LINDSTRÖM'S THEOREM

SALMAN SIDDIQI

Abstract. This paper attempts to serve as an introduction to abstract model theory. We introduce the notion of abstract logics, explore first-order logic as an instance of and as the basis for abstract logics, and end by proving Lindström's theorem.

Contents

0. Introduction

This paper hopes to serve as an introduction to abstract model theory, though I cannot profess to know but very little of the subject. On a general note, we will assume set theory: specifically, Zermelo-Fraenkel set theory with the axiom of choice. We will often use the axiom of choice or Zorn's lemma – particularly in Section 3 – without citation, and we will index sets with ordinals with little guilt.

Abstract model theory concerns itself with extensions of first-order logic, and its starting point and perhaps most well-known result is Lindström's theorem. Lindström's theorem states that first-order logic is completely characterized among its extensions by two of its more basic theorems: the downward Löwenheim-Skolem theorem and the compactness theorem. As a consequence of this, at least one of these must fail in any proper extension of first-order logic.

One of the central tasks of abstract model theory, as noted by Barwise (1974), is to determine the relationship between gaining expressive power in extending first-order logic and losing 'control' over the extension. The motivation for moving to extensions of first-order logic is simple: a significant amount of mathematics cannot be expressed in terms of first-order sentences. Lindström's theorem, however, already places a restriction on the usefulness of any proper extension of first-order logic: at least one of two theorems which help make first-order logic manageable does not hold. We aim to provide a proof for Lindström's theorem, as well as develop the concepts in logic and model theory necessary in doing so.

1. Abstract Logics

Among the basic concepts of model theory are those of a language and a model, and we will require both of these in order to state any other definitions.

Definition 1.1. A *language* $\mathscr{L}$ is the union of three disjoint sets $\mathscr{L} = R \cup F \cup C$, as well as a function g, where R is a set of of *relation symbols*, F is a set of *function symbols* and C is a set of *constant symbols*, and g assigns to each relation and function symbol a natural number called its *arity*. We will require languages to be nonempty.

Example 1.2. The set $\mathscr{L} = \{\leq, \cdot, +, 0, 1\}$ is a language, where $\cdot$ and $+$ are binary function symbols, $\leq$ is a binary relation, and 0 and 1 are constant symbols. So is $\mathscr{L}' = \{\leq, S\}$, where $\leq$ is a binary relation symbol, and S is a unary function symbol.

This definition is relatively simple though, on a cautionary note, the reader should refrain from looking at the relation, function and constant *symbols* as relations, functions or constants. In particular, note that the arity of the relation and function symbols is not actually meaningful in any way, even though we would wish to identify the arity of a relation or function symbol with the arity of its interpretation in a model. Specifically, in the example above, an expression like $+(0, 1)$ or $S(0)$ would be purely formal constructions, which would only be meaningful within an abstract logic which provides syntactical rules to form such constructions. Finally, we will prefer to talk about the number of 'places' of a relation or function symbol; we note that we mean the same thing by the 'number of places' and the 'arity' of a given symbol.

Note that languages are sets, and as such have cardinalities; however, a more useful notion is what we shall call the 'power' of the language. The following definition may seem peculiar at first, but the motivation behind this is that the power of language is precisely the cardinality of the set of all finite strings we can form using symbols in the language. When we begin to discuss sentences in first-order logic, the usefulness of this definition will become clear.

Definition 1.3. The *power* of a language $\mathscr{L}$ is the maximum $\max(\aleph_0, |\mathscr{L}|)$, where $|\mathscr{L}|$ is the cardinality of the language. We often write $\|\mathscr{L}\|$ for the power of $\mathscr{L}$.

Example 1.4. The language $\mathscr{L} = \{0, 1, +, \cdot, \leq\}$ has power $\|\mathscr{L}\| = \aleph_0$. The language $\mathscr{L}'$ which has all real numbers as constant symbols, and no relation or function symbols, has power $\|\mathscr{L}'\| = 2^{\aleph_0}$.

Proposition 1.5. *Consider a language $\mathscr{L}$. The set of finite tuples of symbols in $\mathscr{L}$ has cardinality equal to the power of $\|\mathscr{L}\|$.*

Proof. Let $\alpha = |\mathscr{L}|$. If α is infinite, then for any particular n, the set of n-tuples has cardinality α. The union over n of these sets must then also have cardinality α. If α were finite, there would be only finitely many n-tuples for a given n, and the union over n would have cardinality $\aleph_0$. Hence, the cardinality of the set of finite tuples of symbols in $\mathscr{L}$ is $\max(\aleph_0, \alpha) = \|\mathscr{L}\|$. □

As we discussed before, symbols in a language have no inherent meaning. To assign meaning to the symbols in a language, we use the notion of a model.

Definition 1.6. A *model* $\mathfrak{A}$ for a language $\mathscr{L}$ is an ordered pair $(A, \mathscr{I})$ where A is a set of constants, often referred to as the *universe*, and $\mathscr{I}$ is the *interpretation function* which assigns to each n-placed relation symbol of $\mathscr{L}$ an n-placed relation on A, to each m-placed function symbol of $\mathscr{L}$ an m-placed function mapping from A^m to A, and constant symbol of $\mathscr{L}$ a constant in A.

Example 1.7. The real numbers with multiplication and addition defined normally is a model for the language $\mathscr{L} = \{\mathbf{0}, \mathbf{1}, \oplus, \odot\}$, with the interpretation function assigning $\mathbf{0}$, $\mathbf{1}$, $\oplus$ and $\odot$ to 0, 1, $+$ and $\cdot$ respectively.

If we refer to the 'size', 'power' or 'cardinality' of a model $\mathfrak{A} = (A, \mathscr{I})$, it will be understood that we simply mean the cardinality of A.

Given two models of the same language, there is a natural notion of 'corresponding' relations, functions and constants.

Definition 1.8. Let $\mathfrak{A} = (A, \mathscr{I}_A)$ and $\mathfrak{B} = (B, \mathscr{I}_B)$ be two models of the language $\mathscr{L}$, and consider relations R_A on $\mathfrak{A}$ and R_B on $\mathfrak{B}$. We say that R_A is the *corresponding relation* of R_B in $\mathfrak{A}$ if there is some P in $\mathscr{L}$ for which $\mathscr{I}_A(P) = R_A$ and $\mathscr{I}_B(P) = R_B$. The *corresponding functions* and *corresponding relations* are defined similarly.

It is easy to see that we can, in fact, generalize this concept of a 'corresponding' function to models with differing languages, as long as the intersection of the languages is non-empty.

Sometimes, we may wish to enlarge languages and models. We can define the concept of an 'expansion' of a model or language to help us more easily define what we mean by an abstract logic. Often, it is easier or more productive to deal with expansion or reducts of models or languages than it is to deal with the original models themselves.

Definition 1.9. We say that a language $\mathscr{L}$ is an *expansion* of a language $\mathscr{L}'$ if $\mathscr{L}' \subset \mathscr{L}$; in this case, we sometimes call $\mathscr{L}'$ a *reduction* of $\mathscr{L}$. When $\mathscr{L} \setminus \mathscr{L}'$ consists solely of constant symbols, we say that $\mathscr{L}$ is a *simple expansion* of $\mathscr{L}'$.

Definition 1.10. Let $\mathfrak{A} = (A, \mathscr{I})$ be a model in a language $\mathscr{L}$, and let $\mathscr{L}'$ be some reduction of $\mathscr{L}$. The *reduct* $\mathfrak{A}'$ of $\mathfrak{A}$ to $\mathscr{L}'$ is the restriction $\mathfrak{A}' = (A, \mathscr{I}|_{\mathscr{L}'})$. In this case, we also say that $\mathfrak{A}$ is the *expansion* of $\mathfrak{A}'$ to $\mathscr{L}'$.

Note that if we have a model $\mathfrak{A}$ for a language $\mathscr{L}$, we can always expand this to a model for some expansion $\mathscr{L}' = \mathscr{L} \cup X$ by simply specifying interpretations for the additional symbols in X.

A more general notion than that of an expansion of a model is that of an extension. With extensions, the universes need not be the same, though we still require that one is a subset of the other.

Definition 1.11. A model $\mathfrak{B} = (B, \mathscr{I}_B)$ is said to be an *extension* of a model $\mathfrak{A} = (A, \mathscr{I}_A)$ for a language $\mathscr{L}$ if we have $A \subset B$ and:

(i) For each n-placed relation R_A of $\mathfrak{A}$, there is a corresponding relation R_B of $\mathfrak{B}$ such that R_A is the restriction of R_B to A. As relations are just ordered n-tuples, we could write this as $R_A = R_B \cap A^n$.

(ii) For each m-placed function G_A of $\mathfrak{A}$, there is a corresponding function G_B of $\mathfrak{B}$ such that G_A is the restriction of G_B to A, or $G_A = G_B|_{A^m}$, with the additional requirement that $G_B(a_1, \ldots, a_m)$ be in A if $a_1, \ldots, a_m \in A$.

(iii) For each constant a in A, there is a corresponding constant b in B.

In this case, we also say that $\mathfrak{A}$ is a *submodel* of $\mathfrak{B}$.

When looking at submodels of a given model, there is a subtlety in the second part of the definition worth noting. Rather than simply restrict the domain of functions, notice that we also require the universe of a sub-model to remain closed under the function. The reason behind this is that $\mathfrak{A}$ is itself a model, and we initially required that the co-domain of the interpretations of function symbols remain within the universe. On another note, it should also be easy to see that, in the above definition, any language for which $\mathfrak{B}$ is a model must contain $\mathscr{L}$.

Naturally, we would also want some way of saying that two models are essentially 'the same', and so we obtain the following definition of an isomorphism of models.

Definition 1.12. Two models $\mathfrak{A} = (A, \mathscr{I}_A)$ and $\mathfrak{B} = (B, \mathscr{I}_B)$ for a language $\mathscr{L}$ are *isomorphic* if there is a bijection $f : A \to B$ with the requirement that:

(i) For each n-placed relation R_A of $\mathfrak{A}$ and the corresponding relation R_B of $\mathfrak{B}$, $R_A(x_1, \ldots, x_n)$[1] if and only if $R_B(f(x_1), \ldots, f(x_n))$.

(ii) For each m-placed function G_A of $\mathfrak{A}$ and the corresponding function G_B of $\mathfrak{B}$, $f(G_A(x_1, \ldots, x_n)) = G_B(f(x_1), \ldots, f(x_n))$.

(iii) For each constant x and the corresponding constant x', we have $f(x) = f(x')$.

In this case, we say that there is an isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$, and write $\mathfrak{A} \cong \mathfrak{B}$.

Note that we require there be a unique corresponding relation, function or constant in the above definition.

We now present the relatively lengthy definition of an abstract logic. An abstract logic, being the basis of abstract model theory, should be properly viewed as simply an extension of first-order logic. We do not define first-order logic first because, to some extent, the definition of an abstract logic makes that of first-order logic easier to understand. However, it is advisable to revisit this definition while reading through that of first-order logic.

At this point, we will assume that the languages we are dealing with do not contain function symbols, for the rest of this section as well as the final section. Doing otherwise would introduce needless complications into both our definition and the proof of Lindström's theorem. In particular, in the relativization property below, if we do not ignore function symbols in the language and the definition of a submodel, it is no longer true that a submodel $\mathfrak{B}$ of $\mathfrak{A}$ with the specified universe necessarily exists.

Definition 1.13. An *abstract logic* is an ordered pair $(l, \models_l)$ where l is the class of *sentences* of the logic and $\models_l$ is the *satisfaction relation* of the logic. We require abstract logics to satisfy the following:

(i) *(Occurrence)* For each $\varphi \in l$, there is a finite language $\mathscr{L}_\varphi$ associated to φ (which we call the set of *symbols occurring in* φ). For any model $\mathfrak{A}$ of the language $\mathscr{L}$, we require the statement $\mathfrak{A} \models_l \varphi$ to be either true or false if $\mathscr{L}_\varphi \subset \mathscr{L}$ and undefined otherwise.

---

[1] By the statement $R_A(x_1, \ldots, x_n)$, we mean that the ordered pair $(x_1, \ldots, x_n)$ is in R; we view relations as sets of ordered n-tuples.

(ii) *(Expansion)* If $\mathfrak{A}$ is a model of the language $\mathscr{L}$ containing $\mathscr{L}_\varphi$, and $\mathfrak{B}$ is an expansion of $\mathfrak{A}$ to some language $\mathscr{L}'$ containing $\mathscr{L}$, then if $\mathfrak{A} \vDash_l \varphi$, we must also have $\mathfrak{B} \vDash_l \varphi$.

(iii) *(Isomorphism)* If $\mathfrak{A} \cong \mathfrak{B}$ and $\mathfrak{A} \vDash_l \varphi$, then $\mathfrak{B} \vDash_l \varphi$.

(iv) *(Renaming)* Let f be a bijection between two languages $\mathscr{L}$ and $\mathscr{L}'$ which preserves the number of places for all symbols. For each model $\mathfrak{A}$ with language $\mathscr{L}$, let $\mathfrak{A}'$ be the model for $\mathscr{L}'$ induced by f (in the obvious way). If $\varphi \in l$ and $\mathscr{L}_\varphi \subset \mathscr{L}$, then there must be a sentence $f(\varphi) \in l$ (obtained by substituting symbols in φ by their images under f) with $\mathscr{L}_{f(\varphi)} = f(\mathscr{L}_\varphi)$ such that for each model $\mathfrak{A}$ for $\mathscr{L}$, we have $\mathfrak{A} \vDash_l \varphi$ if and only if $\mathfrak{A}' \vDash_l f(\varphi)$.

(v) *(Closure)* For any language $\mathscr{L}$, l contains all atomic sentences of the language (these will be defined in the next section), and is closed under the connectives $\wedge$ and $\neg$. That is, if $\varphi, \psi \in l$, then $\varphi \wedge \psi \in l$ and $\neg\varphi \in l$. Furthermore, if $\mathfrak{A}$ is a model of the language $\mathscr{L} = \mathscr{L}_\varphi \cup \mathscr{L}_\psi$, we require $\mathfrak{A} \vDash_l (\varphi \wedge \psi)$ if and only if $\mathfrak{A} \vDash_l \varphi$ and $\mathfrak{A} \vDash_l \psi$ are both true, and we require $\mathfrak{A} \vDash_l (\neg\varphi)$ if and only if $\mathfrak{A} \vDash_l \varphi$ is false.

(vi) *(Quantifiers)* For each $\varphi \in l$ and each constant symbol $c \in \mathscr{L}_\varphi$, there is a sentence $(\forall x_c)\varphi(x_c)$ where $\mathscr{L}_{\varphi(x_c)} = \mathscr{L}_\varphi \setminus \{c\}$. Furthermore, if we let $\mathfrak{A} = (A, \mathscr{I})$ be a model of the language $\mathscr{L}_{\varphi(x_c)}$ and $\mathfrak{A}_y = (A, \mathscr{I}')$ for $y \in A$ be the expansion of $\mathfrak{A}$ to $\mathscr{L}_\varphi$ by adding $\mathscr{I}'(c) = y$, we require:

$$\mathfrak{A} \vDash_l (\forall x_c)\varphi(x_c) \text{ if and only if } \mathfrak{A}_y \vDash_l \varphi \text{ for all } y \in A.$$

(vii) *(Relativization)* Let $\varphi \in l$ be a sentence, $\mathfrak{A} = (A, \mathscr{I})$ be a model, R be an $(n+1)$-placed relation on A and $b_1, \ldots, b_n$ be constants in A, with the requirement that neither R nor $b_1, \ldots, b_n$ are interpretations of symbols in $\mathscr{L}_\varphi$. There must then be a sentence $\varphi \mid G(x, c_1, \ldots, c_n)$, called the *relativization* of φ to $G(x, c_1, \ldots, c_n)$ in the expansion $\mathscr{L}' = \mathscr{L}_\varphi \cup \{G, c_1, \ldots, c_n\}$, where G is an $(n+1)$-placed relation symbol and $c_1, \ldots, c_n$ are constant symbols. If we expand the model $\mathfrak{A}$ to a model $\mathfrak{A}'$ of $\mathscr{L}'$ by adding the obvious interpretations for $G, c_1, \ldots, c_n$ and additionally consider a submodel $\mathfrak{B}$ of $\mathfrak{A}$ with universe $B = \{a \in A \mid R(a, b_1, \ldots, b_n)\}$, we further require that:

$$\mathfrak{A}' \vDash_l \varphi \mid G(x, c_1, \ldots, c_n) \text{ if and only if } \mathfrak{B} \vDash_l \varphi.$$

The satisfaction relation should be seen as a declaration of truth in the model. When we write $\mathfrak{A} \vDash \varphi$, we often say $\mathfrak{A}$ *satisfies* or *is a model of* φ, or φ *holds in*, *is true in* or *is satisfied by* $\mathfrak{A}$. We will also use the same expressions for sets of sentences T.

These properties will likely all become clear and intuitive after our definition of first-order logic, with the possible exception of the relativization property, which will likely appear peculiar regardless. The relativization property simply states that for any sentence, any given model of the language of the sentence and any relation on the universe of the model, one can 'carve out' a subset of the universe with the relation and there will then be a sentence which is true in the model if and only if the original sentence is true in the submodel restricted to the carved out subset.

2. First-Order Logic

The definition of an abstract logic, as will soon be apparent, simply aims to capture the structure of first-order logic: in a sense, first-order logic is the most 'basic' abstract logic, with just the bare essentials. The purpose of this section will be to specify what we mean by first-order logic, in an attempt to make the definition of an abstract logic more intelligible. We begin by introducing atomic sentences, with the remark that the following notion of an atomic sentences works just as well for abstract logics in general.

Definition 2.1. Let $\mathscr{L}$ be a language. Then the following are *terms*:

(i) Constant symbols.
(ii) $F(t_1, \ldots, t_m)$ where F is an m-placed function symbol and $t_1, \ldots, t_m$ are terms.
(iii) Any string of symbols that can be shown to be a term by a finite number of applications of the above.

Remark 2.2. Note that, when we write $F(t_1, \ldots, t_m)$, this is purely formal, though our intuitive understanding of F as taking $t_1, \ldots, t_m$ as arguments is often helpful.

The terms of a language form the basis of atomic sentences:

Definition 2.3. Let $\mathscr{L}$ be a language and let $\equiv$ be an identity relation.[2]. Then the following are *atomic sentences*:

(i) $t_1 \equiv t_2$, where t_1 and t_2 are terms.
(ii) $P(t_1, \ldots, t_n)$ where $t_1, \ldots, t_n$ are terms and P is a relation symbol.

When we required all atomic sentences of any given language to be in an abstract logic, this was the definition of an atomic sentence that we had in mind. It should be easy to see that we obtain the following definition of first-order logic if we simply attempt to construct a 'minimal' abstract logic.

Definition 2.4. *First-order logic* is the pair $(l_{\omega,\omega}, \models)$ where $l_{\omega,\omega}$ contains:

(i) All atomic sentences for any given language.
(ii) The sentence $\neg\varphi$ for any $\varphi \in l_{\omega,\omega}$.
(iii) The sentence $\varphi \wedge \psi$ if $\varphi, \psi \in l_{\omega,\omega}$.
(iv) The sentence $(\forall x_c)\varphi(x_c)$ where $\varphi \in l_{\omega,\omega}$ and $\varphi(x_c)$ is obtained by replacing every instance of the constant c by the variable x_c.
(v) Only sentences which can be shown to be sentences by a finite number of applications of the above.

We define the satisfaction relation inductively given a model $\mathfrak{A} = (A, \mathscr{I})$ for a language $\mathscr{L}$. That is, we have:

(i) $\mathfrak{A} \models R(t_1, \ldots, t_n)$ if and only if $(\mathscr{I}(t_1), \ldots, \mathscr{I}(t_n)) \in \mathscr{I}(R)$.
(ii) $\mathfrak{A} \models t_1 \equiv t_2$ if and only if $\mathscr{I}(t_1) = \mathscr{I}(t_2)$.
(iii) $\mathfrak{A} \models \neg\varphi$ if and only if $\mathscr{L}_\varphi \subset \mathscr{L}$ and it is not the case that $\mathfrak{A} \models \varphi$.
(iv) $\mathfrak{A} \models (\forall x_c)\varphi(x_c)$, where $\mathscr{L}_\varphi \subset \mathscr{L}$ and $c \notin \mathscr{L}$, if and only if we have $\mathfrak{A}_y = (A, \mathscr{I}')$ where $\mathscr{I}'$ is obtained from $\mathscr{I}$ by adding an interpretation for c, namely $\mathscr{I}(c) = y$ for $y \in A$, and we require that $\mathfrak{A}_y \models \varphi$ for all $y \in A$.

[2]The relation $\equiv$ should be thought of as a purely formal symbol, just as $\neg$, $\wedge$ and $\forall$ should be thought of as purely formal symbols in the definition of an abstract logic. We generally take the $\equiv$ relation as a given in an abstract logic, just as we take equality as a given in models.

To actually verify that first-order logic is an abstract logic is tedious though straightforward, with the possible exception of the relativization property. The reader should check, however, that the relativization of φ to $R(x, b_1, \ldots, b_n)$ is obtained by replacing each instance of $(\forall x)\psi(x)$ (for any sentence ψ) in φ by the sentence $(\forall x)[R(x, b_1, \ldots, b_n) \to \psi]$.

Having used the connective $\to$, it is worth noting at this point that there are other connectives which are commonly used in first-order logic which we have not discussed explicitly, but are simply combinations of the connectives we have specified, and therefore implicitly determined. We will, for the sake of clarity, define the following:

(i) $\varphi \vee \psi$
(ii) $(\exists x_c)\varphi(x_c)$
(iii) $\varphi \to \psi$

To mean, respectively:

(i) $\neg(\neg\varphi \wedge \neg\psi)$
(ii) $\neg[(\forall x_c)\neg\varphi(x_c)]$
(iii) $\neg(\varphi \wedge \neg\psi)$

Finally, while we did not require a discussion of variables for the quantifier property of abstract logics, it would perhaps be appropriate to do so here. We did not technically define what we mean by a 'variable', though the intuition behind this should be clear. Normally, an infinite set of variables is a part of the abstract logic (and we require only a countably infinite set of variables, because sentences can have at most finitely many constant symbols), but it is more convenient to not make this explicit for notational reasons.

3. Compactness and the Downward Löwenheim-Skolem Theorem

For the rest of this section, it will be assumed that we are working within first-order logic. The purpose of this section will be to equip first-order logic with logical axioms, rules of inference and the concept of a 'deduction'. This will help us define the notion of witnesses for a sentence, with which we can prove two lemmas which have, as a direct consequence, the compactness and the downward Löwenheim-Skolem theorems, as well as the extended completeness theorem.

Definition 3.1. A sentence φ is said to be a *tautology* if and only if for any $\mathfrak{A} = (A, \mathscr{I})$ in a language $\mathscr{L}$ with $\mathscr{L}_\varphi \subset \mathscr{L}$, we have $\mathfrak{A} \models \varphi$.

Intuitively, this means that these sentences are - in some sense - 'structurally' true, because the specific interpretation of the symbols used does not matter. An example of a tautology in first-order logic as we have defined it would be, for instance, the sentence $c \equiv c$ for a constant symbol c, because any model which provides an interpretation for c evidently satisfies the sentence - it is always true that $\mathscr{I}(c) = \mathscr{I}(c)$. Here, it is in some sense the structure of the sentence which makes it true regardless of the specific interpretation of the symbols.

An important reason we single out tautologies as such is precisely because we regard them as being 'inherently true' in some vague sense of the word, and we would therefore like to allow ourselves the ability to use them in making deductions. In particular, tautologies form part of our logical axioms.

Definition 3.2. The following are *logical axioms*:

(i) The sentence φ, where φ is a tautology.
(ii) The sentence $(\forall v)(\varphi \to \psi(v)) \to (\varphi \to (\forall v)\psi(v))$ where φ and ψ are sentences.
(iii) The sentence $(\forall v)\varphi \to \psi$, where ψ is obtained by substituting each instance of v which is not quantified over in φ by the term t (such that all variables x in t are not quantified over in ψ).

Note that the last two are tautologies themselves, but they are used often enough that it is worth mentioning them explicitly.

Definition 3.3. Consider a set of sentences T in a language $\mathscr{L}$. We say that a sentence φ is *deducible* from T, written $T \vdash \varphi$, if and only if there is a finite sequence of sentences $\varphi_1, \ldots, \varphi_n$ where each φ_i is one (or more) of:

(i) A tautology.
(ii) A sentence in T.
(iii) *(Detachment)* A sentence ψ if there are natural numbers $j, k < i$ where φ_j is γ and φ_k is $\gamma \to \psi$, for any sentence γ.
(iv) *(Generalization)* A sentence $(\forall x_c)\psi(x_c)$ if there is a natural number $j < i$ such that φ_j is ψ and $c \in \mathscr{L}_\varphi$, but c does not occur in any sentence in T. Here, $\psi(x_c)$ is obtained by substituting the variable x_c for each instance of c.

The last two are often referred to as *rules of inference*. Such a sequence of sentences is often called a *deduction*.

There is an important notion of consistency which shows up very frequently in model theory:

Definition 3.4. A set of sentences T in a language $\mathscr{L}$ is *inconsistent* if for any sentence φ for a language $\mathscr{L}$, we can deduce φ from T. T is said to be *consistent* otherwise.

Definition 3.5. A set of sentences T in a language $\mathscr{L}$ is said to be *maximal consistent* if it is consistent and the only set of consistent sentences of $\mathscr{L}$ containing T is T itself.

We will now state, without proof, a useful theorem in model theory:

Theorem 3.6. *(Lindenbaum's Theorem) For any consistent set of sentences T, there is a maximal consistent set of sentences T' containing T.*

The idea of the proof uses the fact that there are at most $\|\mathscr{L}\|$ sentences that we can form, and so we can index these by the ordinals less than $\|\mathscr{L}\|$. We start with T and at each successor ordinal, we add the sentence to our current set if the resulting set is consistent, and we take unions at the limit ordinals. A simple transfinite induction shows that this set is consistent, and by construction it is maximal consistent.[3]

Another important concept is that of a set of witnesses.

Definition 3.7. Let T be a set of sentences with the associated language $\mathscr{L}$ and let C be a set of constant symbols of $\mathscr{L}$. We say that C is a *set of witnesses* for

[3]A complete proof can be found in Chang & Keisler, on page 10.

T in $\mathscr{L}$ if and only if for every sentence φ in the language $\mathscr{L}$, there is a constant $c \in C$ such that:

$$T \vdash (\exists x_d)\varphi(x_d) \to \varphi(c)$$

where $\varphi(c)$ is obtained from φ by replacing every instance of d with c.

And this definition helps us prove one of two useful lemmas.

Lemma 3.8. *Let T be a consistent set of sentence of $\mathscr{L}$ and let C be a set of new constant symbols of power $|C| = \|\mathscr{L}\|$. Let $\overline{\mathscr{L}} = \mathscr{L} \cup C$ be a simple expansion of $\mathscr{L}$ formed by adding C. Then T can be extended to a consistent set of sentences $\overline{T}$ in $\overline{\mathscr{L}}$ which has C as a set of witnesses in $\overline{\mathscr{L}}$.*

Proof. Let $\alpha = \|\mathscr{L}\| = |C|$, and index the elements of C as c_β for $\beta < \alpha$, where the c_β are distinct if their subscripts are distinct. The power of $\overline{\mathscr{L}}$ is evidently also α, and so we can index all sentences of $\overline{\mathscr{L}}$ with α.

We define a sequence of sets of sentences of $\overline{\mathscr{L}}$, (T_ξ) for $\xi < \alpha$, recursively:

(i) $T_0 = T$.
(ii) $T_{\xi+1} = T_\xi \cup \{(\exists x_d)\varphi_\xi(x_d) \to \varphi_\xi(c_\xi)\}$ (where $\varphi(c_\xi)$ is constructed as in the definition).
(iii) $T_\xi = \bigcup_{\zeta<\xi} T_\zeta$ when ξ is a limit ordinal different from 0.

We now wish to show that this set of sentences is consistent. We note that T_ξ is evidently consistent at non-zero limit ordinals, and that T_0 is consistent by hypothesis. We will now show that if T_ξ is consistent, then $T_{\xi+1}$ must be consistent. If this were not the case, then:

$$T_\xi \vdash \neg[(\exists x_d)\varphi_\xi(x_d) \to \varphi_\xi(c_\xi)].$$

We can rewrite this as:

$$T_\xi \vdash (\exists x_d)\varphi_\xi(x_d) \wedge \neg\varphi_\xi(c_\xi).$$

And by the rule of generalization, since c_ξ does not occur in T_ξ by construction, we can rewrite this as:

$$T_\xi \vdash (\forall x_d)[(\exists x_d)\varphi_\xi(x_d) \wedge \neg\varphi_\xi(x_d)].$$

We will assume that the following are tautologies, without proof (though this should be intuitively clear and follow from our definitions):

$$(\forall x_d)(\exists x_d)\varphi_\xi(x_d) \wedge \neg\varphi_\xi(x_d) \to (\exists x_d)\varphi_\xi(x_d) \wedge (\forall x_d)\neg\varphi_\xi(x_d)$$

$$(\exists x_d)\varphi_\xi(x_d) \wedge (\forall x_d)\neg\varphi_\xi(x_d) \to (\exists x_d)\varphi_\xi(x_d) \wedge (\exists x_d)\neg\varphi_\xi(x_d).$$

And this allows us to obtain that:

$$T_\xi \vdash (\exists x_d)\varphi_\xi(x_d) \wedge (\exists x_d)\neg\varphi_\xi(x_d).$$

Which shows that T_ξ itself must have been inconsistent, contradicting the hypothesis in our induction. Thus, we have shown that T_ξ is consistent for all $\xi < \alpha$ and we can therefore construct a consistent set of sentences $\overline{T} = \bigcup_{\zeta<\xi} T_\zeta$.

We then see that $\overline{T}$ is a consistent set of sentences in $\overline{\mathscr{L}}$ which, by construction, has C as a set of witnesses since, for any sentence φ with language $\mathscr{L}$, there is a constant $c \in C$ such that $T \vdash (\exists x_d)\varphi(x_d) \to \varphi(c)$. This completes the proof. □

The following lemma is even more tedious and, as such, we provide only a partial proof.

Lemma 3.9. *Let T be a consistent set of sentences and C be a set of witness for T in $\mathscr{L}$. Then there is a model $\mathfrak{A} = (A, \mathscr{I})$ where for any $a \in A$, $a = \mathscr{I}(c)$ for some $c \in C$ and $\mathfrak{A} \vDash T$.*

Proof. If C is a set of witness for T in $\mathscr{L}$, then C is also a set of witnesses for every extension of T (since any sentence that can be deduced from T can also be deduced from any extension of T). Furthermore, if $\mathfrak{A} \vDash T'$ for some extension T' of T, then evidently $\mathfrak{A} \vDash T$. Thus, we can assume that T is a maximal consistent set of sentences.

For any two constants $c, d \in C$, we will define $c \sim d$ if and only if $c \equiv d \in T$. Because T is maximal consistent, we know that $c \equiv c \in T$ and if $c \equiv d$ and $d \equiv e$ are in T, then $c \equiv e$ and $d \equiv c$ are in T. Thus, if $c \sim d$ and $d \sim e$, then $c \sim e$ and $d \sim c$. We see that $\sim$ is an equivalence relation on C as defined, and so for each $c \in C$, we can define $\widetilde{c}$ to be the equivalence class of c in C under this equivalence relation. We construct a model $\mathfrak{A} = (A, \mathscr{I})$ where A is the set of equivalence classes of all $c \in C$.

For the relations on A, we will define R to be an n-placed relation on A for an n-placed relation symbol P such that $R(\widetilde{c_1}, \ldots, \widetilde{c_n})$ if and only if $P(c_1, \ldots, c_n) \in T$. We note that by our axioms of identity, we have:

$$T \vdash P(c_1, \ldots, c_n) \wedge c_1 \equiv d_1 \wedge \ldots \wedge c_n \equiv d_n \to P(d_1, \ldots, d_n)$$

And this establishes that our choice of the representative of the equivalence does not matter, which shows that our interpretation will be well-defined.

For constants, we note that we can consider constant symbols $d, e \in \mathscr{L}$ and form the sentence $(\exists x_e)(d \equiv x_e)$, which must be in T because of . This sentence must have some constant $c \in C$ as a witness, and so $d \equiv c \in T$ (as we assumed T was maximal consistent). The equivalence class of the constant is unique, as our axioms of identity show that:

$$T \vdash (d \equiv c \wedge d \equiv c' \to c \equiv c')$$

We can therefore uniquely interpret any constant d with an equivalence class in C.

Similarly, for functions, we remark that we can perform the same trick to construct the sentence $(\exists x_e)[F(c_1, \ldots, c_m) \equiv x_e] \in T$ for any m-placed function symbol $F \in \mathscr{L}$. Once again, because T has witnesses in C, there must be a constant $c \in C$ such that $F(c_1, \ldots, c_m) \equiv c \in T$. Once again, we can use the axioms of identity to show that while the c's are not necessarily unique, their equivalence classes must be and so we can always define a function G on the set A of equivalence classes.

To complete our proof, it remains to be shown that $\mathfrak{A} \vDash T$. This can be accomplished through a tedious induction on the complexity of sentences in T, but we will omit the proof. □

Theorem 3.10 (Downward Löwenheim-Skolem). *Every consistent theory T in $\mathscr{L}$ has a model of cardinality at most $\|\mathscr{L}\|$.*

Proof. By Lemma 3.8, there is an extension $\overline{T}$ of T and a simple extension $\overline{\mathscr{L}}$ of $\mathscr{L}$ (the latter with $\|\overline{\mathscr{L}}\| = \|\mathscr{L}\|$) such that $\overline{T}$ has witness in $\overline{\mathscr{L}}$. We note that in our constructions, we had $|C| = \|\overline{\mathscr{L}}\|$. Then, by Lemma 3.9, there is a model $\mathfrak{A}$ for $\overline{T}$, and so a model $\mathfrak{A}$ for T, where $|A| \leq |C|$. We can then take the reduct $\mathfrak{B}$ of $\mathfrak{A}$ to the language $\mathscr{L}$, and so we have $|B| = |A| \leq \|\overline{\mathscr{L}}\| = \|\mathscr{L}\|$, completing the proof. □

Theorem 3.11 (Extended Completeness)**.** *Let T be a set of sentence of $\mathscr{L}$. Then T is consistent if and only if T has a model.*

Proof. By Theorem 3.10, if T is consistent then T has a model. Then, if T has a model $\mathfrak{A}$, suppose T were inconsistent. We would then have $T \vdash \varphi$ and $T \vdash \neg\varphi$ for some sentence φ in the same language as T. We would then have $\mathfrak{A} \models \varphi$ and $\mathfrak{A} \models \neg\varphi$, which cannot happen by our definition of the satisfaction relation. □

It turns out extended completeness is equivalent to compactness. This equivalence becomes clear with the following (obvious) proposition.

Proposition 3.12. *A set of sentences is consistent if and only if every finite subset is consistent.*

Proof. Consistency concerns only finite deductions. □

The proof of the compactness theorem is straightforward.

Theorem 3.13 (Compactness)**.** *A set of sentences T has a model if and only if every finite subset of T has a model.*

Proof. By Theorem 3.11, T has a model if and only if T is consistent. However, by Proposition 3.12, T is consistent if and only if every finite subset is consistent. Once again, by Theorem 3.11, every finite subset of T is consistent if and only if every finite subset of T has a model. Hence, T has a model if and only if every finite subset of T has a model. □

4. Lindström's Theorem

Lindström's theorem characterizes first-order logic up to equivalence. To present a proof of the theorem, however, we will need to develop a few definitions and a lemma. Once again, we remind the reader that we are technically ignoring function symbols, though we may occasionally mention functions and function symbols to elucidate the mechanics of some of the proofs (in which case, these should be simply thought of as relations and relation symbols).

Though we have not defined what it means for two abstract logics to be 'equivalent', one might guess that an isomorphism argument could come into play here. Full isomorphisms, however, are too restrictive for our purposes, and so we begin by defining the concept of a partial isomorphism.

Definition 4.1. Let $\mathfrak{A} = (A, \mathscr{I}_A)$ and $\mathfrak{B} = (B, \mathscr{I}_B)$ be models. A *partial isomorphism* between $\mathfrak{A}$ and $\mathfrak{B}$ is an isomorphism between finite submodels $\mathfrak{A}' = (A', \mathscr{I}'_{A'})$ and $\mathfrak{B}' = (B', \mathscr{I}'_{B'})$ of $\mathfrak{A}$ and $\mathfrak{B}$ respectively.

Definition 4.2. Let $\mathfrak{A} = (A, \mathscr{I})$ and $\mathfrak{B} = (B, \mathscr{I}')$ be models for a language $\mathscr{L}$. We say that $\mathfrak{A}$ and $\mathfrak{B}$ are *partially isomorphic* (written $\mathfrak{A} \cong_p \mathfrak{B}$) if there is a nonempty set of partial isomorphisms I with the property that for every partial isomorphism $F \in I$ and $a \in A$, there is a partial isomorphism $G \in I$ such that $G(a)$ is defined and $F \subset G$, and similarly for all $b \in B$. This last property is often called the *back-and-forth* property.

Note that two partially isomorphic models need not necessarily be isomorphic; this can be seen in the case where both models are uncountable, since the union of the domains of the partial isomorphisms can be constructed to be at most countably infinite. However, if we restrict the cardinality of the model to at most $\aleph_0$, then two models are indeed partially isomorphic if and only if they are isomorphic.

Proposition 4.3. *If two at most countable models are partially isomorphic, then they are isomorphic.*

Proof. We will present an outline of the proof. If the set of partial isomorphisms is nonempty, we can find a partial isomorphism which is an isomorphism of the submodels with only one element in each (one can always restrict the domain of any partial isomorphism). Then, however, we could construct an increasing chain of partial isomorphisms using the back-and-forth property. Since the models are countable, for any given relation, function or constant, there is a partial isomorphism which preserves it. Hence, the models must be isomorphic. □

We aim to characterize first-order logic by the downward Löwenheim-Skolem theorem and the compactness theorem, and so we make the following two definitions for abstract logics in general, which allow for more convenient phrasing in the statement and proof of the lemma and the theorem.

Definition 4.4. We define the *Löwenheim number* of an abstract logic $(l, \models_l)$ to be the least cardinal α such that every sentence $\varphi \in l$ which has a model has a model of power at most α.

Definition 4.5. We say that an abstract logic $(l, \models_l)$ is *countably compact* if for every countable set T of sentences of l, T has a model if and only if every finite subset of T has a model.

It is interesting to note the difference between countable compactness and the compactness theorem (which we might call 'full' compactness). While it is true that there is a version of the compactness theorem in first-order logic where the set of sentences need not be countable, it turns out that simply requiring that the compactness theorem holds for countable sets of sentences is sufficient, despite being a strictly weaker condition.

Note that, by Theorems 3.10 and 3.13, first-order logic has Löwenheim number $\aleph_0$ (we sometimes write ω instead) and is countably compact. These two things will characterize first-order logic up to equivalence, but before we specify what we mean when we say that two abstract logics are equivalent, we present the following definition of an equivalence of models.

Definition 4.6. Let $\mathfrak{A}$ and $\mathfrak{B}$ be models, and let $(l, \models_l)$ be an abstract logic. We say that $\mathfrak{A}$ and $\mathfrak{B}$ are *l-elementarily equivalent* if for any sentence $\varphi \in l$, we have $\mathfrak{A} \models_l \varphi$ if and only if $\mathfrak{B} \models_l \varphi$.

At this point, we can prove a useful lemma, which is a somewhat interesting result in its own right. The idea of the proof is that if two models $\mathfrak{A}$ and $\mathfrak{B}$ are partially isomorphic, then there must be a sentence $\psi \in l$ which expresses this (as well as the fact that $\mathfrak{A} \models_l \varphi$ and $\mathfrak{B} \models_l \neg\varphi$). The fact that there must then be a countable model because l has Löwenheim number $\aleph_0$ gives us the contradiction we need.

Lemma 4.7. *Let $(l, \models_l)$ be an abstract logic with Löwenheim number $\aleph_0$. If $\mathfrak{A}$ and $\mathfrak{B}$ are models for a language $\mathscr{L}$ which are partially isomorphic, then $\mathfrak{A}$ and $\mathfrak{B}$ are l-elementarily equivalent.*

Proof. Suppose $\mathfrak{A}$ and $\mathfrak{B}$ are models for $\mathscr{L}$ that were partially isomorphic by some I, but that there were some sentence $\varphi \in l$ such that $\mathfrak{A} \vDash_l \varphi$ but $\mathfrak{B} \vDash_l \neg\varphi$. By the expansion property of abstract logics, we need only consider the reducts of $\mathfrak{A}$ and $\mathfrak{B}$ to $\mathscr{L}_\varphi$, and so we may as well assume $\mathscr{L} = \mathscr{L}_\varphi$ and, therefore, that $\mathscr{L}$ is finite.

We will attempt to construct a model $\mathfrak{C}$ and a sentence $\psi \in l$ such that $\mathfrak{C} \vDash_l \psi$ and ψ implies that that $\mathfrak{A} \cong_p \mathfrak{B}$. To do this, we will let U and W be unary relation symbols, E be a binary relation symbols and p be a binary function symbol. We will use the notation $\langle x, y\rangle$ for $p(x, y)$ and $\langle x_1, \ldots, x_n\rangle$ for $p(\langle x_1, \ldots, x_{n-1}\rangle, x_n)$. We now expand the language $\mathscr{L}$ to a language $\mathscr{L}' = \mathscr{L} \cup \{U, W, E, p\}$ and consider the following sentences of l:

(i) $\varphi \mid U(x)$
(ii) $(\neg\varphi) \mid W(x)$
(iii) $\forall x, y, z, w(\langle x, y\rangle \equiv \langle z, w\rangle \leftrightarrow (x \equiv z \wedge y \equiv w))$
(iv) $\forall x, y, u[E(x, y) \wedge U(u) \to \exists w(W(w) \wedge E(\langle x, u\rangle, \langle y, w\rangle))]$
(v) For each n-placed relation symbol R in $\mathscr{L}$ (where n is allowed to vary), the sentence:
$(\forall x_1, \ldots, x_n, y_1, \ldots, y_n)[E(\langle x_1, \ldots, x_n\rangle, \langle y_1, \ldots, y_n\rangle)$
$\to (U(x_1) \leftrightarrow W(y_1)) \wedge \ldots \wedge (U(x_n) \leftrightarrow W(y_n))$
$\wedge(R(x_1, \ldots, x_n) \leftrightarrow R(y_1, \ldots, y_n))]$
(vi) For each n-placed function symbol F in $\mathscr{L}$ (where n is allowed to vary), the sentence:
$(\forall x_1, \ldots, x_{n+1}, y_1, \ldots, y_{n+1})[E(\langle x_1, \ldots, x_{n+1}\rangle, \langle y_1, \ldots, y_{n+1}\rangle)$
$\to (U(x_1) \leftrightarrow W(y_1)) \wedge \ldots \wedge (U(x_{n+1}) \leftrightarrow W(y_{n+1}))$
$\wedge(F(x_1, \ldots, x_n) \equiv x_{n+1} \leftrightarrow F(y_1, \ldots, y_n) \equiv y_{n+1})]$

Since there are only finitely many relation symbols and function symbols, we can form a sentence ψ which is the conjunction of all of the above. We note that ψ is a sentence in the language $\mathscr{L}'$, and is therefore also a sentence in the language $\mathscr{L}''$, where $\mathscr{L}''$ is obtained by adding the elements of $\mathscr{L}$ indexed with $\mathfrak{B}$ to $\mathscr{L}'$. We form a model $\mathfrak{C} = (C, \mathscr{I})$ of $\mathscr{L}''$ whose universe is the disjoint union of the sets A, B, A' and B', where we let A' and B' be the sets of finite sequences of A and B respectively. For the sake of convenience, we will index the elements of C by $\mathfrak{A}$ if they were elements of A or A', and by $\mathfrak{B}$ otherwise.

We assign to symbols which are in both $\mathscr{L}''$ and $\mathscr{L}$ their interpretation under $\mathscr{I}_A$ if they are indexed by $\mathfrak{A}$, and their interpretation under $\mathscr{I}_B$ otherwise. We then define $\mathscr{I}(U)$ to be the relation X where $X(a)$ if and only if $a \in A$ and, similarly, $\mathscr{I}(W)$ to be the relation Y where $Y(b)$ if and only if $b \in B$. We further define $\mathscr{I}(E)$ to be the relation Z where $Z(c, d)$ if and only if $c \in A \cup A'$, $d \in B \cup B'$ and there is an isomorphism between the submodels of $\mathfrak{A}$ and $\mathfrak{B}$ with universes restricted to the elements of A in c and B in d respectively. Finally, we define $\mathscr{I}(p)$ to be the function P where $P(x, y)$ is defined to be the sequence $\langle x, y\rangle$, $P(\langle x_1, \ldots, x_n\rangle, y)$ is defined to be the sequence $\langle x_1, \ldots, x_n, y\rangle$ and $P(x, \langle y_1, \ldots, y_n\rangle)$ is defined to be the sequence $\langle x, y_1, \ldots, y_n\rangle$. This completes our construction of $\mathfrak{C}$.

We note that because $\mathfrak{A} \cong_p \mathfrak{B}$, $\mathfrak{A} \vDash_l \varphi$ and $\mathfrak{B} \vDash_l \neg\varphi$, we have $\mathfrak{C} \vDash_l \psi$ by construction. We see, however, that because l has Löwenheim number $\aleph_0$, there must be a model $\mathfrak{C}_0$ of power at most $\aleph_0$ such that $\mathfrak{C}_0 \vDash_l \psi$, from which we can obtain models $\mathfrak{A}_0$ and $\mathfrak{B}_0$ which are at most countable (using the interpretations of the relations U and W). We note that we must then have both $\mathfrak{A}_0 \vDash_l \varphi$ and $\mathfrak{B}_0 \vDash_l \neg\varphi$, but also $\mathfrak{A}_0 \cong_p \mathfrak{B}_0$. This, however, is a contradiction by Proposition 4.3

because both $\mathfrak{A}_0$ and $\mathfrak{B}_0$ are now at most countable, and so $\mathfrak{A}$ and $\mathfrak{B}$ must have been l-elementarily equivalent. □

Finally, we require a definition of an equivalence of abstract logics.

Definition 4.8. Let $(l, \vDash_l)$ and $(l', \vDash_{l'})$ be abstract logics. We say that $(l, \vDash_l)$ and $(l', \vDash l')$ are *equivalent* if for every $\varphi \in l$, there is a $\varphi' \in l'$ such that $\mathscr{L}_\varphi = \mathscr{L}'_{\varphi'}$ and for any model $\mathfrak{A}$ for $\mathscr{L}_\varphi$, $\mathfrak{A} \vDash_l \varphi$ if and only if $\mathfrak{A} \vDash_l \varphi'$, and vice-versa.

In essence, we consider two abstract logics 'equivalent' if there is no way to distinguish between them from the point of view of a model. We are finally in a position to prove Lindström's theorem. This central idea of the proof lies in the first three and final paragraphs, but we also need to construct a sentence and a model, as we did in the previous lemma, which requires the better part of a page.

Theorem 4.9 (Lindström)**.** *Let $(l, \vDash_l)$ be a countably compact abstract logic with Löwenheim number $\aleph_0$. Then $(l, \vDash_l)$ is equivalent to first-order logic, $(l_{\omega,\omega}, \vDash)$.*

Proof. We wish to show that for every sentence $\varphi \in l$, there is some sentence $\psi \in l_{\omega,\omega}$ such that for any model $\mathfrak{A}$ for a language $\mathscr{L}$ containing $\mathscr{L}_\varphi$, we have $\mathfrak{A} \vDash_l \varphi$ if and only if $\mathfrak{A} \vDash \psi$. Since the converse follows from the closure property of abstract logics, this will complete the proof.

By the expansion and occurrence properties, we need only consider the case where $\mathscr{L}$ is finite, since the satisfaction relation depends only on the reduct to $\mathscr{L}_\varphi$, which itself must be finite. We define a sequence of relations I_k between finite n-tuples of $\mathfrak{A}$ and $\mathfrak{B}$ and say that $(\langle a_1, \dots, a_n \rangle, \langle b_1, \dots, b_n \rangle) \in I_0$ if the submodels of $\mathfrak{A}$ and $\mathfrak{B}$ obtained by restricting the models to the sets $\{a_1, \dots, a_n\}$ and $\{b_1, \dots, b_n\}$ respectively are $l_{\omega,\omega}$-elementarily equivalent. We further say that $(\langle a_1, \dots, a_n \rangle, \langle b_1, \dots, b_n \rangle) \in I_{m+1}$ if and only if for all $c \in A$, there is a $b \in B$ such that $(\langle a_1, \dots, a_n, c \rangle, \langle b_1, \dots, b_n, d \rangle) \in I_m$. Because $\mathscr{L}$ is finite, and therefore there are only finitely many atomic sentences of $\mathscr{L}$, and we are assuming it contains no function symbols, we see that for each k there is a finite set Γ_k of sentences of first-order logic in the language $\mathscr{L}$ such that $\emptyset I_k \emptyset$ if and only if $\mathfrak{A}$ and $\mathfrak{B}$ satisfy the same sentences of Γ_k, where $\emptyset$ is the empty sequence.

We now consider any $\varphi \in l$ with $\mathscr{L}_\varphi \subset \mathscr{L}$. We define $\mathfrak{A} \equiv_k \mathfrak{B}$ to mean $\emptyset I_k \emptyset$ as above. We suppose that there is no k such that for any two models $\mathfrak{A}$ and $\mathfrak{B}$, if $\mathfrak{A} \equiv_k \mathfrak{B}$ and $\mathfrak{A} \vDash_l \varphi$ then $\mathfrak{B} \vDash_l \varphi$. For each k, we can then find models $\mathfrak{A}_k$ and $\mathfrak{B}_k$ such that $\mathfrak{A}_k \equiv_k \mathfrak{B}_k$ and $\mathfrak{A}_k \vDash_l \varphi$, but $\mathfrak{B}_k \vDash_l \neg\varphi$. We will show that we must then be able to construct models $\mathfrak{A}'_H$ and $\mathfrak{B}'_H$ which are partially isomorphic but for which $\mathfrak{A}'_H \vDash_l \varphi$ and $\mathfrak{B}'_H \vDash_l \neg\varphi$.

By taking a subsequence, we can assume that all the $\mathfrak{A}_k$ satisfy the same atomic sentences of $\mathscr{L}$. Furthermore, by the isomorphism property of abstract logics, we assume that each $\mathfrak{A}_k$ has the same interpretation of the contsants of $\mathscr{L}$. The model $\mathfrak{A}$ formed by the union of all of the models $\mathfrak{A}_k$ then has each model $\mathfrak{A}_k$ as a submodel. We will also, abusing the isomorphism property, take each $\mathfrak{A}_k$ so that their universes are all disjoint from ω, the first infinite ordinal. We use the same construction for $\mathfrak{B}$.

We use the notation $\mathbf{a}$ or $\mathbf{b}$ to mean a finite sequence of elements of A, and $\mathbf{ab}$ to mean their concatenation (by attaching the sequence $\mathbf{b}$ to the end of the sequence $\mathbf{a}$, with a similar construction for attaching elements of A). We will let A' be the set of finite sequences of A and define functions F and F' such that $F(\mathbf{a}, b) = \mathbf{a}b$

and $F(\mathbf{a}, \mathbf{b}) = \mathbf{ab}$. We then construct $\mathfrak{A}''$ which is an extension of $\mathfrak{A}$ with universe $A \cup A'$ and a relation U (where $U(a)$ if $a \in A'$), as well as functions F and F'. We construct $\mathfrak{B}''$ analogously.

We form a model $\mathfrak{C}$ for $\mathscr{L}$ with universe C which is the disjoint union of A, A', B and B'. We add relations R and S such that, for each $k \in \omega$, $A_k = \{a \in A \mid R(a, k)\}$, $A'_k = \{a \in A' \mid R(a, k)\}$, $B_k = \{b \in B \mid S(b, k)\}$ and $B'_k = \{b \in B \mid S(b, k)\}$. We define a relation I such that for $k \in \omega$, we have $I(k, \mathbf{a}, \mathbf{b})$ if and only if $\mathbf{a}\ I_k\ \mathbf{b}$, and the usual order relation $\leq$ on ω. We also retain the functions defined in the previous paragraph. We then consider an expansion $\mathscr{L}'$ of $\mathscr{L}$ which is obtained by adding relation symbols for each of the functions and relations defined above (we look at the binary functions as ternary relations), and we extend the interpretation function of $\mathfrak{C}$ to $\mathscr{L}'$ by adding the obvious interpretations for the added symbols.

We construct a sentence $\psi \in l$ in this expanded language which expresses that ω under $\leq$ is a well-ordering with well-defined successors and predecessors except for the first element, and that for all $k \in \omega$, $\mathfrak{A}_k \equiv_k \mathfrak{B}_k$, $\mathfrak{A}_k \vDash_l \varphi$ but $\mathfrak{B}_k \vDash_l \neg\varphi$ (this sentence is constructed in exactly the same way as the sentence in the preceding lemma was constructed, except we use the quantifier property an additional time to quantify over k, and we add sentences expressing that ω under $\leq$ is a well-ordering with well-defined successors and predecessors except for the first element).

By our construction, this sentence $\psi \in l$ holds in $\mathfrak{C}$. Furthermore, we can consider a countable set of sentences T which contains the sentences:

(i) $(\exists x)(1 \leq x)$
(ii) $(\exists x)(1 \leq x \wedge 2 \leq x)$
(iii) $(\exists x)(1 \leq x \wedge 2 \leq x \wedge 3 \leq x)$
$\vdots$

as well as the sentence ψ (we use $\leq, 1, 2, 3, \ldots$ loosely). We see that $\mathfrak{C}$ is a model for any finite subset of T and that therefore there must be a model $\mathfrak{C}''$ for T. This model then contains a nonstandard element H (nonstandard in the simple order). Then, we have $\mathfrak{A}_H \vDash_l \varphi$, $\mathfrak{B}_H \vDash_l \neg\varphi$ and $\mathfrak{A}_H \equiv_H \mathfrak{B}_H$, where $\mathfrak{A}_H$ and $\mathfrak{B}_H$ have universes A_H and B_H defined analogously to A_k and B_k in the construction of $\mathfrak{C}$ above. Then, we can define a relation J between m-tuples $\mathbf{a}$ in the universe of $\mathfrak{A}_H$ and $\mathbf{b}$ in the universe of $\mathfrak{B}_H$ where $\mathbf{a}\ J\ \mathbf{b}$ if and only if $\mathbf{a}\ I'_{H-m}\ \mathbf{b}$.

This relation is a partial isomorphism between $\mathfrak{A}'_H$ and $\mathfrak{B}'_H$ because $\mathfrak{A}'_H \equiv_H \mathfrak{B}'_H$ and so, by our initial construction, we must have a suitable sequence of isomorphisms such that $\emptyset I_H \emptyset$. Then, however, since $\mathfrak{A}'_H$ and $\mathfrak{B}'_H$ are partially isomorphic and $(l, \vDash_l)$ has Löwenheim number $\aleph_0$, $\mathfrak{A}'_H$ and $\mathfrak{B}'_H$ must in fact be l-elementarily equivalent, which is a contradiction because φ is true in $\mathfrak{A}'_H$ and not in $\mathfrak{B}'_H$. Hence, for any φ, there is a $k \in \omega$ such that if $\mathfrak{A} \equiv_k \mathfrak{B}$ and $\mathfrak{A} \vDash_l \varphi$, then $\mathfrak{B} \vDash_l \varphi$.

However, we remark that for each k, we had a finite set of first-order sentences Γ_k such that for all models $\mathfrak{A}$ and $\mathfrak{B}$, $\mathfrak{A} \equiv_k \mathfrak{B}$ if and only if $\mathfrak{A}$ and $\mathfrak{B}$ satisfied the same sentences of Γ_k. Hence, for every $\varphi \in l$ there is a $k \in \omega$ such that, for all models $\mathfrak{A}$ and $\mathfrak{B}$, if $\mathfrak{A}$ and $\mathfrak{B}$ satisfy the same sentences of Γ_k and $\mathfrak{A} \vDash_l \varphi$, then $\mathfrak{B} \vDash_l \varphi$. We can restate this as: given any sentence $\varphi \in l$ and model $\mathfrak{A}$ such that $\mathfrak{A} \vDash_l \varphi$, there is a finite set of first-order sentences in the same language as φ such that any model $\mathfrak{B}$ which satisfies the same sentences as $\mathfrak{A}$ also satisfies φ. This shows that φ is equivalent to some combination of first-order sentences in Γ_k, completing the proof. $\square$

Acknowledgements. I would like to thank Professor Malliaris for both suggesting the topic and providing some much-needed guidance on and understand of the subject. I would also like to thank Michael Smith, for his help in editing the paper (which would no doubt be unreadable otherwise), as well as my mentors Max Engelstein and Daphne Kao, for their help in editing the paper, helping clarify my thoughts and voluntarily subjecting themselves to a summer of model theory.

The proofs are drawn largely from Chang & Keisler, with some help from Barwise (which provided some much-needed clarification on more than one occasion).

References

[1] C. C. Chang and H. J. Keisler. Model Theory. Elsevier. 1990.
[2] J. Barwise. Axioms for abstract model theory. Annals of Mathematical Logic, v. 7 (1974).