Chapter 6
Completeness Theorem: FO Validity is r.e.

Gödel proved in 1929 his first famous theorem that there is a formal proof system that can prove every valid formula in FOL. As the formal proof system one can choose a variety of proof systems (Gödel showed it for one proposed by Hilbert and Ackermann). A proof system is a formal set of rules of writing down a finite sequence (called a proof) that establishes the validity of a formula/sentence. In fact, a stronger statement is proved (let's call this the strong completeness theorem): there is a formal proof system such that for any set of axioms A, the formal system (incorporating axioms A) can prove any formula/sentence that is semantically entailed by A. In other words, the system can prove any sentence φ where φ holds in *all* models that satisfy the axioms A.

The above result is remarkable. It basically shows that any theorem that can be stated in FOL has a proof. Not only that, for any class of axiomatizable structures, the class of valid FO-formulatable theorems over such structures always has a proof. For instance, take the class of groups— they can be axiomatized using a few axioms, as we saw in Chapter 1. Consequently, every first-order formulatable theorem over groups has a proof.

Given a set of formulae/sentences A, the validity problem for the theory of A is to determine whether, given a formula/sentence φ, whether every model and every interpretation under which A holds also satisfies φ.

Connection to computability

In this book/course, we won't be studying proof systems, and hence won't prove Gödel's completeness theorem. However, we will prove essentially Gödel's completeness theorem, but where we replace proofs with *computation*.

Consider a set of axioms A which is a decidable set (i.e., it is either empty or finite or infinite where a TM can check whether a given sentence is an axiom or not). Then Gödel's completeness theorem says that every logically entailed theorem has a proof. Proofs are generally *finite* objects— they are typically finite sequences over some signature, that closely follow some set of allowed rules, incorporating the axioms when needed, in order to prove a theorem. Whatever the proof system is, it is always true that *checking* whether such a sequence encodes a correct proof is a decidable problem. Consequently, it is easy to see that validity with respect to the axioms is a problem solvable in r.e.. This is because a Turing machine can enumerate all possible finite proofs, systematically, checking if any of them prove a given theorem, and halt if it does. So Gödel's theorem can be seen as saying that the problem of checking validity wrt any recursive set of axioms is recursively enumerable.

Our goal in this chapter is hence to prove this version of completeness. For every formula/sentence, when the TM finds that the sentence is a theorem in the theory of the given axioms, the computation itself can be viewed as a proof of the theorem.

Outline of Proof

The procedure we are going to outline is not entirely the usual one found in standard textbooks, and has a more computational flavor. As we will see, it can also be automated to some extent (we can build an r.e. procedure using calls to an SMT solver that decides the quantifier-free theory of equality).

The rough outline is as follows. We fix a countable signature. We are given a countable decidable set of axioms A and we consider the problem of proving validity of a FOL formula ψ.

1. Our procedure will work through *refutation*— we will show that $\psi' = \neg\psi$ is not satisfiable in any model satisfying the axioms. In other words, we want to show there is no model satisfying $A \cup \{\psi'\}$.
2. We first show that formulas can be translated to equivalent formulas in *prenex normal form*. Then we show that we can convert the negated formula to an *equisatisfiable* formula ψ'' over an expanded signature which has only universal quantification, and is of the form

$$\psi'' : \forall\overline{x}.\ \varphi$$

   This process is called *Skolemization*.
3. We then show *Herbrand's Theorem* for such sets of universally quantified formulas, which roughly says that if the axioms and the formula is satisfiable, then they satisfiable in a universe that is composed of only ground terms over the signature modulo a congruence.
4. The above result will show that the universally quantified formula will be *unsatisfiable* iff replacing variables with all possible terms, which gives an infinite set of formulas, is an unsatisfiable set.
5. We will then use compactness of propositional logic to argue why this instantiated infinite set is unsatisfiable iff there is a finite subset of it that is unsatisfiable.
6. The above gives our r.e. procedure: negate the formula, Skolemize the axioms and formula, and instantiate systematically the formulas by a growing set of terms and

check whether the the resulting set is unsatisfiable. Any instantiation procedure that dovetails between the axioms and term instantiation so that all axioms are instantiated for all terms eventually will do. Each level of instantiation gives a set of quantifier-free formulae in the theory of equality, which is decidable. The algorithm will halt only if it finds that there is some level where the instantiated set is unsatisfiable.

We first show Step 2: Skolemization. Then we prove Herbrand's theorem. We then will use compactness to argue unsatisfiability can be proved using only a finite set of terms. And finally give the r.e. procedure.

6.1 Prenex Normal Form

We assume that the formulae/sentences we are considering for validity/satisfiability have first been convereted to prenex normal form, i.e., to the form:

$$Q_1 x_1 .\ Q_2 x_2\ \ldots Q_n x_n .\varphi$$

where φ is quantifier free, and furthermore, where no variable repeats (for every $i \neq j$, $x_i \neq x_j$).

We refer the reader to a standard text that shows that any formula in FOL can be converted to an equivalent formula in prenex form.

6.2 Skolemization / Herbrandization

Recall that for validity, pure universal quantification was easy to handle (we showed decidability in the last chapter). Hence, for satisfiability, pure existential quantification is easy to handle.

We can in fact eliminate all existential quantification in a satisfiability problem easily.

Consider a formula of the form

$$\psi : \forall x_1, \ldots x_n . \exists y . \varphi(x_1, \ldots, x_n, y)$$

where φ is an arbitrary formula (can have quantifiers). We will show that there is an equisatisfiable formula (over an expanded signature) where we essentially remove the quantified variable x.

The formula roughly says:

For every valuation of $x_1, \ldots, x_n$, there exists a value for y such that φ holds.

Assume there is some model that satisfies the above property. Then for every sequence of values of $x_1, \ldots, x_n$, since there is an element y in the universe such that φ holds, we can fix a particular choice of this element y using a new *function* f. This function in the model takes a tuple of values $(v_1, \ldots, v_n) \in \mathcal{U}^n$ (standing for a valuation of $x_1, \ldots, x_n$, respectively) and maps it to an element in $\mathcal{U}$. Now, instead of saying that there is a value y that satisfies φ, we can instead say that choosing y to be $f(x_1, \ldots, x_n)$ satisfies φ.

More precisely, we can write instead the formula:

$$\psi' : \forall x_1, \ldots x_n.\ \varphi(x_1, \ldots, x_n, f(x_1, \ldots, x_n) \ / \ y)$$

In other words, we remove the existential quantification on y, and instead replace y in φ with $f(x_1, \ldots x_n)$. Here, f is a *new* function symbol introduced specifically for this quantification of y.[1].

It should be clear that the original formula ψ is satisfiable over a signature Σ iff the above formula ψ' is satisfiable over the signature $\Sigma \cup \{f\}$, where f is a function symbol not occurring in Σ. In the forward direction, if ψ is satisfiable in a model M, we construct a model M' over the expanded signature that extends M by interpreting f on an n-tuple of values to some value y that makes φ true when $x_1, \ldots, x_n$ are evaluated as the n-tuple. This extended model M' will satisfy ψ'. In the reverse direction, if there is a model M' for ψ', we can show that the model M which is the same as M' except with the interpretation of f erased, satisfies ψ: for every valuation of $x_1, \ldots, x_n$, if we choose choose y to be $f(x_1, \ldots, x_n)$, then we are guaranteed to satisfy φ.

When a formula has no universal quantification preceding an existential quantifier, the above works too, except that now the function takes *no arguments*, i.e., it is a 0-ary function. A function that takes no arguments and returns an element is essentially a constant. So we can replace such an existentially quantified variable with a new constant symbol.

More precisely, we can show that for any formula $\exists x.\varphi$ over a signature Σ, the formula $\varphi[c/x]$, where c is a new constant symbol that is not in Σ, is equisatisfiable.

The following lemma captures the above:

Lemma 6.1 *For any formula $\psi : \forall x_1, \ldots x_n.\exists y.\varphi(\overline{z}, x_1, \ldots, x_n, y)$ over a signature Σ, let f be a function symbol not in Σ, and let*

$$\psi' : \forall x_1, \ldots x_n.\ \varphi(\overline{z}, x_1, \ldots, x_n, f(x_1, \ldots, x_n) \ / \ y)$$

over the signature $\Sigma \cup \{f\}$, where the arity of f is n. Then ψ and ψ' are equisatisfiable.

Also, for any formula $\psi : \exists y.\varphi(\overline{z}, y)$ over a signature Σ, let c be a constant symbol not in Σ, and let

$$\psi' : \varphi(\overline{z}, c \ / \ y)$$

over the signature $\Sigma \cup \{c\}$. Then ψ and ψ' are equisatisfiable.

Example 6.2 For example, consider the formula

[1] Some readers may wonder if we are using the axiom of choice here; we are.

$$\forall x.\exists y.R(x, y)$$

which says that every element x is R-related to *some element* The above is equisatisfiable to the formula

$$\forall x.R(x, f(x))$$

Intuitively, the function f chooses one of the (potentially several) elements x is R-related to. Such a function exists iff every x is indeed R-related to some element.

We can now apply the above procedure of eliminating existential quantifiers repeatedly to a formula in prenex rectified normal form in order construct a purely universally quantified formula that is equisatisfiable.

Let us call formulas that are purely universally quantified *universal formulas*.

Herbrandization:

The above also shows that we can take any formula ψ and convert it into an equi-valid formula of the form $\exists x_1. \ldots, \exists x_n \varphi$ over an expanded signature. We can simply take $\neg psi$, Skolemize it to derive an equi-satisfiable formula over an expanded signature, and then take its negation, to get a formula with purely existential quantification that is equi-valid to ψ. This process of getting equi-valid formulas with existential quantification only is called Herbrandization.

6.3 Herbrand's theorem

One of the the first hurdles for solving satisfiability or proving unsatisfiability is to figure out what the *universe* for a formula might be. Clearly, we need elements to represent constants as they are terms. And we need elements to represent terms formed by applying functions (any number of times) to terms. But do we need more? Can a formula/sentence talk about elements that are are not *accessible* by using functions that involve constants?

Let us define accessible elements more formally. A *ground term* is a term without variables (it it built only using functions and constants). Let M be a model. An element e in the universe of M is said to be *accessible* if there is a ground term t such that t evaluates to the element e in the model M. A model is said to be *fully accessible* if every element of it is accessible.

We can now ask whether every sentence that is satisfiable has a fully accessible model. It turns out this is not true. For example, consider a signature that has a constant 0 and a function s and the formulae that force a number line from 0:

$$\varphi_0 : \forall x. (\neg s(x) = 0) \wedge \forall x, y. (s(x) = s(y) \Rightarrow x = y)$$

For this formula, it is indeed true that it is satisfied in a fully accessible model, for example a model that contains elements that serve as interpretations for $0, s(0), s(s(0)), \ldots$ only.

However, consider adding a conjunct:

$$\varphi_1 : \ \varphi_0 \wedge \forall x.\ \exists y.(\ f(y) = x \wedge s(y) = y)$$

This formula means that there must be elements whose f images are $0, s(0), s(s(0))$, etc., and hence these elements must all be *distinct* as well. These have to be different from the 0-chain and must be distinct from each other as well (as their f-images are different). Note that there are no *ground terms* that access these (infinitely many) elements. For example, one model that satisfies the above constraints is:

$$U = \mathbb{N} \cup \{i' \mid i \in \mathbb{N}\}$$

$$s(i) = i + 1, \textit{ for every } i \in \mathbb{N}$$

$$s(i') = i', \textit{ for every } i \in \mathbb{N}$$

$$f(i') = i, \textit{ for every } i \in \mathbb{N}$$

$$f(i) = i, \textit{ for every } i \in \mathbb{N}$$

Note that there are no ground terms that "evaluate" to the elements i', where $i \in \mathbb{N}$.

It turns out however that *universal formulae do indeed have the property that satisfiable sentences always have fully accessible models*. This is called *Herbrand's theorem* which we will prove below.

In fact, in the above example, the formula φ_0 is a universal sentence and has a fully accessible model. The sentence φ_1 does not have a fully accessible model, and notice that it uses an existential quantifier. We can, as argued in the last section, Skolemize formulas to have an equisatisfiable formula that has only universal quantification. Skolemizing φ_1 gives:

$$\varphi_1' : \ \varphi_0 \wedge \forall x.\ (\ f(g(x)) = x \wedge s(g(x)) = g(x)\ )$$

Notice that the Skolemization introduces a new function g for the existentially quantified variable y removed. And notice now there is a satisfying fully accessible model. In the model above sketched, we can make $g(i) = i'$ to satisfy the formulas (g for other elements can be defined arbitrarily).

Note that having accessible models is very pleasing. The universe can be thought of as consisting *only* of ground terms in the logic! In fact, we can *name* our elements using the terms in the logic (more precisely, equivalence classes of terms will be the elements in our universe). Also, notice that if ground terms $t_1, \ldots, t_n$ access the elements $e_1, \ldots, e_n$, respectively, in a model M, then clearly $f(t_1, \ldots, t_n)$ accesses the element $f^M(e_1, \ldots, e_n)$. Consequently, in the models we build, the interpretation of functions is *fixed*— the function f will map the ground terms $t_1, \ldots, t_n$ (which are in the universe as the universe consists only of ground terms) to the ground term $f(t_1, \ldots, t_n)$. So, really, the universe, and the interpretation of constants and functions will be *fixed*. The only things to figure out are the interpretation of relations, including equality which will cause the universe to be equivalence classes over ground terms.

We now prove that this is always the case— universal sentences that are satisfiable have fully accessible models. More precisely, we will define Herbrand models where elements *are* equivalence classes of ground terms (with fixed interpretations of functions), and show that satisfiable universal sentences (also called sentences in Skolem form) have Herbrand models.

Universal Formulas and Closed Submodels

The key property that universal sentences satisfy is that they are satisfied in any *submodel* of a satisfying model, as long as the submodel is closed with respect to function applications. Let φ be a universal sentence and M, with universe U, be a model satisfying it. Let $U' \subseteq U$ that satisfies the following properties:

- For every constant c, $c^M \in U'$
- For every function symbol of arity n, if $e_1, \ldots, e_n \in U'$, then $f^M(e_1, \ldots, e_n) \in U'$.

Then the submodel $M' = (U', I')$ define by taking U' as the universe and interpreting all constants, functions, and relation symbols on U' exactly as in M, but restricted to U', is called a *closed submodel*. More precisely, we define the interpretation of the closed submodel with universe U' to be:

- $c^{M'} = c^M$, for every constant symbol c
- For every relation symbol R of arity n, and for every $e_1, \ldots, e_n \in U'$, $R^{M'}(e_1, \ldots, e_n)$ *iff* $R^M(e_1, \ldots, e_n)$
- For every function symbol f of arity n, and for every $e_1, \ldots, e_n \in U'$, $f^{M'}(e_1, \ldots, e_n) = f^M(e_1, \ldots, e_n)$

Note that the properties that U' needs to satisfy is crucial to build the submodel— we cannot build a submodel using any sub-universe of elements (the values that f takes tuples of elements in the sub-universe to must be in the sub-universe as well).

If M' is a closed submodel of M, it turns out that M' will satisfy all the universal sentences that M satisfies (the converse does not hold, of course). Note that a sentence that has an existential quantification, say of the form $\exists x.R(x)$, may hold in M but may not hold in M' (as the elements witnessing the property may be not in the sub-universe, for example). But satisfiability is preserved for submodels on universal formulas. The proof is rather simple:

Lemma 6.3 (Closed submodel property) *Let M be a model and let M' be a closed submodel of M. Let φ be a universal sentence and let $M \models \varphi$. Then $M' \models \varphi'$. Furthermore, every term evaluates to the same element in M' as it does in M.*

Proof Fix a model M, a closed submodel M', and a universal sentence φ : $\forall x_1, \dots, x_n.\varphi'$ where φ' is quantifier free, where $M \models \varphi$. Let $e_1, \dots, e_n$ be the interpretation of the variables $x_1, \dots, x_n$ in the universe of the submodel M'. Then these belong to the universe in M, and since the universe of M' is closed under function applications, and since M' inherits the interpretations of constants and functions from M, it follows that every term t involving constants and these variables evaluate to the same element in M' as they do in M. Since M' also inherits the relations from M (including equality), it follows that all atomic formulas involving these variables evaluate to the same Boolean value in M and M'. Since φ' is quantifier-free, it too will evaluate to the same value in M as in M'. Since M' satisfies φ, for this interpretation of variables, φ' will also evaluate to *true*. Hence we have shown that for all possible interpretations of the variables in the universe of the submodel, φ' evaluates to *true*, which means that $M' \models \varphi$. □

Note in the above that we don't make the claim for *universal formulas* but just for *universal sentences*. The reader should make sure they understand why the lemma does not hold for universal formulas.

Herbrand Models and Herbrand's Theorem

Let us now define Herbrand models.

Definition 6.4 Fix a FO signature Σ. Let GT be the set of all ground terms over Σ. A *functional congruence over ground terms* $\sim$ is an equivalence relation over ground terms such that for every $t_1, \dots, t_n, t'_1, \dots, t'_n$, where for each $1 \leq i \leq n$, $t_i \sim t'_i$, it is the case that $f(t_1, \dots, t_n) \sim f(t'_1, \dots, t'_n)$. For such a congruence $\sim$, we denote the equivalence class containing t with the notation $[\![t]\!]_\sim$.

In the notation for equivalence classes, we sometimes write $[\![t]\!]$, if $\sim$ is clear from context.

Definition 6.5 (Herbrand model with equality) Fix a FO signature Σ with at least one constant symbol (hence the set of ground terms over Σ is non-empty). A Herbrand model (with equality) is one where:

- The universe of the model is $U = \{[\![t]\!] \mid t \in GT\}$ consists of the set of equivalence classes of ground terms of Σ with respect to some functional congruence over ground terms $\sim$.
- Any constant c is interpreted as $[\![c]\!]$.
- Any function symbol f of arity n is interpreted so that for every $t_1, \dots, t_n \in GT$, $f^M([\![t_1]\!], \dots, [\![t_n]\!]) = [\![f(t_1, \dots, t_n)]\!]$.

The first condition says that the universe of a Herbrand model consists of just equivalence classes of terms with respect to a functional congruence $\sim$. The second says that the interpretation of functions is given by the names of the elements themselves— a function f will take the equivalence classes of n terms $t_1, \ldots, t_n$ to the equivalence class of the term $f(t_1, \ldots, t_n)$. This definition of f^M is well-defined since $\sim$ is a functional congruence over terms.[2]

Let us make some simple observations. First, in a Herbrand model, because of the way constants and functions are interpreted, it is easy to show, by induction, that every ground term t evaluates to the equivalence class containing it, i.e., $[\![t]\!]$. It hence follows that in a Herbrand model is fully accessible— every element $[\![t]\!]$ is accessible using the term t.

In fact the converse is also true: every fully accessible model is a Herbrand model, which will be evident in the proof of Herbrand's theorem below.

Let us now prove Herbrand's theorem. [3] Herbrand's theorem states that if a universal sentence has a model, it has a Herbrand model.

The intuition of the proof is quite simple. Let M be a model satisfying a universal sentence φ. Then we can simply take the sub-universe that corresponds to all accessible elements (elements accessible using terms). Clearly, this subset is closed under function applications. And hence it defines a closed submodel that satisfies φ as well. This closed submodel, having accessible elements only, is isomorphic to a Herbrand model— we can relabel every element e using the equivalence class of *all ground terms that evaluate to the element e*, in order to make it a Herbrand model.

Theorem 6.6 (Herbrand's theorem with equality) *Let φ be a universal sentence. Then φ is satisfiable iff it is satisfiable in a Herbrand model.*

Proof We prove the forward direction (the reverse direction is trivial as if φ has a Herbrand model, then it is clearly satisfiable).

Let φ be satisfiable. Let M be a model for φ, with universe U.

Let $U' = \{t^M \mid t \textit{ is a ground term}\}$. Then, clearly, U' satisfies the properties for defining a closed submodel of M— it includes the interpretations of all constant symbols, and for any function symbol of arity n and any n-tuple of elements, say $t_1^M, \ldots, t_n^M$, it clearly contains $f^M(t_1^M, \ldots, t_n^M)$, as that is precisely $f(t_1, \ldots t_n)^M$.

Now let M' be the closed submodel of M defined by U'. By the previous lemma, $M' \models \varphi$.

We now prove M', with its elements renamed, is in fact a Herbrand model. Define a congruence on ground terms as follows: $t \sim t'$ iff $t^M = t'^M$. Verify that this is indeed a congruence on ground terms (proof: verify it is an equivalence relation, and note that if $t_1, \ldots, t_n, t'_1, \ldots, t'_n$ are such that each for each i, $t_i \sim t'_i$, then

[2] If you were a student in elementary school and you knew Herbrand models, and your math teacher asked you what 2 + 3 is, you would say it's "2 + 3"! The value of the function + applied on the terms $\langle 2, 3\rangle$ is simply the term/element $+(2, 3)$. You may not pass your elementary school exams though!

[3] Herbrand's theorem is generally proved in a signature without equality. Then one can show that purely universally quantified formulas have a model where the elements are terms, not equivalence classes of terms. Since we want to treat equality as an interpreted relation that is always in the signature, our treatment has equivalence classes of terms. One could instead take Herbrand's theorem and introduce equality as an uninterpreted relation that satisfies the congruence axioms, and get the same result too.

it follows that $t_i^M = t_i'^M$, and hence $f(t_1, \dots, t_n)^M = f(t_1', \dots, t_n')^M$, and hence $f(t_1, \dots, t_n) \sim f(t_1', \dots, t_n')$).

Let us rename every element e as the nonempty set $[\![t]\!]$, the equivalence class of t wrt $\sim$, where t is some ground term that evaluates to e in M (such a term must exist since every element in U' is accessible). It is easy to prove that no two elements get named by the same equivalence class. Also, every equivalence class $[\![t]\!]$ is the label of some element in U', namely t^M. We can easily prove, by induction on terms, that that every ground term t evaluates to $[\![t]\!]$ in M'.

It immediately follows that this is a Herbrand model satisfying φ. □

Now, notice that in the proof of Herbrand's theorem, given a model that satisfies a formula, we built the submodel *independent* of the formula. The submodel consisted of all elements accessible using any ground term in the signature. Consequently, the same model construction works in showing that if a *set of universal sentences S* has a model, then it has a Herbrand model as well.

Corollary 6.7 (Herbrand's theorem with equality for sets of formulas) *Let S be a set of universal sentences. Then S is satisfiable iff it is satisfiable in a Herbrand model.*

6.4 Some consequences of Herbrand's theorem

Before we move to completeness, let us observe some simple consquences of Herbrand's theorem. First, it follows that if the signature is countable, then a set of sentences S has a model iff it has a *countable* model. In fact, this is true for any set of formulas as well.

Theorem 6.8 (Downward Löwenheim-Skolem Theorem) *Fix a countable signature Σ. If a set of formulas S over Σ has a model then it has a countable model.*

Proof Every formula in S can be made closed (i.e., made into a sentence) and made universal by Skolemizing it (by replacing variables by new constant symbols and removing existential quantification) to result in equisatisfiable formulas. The resulting set S' and S are equisatisfiable. In fact, it is easy to see that models for S work as models for S', and vice-versa (we can keep the universe, and the interpretation of constants, functions, and relations in the common vocabulary the same). Let S be satisfiable. Then S' is satisfiable as well, and by Herbrand's theorem, there is a Herbrand model for S', which, by definition, is countable. This model can be converted back to a model for S (we keep the same universe, we just throw away the interpretations of the added constants and functions during Skolemization, and interpret variables using the interpretations of constants that replaced them). Hence S has a countable model. □

The above is a surprising result. Every set of axioms A (even infinite ones) that has a model also has a countable model. Recall that there are several complete axiomatizations of theories where the *intended* models are uncountable. For example, there is an axiomatization of reals with addition and multiplication, i.e., for the theory of $(\mathbb{R}, 0, 1, +, \cdot)$. How then do they have a countable model? The only explanation is that even for such theories, there is a countable model that is elementary equivalent (which means it satisfies the same first-order sentences) as the model of reals with addition and multiplication! This is truly bizarre, but true!

There is n generalization of the Downward Löwenheim-Skolem Theorem, called the Löwenheim-Skolem Theorem, which we will not prove in this book, that says that if a formula over a countable signature has a satisfying model that is infinite, then it has models satisfying it of cardinalities κ. In particular, there will always be an uncountable satisfying model. This result is surprising too, as there are complete axiomatizations for certain countable models, like $(\mathbb{N}, 0, 1, +)$. The result then says that this set of axioms also has uncountable satisfying models! These are often referred to as nonstandard models of arithmetic. Again, the key thing is though such models exist, they agree with the standard model on all *first-order expressible properties*.

The above results can also be seen as saying that first-order logic is not powerful enough to talk about infinite cardinalities. A set of first-order sentences can say that the model has at most k elements, for any particular $k \in \mathbb{N}$ $(\exists x_1, \ldots x_k. \forall y \bigvee_{i \in [1,k]} (y = x_i))$. However, there is no set of first-order formulae that ensure that the models that satisfy it are countable, or have any particular cardinality. First-order logic also cannot ensure that satisfying models are *finite*—this is true since validity of first-order logic over finite models is not in r.e., but validity over arbitrary models is in r.e. (as we shall see soon in this chapter).

6.5 Gödel's completeness theorem: FO Validity is recursively enumerable

Let us fix a countable signature Σ.

An instance of the validity problem is a set (finite or infinite, but recursive) A of axioms, which are FO sentences, and a sentence φ. Our goal is to show that the problem of checking validity of such instances, i.e., checking if $A \models \varphi$, is recursively enumerable.

We first negate the formula φ. $A \models \varphi$ iff $A \cup \{neg\varphi\}$ is unsatisfiable. Hence our goal is to prove that $S = A \cup \{\neg\varphi\}$ is unsatisfiable. We convert each formula in S to prenext rectified normal form. We then Skolemize the sentences in S to obtain a set X of sentences over an expanded signature Σ' such that S and X are equisatisfiable. Note that X is itself a recursive set. Our goal is now to show that checking whether X is unsatisfiable is recursively enumerable.

Since X has only universal formulas, we know by Herbrand's theorem that to prove X is unsatisfiable iff X has no satisfying Herbrand model.

Since the signature is countable, the set of all formulas is countable, and hence either X is finite or countable. Let us fix an enumeration of X: $\varphi_1, \varphi_2, \ldots$.

Now any universal formula $\forall\bar{x}\psi$ is true in a Herbrand model iff it is true when the variables $\bar{x}$ are interpreted to be elements corresponding to all possible *ground terms*, since Herbrand models have only interpretations of ground terms in their universe. Consequently, such a universal formula is true in a Herbrand model iff for every tuple of ground terms $\bar{t}$, $\psi[\bar{t} \,/\, \bar{x}]$ holds in the model.

Consequently, it is easy to see that X is satisfied in a Herbrand model iff $\{\psi[\bar{t}] \,/\, \bar{x} \mid \forall\bar{x}\; \psi \in X, t \in GT(\Sigma)\}$ is satisfied in the Herbrand model. This leads us to:

Lemma 6.9 (Term Expansion Lemma) *A set of universal formulas* Γ *is satisfiable iff* $\Gamma^* = \{\psi[\bar{t} \,/\, \bar{x}] \mid \forall\bar{x}\; \psi \in \Gamma, t \in GT(\Sigma)\}$ *is satisfiable.*

Proof If Γ is satisfiable, then clearly Γ^* is satisfied in any model satisfying Γ, and hence is satisfiable as well. Conversely, if Γ^* is satisfiable, then consider a Herbrand model satisfying it (which must exist since the sentences are universal, in fact quantifier-free). Clearly, in this Herbrand model, since all elements are accessible using terms, the formulas in Γ are satisfied as well, and hence Γ is satisfiable. □

Due to the above lemma, we can now take

$$X^* = \{\psi[\bar{t} \,/\, \bar{x} \mid \forall\bar{x}\; \psi \in X, t \in GT(\Sigma)\}$$

and our problem now reduces to showing X^* is unsatisfiable. Note that formulas in X^* are quantifier-free. And quantifier-free formulas admit a decidable satisfiability procedure (see previous chapter). However, even if $A = \emptyset$, X^* can be infinite. Consequently, we cannot subject the X^* to a satisfiability decision procedure.

We now want to show a *compactness theorem* for such quantifier-free sets of formulas. This will allow us to prove unsatisfiability of X^* by just looking at finite subsets of it for unsatisfiability. Note that finite subsets of X^* can be conjuncted and subject to a satisfiability decision procedure, as given in the previous chapter.

Compactness Theorem for quantifier-free grounded formulas

We want to show the following lemma:

Lemma 6.10 *Let* Γ *be a set of quantifier-free grounded sentences. Then* Γ *is satisfiable iff every finite subset of* Γ *is satisfiable.*

Proof If Γ is satisfiable, then, of course, every finite subset of Γ is satisfiable. We hence need to show only the converse.

We will use the propositional compactness theorem to prove this lemma. Note that since every sentence is Γ is grounded, each atomic formula is of the form $t = t'$ or $R(t_1, \ldots, t_n)$, where $t, t', t_1, \ldots, t_n$ are grounded terms.

Let us introduce a set of propositions p_a for every atomic grounded formula a. We can now form a set Γ_p that contains the propositional abstraction of formulas in Γ, obtained by replacing every atomic formula a in any formula in Γ with the proposition p_a.

Now, of course, an arbitrary satisfying assignment of Γ_p may not correspond to a way of satisfying Γ, since equalities obey a set of properties, namely the *congruence axioms* detailed in the last chapter. Let us now introduce a set of propositional constraints that capture these axioms, called Δ.

Δ contains the following formulae:

- $p_{t=t}$ for every $t \in GT(\Sigma)$
- $p_{t=t'} \Rightarrow p_{t'=t}$, for every $t, t' \in GT(\Sigma)$
- $(p_{t_1=t_2} \wedge p_{t_2=t_3}) \Rightarrow p_{t_1=t_3}$, for every $t_1, t_2, t_3 \in GT(\Sigma)$
- $$\left(\bigwedge_{i\in[1,n]} p_{t_i=t_i'}\right) \Rightarrow \left(p_{R(t_1,\ldots,t_n)} \Leftrightarrow p_{R(t_1',\ldots,t_n')}\right)$$
  for every relation symbol R of arity n.
- $$\left(\bigwedge_{i\in[1,n]} p_{t_i=t_i'}\right) \Rightarrow p_{f(t_1,\ldots,t_n)=f(t_1',\ldots,t_n')}$$
  for every function symbol f of arity n.

It is now easy to show that Γ is satisifiable iff $\Gamma_p \cup \Delta$ is satisfiable. (Proof: If Γ is satisfied in a model M, define a valuation that assigns the propositions p_a to true iff a is true in the model, and argue that Γ_p and Δ will be satisfied under this valuation. Conversely, if $\Gamma_p \cup \Delta$ is satisfied by a propositional valuation, it is easy to see that the equality relation defined by the propositional formulas is a functional congruence over ground terms, and hence defines a Herbrand model of equivalence classes of terms. Interpreting each relation according to the propositional valuation will satisfy the formulas in Γ.

Now, using compactness theorem for propositional logic, we know that $\Gamma_p \cup \Delta$ is satisfiable iff every finite subset of $\Gamma \cup \Delta$ is satisfiable.

Now let us show the required property. If Γ is unsatisfiable, then $\Gamma_p \cup \Delta$ is unsatisfiable, and hence there is a finite subset F of $\Gamma \cup \Delta$ that is satisfiable. Consider $F' = \Gamma_p \cap F$, which is finite. Then the set of FO formulas corresponding to F' in Γ, i.e., the set of formulas whose propositional abstractions are in F', is unsatisfiable (since $F' \cup \Delta$ is unsatisfiable). Hence there is a finite subset of Γ that is unsatisfiable.$\square$

The Algorithm

We now continue and finish our recursively enumerable procedure. Recall that we had left off in showing X^* is unsatisfiable, where X^* was obtained by replacing each universally quantified sentence with all possible instantiations of ground terms.

Using the above lemma, we know that X^* is unsatisfiable iff there is some finite subset of X^* that is unsatisfiable.

We can now build a procedure to find such a finite subset. Recall that for any finite subset of quantifier-free formula, there is a decision procedure (that always halts) whether the set is satisfiable, from the previous chapter. Let's call this decision procedure $DP\text{–}QFE$ (decision procedure for quantifier-free equality).

6.5.1 The case of finite sets of formulas

We first consider the case when the set of axioms is finite, and hence X is finite. Note that in this case, we can assume the signature is finite too (as the functions/relations not mentioned in the set of formulas clearly do not matter). Note that X is finite, but X^* can be, however, infinite.

Let us consider the following increasing sets that cover X^*. For any $d \in \mathbb{N}$, let T_d denote the ground terms of depth at most d. Formally, these sets are defined recursively as:

- $T_0 = \{c \mid c \ \textit{is a constant symbol in}\ \Sigma\}$
- $T_{d+1} = T_d \cup \{f(t_1, \ldots, t_n) \mid t_1, \ldots, t_n \in T_d, f \ \textit{is a function symbol of arity}\ n\}$

Note that $T_i \subseteq T_j$ for any $i \leq j$, and $\bigcup_{i \in \mathbb{N}} T_i$ is the set of all ground terms.
Our procedure is as follows: Given X, a finite set of universal sentences, we do the following:

1. Set $i := 0$;
2. Repeat forever: {
3. $R := \{\psi[\bar{t} \,/\, \bar{x}] \mid \bar{t} \ \textit{is a tuple of elements in}\ T_i\}$.
4. Check if R is satisfiable, by calling $DP\text{-}QFE(R)$.
   If it is not satisfiable, then report X is unsatisfiable and exit (concluding $A \models \varphi$).
5. Increment i;
6. }

The correctness of the algorithm is straightforward to see. If $A \models \varphi$, then $A \cup \{\neg\varphi\}$, and hence X would be unsatisfiable. Hence X^* is unsatisfiable. Hence there is a finite subset of $F \subseteq X^*$ that is unsatisfiable. Let FT be the set of terms mentioned in F; then FT is finite. Let i be the maximum depth of the terms in FT. Then in iteration i, the algorithm will instantiate X with all terms of depth i, and hence the set R it constructs will be a superset of F, and hence will be unsatisfiable. Hence the decision procedure call to $DP\text{–}QFE$ will report unsatisfiable, and the algorithm will halt and report $A \models \varphi$.

On the other hand, if $A \not\models \varphi$, then $A \cup \{\neg\varphi\}$, and hence X would be satisfiable. Hence X^* is satisfiable. In each iteration, the algorithm constructs R which is a subset of X^*, and hence the call to $DP\text{–}QFE$ will report satisfiable in each round. Hence the algorithm will not halt, and will never declare $A \models \varphi$ holds.

6.5.2 The case for infinite sets of formulas

Let us assume the signature is finite. Let us assume we are asked whether $A \models \varphi$, where A is infinite, but recursive. Again, we know that $A \models \varphi$ iff $A \cup \{\neg\varphi\}$ is unsatisfiable iff the set X constructed by converting formulas in the set to universal formulas is unsatisfiable. This set X is unsatisfiable iff X^* is unsatisfiable. And X^* is unsatisfiable iff there is a finite subset of X^* that is unsatisfiable. The key difficulty is to explore larger and larger finite subsets of X^* systematically such that for every finite subset F of X^*, we eventually will explore a superset of F. There are two infinities to consider here— the set of formulas in X is infinite and the set of terms to instantiate the formulas is also infinite. We need to *dovetail* through the two infinities in order to build our procedure.

Let $En : Y_0, Y_1, Y_2, \ldots$ be an enumeration of certain finite subsets of X^*. Such an enumeration is said to be *fair* if for every finite subset $F \subseteq_{fin} X^*$, there is some $i \in \mathbb{N}$ such that $F \subseteq Y_i$.

There are several ways to achieve a fair enumeration. We give just one example. Consider the enumeration where Y_i consists of the first i sentences in X enumerated by all possible ground terms of depth i. Then clearly this is a fair enumeration. Let F be a finite subset of X^*. Let i be the largest number such that the i'th formula in X, instantiated in some way, belongs to F. Let j be the depth of the largest term that was used to instantiate some element in F. Now, let $k = max(i, j)$. Then it follows that $F \subseteq Y_i$.

For any fair enumeration, we have the following semi-algorithm: Given X, a recursive but infinite set of universal sentences, we do the following:

1. Fix a fair enumeration $Y_0, Y_1, \ldots$ of X^*
2. Set $i := 0$;
3. Repeat the following forever: {
4. Check if Y_i is satisfiable, by calling *DP-QFE*(Y_i).
   If it is not satisfiable, then report X is unsatisfiable and exit (concluding $A \models \varphi$).
5. Increment i;
6. }

Again, the proof that the above algorithm always halts when $A \models \varphi$ and reports that it is so, and the proof that when $A \not\models \varphi$, the algorithm runs forever, is easy to see.

We can extend the above argument also to *countably infinite signatures* and infinite but recursive set of axioms. In this case, we need to dovetail between several infinities— exploring more symbols in the signature, exploring more axioms involving this expanding signature, and systematic term instantiation involving this expanding signature. Again, any fair enumeration will give an r.e. procedure.

6.5.3 Completeness Theorem

We can now phrase our completeness theorem, which follows from the above results.

Theorem 6.11 (Completeness) *Let Σ be a finite or countable signature. Let A be a finite set of sentences or an infinite recursive set of sentences over Σ, and let φ be a sentence over Σ. Then the problem of checking whether $A \models \varphi$, is recursively enumerable.*

Let us now work out an example.

Example 6.12 Consider the group axioms, where we have a special constant e for the identity element:

- Associativity: $\forall x, y, z.\ f(f(x, y), z) = f(x, f(y, z))$
- Identity: $\forall x. f(x, e) = x \wedge f(e, x) = x$
- Inverse: $\forall x \exists y. f(x, y) = e \wedge f(y, x) = e$

Let us now take the above three sentences as the set of axioms A. And let us try to prove the following formula, which says the identity is unique, i.e.,

$$\varphi : \forall e'.\ ((\forall x.(f(x, e') = x \wedge f(e', x) = x) \Rightarrow (e = e')))$$

Of course, the above property is true even of monoids, i.e., even when the first two axioms hold. However, let's consider all group axioms for this example.

The first two formulas are already in prenex rectified normal form and universal. Skolemizing the third axiom using a new function g gives:

$$\forall x f(x, g(x)) = e \wedge f(g(x), x) = e$$

The new function symbol g intuitively corresponds to a function that provides the inverse of an element. (We don't need to know it is unique in order to ask that such a function exists.)

The formula φ is not in prenex form; bringing it to prenex form gives:

$$\begin{aligned}\varphi &\equiv \forall e'.\ (\neg(\forall x.(f(x, e') = x \wedge f(e', x) = x) \vee (e = e'))) \\ &\equiv \forall e'.\ ((\exists x.(\neg(f(x, e') = x) \vee \neg(f(e', x) = x)) \vee (e = e'))) \\ &\equiv \forall e'.\ \exists x.(\neg(f(x, e') = x) \vee \neg(f(e', x) = x) \vee (e = e'))\end{aligned}$$

The negation of φ is hence:

$$\neg\varphi \equiv \exists e'.\ \forall x.(f(x, e') = x \wedge f(e', x) = x \wedge \neg(e = e'))$$

Skolemizing the above, by replacing the quantified variable e' by a new constant symbol c gives:

$$\forall x.(f(x, c) = x \wedge f(c, x) = x \wedge \neg(e = c))$$

We now have a set X containing four universal formulas:

- $\forall x, y, z.\ f(f(x, y), z) = f(x, f(y, z))$
- $\forall x. f(x, e) = x \wedge f(e, x) = x$
- $\forall x f(x, g(x)) = e \wedge f(g(x), x) = e$
- $\forall x.(f(x, c) = x \wedge f(c, x) = x \wedge \neg(e = c))$

And our task is to check whether they are simultaneously satisfiable.

Let us instantiate with the depth 0 ground terms, i.e., by constants e and c. Then we get the formulae where all quantified variables are replaced by all possible combinations of e and c. That's 14 quantifier-free formulas!

Note that this set includes the following formulae:

- The second formula with x replaced by c:
  $f(c, e) = c \wedge f(e, c) = c$
- The fourth formula with x replaced by e:
  $(f(e, c) = e \wedge f(c, e) = e \wedge \neg(e = c))$

Clearly these two formulas are not satisfiable in any model. If $f(c, e) = c$ and $f(c, e) = e$, then we must have $c = e$, which contradicts the conjunction $\neg(e = c)$.

Hence when we ask the decision procedure for quantifier-free formulae whether the 14 formulas have a model, it will report unsatisfiable, and the algorithm above would conclude $A \models \varphi$.

We invite the reader to in fact generate the above formulae, and give them to an SMT solver, like Z3 or CVC4, in order to check that the quantifier-free formulae are unsatisfiable.

Example 6.13 We can take the same axioms above, and try to show that the following holds, which says that inverses are unique. Since we have used the function g, during Skolemization of the axioms, to give us the inverse of elements, let's use the same function g (for brevity).

$$\varphi : \forall x, y.\ (f(x, y) = e \wedge f(y, x) = e) \Rightarrow (y = g(x))$$

Negating the above and Skolemizing using two new constant symbols c and d gives:

$$(f(c, d) = e \wedge f(d, c) = e) \wedge \neg(d = g(c))$$

Instantiating the Skolemized axioms and the above formula with depth 0 terms (i.e., by the constants e,c, and d) gives a large set of quantifier-free formulae, and it turns out that they are already unsatisfiable. We encourage the reader to write these formulae and feed it to an SMT solver to check that this is indeed so. Consequently, $A \models \varphi$.

6.6 Observations and Consequences

Using SMT solvers:

The above presentation was carefully done so that we get an r.e. procedure that repeatedly calls a solver to check satisfiability of quantifier-free formulae with equality. One can instead also go all the way down to propositional logic satisfiability, and implement the satisfiability of quantifier-free formulae using satisfiability of a propositional encoding of it. This was in fact proposed by Gilmore in 1960! Since SMT solvers already implement satisfiability of quantifier-free formulae with equality, and avoids the blow-up that the propositional encoding entails, we prefer this technique. Furthermore, we will see another application of this term instantiation in a later chapter that allows us to *combine* quantified theories.

The Bernays-Schönfinkel-Ramsey/EPR class

Let us now consider a signature without any function symbols, and a finite set S of formulas of the form $\exists\overline{x}\forall\overline{y}\varphi$. We are asked to check if S is satisfiable. Skolemizing these formulas could introduce new constants but *no new functions*. Consequently, we end up with a set of universal formulas X that we need to check for satisfiability. Since there are no function symbols, the only ground terms are the constants, and we can assume that the constants are only those that occur in the formula, without loss of generality. Consequently, the r.e. procedure outlined earlier in this section can stop after the first instantiation of constants! Hence it is a decision procedure (which always halts on all inputs) and decides satisfiability of such formulae. This fragment of FO formulae, namely $\exists^*\forall^*$ sentences over a signature that has no function symbols, hence admits a decidable satisfiability problem, and is called the *Bernays–Schönfinkel-Ramsey* class or the *effectively propositional reasoning* (EPR) class. Note that for validity, the fragment that is decidable is the $\forall^*\exists^*$ fragment where the signature has no function symbols. This is one of the few quantified fragments of first-order logic that admits decidable validity.

Decidability when Axioms are Negation Complete

A set of axioms A is said to be *consistent* (without contradiction) if there is no sentence such that $A \models \varphi$ and $A \models \neg\varphi$, i.e., $\varphi, \neg\varphi \in Th(A)$. Note that a set of axioms A is consistent iff there is at least one model satisfying the axioms A.

A set of axioms A is said to be *complete* (or *negation complete*) if for every sentence φ, either $A \models \varphi$ or $A \models \neg\varphi$. In other words, the theory of A, $Th(A)$, contains either φ or $\neg\varphi$.

For example, the set of axioms of Presburger arithmetic is consistent and complete. The set of axioms of groups is consistent but not complete.

A consequence of the results of this section is that the theory any complete and consistent axiomatizations is *decidable*. Given a sentence φ, we can execute two copies of the r.e. procedure defined in this section to check whether $A \models \varphi$ and whether $A \models \neg\varphi$. These two executions must be simulated essentially in parallel— for example, running one procedure k steps and then switching to the other for k steps, and then switching back, forever, for some fixed k. Since either $A \models \varphi$ or $A \models \neg\varphi$, one of these procedures will terminate, in which we can halt, and report whether φ is in the theory or not.

Theorem 6.14 *Let A be a recursive set of sentences that is complete. Then the theory of A, $Th(A)$ is decidable.*

The above also means that if the theory of a single structure is *undecidable*, then it is not axiomatizable. We will prove (see next chapter) that the theory of $(\mathbb{N}, 0, 1, +, \times)$ is undecidable. This means that there is *no recursive set of FO axioms A* such that the theory of A is identical to the theory of this model! This is in fact a version of Gödel's first incompleteness theorem.

Axiomatizability and recursive enumerability

We proved completeness for any recursive set of axioms. However, it is easy to extend the result even when the axioms are recursively enumerable— the procedure will enumerate axioms and instantiate them systematically.

Consider a class of structures C. The notions of having a recursively enumerable set of axioms A that characterize the theory (i.e., $Th(A) = Th(C$ and having $Th(C)$ itself being recursively enumerable are synonymous. If a r.e. set of axioms A exists characterizing C, then by the completeness theorem, $Th(\mathcal{A})$ is r.e. as well. On the other hand, if $Th(C)$ is r.e., then we can choose as axioms this theory itself.

Axiomatic Systems

The most important consequence of the completeness theorem is that it justifies the axiomatic approach. We are typically interested in logic over a particular single structure, or interested in a class of structures. There are many ways to define such a single structure or a class of structures, even using finite means (for example, we can define them using computable functions— giving functions that decide which strings over an alphabet are the elements of a univers, and providing programs that oeprationally define functions and relations). The axiomatic method, in contrast, asks the class of structures to be defined using properties which are themselves written in FOL. And the completeness theorem gives the guarantee that validity of such a set of axioms is always r.e., which roughly means that every theorem has a proof.

Compactness Theorem for FOL

Another consequence of the results of this section is that the compactness theorem holds for first-order logic sentences as well.

Theorem 6.15 *Let Γ be a set of first-order sentences over a countable signature Σ. Γ is satisfiable iff every finite subset of Γ is satisfiable.*

Proof The forward direction is trivial. For the converse, assume Γ is unsatisfiable. Then, by the results of this section, we can assume Γ is a set of universal sentences. Then, by Lemma 6.3 (Term Expansion Lemma), $\Gamma^* = \{\psi[\bar{t}/\bar{x} \mid \forall x\psi \in \Gamma, t \in GT(\Sigma)\}$ is unsatisfiable. In other words, the set of quantifier-free sentences obtained by instantiating variables by all possible ground terms is unsatisfiable. By Lemma 6.4, there exists a finite subset F^* of Γ^* that is unsatisfiable. Let $F \subseteq \Gamma$ be a finite subset of Γ from which the elements of F^* were obtained (using term instantiation). Then F is unsatisfiable as well (since even instantiations of variables by ground terms make it unsatisfiable). Hence there is a finite subset of Γ that is unsatisfiable. $\square$