THE COMPACTNESS THEOREM AND APPLICATIONS

BEN CALL

Abstract. In this paper we develop the basic principles of first-order logic, and then seek to prove the Compactness Theorem and examine some of its applications. We then seek to provide further areas for an interested reader to study.

Contents

1. Introduction

In this paper, we seek to provide a basic understanding of the mechanics of first-order logic necessary to prove the Compactness Theorem, which states that a set of sentences Σ has a model if and only if every finite subset of Σ has a model. After proving this, we will examine some of the applications of this theorem, specifically, the fact that the Archimedean Property of the real numbers is not a property of ordered fields in general. This paper is meant to be accessible to those who have no prior experience studying logic, so at times, we will sacrifice formalism in favor of intuition and clarity.

2. First-Order Logic

In order to use the Compactness Theorem, and in fact, even to state it, we must first develop the logical language to which it applies. In this case, we will use first-order logic. We will begin by listing the requisite definitions.

Definition 2.1. A **language** $\mathscr{L}$ is a not necessarily countable collection of relation symbols P, function symbols G, and constant symbols c.

Definition 2.2. A **universe** A is a nonempty set.

Definition 2.3. An **interpretation function** $\mathscr{I}$ is a function such that:

(1) For each n-place relation symbol P of $\mathscr{L}$, $\mathscr{I}(P) = R$ where $R \subset A^n$.

Date: August 2, 2013.

(2) For each m-place function symbol G of $\mathscr{L}$, $\mathscr{I}(G) = F$ where $F : A^m \longrightarrow A$ is an m-place function on A.
(3) For each constant symbol c, $\mathscr{I}(c) = x$ for some $x \in A$.

We now have enough basic machinery to give one of the central definitions used in the Compactness Theorem.

Definition 2.4. A **model** $\mathfrak{U}$ for a language $\mathscr{L}$ consists of a universe A and an interpretation function $\mathscr{I}$, which we denote

$$\mathfrak{U} = \langle A, \mathscr{I} \rangle.$$

For those occasions where the interpretation function is general knowledge, such as interpreting the two-place function $+$, we simply list the functions, relations, and constants used without specifying what they are. In this paper, this only occurs in Section 4.

Before moving on, we will attempt to provide some sense of intuition of what a model really is. This intuition can be derived simply from examining the name itself – "model". A model provides a world, or universe, for a language to act upon, and then provides the structure of that universe when it is acted upon by that language with the interpretation function. Thus, there are many possible models for any given language $\mathscr{L}$ because each different universe and interpretation function creates a new one.

This concept of having multiple models for a language does us no good by itself. Instead, we seek to find some way to compare these models to each other, and to do that, we turn to the concept of formulas and sentences, typically denoted by lower-case Greek letters, such as φ, ψ, and θ.

In order to define a formula, we must first provide some more basic terminology.

Definition 2.5. A **term** is one of four things:

i) A variable is a term.
ii) A constant symbol is a term.
iii) If F is an m-placed function symbol, and $t_1, \cdots, t_m$ are terms, then $F(t_1 \cdots t_m)$ is a term.
iv) A string of symbols is a term only if it can be shown to be a term by a finite number of applications of (i) - (iii).

The purpose of (iv) is to ensure that there are no infinite terms.

Now before we continue, we note that there is one two-place relation symbol which always belongs to first-order logic, even though it does not belong to $\mathscr{L}$. This relation is called the identity relation, and is denoted by $\equiv$.

Definition 2.6. An **atomic formula** of $\mathscr{L}$ is a string of the form:

i) $t_1 \equiv t_2$ where t_1 and t_2 are terms of $\mathscr{L}$.
ii) $P(t_1 \cdots t_n)$, where P is an n-placed relation and $t_1, \cdots, t_n$ are terms of $\mathscr{L}$.

With these definitions, we are finally able to define a formula.

Definition 2.7. A **formula** of $\mathscr{L}$ is defined as follows:

i) An atomic formula is a formula.
ii) If φ and ψ are formulas, then $(\varphi \wedge \psi)$ and $(\neg\varphi)$ are formulas.
iii) If v is a variable and φ is a formula, then $(\forall v)\varphi$ is a formula.

iv) If a string of symbols can be shown to be a formula by a finite number of applications of (i) - (iii), then it is a formula.

Definition 2.8. A formula is a **sentence** if every variable in the formula is bound by the quantifier $\exists$ or $\forall$.

In the following example, P is an n-place relation, x is a variable, and $t_1, \cdots, t_n$ are terms of $\mathscr{L}$ that have no variables.

Example 2.9. By Definition 2.6, both $P(x, t_2, \cdots, t_n)$ and $P(t_1, \cdots, t_n)$ are atomic formulas of $\mathscr{L}$, so by 2.7, they are formulas as well. However, only $P(t_1, \cdots, t_n)$ is a sentence, because there are no variables that need to be bound by a quantifier.

However, $(\forall x)P(x, t_2, \cdots, t_n)$ and $(\exists x)P(x, t_2, \cdots, t_n)$ are both sentences of $\mathscr{L}$. This is because $(\forall x)P(x, t_2, \cdots, t_n)$ is a formula by 2.7.iii, but its variable is bound by the quantifier $\forall$, so it is a sentence as well. We use a similar argument for $(\exists x)P(x, t_2, \cdots, t_n)$.

Notice that a sentence depends only on the language, and not any specific model of a language. However, with one simple concept, we may use sentences to compare models. This concept is "truth". We neglect to give a formal definition in favor of providing an intuitive definition:

Nondefinition 2.10. A sentence φ is **true in a model** $\mathfrak{A}$, or alternatively, $\mathfrak{A}$ **is a model of** φ, denoted $\mathfrak{A} \models \varphi$, if for every possible sequence of elements in A, substituting these elements in for the variables present in φ yields a true sentence.

Note that this idea of truth precludes the possibility of both φ and $\neg\varphi$ holding in $\mathfrak{A}$.

Now that we have this definition, we can begin to characterize the way in which a model interacts with a language. We now extend this definition of truth to a set of sentences, Σ. We say that $\mathfrak{A}$ is a model of a set of sentences Σ if $\mathfrak{A}$ is a model of φ for all $\varphi \in \Sigma$. We now come to another definition, which allows us to relate two sentences of a language with respect to models:

Definition 2.11. A sentence φ is a **consequence** of a set of sentences Σ, denoted $\Sigma \models \varphi$, if every model of Σ is a model of φ.

While this definition is informative, we must provide what we will eventually show to be an equivalent definition. However, by using the following definition, we only have to deal with a finite number of sentences, which is key to our proof of the Compactness Theorem.

Definition 2.12. A sentence φ is **deducible** from Σ, expressed $\Sigma \vdash \varphi$, if there exists a finite chain of sentences $\psi_0, \psi_1, \cdots, \psi_n$ where ψ_n is φ and each previous sentence in the chain either belongs to Σ, follows from one of the logical axioms, or can be inferred from previous sentences.[1]

Using this idea of deducibility, we can say something about sets of sentences as a whole – namely, whether or not they are consistent.

Definition 2.13. A set of sentences Σ is consistent if and only if there does not exist a sentence φ such that $\Sigma \vdash \varphi$ and $\Sigma \vdash \neg\varphi$.

[1]We omit the logical axioms and rules of inference specific to first-order logic as they do not play any prominent role in any of the following proofs. If the reader wishes to learn more, they can be found in more detail on pages 24 and 25 of Chang and Keisler's *Model Theory*.

We now have all of the requisite definitions to construct a proof of the Compactness Theorem of first-order logic, so we give one final necessary lemma, which is actually an equivalent formulation of the Soundness Theorem. For brevity, we omit the proof.

Lemma 2.14. *Let Σ be a set of sentences. If $\Sigma \vdash \varphi$, then $\Sigma \models \varphi$.*

3. Proof of the Compactness Theorem

Before we prove the Compactness Theorem, we will first discuss the Completeness and Soundness Theorems, which will give us the tools necessary to prove the Compactness Theorem.

Theorem 3.1 (Soundness Theorem)**.** *Let Σ be a set of formulas. If Σ has a model, then it is consistent.*

Proof. Suppose that Σ is not consistent. Then there exists some φ such that $\Sigma \vdash \varphi$ and $\Sigma \vdash \neg\varphi$. By Lemma 2.14, we have $\Sigma \models \varphi$ and $\Sigma \models \neg\varphi$. This means that Σ cannot have a model $\mathfrak{U}$, because if it did, then $\mathfrak{U} \models \varphi$ and $\mathfrak{U} \models \neg\varphi$, which is impossible. Thus, we have proved our statement. $\square$

The following theorem was first proved by Gödel in 1930 for countable languages, and then was later generalized to uncountable languages as well. In this paper, we omit the proof, and give an outline of the general steps that the proof takes.[2]

Theorem 3.2 (Gödel's Completeness Theorem)**.** *Let Σ be a set of formulas. If Σ is consistent, then it has a model.*

Proof. Let Σ be an arbitrary consistent set of sentences of some language $\mathscr{L}$. Let $\overline{\mathscr{L}}$ be an expansion of $\mathscr{L}$ created by adding a set of new constant symbols not in $\mathscr{L}$ that has the same cardinality as $\mathscr{L}$. The first step is to add sentences to Σ to create a consistent set of sentences $\overline{\Sigma}$ in the language $\overline{\mathscr{L}}$. It is possible to show that $\overline{\Sigma}$ has a model $\mathfrak{U}$ which is a model for $\overline{\mathscr{L}}$. Now if we let $\mathfrak{B}$ be the reduction of $\mathfrak{U}$ to only involve the original language $\mathscr{L}$, it is possible to show that $\mathfrak{B}$ is a model for Σ, because the sentences in Σ do not involve any constants which belonged to $\overline{\mathscr{L}}$, so the reduction of $\mathfrak{U}$ to $\mathfrak{B}$ did not affect its ability to model Σ. $\square$

Finally, we reach the main goal of this paper: to prove the Compactness Theorem of first-order logic.

Theorem 3.3 (Compactness Theorem)**.** *A set of sentences Σ has a model if and only if every finite subset of Σ has a model.*

Proof. The forward direction of this proof, that if Σ has a model then every finite subset of Σ has a model, is trivial. Thus, we will turn directly to the more interesting direction.

Suppose for contradiction that every finite subset of Σ has a model but Σ does not have a model. Then by the contrapositive of the Completeness Theorem, we see that Σ is inconsistent. Consequently, there exists a sentence φ such that $\Sigma \vdash \varphi$ and $\Sigma \vdash \neg\varphi$. Now, each of these deductions consists of a finite chain of sentences.

[2]Should the reader wish to examine the specific proof, Enderton gives the original proof on page 135, while Chang and Keisler give a proof by Henkin on page 61. The former proof is more immediately accessible, but the second proof provides an interesting model-theoretic perspective. The outline we give here is due to Chang and Keisler.

Therefore, if we let Σ' be the set of sentences of Σ involved in the deduction of φ and Σ'' be the set of sentences involved in the deduction of $\neg\varphi$, then we have two finite subsets of Σ such that $\Sigma' \vdash \varphi$ and $\Sigma'' \vdash \neg\varphi$. It is clear that $\Sigma' \cup \Sigma'' \vdash \varphi$ and $\Sigma' \cup \Sigma'' \vdash \neg\varphi$, because the sentences involved in the original deductions still belong to the union of the two finite subsets. Thus, $\Sigma' \cup \Sigma''$ is inconsistent, and consequently, it does not have a model. We've produced a finite subset of Σ which does not have a model, so we've reached the desired contradiction.

□

4. Applications for the Compactness Theorem

With the proof of our theorem complete, we now continue with some of the applications for which it can be used. However, we must first provide some basic ideas about what sentences and models can do.

Definition 4.1. A **theory** of $\mathscr{L}$ is a collection of sentences of $\mathscr{L}$.

Definition 4.2. A set of **axioms** for a theory is a set of sentences which has the same exact set of consequences in first-order logic as the theory itself.

Using this definition, we see that if we are given some theory Σ which has Δ as a set of axioms, then Δ, to some extent, characterizes Σ.

Example 4.3. Every one of the field axioms is expressible in first-order logic. Consequently, we consider every field to be an example of a theory which has the field axioms as a set of axioms. By adding the axioms for an ordering, which are also expressible in first-order logic, we can obtain a set of axioms for an ordered field.

Suppose that Δ is a set of axioms for two theories, Σ and Σ'. From our definition, we know that Σ and Σ' have the same exact set of consequences in first-order logic. Thus, if they differ, we know two things: first, the properties that they differ by are not expressible in first-order logic, and second, the properties that they differ by are not consequences of Δ in first-order logic.

Consider the Archimedean Property of the real numbers. This states:

$$\text{For all real } a, b > 0, \text{ there exists a natural number } n \text{ such that } na \geq b.$$

We will show in the following example that it is possible to construct a theory which has the same consequences in first-order logic as the axioms for an ordered field, but does not have the Archimedean Property. This demonstrates that the Archimedean Property is not expressible in first-order logic and it is not a consequence of the field axioms.

Example 4.4. There exists an ordered field $\mathbb{F}$ that is not Archimedean.

Proof. Let $\mathscr{L} = \{+, \cdot, 0, 1, \leq\}$ and let Γ be the set of all sentences which hold in the ordered field of real numbers. Now let x be a constant symbol which is not 0 or 1, and let $\Sigma = \Gamma \cup \{0 < x\} \cup \{1 \leq x\} \cup \{2 \leq x\} \cup \{3 \leq x\} \cup \cdots$. This set of additional sentences is a way of stating in first-order logic that x is larger than every natural number. This is also an infinite collection of sentences, which is where the Compactness Theorem comes into play, by showing that Σ has a model.

Let Σ' be an arbitrary finite subset of Σ. Now, we must show that Σ' has a model. It is clear that $\Sigma' \cap \Gamma$ has a model because it is simply a subset of Γ, which is modeled by the axioms for the real numbers. Now for all sentences $\varphi \in \Sigma' \setminus \Gamma$, we can expand the model of the real numbers to include φ, because the sentence φ does not have any interaction with any of the consequences of Γ, because it involves a constant symbol which is not in Γ, namely x. Thus, every finite subset of Σ has a model, so by the Compactness Theorem, Σ has a model which clearly satisfies all of the properties of an ordered field. However, if we examine Σ, we see that while both 1 and x are positive, by construction there does not exist a natural number n such that $n \cdot 1 \geq x$. Therefore, in this model, we have an ordered field which does not have the Archimedean Property. □

Another powerful result whose proof can be derived using the Compactness Theorem relates algebraically closed fields to one another. It states that every sentence φ of first-order logic in the language of fields is true in every algebraically closed field of characteristic 0 if and only if there exists $p > 0$ such that φ holds in every algebraically closed field of characteristic m where m is larger than p [3, p.641]. We omit the proof because it would take up too much space in this paper.

5. Where to go from here?

Thus far, we have provided the basics of first-order logic, the beginnings of model theory, a proof of the Compactness Theorem, and some of its applications. However, this paper leaves many areas for the reader to explore later, which we seek to discuss briefly in this section.

First and foremost, for a more in-depth discussion of various forms of logic, whether it is sentential, first-order, or second-order, Enderton provides a good reference in *An Introduction to Mathematical Logic.* For a very rigorous introduction to model theory, *Model Theory* by Chang and Keisler is an outstanding reference. However, it is very helpful, although not necessary, to have some background in logic before reading.

Readers who found the applications of the Compactness Theorem interesting may wish to look into the field of nonstandard analysis, which revolves around creating nonstandard models that are isomorphic to the standard models and proving results in this new model. These results can then sometimes be transported between models, with the alternate proofs providing deeper insights. For example, Abraham Robinson used a non-standard model of the reals to provide a rigorous justification for the use of infinitesimals in calculus as opposed to limits [2, p.173].

Finally, examining alternate proofs of the Completeness Theorem provides an interesting perspective on the different approaches and uses of logic. Additionally, for those who wish to study why exactly first-order logic is used as opposed to some other logical system, Lindström's Characterization of first-order logic is extremely interesting and rewarding. A good explanation and proof can be found in section 2.5 of *Model Theory.*

Acknowledgements. I would like to thank my mentor Olga Turanova for all of her help in guiding me through my first exposure to logic as well as the process of writing and revising this paper. I would also like to thank Maryanthe Malliaris and Jonathan Stephenson for their suggestions of books for me to read, and finally Peter May for organizing this REU.

References

[1] Chang, Chen Chung, and H. Jerome Keisler. *Model Theory*. Amsterdam: North-Holland Pub. 1973.

[2] Enderton, Herbert B. *A Mathematical Introduction to Logic*. San Diego, CA: Academic, 2001.

[3] Gowers, Timothy, June Barrow-Green, and Imre Leader. *The Princeton Companion to Mathematics*. Princeton: Princeton UP, 2008. 641-42.