{\mathcal{M}} &= \llbracket f \rrbracket^{\mathcal{M}}(\llbracket t_1 \rrbracket^{\mathcal{M}}, \dots, \llbracket t_n \rrbracket^{\mathcal{M}}) && \text{by def. ??} \\
 &= \llbracket f \rrbracket^{\mathcal{M}}([t_1], \dots, [t_n]) && \text{by ind. hyp.} \\
 &= [f(t_1, \dots, t_n)] && \text{by def. 1.4. } \square
 \end{aligned}$$

Lemma 1.7

If \mathcal{M} is a Henkin model and B a formula then $\mathcal{M} \models \forall x B$ iff $\mathcal{M} \models B(x/c)$ for all individual constants c .

Proof. We first show that if $\mathcal{M}^{d:[c]}$ is just like \mathcal{M} except that it assigns $[c]$ to d , and d does not occur in B , then

$$\mathcal{M}^{d:[c]} \models B(x/d) \text{ iff } \mathcal{M} \models B(x/c). \quad (1)$$

There are two cases to consider. If d is c then $\mathcal{M}^{d:[c]}$ is \mathcal{M} and (1) holds trivially. If d is a constant other than c , then d does not occur in $B(x/c)$. So $\mathcal{M}^{d:[c]}$ and \mathcal{M} agree on the interpretation of all symbols in $B(x/c)$, and so (1) holds by lemma ??.

Now, by definition ??, $\mathcal{M} \models \forall x B$ iff $\mathcal{M}' \models B(x/d)$ for every model \mathcal{M}' that differs from \mathcal{M} at most in the object assigned to d , where d is the alphabetically first constant that does not occur in B . Since each such \mathcal{M}' has the same domain as \mathcal{M} , the set of \mathcal{M}' is precisely the set of variants $\mathcal{M}^{d:[c]}$ of \mathcal{M} . So $\mathcal{M} \models \forall x B$ iff $\mathcal{M}^{d:[c]} \models B(x/d)$ for every constant c , which, by (1), is the case iff $\mathcal{M} \models B(x/c)$ for every constant c . \square

Lemma 1.8: Truth Lemma

If \mathcal{M} is the Henkin model of a Henkin set Γ in a language \mathcal{L} , then $\mathcal{M} \models A$ iff $A \in \Gamma$, for every \mathcal{L} -sentence A .

By induction on the complexity of A .

Base case: A is atomic. There are two subcases.

Assume that A is an identity sentence $s = t$. Then $\mathcal{M} \models s = t$ iff $\llbracket s \rrbracket^{\mathcal{M}} = \llbracket t \rrbracket^{\mathcal{M}}$ by definition ??, iff $[s] = [t]$ by lemma ??, iff $s = t \in \Gamma$, by definition of the equivalence classes.

Assume next that A is an atomic sentence $P(t_1, \dots, t_n)$, where P is non-logical. Then $\mathcal{M} \models P(t_1, \dots, t_n)$ iff $(\llbracket t_1 \rrbracket^{\mathcal{M}}, \dots, \llbracket t_n \rrbracket^{\mathcal{M}}) \in \llbracket P \rrbracket^{\mathcal{M}}$ by definition ??, iff $([t_1], \dots, [t_n]) \in \llbracket P \rrbracket^{\mathcal{M}}$ by lemma ??, iff $P(t_1, \dots, t_n) \in \Gamma$ by definition 1.4.

Inductive step: A is composed of other sentences. We have three subcases. The first two, where A is $\neg B$ or $B \rightarrow C$, are easy. I'll skip them. (But see the proof of lemma ?? if you want to see the details. The proof is essentially the same.)

For the final subcase, assume that A is $\forall x B$, for some formula B . We have $\mathcal{M} \models \forall x B$ iff $\mathcal{M} \models B(x/c)$ for every constant c by lemma 1.7, iff $B(x/c) \in \Gamma$ for every constant c by induction hypothesis, iff $\forall x B \in \Gamma$ by exercise 1.7. \square

Let's put the pieces together

Theorem 1.3: Completeness of the first-order calculus (Gödel 1929)

If $\Gamma \models A$ then $\Gamma \vdash A$.

Proof. We argue by contraposition. Assume $\Gamma \not\models A$. Then $\Gamma \cup \{\neg A\}$ is consistent, by lemma ???. By lemma 1.3, $\Gamma \cup \{\neg A\}$ can be extended to a Henkin set Γ^+ in an extended language \mathcal{L}^+ . By the Truth Lemma, the Henkin model \mathcal{M}_{Γ^+} of Γ^+ satisfies every sentence in Γ^+ , and therefore every sentence in $\Gamma \cup \{\neg A\}$. So there is a model that satisfies all sets in Γ but doesn't satisfy A . So $\Gamma \not\models A$. \square

Technically, the model \mathcal{M}_{Γ^+} that figures in this proof is not an model for the original language \mathcal{L} of Γ and A , because it also interprets the added individual constants. We can get an \mathcal{L} -model that falsifies $\Gamma \models A$ by restricting the interpretation function of \mathcal{M}_{Γ^+} to \mathcal{L} -constants, without changing the domain. (This is called the *reduct* of \mathcal{M}_{Γ^+} to \mathcal{L} .)

Exercise 1.8 Assume that $\Gamma_0, \Gamma_1, \dots$ are consistent sets of sentences in a first-order language \mathcal{L} , and that each Γ_i is a subset of Γ_{i+1} . Show that their union $\bigcup_i \Gamma_i$ is consistent.

1.4 Unintended Models

Now return to first-order logic. Suppose we want to formalize reasoning about, say, the natural numbers, in a suitable first-order language \mathcal{L} . This language might have a constant '0', function symbols '+' and '×', and so on. We can say things like

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

or

$$\forall x(x \times 0 = x).$$

Intuitively, the first sentence is true and the second false. But wait! Officially, truth and falsity are defined only relative to a model. It is easy to construct a model in which the second sentence is true.

Exercise 1.9 Do it.

Logic abstracts away from the meaning of non-logical expressions. '0', '+', '×' are

non-logical. Neither of the two sentences is *logically true*, true in all models.

Still, there's a good sense in which the first sentence is true and the second false. When we use the language, we have a particular interpretation in mind. We use '0' for the number zero, '+' for addition, '×' for multiplication. We also assume that the quantifiers range over the set of natural numbers. So understood, the first sentence says that adding zero to any number yields that same number, and the second says that multiplying any number by zero yields that same number.

We'll say that *intended model* of the language \mathcal{L} is the model \mathcal{N} whose domain is the set of natural numbers, and in which '0' refers to zero, '+' to addition, and '×' to multiplication.

We might hope that we can rule out non-intended models by laying down some postulates in the language. For example, we can rule out models with only one object by the following postulate:

$$\exists x \exists y (x \neq y).$$

We can similarly find a postulate that rules out models with only two objects. And so on.

there are at least two objects.

of course, there will always be non-intended models because the nature of the elements in D doesn't matter. That is, given a model, I can define a new model in which the number 7 is swapped by Caesar etc. The best we can hope for is to narrow the models down to isomorphism.

Define Isomorphism.

Define elementary equivalence.

Proof sketch: isomorphic models are elementarily equivalent.

If we had the converse – which we don't – we'd know that non-isomorphic models can always be distinguished by some sentence. So that we can always rule out unintended non-isomorphic models by laying down another postulate, although we might need infinitely many postulates, but never mind that for now. In fact, for present purposes we can consider whether it would suffice to lay down as postulate every sentence in the language that's true in the intended model.

Definition: If M is a model for a language L , the /theory of M / is the set of L -sentences A with $M \models A$.

(Note that elementary equivalent models are precisely models with the same theory.)

Consider the language of arithmetic. Its intended model is called \mathcal{A} . Are all models of the theory of arithmetic isomorphic to the intended model?

No. We can show this by compactness. First, extend the language by a new constant c . Add $c \neq 0$ and $c \neq 1$ etc. Every finite subset is satisfiable in a model based on \mathcal{A} , interpreting c as a sufficiently large number. So the whole theory is satisfiable in some model M . This model must have an extra element besides the numbers. Now discard the interpretation of c from the model. (This is called the *reduct* of M to the original language.) The resulting model is still a model of $Th(\mathcal{A})$. But it has an extra element. It is called a *non-standard model of arithmetic*. And yet it is elementarily equivalent to \mathcal{A} ! There is no sentence in the original language that distinguishes the two models.

Exercise 1.10 Show by compactness that if Γ is a set of first-order sentences with arbitrarily large finite models, then Γ has an infinite model.

The Skolem-Loewenheim theorem makes things much worse.

Are these sentences true?

The second sentence

Consider $\exists x \exists y (x \neq y)$. Any model of this must have at least size 2. Its negation only has models with size 1.

Exercise 1.11 Find a sentence with model size at least 3 and at most 2.

Here is a set of sentences with only infinite models:

xxxx [e.g. BBJ 138]

Corollary: Every satisfiable set of sentences has a model whose domain is a set of natural numbers.

Here are some easy consequences of the completeness theorem.

Definition 1.5

A set of first-order sentences is *satisfiable* if it has a model.

Theorem 1.4: Compactness (Gödel 1929)

If every finite subset of a set is satisfiable then the set itself is satisfiable.

Proof. Let Γ be a set of first-order sentences. Assume for contraposition that Γ is not satisfiable. So $\Gamma \models \perp$. By the Compactness Theorem, it follows that $\Gamma \vdash \perp$: there is a deduction of \perp from Γ . This deduction can only use finitely many sentences from Γ . So there is a finite subset Γ_0 of Γ for which $\Gamma_0 \vdash \perp$. By the Soundness Theorem, $\Gamma_0 \models \perp$.

Corollary 1.1

If $\Gamma \models A$ then there is a finite subset Γ_0 of Γ such that $\Gamma_0 \models A$.

Theorem 1.5: (Downward) Löwenheim-Skolem

If a set Γ of countable first-order sentences has a model, then it has a countable model.

The Henkin model M_Γ for Γ is countable, as its domain consists of equivalence classes of closed terms in the language, of which (by assumption) there are countably many.

Theorem 1.6: Upward Löwenheim-Skolem

If a set Γ of first-order sentences has a countably infinite model, then it has a model at least as large as any larger cardinality.

Let κ be any uncountable cardinal. Expand the language of Γ by κ new individual constants. For each pair of these constants c_i and c_j , add the sentence $c_i \neq c_j$ to Γ , thereby creating the set Γ^+ . If Γ has a countably infinite model M , then all finite subsets of Γ^+ are satisfied in the same model, with the (finitely many) new constants in the set interpreted as distinct objects in the domain of M . By the Compactness Theorem, Γ^+ has a model M^+ . All the κ new constants denote distinct objects in this model. So the model has at least κ objects. If we restrict M^+ to the original language of Γ , we get a model of Γ with at least κ objects.

The downward theorem, as stated, was first proved by Thoralf Skolem in 1920, building on an incomplete 1915 sketch by Leopold Löwenheim. The upward theorem was proved by Alfred Tarski in 1935, who also showed that one can find a model whose cardinality is exactly κ .

Let's think about what these theorems mean, starting with compactness. It is easy to

find cases where a conclusion is entailed by infinitely many premises, but not by any finite subset of those premises. For example, from the premises:

I like the number 0;
I like the number 1;
I like the number 2;
...

and so on, for all natural numbers, one can infer

I like all natural numbers.

But this conclusion is not entailed by any finite subset of the premises. Similarly, one can infer that some person X is not an ancestor of some person Y from the assumptions that

X is not a parent of Y;
X is not a parent of a parent of Y;
X is not a parent of a parent of a parent of Y;
...

and so on, but not from any finite subset of those assumptions.

The compactness theorem shows that this kind of situation never arises for *first-order logical entailment*. The two inferences just described are valid, but their validity depends on the meaning of non-logical terms: ‘1’, ‘2’, ‘3’, ..., and ‘number’ in the first case; ‘parent’ and ‘ancestor’ in the second. From the compactness of first-order logic, it immediately follows that one can’t define these concepts using only the connectives, the first-order quantifiers, and identity.

What if we declare the relevant terms to be logical? Let’s see what this would mean for a formalized version of the first inference. Take a first-order language with infinitely many constants ‘0’, ‘1’, ‘2’, ..., and predicate symbols ‘N’ and ‘L’. The premises can then be formulated as: $L0, L1, L2, \dots$. The conclusion is: $\forall x(Nx \rightarrow Lx)$. Obviously, the conclusion is not entailed by the premises in the ordinary sense of first-order logical entailment: we can easily find an interpretation of ‘0’, ‘1’, ‘2’, ..., ‘N’, and ‘L’ that makes the premises true and the conclusion false. But now the idea is that we can define a new kind of entailment – call it “arithmetical entailment” – by fixing the interpretation of the ‘0’, ‘1’, ‘2’, ..., and ‘N’: ‘0’ must denote the number zero, ‘1’ the number one, ..., and the extension of ‘N’ must be the set of natural numbers. Some premises Γ “arithmetically

entail” a conclusion A if any model of Γ that conforms to these rules is a model of A . Now $\{L1, L2, \dots\}$ arithmetically entails $\forall x(Nx \rightarrow Lx)$. But no finite subset of $\{L0, L1, \dots\}$ arithmetically entails $\forall x(Nx \rightarrow Lx)$. So arithmetical entailment is not compact. Since compactness is a consequence of soundness and completeness, it follows that there can’t be a sound and complete proof system for arithmetical validity. That is, there is no way to spell out axioms and rules of inference that define a proof relation $\vdash_{\mathcal{A}}$ so that $\Gamma \vdash_{\mathcal{A}} A$ iff Γ arithmetically entails A – as long as proofs remain finite.

Exercise 1.12 Explain how the argument can be run with a single non-arithmetical constant c instead of the predicate symbol L .

Exercise 1.13 Consider the “proof system” that has all arithmetical truths as axioms, in addition to the axioms and rules of first-order predicate calculus. Can all arithmetical truths be proved in this system? Why doesn’t it follow that the system is complete?

Thus compactness reveals a deep expressive limitation of first-order logic. It shows, for example, that one cannot express the concept of finiteness. That is, there is no first-order sentence that is true in all and only the models with a finite domain. [Explain.]

Exercise 1.14 Can you find a sentence that is only true in infinite models? Can you find a sentence that is true in all and only the infinite models?

Here’s another example, from the theory of orders.

A relation \leq on a set D is a *linear order* if it is

- *reflexive*: for every x in D , $x \leq x$; - *transitive*: if $x \leq y$ and $y \leq z$ then $x \leq z$; - *anti-symmetric*: if $x \leq y$ and $y \leq x$ then $x = y$; - *total*: for every x, y in D , either $x \leq y$ or $y \leq x$.

A linear order \leq is a *well-order* if every non-empty subset of D has a \leq -least element. That is, every non-empty subset S of D has some element $x \in S$ such that for every $y \in S$, $x \leq y$. For example, the less-than-or-equal relation \leq on the natural numbers \mathbb{N} is a well-order: any set of natural numbers has a least element.

Obviously, the conditions of reflexivity, anti-symmetry, and totality can be expressed in first-order logic. So we one can write down a set of first-order sentences with a predicate symbol ‘ \leq ’ that is true in a model iff the model interprets ‘ \leq ’ as a linear order on

its domain. We say that these sentences *axiomatize* the concept of a linear order. Can we also axiomatize the concept of a well-order?

The answer is no, because of compactness. Suppose for reductio that some set of first-order sentences Γ is true in a model iff it interprets ' \leq ' as a well-order. In particular, Γ is true in the model M whose domain are the natural numbers \mathbb{N} and that interprets ' \leq ' as the less-than-or-equal relation on \mathbb{N} . Now let c_1, c_2, \dots be a list of individual constants. Construct an extension Γ^+ of Γ by adding, for each c_i , the sentence

$$c_{i+1} \leq c_i \wedge c_{i+1} \neq c_i,$$

which we may abbreviate as

$$c_{i+1} < c_i.$$

Every finite subset of Γ^+ is still satisfiable: it is true in the model M just mentioned. By compactness, Γ^+ has a model. But that model must have an infinite descending chain $\dots c_{i+1} < c_i \dots < c_1$. So it is not a well-order.

Compactness is a fundamental theorem in model theory. Like Gödel, we've derived it from completeness, but note that it has nothing to do with proofs. It's a purely model-theoretic result, that can also be proved without any reference to proofs.

In fact, we can use a version of our Henkin proof of Completeness. Call a set Γ *finitely satisfiable* if every finite subset of Γ is satisfiable. Assume that Γ is finitely satisfiable. We would show that Γ can be extended to a maximal finitely satisfiable set Γ^+ in a language \mathcal{L}^+ with new constant symbols. The Henkin model of Γ^+ is then a model of Γ .

Exercise 1.15 A proof system \vdash is called *weakly complete* if it can prove all validities: if $\models A$ then $\vdash A$. In the previous section, we've shown that first-order logic is *strongly complete*: if $\Gamma \models A$ then $\Gamma \vdash A$. Strong completeness obviously implies weak completeness. We've seen that it also implies compactness. Show that, conversely, weak completeness and compactness together imply strong completeness, given that \vdash has the properties ID and MP.

Exercise 1.16 Compactness is an immediate consequence of Completeness. Can you argue in the other direction from Compactness to Completeness?