SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Three Proofs of the Compactness Theorem of First Order Logic

av

Maximilian Vranjes

2025 - No K38

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET, 106 91 STOCKHOLM

Three Proofs of the Compactness Theorem of First Order Logic

Maximilian Vranjes

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Anders Mörtberg

2025

Abstract

Abstract: The goal of this thesis will be to explore three areas of mathematics through the lens of a single theorem - the compactness theorem. These areas include filters and ultraproducts, Boolean algebras and topologies, and Henkin constructions. The thesis will also be clear on the role of the axiom of choice in each of these methods.

Syftet med denna uppsats är att utforska tre områden inom matematiken genom linsen av en sats - kompakthetssatsen. Dessa områden inkluderar filter och ultraprodukter, booleska algebror och topologier samt Henkin-konstruktioner. Uppsatsen kommer också att tydligt belysa urvalsaxiomets roll i var och en av dessa metoder.

AI Statement

An LLM based translation tool, DeepL, was used to assist in translating the abstract to Swedish. This is the full extent of the use of AI in writing this thesis.

Contents

Chapter 1

Introduction

The compactness theorem is one of the central theorems in the study of first-order logic. In fact it can be considered one of the defining properties of first order logic trough Lindström's theorem, which states that first order logic is the strongest logic with countable compactness and downwards Löwenheim-Skolem.

This importance has led to the compactness theorem drawing a lot of attention which in turn means that many proofs covering different concepts in mathematics have been created for this theorem. This thesis will cover three examples of such proofs with a focus towards order theory, topology, and model theoretic techniques, which allows for an overview of a number of important mathematical concepts in a coherent manner. These concepts include filters, ultraproducts, Boolean algebras, Stone spaces, and Henkin constructions.

The knowledge assumed of the reader is knowledge of first-order logic equivalent to any undergraduate course in mathematical logic as well as knowledge of basic concepts in topology, notably the notions of topological bases and compact spaces.

It is common in the study of mathematics - specifically in mathematics close to foundations - to put effort into determining what specific axioms are necessary for certain theorems and proofs. This thesis in particular will briefly consider the roles of the axiom of choice and the ultrafilter lemma in the proofs discussed.

Chapter 2

Filters and Boolean Algebras

The knowledge discussed in this chapter is largely informed by [1].
The intuitive idea that the notion of a filter attempts to capture is the notion of a "large" set. In particular, the formal definition of a filter is as follows:

Definition 2.1. Let S be a set. A subset F of the power set of S is called a filter on S if

1. F is nonempty
2. $\varnothing \notin F$
3. If $A \in F$ and $B \in F$, then $A \cap B \in F$
4. If $A \in F$ and $A \subseteq B \subseteq S$, then $B \in F$

These rules follow the desired intuition since clearly the empty set is not large, a superset of a large set is large, and the intersection of two large sets is large.

Definition 2.2. A filter F on S is called an ultrafilter if there does not exist a filter G on S s.t. $F \subseteq G$ and $F \neq G$.

Lemma 2.3. *Let $\{F_i\}$ be a set of filters ordered by some ordinal α such that if $\beta < \lambda \leq \alpha$, then $F_\beta \subset F_\lambda$. The union $F = \bigcup\{F_i\}$ is a filter.*

Proof. Clearly F is nonempty and does not contain $\varnothing$. If A and B are in F, then by the well ordering of ordinals, there must exist a least ordinal α' s.t. $F_{\alpha'}$ contains both A and B. By definition $F_{\alpha'}$ contains $A \cap B$ so F does as well. The same argument shows that F is closed under supersets. $\square$

Lemma 2.3 immediately shows that the axiom of choice implies that all filters can be extended to an ultrafilter through a simple application of Zorn's lemma. The statement that all filters can be extended to an ultrafilter is known as the ultrafilter lemma (UF). It is strictly weaker than the axiom of choice but showing that is beyond the scope of this text [2, p. 17].

There exists another useful characterization of ultrafilters.

Proposition 2.4. *A filter F is an ultrafilter on X if and only if for all $S \subset X$, $S \in F$ or $X \setminus S \in F$.*

Proof. The if direction is trivial, as if $S \notin F$, then clearly $F \cup S$ cannot generate a filter as $X \setminus S \in F$.
Assume F is an ultrafilter on X and $S \subset X$ is such that $S \notin F$ and $X \setminus S \notin F$. If there exists an $f \in F$ such that $f \cap S = \varnothing$, then $f \subset X \setminus S$, which contradicts the closure of F under supersets. Thus for all $f \in F$, $f \cap S \neq \varnothing$. Since all finite intersections in F exist as an element in F, this implies that $F \cup \{S\}$ has the finite intersection property and extends to a filter. This contradicts the assumed maximality of F, thus no such S can exist. □

There exists one case where filters fail to meet the intuition of "a set of large sets" known as principal filters.

Definition 2.5. A filter F on S is called principal if there exists an $X \subseteq S$ s.t. $F = \{A \subseteq S : X \subseteq A\}$.

A principal ultrafilter is of the form $\{A \subset S : x \in A\}$ for some $x \in S$.

Theorem 2.6. *Given any infinite set X, There exists a nonprincipal ultrafilter on X.*

Proof. Consider the set F of cofinite subsets of X, that is all subsets with finite complement. Clearly this set is closed under supersets, which result in smaller complements, and binary intersections since the union of two finite sets is finite. Since X is assumed to be infinite, F is nonempty and does not contain the empty set since the complement of $\varnothing$ is X which is infinite. Thus it is a filter. By lemma 2.3 there exists an ultrafilter U extending F. Assume U is principal, this means that for some $x \in X$, U is of the form $\{A \subseteq X : x \in A\}$. However, the set $X \setminus \{x\}$ is clearly cofinite and as such is also in U. This means both $\{x\}$ and $X \setminus \{x\}$ are in U, thus their intersection, $\varnothing$, is also in U. This contradicts the fact that U is a filter, so U must not be principal. □

The following definition and result will provide an important method for constructing filters.

Definition 2.7. A set S is said to have the finite intersection property if for all finite nonempty subsets T of S, $\bigcap T \neq \varnothing$

Theorem 2.8. *If S has the finite intersection property, then S extends to a filter.*

Proof. Clearly S does not contain $\varnothing$. First create S' closing S under finite intersections. By assumption S' will still not contain $\varnothing$, and will still have the finite intersection property as any element of S' is a finite intersection of elements of S. Then create S'' by closing S' under supersets. Clearly S'' will not contain any smaller intersections than S', it will be closed under supersets, intersections, and will not contain $\varnothing$. Thus it will be a filter. In fact this is the smallest filter containing S. □

2.1 Boolean Algebras

Definition 2.9. A Boolean algebra is a set S equipped with two binary operations, $+, \cdot$, one unary operation $-$, and two constants 0 and 1 satisfying the following properties:

1. Commutativity and associativity of $+$ and $\cdot$
2. Distributativity of $+$ over $\cdot$ and of $\cdot$ over $+$
3. $x + (x \cdot y) = x \cdot (x + y) = x$
4. $x + -x = 1$
5. $x \cdot -x = 0$
6. $x + 0 = x$
7. $x \cdot 1 = x$

Lemma 2.10. *$-x$ is the unique element that satisfies properties 4 and 5.*

Proof. Assume z is such that $x+z = 1$ and $x \cdot z = 0$, then $z = z+0 = z+(x \cdot -x) = (z+x) \cdot (z+-x) = 1 \cdot (-x+z) = (-x+x) \cdot (-x+z) = -x+(x \cdot z) = -x+0 = -x$ □

An important property of Boolean algebras is that $x \leq y \Leftrightarrow x = x \cdot y$ defines a partial ordering on the underlying set. The following few propositions will show that the ordering is an equivalent way of defining a Boolean algebra.

Definition 2.11. The meet of a pair of elements in a partial order, denoted $x \wedge y$, is the infimum of x and y. The join, denoted $x \vee y$, is the supremum of x and y. Both of these definitions extend to arbitrary sets of elements. The meet of a set S of elements is the infimum of S and the join is the supremum of S

Proposition 2.12. *Let B be a Boolean algebra, then $x \cdot y$ is the meet of the elements x and y wrt the $\leq$ ordering.*

Proof. Let z be such that $z = z \cdot x$ and $z = z \cdot y$. Consider $z \cdot (x \cdot y)$. By associativity this equals $(z \cdot x) \cdot y = z \cdot y = z$ by construction. Thus $x \cdot y \leq z$. □

Similarly $x + y$ is the join of x and y.

Theorem 2.13. *Two Boolean algebras X and Y are isomorphic if and only if their respective partial orders are isomorphic.*

Proof. If X and Y are isomorphic as Boolean algebras under the isomorphism f, then assume $x \leq y$, the goal is to prove that $f(x) \leq f(y)$. In the Boolean algebra this is written as $f(x) = f(y) \cdot f(x)$, and by isomorphism this follows from $x = x \cdot y$, since $f(x \cdot y) = f(x) \cdot f(y)$. Assume instead that it is false that $x \leq y$, then in particular $x \neq x \cdot y$, so since f is an isomorphism $f(x) \neq f(x) \cdot f(y)$ and as such it is untrue that $f(x) \leq f(y)$
If instead X and Y are isomorphic as partial orders under the isomorphism f, then

1. $f(0_X) = f(0_Y)$ since 0 is the minimal element wrt $\leq$, similarly $f(1_X) = f(1_Y)$ since it is the maximal element.

2. $f(x + y) = f(x) + f(y)$ since order isomorphisms send supremums to supremums. Similarly $f(x \cdot y) = f(x) \cdot f(y)$.

3. By uniqueness of $-f(x)$ and preservation of supremums and infimums, $f(-x) = -f(x)$.

□

This partial order allows us to consider filters within Boolean algebras.

Definition 2.14. A subset F of a Boolean algebra B is called a filter if

1. F is nonempty

2. If $a \in F$ and $b \in F$ then $a \wedge b \in F$

3. If $a \in F$ and $a \leq b$ then $b \in F$

4. $0 \notin F$

All the properties that hold about filters in algebras of sets hold in filters in Boolean algebras. This can either be proven individually in each case or by appealing to Stone's representation theorem [1], which states that every Boolean algebra is order-isomorphic to an algebra of sets. This means that discussing things like ultrafilters makes sense in this context.
The Boolean algebra that is relevant to this paper is called the Lindenbaum-Tarski algebra.

Definition 2.15. The Lindenbaum-Tarski algebra has as the underlying set the set of all formulas in some first order language, quotiented by the equivalence relation $\varphi \sim \psi \leftrightarrow \vdash \varphi \Leftrightarrow \psi$. The operations are defined as follows:

1. $x \cdot y$ is the equivalence class of $x' \wedge y'$ where x' and y' are representatives of x and y. This is independent of choice of representatives as formula equivalence is preserved by substitution of subformulas with equivalent subformulas.

2. Similarly $x + y$ is the equivalence class of $x' \vee y'$.

3. Similarly $-x$ is the equivalence class of $\neg x$

4. 0 is the equivalence class of $\perp$

5. 1 is the equivalence class of $\top$

From this one can conclude that $x \leq y$ iff $x = y \wedge x$ iff $x \vdash y$

Chapter 3

Proof by Ultraproducts

3.1 Ultraproducts

Ultraproducts are a way to go from various types of finite structures, whether the finiteness be in cardinality, generating sets, dimension, characteristic, or something else, to infinite structures. The main reason they are important in model theory is Łoś's theorem, which in particular leads to a very simple proof of the compactness theorem through this connection between the finite and the infinite.

Definition 3.1. Let L be a first-order language, I be some infinite set, U an ultrafilter on I, and M_i an I-indexed set of nonempty L-structures. The ultraproduct of M_i with respect to U, denoted $\prod M_i/U$, is defined as followed:

1. The underlying set is the regular set-theoretic product quotiented by the equivalence relation $(x_1, x_2, ...) \cong (y_1, y_2, ...)$ if the set of $i \in I$ s.t. $x_i = y_i$ is in U. An element f in each equivalence class can be understood as a function $f : I \to \bigcup M_i$ with the property that $f(i) \in M_i$.
2. If c is a constant symbol in L, then c in the ultraproduct is the equivalence class of $(c^{M_1}, c^{M_2}, ...)$ where c^{M_i} is the interpretation of c in M_i.
3. If f is an n-ary function symbol in L with inputs $x_1, ..., x_n$, and for each x_i, y_i is a representative of it as an equivalence class. Define $z \in \Pi M_i/U$ and z_0 as a representative of z s.t. $z_0(x) = f^{M_x}(y_1(x), y_2(x), ...)$, then f sends $(x_1, ..., x_n)$ to z.
4. If R is an n-ary relation symbol in L with inputs $x_1, ..., x_n$, and for each x_i, y_i is a representative of it as an equivalence class. $R(x_1, ..., x_n)$ holds iff $\{x : R^{M_x}(y_1(x), ..., y_n(x))\}$ is in U.

Proposition 3.2. *Ultraproducts are well defined with this definition.*

Proof. 1. The underlying set being well defined is equivalent to the relation actually being an equivalence relation. Reflexivity is trivial, symmetry follows from symmetry of equality, and transitivity follows form the closure of U under binary intersection

2. The interpretation of constant symbols is clearly well defined as it explicitly constructs a single representative of a single equivalence class for each symbol.

3. Let f be an n-ary function symbol. Again let $x_1, ..., x_n$ be inputs with two collections of representatives $y_1, ..., y_n$ and $y'_1, ..., y'_n$. Let $z_0(x)$ be as in the definition for the representatives $y_1, ..., y_n$ and $z'_0(x)$ for the representatives $y'_1, ..., y'_n$. Each y_i agrees with y'_i on a U-large set by construction. By the closure under finite intersections, this means that $(y_1, ..., y_n)$ agrees with $(y'_1, ..., y'_n)$ on a U-large set, thus so does $z_0(x)$ and $z'_0(x)$. This means both are representatives of the same z.

4. Let R be an n-ary relation symbol. Again let $x_1, ..., x_n$ be inputs with two collections of representatives $y_1, ..., y_n$ and $y'_1, ..., y'_n$. Assume $\{x : R^{M_x}(y_1(x), ..., y_n(x))\} \in U$. As in the previous case $(y_1, ..., y_n)$ agrees with $(y'_1, ..., y'_n)$ on a U-large set S, thus $\{x : R^{M_x}(y'_1(x), ..., y'_n(x))\}$ holds on $S \cap \{x : R^{M_x}(y_1(x), ..., y_n(x))\}$, which by closure of U under finite intersection is a U-large set. Since the representatives were chosen arbitrarily this argument can be done with the choices swapped, thus well-definedness is proven.

□

3.2 Łoś's Theorem

The following theorem, Łoś's theorem in essence states that some first order sentence is true in an ultraproduct iff it is true in a U-large set of factors.

Theorem 3.3. *Let M_i be an I-indexed sequence of L-structures, let U be an ultrafilter on I, and let $\Pi M_i/U$ be the ultraproduct of M_i with respect to U. Let $t_1, ..., t_n$ be elements of the ultraproduct with representatives $f_1, ..., f_n$ Let $\varphi(x_1, ..., x_n)$ be an L-formula. $\Pi M_i/U \models \varphi(t_1, ..., t_n)$ iff $\{i : \varphi(f_1(i), ..., f_n(i)\}$ is in U. This theorem is known as Łoś's theorem.*

Proof. We proceed by induction on the formula φ:

1. $\varphi(x_1, ..., x_n)$ is atomic. Only the case when $\varphi(x_1, ..., x_n)$ is of the form $R(g_1(x_1), ..., g_n(x_n))$ for terms g_i will be covered. In the case of equality one can simply treat it as a 2-ary relation.
   $\Pi M_i/U \models R(g_1(t_1), ..., g_n(t_n))$
   $\iff \{x : R^{M_x}(g_1(f_1(x)), ..., g_n(f_n(x))\} \in U$
   $\iff \{x : \varphi(f_1(x), ..., f_n(x)\} \in U.$

2. If $\varphi = \neg\psi$ with the theorem holding for ψ, then
$\Pi M_i/U \models \varphi(t_1, ..., t_n)$
$\iff \Pi M_i/U \models \neg\psi(t_1, ..., t_n)$
$\iff \{x : M_x \models \neg\psi(f_1(x), ..., f_n(x))\} \in U$
$\iff \{x : M_x \models \varphi(f_1(x), ..., f_n(x))\} \in U$

3. If $\varphi = \psi \wedge \sigma$ where the theorem holds for ψ and σ then
$\Pi M_i/U \models \varphi(t_1, ..., t_n)$
$\iff \Pi M_i/U \models \psi(t_1, ..., t_n) \wedge \sigma(t_1, ..., t_n)$
$\iff \{x : M_x \models \psi(f_1(x), ..., f_n(x)) \wedge \sigma(f_1(x), ..., f_n(x))\} \in U$
$\iff \{x : M_x \models \varphi(f_1(x), ..., f_n(x))\} \in U$

4. If $\varphi = \exists x\psi$ where the theorem holds for ψ, then there exists some $t \in \Pi M_i/U$ with representative f s.t.
$\Pi M_i/U \models \psi(t, t_1, ..., t_n)$
$\iff \{x : M_x \models \psi(f(x), f_1(x), ..., f_n(x)\} \in U$
$\implies \{x : M_x \models \exists f\psi(f, f_1(x), ..., f_n(x)\} \in U$.
This shows one direction, but doesn't immediately show the other as the last step is a one-way implication. To show the other direction assume $\implies \{x : M_x \models \exists f\psi(f, f_1(x), ..., f_n(x)\} \in U$, and choose, from each M_x, some witness g_x to the existential if it exists, choose it arbitrarily otherwise. Then $f(x) = g_x$ defines an element of an equivalence class t in $\Pi M_i/U$ which satisfies $\Pi M_i/U \models \psi(t, t_1, ..., t_n)$. The use of choice in this step cannot be avoided.[3]

The induction ends here as every formula can be obtained by only negation and conjunction of atomic formulas. □

Some steps in this induction took reference from [1, p. 160] With this, we now have enough infrastructure to prove the compactness theorem of first-order logic.

Theorem 3.4. *Any finitely satisfiable theory has a model.*

Proof. Let T be a finitely satisfiable theory. If it is finite the result follows trivially. Assume now that T is infinite. By finite satisfiability, there exists, for each finite $S \subset T$, some M_S that models S. For each formula $\varphi \in T$, let $X_\varphi = \{S : M_S \models \varphi\}$. Let X=$\{X_\varphi : \varphi \in T\}$. Clearly X is a subset of the powerset of T. Consider the intersection of finitely many elements of X: $\bigcap X_{\varphi_i}$. By finite satisfiability of T, there must exist some model of $\bigwedge \varphi_i$, so this intersection is nonempty. Thus X has the finite intersection property and extends to a filter, then to an ultrafilter U. Consider the ultraproduct $\Pi M_S/U$. By Łoś's theorem, if $\varphi \in T$, then since $X_\varphi \in U$, $\Pi M_S/U \models \varphi$. Thus $\Pi M_S/U \models T$. □

Chapter 4

Proof by Topology

The core idea behind this proof will be that of Stone spaces[4, p. 135].

Definition 4.1. Let B be a Boolean algebra. The Stone space of B, denoted $S(B)$ has as points ultrafilters in B. The topology of $S(B)$ is generated by the basis of sets of the form $U_x = \{F : x \in F\}$ where x is some element in B.

Proposition 4.2. *The claimed basis is actually a basis.*

Proof. A subset S of the powerset of some set X is a basis if and only if S covers X and for all $A \in S, B \in S$ and all $x \in A \cap B$, there exists a $C \in X$ such that $x \in C \subset A \cap B$.
Clearly the sets cover the space as if $F \subset B$ is a filter it is nonempty so it contains some element x, thus it is in U_x.
Let U_x and U_y be basis elements such that there exists an $F \in U_x \cap U_y$. Since by definition F must contain both x and y, it must also by the definition of a filter contain $x \vee y$ and as such is in $U_{x \vee y}$, which is a subset of both U_x and U_y, thus it is a subset of $U_x \cap U_y$. □

The structure of this proof will be to prove that these spaces are compact in the topological sense and then to apply that using the Stone space of the Lindenbaum-Tarski algebra to prove that the compactness theorem holds using an argument heavily inspired by [5].

Lemma 4.3. *If F is an ultrafilter on a boolean algebra B, then for all $x \in B$ either $x \in F$ or $-x \in F$.*

Proof. Since F is an ultrafilter it cannot be extended to a larger filter. Assume $x \notin F$, then by closure of finite joins, it must be the case that there exists an element $y \in F$ such that $y \vee x = 0$ as otherwise $F \cup \{x\}$ is a filter. By uniqueness of complement $y = -x$. □

Theorem 4.4. *The Stone space of a Boolean algebra B is compact.*

Proof. It is sufficient to show that any basic open cover has a finite subcover. Let $S = \{U_i\}_{i \in I}$ be a basic open cover of B, assume S does not have a finite subcover. This means that for any finite subset $S' = \{U_i\}_{i \in I'}$ where I' is a finite subset of B, there must exist some ultrafilter F containing none of the elements in I', in particular this means it must contain $-i$ for all $i \in I'$. By closure under finite joins this means that $\bigvee_{i \in I'} -i \neq 0$, in particular this means that the set $\{-x \in B : x \in I\}$ has the finite intersection property translated to the language of boolean algebras. From this it follows that it extends into an ultrafilter F'. Assume there exists a $U_x \in S$ such that $F' \in U_x$, this means $x \in F'$ but by construction $-x \in F'$ so $0 \in F'$ which contradicts it being a filter. Thus S is not actually an open cover, since it does not cover F'. □

Notably theorem 4.4 uses the ultrafilter lemma when constructing F', however it does not use the axiom of choice.

Lemma 4.5. *An ultrafilter in the Lindenbaum-Tarski algebra is a complete, deductively closed, and consistent theory.*

Proof. Let F be an ultrafilter in the Lindenbaum-Tarski algebra. Clearly it is complete as for all equivalence classes φ either $\varphi \in F$ or $\neg\varphi \in F$. It is deductively closed as all finite sets of formulas $\{\varphi_n\}_{n \in N}$ for some arbitrary $N \in \mathbb{N}$ can be written as a single formula $\varphi_0 \wedge ... \wedge \varphi_{N-1}$, and if $\varphi \leq \psi$ and $\varphi \in F$ then $\psi \in F$. Since F is deductively closed, consistency follows form it not containing 0. □

Theorem 4.6. *Any finitely satisfiable theory is satisfiable.*

Proof. Assume $S = \{\varphi_i\}_{i \in I}$ is an inconsistent theory, then the set $S' = \{U_{\neg\varphi_i}\}_{i \in I}$ is an open cover since for all ultrafilters F, there must exist at least one element of S not in F as otherwise F would be inconsistent and contain 0. By topological compactness, this means there must exist a finite subcover of S', which in particular means that S is not finitely satisfiable. By contraposition this proves the compactness theorem. □

Since this proof does not use the axiom of choice, rather only using the ultrafilter lemma, this shows that compactness is strictly weaker than choice as the ultrafilter lemma is strictly weaker than choice.

Chapter 5

Proof by Henkin constructions

The idea behind this proof is to construct a model of a finitely satisfiable theory by expanding the language to include all the elements of this model as constants. This proof closely follows structure of [6, p. 35-38].

Definition 5.1. A theory T in the language L is said to have the witness property if for all formulas $\varphi(x)$ with exactly 1 free variable, there exists some constant symbol c in L s.t. $\exists x\varphi(x) \to \varphi(c) \in T$.

Definition 5.2. A theory T is called maximal if for all sentences φ, T includes φ or T includes $\neg\varphi$

Lemma 5.3. *Assume T is maximal and finitely satisfiable. Let T' be a finite subset of T and φ some sentence s.t. $T' \models \varphi$. Then $\varphi \in T$.*

Proof. Assume $\varphi \notin T$. By maximality, $\neg\varphi \in T$. Since T is finitely satisfiable T' must be satisfiable, as must $T' \cup \{\varphi\}$. However by assumption all models of T' satisfy φ, so as must all models of $T' \cup \{\varphi\}$, however this shows that $T' \cup \{\varphi\}$ is inconsistent which contradicts it being satisfiable. By contradiction this implies $\varphi \in T$. □

Lemma 5.4. *Let L be a language and T be a maximal, finitely satisfiable L theory with the witness property. Then T has a model.*

Proof. Let c, c' be constant symbols in L. Let $c \sim c'$ if $c = c' \in T$.
Clearly for all constant symbols c, c', c''; $c = c \in T$, $c = c' \to c' = c \in T$, and $c = c' \wedge c' = c'' \to c = c'' \in T$, as all of these hold in all structures, so maximality and finite satisfiability guarantee they must be in T.
The goal will be to explicitly construct a model M. Let the underlying set of M be $C/\sim$ where C is the set of constant symbols in L.
Let R be an n-ary relation symbol in L. $R^M(c_1, ..., c_n)$ is true iff T contains $R(c'_1, ..., c'_n)$ for some representatives $c'_1, ..., c'_n$. This is clearly well defined as

relational truth is preserved under substitution by equal terms.
Let f be an n-ary function symbol in L, $f^M(c_1, ..., c_n) = c$ iff T contains $f(c'_1, ..., c'_n) = c'$ for some representatives $c'_1, ..., c'_n, c'$. This is well defined as function value is preserved under substitution by equal terms, and such a value always exists by the witness property applied to the formula $\varphi(x) = f(c'_1, ..., c'_n) = x$.
Let t be an L-term with n free variables. The next goal will be to prove that $t^M(c_1, ..., c_n) = c$ iff $T \models t(c'_1, ..., c'_n) = c'$ for some choice of representatives (as in the previous case the choice is irrelevant as term value is preserved under substitution by equal terms).
Assume that $t(c'_1, ..., c'_n) = c' \in T$, proceed by induction on the construction of terms.

1. If t is a single constant, then t is in the same $\sim$ equivalence class as c', so by construction $t^M = c$
2. If t is a single variable, then $t^M(c_1)$ is the case above.
3. If t is of the form $f(t_1, ..., t_k)$ for a k-ary function symbol f. Applying the witness property to $\varphi_i(t_i(c'_1, ..., c'_n) = x)$ shows that there exists some d'_i s.t. $t_i(c'_1, ..., c'_n) = d'_i \in T$. Since this is a finite set of formulas, one can apply lemma 5.3 to show that $t(c'_1, ..., c'_n) = f(d'_1, ..., d'_k) \in T$. By transitivity this means that $c' = f(d'_1, ..., d'_k) \in T$, call this ψ. By the induction hypothesis $t_i^M(c_1, ..., c_n) = d_i$, so by the recursive definition of term interpretation $t^M(c_1, ..., c_n) = f^M(d_1, ..., d_k)$. By ψ, $t^M(c_1, ..., c_n) = c$, which is the desired result.

Assume $t^m(c_1, ..., c_n) = c$, by the witness property applied to $\varphi(x) = t(c_1, ..., c_n) = x$, there must exist some d' s.t. $t(c_1, ..., c_n) = d' \in T$. By the previous induction, $t^M(c_1, ..., c_n) = d$, thus by transitivity $c' = d'$. Applying lemma 5.3 on the theory $T' = \{c' = d', t(c_1, ..., c_n) = d'\}$, we have that $t(c_1, ..., c_n) = c' \in T$.
Finally it is time to prove that this structure actually models T. This is done by induction on formulas with the goal to prove that for all formulas $\varphi(x_1, ..., x_n)$, and all choices of constants $c'_1, ..., c'_n$ $M \models \varphi(c_1, ..., c_n)$ iff $\varphi(c'_1, ..., c'_n) \in T$.

1. Let $\varphi = (t_1 = t_2)$. The witness property shows that there exists d'_1, d'_2 s.t. $t_i(c'_1, ..., c'_n) = d'_i \in T$. By the previous step, $t_i^M(c_1, ..., c_n) = d_i$, so
   $M \models \varphi(c_1, ..., c_n)$
   $\Leftrightarrow d_1 = d_2$
   $\Leftrightarrow d'_1 = d'_2 \in T$
   $\Leftrightarrow t_1(c'_1, ..., c'_n) = t_2(c'_1, ..., c'_n) \in T$.
2. Let $\varphi = R(t_1, ..., t_k)$. The witness property shows that there exists $d'_1, ..., d'_k$ s.t. $t_i(c'_1, ..., c'_n) = d'_k \in T$. By the previous step, $t_i^M(c_1, ..., c_n) = d_i$, so
   $M \models \varphi(c_1, ..., c_n)$
   $\Leftrightarrow R^M(d_1, ..., d_k)$
   $\Leftrightarrow R(d'_1, ..., d'_k) \in T$
   $\Leftrightarrow R(t_1, ..., t_k) \in T$.

3. Assume $\varphi = \neg\psi$.
   $M \models \neg\psi(c_1, ..., c_n)$
   $\Leftrightarrow M \not\models \psi(c_1, ..., c_n)$
   $\Rightarrow$ (by induction) $\psi(c_1', ...c_n') \notin T$
   $\Rightarrow$ (by maximality $\neg\psi(c_1', ..., c_n') \in T$
   If instead $\neg\psi(c_1', ..., c_n') \in T$, by finite satisfiability $\psi(c_1', ..., c_n') \notin T$, thus by induction hypothesis $M \not\models \psi(c_1, ..., c_n)$, so $M \models \varphi(c_1, ..., c_n)$.
   Assume $\varphi = \psi \wedge \sigma$.
   $M \models \psi(c_1, ..., c_n) \wedge \sigma(c_1, ..., c_n)$
   $\Leftrightarrow \psi(c_1', ..., c_n') \in T$ and $\sigma(c_1', ..., c_n') \in T$
   $\Leftrightarrow$ (by lemma 7.3($\psi(c_1', ..., c_n') \wedge \sigma(c_1', ..., c_n') \in T$
   Assume $\varphi = \exists x\psi(x)$.
   $M \models \psi(c, c_1, ..., c_n)$
   $\Rightarrow$ (by induction) $\psi(c', c_1', ..., c_n') \in T$, thus by lemma 7.3, $\varphi(c_1', ..., c_n') \in T$.
   If instead $\psi(c', c_1', ..., c_n') \in T$, by the induction hypothesis $M \models \psi(c_1, ..., c_n)$, thus $M \models \varphi(c_1, ..., c_n)$.

Thus $M \models T$. □

What remains is to show that every finitely satisfiable theory in any language can extend to this desired form.

Lemma 5.5. *Let L_0 be a language and T_0 be a finitely satisfiable theory over L_0. There exists a language L and a finitely satisfiable theory T over L such that $L_0 \subset L$, $T_0 \subset T$, and such that any extension T' of T has the witness property.*

Proof. This language and theory will be constructed as the unions of infinite sequences. Given L_n and T_n one constructs L_{n+1} and T_{n+1} as follows:
For each formula φ in L_n with one free variable, let c_φ be a constant symbol not in L_n, let $\psi_\varphi = \exists x\varphi(x) \to \varphi(c_\varphi)$. Let $L_{n+1} = L_n \cup \{c_\varphi : \varphi$ is a formula with 1 variable in $L_n\}$. Let $T_{n+1} = T_n \cup \{\psi_\varphi : \varphi$is a formula with 1 free variable in $L_n\}$.
It is important to show that T_n being finitely satisfiable implies T_{n+1} is finitely satisfiable. Let F be a finite subset of T_{n+1}, then $F = F' \cup \{\psi_{\varphi_i}\}$ for some finite subset F' of T_n and some finite collection of formulas in L with 1 free variable. Because T_n is finitely satisfiable F' has a model M. M also extends to a L_{n+1}-structure M' by interpreting c_φ as follows:

$$\begin{cases} \text{An arbitrary witness} & \text{If } \exists x\varphi(x) \\ \text{An arbitrary element} & \text{else} \end{cases}$$

Clearly $M' \models \psi_\varphi$ for all such formulas φ, thus $M' \models F$, so T_{n+1} is finitely satisfiable.
As stated before, $L = \bigcup L_n$ and $T = \bigcup T_n$. Clearly T has the witness property and is finitely satisfiable as any finite subset $F \subset T$ is a subset of some T_i. Any extension of T keeps all of the formulas relevant to the witness property. □

Notably since c_φ is chosen arbitrarily either among all elements or witnesses, this proof utilizes the axiom of choice.

Lemma 5.6. *Let T be a finitely satisfiable theory and φ a sentence, then either $T \cup \{\varphi\}$ or $T \cup \{\varphi\}$ is finitely satisfiable.*

Proof. Assume $T \cup \{\varphi\}$ is not finitely satisfiable, then by assumption of finite satisfiability of T, there must exist a finite $F \subset T$ such that $F \models \neg\varphi$, but then if $T \cup \{\varphi\}$ is not finitely satisfiable, then there exists some finite subset $F' \subset T$ such that $F' \cup \{\neg\varphi\}$, however then $F \cup F'$ is not satisfiable which contradicts finite satisfiability of T, □

Lemma 5.7. *Let T be a finitely satisfiable theory, there exists a maximal finitely satisfiable extension of T*

Proof. Consider the poset of all finitely satisfiable extensions of T ordered by the subset relation. Consider some infinite chain $T_0, ...$ Let $T' = \bigcup T_i$. T' is finitely satisfiable as any finite subset of T exists in some T_i, thus any chain has an upper bound so by Zorn's lemma there exists a maximal finitely satisfiable extension of T. □

Given all of this the compactness theorem follows immediately.

Theorem 5.8. *Any finitely satisfiable theory T is satisfiable.*

Proof. Utilize lemma 5.5 to extend T to a finitely satisfiable theory whose extensions all have the witness property, then use lemma 5.7 to extend this theory to a maximal finitely satisfiable theory, since this is an extension it still has the witness property. Finally use lemma 5.4 to construct a model of this theory. This model will also be a model of the original theory. □

Chapter 6

Applications of Compactness and Conclusion

6.1 Applications of Compactness

The most well known application of compactness is likely the following

Theorem 6.1. *Any theory with arbitrarily large finite models has infinite models.*

Proof. Let $\varphi_n := \exists x_0, ..., \exists x_{n-1}(\bigwedge_{n,m<n} \neg(x_n = x_m))$. In natural language, φ says that there exists at least n many elements. If theory T has arbitrarily large finite models, then $T \cup \{\varphi_n : n \in \mathbb{N}\}$ is finitely satisfiable as any finite subset S has a maximal k such that $\varphi_k \in S$, thus the full theory is satisfiable, thus there exists a model with at least n elements for all naturals n. This means there exists some infinite model of T. $\square$

This general structure of proof can be used to prove similar things for any type of size constraint that can be bounded below at arbitrarily high values by first order formulas. Above is the case for finite cardinalities, but this also works for dimensions of vector spaces, characteristics of fields although in this case infinite is replaced with 0, the amount of edges an element in a graph can have, and more.
A particularly interesting way to extend this method is illuminated by the following example.

Theorem 6.2. *Let $Th(\mathbb{N})$ be the first order theory of the naturals as an ordered semiring, there exists a model of $Th(\mathbb{N})$ that has elements larger than any standard natural number, where a standard natural number is defined as the interpretation of a term of the form $0 + 1 + ... + 1$ for a possibly 0 finite number of $+1$s.*

Proof. First we add a new constant symbol c to the language, then we consider the sequence of formulas $\varphi_n := (c > 1+...+1)$ with n 1s, clearly $Th(\mathbb{N}) \cup \{\varphi_n\}_{n\in\mathbb{N}}$ is finitely satisfiable, as for the maximal φ_k in any finite subset, one can simply interpret c as $n+1$. From here compactness tells us that there exists a model in this new language where c acts as a nonstandard natural number. At this point one can simply forget the interpretation of c and get the same model, with the same nonstandard natural, in the desired language. □

There also exists an interesting application of this technique where one moves into an expanded language to showcase the existence of a very small number.

Theorem 6.3. *Let $Th(\mathbb{R})$ be the first-order theory of the real numbers as an ordered field. There exists a model of $Th(\mathbb{R})$ with a positive element smaller than the inverses of all natural numbers, where natural numbers are interpreted as before.*

Proof. We begin by extending the language with a constant symbol c. Consider the sequence of formulas $\varphi_n := c > 0 \wedge \forall x(x \cdot (1 + ... + 1) = 1 \Rightarrow c < x$ where there are n 1s in the sum. In natural language this formula says that c is positive and is less than any multiplicative inverse of n. This is finitely satisfiable as for a finite subset one can consider the maximal k such that φ_k is in the subset, and simply choose $\frac{1}{k+1}$. Using compactness then creates a model with a single element c that is both positive and less than all inverse naturals, and once the interpretation of c is forgotten this is a model of $Th(\mathbb{R})$ in the desired language. □

6.2 Conclusion

In this thesis three proofs of the Compactness theorem for First-Order Logic have been provided, showcasing highly different approaches to this theorem. This process has acted as both a way to explore numerous different fields of mathematics: order theory, topology, and model theory; and as a way to compare different proofs in terms of their use of the axiom of choice. In particular it was found that the proofs using Łoś's theorem and Henkin constructions require the full power of the axiom of choice - at least without nontrivial modification - while the proof using Stone spaces of Boolean algebras only requires the ultrafilter lemma. Finally the thesis has presented a number of standard use cases of the Compactness theorem in order to show why it is considered one of the most important theorems in the field of mathematical logic.

Bibliography

[1] Thomas Jech. *Set Theory: The Third Millennium Edition, revised and expanded*. Springer, 2003.

[2] Thomas Jech. *Axiom of Choice*. Springer, 2006.

[3] Paul E. Howard. Łoś' theorem and the boolean prime ideal theorem imply the axiom of choice. *Proceedings of the American Mathematical Society*, 49:426–428, 1975.

[4] Stanley N. Burris and H.P. Sankappanavar. *A Course in Universal Algebra*. Burris Sankappanavar, 2012.

[5] François G. Dorais. Why is compactness in logic called compactness? Mathematics Stack Exchange, 2010.

[6] David Marker. *Model theory: an Introduction*. Springer, 2002.