Henkin's Method and the Completeness Theorem

Guram Bezhanishvili*

1 Introduction

Let $\mathcal{L}$ be a first-order logic. For a sentence φ of $\mathcal{L}$, we will use the standard notation "$\vdash \varphi$" for φ is *provable* in $\mathcal{L}$ (that is, φ is derivable from the axioms of $\mathcal{L}$ by the use of the inference rules of $\mathcal{L}$); and "$\models \varphi$" for φ is *valid* (that is, φ is satisfied in every interpretation of $\mathcal{L}$). The *soundness theorem* for $\mathcal{L}$ states that if $\vdash \varphi$, then $\models \varphi$; and the *completeness theorem* for $\mathcal{L}$ states that if $\models \varphi$, then $\vdash \varphi$. Put together, the soundness and completeness theorems yield the *correctness theorem* for $\mathcal{L}$: a sentence is derivable in $\mathcal{L}$ iff it is valid. Thus, they establish a crucial feature of $\mathcal{L}$; namely, that syntax and semantics of $\mathcal{L}$ go hand-in-hand: every theorem of $\mathcal{L}$ is a logical law (can not be refuted in any interpretation of $\mathcal{L}$), and every logical law can actually be derived in $\mathcal{L}$.

In fact, a stronger version of this result is also true. For each first-order theory $\mathcal{T}$ and a sentence φ (in the language of $\mathcal{T}$), we have that $\mathcal{T} \vdash \varphi$ iff $\mathcal{T} \models \varphi$. Thus, each first-order theory $\mathcal{T}$ (pick your favorite one!) is sound and complete in the sense that everything that we can derive from $\mathcal{T}$ is true in all models of $\mathcal{T}$, and everything that is true in all models of $\mathcal{T}$ is in fact derivable from $\mathcal{T}$. This is a very strong result indeed. One possible reading of it is that the first-order formalization of a given mathematical theory is *adequate* in the sense that every true statement about $\mathcal{T}$ that can be formalized in the first-order language of $\mathcal{T}$ is derivable from the axioms of $\mathcal{T}$.

It is relatively easy to prove the soundness theorem. All it takes is to verify that all axioms of $\mathcal{L}$ (or, more generally, of $\mathcal{T}$) are valid, and that the inference rules preserve validity. This can be done by a routine use of mathematical induction. On the other hand, it is much more challenging to prove the completeness theorem. It requires a construction of a counter-model for each non-theorem φ of $\mathcal{L}$. More generally, the strong completeness theorem requires, for each non-theorem φ of a first-order theory $\mathcal{T}$, a construction of a model of $\mathcal{T}$ which is a counter-model of φ. This is by no means an obvious task.

The importance of the completeness theorem was first realized by David Hilbert (1862–1943), who posed it as an open problem in 1928 in the influential book [10], which he coauthored with Wilhelm Ackermann (1896–1962). The first proof of the completeness theorem was given by Kurt Gödel (1906–1978) in his dissertation thesis the following year.

*Mathematical Sciences; Dept. 3MB, Box 30001; New Mexico State University; Las Cruces, NM 88003; gbezhani@nmsu.edu.

Gödel's proof appeared in print in 1930 in [3]. An English translation of it can be found in the Collected Works of Kurt Gödel [5].

Gödel's proof was rather complicated and was understood by only a few mathematicians. It was not until Leon Henkin's (1921-2006) ingenious construction in the late 1940s, which became known as the *Henkin method*, that the completeness theorem became widely understood. Henkin's method became an instant classic, and the standard way to prove completeness, taught in almost every course on mathematical logic worldwide. An entertaining account of how he discovered his method is described in Henkin's "The discovery of my completeness proofs" [7]. In this project we will study Henkin's method from the original historical source by Henkin [6].

2 Leon Henkin

Leon Henkin was born on 19 April 1921 in Brooklyn, into a Russian Jewish immigrant family. In 1937 he enrolled at Columbia University, where he majored in Mathematics and Philosophy. His first Logic teacher was Ernest Nagel (1901–1985), a distinguished philosopher of Science and one of the major figures of the logical positivist movement. Nagel is also widely known for his popular account [11] of Gödel's celebrated incompleteness theorems, which he coauthored with James R. Newman (1907-1966).

During his studies at Columbia, Leon studied the works of Bertrand Russell (1872–1970) and Alfred North Whitehead (1861–1947), Hilbert and Ackermann, and Willard Quine (1908–2000). In 1939 he also attended a lecture by Alfred Tarski (1901–1983), who later hired him at the University of California, Berkeley. Henkin was also greatly influenced by the book "Projective Geometry" [12, 13] by Oswald Veblen (1880–1960) and John W. Young (1879–1932). But, as he points out himself [7, p. 131], his most important learning experience was studying Gödel's consistency proof of the Continuum Hypothesis [4].

Henkin graduated from Columbia University in 1941 with a degree in Mathematics and Philosophy. The same year he was admitted in the graduate program of Princeton University. His advisor was Alonzo Church (1903–1995). But only a year later the US entered the Second World War. Because of this, Leon passed his qualifying exams in a rush, received an M.A. degree, and left Princeton to join the military. He spent the next four years working in the Signal Corps Radar Laboratory, Belmar, New Jersey. As participant in the Manhattan project, Henkin worked on isotope diffusion. Most of his work involved numerical analysis to obtain solutions of certain partial differential equations. During this period Henkin neither read nor thought about logic.

Leon returned to Princeton in 1946 to complete his Ph.D. degree. He was awarded a predoctoral fellowship by the National Research Council, and continued his work with Church. It was at Princeton, one year later, that he discovered his method, which became the basis of his Ph.D. thesis, and allowed him to obtain a new and elegant proof of the completeness theorem, as well as many other useful corollaries, including completeness of higher order logics with respect to what later became known as *Henkin models*. A detailed account of his discoveries can be found in [7].

Henkin received his Ph.D. in 1947 from Princeton University. In 1953 he was hired by Tarski at the U.C. Berkeley, where he became a close collaborator and an ally of Tarski in promoting logic. Henkin stayed at Berkeley for the rest of his life, where he became a highly regarded professor and social activist. During his long and fruitful career Henkin was regarded as a fine logician and an extraordinary human being. With Tarski and J. Donald Monk of the University of Colorado, Boulder, he coauthored an influential book "Cylindric Algebras" [8, 9], the second volume of which was completed after Tarski's death. Henkin was a recipient of many awards, among which were Chauvenet Prize (1964), Lester R. Ford Award (1971), Gung and Hu Award (1990), Berkeley Citation (1991), and Leon Henkin Citation (2000). But he will be most remembered for the discovery of his method, which we now turn to.

3 First-Order Logic

In this section we will develop all the background from Henkin's source necessary for his ingenious method. In particular, we will learn about the first-order language Henkin works with, its syntax and semantics. We will also study the axiomatic system Henkin introduces, learn how to prove theorems in the system, and describe what it means for the system to be complete.

3.1 Syntax

We start by getting acquainted with the notation of Henkin's original source. Henkin mainly uses the notation from the classic monograph by Church [1], which was considered as standard by the time Henkin's paper appeared in print. At the time, the first-order logic was called *the first-order functional calculus*. Henkin starts out by fixing the following notation:

The system with which we shall deal here will contain as primitive symbols

$$(\quad) \quad \supset \quad f \quad ,$$

and certain sets of symbols as follows:

(i) *propositional symbols* (some of which may be classed as *variables*, others as *constants*), and among which the symbol "f" above is to be included as a constant;

(ii) for each number $n = 1, 2, \cdots$ a set of *functional symbols of degree* n (which again may be separated into *variables* and *constants*); and

(iii) *individual symbols* among which *variables* must be distinguished from *constants*. The set of variables must be infinite.

The alphabet that Henkin uses is slightly different from what became standard in later years. He uses "f" to denote the logical constant "False" and "$\supset$" to denote "implication." He also has left and right parentheses in the alphabet. In addition, by *propositional symbols* Henkin means a denumerable set of propositional letters, which, as we will see shortly, will be interpreted as either true or false statements. This was standard practice at the time, mostly because of propositional logic. But more recent authors, such as Ebbinghaus, Flum, and Thomas [2], refrain from including propositional symbols in the alphabet. Moreover, by *functional symbols of degree n* Henkin means relation symbols of arity n, for each natural number n. Note that Henkin separates relation symbols into constants and variables. Most likely, this is because he always had in mind the treatment of higher-order quantification. More recent authors, however, refrain from the distinction (see, e.g., [2]). Finally, by *individual symbols* Henkin means constants and infinitely many variables.

It is worth pointing out that Henkin does not have the equality symbol in the alphabet. But we will see that this issue is addressed later in the paper.

Exercise 1 Note that Henkin does *not* use any function symbols, which became customary in later developments. The reason, of course, is that we can always think of a function as a special relation, and so having only relation symbols in the alphabet is sufficient. As a result, Henkin does not have a need to define *terms*. Give a definition of a term in Henkin's alphabet. Based on your definition, give an explanation of why there was no need for Henkin to define terms.

Henkin's next task is to construct formulas in the alphabet he has described, which he does by induction.

> *Elementary well-formed formulas* are the propositional symbols and all formulas of the form $G(x_1, \ldots, x_n)$ where G is a functional symbol of degree n and each x_i is an individual symbol.
>
> *Well-formed formulas* (wffs) consist of the elementary well-formed formulas together with all formulas built up from them by repeated application of the following methods:
>
> (i) If A and B are wffs so is $(A \supset B)$;
>
> (ii) If A is a wff and x an individual variable then $(x)A$ is a wff. Method (ii) for forming wffs is called *quantification with respect to the variable x*. Any occurrence of the variable x in the formula $(x)A$ is called *bound*. Any occurrence of a symbol which is not a bound occurrence of an individual variable according to this rule is called *free*.

Note that what Henkin calls *elementary well-formed formulas* are often called *atomic formulas*; and what Henkin calls *well-formed formulas* are often called simply *formulas*. Also, Henkin uses the symbol $(x)A$ for the *universal quantification*. Of course, it has since become customary to use $(\forall x)A$ instead of $(x)A$. But observe that the usage of $\forall$ requires to treat it as part of the alphabet. Therefore, Henkin's approach to the universal quantification is simpler because it does not require an extra symbol in the alphabet.

Exercise 2 Use induction to give your own definition of a formula in Henkin's alphabet.

3.2 Semantics

Having built a first-order language, Henkin's next task is to define an interpretation of the language.

> In addition to formal manipulation of the formulas of this system we shall be concerned with their *meaning* according to the following interpretation. The propositional constants are to denote one of the truth values, T or F, the symbol "f" denoting F, and the propositional variables are to have the set of these truth values as their range. Let an arbitrary set, I, be specified as a domain of individuals, and let each individual constant denote a particular element of this domain while the individual variables have I as their range. The functional constants (variables) of degree n are to denote (range over) subsets of the set of all ordered n-tuples of I. $G(x_1, \cdots, x_n)$ is to have the value T or F according as the n-tuple $\langle x_1, \cdots, x_n \rangle$ of individuals is or is not in the set G; $(A \supset B)$ is to have the value F if A is T and B is F, otherwise T; and $(x)A$ is to have the value T just in case A has the value T for every element x in I.

Exercise 3 Give your own definition of interpretation. Try to be as formal as possible.

> If A is a wff, I a domain, and if there is some assignment of denotations to the constants of A and of values of the appropriate kind to the variables with free occurrences in A, such that for this assignment A takes on the value T according to the above interpretation, we say A is *satisfiable with respect to* I. If *every* such assignment yields the value T for A we say that A *is valid with respect to* I. A is *valid* if it is valid with respect to every domain.

Exercise 4 Today we write $I \models A$ whenever A is satisfiable with respect to I, and $\models A$ whenever A is valid. Explain in your own words what it means for A to be "satisfiable with respect to I," to be "valid with respect to I," and to be valid.

Having given the definition of valid formulas, Henkin's next task is to give an axiomatization of all valid formulas. But before he embarks on this task, he describes several abbreviations he will use in order to simplify notation.

∞∞∞∞∞∞

If A is any wff and x any individual variable we write

$$\sim A \text{ for } (A \supset f)$$

,

$$(\exists x)A \text{ for } \sim (x) \sim A.$$

From the rules of interpretation it is seen that $\sim A$ has the value T or F according as A has the value F or T, while $(\exists x)A$ denotes T just in case there is some individual x in I for which A has the value T.

Furthermore we may omit outermost parentheses, replace a left parenthesis by a dot omitting its mate at the same time if its mate comes at the end of the formula (except possibly for other right parentheses), and put a sequence of wffs separated by occurrence of "$\supset$" when association to the left is intended. For example,

$$A \supset B \supset . C \supset D \supset E \text{ for } ((A \supset B) \supset ((C \supset D) \supset E)),$$

where A, B, C, D, E may be wffs or abbreviations of wffs.

∞∞∞∞∞∞

Note that by $\sim A$ Henkin denotes the *negation* of A. Therefore, negation and existential quantification are merely abbreviations for Henkin. Today we often use $\neg A$ instead of $\sim A$.

Exercise 5 Give a formal definition of $I \models \sim A$ and $I \models (\exists x)A$.

3.3 Axiomatization

Having prepared all the necessary background, Henkin goes on to describe the axiomatic system, which later he will show to be capable of proving all valid formulas.

∞∞∞∞∞∞

If A, B, C are any wffs, the following are called *axioms*:

1. $C \supset . B \supset C$
2. $A \supset B \supset . A \supset (B \supset C) \supset . A \supset C$
3. $A \supset f \supset f \supset A$
4. $(x)(A \supset B) \supset . A \supset (x)B$, where x is any individual variable with no free occurrence in A.
5. $(x)A \supset B$, where x is any individual variable, y any individual symbol, and B is obtained by substituting y for each free occurrence of x in A, provided that no free occurrence of x in A is in a well-formed part of A of the form $(y)C$.

There are two formal rules of inference:

I (*Modus Ponens*). To infer B from any pair of formulas $A, A \supset B$.

II (*Generalization*). To infer $(x)A$ from A, where x is any individual variable.

Exercise 6 Rewrite the axioms in the dot-free notation and explain their meaning. In particular, explain why in axiom (4) Henkin requires that "x is any individual variable with no free occurrence in A." Also explain why in axiom (5) he requires that "no free occurrence of x in A is in a well-formed part of A of the form $(y)C$." In addition, explain the meaning of the two inference rules.

Exercise 7 How many axioms does Henkin have? Justify your answer.

Exercise 8 Verify that each axiom is valid.

Exercise 9 Verify that the inference rules preserve validity. That is, for Modus Ponens, verify that if A and $A \supset B$ are valid, then so is B; and for Generalization, verify that if A is valid, then so is $(x)A$.

3.4 Formal concept of proof

Henkin's next task is to define the concept of *derivation* or *proof*.

A finite sequence of wffs is called a *formal proof from assumptions* Γ, where Γ is a set of wffs, if every formula of the sequence is either an axiom, an element of Γ, or else arises from one or two previous formulas of the sequence by *modus ponens* or generalization, except that no variable with a free occurrence in some formula of Γ may be generalized upon. If A is the last formula of such a sequence we write $\Gamma \vdash A$. Instead of $\{\Gamma, A\} \vdash B$ ($\{\Gamma, A\}$ denoting the set formed from Γ by adjoining the wff A), we shall write $\Gamma, A \vdash B$. If Γ is the empty set we call the sequence simply a *formal proof* and write $\vdash A$. In this case A is called a *formal theorem*.

Exercise 10 Explain why in the definition of formal proof Henkin requires that "no variable with a free occurrence in some formula of Γ may be generalized upon."

Exercise 11 Given a set of formulas Γ and a formula A, give a formal definition of $\Gamma \vdash A$. Also, give a formal definition of $\vdash A$.

Exercise 12 Use your answers to Exercises 8–11 to show that if $\vdash A$, then $\models A$.

Exercise 13 Use your answers to Exercises 8–12 to show that for a set of formulas Γ and a formula A, if $\Gamma \vdash A$, then $\Gamma \models A$.

What you have just verified is called the *Soundness Theorem*. At this stage it is useful to reread the introduction to have a better perspective on what we have achieved already and what is still ahead of us, especially with respect to the *Completeness Theorem*.

3.5 Formal Proofs. The Deduction Theorem

After giving the definition of a theorem, Henkin states the main goal of the paper.

> Our object is to show that every valid formula is a formal theorem, and hence that our system of axioms and rules is *complete*.

To achieve this task, Henkin needs more information about how formal proofs are derived in the axiomatic system he has just presented. But this has already been known by the time of the publication of Henkin's treatise. Indeed, he only states several facts that he will need to achieve completeness and refers the interested reader to Church's monograph [1].

> The following theorems about the first-order functional calculus are all either well-known and contained in standard works, or else very simply derivable from such results. We shall use them without proof here, referring the reader to Church [1] for a fuller account.
>
> III (*The Deduction Theorem*). If $\Gamma, A \vdash B$ then $\Gamma \vdash A \supset B$ (for any wffs A, B and any set Γ of wffs).
>
> 6. $\vdash B \supset f \supset . B \supset C$
> 7. $\vdash B \supset . C \supset f \supset . B \supset C \supset f$
> 8. $\vdash (x)(A \supset f) \supset . (\exists x)A \supset f$
> 9. $\vdash (x)B \supset f \supset . (\exists x)(B \supset f)$.
>
> IV. If Γ is a set of wffs no one of which contains a free occurrence of the individual symbol u, if A is a wff and B is obtained from it by replacing each free occurrence of u by the individual symbol x (none of these occurrences of x being bound in B), then if $\Gamma \vdash A$, also $\Gamma \vdash B$.

Exercise 14 Prove the Deduction Theorem. (Hint: since $\Gamma, A \vdash B$, there exists a formal proof of B from $\Gamma \cup \{A\}$; let $B_1, \ldots, B_n$ be the formal proof, where $B_n = B$; the idea is to turn the sequence $A \supset B_1, \ldots, A \supset B_n$ into a formal proof of $A \supset B$ from Γ; for this you need to show that $\Gamma \vdash A \supset B_i$ for each $i \leq n$; you will need to consider four cases: (i) B_i is an axiom or a member of Γ, (ii) $B_i = A$ (for this case you will need to show that $A \supset A$ is a theorem of Henkin's axiomatic system), (iii) B_i is obtained from B_j and B_k by Modus Ponens, where $j, k < i$, and (iv) B_i is obtained from B_j by Generalization, where $j < k$.)

Exercise 15 Rewrite formulas (6) and (7) in the dot-free notation. Prove that formulas (6) and (7) are theorems of Henkin's axiomatic system. (Hint: use the Deduction Theorem; more specifically, to show, e.g. (6), first show that $B \supset f, B \vdash C$, and then apply the Deduction Theorem twice.)

Exercise 16 Rewrite formulas (8) and (9) in the dot-free and abbreviation-free notation. Prove that formulas (8) and (9) are theorems of Henkin's axiomatic system. (Hint: use the Deduction Theorem.)

Exercise 17 Prove (IV).

4 Henkin's Method and the Completeness Theorem

Now that we have developed all the necessary background, we are ready to describe Henkin's ingenious method, which is the main topic of this section.

Let $\mathbf{S}_0$ be a particular system determined by some definite choice of primitive symbols. A set Λ of wffs of $\mathbf{S}_0$ will be called *inconsistent* if $\Lambda \vdash f$, otherwise *consistent*. A set Λ of wffs of $\mathbf{S}_0$ will be said to be *simultaneously satisfiable* in some domain I of individuals if there is some assignment of denotations (values) of the appropriate type of the constants (variables) with free occurrences in formulas of Λ, for which each of these formulas has the value T under the interpretation previously described.

Exercise 18 Using modern notation, describe what it means for a set of formulas to be consistent, inconsistent, and simultaneously satisfiable.

Given a consistent set of sentences Λ, Henkin's main task is to construct a domain which simultaneously satisfies Λ. More precisely:

THEOREM. *If Λ is a set of formulas of $\mathbf{S}_0$ in which no member has any occurrence of a free individual variable, and if Λ is consistent, then Λ is simultaneously satisfiable in a domain of individuals having the same cardinal number as the set of primitive symbols of $\mathbf{S}_0$.*

This is a difficult task, because it is unclear where the domain I should come from. This Henkin achieves by an ingenious construction he discovered, which became known as the *Henkin method*. Note that the theorem is not restricted to the case where the set of primitive symbols of $\mathbf{S}_0$ is countably infinite. Indeed, it is true no matter what the cardinality of the set of primitive symbols of $\mathbf{S}_0$ is. But Henkin only demonstrates his construction for the case where the set of primitive symbols of $\mathbf{S}_0$ is countably infinite, and mentions that a simple modification of the proof is sufficient for the general case. (The modification, however, requires the Axiom of Choice!) Let us follow Henkin's proof:

Let u_{ij} $(i, j = 1, 2, 3, \cdots)$ be symbols not occurring among the symbols of $\mathbf{S}_0$. For each i $(i = 1, 2, 3, \cdots)$ let $\mathbf{S}_i$ be the first-order functional calculus whose primitive symbols are obtained from those of $\mathbf{S}_{i-1}$ by adding the symbols u_{ij} $(j = 1, 2, 3, \cdots)$ as individual constants. Let $\mathbf{S}_\omega$ be the system whose symbols are those appearing in any one of the systems $\mathbf{S}_i$. It is easy to see that the wffs of $\mathbf{S}_\omega$ are denumerable, and we shall suppose that some particular enumeration is fixed on so that we may speak of the first, second, $\cdots$, nth, $\cdots$ formula of $\mathbf{S}_\omega$ in the standard ordering.

Exercise 19 Describe in your own words each system $\mathbf{S}_i$ and give an argument as to why there are countably infinite formulas of $\mathbf{S}_i$. Also, describe in your own words the system $\mathbf{S}_\omega$ and give an argument as to why there are countably infinite formulas of $\mathbf{S}_\omega$.

We can use this ordering to construct in $\mathbf{S}_0$ a maximal consistent set of cwffs, Γ_0, which contains the given set Λ. (We use "cwff" to mean *closed* wff: a wff *which contains no free occurrences of any individual variable*.) Γ_0 is maximal consistent in the sense that if A is any cwff of $\mathbf{S}_0$ which is not in Γ_0, then $\Gamma_0, A \vdash f$; but not $\Gamma_0 \vdash f$.

Exercise 20 Explain in your own words what it means for a set of sentences to be maximal consistent.

To construct Γ_0 let Γ_{00} be Λ and let B_1 be the first (in the standard ordering) cwff A of $\mathbf{S}_0$ such that $\{\Gamma_{00}, A\}$ is consistent. Form Γ_{01} by adding B_1 to Γ_{00}. Continue this process as follows. Assuming that Γ_{0i} and B_i have been found, let B_{i+1} be the first cwff A (of $\mathbf{S}_0$) after B_i, such that $\{\Gamma_{0i}, A\}$ is consistent; then form Γ_{0i+1} by adding B_{i+1} to Γ_{0i}. Finally, let Γ_0 be composed of those formulas appearing in any Γ_{0i} $(i = 0, 1, \cdots)$. Clearly Γ_0 *contains* Λ. Γ_0 is *consistent*..

Exercise 21 Give the definition of Γ_0 in your own words. Show that Λ is contained in Γ_0 and that Γ_0 is consistent. (Hint: observe that if Γ_0 is inconsistent, then $\Gamma_0 \vdash f$; now analyze what it takes for f to be derivable from Γ_0.)

..Finally, Γ_0 is *maximal* consistent.

Exercise 22 Justify this claim of Henkin. (Hint: if $\{\Gamma_0, A\}$ is consistent for some cwff A, then what can you conclude about $\{\Gamma_{0i}, A\}$ for each i? What does it say about A?)

Having obtained Γ_0 we proceed to the system $\mathbf{S}_1$ and form a set Γ_1 of its cwffs as follows. Select the first (in the standard ordering) cwff of Γ_0 which has the form $(\exists x)A$ (unabbreviated: $((x)(A \supset f) \supset f)$), and let A' be the result of substituting the symbol u_{11} of $\mathbf{S}_1$ for all free occurrences of the variable x in the wff A. The set $\{\Gamma_0, A'\}$ must be a consistent set of cwffs of $\mathbf{S}_1$.

Exercise 23 Justify this claim of Henkin. (Hint: assume that $\{\Gamma_0, A'\}$ is inconsistent; then $\Gamma_0, A' \vdash f$; use the Deduction Theorem to obtain $\Gamma_0 \vdash A' \supset f$; use (IV) to deduce $\Gamma_0 \vdash A \supset f$; next use (II) to obtain $\Gamma_0 \vdash (x)(A \supset f)$; now use (8) and (I) to deduce $\Gamma_0 \vdash (\exists x)A \supset f$; finally, use Henkin's assumption and Modus Ponens to conclude that $\Gamma_0 \vdash f$, which is a contradiction.)

We proceed in turn to each cwff of Γ_0 having the form $(\exists x)A$, and for the j^{th} of these we add to Γ_0 the cwff A' of $\mathbf{S}_1$ obtained by substituting the constant u_{1j} for each free occurrence of the variable x in the wff A. Each of these adjunctions leaves us with a consistent set of cwffs of $\mathbf{S}_1$ by the argument above. Finally, after all such formulas A' have been added, we enlarge the resulting set of formulas to a maximal consistent set of cwffs of $\mathbf{S}_1$ in the same way that Γ_0 was obtained from Λ in $\mathbf{S}_0$. This set of cwffs we call Γ_1.

After the set Γ_i has been formed in the system $\mathbf{S}_i$ we construct Γ_{i+1} in $\mathbf{S}_{i+1}$ by the same method used in getting Γ_i[1] from Γ_0 but using the constants u_{i+1j} $(j = 1, 2, 3, \cdots)$ in place of u_{1j}. Finally we let Γ_ω be the set of cwffs of $\mathbf{S}_\omega$ consisting of all those formulas which are in any Γ_1.[2] It is easy to see that Γ_ω possesses the following properties:

i) Γ_ω is a maximal consistent set of cwffs of $\mathbf{S}_\omega$.
ii) If a formula of the form $(\exists x)A$ is in Γ_ω then Γ_ω also contains a formula A' obtained from the wff A by substituting some constant u_{ij} for each free occurrence of the variable x.

Exercise 24 Justify this claim of Henkin.

This concludes the first half of the Henkin construction. Henkin's final task is to construct a countably infinite domain, which will simultaneously satisfy Γ_ω. For this he will use the very constants he has enriched the language with!

Our entire construction has been for the purpose of obtaining a set of formulas with these two properties; they are the only properties we shall use now in showing that the elements of Γ_ω are simultaneously satisfiable in a denumerable domain of individuals.

In fact we take as our domain I simply the set of individual constants of $\mathbf{S}_\omega$, and we assign to each such constant (considered as a symbol in an interpreted system) itself (considered as an individual) as denotation. It remains to assign values in the form of truth-values to propositional symbols, and sets of ordered n-tuples of individuals

[1] This is a typo. Henkin really means Γ_1, not Γ_i.
[2] This is another typo. Henkin really means Γ_i, not Γ_1.

to functional symbols of degree n, in such a way as to lead to a value T for each cwff of Γ_ω.

Every propositional symbol, A, of $\mathbf{S}_0$ is a cwff of $\mathbf{S}_\omega$; we assign to it the value T or F according as $\Gamma_\omega \vdash A$ or not. Let G be any functional symbol of degree n. We assign to it the class of those n-tuples $\langle a_1, \cdots, a_n \rangle$ of individual constants such that $\Gamma_\omega \vdash G(a_1, \cdots, a_n)$.

∞∞∞∞∞∞

Exercise 25 Describe in your own words Henkin's interpretation.

∞∞∞∞∞∞

This assignment determines a unique truth-value for each cwff of $\mathbf{S}_\omega$ under the fundamental interpretation prescribed for quantification and "$\supset$". (We may note that the symbol "f" is assigned F in agreement with that interpretation since Γ_ω is consistent.) We now go on to show the

LEMMA: *For each cwff A of $\mathbf{S}_\omega$ the associated value is T or F according as $\Gamma_\omega \vdash A$ or not.*

The proof is by induction on the length of A. We may notice, first, that if we do not have $\Gamma_\omega \vdash A$ for some cwff A of $\mathbf{S}_\omega$ then we do have $\Gamma_\omega \vdash A \supset f$.

∞∞∞∞∞∞

Exercise 26 Justify this claim of Henkin. (Hint: use that Γ_ω is maximal consistent.)

∞∞∞∞∞∞

In case A is an elementary cwff the lemma is clearly true from the nature of the assignment.

Suppose A is $B \supset C$.

∞∞∞∞∞∞

Exercise 27 Prove that $B \supset C$ has the associated value T if and only if $\Gamma_\omega \vdash B \supset C$. (Hint: first assume that $B \supset C$ has the value T; this means that C has the value T or B has the value F; for each of these cases show that $\Gamma_\omega \vdash B \supset C$; next assume that $B \supset C$ has the value F; this means that B has the value T and C has the value F; for this case show that $\Gamma_\omega \nvdash B \supset C$; when C has the value T, you will need to use the induction hypothesis, (1), and (I); when B has the value F, you will need to use the induction hypothesis, (6), and (I); finally, when B has the value T and C has the valued F, you will need to use the induction hypothesis, (7), and (I).)

Suppose A is $(x)B$.

Exercise 28 Prove that $(x)B$ has the associated value T if and only if $\Gamma_\omega \vdash (x)B$. (Hint: first assume that $\Gamma_\omega \vdash (x)B$ and show that $(x)B$ has the value T; for this you will need to use (5), (I), and the induction hypothesis; next assume that $\Gamma_\omega \nvdash (x)B$ and show that $(x)B$ has the value F; for this you will need to use (9), (I), property (ii) of Γ_ω, and the induction hypothesis.)

This concludes the inductive proof of the lemma. In particular the formulas of Γ_ω all have the values T for our assignment and so are simultaneously satisfiable in the denumerable domain I. Since the formulas of Λ are included among those of Γ_ω our theorem is proved for the case of a system $\mathbf{S}_0$ whose primitive symbols are denumerable.

As a result, Henkin has finally completed the proof of his theorem. Now we know what the *Henkin method* is all about: it starts from a consistent set of sentences Λ, and extends it to a set of sentences Γ_ω, in the language enriched with new constants, which has properties (i) and (ii). Having all the new constants available, it is easy to construct a domain I and an interpretation into the domain such that a sentence (of the extended language) is satisfied in I if and only if it is derivable from Γ_ω.

Having mastered Henkin's method, we are one step away from establishing the completeness theorem.

The completeness theorem of the system $\mathbf{S}_0$ is an immediate consequence of our theorem.

COROLLARY 1: *If A is a valid wff of $\mathbf{S}_0$ then $\vdash A$.*

Exercise 29 Prove the above corollary. (Hint: first consider the case where A is a cwff; observe that since A is valid, then $A \supset f$ is not satisfiable; use Henkin's theorem to deduce that $\{A \supset f\}$ is inconsistent; next use the Deduction theorem to obtain that $\vdash A \supset f \supset f$; now use (3) and (I) to conclude that $\vdash A$; now reduce the case of a wff to that of a cwff by taking the *closure* A' of a wff A, where A' is obtained by prefixing to A universal quantifiers with respect to each individual variable with free occurrences in A in the order in which they appear; for this you will need to use successive applications of (5) and (I).)

Exercise 30 Prove that if $\Gamma \models A$, then $\Gamma \vdash A$. Hint: modify your argument for Exercise 29.

5 The Löwenheim–Skolem Theorem

Now that Henkin established the completeness theorem, his next task is to establish the Löwenheim–Skolem theorem, which appears as Corollary 2 in Henkin's treatise. The celebrated Löwenheim–Skolem theorem was first proved by Leopold Löwenheim (1878–1957) in 1915, and greatly simplified by Thoralf Skolem (1887-1963) in 1920. In this section we will learn that, similar to the completeness theorem, the Löwenheim–Skolem theorem also follows easily from Henkin's construction.

We recall that the Löwenheim–Skolem theorem states that if a first-order theory in a countable language has a model, then it also has a countably infinite model. Since Henkin's method works for first-order languages of any cardinality, Henkin is able to prove a more general version of the Löwenheim–Skolem theorem.

> COROLLARY 2: *Let $\mathbf{S}_0$ be a functional calculus of first order and $\mathbf{m}$ the cardinal number of the set of its primitive symbols. If Λ is a set of cwffs which is simultaneously satisfiable then in particular Λ is simultaneously satisfiable in some domain of cardinal $\mathbf{m}$.*

Exercise 31 Prove the above corollary. (Hint: first observe that if Λ is simultaneously satisfiable, then Λ is consistent; then use Henkin's Theorem.)

Exercise 32 Deduce the Löwenheim–Skolem theorem from Corollary 2 of Henkin's treatise.

Thus, we obtained that whenever a first-order theory in a countable language has a model, it also has a countably infinite model. Henkin also points out that whenever a first-order theory has a countably infinite model, it also has models of any higher cardinality. However, Henkin notices that it is not guaranteed that if a first order-theory has infinite models (of any cardinality), then it also has a finite model.

> It should be noticed, for this case, that the assertion of a set of cwffs Λ can no more compel a domain to be finite than non-denumerably infinite: there is always a denumerably infinite domain available. There are also always domains of any cardinality greater than $\aleph_0$ in which a consistent set Λ is simultaneously satisfiable, and sometimes finite domains. However for certain Λ no finite domain will do.

Exercise 33 Show that if Λ has a model, then Λ has models of any infinite cardinality. (Hint: use Henkin's theorem.)

Exercise 34 Give an example of a consistent Λ such that Λ has no finite models. (Hint: play with first-order theories of orderings.)

6 Adding Equality

So far Henkin has not addressed the case when the equality is part of the language. As we will see in this section, the situation with equality is very much similar to the case without equality: we can still prove the completeness theorem (and hence the Löwenheim–Skolem theorem) by Henkin's method. However, a little more care is required.

Along with the truth functions of propositional calculus and quantification with respect to individual variables the first-order functional calculus is sometimes formulated so as to include the notation of equality as between individuals. Formally this may be accomplished by singling out some functional constant of degree 2, say Q, abbreviating $Q(x, y)$ as $x = y$ (for individual symbols x, y), and adding the axiom schemata

E1. $x = x$

E2. $x = y \supset . A \supset B$, where B is obtained from A by replacing some free occurrences of x by a free occurrence of y.

For a system $\mathbf{S}_0'$ of this kind our theorem holds if we replace "the same cardinal number as" by "a cardinal number not greater than," where the definition of "simultaneously satisfiable" must be supplemented by the provision that the symbol "=" shall denote the relation of equality between individuals. To prove this we notice that a set of cwffs Λ in the system $\mathbf{S}_0'$ may be regarded as a set of cwffs (Λ, E_1, E_2) in the system $\mathbf{S}_0$, where E_1[3] is the set of closures of axioms E_i $(i = 1, 2)$. Since $E_1, E_2 \vdash x = y \supset y = x$ and $E_1, E_2 \vdash x = y \supset . y = z \supset x = z$ we see that the assignment which gives a value T to each formula of Λ, E_1, E_2 must assign some equivalence relation to the functional symbol Q. If we take the domain I' of equivalence classes determined by this relation over the original domain I of constants, and assign to each individual constant (as denotation) the class determined by itself, we are led to a new assignment which is easily seen to satisfy Λ (simultaneously) in $\mathbf{S}_0'$.

Exercise 35 Explain in your own words what Henkin means by "..a set of cwffs Λ in the system $\mathbf{S}_0'$ may be regarded as a set of cwffs (Λ, E_1, E_2) in the system $\mathbf{S}_0$.."

Exercise 36 Prove that $E_1, E_2 \vdash x = y \supset y = x$ and $E_1, E_2 \vdash x = y \supset . y = z \supset x = z$. Hint: do not confuse E_i and E_i for $i = 1, 2$.

Exercise 37 Based on Henkin's argument above, give your own proof of Henkin's Theorem for languages with equality. Use modern notation in your proof.

[3] This is a typo. Henkin really means E_i.

At the end of his treatise Henkin discusses the Löwenheim–Skolem theorem for languages with equality.

..Every axiom set with an infinite model has models with arbitrary infinite cardinality. For if α, β range over any set whatever the set of all formulas $\sim (x_\alpha = x_\beta)$ for distinct α, β will be consistent (since the assumption of an infinite model guarantees consistency for any finite set of these formulas) and so can be simultaneously satisfied.

Exercise 38 State the Löwenheim–Skolem theorem for languages with equality.

Exercise 39 Formalize Henkin's argument above and provide the missing details for the proof of the Löwenheim–Skolem theorem for languages with equality.

Exercise 40 Compare the Löwenheim–Skolem theorem for languages without equality with the Löwenheim–Skolem theorem for languages with equality.

7 Notes to the Instructor

The project is designed for an upper level undergraduate course in mathematical logic. It covers the core material of the course, and the whole course may be designed around the project. Although the beginning of the project addresses basic facts about first-order languages, their syntax and semantics, it is expected that students are already familiar with these topics by the time the project is assigned. Instructors may wish to spend about three to four weeks early in the semester covering these topics before assigning the project. The project itself may take anywhere from four to seven weeks depending on the pace of the instructor. The only core topics of the course not covered by the project are the compactness theorem, elementary classes and elementary equivalence, and Peano Arithmetic. These topics may be assigned in the last three to four weeks of the semester, right after finishing the project. The notation Henkin uses in his treatise was common at the time of writing the treatise, but it is slightly outdated from today's point of view. Instructors may wish to spend some time comparing Henkin's notation to the contemporary one.

References

[1] Alonzo Church, *Introduction to Mathematical Logic. I*, Annals of Mathematics Studies, no. 13, Princeton University Press, Princeton, N.J., 1944.

[2] H. Ebbinghaus, J. Flum, and W. Thomas, *Mathematical logic*, second ed., Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1994, Translated from the German by Margit Meßmer.

[3] Kurt Gödel, *Die Vollständigkeit der Axiome des logischen Funktionenkalküls*, Monatsh. Math. Phys. **37** (1930), 349–360.

[4] ———, *The Consistency of the Continuum Hypothesis*, Annals of Mathematics Studies, no. 3, Princeton University Press, Princeton, N. J., 1940.

[5] ———, *Collected works. Vol. I*, The Clarendon Press Oxford University Press, New York, 1986, Publications 1929–1936, Edited and with a preface by Solomon Feferman.

[6] Leon Henkin, *The completeness of the first-order functional calculus*, J. Symbolic Logic **14** (1949), 159–166.

[7] ———, *The discovery of my completeness proofs*, Bull. Symbolic Logic **2** (1996), no. 2, 127–158.

[8] Leon Henkin, J. Donald Monk, and Alfred Tarski, *Cylindric algebras. Part I. With an introductory chapter: General theory of algebras*, North-Holland Publishing Co., Amsterdam, 1971, Studies in Logic and the Foundations of Mathematics, Vol. 64.

[9] ———, *Cylindric algebras. Part II*, Studies in Logic and the Foundations of Mathematics, vol. 115, North-Holland Publishing Co., Amsterdam, 1985.

[10] David Hilbert and Wilhelm Ackermann, *Grundzüge der theoretischen Logik*, Springer–Verlag, Berlin, 1928.

[11] Ernest Nagel and James R. Newman, *Gödel's proof*, New York University Press, New York, 1958.

[12] Oswald Veblen and John W. Young, *Projective geometry. Vol. I*, Ginn and Company, New York, 1910.

[13] ———, *Projective geometry. Vol. II*, Ginn and Company, New York, 1918.