How first order is first order logic?*

Juliette Kennedy
Department of Mathematics and Statistics
University of Helsinki, Finland

Jouko Väänänen
Department of Mathematics and Statistics
University of Helsinki, Finland
ILLC, University of Amsterdam
Amsterdam, Netherlands

June 14, 2023

Contents

*The first author would like to thank the Academy of Finland, grant no: 322488. The second author would like to thank the Academy of Finland, grant no: 322795. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 101020762).

1 Introduction

Fundamental to the practice of logic is the dogma regarding the first order/second order logic distinction, namely that it is *ironclad*. Was it always so? The emergence of the set theoretic paradigm is an interesting test case. Early workers in foundations generally used higher order systems in the form of type theory; but then higher order systems were gradually abandoned in favour of first order set theory—a transition that was completed, more or less, by the 1930s.[1]

Gödel [*1933o] describes the higher order "provenance" of (first order) set theory—*the fact that set theory lends itself to being viewed in a natural way as a higher order system*—as follows:

> It may seem as if another solution were afforded by the system of axioms for the theory of aggregates, as presented by Zermelo, Fraenkel and von Neumann; but it turns out that this system of axioms is nothing else but a natural generalization of the theory of types, or rather, it is what becomes of the theory of types if certain superfluous restrictions are removed.[2]

Set theory, then, has a double nature—logically speaking.

Thinking of logic in general, there is some tension in what the phrase "first order," as in "first order logic," really means. It is clear what it means for a model class K to be first order definable.[3] It means that there is a first order formula ϕ such that for all models M

$$M \in K \iff M \models \phi. \tag{1}$$

[1] According to Hodges the transition in Tarski's early work, at least, from simple type theory to informal set theory, was in place by then. As Hodges puts it: "The deductive theories in question (such as RCF) are formulated in simple type theory; by 1935 the axioms for RCF is regarded as a definition within set theory." [14], p. 118. See Ewald's [9] for the emergence of first order logic.

[2] [11], p. 45. A similar point is made in the second author's [37]: "First-order set theory is merely the result of extending second order logic to transfinitely high types."

[3] By a model class we mean a class of models, closed under isomorphisms, of a fixed vocabulary, see Section 2.

On the other hand, K being merely a model class means, from the point of view of set theory[4], that there is a first order formula $\Phi(x)$, perhaps with set parameters, such that for all models M

$$M \in K \iff \Phi(M). \tag{2}$$

While (1) holds for a very restricted collection of model classes only, (2) holds for *all* model classes. Still both (1) and (2) seem to be based on first order logic. Obviously first order logic is playing a different role in (1) and (2): in (1) first order logic "views" the model M in some sense from the *inside*, while in (2) the perspective taken is from the *outside*, how M sits in the universe of sets.

This simple observation suggests that the concept of a logic being first order is not only about whether the variables range over the elements of a given domain, or over sets of elements, or over sets of sets of elements, and so on, it is also about the *context*. In (1) the first order variables range over elements of M, while in (2) the first order variables range over the universe of set theory V, which contains sets generated by unbounded, even transfinite, iterations of the power set operation, i.e. objects of all orders. This means that higher order quantification (over set-size domains) is in a clear sense allowed in (first order) set theory.

Of course, set theory is a *theory* and second order logic is a *logic*, at least that is the common understanding. We will argue in this paper that if one cares to view set theory as a logic—and if we do think of set theory as a logic, it is a logic with the cumulative hierarchy V as its standard (class) model—then set theory turns out to be a stronger logic than second order logic. This is perhaps as it should be, given that the latter restricts the domain of quantifiable objects to those generated by (at most) a *single* iteration of the power set operation, while set theory allows for arbitrary iterations of the power set operation, as was mentioned.

It is useful at the beginning to recall an elementary observation about the first order/second order distinction. Consider the structure

$$\mathcal{M} = (\mathbb{R}, +, \times, \mathbb{N}, <, 0, 1).$$

So here we have the all-important structure of the real numbers together with the structure of natural numbers conceived of as a substructure of the reals (as can be expected). First order quantification over this structure involves quantification over the real numbers as we see. Via their binary

[4]See e.g. [15, Page 5.].

representation, definable in this structure, every real number corresponds canonically to a subset of $\mathbb{N}$. Thus when we quantify in a first order way over the real numbers we are implicitly quantifying in a second order way over natural numbers, because we can identify a real number with a subset of $\mathbb{N}$. Thus first order quantification over the reals can be viewed as second order quantification over the naturals.

One may ask, does first order logic over the reals in $\mathcal{M}$ benefit—or suffer?—from all the special properties of second order logic, it being (implicitly) second order logic over the integers? It both benefits and suffers, depending on what we are after. The first order theory of $\mathcal{M}$ is extraordinarily complex from the point of view of computability theory as it encodes the entire second order theory of $(\mathbb{N}, <, 0, 1)$, well-known to be non-computable in the extreme. This should be contrasted with the fact that the arithmetic of the reals alone is decidable [32]. The point is that the decidability concerns the structure $(\mathbb{R}, +, \times, <, 0, 1)$, in which $\mathbb{N}$ is not a part of the structure (it is not even a *definable* subset). Thus the presence or non-presence of $\mathbb{N}$ as part of the structure $\mathcal{M}$ decides whether first order quantification over the model is truly first order or implicitly second order over an infinite substructure. This phenomenon is an important aspect of second order logic, namely that second orderness may be present in a hidden way. Quine made a similar point in [29] when he suggested that using second order predicate symbols as schematic letters masks the set theoretic content of second order logic; one should rather include the membership relation in the given signature.

Another ambiguity in the notion of "first order" is due to the fact that there is a whole spectrum of logics which extend first order logic in the sense of (1) but are sublogics of first order logic in the sense of (2).[5] *In fact, every (abstract) logic is first order from the point of view of set theory.* This is because an *abstract logic* is given by two predicates of set theory, namely the set (or class) of formulas and the truth predicate, where the latter is a predicate of set theory holding between structures and sentences of the logic. Such predicates are given in set theory by a first order formula. Of course the first order formula involves the epsilon relation. Hence in first order logic together with the epsilon relation one can define every logic.

Let us look at a very simple example:

Example 1. *A typical non-first order property of a model is its finiteness. Let K be the class of finite models of some vocabulary L. It is a familiar*

[5] See below Section 3.1 for the definition of abstract logic and some examples.

consequence of the Compactness Theorem[6] *that there is no first order sentence ϕ of any vocabulary L such that for all models M the equivalence (1) holds (for the class K). On the other hand, if $\Phi(x)$ is the familiar*[7] *set-theoretical formula which says that the set x is a finite model of vocabulary L, then all models M of vocabulary L satisfy (2). First order logic tries to express finiteness from inside the model and fails; first order set theory, just like second order logic, by virtue of the power set operation, easily expresses finiteness from the outside.*

So what does "first order" mean, after all, if first order logic can appear in such different roles as (1) and (2)? In this paper we will look at first order logic from various points of view, arguing that the distinction between first order and higher order logics, such as second order logic, is somewhat context dependent. From the philosophical or foundational point of view this complicates the picture of first order logic as a canonical logic.

2 Absoluteness and symbiosis

Barwise [4] pinned the canonicity of a logic to its *absoluteness*:

> When is it reasonable for us, as outsiders looking on, to call [an abstract logic] L^* a "first order" logic? If the words "first order" have any intuitive content it is that the truth or falsity of $M \models^* \phi$ should depend only on ϕ and M, not on what subsets of M may or may not exist in [the logician] k's model of his set theory T. In other words, the relation $\models^*$ should be absolute for models of T.

First order logic is of course absolute, but there are many extensions of first order logic that are absolute in Barwise's sense, and therefore, in Barwise's sense, first order, or at least close to first order. Examples of absolute logics include $L(Q_0)$ [28], $L(Q_0^{\mathrm{MM}})$ [24], $L_{\infty\omega}$ [16], and the logic $L_{\infty G}$ [12, 27] with the closed game-quantifier

$$\forall x_0 \exists x_1 \forall x_2 \exists x_3 \ldots \bigwedge \phi_n(x_0, \ldots, x_{2n+1}). \tag{3}$$

[6] Every theory, which has no models, has a finite subtheory without models.

[7] There are many different definitions of finiteness in set theory, all equivalent if the Axiom of Choice is assumed. The most common definition says that there is no one-one function from x into a proper subset of x.

In order to talk about the absoluteness of a logic more exactly we first define the concept of an abstract logic: A *logic* (a.k.a. abstract logic) in the sense of [22] is a pair $L^* = (\Sigma, T)$, where Σ is an arbitrary set (sometimes also a class) and T is a binary relation between members of Σ on the one hand and structures on the other. Members of Σ are called L^*-sentences. Classes of the form

$$\mathrm{Mod}(\phi) = \{\mathcal{M} : T(\phi, \mathcal{M})\},$$

where ϕ is an L^*-sentence, are called L^*-characterizable, or L^*-definable, classes. Abstract logics are assumed to satisfy five axioms expressed in terms of L^*-characterizable classes, corresponding to being closed under isomorphism, conjunction, negation, permutation of symbols, and "free" expansions.[8] A model class is the abstraction of a sentence. We conceive of formulas with free variables by considering model classes in which the vocabulary has (new) constant symbols that play the role of free variables. This makes it possible to define what it means for an abstract logic to be closed under first order quantification. A class $\mathcal{K}$ of models is said to be *definable* in a logic L^* if there is a sentence ϕ in L^* such that

$$\mathcal{K} = \mathrm{Mod}(\phi).$$

A logic L^* is a *sublogic* of another logic L^+, $L^* \leq L^+$, if every L^*-definable model class is also L^+-definable. The logics are *equivalent*, $L^* \equiv L^+$, if $L^* \leq L^+$ and $L^+ \leq L^*$.

Why consider model classes, if we are thinking about logics? Tarski [33] introduced the concept of an *elementary* class. This refers to the class of all models of a given first order sentence ϕ with vocabulary L. Thus elementary classes K satisfy:

1. The elements of K are models (i.e. structures).
2. All models in K have the same vocabulary L.
3. K is closed under isomorphisms.

Generalizing from this, we call any class K a *model class* if it satisfies the above three conditions. Elementary classes are examples of model classes but there are many more. For example, the classes of all groups, all well-orders, all equivalence relations, all algebraically closed fields, all models of Peano arithmetic, all models of ZFC set theory, and the class of all models

[8] The free expansion to vocabulary L of a model class K of a smaller vocabulary is the class of all expansions of elements of K to the vocabulary L.

isomorphic to some $(V_\alpha, \in)$, where α is an ordinal, are model classes. If ϕ is a sentence of any logic whatsoever, be it second order logic, logic with generalized quantifiers, or infinitary logic, the class of models of ϕ is a model class.

An interesting fact about model classes is that every model class is definable in *some* logic, because we can take the model class as a generalized quantifier in the sense of [21]: Suppose $\mathcal{K}$ is a model class with vocabulary L. For simplicity we assume $L = \{R\}$ where R is a binary predicate symbol. We can associate with $\mathcal{K}$ the generalized quantifier $Q_\mathcal{K}$ in the sense of [21] with the semantics

$$\mathcal{M} \models Q_\mathcal{K} xy\phi(x, y, \vec{a}) \iff (M, \{(b, c) \in M^2 : \mathcal{M} \models \phi(b, c, \vec{a})\}) \in \mathcal{K}.$$

Now $\mathcal{K}$ is trivially definable in the extension $L_{\omega\omega}(Q_\mathcal{K})$ of first order logic by the quantifier $Q_\mathcal{K}$ by the sentence

$$Q_\mathcal{K} xyR(x, y).$$

Conversely, every class of models definable in $L_{\omega\omega}(Q_\mathcal{K})$, or indeed in any abstract logic, is a model class i.e. is closed under isomorphisms.

To summarize the discussion of the previous paragraphs: *talking about model classes is tantamount to talking about sentences in arbitrary logics.*

Returning to absoluteness, the absoluteness of a logic means that its set theoretical definition is absolute in the usual set theoretical sense, i.e. persisting upwards and downwards across transitive models of (of a finite part of) ZFC. More precisely a logic is said to be *absolute* [4] if the satisfaction predicate $M \models \phi$ is Δ_1 in the Levy hierarchy, and the property of being a formula of the logic is Σ_1 in the Levy hierarchy.[9]

Essentially a logic L is absolute if the truth of a sentence in an L-structure depends only on the elements of the domain, not on what kind of subsets it has.

In contrast to first order logic, second order logic is famously nonabsolute, because of its entanglement with set theory. For example, one can easily write a second order sentence Φ which is true in the ordered field

[9] The Levy hierarchy is defined as follows: Σ_0-formulas of set theory are formulas in which all quantifiers are bounded i.e. of the form $\forall x \in y$, $\exists x \in y$. Π_0-formulas are the same as Σ_0-formulas. A formula is Σ_{n+1} if it is of the form $\exists x\phi$, where ϕ is Π_n. A formula is Π_{n+1} if it is of the form $\forall x\phi$, where ϕ is Σ_n. A property of sets is Δ_n if it can be defined both by a Σ_n-formula and a Π_n-formula. If the equivalence can be proved in the theory T, typically Kripke-Platek set theory KP or ZFC, the property is called Δ^T_n. The Kripke-Platek axioms KP consist of some elementary axioms plus the Σ_0-separation and Σ_0-collection schemas.

$(\mathbb{R}, +, \cdot, 0, 1, <)$ of real numbers if and only if the Continuum Hypothesis CH holds,[10] and this equivalence is provable in ZFC. Now the predicate $(\mathbb{R}, +, \cdot, 0, 1, <) \models \Phi$ is non-absolute, for we can start with a model in which CH is false and then collapse the continuum to $\aleph_1$ without changing the reals. First $(\mathbb{R}, +, \cdot, 0, 1, <) \models \phi$ is false but after the collapse it is true. Such a "change of heart" on the part of Φ would not be allowed if Φ were a sentence of an absolute logic.[11]

2.1 Symbiosis

One may ask, what prevents second order logic, or indeed any logic, from being absolute? The answer to this question must lie in the nature of the connection of the logic to set theory, that is to say the answer to the question must lie in the set-theoretical content of the logic—but what exactly is the set-theoretical content of e.g. second order logic? *Symbiosis*, introduced in [35], was designed exactly in order to bring the set theoretical content of a logic to the fore; to "expose the nature of the logic, to uncover the set-theoretical commitments of the logic, its content, its strength, even its reference."[12]

Precisely, in symbiosis one finds a set-theoretical predicate or operation P such that in any situation where P is absolute the logic $\mathcal{L}$ is, and (roughly) vice versa. This means that one is able to detect, on the one hand, whether a logic "sees" the invariant content of a given set theoretic predicate; while on the other hand the absoluteness of the *logic* is pinned to the absoluteness of the *predicate*—whence the name "symbiosis."

So why is second order logic nonabsolute? Or to put it another way, what is the nature of this entanglement of second order logic with set theory, an entanglement so decried by Quine?[13] It is expressed in symbiosis this way: second order logic is actually *symbiotic* with the power set operation of set theory. That is to say, once we hold the power set operation fixed, second

[10] The CH says that every uncountable set of reals has the same cardinality as the set of reals itself. This is expressible in second order logic because we can quantify over all subsets of the domain, we can express countability and we can express being of the same cardinality as the entire domain. Let Φ be the second order sentence $\forall P(\Psi(P) \rightarrow \exists F\Theta(F, P))$, where $\Psi(P)$ says "P is uncountable" and $\Theta(F, P)$ says "F is a one-one function from elements of the domain into P." Then Φ holds in $(\mathbb{R}, +, \cdot, 0, 1, <)$ iff CH is true.

[11] Our example of non-absoluteness of second order logic is an overkill. A simpler example is uncountability. No absolute logic can express uncountability, as shown in [4].

[12] See [18].

[13] For Quine's view of the entanglement of second order logic with set theory see the section entitled "Set theory in sheep's clothing," [29], p. 66.

order logic becomes absolute. On the other hand, second order logic "sees" the power set operation and can talk about it and everything else that is "absolute relative to the power set operation," (see Definition 2 below) via its definable model classes.

The symbiosis between second order logic and the power set operation is described in [19, p. 173] thus:

> When we talk about mathematics based on the power set operation in some vague sense, it seems immaterial whether we use second order logic or predicates that are absolute with respect to the power set operation, P-absolute. For second order logic we have the formal counterpart in the formal second order language. Likewise, for P-absoluteness we have $\Delta_1(P)$- formulas which capture the P-absolute properties, if we want to use a formal concept. Of course the predicate "x is the power set of y" is absolute in this sense only with respect to a special class of models of set theory. But the point here was not to study this particular binary predicate on its own; the aim was to determine whether a given logic could "see" the predicate through its definable model classes.

Of course it is not surprising that the nonabsoluteness of second order logic should be tied to the power set operation, somehow. Symbiosis makes this exact and explicit. Moreover, from the symbiosis point of view, the moral is clear: *it is useless to try to separate second order logic from set theory.*

Before giving the technical definition of symbiosis, we require the auxiliary concept of R-absoluteness, mentioned above:

Definition 2. *Suppose R is a predicate (i.e. a first order formula) in the language $\{\epsilon\}$. A predicate P is absolute w.r.t. R, or R-absolute, if it is absolute with respect to transitive extensions preserving the predicate R, i.e. if the predicate is preserved by extensions of the universe as long as R itself is preserved, and no new elements are added to old elements (technically: extensions of transitive models); and the same is true of restrictions of the universe.*

Technically this is the same as P being $\Delta_1(R)$ i.e. Δ_1 in the extended language $\{\in, R\}$ [10].

Intuitively, a predicate P is R-absolute if whenever we add sets to the universe or take sets away, without changing R, also P remains unchanged. For example, if $R(x)$ is the predicate "x is countable," then the predicates

- "x is uncountable,"
- "x is a countable ordinal,"
- "x is a countable set of singletons,"
- "$(A, <)$ is a linear order in which every initial segment is countable,"
- "G is a graph in which every node has uncountably many neighbours,"

are all R-absolute.

Finally, we need the notion of Δ-operation. We refer to the Appendix for the definition and properties of this operation on logics. In principle the Δ-extension of a logic L, denoted $\Delta(L)$, uncovers the hidden power of a logic. It preserves properties like compactness, axiomatizability, Hanf and Löwenheim numbers; it "fills the gaps" left by explicit definability in the sense that if a model class is "implicitly" definable in the logic then it is explicitly definable in the Δ-extension. For example, $L(Q_0)$ cannot say that an equivalence relation has infinitely many equivalence classes,[14] although "morally" it should be able to do so, whereas $\Delta(L(Q_0))$ can say it easily. Essentially, when we consider $\Delta(L)$ rather than the logic L itself, we focus on what the logic becomes when some accidental weaknesses are removed. Maybe $L(Q_0)$ was defined as it was because the generalized quantifier Q_0 is appealing in its simplicity. But it turns out that $\Delta(L(Q_0))$ is an infinitary logic, viz. the logic[15] L_{HYP}, the smallest admissible fragment of $L_{\omega_1\omega}$ [5]. For a proper treatment of $L(Q_0)$ we have to consider the entire $\Delta(L(Q_0))$, otherwise our investigation may be centred around some accidental properties of $L(Q_0)$ with no general interest.

We now define the notion of symbiosis:

Definition 3. *An n-ary predicate R and a logic $\mathcal{L}^*$ are* symbiotic *if the following conditions are satisfied:*

1. *Every $\mathcal{L}^*$-definable model class is absolute w.r.t. R.*
2. *Every model class which is absolute w.r.t. R is $\Delta(\mathcal{L}^*)$-definable.*

Theorem 4. *If L^* and P are symbiotic, then an arbitrary model class is $\Delta_1(P)$ in set theory if and only if it is $\Delta(L^*)$-definable in model theory.*

[14] The proof of this is an easy application of the method of Ehrenfeucht-Fraïssé games.

[15] See Subsection 2.3.

In terms of the inside/outside metaphor, here the model theoretic definition of the class corresponds to the inside view, whereas the set theoretic definition of the class corresponds to viewing the class from the outside.[16]

We suggested above that because set theory (expressed in a first order language) provides a first order way to quantify in any given structure over not only elements of the structure but also over subsets (second order logic!), sets of subsets (third order logic!), etc., in this sense (first order) set theory is a very strong, indeed the strongest logic. Symbiosis is an exact analysis of this relationship between strong logics and predicates of set theory—an analysis revealing, again, the set-theoretical content of the logic in question. Such an analysis could arguably be done on the level of axioms. For most abstract logics some natural axioms are known even if the axioms are not complete, and even cannot be complete because the set of valid sentences of the logic is not recursively enumerable (as in the case of second order logic). But a more general approach is to work semantically and thereby avoid all questions of axiomatizabilty.

The most blatant example of symbiosis is that already mentioned, between second order logic and the power set predicate "x is the power-set of y". Another is the symbiosis between the Härtig (or equicardinality) quantifier and the predicate $\mathrm{Cd}(x)$ i.e. "x is a cardinal number". Table 1 displays various logics together with the predicates of set theory with which the logic is symbiotic.[17]

2.2 Sort logic

Before commenting on the very important case of the symbiosis of sort logic with full set theory, we need to define sort logic. Recall that a relational structure, i.e. a model, has a domain and relations, functions and constants on that domain. A modification is a *many-sorted structure* [13] in which there are several domains and relations, functions and constants on those domains or between the domains. A good example is a vector space, where there is a domain for scalars and a domain for vectors. A vector multiplied by a scalar is again a vector. To be able to talk about many-sorted structures in logic one adopts variables of different *sorts*, one sort for each domain. Thus in a language for vector spaces there is a sort for scalars and a sort for vectors. In other words, every individual variable has a sort attached to it

[16] In the appendix we explain the meaning of symbiosis in the special case of second order logic and the operation with which it is symbiotic, namely the power set operation.

[17] In the table the theory KPU^- is the theory KP with urelements and without the Axiom of Infinity.

Inside: $\mathcal{M} \models \phi$	Outside: $\Phi(\mathcal{M})$	Reference
Sort logic	First order logic	[38]
Δ(Second order logic)	Δ_1(Pw) in the Levy-hierarchy	[35]
Δ(First order logic with the Härtig-quantifier)	Δ_1(Cd) in the Levy-hierarchy	[35]
Δ(First order logic with the game quantifier)	Δ_1 in the Levy-hierarchy	[8]
Infinitary logic L_{HYP}	Δ_1^{KP} in the Levy-hierarchy	[6]
First order logic	$\Delta_1^{KPU^-}$ in the Levy-hierarchy	[1]

Figure 1: Manifestations of symbiosis

and it is supposed to range over elements of the domain of that sort. Thus in a language for vector spaces there are variables for scalars and variables for vectors.

In second order many-sorted logic we have to declare for each second order variable what are the sorts involved. Is it a relation symbol between elements of sort s and sort s'? Is it a function from arguments of sort s to a value of a sort s', etc. An example of a second order sentence in many sorted logic could be a sentence saying that there is a bijection between elements of sort s and elements of sort s'.

Sort logic [38] arises when we are allowed to quantify over a variable of a *new* sort i.e. a sort not present in our vocabulary. Semantically this means that we claim there is a new domain that can be added as a new sort into our model and the expanded model satisfies what we want to say. For example, we may want to say of a group that it is the multiplicative group of a field. We have to say that there is a zero-element outside the group so that the group together with the new element and a new addition form a field. This is particularly significant in many-sorted second order logic. We may ask whether there is a new domain with relations making the new domain essentially the power-set of the power-set of the union of the old domains. This allows us to reduce third order logic to sort logic. In fact sort logic contains higher order logics of all orders definable in second order logic.

A particularly striking instance of symbiosis is that between sort logic ([38]) and first order set theory. In this case the symbiosis means that every model class definable by a sentence of sort logic is (a fortiori) definable in first order set theory, and, conversely, any model class definable in set theory by a first order formula is definable in sort logic. This is means that sort logic is the strongest logic; and thus, by symbiotic correspondence, set theory is too.

2.3 $L_{\mathbf{HYP}}$

Another logic appearing in the table is the logic L_{HYP}. Barwise introduced in [3] the concept of an "admissible" fragment of $L_{\omega_1\omega}$: If A is a countable transitive set, the fragment L_A is simply $L_{\omega_1\omega}\cap A$, that is, we take from $L_{\omega_1\omega}$ the sentences that are, as set-theoretical objects[18], elements of A, together with their truth-definition as formulas of $L_{\omega_1\omega}$. If A (or rather $(A,\in)$) is not only transitive but also satisfies the so-called Kripke-Platek axioms[19] of set theory, making it an *admissible set*, then L_A is an interesting infinitary logic with a completeness theorem, interpolation theorem, etc. It is called an *admissible fragment*. The smallest non-trivial admissible fragment L_{HYP} is obtained from the smallest non-trivial admissible set HYP. This set is actually the fragment $L_{\omega_1^{\mathrm{CK}}}$ of Gödel's constructible hierarchy, where ω_1^{CK} is the smallest non-recursive countable ordinal. Informally L_{HYP} can be described as the closure of first order logic under recursive, rather than countable, conjunctions and disjunctions. The extension $L(Q_0)$ of first order logic by the generalized quantifier Q_0 ("there are infinitely many") is, of course, a sublogic of L_{HYP} because

$$Q_0x\phi(x) \equiv \bigwedge_{n<\omega} \exists x_0 \ldots \exists x_{n-1} \bigwedge_{i<j<n} (x_i \neq x_j \wedge \phi(x_i)).$$

Interestingly ([6])

$$\Delta(L(Q_0)) \equiv L_{\mathrm{HYP}}.$$

Thus the recursive conjunctions and disjunctions of L_{HYP} "fill" all the "definability-gaps" that are generated into $L(Q_0)$ when Q_0 is added to first order logic.

[18]Ordinarily we think of sentences as finite strings of symbols. However, in infinitary logic it is more natural to think of sentences as sets. For example, $\bigvee_{i\in I}\phi_i$ is the set $\{\bigvee, \{\phi_i : i \in I\}\}$, see e.g. [17, Page 6.]

[19]Recall that the Kripke-Platek axioms KP consist of some elementary axioms plus the Σ_0-separation and Σ_1-collection schemas.

2.4 Applications of symbiosis

Symbiosis has applications also outside of philosophy, i.e. beyond is its ability to calibrate the set-theoretic content of a logic. Symbiosis can be used for:

1. The non-absoluteness of a logic can be delineated in terms of predicates of set theory. This may be useful for a better understanding of, for example, the behavior of the logic in forcing extensions. There are many examples of this, e.g. [36], [25].

2. One can relate Löwenheim-Skolem type model theoretic properties of logics with reflection or large cardinal properties of cardinals in set theory [2]. An early example of this is the fact that the smallest κ for which second order logic satisfies the Löwenheim-Skolem-Tarski Theorem at κ (i.e. for every second order sentence ϕ every model has an elementary submodel of cardinality $< \kappa$ in which ϕ is true) is exactly the same as the smallest supercompact cardinal[20] [23].

3. On can relate the complexity of the decision problem of a logic[21] with set-theoretic definability criteria. An example of this is the result that the decision problem of second order logic is the complete Π_2-definable set of integers [34]. Another example is that the decision problem of the extension of first order logic with the Härtig-quantifier can be Π_2-complete but also can be Δ^1_3 [36].

In conclusion, philosophically speaking, symbiosis lays down a bridge between the interior (1) and the exterior (2) view of a logic. In both perspectives first order logic is in a central role. From the interior point of view it is the weakest logic; from the exterior point of view it is the strongest.

3 Other ways in which the first order/second order distinction (nearly) collapses

We discuss three ways, other than the above, how first order logic and second order logic may resemble each other, sometimes in unexpected ways.

[20]A cardinal κ is supercompact if for every λ there is an elementary embedding $i : V \to M$, M transitive, with critical point κ such that $M^\lambda \subseteq M$.

[21]The decision problem of a logic is the set of Gödel numbers of the valid sentences of the logic. Obviously this only makes sense, and is used, for logics with finite formulas.

3.1 Strong logics coming close to being first order

We mentioned the spectrum of logics which extend first order logic in the sense of (1) but are sublogics of first order logic in the sense of (2). This reveals another ambiguity in the notion of "first order," in that from the point of view of Lindström's theorem, characterizing first order logic in terms of certain canonical model theoretic properties,[22] some of these strong logics come very close to being first order by virtue of these properties; the logic satisfies a Compactness Theorem and a Downward Löwenheim-Skolem theorem in the same spirit as first order logic.

Example 5 (Cofinality logic). *[31] Consider the generalized quantifier*

$$M \models Q^{cof}_{\omega} xy\phi(x, y, \vec{a}) \iff$$
$$\{(b, c) \in M^2 : M \models \phi(b, c, \vec{a})\} \text{ is a linear order of cofinality } \omega.$$

The extension $L(Q^{cof}_{\omega})$ of first order logic by the quantifier Q^{cof}_{ω} is fully compact ([31]), meaning that it satisfies the Compactness Theorem in vocabularies of any cardinality. In this respect it is "close" to first order logic. It satisfies also the following strong form of the Downward Löwenheim-Skolem Theorem: Given any model M and a subset X of M of cardinality $\aleph_1$, there is a submodel N of M containing X such that the cardinality of N is $\aleph_1$ and N is an elementary submodel of M in the sense of the logic $L(Q^{cof}_{\omega})$. First order logic satisfies the same Downward Löwenheim-Skolem Theorem but with $\aleph_1$ replaced by $\aleph_0$. On the surface, $L(Q^{cof}_{\omega})$ looks very much like first order logic, but of course it is a proper extension.

Example 6 (Stationary logic). *[7][31] Consider the second order generalized quantifier*

$$M \models aas\phi(s, \vec{a}) \iff$$
$$\{P \subseteq M : |M| \leq \omega, (M, P) \models \phi(P, \vec{a})\} \text{ contains a club.}$$

Here a club means a closed unbounded set of countable subsets of M. The extension $L(aa)$ of first order logic by the quantifier aa is countably compact, meaning that it satisfies the Compactness Theorem in countable vocabularies. In this respect it is "close" to first order logic although perhaps not as "close" as its sublogic $L(Q^{cof}_{\omega})$. It satisfies also the following Downward Löwenheim-Skolem Theorem: Given any model M, a subset X of M of cardinality $\aleph_1$,

[22] Lindström's theorem states that first order logic is, up to equivalence of logics, the only logic closed under some elementary operations and satisfying the Compactness Theorem as well as the Downward Löwenheim-Skolem theorem (every sentence which has a model has a countable model).

and a sentence ϕ from $L(aa)$, there is a submodel N of M containing X such that the cardinality of N is $\aleph_1$ and N is a model of ϕ. This is a slightly weaker property than what $L(Q_\omega^{cof})$ has. Still, on the surface, $L(aa)$ looks very much like first order logic, albeit in a slightly different sense than $L(Q_\omega^{cof})$ (both the Compactness and Downward Löwenheim-Skolem Theorems are slightly weaker), but of course it is again a proper extension. It has $L(Q_\omega^{cof})$ as a sublogic and deeper analysis shows that it is much stronger than $L(Q_\omega^{cof})$.

The two logics defined above are properly between first and second order logics, however manifesting properties typical of first order logic rather than second order logic. An interesting question is whether the isolation of these logics helps us understand cofinality or stationarity better than simply working in first order set theory, set theory being the usual home territory of these concepts. Recent work in inner model theory suggests that these logics really contribute something over and beyond their set-theoretical analogues [20]. Interestingly, some other logics that otherwise seem non-first order, behave like first order logic in this inner model context [20]. There is also the question whether the existence of strong logics that resemble first order logic in their model-theoretic properties, as in the above examples, may help to shed light of the nature of *first* order logic.

3.2 Second order logic and internal categoricity

The prime example of a non-first order logic is second order logic L^2. Can L^2 be seen as a first order logic? It is, via symbiosis, a fragment of set theory and in that sense it can be represented in first order logic if we are granted $\in$. On the other hand, we can treat L^2 as a two-sorted logic with an individual-sort for elements and a set-sort for sets, relations and functions. This only makes sense if the Comprehension Schema is assumed in order that L^2 has some second order content.[23] For example, without the Comprehension Schema we do not know whether $\forall x \exists X \forall y (yEX \leftrightarrow x = y)$ is valid, although it clearly should be valid. In this two-sorted version second order logic is really just two-sorted first order logic. However, it is not a completely general two-sorted first order logic because the Comprehension Schema creates highly non-trivial structure into the set-sort. Since the relation xEX between a set-sort variable X and an individual-sort variable x has affinity with the $\in$-relation of set theory, second order logic as a two-sorted first order logic

[23]The Comprehension Schema states that every definable (with parameters) set of subsets (or relations or functions) is in the range of the second order variables. There is a natural restriction to prevent circular definitions.

still retains some of the strength of the original L^2. We cannot characterise finiteness, because first order logic satisfies the Compactness Theorem, also in its many-sorted version. But there is enough of a remnant of full second order logic in the two-valued first order version of L^2 that the Comprehension Schema as a first order theory can be considered an unclassifiable first order theory in the sense of stability theory [30]. If, to avoid unclassifiability, the Comprehension Schema is abandoned, there is nothing left of second order logic[24] and we are solidly back in full-blooded first order logic.

Just as first order set theory is investigated by the method of transitive models of first order ZFC, second order logic, in its original syntax or alternatively as a two-sorted first order language, can be investigated by the method of Henkin models i.e. sufficiently large collections of sets of urelements, relations between urelements and functions between urelements in order that the Comprehension Schema holds. In fact, it is difficult to imagine how second order logic could be developed otherwise. There is no miraculous oracle that can reveal to us properties of sets, relations and functions. We have to study the axioms and their models and thereby come to isolate new axioms. In set theory this has led to the introduction of Large Cardinal Axioms, Forcing Axioms and the Inner Model Program.

But what happens to the cherished categoricity results of second order logic, if second order logic is interpreted as first order many-sorted logic? Recent (and in some cases not so recent) results show that categoricity results hold also in the many-sorted version of second order logic, and can be proved from the Comprehension Schemas [40]. They even hold for first order Peano arithmetic and first order ZFC [39]. So the categoricity results of second order logic, despite their smooth formulation in second order logic, turn out to be results about first order logic. The "second-orderness" of *second* order logic is hereby somewhat undermined.

3.3 The metatheory problem

The received wisdom is that second order logic is a powerful logic capable of expressing finiteness, countability, well-foundedness, etc. In comparison, the received wisdom regarding first order logic is that it is very weak in expressing exactly mathematical concepts such as the ones mentioned: finiteness, countability, and well-foundedness. According to this wisdom there is a huge difference between first and second order logic: first order logic very weak, second order logic very strong.

[24]For example, $\exists X \forall x X(x)$ and $\exists X \forall x \neg X(x)$ would not be valid.

We argue that this difference is visible in its full strength only if we look at it from the perspective of a background formal language $\mathcal{F}$, in which we can formalize both first and second order logic and compare these "object languages" to each other. Then we can see the perceived huge difference between the expressive powers of the two. But we should not forget to ask, what is the metalanguage of $\mathcal{F}$? An "infinite regress" problem arises. If the metalanguage is first order set theory we are essentially using first order logic to define the meaning of second order logic. Then which is stronger? Is it not first order logic? This seems like a conundrum.

For example, if we say that second order logic can express wellfoundedness, we are saying that there is a sentence $\phi(E) \in L^2$ such that for all models (M, E), $E \subseteq M \times M$,

$$(M, E) \text{ is well-founded} \iff (M, E) \models \phi(E). \tag{4}$$

Here the left-hand side is thought as being understood from the "outside", in the metatheory, whatever that means.

Let us analyse the situation. The equivalence (4) informs us, or even defines, the meaning of $\phi(E)$. But what is the meaning of the equivalence (4) *itself*? In particular, what is the meaning of the left hand side of (4)? What criterion are we using to judge whether (M, E) is wellfounded or not on the left side of (4)? If ϕ is the *usual* second order sentence saying that the binary relation E is wellfounded, we can use the same sentence in the left side of (4), except that then (4) becomes a tautology i.e. it says nothing.

Probably most people would say that we should use the *set-theoretical* definition of wellfoundedness in the left side of (4). That is, we should use first order set theory as the ultimate arbiter that decides the meaning of second order sentences. But how is the meaning of *set theory* on the left hand side of (4), i.e. of "(M,E) is well-founded" defined? This is a set-theoretical statement and, barring reference to metatheory, the best description of its meaning is derived from the axioms of set theory.

Why not take "(M,E) is well-founded in V" as the criterion of truth of "(M,E) is well-founded"? This is what we do intuitively, but to make the intuition exact we resort to axioms. As to truth in V we say that *at least* what we can derive from the axioms we accept as true in V. What we cannot yet derive from the axioms, remains unsettled so far, and will perhaps be resolved in future research that yields new generally accepted axioms.

Curiously, it is the same with the right hand side of (4). We can derive the meaning of "$(M, E) \models \phi(E)$" from the axioms of second order logic. Again, the axioms are incomplete but maybe the future will bring new axioms. It would be rather surprising if such new axioms arose from somewhere other than set theory but in principle it is possible. But if we use the axioms of second order logic to define the meaning of second order logic, we are really talking about second order logic as a two-sorted first order logic or, in other words, second order logic with Henkin semantics, rather than second order logic with full semantics.

One often takes first order set theory as a foundation for mathematics. To say that there is something more primitive begs the question of a foundation. Since e.g. first order set theory, viewed as a foundation, does not have its own metalanguage, we cannot have a truth definition in the usual sense for this language, as truth definitions are given in the metalanguage. Instead, we formalize the language, define in the language the concept of a model, and then give an inductive truth definition for an arbitrary formalized sentence in a given model. This works perfectly and is very useful even though these models for which we know what truth means are not the "reality" that our logic (e.g. first order set theory) talks about. Fortunately these models are able to capture large chunks of "reality", as in the V_α's or the $H(\kappa)$'s. Of course, "truth in V" is not the same as truth in some V_α. To reach truth in V one would have to capture truth in V_α for a club of ordinals α but "for a club of α" is not an expression of first order set theory, so it goes beyond our metatheory and calls for a metametatheory which we have had reasons to abandon.

4 Conclusion

First order logic alone is expressively weak, but when it is combined with $\in$, yielding first order set theory, it is suddenly the strongest logic. Something magical happens when $\in$ appears! We argued in this paper that this resembles the case of second order logic, which arises when "$x \in Z$" for first order variables x and second order set-variables Z are incorporated into first order logic. One might ask, if second order logic is thought of as a "logic" shouldn't we think of first order set theory as a logic, too? If we do, then it is a very high order logic—in fact it is the *strongest* logic. There is a way to view second order logic as merely a two-sorted first order logic with the so-called Henkin semantics. While this does not quite make second order logic the same as first order logic, it comes close. We can similarly consider the first order language of set theory to be a many sorted logic ("sort logic" [38]) with a variant of Henkin semantics. The Henkin models of set theory are simply the transitive models that are commonly studied in axiomatic set theory. Thus the adoption of Henkin semantics, which is the *modus operandi* of set theory anyway, brings the a priori incredibly strong first order logic with $\in$ into the same family as the received famously weak first order logic.

In our end is our beginning.

5 Appendix

In this appendix we will prove that second order logic is symbiotic with the power set operation. The proof is adapted from [35], where the concept of symbiosis is presented in full generality, for absolute as well as non-absolute logics.

First we need the concept of relativised reduct. To this end, let $\mathcal{M}$ be a model with vocabulary L and $L_1 \cup \{P\} \subseteq L$, where P is unary. Then the *relativised reduct* $\mathcal{M}^P \restriction L_1$ of $\mathcal{M}$ to the vocabulary L_1 and to the predicate P is the model $\mathcal{M}_1$ with $P^{\mathcal{M}}$ as the domain and the following structure:

- $R^{\mathcal{M}_1} = R^{\mathcal{M}} \cap (P^{\mathcal{M}})^n$ if $R \in L_1$ is of arity n.
- $f^{\mathcal{M}_1} = f^{\mathcal{M}} \restriction (P^{\mathcal{M}})^n$ if $f \in L_1$ is of arity n.
- $c^{\mathcal{M}_1} = c^{\mathcal{M}}$ if $c \in L_1$.

The scalar field of a vector space E is an example of a relativised reduct, as is the additive group of the vectors of E.

We now extend the above definition of relativised reduct of an individual model to the relativised reduct of a *class* of models as follows: If K is a model class with vocabulary L, and $L_1 \cup \{P\} \subseteq L$, then $K^P \restriction L_1$ is the class of relativised reducts $\mathcal{M}^P \restriction L_1$ of models $\mathcal{M} \in K$.

Thus we can obtain from, e.g., a class of vector spaces, their relativised reduct: a class of (scalar) fields, a class of (additive) groups, and a class of multiplicative groups.

We can now define our principal auxiliary technical notion, that of a Δ-extension:

Definition 7 ([5])**.** *The Δ-extension $\Delta(\mathcal{L}^*)$ of a logic $\mathcal{L}^*$ is the logic the definable model classes of which are such K that both K and the complement of K are relativised reducts of $\mathcal{L}^*$-definable model classes.*

$\Delta(\mathcal{L}^*)$ is equivalent to $\mathcal{L}^*$ for some logics, e.g. for first order logic, for L_{HYP}, and for $L_{\omega_1\omega}$. However often there are properties of models that the logic $\mathcal{L}^*$ cannot express, but for trivial reasons, while they are expressible in $\Delta(\mathcal{L}^*)$. This anomaly is repaired by taking the Δ-extension. As an example, recall that the logic $L(Q_0)$ denotes first order logic with the generalized quantifier “there are infinitely many,” and L^2_w denotes weak second order logic with second order quantifiers for second order variables that range over finite subsets and relations. Then $\Delta(L(Q_0)) \equiv \Delta(L^2_w)$, i.e. the logics $\Delta(L(Q_0))$ and $\Delta(L^2_w)$ have the same definable model classes [5]. However the two logics are not equivalent, as $L(Q_0)$ is strictly stronger than L^2_w. See [26] for properties of this Δ-operation.

The Δ-operation is a closure operation in the sense that for all logics $\mathcal{L}^*$ [26]:

1. $\mathcal{L}^* \leq \Delta(\mathcal{L}^*)$. (Increasing)
2. $\mathcal{L}^* \leq \mathcal{L}^{**}$ implies $\Delta(\mathcal{L}^*) \leq \Delta(\mathcal{L}^{**})$. (Monotone)
3. $\Delta(\Delta(\mathcal{L}^*)) \equiv \Delta(\mathcal{L}^*)$. (Idempotent)

Moreover, the Δ-operation preserves the essential properties of a logic [26]:

1. If every sentence of $\mathcal{L}^*$ which has a model has also a countable model, then the same is true of sentence of $\Delta(\mathcal{L}^*)$. (Löwenheim-Skolem Property)
2. If every set of sentences of $\mathcal{L}^*$, every finite subset of which has a model, itself has a model, then the same is true of $\Delta(\mathcal{L}^*)$. (Compactness Property)
3. If there is a recursive complete axiomatization for (validity in the logic) $\mathcal{L}^*$, then the same is true of $\Delta(\mathcal{L}^*)$. (Completeness Property)

This is the reason why $\Delta(\mathcal{L}^*)$ is considered a “minor” modification of $\mathcal{L}^*$.

We shall now establish symbiosis in the important special case of second order logic. Let Pw be the relation $\mathrm{Pw}(x, y) \leftrightarrow y = \mathcal{P}(x)$. Let $\mathcal{L}^2$ be second-order logic.

Theorem 8. *The relation Pw and the logic $\mathcal{L}^2$ are symbiotic.*

Proof. We need to prove the following

1. Every $\mathcal{L}^2$-definable model class is $\Delta_1(\mathrm{Pw})$-definable.
2. Every $\Delta_1(\mathrm{Pw})$-definable model class is $\Delta(\mathcal{L}^2)$-definable.

To statement 1, suppose $\mathcal{K} = Mod(\varphi)$, for some $\mathcal{L}^2$-sentence φ. Then $\mathcal{A} \in \mathcal{K}$ if and only if $\langle V_\alpha, \in, \mathcal{A} \rangle \models$ "$\mathcal{A} \models \varphi$", for some (any) α strictly greater than the smallest β such that $\mathcal{A} \in V_\beta$. Since the predicate $x = V_\alpha$ is $\Delta_1(\mathrm{Pw})$, $\mathcal{K}$ is $\Delta_1(\mathrm{Pw})$-definable.

To statement 2, suppose $\mathcal{K}$ is a $\Delta_1(\mathrm{Pw})$-definable model class that is closed under isomorphisms. Suppose the vocabulary of $\mathcal{K}$ is L_0 which we assume for simplicity to have just one sort s_1 and one binary predicate P. Let $\Phi(x)$ be a $\Sigma_1(\mathrm{Pw})$-formula of set theory such that $\mathcal{A} \in \mathcal{K}$ if and only if $\Phi(\mathcal{A})$. Let s_0 be a new sort, E a new binary predicate symbol of sort s_0, F a new function symbol from sort s_1 to sort s_0, and c a new constant symbol of sort s_0. Consider the class $\mathcal{K}_1$ of models

$$\mathcal{N} = (N, B, E^{\mathcal{N}}, c^{\mathcal{N}}, F^{\mathcal{N}}, P^{\mathcal{N}}),$$

where N is the universe of sort s_0 and B the universe of sort s_1, that satisfy the sentence φ given by the conjunction of the following sentences of second order logic:

(i) The conjunction θ of a finite part of the first-order axiom system ZFC^- (ZFC axioms except the power-set axiom) written in the vocabulary $\{E\}$ (instead of $\in$) in sort s_0. This finite part of ZFC^- is chosen so that everything that is needed is included.

(ii) The first-order sentence $\Phi(c)$ written in the vocabulary $\{E\}$ in sort s_0.

(iii) The second-order sentence which says that E is well-founded.

(iv) The following second-order sentence: $\forall x \forall W(\forall y(W(y) \to yEx) \to \exists z \forall y(W(y) \leftrightarrow yEz))$.

(v) A first-order sentence saying that c is a pair (a, b), where $b \subseteq a \times a$, all written in the vocabulary E in sort s_0.

(vi) A first-order sentence saying that F is an isomorphism between the s_1-part $(B, P^{\mathcal{N}})$ of the model and the structure $(\{x \in N : xE^{\mathcal{N}}a\}, \{(x, y) \in N^2 : (x, y)E^{\mathcal{N}}b\})$.

Claim: $\mathcal{A} \in \mathcal{K}$ if and only if $\mathcal{A} = \mathcal{N} \restriction \{s_1, P\}$, for some $\mathcal{N} \in \mathcal{K}_1$.

Suppose first $\mathcal{A} \in \mathcal{K}$. Note that Pw is Σ_2-definable. Let n be such that θ is a Σ_n-sentence. By the Reflection Principle ([15, Theorem 7.4]), let $V_\alpha \preceq_{n+2} V$, with $\mathcal{A} \in V_\alpha$. Then, $\mathcal{A} = \mathcal{N} \restriction \{s_1, P\}$, where

$$\mathcal{N} := (V_\alpha, A, \in, \mathcal{A}, id, P^{\mathcal{A}}) \in \mathcal{K}_1.$$

Conversely, suppose $\mathcal{N} := (N, B, E^{\mathcal{N}}, c^{\mathcal{N}}, F^{\mathcal{N}}, P^{\mathcal{N}}) \in \mathcal{K}_1$ with $\mathcal{A} = \mathcal{N} \upharpoonright \{s_1, P\}$. Clearly, the structure $(N, E^{\mathcal{N}})$ is extensional and well-founded (by condition (iii)). Moreover, $(N, E^{\mathcal{N}}) \models \Phi(c^{\mathcal{N}})$. Since $\mathcal{K}$ is closed under isomorphisms, we may assume, w.l.o.g., that N is a transitive set and $E^{\mathcal{N}} = \in \upharpoonright N$ (We are appealing here to the so-called Mostowski's Collapsing Lemma [15, page 65]). Now, Pw is absolute for N: for every a, b in N, we know that $\mathrm{Pw}(a, b)$ iff $N \models \mathrm{Pw}(a, b)$. Since $(N, \in) \models \Phi(c^{\mathcal{N}})$ and N is transitive, and since Φ is $\Sigma_1(\mathrm{Pw})$, we have that $\Phi(c^{\mathcal{N}})$ is true, i.e., it holds in V. Thanks to condition (vi), $c^{\mathcal{N}}$ is a binary structure isomorphic to $\mathcal{A}$. Since $\mathcal{K}$ is closed under isomorphism, $\mathcal{A} \in \mathcal{K}$. Thus $\mathcal{K}$ is a projection of the $\mathcal{L}^2$-definable model class $\mathcal{K}_1$. The Claim is proved.

We can do the same for the complement of $\mathcal{K}$. Hence $\mathcal{K}$ is $\Delta(\mathcal{L}^2)$-definable. Statement 2 is proved. □

References

[1] Jyrki Akkanen. Absolute logics. *Ann. Acad. Sci. Fenn. Ser. A I Math. Dissertationes*, (100):83, 1995.

[2] Joan Bagaria and Jouko Väänänen. On the symbiosis between model-theoretic and set-theoretic properties of large cardinals. *J. Symb. Log.*, 81(2):584–604, 2016.

[3] Jon Barwise. Infinitary logic and admissible sets. *J. Symbolic Logic*, 34(2):226–252, 1969.

[4] Jon Barwise. Absolute logics and $L_{\infty\omega}$. *Ann. Math. Logic*, 4:309–340, 1972.

[5] Jon Barwise. Axioms for abstract model theory. *Ann. Math. Logic*, 7:221–265, 1974.

[6] Jon Barwise. Axioms for abstract model theory. *Ann. Math. Logic*, 7:221–265, 1974.

[7] Jon Barwise, Matt Kaufmann, and Michael Makkai. Stationary logic. *Ann. Math. Logic*, 13(2):171–224, 1978.

[8] John P. Burgess. Descriptive set theory and infinitary languages. *Zb. Rad. Mat. Inst. Beograd. (N.S.)*, 2(10):9–30, 1977.

[9] William Ewald. The Emergence of First-Order Logic. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2019 edition, 2019.

[10] S. Feferman and G. Kreisel. Persistent and invariant formulas relative to theories of higher order. *Bull. Amer. Math. Soc.*, 72:480–485, 1966.

[11] Kurt Gödel. *Collected works. Vol. III.* The Clarendon Press, Oxford University Press, New York, 1995. Unpublished essays and lectures, With a preface by Solomon Feferman, Edited by Feferman, John W. Dawson, Jr., Warren Goldfarb, Charles Parsons and Robert M. Solovay.

[12] L. Henkin. Some remarks on infinitely long formulas. In *Infinitistic Methods (Proc. Sympos. Foundations of Math., Warsaw, 1959)*, pages 167–183. Pergamon, Oxford; Państwowe Wydawnictwo Naukowe, Warsaw, 1961.

[13] J. Herbrand. Recherches sur la théorie de la démonstration. Travaux de la Société des Sciences et des Lettres de Varsovie. Nr. 33, 128 p (1930)., 1930.

[14] Wilfrid Hodges. Tarski's Theory of Definition. In *New essays on Tarski and philosophy*, pages 94–132. Oxford Univ. Press, Oxford, 2008.

[15] Thomas Jech. *Set theory.* Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2003. The third millennium edition, revised and expanded.

[16] Carol R. Karp. *Languages with expressions of infinite length.* North-Holland Publishing Co., Amsterdam, 1964.

[17] H. Jerome Keisler. *Model theory for infinitary logic. Logic with countable conjunctions and finite quantifiers.* Studies in Logic and the Foundations of Mathematics, Vol. 62. North-Holland Publishing Co., Amsterdam-London, 1971.

[18] Juliette Kennedy. On the Logic without Borders Point of View. In *Logic without Borders: Essays on Set Theory, Model Theory, Philosophical Logic and Philosophy of Mathematics*, pages 1–14. de Gruyter, 2015.

[19] Juliette Kennedy. *Gödel, Tarski and the lure of natural language—logical entanglement, formalism freeness.* Cambridge University Press, Cambridge, 2021.

[20] Juliette Kennedy, Menachem Magidor, and Jouko Väänänen. Inner models from extended logics: Part 1. *J. Math. Log.*, 21(2):Paper No. 2150012, 53, 2021.

[21] Per Lindström. First order predicate logic with generalized quantifiers. *Theoria*, 32:186–195, 1966.

[22] Per Lindström. On extensions of elementary logic. *Theoria*, 35:1–11, 1969.

[23] M. Magidor. On the role of supercompact and extendible cardinals in logic. *Israel J. Math.*, 10:147–157, 1971.

[24] Menachem Magidor and Jerome Malitz. Compact extensions of $L(Q)$. Ia. *Ann. Math. Logic*, 11(2):217–261, 1977.

[25] Menachem Magidor and Jouko Väänänen. On Löwenheim-Skolem-Tarski numbers for extensions of first order logic. *J. Math. Log.*, 11(1):87–113, 2011.

[26] J. A. Makowsky, Saharon Shelah, and Jonathan Stavi. Δ-logics and generalized quantifiers. *Ann. Math. Logic*, 10(2):155–192, 1976.

[27] Yiannis N. Moschovakis. The game quantifier. *Proc. Amer. Math. Soc.*, 31:245–250, 1972.

[28] A. Mostowski. On a generalization of quantifiers. *Fund. Math.*, 44:12–36, 1957.

[29] W. V. Quine. *Philosophy of logic.* Harvard University Press, Cambridge, MA, second edition, 1986.

[30] S. Shelah. *Classification theory and the number of nonisomorphic models*, volume 92 of *Studies in Logic and the Foundations of Mathematics.* North-Holland Publishing Co., Amsterdam, second edition, 1990.

[31] Saharon Shelah. Generalized quantifiers and compact logic. *Trans. Amer. Math. Soc.*, 204:342–364, 1975.

[32] Alfred Tarski. *A Decision Method for Elementary Algebra and Geometry.* The Rand Corporation, Santa Monica, Calif., 1948.

[33] Alfred Tarski. Some notions and methods on the borderline of algebra and metamathematics. In *Proceedings of the International Congress of*

Mathematicians, Cambridge, Mass., 1950, vol. 1, pages 705–720. Amer. Math. Soc., Providence, R.I., 1952.

[34] Leslie H. Tharp. The characterization of monadic logic. *J. Symbolic Logic*, 38:481–488, 1973.

[35] Jouko Väänänen. Abstract logic and set theory. I. Definability. In *Logic Colloquium '78 (Mons, 1978)*, volume 97 of *Studies in Logic and the Foundations of Mathematics*, pages 391–421. North-Holland, Amsterdam-New York, 1979.

[36] Jouko Väänänen. Boolean-valued models and generalized quantifiers. *Ann. Math. Logic*, 18(3):193–225, 1980.

[37] Jouko Väänänen. Second-order logic and foundations of mathematics. *Bull. Symbolic Logic*, 7(4):504–520, 2001.

[38] Jouko Väänänen. Sort logic and foundations of mathematics. In *Infinity and truth*, volume 25 of *Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.*, pages 171–186. World Sci. Publ., Hackensack, NJ, 2014.

[39] Jouko Väänänen. An extension of a theorem of Zermelo. *Bull. Symb. Log.*, 25(2):208–212, 2019.

[40] Jouko Väänänen. Tracing internal categoricity. *Theoria*, 87(4):986–1000, 2021.