THE REVIEW OF SYMBOLIC LOGIC
Volume 10, Number 1, March 2017

# LOGIC IN THE *TRACTATUS*

MAX WEISS

**Abstract.** I present a reconstruction of the logical system of the *Tractatus*, which differs from classical logic in two ways. It includes an account of Wittgenstein's "form-series" device, which suffices to express some effectively generated countably infinite disjunctions. And its attendant notion of structure is relativized to the fixed underlying universe of what is named.

There follow three results. First, the class of concepts definable in the system is closed under finitary induction. Second, if the universe of objects is countably infinite, then the property of being a tautology is $\Pi^1_1$-complete. But third, it is only granted the assumption of countability that the class of tautologies is $\Sigma_1$-definable in set theory.

Wittgenstein famously urges that logical relationships must show themselves in the structure of signs. He also urges that the size of the universe cannot be prejudged. The results of this paper indicate that there is no single way in which logical relationships could be held to make themselves manifest in signs, which does not prejudge the number of objects.

We have by now a quite systematic and rigorous grasp of the logical work of two of Wittgenstein's teachers, Frege and Russell. This is thanks in part to decades of flourishing scholarship, and thanks also to Frege and Russell themselves. In contrast, despite comparably voluminous commentary there is still no received understanding of anything describable as the logical system of the *Tractatus* (Wittgenstein, 1921). It is hard to resist the conclusion that the *Tractatus* did not, despite its professed program and its large reputation, offer any systematic alternative conception of the nature of logic.

But the conclusion is mistaken. To the contrary, there is a system, or a class of similar systems, which can be understood to explicate the development of logic in the *Tractatus*. They differ rather sharply from those of Frege or Russell, as well as from classical first- or second-order logic. Nonetheless they can be exactly described and investigated metamathematically, for epistemological and metaphysical evaluation. In this paper, I will present one such system, and investigate some of its properties. These turn out to be of surprising and indeed independent interest.

In seeking to understand what logic is supposed to be according to the early Wittgenstein, we may distinguish two kinds of evidence. First, there are the contours of his own construction, famously, for example, in the truth-functionality thesis and its enactment through iterated joint denial. Second, there are in the *Tractatus* apparent declarations of epistemological constraints on the nature of logic: some of these, for example, have been taken to suggest that according to Wittgenstein, logic must be decidable.

I wish to separate these two strands. In the early decades of the 20th century, it was entirely possible for a proficient researcher to develop a computationally intractable system

Received: July 26, 2014.
2010 *Mathematics Subject Classification*: 03A05, 03C80, 03C70, 03D70, 01A60, 03E47, 03D80.
*Key words and phrases*: logic, induction, consequence, Wittgenstein, Frege, Russell, Tarski.


doi:10.1017/S1755020316000472

under the misapprehension of its decidability. Of course, it might be worthwhile to investigate systems which now can be seen to resonate with purported epistemological declarations of the *Tractatus*. The question still remains: what system, if any, did Wittgenstein in fact describe? The undecidability of logic being so deeply rooted in our own understanding, it is hard to take up in imagination the computational intuitions of researchers in the era which preceded its discovery. But in seeking to understand how someone tried to climb a mountain, we should study the climber's movements and the mountain's contours, not transpose the climber to a molehill.

A large body of literature exhibits accelerating progress in our understanding of the development of logic in the *Tractatus*. Thanks especially to Geach (1981) and Soames (1983) and more recently Wehmeier (2004), (2008), (2009), and Rogers and Wehmeier (2012), it has become nearly received wisdom that Wittgenstein both intended and managed, if only haphazardly, to accommodate the expressive resources of first-order logic with equality.

However, when the *Tractatus* was in preparation, first-order logic had not attracted much attention as an autonomous logical system. So it would be surprising to say the least to find that something like first-order logic, as it might be now understood, is the logic of the *Tractatus*. I want to explore two respects in which Wittgenstein's logic differs. They can each be understood to characterize a conception of logic which is a kind of intermediate between the logicisms of Frege and Russell on the one hand, and what became classical logic on the other.

First, Wittgenstein's conception of logic differs from what became classical logic in regard to what is now understood as interpretation. Although the exact height of metalogical perspective reached by Frege or Russell has been a matter of some scholarly controversy, it seems nonetheless safe to deny that anything like the modern notion of first-order satisfiability or logical consequence plays a central role in their logical writings. Rather, for Frege and Russell, the notion of axiomatic derivability sets for logic a basic standard of rational organization, and thence also of logical correctness. But Wittgenstein denied such a role for proof. So it remains to be said just how the signs of a formalism could be subject to some standard of logical correctness by the lights of the *Tractatus*. Here, Wittgenstein propounds truth-conditional analyses of the notions of logical validity and of logical consequence. These analyses resemble the now classical reduction of consequence to the notions of truth-in-a-structure and class-of-all-structures. But while classically, a structure for a language may have as its domain any nonempty set, for Wittgenstein all structures relevant to the determination of validity have an underlying domain in common, the universal collection of all objects. It has been both claimed and denied, in commentary on Tarski (1936)—see Mancosu (2010) for a review—that this feature disrupts certain rudiments of model theory. But as we'll see, Wittgenstein's conception is yet further from the classical one, since like Russell he holds that each object has its own proper representative in a proposition.

Second, Frege and Russell understood logic to include what we recognize as non-first-order resources. Frege's 1879 *Begriffsschrift* culminates with what, by Frege's lights, would be a purely logical analysis of the concept of the ancestral. Likewise, Russell's introduction of the axiom of reducibility was motivated in part by the desire to reduce to logic the principle of mathematical induction. First-order logic not having been isolated as an autonomous logical system, it is natural to suppose that Wittgenstein inherited the expectation, regarding such non-first-order notions, that they were nonetheless logical. As Geach (1981), Potter (2009), and Ricketts (2012) among others have observed, he introduced a notion of "form-series" variable, which permits the expression, in finite space, of some countably infinite disjunctions. However, the literature contains no attempt at an exact reconstruction of the device. For example, although it's agreed that the disjuncts

https://doi.org/10.1017/S1755020316000472 Published online by Cambridge University Press

must be generated in some vaguely "effective" manner, it's not at all clear what this is supposed to mean. Moreover, the logical significance of the form-series device depends on the intelligibility of quantifying into the contexts it creates, but this in turn involves subtleties hitherto unaddressed in the secondary literature.

I will construct a notion of formal series exactly. As an approximation from below, my assumption will be that it lies well within the system of constructions Wittgenstein did intend to admit. But then I will show that adding to first-order logic this approximation of the form-series device yields an expression of every concept definable by finitary induction. I'll also give some interpretive evidence to suggest that this was the point.

So for present purposes, Wittgenstein's two main divergences from classical logic are these. The domain of every structure is one and the same collection of uniquely named objects. And, basic logical resources include the notion of iterated formal procedure. By making these divergences explicit, it becomes possible to give a mathematically definite characterization of the complexity of Wittgenstein's logical system. Let me quickly summarize the results to be established here; the technical notions are standard but will be defined in later sections. It has been widely recognized that Wittgenstein could respect a commitment to decidability of logic by presupposing that the domain of quantification is finite. This already implies that the complexity of logic depends on the cardinality of the domain; in §4 I'll evaluate the dependence precisely. I will show that if the domain is countably infinite, then the property of being a tautology in the logical system of the *Tractatus* is $\Pi^1_1$-complete in a suitable analogue of the analytical hierarchy. But moreover, it is only granted the assumption of countability that the concept of tautology is $\Sigma_1$ in the Levy hierarchy of formulas of set theory. Since, in any case, the notions of countable- and of arbitrary-domain tautology are $\Pi^1_1$ and $\Pi^{ZF}_1$ *a priori*, these results are just about as strong as possible. When the form-series device is dropped from the system, then the notion of tautology returns to its familiar $\Sigma_1$ position in the arithmetical hierarchy; however, the underlying notion of consequence remains just as intractable.

There are several further respects in which logic as elaborated in the *Tractatus* may seem to differ from classical first-order logic. First, Wittgenstein takes as primitive not ordinary connectives like negation, disjunction or existential generalization, but rather a truth-functional operator $N$ together with a variety of means of specifying the multiplicities of formulas to which it may be applied. Second, through a nonstandard interpretation of the objectual variables, Wittgenstein tries to eliminate the equality predicate. Third, the role of higher logical types in the *Tractatus* is a matter of some controversy. The first two of these features do not essentially alter the complexity-theoretic situation, but it considerably expedites the mathematics to abstract from them. On the other hand, the introduction of higher-order variables can only raise the complexity of the system. My primary aim is to establish lower bounds, so I could simply duck the controversy by introducing no higher types. Nonetheless, the complexity of, for example, the class of valid second-order formulas is far above the lower bounds established here (Väänänen, 2001). So, I argue on textual and systematic grounds that higher-order quantification does not figure centrally in Wittgenstein's account of the expressive resources of logic, and that impredicative quantification cannot figure at all. Thus, my contention will not just be that such-and-such are lower bounds, but moreover that significant strengthenings do not warrant comparable credence.

A project to understand the purported nature of logic in the *Tractatus* encounters two kinds of difficulty. First, there are the well-known exegetical difficulties raised by the text itself. It is highly compressed, with its use of logical notation somehow both telegraphic and inconsistent. But a little reflection reveals technical subleties too, even within relatively uncontroversial features of the program. Of particular concern here will be the problem of implementing within quantificational logic Wittgenstein's proposal to compress some infinitary disjunctions into finite expressions. So in this paper, I will confine exegetical discussion to §1, with the aim there to justify interpretive hypotheses which underwrite the subsequent complexity-theoretic analysis. The point of the remainder of the paper is to give a mathematical explication of the hypotheses and then to investigate their consequences. I freely use notation and techniques which did not mature until after Wittgenstein's death.

In outline, the rest of the paper runs as follows. §1 lays out some interpretive background. After briefly sketching the outlines of the *Tractatus* system, I introduce the two departures from classical logic under investigation, summarizing relevant literature and briefly sketching the importance of these eccentricities for the philosophical project of the *Tractatus*. §2 opens the mathematical developments. I begin by explicating Wittgenstein's notion of a structure as truth-possibility for elementary propositions, and show that it yields an embedding of the concept of truth in the concept of consequence. I then introduce an approximation of the form-series device, and propose an analysis of quantification into form-series contexts. Thanks to this analysis, it becomes possible to establish a foundational result, that the addition of the form-series device to first-order logic does preserve a decent form of extensionality. In §3, I turn to a question of the power of the form-series device, and show that as explicated here, it yields the effect of adding to first-order logic an operator for expressing finitary inductive definitions. At this point, the stage is set for the investigation of complexity. In §4 I present a detailed characterization of the complexity-theoretic effects of Wittgenstein's conception of structure and of his introduction of the form-series device, both separately and jointly: §4.1, §4.2, and §4.3 respectively treat the cases in which the underlying domain is supposed to be finite, or some fixed infinite cardinality, or considered in advance of any cardinality constraints at all. After sketching in §5 what I conjecture to be some philosophical significance for these results, the paper concludes in §6 with suggestions for future work.

**§1. Interpretive background.** In broad outline, Wittgenstein's conception of logical structure in the *Tractatus* is straightforward. The upshot is a collection of sentences, together with a relation on the collection which might be called "direct denial". The direct denial relation distinguishes some sentences as atomic, namely those which directly deny no sentences. A nonatomic sentence is to be true if and only if each of the sentences it directly denies is not true.

The relation of direct denial is supposed to secure that the truth or falsehood of each sentence be determined by the truth or falsehood of the sentences which are atomic. To this end, Wittgenstein prescribes that the collection of sentences and the relation of direct denial must together satisfy a certain structural condition, a so-called "general propositional form". Accordingly, atomic sentences may be regarded as having been given initially, and any other sentence must be presented as a joint denial of some sentences presented before it. As Geach (1981, 170) pointed out, the structural condition cannot be understood to require that anybody actually construct all sentences which precede a given sentence in the ordering, because in general, a sentence will have infinitely many antecedents. Rather, the condition purports to indicate when some finite manipulation of signs could secure for them a sense. The crux of the condition is that the relation of direct denial on the class of sentences be wellfounded.

It is clear that if the relation of direct denial exhibited circularities, then it might not be intelligible as realizing a logical relationship of denial at all. For example, a sentence which appears to deny itself could be so understood only by taking it both to be both true and to be not true, which for early Wittgenstein would not be an understanding. But although Russell claimed to locate the origin of paradoxes in circularity, the mere exclusion of circularity does not suffice to secure the coherent interpretability of direct denial as denial: for example, it would still admit the construction of a sequence of sentences $A_0, A_1, \ldots$ each of whose terms is the direct denial of all its successors (see Yablo (1993)). In practice, Russell enforced the acyclicity of logical dependence by a metaphor of bottom-up construction in the ramified hierarchy of types; in imposing the stronger condition of wellfoundedness, this hierarchy is a clear conceptual antecedent of Wittgenstein's general propositional form. In neither *Principia* nor the *Tractatus* does the notion of wellfoundedness occur explicitly; so far as I know the concept is first explicitly formulated in Mirimanoff (1917).

The existence of a wellfounded direct denial relation is supposed to give a condition on the construction of signs, according to which signs would be capable of expressing a sense. This means that wellfoundedness of direct denial should characterize the signs themselves, so that the relation can be determined from the mere signs without reference to their sense (cf. 3.33–3.331). To this end Wittgenstein prescribes that each nonatomic sentence have two parts. One part, which at 5.501 Wittgenstein calls a "bracket-expression", serves to present some possibly infinite multiplicity of other sentences. The other part is the famous operator $N$, which, attached to the bracket-expression, yields a sentence which directly denies precisely those sentences which the bracket-expression presents. Thus, to determine for any two sentences whether one directly denies the other, it should suffice to check whether the second is among the sentences presented by the bracket-expression of the first.

It is clear that the complexity of logical dependence exhibited by the system will depend on the methods available for constructing bracket-expressions. Wittgenstein describes three methods at T5.501. The first method is simply to make a list of sentences outright; the resulting expression presents the sentences listed. Using joint denial on the lists, this obviously yields an analysis of negation and disjunction: the negation of a sentence is a sentence which directly denies just it, and the disjunction of two sentences is the negation of a direct denial of just those two. The second method is intended to yield an analysis of generalization over objects, and ultimately together with the first method to recover the expressive power of first-order logic with equality, at least under a certain "Russellian" construal. The third method goes rather farther, and was billed by Wittgenstein to yield at least a means of expressing the ancestral, or transitive closure, of expressible relations. In the rest of this section I will briefly sketch a somewhat anachronistic and oversimplified account of these two further methods which will suffice for the main purpose of this paper, which, again, is primarily analytical.

***1.1. Objectual generality.*** Wittgenstein's account of quantification is rather sparse, mentioning only the $N$-operator, its application to finite lists, and its application to the collections of terms of bracket expressions constructed by a second method. According to T5.501, this second method consists in "giving a function *fx*, whose values for all values of $x$ are the propositions to be described." Suppose, then, that some bracket expression presents the values of *fx*. Then, the result of attaching the sign of joint-denial to the bracket expression amounts, in some sense, to the same as the ostensibly Principian formula $\sim \exists x.fx$. Wittgenstein seems to intend this mysterious "function" *fx* to be something like a propositional function in the sense of *Principia*. However, it would appear to be only in the 3.31s that there appears anything approaching an explanation: "if we turn a constituent of a proposition into a variable, there is a class of propositions which are values of the resulting variable proposition" (3.315). Very roughly speaking, the second method might now be summarized as follows. The result of turning a constituent of a proposition into a variable is supposed to determine a propositional function; then an existential generalization is to be analyzed as the negation of the direct denial of the totality of the function's values.

*1.1.1. The grammar of quantification.* The abstractness of this account has abetted some controversy. In particular, Fogelin (1976) argued that since Wittgenstein does not mention any device for indicating the scope of the generalization, therefore he cannot distinguish between, say, the negation of the joint denial of the values of a function, and the joint denial of their negations. However, Soames (1983, 583ff) responded that Wittgenstein does not in his brief remarks purport to analyze quantification by supplying a notational system with definite syntax. Rather, the proposal purports merely to schematize the construction of one truth-condition from others. Thus, Fogelin's argument can be taken to show, unsurprisingly, that any particular instance of Wittgenstein's constructional schema must include a device for the identification of those subformulas whose semantic role is to present the totality of values of a propositional function. Geach and Soames conclude the debate by observing that the schema for construction of truth-conditions is satisfied by a notation tricked out with *Principia*-like scope indicators. Such a notation can be understood as an "intended model" of the schematic description. Thus, for example, within the model one might distinguish elements $NNxfx$ and $NxNfx$ as respectively truth-conditionally equivalent to the disjunction and conjunction of the collection of values of $fx$. Henceforth, I will simply work in such an intended model. In fact, I will assume that the model includes an image of a classical syntax of quantificational logic, in a signature which contains at least a few individual constants and at least a few monadic and dyadic predicate letters. This assumption can actually be weakened in some philosophically interesting respects, but I will not explore the possibility here.

We can now similarly explicate the notion of propositional function. Recall that a procedure of turning a constituent of a proposition into a variable is supposed to determine a function, which given some argument, returns a proposition. Without reference to any particular model of syntax, it is not at all clear just what could be meant by "constituent" (*Bestandteil*), let alone by turning one into a variable. But having assumed a classical syntax, then propositions may be identified with closed formulas. The notion of constituent may then be taken to include at least the individual constants which occur in the formula.

Wittgenstein's construction by way of propositional variables and the $N$-operator establishes an interesting asymmetry between disjunctive and conjunctive truth-operations. The signs $\vee$ or $\exists$ of a disjunctive truth-operation can be disabbreviated trivially as $NN$. In contrast, conjunctive expressions can be disabbreviated only in the context of their occurrence with propositional variables. Fogelin has therefore correctly noted that in Wittgenstein's framework, the admissibility of conjunctive truth-operations entails nontrivial constraints on the class of propositional variables. Suppose, for example, that there were added to finitary truth-functional logic a one-off propositional variable whose range were the totality of elementary propositions. The resulting system could express the disjunction but not the conjunction of the totality of elementary propositions. Now about the more interesting system developed in §2, it does hold that conjunctions can be taken wherever disjunctions can, but this must be proven (see Proposition 3.3 of §3.1).

*1.1.2. Higher-order logic?* One might wonder if the notion of constituent might not best be understood to include other syntactically related items as well, for example complex expressions constructible from a proposition by abstraction. Michael Potter, for example, writes:

> Take some sign expressing a proposition and single out part of the sign. Now keep this part fixed and let the rest of the sign vary. All the propositions that can be obtained by varying the sign in this way form a class. The variable used to pick out this class is called a propositional function, but Wittgenstein invariably refers to it simply as a function. (Potter, 2009, 269)

Although Potter cites only 5.501 in the vicinity of this proposal, its strongest support may come from 3.315. According to the proposal, it is not a singled-out "part" which is initially allowed to vary, but rather "the rest" of a proposition. Potter does not actually say that the rest is a part at all, nor does he explicitly link his use of the term "part" directly to any usage in the text. But, since Wittgenstein only says that parts can be varied (or really: "turned into variables"), therefore Potter's gloss appears to presume that for any part of a proposition, the "rest" of the proposition besides it is also a part. However, in general, it is not the case that for any part of a given proposition, the rest of the proposition is a simple or elemental part. Thus, Potter would appear to understand the usage of *Bestandteil* to include parts or aspects of propositions which are other than simple parts.

Both Ogden-Ramsey (Wittgenstein, 1922) and Pears-McGuiness (Wittgenstein, 1961) translate *Bestandteil* not as "part" but as "constituent". As with *Bestandteil*, one might say that a thing which has constituents is somehow formed by them in a naturally privileged way. For example, molecules are constituted by atoms. In contrast, a mere part can be understood as a result of projection or abstraction, as with the northern hemisphere of the earth. So, one might say that something which has constituents is already "articulated" in terms of them, for they are simple or elemental in contrast to it; on the other hand a decomposition into mere parts or segments in general requires choosing one path rather than another. Moreover, one might say that something depends on its constituents, but its mere parts—like the top half of the earth—conversely depend on it (Fine, 2010, 586).

The heuristic metaphysics of Russell's ramified hierarchy of propositions provides a historically apposite example. Russell motivates the construction of the ramified hierarchy along an ordering of metaphysical dependence, if only as a convenient fiction. The ordering begins with the constitution of the universe by various objects and relations (Whitehead & Russell, 1913, 45); these objects and relations are the constituents of the earliest propositions (57ff). In contrast, a proposition can be seen to fall in the range of various propositional functions, these functions being obtained from the proposition by the abstraction, or "turning into variables" of constituents. Since propositional functions follow their values in the order of dependence, therefore propositional functions are not constituents of their values. This motivates the requirement in Russell that a propositional function cannot take itself as argument: the function depends on its values, the argument of the function which yields a value is a constituent of that value, and the value depends on its constituents.

There are a couple of reasons to suppose that at 3.315 and elsewhere, by *Bestandteil* Wittgenstein doesn't just mean part, but part which is simple or elemental. The first is circumstantial. Wittgenstein himself accepted the rendering by Ogden-Ramsey of *Bestandteil* as "constituent". As I've just sketched, a usage of the latter term was already established by Russell in the exposition of the ramified theory of types. And in claiming at 3.333 that "the sign for a function already contains the prototype of its argument", Wittgenstein borrows a plank of that theory.

The second reason is internal. Wittgenstein identifies what he refers to as *Bestandteile* with those expressions whose meaning is chosen, or determined by arbitrary convention. Thus, at 3.315 he says what might be turned into a variable are "those signs whose meaning is arbitrarily determined" (*jene Zeichen, deren Bedeutung willkürlich bestimmt wurde*), and indeed so determined "by arbitrary convention" (*willkürlich Übereinkunft*). In contrast, the meaning of expressions which are not *Bestandteile* is determined not by arbitrary convention but by their structure, given the choice of meanings for the *Bestandteile*. Thus, at 4.024 he says that to understand a sentence it suffices to understand its *Bestandteile*. And at 4.025 he remarks, heuristically, that in translation only the propositional—or perhaps here sentential—constituents (*Satzbestandteile*) are translated; presumably *Satzbestandteil* is supposed to pick out the sort of thing that is listed in a dictionary, and so is intended, without modification by any adjective like "simple", to evoke a contrast with expressions composed of several words. Now, at 4.025 Wittgenstein does mention *Bindewörter*—translated "conjunctions"—among the *Satzbestandteile* which are translated. But, this is because 4.025 really presents an heuristic analogy of the dependence on convention of propositions with the dependence on convention of unanalyzed sentences of English or German. Ultimately, there is only one essential respect in which the identity of a proposition must depend on a conventional choice of meanings, and this is the assignment of meanings to names. Thus, at 4.5, Wittgenstein proposes to give a general description of symbols, such that everything satisfying the description can express a sense given only a suitable choice of meanings of names. And this matters philosophically, because it's by boiling down the choice of meaning to the choice of meanings of names that Wittgenstein works out what's announced as his "fundamental thought", that logical constants are not representatives (4.023).

So, *Bestandteile* are the parts of a proposition whose meaning is chosen, or immediately fixed by convention. And the parts of a proposition whose meaning is chosen are precisely the names, i.e., the simple propositional constituents. But, at 3.315 Wittgenstein describes propositional functions as the results of turning into variables not mere parts, i.e., *Teile*, but *Bestandteile*, or constituents. Hence, the sort of thing which yields a propositional function upon being turned into a variable is a name.

Now, it's true that some kind of higher-type quantification appears to be mentioned twice in the *Tractatus* (3.333 and 5.5261). The two mentions can both be glossed with multisorted first-order quantification, which entails no increase in logical complexity (Henkin, 1950). Still, Wittgenstein does leave room for some kind of higher-order generality. For, at 5.501, Wittgenstein introduces his three ways of presenting propositional multiplicities without clearly suggesting that the ways are exhaustive. In particular, predicative higher-order generality might be introduced through some fourth or fifth method of presentation of propositional multiplicities.

However, the range of any such higher-order generality would be constrained by the requirement of wellfoundedness of logical evaluation. To see this, consider, for example, "Napoleon has all the qualities of a great general." This sentence can be seen as predicating something, say $\psi$, of Napoleon; let's call the sentence $\psi(n)$. Now, further suppose that $\psi(n)$ generalizes over everything predicable of Napoleon. Then, the truth-value of $\psi(n)$ would be determined as the conjunction of all sentences $f(\phi) \supset \varphi(n)$, for $\varphi$ anything predicable of Napoleon. So $\psi$ itself falls in the range of $\varphi$, and the determination of the truth-value of $\varphi(n)$ would depend on that very determination. In other words: the conception of generality as depending for its truth-value on those of its instances rules out higher-order quantification which is impredicative.

So, while Wittgenstein's remarks do leave room for some kind of "higher-order" quantification, this would have to be interpreted predicatively. Precisely what expressive power

Wittgenstein might've sought from such supplementations of the three stated methods of construction remains, to my knowledge, an open question in the literature.

In sum, a cursory examination of the text suggests that the notion of turning a constituent into a variable is reasonably interpreted on the syntactical model of propositions as replacing a constant term $a$ with a variable term $x$. Just as propositions become closed formulas, propositional functions are canonically explicated here as formulas containing at least one free variable.

*1.1.3. Identity.* This syntactical model of propositional functions suggests a natural account of their courses of values. We might identify the application of a propositional function with the instantiation of a formula. So the totality of values of a propositional function would become the class of closed instances of a formula, and an existential generalization becomes the disjunction of the elements of such a class. However, the reality is not quite so simple. As Hintikka (1956) and Wehmeier (*op. cit.*) have shown, Wittgenstein introduced a reinterpretation of the objectual variable, which helps to sustain a contention that the equality predicate is dispensable. This reinterpretation, which I'll call the "sharp" as opposed to the "natural" reading, amounts to requiring that a variable omit from its range those objects which are mentioned in its scope. Thus, the result of replacing a free variable with an appropriately sorted constant in a formula is a value of the function if and only if the constant does not already occur in the formula. Or equivalently, a proposition is a value of a propositional function if and only if that function is the result of turning some constant into a variable in the proposition.

This amendation yields a sharp divergence from classical first-order semantics. Hintikka and Wehmeier have shown, in a sense to be made precise, that every truth-condition of a formula of first-order logic with equality under the natural interpretation is the truth-condition of a formula of first-order logic without equality under the sharp interpretation. The translation of a naturally into a sharply read formula requires just two changes. Each existential quantification is disjoined with its omitted instances. And, each predication of equality is replaced with either a tautology or a contradiction containing the same constants and free variables, according as the two arguments of the predication, as linguistic expressions, are the same or distinct.

The Hintikka-Wehmeier result certainly helps to justify Wittgenstein's claim that the equality predicate is dispensable. But the justification is not obviously complete, for as Rogers and Wehmeier (2012, 547) point out, the result has a seemingly important qualification: the sharp reading affords an equality-free rephrasing of truth-conditions only when the class of all structures is restricted to those in which no distinct constants codenote. So, the translations eliminate the equality predicate only if there is no nontrivial distinction to be drawn between structures according as they do or do not assign the same denotation to a given term. In response to this apparent difficulty, Rogers and Wehmeier (2012, 546) cite Wittgenstein's remark (5.526) that "the world can be completely described by means of completely generalized propositions". But this remark could be used to show that there don't need to be any simple predications of equality only by being taken to show that there don't need to be any elementary propositions. It's more likely that at 5.526 means just that there are some completely generalized propositions such that their truth leaves no further question how the world is. It doesn't follow that generalities could be susceptible to truth and falsehood independently of their instances. For example, waxing psychologistic Wittgenstein says "the understanding of general propositions depends palpably on that of the elementary propositions" (4.411). It's unlikely that 5.526 is an offhand remark that the truth-functionality thesis is optional. Rather, the purported elimination of the equality predicate would appear to require some independent justification for the claim that simple predications of equality do not distinguish between structures.

Wittgenstein's position regarding the equality predicate cannot be just that the equality predicate is, like disjunction, not primitive, for its purported uses are not all accommodated. But nor does Wittgenstein hold that the conventions governing the sign underwrite no symbols at all; one might express by their means a proposition to the effect that there are at least two authors of *Principia Mathematica*. The position is rather that the conventions do not suit the logicosyntactic role of a predicate. Specifically, the basic function of an equality predicate would have to be, completed with some other occurrences of terms, to yield a proposition which is true iff those terms denote the same object. But, claims Wittgenstein, there is no such function: what would fulfill such a purported function must be either nonsensical or empty (5.5303).

I suggest that the situation should be understood like this. Wittgenstein derived his notion of proposition from early Russell, for whom at the end of analysis each object has only one representative, which is that very object. Then for anything at all, there is only one true proposition to the effect that it is the same as a given object, namely the proposition that the thing is the same as itself. And as Russell remarked: "when any term is given, the assertion of its identity with itself, though true, is perfectly futile" (Russell, 1903, 65). Famously, he proceeded in "On Denoting" (1905) to explain the value of apparent statements of identity on the ground that they aren't identities after all. The hypothesis that each object has one and only one representative reduces the truth or falsehood of simple predications of equality to their own mere identity as propositions, making it plausible that they draw no genuine distinctions between structures.

Now, the Russellian conception of propositions also makes it intelligible to suppose that "the specification of all true elementary propositions describes the world completely" (4.26), even if there are no elementary predications of equality. For example, if it is given that there are just two elementary propositions, that Lisa and Lucy are cats on the mat, then since the propositions themselves contain cats, the specification that both are true determines that there are two cats on the mat. If, on the other hand, the propositions instead contained proxies not in one-one correspondence with the objects, then knowing the truth-values of elementary propositions would leave open the question of how crowded the mat is. For Wittgenstein maintains that two objects might have all their properties in common (5.5302).

But as is well known, Wittgenstein did pull away from the Russellian conception. On the one hand he maintained that the proposition must by itself determine how things must stand if it is true. Furthermore, he supposed that this could be the case only if the way constituents stand to each other in the proposition were the same as the way that objects are thereby said to stand. Yet if the constituents of the proposition just are what is mentioned, it follows that every proposition is true. In the *Tractatus*, Russellian objectual propositional constituents must therefore give way to proxies.

Surely there is no initial plausibility to the suggestion that each object has one and only one representative, though things do not self-represent. But Wittgenstein fully imbibed the doctrine that the world is completely determined by the truth and falsehood of elementary propositions. So even as objects do not represent themselves, propositions retain a role in Wittgenstein's logical thinking which requires that each object has one and only one representative. Or as Russell himself put it: "there will be one word and no more for every simple object" (Russell, 1918, 198).

In summary, then, a simple predication of equality could distinguish between structures only if it has both possibilities of truth and falsehood. But it can't have both possibilities, if denotation essentially puts names and objects in one-one correspondence. Yet only if names and objects are in one-one correspondence could a proposition be a truth-function of elementary propositions. On pain of breaking the truth-functionality thesis, Wittgenstein must hold in the *Tractatus* that there is no room for the equality predicate to distinguish nontrivially between structures. In this way, the proviso on the Hintikka-Wehmeier result is discharged. In the context of Wittgenstein's other commitments, these authors have indeed justified his claim to have exhausted what could count as the genuine expressive contribution of the equality predicate.

*1.1.4. The Russellian constraint and the fixed-domain conception.* On the strength of the arguments of Geach, early Soames, and Wehmeier, we can conclude that Wittgenstein manages to accommodate the basic notions of first-order logic. But as we've seen, the accommodation induces four eccentricities. First, Wittgenstein rejects the equality predicate. Second, he recaptures some lost expressiveness by stipulating that a variable omits from its range what is mentioned in its scope. Third, he maintains that no further first-order expressiveness remains uncaptured, by stipulating that no two constants codenote. Finally, the analysis of quantified propositions as truth-functions requires that every element of the domain is denoted by a constant.

The point of this paper is to develop a complexity-theoretic analysis of Wittgenstein's logical system. From this point of view, the work of earlier interpreters licenses abstraction from the first two eccentricities. For the translations of Hintikka and Wehmeier establish that the consequence relations determined by the sharp and natural semantics of first-order formulas are mutually Turing-reducible. It is not straightforward to extend these translations to the outer reaches of Wittgenstein's system; but enough will be clear for present purposes. The second pair of eccentricities reflects Wittgenstein's conception of the relationship between names and the universe, which differs in two fundamental respects from the classical model-theoretic treatment of the relationship between constants and the domain of a structure. The argument that the equality predicate is dispensable requires that simple statements of equality draw no nontrivial distinctions in the class of all structures, and hence that there is no structure in which distinct constants codenote. The proposal to analyze an existentially generalized proposition as the disjunction of the values of a propositional function is extensionally adequate only to structures in which each element of the domain is denoted by a constant.

Together these requirements are equivalent to what I'll call a "Russellian" constraint on structures, that the denotation relation be a one-one correspondence between the class of names and the domain. The reason for the terminology is simply that the constraint is canonically satisfied by a structure such that every element of its domain is the one and only name of itself. Of course, the terminology isn't intended to convey any characterization of Russell himself. Rather, it's intended to mark Russell's deep effect on Wittgenstein.

We've seen that the Russellian constraint is well-rooted in the *Tractatus*, namely in the doctrine that a proposition is a truth-function of elementary propositions. To say that a proposition is a truth-function of some others is to say that, logically, its truth-value is a function of theirs, so that any maximal consistent choice of elementary propositions and their negations entails either the proposition or its negation. But this claim has counterexamples if distinct constants codenote, and has counterexamples if not all elements of a domain must be denoted by constants. Hence the truth-functionality thesis evidently presupposes the Russellian constraint.

It might be wondered whether the truth-functionality thesis somehow obfuscates the very notion of a quantifier. Doesn't it just turn a universal generalization into a "big

https://doi.org/10.1017/S1755020316000472 Published online by Cambridge University Press

conjunction"? However, the truth-functionality thesis follows from the Russellian constraint. And the Russellian constraint is only a restriction on the class of all structures. Hence, the truth-functionality thesis cannot by itself entail any substantive change in the theory of truth-conditions. It modifies only the universe of relata to which that theory might be applied, namely the universe of structures. So, the truth-functionality thesis does not affect the meaning of quantifiers, if the meaning of quantifiers depends only on their role in fixing truth-conditions.

On the other hand, the truth-functionality thesis does modify the notion of consequence, which Wittgenstein seems to give a proto- model-theoretic analysis (5.12). Say that an elementary truth-possibility is a maximal consistent set of elementary propositions and negations thereof, and that the truth-grounds of a proposition are the elementary truth-possibilities which entail it. Then a proposition is a consequence of some others iff all their common truth-grounds are truth-grounds of it. By the Russellian constraint, every truth-ground common to all instances of a universal generalization is a truth-ground of the generalization. So the generalization is a logical consequence of the class of its instances. Thus, the truth-functionality thesis does modify the meaning of quantifiers if their meaning is taken to depend essentially on their role in constituting the consequence relation. However, a defender of the Russellian constraint might respond that the role of subsentential expressions in constituting the consequence relation is exhausted by their contribution to truth-conditions. Thus, the charge that truth-functionality obfuscates the quantifiers depends on a controvertible philosophical presumption.

Wittgenstein's analysis of the consequence relation resembles that of Tarski (1936). The similarity is not just that both analyses generalize over something like logically possible distributions of truth and falsehood. For Tarski seems furthermore to have defined the consequence relation under a "fixed-domain" conception of structure, according to which they all have the same domain in common (Bays, 2001). The fixed-domain conception and the Russellian constraint are closely related, for they both require that the domains of all structures have the same cardinality. However, the fixed-domain conception of structure is not as strong. For it implies neither that every object be denoted by a constant, nor that no constants codenote.

The two conceptions differ significantly in their effects on the complexity of logic. As with Russellian consequence, the complexity of fixed-domain consequence depends on whether the domain is infinite. In case the domain is finite, the fixed-domain and Russellian notions pretty much coincide. But suppose that the domain $\mathcal{D}$ is infinite, and consider the corresponding notion of fixed-domain consequence for a countable signature. Let $I$ be the set of formulas to the effect that there are at least $n$ objects for each $n$. Since $\mathcal{D}$ is infinite, $A$ must be a fixed-domain consequence of $T$ if $A$ is a classical consequence of $T \cup I$. Conversely, suppose that some classical structure satisfies $T \cup I$ but not $A$. Since the signature is countable and $\mathcal{D}$ is infinite, the Lowenheim-Skolem theorems imply that some structure with domain $\mathcal{D}$ satisfies $T$ but not $A$. Therefore, if the signature is countable, then the fixed-domain consequences of $T$ are precisely the classical consequences of $T \cup I$. So fixed-domain consequence is recursively enumerable on any infinite domain. In contrast, we'll see in §2.2 and §4.3 that Russellian consequence permits categorical axiomatizations of rich countable structures. On no infinite domain is Russellian consequence even arithmetically definable, at least if the signature contains a dyadic predicate. Finally, it should be noted that a fixed-domain theorist can follow normal mathematical practice in taking uncountable structures to have a countable signature. So the fixed-domain theorist can generally avoid difficulties associated with uncountable signatures.

***1.2. Formal generality.*** We have now considered two of the three methods of presenting propositional multiplicities which Wittgenstein sketches at 5.501. I've indicated how they can together be understood to yield something like the expressive resources of first-order logic, but also argued that this understanding must inflect the underlying proto-model-theoretic analyses of validity and consequence. Let's now turn to the third method, which Wittgenstein describes as "giving a formal law" according to which the presented sentences are constructed. The sentences so presented constitute the terms of what he calls a "form-series".

> Series which are ordered by *internal* relations I call formal series. The series of numbers is ordered not by an external relation, but by an internal relation. Similarly the series of propositions $aRb$, $(\exists x) : aRx.xRb$, $(\exists x, y) : aRx.xRy.yRb$, etc. (4.1252)

> The general term of the formal series $a$, $O'a$, $O'O'a$, ... I write like so: "$[\![a, x, O'x]\!]$". This expression in brackets is a variable. The first term of the bracket expression is the beginning of a formal series, the second the form of an arbitrary term $x$ of the series, and the third the form of that term of the series which immediately follows $x$. (5.2522)

As Geach (1981, 171) observed, Wittgenstein here announces an intention, by means of a so-called formal series, to construct an expression of the ancestral. Frege's analysis of the ancestral required second-order quantification, indeed impredicatively. In contrast, Wittgenstein begins with the natural idea of constructing the countably infinite disjunction of all propositions to the effect that $b$ is connected to $a$ by $R$ through this or that number of steps. However, he then proposes a notation—the third kind of bracket-expression—by means of which a series of disjuncts would be presented in finite space. The ancestral would be expressed by the negation of the joint denial of the sentences which the bracket expression presents.

So, the form-series variable extends first-order logic to include the simulated presence of some countably infinite disjunctions. Clearly, not every countably infinite set of formulas will correspond to some such simulated disjunction, for the disjuncts must be generated by means of what Wittgenstein calls an "operation". In turn, an operation should return some sentence $B$ when applied to a sentence $A$ only in virtue of some "internal relation" which $B$ bears to $A$. Sundholm (1992, 61) points out that the apostrophe in Wittgenstein's notation $O'a$ apparently borrows from the description in *Principia* of the object to which $a$ bears the relation $O$. Thus, the question what counts as an operation reduces to the question which relations are "internal".

Wittgenstein's handling of technical matters is not widely celebrated for its precision. But the remarks about quantifiers borrow some clarity from Russell and Frege. And of course quantifiers became a standard part of the logical education of philosophers. On the other hand, the origins of the concept of operation are obscure, and corresponding extensions of quantificational logic never got much traction.

What is widely acknowledged is that the form-series device is supposed to yield an analysis of the ancestral in terms of an idea of iterated operation. There are few published attempts to propose any general comprehension principles for operations to do the job. Geach (1981, 170) suggests "the notation $[aRb, aSb, aR/Sb]$ gives us the series of propositions $aRb, aR/Rb, aR/(R/R)b$ *etc. ad inf.*..." Here, an expression like $aR/Rb$ is presumably supposed to abbreviate an ordinary first-order formula. However, the disabbreviation for nonelementarily expressed relations is left obscure; and the proposal makes

https://doi.org/10.1017/S1755020316000472 Published online by Cambridge University Press

the concept of operation seem tailored to a particular one of what are presumably various possible uses. In a similar vein, Potter (2009, 272) writes: "the formal series [of 4.1252] is expressed by the variable $[\, aRb,\ a\chi b,\ (\exists x) : aRx\, .\, x\chi b\,]$". The grounds of this declaration aren't made explicit: it's not clear why, for example, the third term of the series of indicated propositions shouldn't be something like $(\exists x) : aRx . (\exists x) : xRx . xRb$. The notation proposed by Potter appears to indicate what's intended only granted an understanding of what the whole thing is supposed to mean. So the proposal runs afoul of Wittgenstein's complaint about Russell, that "he had to mention the meaning of signs when establishing the rules for them" (3.331). Of course, Wittgenstein himself never established those rules clearly. But as I've urged, the question whether anything could realize the conception of logic in the *Tractatus* can't be answered without going beyond what Wittgenstein actually did. For better or for worse, the job is ours to make something clear.

There has also been disagreement about what should count as an operation and what shouldn't. For example, Sullivan (2004) contends that no internal relation distinguishes nontrivially between names, or—though this terminology is neither Sullivan's nor Wittgenstein's—that internal relations are "permutation-invariant" (see also Sundholm, 1992, 69). In contrast, Ricketts (2012) conjectures that form-series disjunctions might serve to simulate predicative higher-order existential generalization. Textual issues aside, the proposal would require enumerating the class of sentences $f(a)$, $f(b)$, ... for all names $a, b, \ldots$, by means of logically insignificant features like spelling. However, it is reasonable to assume that the class of operations should be closed under composition. But then supposing $O$ to enumerate the names as in Ricketts' proposal, the result of composing $O$ once with itself would enumerate, as it were, "every other name". It may then so happen to turn out that in the actual world, the thereby enumerated names would denote just those objects which are spacetime points interior to Russell's left hand. In another language, differing only in what would seem to be a logically insignificant manner, the probability would be zero that Russell's left hand is definable, but his nose might be definable instead. Ricketts' proposal implies that languages which differ over logically insignificant features are not intertranslatable.

Textual evidence suggests that Wittgenstein does countenance a formal enumeration of names in the *Notebooks* (e.g., 23.11.16d). But such suggestions largely disappear in the *Tractatus*. Disagreements about the precise extent of the concept of operation may need to be resolved on systematic grounds, which I leave to further work. In this paper's formal development, I will adopt a fairly strict policy: operations will be invariant under permutations of all names except those which occur as parameters in the operator-sign. The key mathematical results depend only on operations which are fully permutation invariant.

*1.2.1. Motivations for form-series.* Textual evidence suggests that two motives drove Wittgenstein to supplement the logic of objectual generality and truth-functions with formal series. Wittgenstein seems to have postulated quite early the presence of some propositional structure which shares the necessities inherent in a fact. The postulation gets expressed in a remark, at 4.023, that thanks to some "logical scaffolding", a person can "see" how things must be if a proposition is true. An example from the *Notebooks* of this sort of visibility might be the following:

> For example, perhaps we assert of a patch in our visual field that it is to the right of a line, and we assume that every patch in our visual field is infinitely complex. Then if we say that a point in that patch is to the right of the line, this proposition follows from the previous one, and if there are infinitely many points in the patch then infinitely many propositions

> of different content follow LOGICALLY from that first one. And this of itself shows that the proposition itself was as a matter of fact infinitely complex. That is, not the propositional sign by itself, but it together with its syntactical application. (Wittgenstein, 1979, 18.6.15g)

A year earlier, Wittgenstein had identified a similar complexity in the fact that a chair is brown, and sketched a formal series of propositions which would reproduce it (1979, 19.9.14, pp. 5 and 134).

The second motivation for the form-series device stems from problem of giving the general propositional form. Although the problem is mentioned already in the 1914 notes of Moore (1979, 113) it is not until April 1916 that there appears a sketch of Wittgenstein's eventual approach: "suppose that all simple propositions were given me: then it can simply be asked what propositions I can construct from them. And these are all propositions and this is how they given" (16.4.16; cf. T4.51). So, the notion of form-series here finds a second use, in an account of the way in which propositions are constructed by elementary propositions through repeated application of a formal procedure. As Avron (2003) argues, this technical task provides a natural motivation for adding to first-order logic a device to express the ancestral. Goldfarb (2012) speculates that it was a conviction that logic must comprehend the grounds of people's understanding of logic, and particularly the underpinnings of recognition of formulas and proofs, which sustained in Wittgenstein the view that induction is logical.

The two roles of formal series in the pre-*Tractatus* manuscripts lead to two roles in the *Tractatus*. On the one hand, T4.1252, T5.252 and T5.501 invoke and explain an "immanent" use of formal series in articulating the structure of facts; paradigmatic of such use is the expression of the ancestral. On the other hand, a "transcendent" use at T6 purportedly fulfills the promise of 4.51 to say how all propositions can be constructed from certain simple ones.

Sundholm (1992, 66) remarks of the transcendent use that it is at least superficially more complicated than the immanent one: it acts on a higher-order relation which takes as one argument not a single item but a potentially infinite multiplicity. In consequence, the use of a form-series variable to specify the totality of propositions flirts with impredicativity (Sundholm, 1992, 70). Why couldn't that variable be used to specify the basis of a truth-operation?

Here it may help to note a textual detail. The *Prototractatus* (Wittgenstein, 1971) antecedent of the explanation in the T5.252s addresses not the simpler, immanently admitted form but the higher-order form instead. So, in deriving the *Tractatus* from this earlier draft, Wittgenstein restricted his explanation of formal series to the immanent use. This clearly deliberate reversal suggests that Wittgenstein eventually decided to admit only the lower-level form under constructional procedures described at T6 by a use of the higher-level form.

The separation of "immanent" and "transcendent" form-series variables introduces a couple of potential responses to the problem about impredicativity raised by Sundholm. First, it might be proposed that there is some fixed type to which all meaningful uses of the form-series device must conform, while allowing that the expression at T6 does not conform to that type. To this proposal, it might be objected that once a system of propositional construction has become surveyable by means of a definite, though higher-level form-series variable, the higher-level expression should itself become susceptible to significant use in a yet broader collection of sentences. On this second proposal, any characterization of a system of propositions would yield *eo ipso* a further method of propositional construction, which the whole would still transcend. In that case, the general form of the truth-function given at T6 would be understood as in some sense open-ended (cf. Floyd, 2001). A difficulty for this proposal is that Wittgenstein repeatedly attempts to justify the postulation of a general propositional form precisely on the grounds that every possible form of a proposition must be foreseeable (4.5c; see also Wittgenstein (1979, 21.11.16)). Developing the open-endedness proposal would require clarifying Wittgenstein's notion of foreseeability.

*1.2.2. Toward an explication.* Luckily, the goals of this paper do not extend to explicating the general propositional form in detail. But the above gloss would license one helpful assumption, that the form-series method mentioned at 5.501 isn't intended to handle the apparently more complicated induction of T6. Rather, following the 5.2522s, we might naturally consider only form-series variables constructed from unary operations. However, even the notion of unary operation is of course still unclear, and no purportedly exhaustive reconstruction could be uncontroversial.

For several reasons I will take a minimalist approach to the positing of operations. First, the goal of the paper is to establish lower bounds on the complexity of the system, not upper bounds. Second, it is nice to see how much can be got from how little. Third, details of implementation do not just take care of themselves, but become complicated quickly (see §2.3, particularly Lemma 2.13).

The minimal analysis of operation I propose to investigate is roughly this. An operation is presented by a schematic letter together with a formula; the presented operation returns the result of substituting the operand for the letter in the presenting formula.

A bit more concretely, suppose that

$$\xi \mapsto \Omega(\xi)$$

is a procedure which, applied to a formula, returns another formula. Then in something like Wittgenstein's notation,

$$[\![A, \xi, \Omega(\xi)]\!]$$

would signify the series of formulas

$$A, \Omega(A), \Omega(\Omega(A)) \dots.$$

Now, consider a first-order formula $B$ which is ordinary except in that someplace where an atomic formula might have occurred, there occurs instead a schematic letter $p$. This can be supposed to determine, with respect to $p$, a function

$$\xi \mapsto B[p/\xi] \tag{1}$$

which, applied to the formula $A$, returns the result $B[p/A]$ of replacing each occurrence of $p$ in $B$ with $A$. In turn, the notation

$$[\![A, \xi, B[p/\xi]]\!] \tag{2}$$

can be understood to signify the series of formulas

$$A, B[p/A], B[p/B[p/A]], \dots.$$

Of course, there are many other reasonable notions of operation besides those of the form of (1). But I'll consider just these. This justifies a notational economy: instead of (2) I'll just write

$$[\![A, p, B]\!].$$

As we'll see in §3, this ostensibly weak reconstruction suffices to express all notions which can be defined by induction, including in particular the ancestral. There will actually be various ways of doing it, but they'll all have in common that they contain only finitely many variables. Thus, it's clear that they don't express the ancestral in quite the way that Wittgenstein envisaged at 4.1252. And more generally, it is certainly not the case that every formal procedure hazarded by Wittgenstein has the form $A \mapsto B[p/A]$. Conversely, however, it is plausible that everything of that form should count as a formal procedure.

On the present reconstruction, the form-series device does not simulate the occurrence of infinitely many variables in a formula. The form-series Wittgenstein gives at 4.1252 to express the ancestral does use infinitely many variables, though of course inessentially. One might take this to license essential uses too. A familiar concept which might then become expressible is the quantifier "there are infinitely many". The interpretive grounds for such a strengthened line of reconstruction should be supplemented with some systematic considerations. Wittgenstein does say that number is a "formal" concept (4.1272); perhaps then "there are infinitely many" ought to be a logically definable. However, by Wittgenstein's lights a concept's being formal doesn't imply that it's definable at all. For example, identity is supposed to be a formal concept but is not definable (see §1.1.3). Although the logicality of an infinity quantifier doesn't seem to be precluded by the text, the quantifier is not logically definable by induction, yet induction handles all the envisaged applications of the form-series device for which there is textual evidence.

*1.2.3. Programmatic background.* Let me conclude this section by sketching some interpretive context for the two logical eccentricities to be investigated in this paper—the restriction to Russellian structures and the addition of the form-series device. In the preface to the book Wittgenstein promises to solve philosophical problems by clarifying the nature of logic. A superficial survey indicates that the centrally organizing task of the book is to give a general propositional form, a purported common nature of whatever can be said or thought. It is not clear at first blush how such a task could itself complete the understanding of logic which is needed for the solution to philosophical problems.

In the *Tractatus*, the purported general propositional form gets presented twice. At first, Wittgenstein keeps it short:

> [. . . ] The general propositional form is: *es verhält sich so und so*. (4.5c)

It's not really clear how this could be a philosophical cure-all. In fact there is more to consider:

> A proposition constructs a world with the help of a logical scaffolding, so that one can actually *see* from the proposition how everything stands logically *if* it is true. (T4.023e)

Such scaffolding may be supposed to constitute the sort of logic whose misunderstanding is the purported source of philosophical problems. If that's right, the solution to philosophical problems might be seen to require attention to the nature of those manipulations of signs through which signs come to have sense: that is, to lay out the pieces of scaffolding clearly. Wittgenstein announces such a consideration of details at T5, which leads at T6 to a refinement of the general propositional form:

> A proposition is a truth-function of elementary propositions. (T5)
>
> The general form of the truth-function is $[\overline{p}, \overline{\xi}, N'(\overline{\xi})]$. That is the general form of the proposition. (T6)

So, Wittgenstein's promised solution to philosophical problems would appear to depend, at least in part, on the progress from T4.5 to T6. In particular, the progress is supposed to

https://doi.org/10.1017/S1755020316000472 Published online by Cambridge University Press

clarify just how it could be of necessity that if some propositions are true, then some other propositions are true as well. Accordingly, possibilities reappear as distributions of truth-value over propositions. And a fundamental problem of logic becomes to clarify how the structure of propositions constrains what distributions of truth and falsehood are possible for them.

As we'll see, the program leads to a realization of logic whose complexity turns on the number of objects that exist. In particular, if the number of objects might be uncountably infinite, then the complexity of the resulting notion of tautology seems to rule out any reasonable sense in which, after all, logical validity or logical consequence could be a mere matter of propositional structure. Under the assumption that validity and consequence can depend only on propositional structure, the significance of the main results of this paper can be stated as follows: it is only if the universe is countable that the conception of logic in the *Tractatus* might be coherent.

**§2. Framework.** We've now seen that Wittgenstein's conception of logic differs from the classical understanding of first-order logic in at least the following two ways. First, the notion of structure is subjected to a Russellian constraint, so that the elements of its domain are in one-one correspondence with the constants of its signature. Second, the logical vocabulary contains a device for the expression in finite space of countably infinite disjunctions.

In this section, I'll assemble a framework for studying those features. After sketching in §2.1 a suitable variant $\mathcal{L}$ of first-order logic, I'll show in §2.2 that the Russellian constraint yields an interpretation of some familiar slogans from the *Tractatus*. In §2.3, I'll show that $\mathcal{L}$ can be extended to a system $\mathcal{LF}$ which implements an analysis of Wittgenstein's form-series device. But first some common background.

A signature $\mathcal{S}$ consists, for all $k$, of a set of $k$-place predicates and a set of $k$-place function symbols. I will say that a signature $\mathcal{R}$ is Russellian provided that $\mathcal{R}$ contains at least one predicate and at least one function symbol, but no function symbol of arity greater than zero. Function symbols of arity zero are also known as constants.

A structure $\mathcal{M}$ for $\mathcal{S}$ consists of a nonempty set, its domain $\mathcal{D}$, plus a mapping which takes $k$-place function symbols and $k$-place predicates to $k$-ary functions and $k$-ary relations on $\mathcal{D}$. Let $\mathcal{R}$ be a Russellian signature. I will say that a structure $\mathcal{M}$ for $\mathcal{R}$ is Russellian, or is a $\mathcal{D}$-structure, if it meets the following further conditions:

- every constant of $\mathcal{R}$ is its own image under $\mathcal{M}$, and
- every element of $\mathcal{D}$ is the image of a constant.

The convention that constants denote themselves is purely for notational convenience, and all subsequent results will extend transparently to the general situation in the *Tractatus* where denotation is any bijection between the constants and the domain.

Finally, a logic maps a signature to the collection of classes of its structures. Typically the mapping is determined in two steps. First a syntax generates from the signature a class of formulas. Then a semantics associates with a formula the class of structures in which it is true. Let's now see how this works for $\mathcal{L}$ and $\mathcal{LF}$.

***2.1. The system $\mathcal{L}$.*** As indicated in §1.1, the system of the *Tractatus* can be understood as including a variant of first-order logic. I'll refer to this variant as $\mathcal{L}$. Let me begin by rehearsing its mostly familiar construction.

Given a Russellian signature, first-order logic determines a language like this. There is a set of individual variables $x, y, \ldots, x_0, y_0, \ldots$, which, given the function symbols of $\mathcal{S}$,

https://doi.org/10.1017/S1755020316000472 Published online by Cambridge University Press

determines the class of terms of the language. Second, the class of $\mathcal{L}$-formulas is given by an induction with just these clauses:

- $R\vec{t}$ is a formula if $R$ is a $k$-ary predicate and $\vec{t}$ is a list of terms of length $k$
- $t = u$ is a formula if $t, u$ are terms
- $\neg A$, $(A \vee B)$, and $\exists x\, A$ are formulas whenever $A$, $B$ are formulas and $x$ is a variable.

I will refer to formulas so generated as the proper formulas of $\mathcal{L}$.

It will be useful to have extended $\mathcal{L}$ with a category of schematic letters, $p, q, r, p_1, \dots,$ adding the inductive clause

- every schematic letter is a formula.

Improper formulas will not be assigned semantic values directly. Rather, they support the implementation of the form-series device in $\mathcal{LF}$ (see §2.3).

Just to fix notation, here are some fairly standard niceties. Free, bound, and binding occurrences of a variable should be taken in the usual way. I'll write $A[x/t]$ for the result of replacing each free occurrence of $x$ in $A$ with the term $t$, unless this replacement introduces new bound occurrences. Vectors over terms for expressions indicate finite sequences of expressions of the indicated type. Then $A[\vec{x}/\vec{t}]$ signifies the result of simultaneously substituting each element of the sequence $\vec{t}$ for the corresponding element of the sequence $\vec{x}$, relettering variables bound in $A$ where necessary. Finally, $A[\vec{a}]$ abbreviates $A[\vec{x}/\vec{a}]$ where $\vec{x}$ is the canonically ordered sequence of all variables occurring free in $A$.

The definition of truth for $\mathcal{L}$ is straightforward. I will specialize for the notion of Russellian structure the approach of Shoenfield (1967). Let $\mathcal{R}$ be a Russellian signature with $\mathcal{D}$ its class of constants, and let $\mathcal{M}$ be a $\mathcal{D}$-structure for $\mathcal{R}$. Now suppose that $A$ is a proper closed formula of $\mathcal{L}$. The atomic, equality, and truth-functional cases are handled as usual. But suppose that $A$ is $\exists x\, B$. In this case we stipulate that

- $\mathcal{M} \models \exists x\, B$ iff $\mathcal{M} \models B[a]$ for some $a$ in $\mathcal{D}$.

***2.2. Effects of the Russellian constraint.*** The specialization to Russellian structures has deep effects on the notions of validity and consequence, which are nonetheless defined in a familiar-looking way.

DEFINITION 2.1. *Let $\mathcal{R}$ be a Russellian signature with domain $\mathcal{D}$. A proper closed formula $A$ over the signature $\mathcal{R}$ is $\mathcal{D}$-*valid *if $\mathcal{M} \models A$ for all $\mathcal{D}$-structures $\mathcal{M}$ for $\mathcal{R}$. Likewise, $A$ is a $\mathcal{D}$-*consequence *of the set $X$ of proper closed formulas, or $X \Rightarrow_{\mathcal{D}} A$, provided that there is no $\mathcal{D}$-structure $\mathcal{M}$ for $\mathcal{R}$ such that $\mathcal{M} \not\models A$ while $\mathcal{M} \models B$ for all $B$ in $X$.*

With the notion of $\mathcal{D}$-consequence in hand, we can begin to interpret some basic slogans of the *Tractatus*. According to 4.26, "the world" is supposed to be completely described by specifying which elementary propositions are true and which false. Say that the diagram $\Delta(\mathcal{M})$ of a structure consists of the atomic sentences true in $\mathcal{M}$, plus the negations of the atomic sentences false in $\mathcal{M}$. The remark of 4.26 now becomes that a structure is axiomatized up to identity by its diagram.

PROPOSITION 2.2. *Let $\mathcal{M}$ and $\mathcal{M}'$ be $\mathcal{D}$-structures. If $\mathcal{M}' \models \Delta(\mathcal{M})$, then $\mathcal{M}' = \mathcal{M}$.*

*Proof.* Since $\mathcal{M}$ and $\mathcal{M}'$ are both $\mathcal{D}$-structures, they have the same domain. So suppose $R$ is an $\mathcal{R}$-predicate. Then $(\vec{a}) \in R^{\mathcal{M}}$ iff $R\vec{a} \in \Delta(\mathcal{M})$, and $R\vec{a} \in \Delta(\mathcal{M})$ iff $(\vec{a}) \in R^{\mathcal{M}'}$. □

Note that if the notion of $\mathcal{D}$-structure is relaxed so as to require only that denotation of constants be a bijection, then diagrams axiomatize structures only up to isomorphism. Subsequent results don't depend on this difference.

When every structure is axiomatized by its diagram, consequence becomes a generalization of truth. Or in other words, truth is what follows from a diagram.

PROPOSITION 2.3. *Let $\mathcal{M}$ be a $\mathcal{D}$-structure. Then $\mathcal{M} \models A$ iff $\Delta(\mathcal{M}) \Rightarrow_{\mathcal{D}} A$.*

*Proof.* Since $\mathcal{M} \models \Delta(\mathcal{M})$, therefore $\Delta(\mathcal{M}) \not\Rightarrow_{\mathcal{D}} A$ if $\mathcal{M} \not\models A$. Conversely, suppose $\mathcal{M} \models A$. By Proposition 2.2, if $\mathcal{M}' \models \Delta(\mathcal{M})$, then $\mathcal{M}' = \mathcal{M}$, so $\mathcal{M}' \models A$. Hence $\Delta(\mathcal{M}) \Rightarrow_{\mathcal{D}} A$. □

Proposition 2.3 records a radical departure from the classical concept of consequence: for Wittgenstein, the concept of consequence embeds the concept of truth. See Martin-Löf (1996) for discussion of this approach. Of course, it is the left-to-right direction of Proposition 2.3 which does not hold for the classical or fixed-domain consequence relations.

Another slogan of the *Tractatus* is that a proposition is a "truth-function" of elementary propositions. In the present framework, this slogan becomes that each sentence or its negation is a Russellian consequence of the diagram of a structure.

PROPOSITION 2.4. $\Delta(\mathcal{M}) \Rightarrow_{\mathcal{D}} A$ *iff* $\Delta(\mathcal{M}) \not\Rightarrow_{\mathcal{D}} \neg A$.

*Proof.* Of course $\mathcal{M} \models A$ iff $\mathcal{M} \not\models \neg A$. But by Proposition 2.3, $\Delta(\mathcal{M}) \Rightarrow_{\mathcal{D}} A$ iff $\mathcal{M} \models A$, and $\mathcal{M} \not\models \neg A$ iff $\Delta(\mathcal{M}) \not\Rightarrow_{\mathcal{D}} \neg A$. □

Consider, for example, the concept of arithmetical truth. Let $\mathcal{R}$ be a Russellian signature with a couple of three-place predicates, while the set $\mathcal{D}$ of constants of $\mathcal{R}$ is countably infinite. Now, let $\mathcal{N}$ be a $\mathcal{D}$-structure for $\mathcal{R}$ which is isomorphic to the structure of the natural numbers under the ternary relations of sum and product.

PROPOSITION 2.5. *Under the $\mathcal{D}$-consequence relation, the set $\Delta(\mathcal{N})$ axiomatizes $\mathcal{N}$ up to identity. Thus, the set of arithmetical truths is decidable relative to the set of $\mathcal{D}$-consequences of $\Delta(\mathcal{N})$.*

*Proof.* Immediate from Propositions 2.2 and 2.3. □

Of course, there's nothing special about the structure $\mathcal{N}$ in this context; Proposition 2.5 holds for any structure whatsoever.

***2.3. The system $\mathcal{LF}$.*** In §1.2, I argued that Wittgenstein envisions logic also to include a method of expressing some countably infinite disjunctions, though in finite space. Let's now develop an extension $\mathcal{LF}$ of $\mathcal{L}$, which implements this idea. The goal is to become able to express a disjunction whose disjuncts are the results of applying repeatedly some finitely presented "operation". We will pursue the proposal of §1.2.2, to consider only operations of substituting the operand for a schematic letter in a formula.

*2.3.1. $\mathcal{LF}$-syntax.* The rules for building formulas of $\mathcal{LF}$ are precisely those of $\mathcal{L}$, plus just one more:

- $\bigvee[\![A, p, B]\!]$ is an $\mathcal{LF}$-formula provided that $A$ is an $\mathcal{LF}$-formula and $B$ is an $\mathcal{L}$-formula.

Let's call the indicated occurrence of $p$ a binding occurrence of $p$; its scope will be said to be the formula $B$. An occurrence of $p$ is bound iff it falls in the scope of a binding occurrence of $p$.

https://doi.org/10.1017/S1755020316000472 Published online by Cambridge University Press

I'll say that a formula of $\mathcal{LF}$ is proper iff it contains no free occurrence of a schematic letter. Note that an improper formula $B$ may be a subformula of a proper formula iff $B$ is actually a formula of $\mathcal{L}$. Improper formulas which don't belong to $\mathcal{L}$ remain only for theoretical convenience. It will be just the proper closed formulas which receive a truth-value in a structure.

*2.3.2. $\mathcal{LF}$-operations.* A formula $\bigvee[\![A, p, B]\!]$ is to become understood as a disjunction. Its terms are to be the results of repeatedly applying the operation presented by $p$, $B$ to the initial formula $A$. The pair $p$, $B$ presents the operation which substitutes the operand for each occurrence of $p$ in $B$. To spell this out, let's fix some notation.

DEFINITION 2.6. (*i*) *The formula $B[p/A]$ is the result of substituting $A$ for every free occurrence of $p$ in $B$.* (*ii*) *If $C = \bigvee[\![A, p, B]\!]$, then write $C^0 = A$ and $C^{i+1} = B[p/C^i]$.*

Regarding clause (i), note that operator-signs contain no bound schematic letters, so the qualifier "free" isn't necessary for intepreting form-series formulas. But it will be useful to have the more general concept of substitution on hand throughout. The clause (ii) begins to spell out how form-series disjunctions present a series of disjuncts. I will sometimes write $[\![A, p, B]\!]^k$ for $(\bigvee[\![A, p, B]\!])^k$.

An alternative characterization of iterated substitution will become useful shortly, in proving the fundamental Lemma 2.13. Write $[p/B]^k$ for $k$ successive applications of the substitution $[p/B]$.

LEMMA 2.7. *Let $\bigvee[\![A, p, B]\!]$ be a formula of $\mathcal{LF}$. Then*

$$[\![A, p, B]\!]^k = p[p/B]^k[p/A].$$

*Proof.* Let's use induction on $k$. The result is clear for $k = 0$. So suppose it to hold for $k$. Then

$$\begin{aligned}
[\![A, p, B]\!]^{k+1} &= B[p/[\![A, p, B]\!]^k] \\
&= B[p/p[p/B]^k[p/A]] \\
&= B[p/p[p/B]^k][p/A] \\
&= B[p/B]^k[p/A] \\
&= p[p/B]^{k+1}[p/A].
\end{aligned}$$ □

Before proceeding to the semantics, let's consider a nontrivial application of the form-series device. Let $R$ be a dyadic predicate. The ancestral of the relation expressed by $R$ can be expressed by an infinite disjunction of a series of formulas like

$$\begin{gathered}
Rxy \\
\exists z_0(Rxz_0 \wedge Rz_0y) \\
\exists z_1(\exists z_0(Rxz_0 \wedge Rz_0z_1) \wedge Rz_1y) \\
\exists z_2(\exists z_1(\exists z_0(Rxz_0 \wedge Rz_0z_1) \wedge Rz_1z_2) \wedge Rz_2y) \\
\vdots
\end{gathered}$$

These formulas altogether contain infinitely many variables $z_0, z_1, \ldots$, and so they can't be generated merely by an operation of the sort postulated here. But, each quantifier $\exists z_{k+2}$ doesn't bind any variable in the scope of $\exists z_k$. So the indices should be recyclable by some trick or other. For example, write $A_0, A_1, \ldots$ for the first, second, ... terms of the series

https://doi.org/10.1017/S1755020316000472 Published online by Cambridge University Press

$$Rxy$$
$$\exists z(Rxz \wedge Rzy)$$
$$\exists z(\exists y(Rxy \wedge Ryz) \wedge Rzy)$$
$$\exists z(\exists y(\exists z(Rxz \wedge Rzy) \wedge Ryz) \wedge Rzy)$$
$$\vdots$$

Let $B = \exists z(\exists y(p \wedge Ryz) \wedge Rzy)$. Then

$$A_{2k} = p[p/B]^k[p/A_0], \quad \text{and}$$
$$A_{2k+1} = p[p/B]^k[p/A_1]$$

for all $k < \omega$. So by Lemma 2.7, the disjuncts $A_i$ of an expression of an ancestral are precisely the 'disjuncts' of

$$\left(\bigvee[\![A_0, p, B]\!]\right) \vee \left(\bigvee[\![A_1, p, B]\!]\right).$$

Of course, I have not yet said what it is for a formula of $\mathcal{LF}$ to express something. Let's fix that.

*2.3.3. $\mathcal{LF}$-semantics.* Developing a semantics for $\mathcal{LF}$ will take a little more effort. To construct a definition of truth it is tempting simply to extend the classical definition directly, evaluating bound variables by assignment or instantiation. Unfortunately, this does not work. The reason is that, roughly speaking, we would like form-series formulas to behave like countably infinite disjunctions—but that's not what they are.

To see the difficulty, note that the substitution of terms does not commute with the expansion of form-series formulas into their infinitary counterparts. Consider, for example, a formula $A = \bigvee[\![Fx, p, \exists x(Gx \wedge p)]\!]$. Applying directly the usual notion of substitution of terms would yield $A[x/a] = \bigvee[\![Fa, p, \exists x(Gx \wedge p)]\!]$, and so, for example, $A[x/a]^1 = \exists x(Gx \wedge Fa)$. On the other hand, $A^1 = \exists x(Gx \wedge Fx)$, and so $A^1[x/a] = \exists x(Gx \wedge Fx)$. Of course, quantification doesn't quite require the good behavior of substitution, but here the mischief is deeper: the introduction of form-series into first-order logic disrupts the expected notion of free occurrence of a variable.

To solve this problem, let's play a trick: run two stages of evaluation. First expand form-series variables into the formal series they present, and only then evaluate the quantifiers. Formulas of $\mathcal{LF}$ thus get expanded into an infinitary extension of first-order logic $\overline{\overline{\mathcal{LF}}}$. This results by adding, to the formation rules of $\mathcal{L}$, the clause

- $\bigvee_{k<\omega} A_k$ is a formula if each $A_k$ is a formula,

and to the evaluation rules of $\mathcal{L}$ the corresponding clause

- $\mathcal{M} \models \bigvee_{k<\omega} A_k$ iff $\mathcal{M} \models A_k$ for some $k < \omega$.

As with $\mathcal{L}$, I'll say that a formula of $\overline{\overline{\mathcal{LF}}}$ is proper if it contains no schematic letters.

We want to define recursively a mapping $\overline{\overline{\cdots}}$ which expands a formula of $\mathcal{LF}$ into some corresponding formula of $\overline{\overline{\mathcal{LF}}}$. To justify the recursion we need a further idea.

DEFINITION 2.8. *The* level *of a formula $A$ of $\mathcal{LF}$ is the pair $(i_A, j_A)$, where*

- $i_A = 0$, *if $A$ is atomic or schematic; $i_A = i_B + 1$ if $A = \bigvee[\![B, p, C]\!]$, and otherwise $i_A$ is the maximum of the $i_B$ for all immediate subformulas $B$ of $A$;*

https://doi.org/10.1017/S1755020316000472 Published online by Cambridge University Press

- $j_A = 0$ *if* $A$ *is atomic or schematic, and otherwise* $j_A = k+1$, *where* $k$ *is the maximum of the* $j_B$ *for* $B$ *an immediate subformula of* $A$.

*The level of* $A$ *is* lower *than the level of* $B$ *iff either* $i_A < i_B$ *or* $i_A = i_B$ *and* $j_A < j_B$.

The ordering of levels is clearly wellfounded. Unlike ordinary syntactical complexity, it fulfils the following useful condition.

LEMMA 2.9. *Suppose that* $C = \bigvee[\![A, p, B]\!]$ *is a formula of* $\mathcal{LF}$. *Then* $C^k$ *has lower level than* $C$.

*Proof.* The result is clear since $i_{C^k} \leq i_{C^0}$ and $i_{C^0} < i_C$. □

Lemma 2.9 now justifies the desired analysis of form-series expansion.

DEFINITION 2.10. *The* form-series expansion *of a formula* $A$ *is the* $\overline{\overline{\mathcal{LF}}}$*-formula*

$$\overline{\overline{A}} = \begin{cases} \bigvee_{k<\omega} \overline{\overline{A^k}}, \text{ if } A \text{ is } \bigvee[\![B, p, C]\!] \\ \exists x \overline{\overline{B}}, \text{ if } A \text{ is } \exists x B \\ \overline{\overline{B}} \vee \overline{\overline{C}}, \text{ if } A \text{ is } B \vee C \\ \neg \overline{\overline{B}}, \text{ if } A \text{ is } \neg B \\ A, \text{ if } A \text{ is atomic or schematic.} \end{cases}$$

In ordinary quantificational logic, a formula is truth-evaluable if closed. That is not the case in $\mathcal{LF}$, because a formula with no free variables may still contain an unbound schematic letter. But a proper closed formula is one that contains neither free variables nor unbound schematic letters. The following somewhat tedious lemma guarantees that proper closed formulas expand into proper closed formulas of $\overline{\overline{\mathcal{LF}}}$.

LEMMA 2.11. *Suppose that* $A$ *is a merely proper formula of* $\mathcal{LF}$. *Then* $\overline{\overline{A}}$ *is a proper formula of* $\overline{\overline{\mathcal{LF}}}$ *whose free variables are among those of* $A$.

*Proof.* We argue by induction on the level of $A$. Only two cases are nontrivial.

First suppose that $A$ is $\exists x B$. By induction hypothesis, $\overline{\overline{B}}$ is a proper formula of $\mathcal{LF}$ whose free variables are among those of $B$. Then $\overline{\overline{A}} = \exists x \overline{\overline{B}}$ is a proper formula of $\mathcal{LF}$, whose free variables are among those of $A$.

Now suppose $A = \bigvee[\![B, p, C]\!]$. By Definition 2.10, $\overline{\overline{A}} = \bigvee_{k<\omega} \overline{\overline{A^k}}$. Clearly $B = A^0$ must be proper if $A$ is, and the free variables of $B$ are among those of $A$. Suppose now that $A^k$ is proper and that its free variables are among those of $A$. Then $A^{k+1} = C[p/A^k]$. The free variables of $C$ must be among those of $A$. And since $A$ is proper, a schematic letter $q$ can occur free in $C$ only if $q = p$. Hence $A^{k+1}$ is proper, and its free variables are among those of $A$. Since each $A^i$ has lower level than $A$, it follows by outer induction hypothesis that each $\overline{\overline{A}}^i$, and hence $\overline{\overline{A}}$ itself, is a proper formula of $\overline{\overline{\mathcal{LF}}}$ whose free variables are among those of $A$. □

After that bit of ground-clearing we can extend the definition of truth to $\mathcal{LF}$.

DEFINITION 2.12. *Let* $A$ *be a proper closed formula of* $\mathcal{LF}$. *Then* $\mathcal{M} \models A$ *iff* $\mathcal{M} \models \overline{\overline{A}}$.

*2.3.4. Extensionality.* A primary goal of this paper is to determine the expressiveness of $\mathcal{LF}$. We have just seen a definition of truth which does not proceed by syntactical complexity. So it is not at all obvious that $\mathcal{LF}$ can be given a robust semantics.

https://doi.org/10.1017/S1755020316000472 Published online by Cambridge University Press

I'll now show, however, that $\mathcal{LF}$ is extensional. Toward this end, let me first record a basic lemma, that the expansion of a form-series formula into its infinitary counterpart commutes with substitution of formulas.

LEMMA 2.13. *Suppose that $A$ is proper, and that $B[p/A]$ is a formula of $\mathcal{LF}$. Then*

$$\overline{\overline{B[p/A]}} = \overline{\overline{B}}[p/\overline{\overline{A}}].$$

Let's now use the lemma to derive the promised extensionality. Then I'll prove the lemma to conclude the section.

PROPOSITION 2.14. *Suppose that $C[p/A], C[p/B]$ are proper formulas of $\mathcal{LF}$, and that $A$ and $B$ are proper. If*

$$\mathcal{M} \models \forall\vec{u}(A \leftrightarrow B)$$

*then*

$$\mathcal{M} \models C[p/A] \leftrightarrow C[p/B].$$

*Proof.* By the definition of $\models$ and $\overline{\overline{\cdots}}$, the condition $\mathcal{M} \models \forall\vec{u}(A \leftrightarrow B)$ implies

$$\mathcal{M} \models \forall\vec{u}(\overline{\overline{A}} \leftrightarrow \overline{\overline{B}}).$$

By the extensionality of infinitary logic it follows that

$$\mathcal{M} \models \overline{\overline{C}}[p/\overline{\overline{A}}] \leftrightarrow \overline{\overline{C}}[p/\overline{\overline{B}}].$$

So by Lemma 2.13,

$$\mathcal{M} \models \overline{\overline{C[p/A]}} \leftrightarrow \overline{\overline{C[p/B]}}.$$

The result now follows by the definition of $\models$. □

Lemma 2.13 owes its relatively straightforward proof to the parsimonious implementation of formal series in $\mathcal{LF}$.

*Proof of Lemma.* The goal is to prove that $\overline{\overline{B[p/A]}} = \overline{\overline{B}}[p/\overline{\overline{A}}]$, under the assumption that $A$ is proper. We argue by induction on the level of $B$. It suffices to handle the only nontrivial situation, where $B = \bigvee[\![C, q, D]\!]$.

Let's distinguish cases according to whether or not $p = q$. First suppose that $p = q$. It is clear that

$$p[p/D]^k[p/C[p/A]] = p[p/D]^k[p/C][p/A].$$

By Lemma 2.7 together with Definitions 2.6 and 2.10, it follows that

$$\begin{aligned}
\overline{\overline{B[p/A]}} &= \overline{\overline{\bigvee[\![C, p, D]\!][p/A]}} = \overline{\overline{\bigvee[\![C[p/A], p, D]\!]}} \\
&= \bigvee_{k<\omega} \overline{\overline{p[p/D]^k[p/C[p/A]]}} = \bigvee_{k<\omega} \overline{\overline{p[p/D]^k[p/C][p/A]}} \\
&= \bigvee_{k<\omega} \overline{\overline{B^k[p/A]}}.
\end{aligned} \tag{3}$$

On the other hand, suppose that $p \neq q$. Using By Lemma 2.7 together with Definitions 2.6 and 2.10 again,

https://doi.org/10.1017/S1755020316000472 Published online by Cambridge University Press

$$\begin{aligned}\overline{\overline{B[p/A]}} \;&=\; \overline{\overline{\bigvee[\![C, q, D]\!][p/A]}} \;=\; \overline{\overline{\bigvee[\![C[p/A], q, D[p/A]]\!]}}\\ &=\; \bigvee_{k<\omega} \overline{\overline{q[q/D[p/A]]^k[q/C[p/A]]}}. \end{aligned} \tag{4}$$

Let's now argue by induction on $k$ that

$$q[q/D[p/A]]^k[q/C[p/A]] \;=\; q[q/D]^k[q/C][p/A]. \tag{5}$$

The case of $k = 0$ is trivial. So assume (5) to hold for $k \geq 0$. Using $p \neq q$ and the propriety of $A$, straightforward considerations about substitution imply that

$$\begin{aligned} q[q/D[p/A]]^{k+1}[q/C][p/A] \;&=\; q[q/D[p/A]][q/D]^k[q/C[p/A]]\\ &=\; q[q/D][p/A][q/D]^k[q/C][p/A]\\ &=\; q[q/D]^{k+1}[q/C][p/A]. \end{aligned} \tag{6}$$

This proves (5). But (5) together with (4) imply by Lemma 2.7 that

$$\overline{\overline{B[p/A]}} \;=\; \bigvee_{k<\omega} \overline{\overline{q[q/D]^k[q/C][p/A]}} \;=\; \bigvee_{k<\omega} \overline{\overline{B^k[p/A]}} \tag{7}$$

holds in the case that $p \neq q$ as well as in the case that $p = q$.

However, $B^k$ has lower level than $B$. So the outer induction hypothesis applies:

$$\overline{\overline{B^k[p/A]}} \;=\; \overline{\overline{B^k}}[p/\overline{\overline{A}}]. \tag{8}$$

From (3), (7), and (8), we obtain the desired conclusion

$$\overline{\overline{B[p/A]}} \;=\; \bigvee_{k<\omega} \overline{\overline{B^k[p/A]}} \;=\; \bigvee_{k<\omega} \overline{\overline{B^k}}\,[p/\overline{\overline{A}}] \;=\; \overline{\overline{B}}[p/\overline{\overline{A}}]. \qquad \square$$

**§3. Expressiveness.** We've now got a reconstruction $\mathcal{LF}$ of Wittgenstein's logical system. Let's see what it can do. In §3.1, I'll begin by developing the semantics a bit further, spelling out what it is for formulas to define relations and operators. I'll then show that $\mathcal{LF}$ has some further nice properties: one result settles an issue about definability of infinitary conjunctions, and another gives a semantic characterization of form-series formulas. In §3.2, I'll apply that semantic machinery to show that $\mathcal{LF}$ is capable of expressing finitary inductive definitions; we'll see in particular that the form-series device yields a categorical analysis of arithmetic. Finally, in §3.3 I'll sketch the rather diffuse system of relationships, in point of expressiveness, between $\mathcal{LF}$ and other extensions of first-order logic.

***3.1. Definability of relations and operators.*** Recall that by Proposition 2.14 of §2.3, the system $\mathcal{LF}$ enjoys a strong form of extensionality: in any reasonable situation, a subformula can be exchanged with any of its logical equivalents without change of truth-value. Exploiting extensionality everywhere, we can now complete the task of wrangling the unruly syntactical manipulations into a denotational semantics. This will sharpen later talk of expressiveness, for example in the treatment of induction in §3.2.

By Lemma 2.11, a formula which is closed and proper has a truth-value. So such formulas have truth-values as suitable semantic values. What about other formulas? Regarding those which are proper but not closed, we can follow an approach which is typical to first-order logic. Take an $n$-ary relation over a domain $\mathcal{D}$ to be a set of $n$-tuples of elements of $\mathcal{D}$. As a limiting case, assume that there is exactly one 0-tuple $\emptyset$. Now, suppose that $A$

is any proper formula containing exactly $k \geq 0$ free variables. Then the formula $A$ defines over $\mathcal{M}$ a $k$-ary relation $|A|^{\mathcal{M}}$ over $\mathcal{D}$ as follows:

$$\left|A\right|^{\mathcal{M}} = \{(a_1, \ldots, a_k) : \mathcal{M} \models A[a_1, \ldots, a_k]\}.$$

This stipulation subsumes the treatment of closed formulas, once the truth-values are identified with the zeroary relations $\{\emptyset\}$ and $\emptyset$ (Krivine, 1998, 63).

To interpret form-series expressions, we also need to make some sense of improper formulas. Here it is easiest to proceed indirectly. Say that a operator on $\mathcal{D}$ is a function from $j$-ary relations on $\mathcal{D}$ to $k$-ary relations on $\mathcal{D}$; let's call the pair $j, k$ its type. Now, let $\mathcal{M}$ be a structure. Where $R, S \ldots$ are predicates not in the signature of $\mathcal{M}$, let $\mathcal{M}, X, Y, \ldots$ be the result of expanding $\mathcal{M}$ to interpret $R$ as $X$, $S$ as $Y$, and so on. As with the treatment of satisfaction, the notation $\mathcal{M}, X, Y, \ldots$ tacitly assumes a lexicographic ordering of the predicates interpreted.

Suppose that $A$ is a formula in the signature of an expansion $\mathcal{M}, X$ of $\mathcal{M}$ to $R$. Then $A$ defines through $R$ over $\mathcal{M}$ the operator

$$\left|A\right|^{R:\mathcal{M}} : X \mapsto \left|A\right|^{\mathcal{M},X}.$$

The type of $A$ is $j, k$, with $j$ the arity of $R$, and $k$ the arity of $|A|^{\mathcal{M},X}$. I will say that an operator is $\mathcal{L}$- or $\mathcal{LF}$-definable if some formula of $\mathcal{L}$, or of $\mathcal{LF}$, defines $\Gamma$.

For example, the formula $\exists y Rx_1 \cdots x_k y$ defines through $R$ the operator of type $k+1, k$ such that

$$\begin{aligned} \left|\exists y R\vec{x} y\right|^{R:\mathcal{M}}(X) &= \{(\vec{a}) : \mathcal{M}, X \models \exists y R\vec{x} y[\vec{a}]\} \\ &= \{(\vec{a}) : (\vec{a}, b) \in X \text{ for some } b \in \mathcal{D}\}. \end{aligned}$$

We can now give a method of interpreting improper formulas. It will only be needed here to interpret an improper formula $B$ with at most one free letter, say $p$. Suppose that $R$ is not in the signature of $\mathcal{M}$. The above stipulations imply that the formula $B[p/R\vec{x}]$ defines through $R$ over $\mathcal{M}$ the operator

$$\left|B[p/R\vec{x}]\right|^{R:\mathcal{M}} : X \mapsto \left|B[p/R\vec{x}]\right|^{\mathcal{M},X}.$$

In this way, an improper formula $B$ is interpreted relative to a substitution of some $R\vec{x}$ for $p$. The type of the defined operator is $j, k$ where $j$ is the length of $\vec{x}$, and $k$ is the number of variables free in $B[p/R\vec{x}]$. Clearly, the choice of $\vec{x}$ affects the type of the operator defined. In form-series contexts, normal uses will involve an operator of type $j, j$ for some $j$.

The following lemma will clarify the semantics of $\mathcal{LF}$'s form-series device.

LEMMA 3.1. *Suppose that $B[p/C]$ is a proper $\mathcal{LF}$-formula, that $C$ is proper, and that the variables free in $C$ are precisely $\vec{x}$. Then*

$$\left|B[p/C]\right|^{\mathcal{M}} = \left|B[p/R\vec{x}]\right|^{R:\mathcal{M}}(|C|^{\mathcal{M}})$$

*Proof.* Immediate from extensionality plus the notion of defined operator. □

Before applying Lemma 3.1 to interpret formal series, we just need to smooth out one wrinkle. Say that a form-series expression $[\![A, p, B]\!]$ is normal if the variables free in each $[\![A, p, B]\!]^k$ are precisely the variables free in $A$. It suffices to interpret normal expressions: $\bigvee[\![A, p, B]\!]$ is equivalent to $A \vee \bigvee[\![B[p/A], p, B]\!]$, and the latter is normal. At last, we interpret a normal form-series disjunction as the union of results of applying a $\mathcal{L}$-defined operator zero or more times to a defined relation.

https://doi.org/10.1017/S1755020316000472 Published online by Cambridge University Press

PROPOSITION 3.2. *Suppose that* $[\![A, p, B]\!]$ *is normal and proper, and that the variables free in* $A$ *are* $\vec{x}$. *Then*

$$\left|\bigvee[\![A, p, B]\!]\right| \;=\; \bigcup_{k<\omega}\left(\left|B[p/R\vec{x}]\right|^{R:\mathcal{M}}\right)^{k}\left(\left|A\right|^{\mathcal{M}}\right).$$

*Proof.* Lemma 3.1 implies for all $k$ that

$$\left|[\![A, p, B]\!]^{k+1}\right| \;=\; \left|B[p/R\vec{x}]\right|^{R:\mathcal{M}}\left(\left|[\![A, p, B]\!]^{k}\right|\right)$$

from which the claim is immediate by induction and the definition of truth. □

We can now take care of a puzzle which has been left hanging for a while. Recall that in building up $\mathcal{LF}$, form-series disjunctions were adopted as primitive. But what about conjunctions? As noted in §1.1.1, an adequate reconstruction of Wittgenstein's system ought to predict the question is not trivial.

Note that by Proposition 3.2, it is reasonable to understand a formula $C$ to express the conjunction of $[\![A, p, B]\!]$ provided that

$$\left|C\right|^{\mathcal{M}} \;=\; \bigcap_{k<\omega}\left(\left|B[p/R\vec{x}]\right|^{R:\mathcal{M}}\right)^{k}\left(\left|A\right|^{\mathcal{M}}\right)$$

for all $\mathcal{M}$. We now have the following.

PROPOSITION 3.3. *Suppose that* $[\![A, p, B]\!]$ *is normal and proper. Then*

$$\neg\bigvee[\![\neg A, p, \neg B[p/\neg p]]\!]$$

*expresses the conjunction of* $[\![A, p, B]\!]$.

*Proof.* Suppose that $\Gamma = |B[p/R\vec{x}]|^{R:\mathcal{M}}$. Let's write $\overline{X}$ for the complement of $X$ relative to the domain of $\mathcal{M}$. Let $\hat{\Gamma}$ be defined by $\hat{\Gamma}(X) = \overline{\Gamma\left(\overline{X}\right)}$.

Applying Lemma 3.1 twice,

$$\left|\neg B[p/\neg R\vec{x}]\right|^{R:\mathcal{M}}(X) \;=\; \overline{\left|B[p/R\vec{x}]\right|^{\mathcal{M},\overline{X}}} \;=\; \hat{\Gamma}(X).$$

By Proposition 3.2, it follows that

$$\left|\neg\bigvee[\![\neg A, p, \neg B[p/\neg p]]\!]\right| \;=\; \overline{\bigcup_{k<\omega}\hat{\Gamma}^{k}\left(\overline{X}\right)}.$$

It therefore suffices to show that

$$\overline{\bigcup_{k<\omega}\hat{\Gamma}^{k}\left(\overline{X}\right)} \;=\; \bigcap_{k<\omega}\Gamma^{k}(X).$$

To this end, let's argue by induction on $k$ that $\hat{\Gamma}^{k}\left(\overline{X}\right) = \overline{\Gamma^{k}(X)}$. The claim is trivial for $k = 0$. So suppose it to hold for $k$. Then

$$\hat{\Gamma}^{k+1}\left(\overline{X}\right) \;=\; \hat{\Gamma}\left(\hat{\Gamma}^{k}\left(\overline{X}\right)\right) \;=\; \hat{\Gamma}\left(\overline{\Gamma^{k}(X)}\right) \;=\; \overline{\Gamma\left(\Gamma^{k}(X)\right)} \;=\; \overline{\Gamma^{k+1}(X)}.$$ □

Before pressing further, let's get a routine syntactical lemma out of the way. In §2.3, I sketched a method of expressing the ancestral of a relation, which relies on a trick for recycling bound variables. The following lemma generalizes that trick. Note that it relies on the presence of an equality predicate. I've found this to be the main obstacle to interpreting $\mathcal{LF}$ under Wehmeier's 2004 semantics of variables.

LEMMA 3.4. *Each formula $A$ of $\mathcal{LF}$ is equivalent to a formula of the form $A'[p/R\vec{x}]$ such that $R$ does not occur in $A'$, for any predicate $R$.*

*Proof.* Let $A$ be a formula of $\mathcal{LF}$. Then there are some sequences $\vec{t}_1, \ldots, \vec{t}_k$ of terms such that $A$ has the form $A'[p_1/R\vec{t}_1]\cdots[p_k/R\vec{t}_k]$, where $R$ does not occur in $A'$. Let $R_{\vec{x}}\vec{t}_i$ be the formula

$$\exists\vec{x}(t_{i,1} = x_1 \wedge \cdots \wedge t_{i,n} = x_n \wedge R\vec{x}),$$

where $x_1, \ldots, x_n = \vec{x}$ and $t_{i,1}, \ldots, t_{i,n} = \vec{t}_i$. Then, $R_{\vec{x}}\vec{t}_i$ is logically equivalent to $R\vec{t}_i$. So by Proposition 2.14, the formula $A$ is equivalent to $A'[p_1/R_{\vec{x}}\vec{t}_1]\cdots[p_k/R_{\vec{x}}\vec{t}_k]$. □

***3.2. Induction.*** Let's now turn to the main goal of this section: to clarify the relationship between form-series disjunctions and finitary inductive definitions. The essential idea, attributed by Barwise (1977) to Arthur Rubin, is that a finitary inductive definition can be regarded as a countably infinite disjunction which contains only finitely many variables. I'll aim to show that every finitary inductive definition can be expressed by a form-series disjunction. To this end, we need first to spell out what is a finitary inductive definition and what it is for a logic to express one.

Consider some examples.

(A) the empty set is hereditarily finite; every finite set of hereditarily finite sets is hereditarily finite.

(B) the empty set is hereditarily countable; every countable set of hereditarily countable sets is hereditarily countable.

Each of these clauses can be used to specify the totality of results of repeatedly applying a rule to some initially given objects. The rule determines an operation, $\Gamma$, which takes a class $X$ and returns the class $\Gamma(X)$ of all results of once applying the rule to any of its elements. The rule is applied again to $X \cup \Gamma(X)$, and so on. Write $\Gamma_\cup : X \mapsto X \cup \Gamma(X)$. Then the inductively specified totality should certainly include all elements of the classes $X, \Gamma_\cup(X), \Gamma_\cup(\Gamma_\cup(X)), \ldots$, each class being obtained by some finite number of applications of $\Gamma$. Consider the union of all those classes. Does it contain every result of applying the rule to its elements? Say that the induction is finitary if this is so. Thus, example (A) is finitary but (B) is not.

Now for some notation. Let $\mathrm{I}^0_\Gamma = \Gamma(\emptyset)$, let $\mathrm{I}^{\alpha+1}_\Gamma = \mathrm{I}^\alpha_\Gamma \cup \Gamma(\mathrm{I}^\alpha_\Gamma)$, and for $\lambda$ a limit ordinal let $\mathrm{I}^\lambda_\Gamma = \bigcup_{\alpha<\lambda} \mathrm{I}^\alpha_\Gamma$. If $\alpha$ is the least ordinal such that $\mathrm{I}^{\alpha+1}_\Gamma = \mathrm{I}^\alpha_\Gamma$, then write $\mathrm{I}_\Gamma = \mathrm{I}^\alpha_\Gamma$. Induction transforms the operator $\Gamma$ into the class $\mathrm{I}_\Gamma$. So the induction determined by $\Gamma$ is finitary if $\mathrm{I}_\Gamma = \mathrm{I}^\omega_\Gamma$.

The notation just introduced assumes induction always to take the empty class as a "base case". Of course this implies no loss of generality. Let $\Gamma_X(\emptyset) = X$ and $\Gamma_X(Y) = \Gamma(Y)$ otherwise. I'll sometimes write $\mathrm{I}^\alpha_\Gamma(X)$ for $\mathrm{I}^\alpha_{\Gamma_X}$.

PROPOSITION 3.5. *Suppose that $\Gamma$ is $\mathcal{L}$-definable. Then $\mathrm{I}^\omega_\Gamma$ is $\mathcal{LF}$-definable. And so is the operator $X \mapsto \mathrm{I}^\omega_\Gamma(X)$.*

*Proof.* Write $\Gamma = |A|^{R:\mathcal{M}}$, where the variables free in $A$ are $\vec{x}$. Also suppose that $A = A'[p/R\vec{x}]$, with $R$ not in $A'$. Clearly, $\left|(p \vee A')[p/R\vec{x}]\right|^{R:\mathcal{M}} = \Gamma_\cup$. So using Proposition 3.2,

$$\left|\bigvee[\![\bot, p, p \vee A']\!]\right|^{\mathcal{M}} = \bigcup_{k<\omega} \mathrm{I}^k_\Gamma = \mathrm{I}^\omega_\Gamma,$$

https://doi.org/10.1017/S1755020316000472 Published online by Cambridge University Press

where $\perp$ is any unsatisfiable $\mathcal{L}$-formula. For the second part, take the defining formula to be $\bigvee[\![R\vec{x}, p, p \vee A']\!]$. $\square$

As a first application of Proposition 3.5, let's reconsider the ancestral. Suppose that $X$ is a two-place relation. Following Russell (1903, 210), let's write $X^1 = X$, and write $X^{k+1}$ for the class of all $(a, b)$ such that $(a, c) \in X^k$ and $(c, b) \in X$ for some $c$. The ancestral can now be interpreted as an operator $\Gamma^* : X \mapsto \bigcup_{0<k<\omega} X^k$.

PROPOSITION 3.6. *In the logic $\mathcal{LF}$, the ancestral operator is definable.*

*Proof.* Let $A$ be the formula $Rxy \vee \exists z(Rxz \wedge Rzy)$. Then $|A|^{R:\mathcal{M}} : \bigcup_{i=1}^{k} X^i \mapsto \bigcup_{i=1}^{2k} X^i$. So $\mathrm{I}^{\omega}_{|A|^{R:\mathcal{M}}}(X) = \bigcup_{k<\omega} X^k = \Gamma^*(X)$. By Proposition 3.5, it follows that there is a formula, call it *Ancestral*, such that $|\mathit{Ancestral}|^{R:\mathcal{M}} = \Gamma^*$. $\square$

Toward a second example, let's develop an analysis of arithmetical concepts within $\mathcal{LF}$. For Wittgenstein unlike for Frege and Russell, this will not be a matter of specifying relations between distinctively numerical objects. Rather, arithmetical concepts are techniques for reporting factual relationships (see *Tractatus* 6.2ff). Accordingly, let's now build a method of arithmetical reporting.

Let $X$ be an ordinary two-place relation on $\mathcal{D}$. Then $X$ can be understood to determine counterparts of arbitary arithmetical relations. For example, consider the one-place arithmetical relation which holds just of the natural number $k$. To this there will correspond a two-place relation on $\mathcal{D}$ which holds of just those $a, b \in \mathcal{D}$ such that $(a, b) \in X^k$. Similarly, if $\rho$ is the arithmetical relation which holds of just those $j, k, l$ such that $l = j + k$, then the choice of $X$ determines as corresponding to $\rho$ the four-place relation on $\mathcal{D}$ which holds of just those $a, b, c, d$ such that $(a, b) \in X^j$, $(a, c) \in X^k$, and $(a, d) \in X^{j+k}$.

Now since numbers are not objects, arithmetical relations are not relations on objects. Therefore no particular choice of material relation $X$ belongs in the analysis of arithmetic. Instead, an arbitrary $k$-place arithmetical relation $\rho$ will be analyzed as an operator $\Gamma^\rho$ of type $2, k + 1$:

$$\Gamma^\rho(X) = \{(a, b_1, \ldots, b_k) : (a,b_1) \in X^{n_1}, \ldots, (a, b_k) \in X^{n_k} \text{ for some } (n_1, \ldots, n_k) \in \rho\}. \tag{9}$$

Specifically, addition and multiplication determine operators $\Gamma^+$ and $\Gamma^\times$ of type 2, 4.

PROPOSITION 3.7. *The addition and multiplication operators are $\mathcal{LF}$-definable.*

*Proof.* In the case of addition, the idea is to show how to build a minimal four-place relation between initial object, left and right summands, and sum, so that (i) any initial object has itself as a sum of itself with itself and (ii) any successor of a sum of left and right summands is a sum both of the left summand with the successor of the right, and of the right summand with the successor of the left.

Let $R, S$ be two- and four-place predicates not in the signature of $\mathcal{M}$. For the base case, let $A$ be the formula $w = x = y = z$. For the induction step, let $B$ be the formula $\exists y \exists z(Swxyz \wedge Ryy_1 \wedge Rzz_1)$. Then

$$|B|^{S:\mathcal{M},X} : Y \mapsto \{(a, b, c_1, d_1) : (a, b, c, d) \in Y \text{ for some } c, d \text{ with } (c, c_1) \in X \text{ and } (d, d_1) \in X\}.$$

Now similarly, let $C$ be an $\mathcal{L}$-formula which describes the effect of at once taking successor in the left summand and the sum. Then

$$\mathrm{I}^{\omega}_{|B \vee C|^{S:\mathcal{M},X}}(|A|^{\mathcal{M}}) = \Gamma^+(X).$$

https://doi.org/10.1017/S1755020316000472 Published online by Cambridge University Press

By Proposition 3.5, it follows that in the signature of an expansion of $\mathcal{M}$ to $R$, there is a formula *Plus* such that $|Plus|^{R:\mathcal{M}} = \Gamma^{+}$.

Let's now turn to multiplication. In handling the successor steps, the strategy will be to refer to addition by passing the buck to the base case. More precisely, I'll prove the definability of an operator $\Gamma^{+\times}$ of type 2, 7 such that

$$\Gamma^{+\times} : X \mapsto \{(a, b, c, d, e, f, g) : (a, b, c, d) \in \Gamma^{+}(X) \wedge (a, e, f, g) \in \Gamma^{\times}(X)\}.$$

For the base case let $A$ be an $\mathcal{LF}$-formula, in the signature of an expansion of $\mathcal{M}$ to a new dyadic predicate $R$, such that

$$|A|^{\mathcal{M},X} = \{(a, b, c, d, a, a, a) : (a, b, c, d) \in |Plus|^{\mathcal{M},X}\}.$$

Now let $S$ be a new seven-place predicate. Then there is a first-order formula $B$ in the signature of an expansion of $\mathcal{M}$ to $R$, $S$ such that

$$|B|^{S:\mathcal{M},X} : Y \mapsto \{(a, b, c, d, e, f_1, g_1) : (a,b, c, d, e, f, g) \in Y \text{ for some } f, g \\ \text{with } (f, f_1) \in X \text{ and } (a, g, e, g_1, a, a, a) \in Y\}.$$

Similarly, let $C$ be a formula which specifies the effect of taking successor in the left multiplicand rather than in the right. Then

$$\mathrm{I}^{\omega}_{|B\vee C|^{S:\mathcal{M},X}}(|A|^{\mathcal{M}}) = \Gamma^{+\times}(X).$$

By Proposition 3.5 there is, in the signature of an expansion of $\mathcal{M}$ to $R$, a formula *PlusTimes* such that $|PlusTimes|^{R:\mathcal{M}} = \Gamma^{+\times}$. Existentially generalizing out the three addition placeholders of *PlusTimes* gives a formula *Times* such that $|Times|^{R:\mathcal{M}} = \Gamma^{\times}$, as desired. □

PROPOSITION 3.8. *If $\rho$ is an arithmetically definable relation on the natural numbers, then $\Gamma^{\rho}$ is an $\mathcal{LF}$-definable operator.*

*Proof.* It suffices to show that for any arithmetically definable relation $\rho$, there is an operator $\Gamma^{\rho}$ satisfying (9) above. We argue by induction on the complexity of formulas in a language of arithmetic with signature $0,', +, \times$, where $', +, \times$ are treated as relations.

The proof is trivial. Let *WeakAncestral* be the formula $x = y \vee Ancestral$. Let $\phi^{\star}$ be the image of $\phi$ under the translation

- $0(x) \mapsto x = w$; $'(x, y) \mapsto Rxy$; $+(x, y, z) \mapsto Plus[w, x, y, z]$; $\times(x, y, z) \mapsto Times[w, x, y, z]$;
- $x = y \mapsto x = y$; $\neg\phi \mapsto \neg\phi^{\star}$; $\phi \vee \psi \mapsto \phi^{\star} \vee \psi^{\star}$; $\exists x\phi \mapsto \exists x(WeakAncestral[w, x] \wedge \phi^{\star})$.

Now define $A^{\phi}$ to be the formula $WeakAncestral[w, x_1] \wedge \cdots \wedge WeakAncestral[w, x_k] \wedge \phi^{\star}$, where $x_1, \ldots, x_k$ are the variables free in $\phi$. If $\phi$ defines $\rho$ over a standard model of arithmetic, then $\left|A^{\phi}\right|^{R:\mathcal{M}} = \Gamma^{\rho}$. □

From a given relation $X$, the addition and multiplication operators determine relations which look only as much like addition and multiplication as $X$ looks like successor. For example, if $X$ looks like the successor relation on the hours of a clock, then the corresponding addition and multiplication relations look like those of arithmetic modulo twelve. However, in the odd event that $\omega$ itself happens to be lying around somewhere, then that can certainly be reported. The proposal that Wittgenstein's form-series operator ought to yield a categorical axiomatization of arithmetic is due to Goldfarb (2012).

PROPOSITION 3.9. *Suppose that the domain of $\mathcal{M}$ is infinite, and that $\mathcal{M}$ contains a dyadic predicate. There is a single formula $A$ of $\mathcal{LF}$ such that $\mathcal{M} \models A$ iff $\mathcal{M}$ contains a corresponding copy of the natural numbers under the successor relation. Moreover, if $\mathcal{M} \models A$, then any counterpart of an arithmetically definable relation is definable on $\mathcal{M}$.*

*Proof.* Immediate from Propositions 3.6 and 3.8. □

After that foray into applications, let's return to the general theory. The operator of an inductive definition is often specified by reference to a previously inductively defined concept: in this way, for example, the usual definition of multiplication invokes the concept of addition. But the syntax of $\mathcal{LF}$ does not allow operator-signs to contain form-series expressions. So it might be wondered whether this technique can always be simulated in $\mathcal{LF}$. In the proof of Proposition 3.7, we saw that multiplication can be defined from addition by folding the concept of addition into a second-order parameter of the base case. Let me conclude this section by sketching a result to the effect that the method is general. Any $\mathcal{LF}$-defined relations can be treated like primitives of the underlying structure.

PROPOSITION 3.10. *Suppose that $\Gamma$ is $\mathcal{L}$-definable over $\mathcal{M}, Y_1, \ldots, Y_n$ and that $Y_1, \ldots, Y_n$ are $\mathcal{LF}$-definable over $\mathcal{M}$. Then $\mathrm{I}^{\omega}_{\Gamma}$ is $\mathcal{LF}$-definable over $\mathcal{M}$.*

*Proof.* By hypothesis, there is an $\mathcal{L}$-formula $A$ such that $|A|^{R:\mathcal{M},Y_1,\ldots,Y_k} = \Gamma$; and there are $\mathcal{LF}$-formulas $B_1, \ldots, B_n$ with $|B_i|^{\mathcal{M}} = Y_i$ for each of the $B_i$. We may assume $A$ to have the form $A'[p/R\vec{x}, q_1/S_1\vec{y}_1, \ldots, q_n/S_n\vec{y}_n]$, where $A'$ is an $\mathcal{L}$-formula containing none of $R, S_1, \ldots, S_n$. Indeed, suppose that the free variables of $A$ are precisely $\vec{x}$, and that the free variables of each $B_i$ are $\vec{y}_i$, with the $\vec{y}_i$ distinct from all $\vec{x}$ and from all $\vec{y}_j$ with $j \neq i$.

For any relations $V, W$, write $V \times W$ for set of all $(\vec{a}, \vec{b})$ with $(\vec{a}) \in V$ and $(\vec{b}) \in W$. And write $\Pi_i V_i$ for $V_1 \times V_2 \times \ldots$. Now it would be nice to define in $\mathcal{L}$ an operator satisfying

$$\mathrm{I}^{k}_{\Gamma} \times \Pi_i Y_i \quad \mapsto \quad \mathrm{I}^{k+1}_{\Gamma} \times \Pi_i Y_i \tag{10}$$

for all $k$. Sadly, such an operator will not in general exist unless $\mathrm{I}^{k}_{\Gamma}$ and each of the $Y_i = |B|^{\mathcal{M}}$ are nonempty. So, the construction must be split into cases.

Note that $\mathrm{I}^{\omega}_{\Gamma} = \mathrm{I}^{\omega}_{\Gamma_{\cup}}$. So we can assume that $\Gamma = \Gamma_{\cup}$. We can at once get a trivial case out of the way, namely where $\Gamma(\emptyset) = \emptyset$. For in this case, $\mathrm{I}^{\omega}_{\Gamma} = \emptyset$, and so $|\bot|^{\mathcal{M}} = \mathrm{I}^{\omega}_{\Gamma}$.

To motivate the handling of the other cases, let's first work the end of the proof. Let $e$ be a selection $e_1, \ldots, e_m$ of numbers from $1, \ldots, n$. For every such $e$, we'll aim to have constructed a formula $D_e$ such that if it is precisely $B_{e_1}, \ldots, B_{e_m}$ amongst $B_1, \ldots, B_n$ which have nonempty extension, then $|D_e|^{\mathcal{M}} = \mathrm{I}^{\omega}_{\Gamma}$. Now let $E_e$ be a formula which says that it is precisely the $B_{e_1}, \ldots, B_{e_m}$ with nonempty extension. Take $C_e$ to be the formula $E_e \rightarrow D_e$. For the trivial case, we can choose $C_0$ to be the formula $\neg\exists\vec{x}A'[p/\bot, q_1/B_1, \ldots, q_n/B_n] \rightarrow \bot$. Finally, let $C$ be the conjunction of $C_0$ with each of the $2^n$ formulas $C_e$. The construction of the $D_e$ will then give the desired result that $|C|^{\mathcal{M}} = \mathrm{I}^{\omega}_{\Gamma}$.

It remains to build each formula $D_e$, under the hypothesis that $\Gamma(\emptyset)$ and precisely $Y_{e_1}, \ldots, Y_{e_m}$ are nonempty. Write $\vec{y}_e$ for $\vec{y}_{e_1}, \ldots, \vec{y}_{e_m}$. Let $T_e$ be a new predicate whose arity is the length of $\vec{x}, \vec{y}_e$. Now for $0 \leq j \leq n$, let

$$G_{e,j} = \begin{cases} \exists\vec{y}_e T_e\vec{x}\vec{y}_e, \text{ if } j = 0; \\ \exists\vec{x}\vec{y}_{e_1}\cdots\vec{y}_{e_{i-1}}\vec{y}_{e_{i+1}}\cdots\vec{y}_{e_m} T_e\vec{x}\vec{y}_e, \text{ if } j = e_i \\ \bot, \text{ otherwise.} \end{cases}$$

Then $|G_{e,0}|^{\mathcal{M},\Pi_i Y_{e_i}} = \Gamma(\emptyset)$, and $|G_{e,j}|^{\mathcal{M},\Pi_i Y_{e_i}} = Y_j$ if $1 \leq j \leq n$.

Let $F_e$ be the formula $A'[p/G_{e,0}, q_1/G_{e,1}, \ldots, q_n/G_{e,n}] \wedge G_{e,e_1} \wedge \cdots \wedge G_{e,e_m}$. Then $F_e$ defines over $\mathcal{M}$ an approximation of the operator (10), so that

$$\mathrm{I}^{\omega}_{|F_e|^{T_e:\mathcal{M}}} : \Gamma(\emptyset) \times \Pi_i Y_{e_i} \;\mapsto\; \mathrm{I}^{\omega}_{\Gamma} \times \Pi_i Y_{e_i}. \tag{11}$$

Since $F_e$ is a formula of $\mathcal{L}$, we can apply the induction of Proposition 3.5 to define the operator mentioned in (11). Using $A'$ and the $B_i$, it is also clear that $\Gamma(\emptyset) \times \Pi_i Y_{e_i}$ is $\mathcal{LF}$-definable over $\mathcal{M}$. So, the relation $\mathrm{I}^{\omega}_{\Gamma} \times \Pi_i Y_{e_i}$ is itself definable over $\mathcal{M}$. Existentially generalizing out the $\vec{y}_e$ yields the formula $D_e$ desired for the case in question, namely such that $|D_e|^{\mathcal{M}} = \mathrm{I}^{\omega}_{\Gamma}$. □

***3.3. Comparisons.*** Let's now zoom way out, and compare $\mathcal{LF}$ with some other extensions of first-order logic. First, it is obvious from the double-bar semantics that $\mathcal{LF}$ is a subsystem of the infinitary logic $\mathcal{L}_{\omega_1\omega}$ which results by adding to first-order logic countably infinite disjunctions. However, even $\mathcal{L}_{\infty\omega}$ has no wellfoundedness quantifier (Lopez-Escobar, 1966). Indeed, $\mathcal{LF}$ is a subsystem of the fragment $\mathcal{L}^{\omega}_{\omega_1\omega}$ in which no formula contains infinitely many variables (Barwise, 1977); and $\mathcal{L}^{\omega}_{\omega_1\omega}$ lacks a quantifier "there are infinitely many". The system $\mathcal{LF}$ is also related to some subsystems of second-order logic. For example, the $\Pi^1_1$ fragment of second-order logic ($\Pi^1_1$*SOL*, as defined in Heck (2011)) expresses finitary inductive definitions, so it is at least as expressive as $\mathcal{LF}$. But it also has an infinity quantifier, so it is strictly more expressive. The situation is different with monadic second-order logic (*MSOL*), which *a priori* expresses finitary inductive definitions only of monadic properties. And indeed, since the monadic second-order theory of the successor relation is decidable (Büchi, 1960), it follows that unlike $\mathcal{LF}$, monadic second-order logic cannot define both addition and multiplication from successor. Conversely, the monadic fragment, like the $\Pi^1_1$ fragment, does have a wellfoundedness quantifier. Thus monadic second-order logic and $\mathcal{LF}$ are incomparable. Strictly weaker than all of these systems is the result $\mathcal{LA}$ of adding an ancestral operator to first-order logic. In sum, we have

PROPOSITION 3.11. *Write $X \to Y$ to mean that logic $Y$ is strictly more expressive than logic $X$. Then*

$$\begin{tikzcd}
\mathcal{L}_{\omega_1\omega} & & \Pi^1_1 SOL & & MSOL \\
\mathcal{L}^{\omega}_{\omega_1\omega} \arrow[u] & \mathcal{LQ}_{\infty} \arrow[ul] \arrow[ur] & & \mathcal{LQ}_{\mathrm{WF}} \arrow[ul] \arrow[ur] & \\
 & & \mathcal{LF} \arrow[uu] \arrow[ull] & & \\
 & & & & \mathcal{LA}. \arrow[ull] \arrow[uuu]
\end{tikzcd}$$

**§4. Definability.** One way to characterize the consequence relation in first-order logic is through a universal generalization over structures. Since a structure consists of a maybe infinite collection of objects plus some relations on the collection, this analysis makes the consequence relation look unmanageable: to verify that the relation obtains, it looks as though we'd have to run through an infinity of in general infinite structures and determine that each is not a countermodel. However, first-order logic admits a complete notion of proof: accordingly, whenever some formula is a first-order consequence of some others, some finite pattern of formulas gives an effective witness. So, it turns out to suffice instead to enumerate the finite patterns of formulas until a proof appears. This collapse in the complexity of the consequence relation is a fairly special property of first-order logic. Slight enrichments of the logic tend to complicate the consequence relation and outrun any system of finite witnesses.

In this section, we turn to the problem of characterizing the complexity-theoretic effects of the Russellian semantics and of the form-series device. On the one hand, the Russellian constraint winnows the class of countermodels, so that the answers to old questions may change. On the other hand, the form-series device introduces new formulas, hence raising more questions.

In §4.1, I'll develop an appropriate framework for measuring the complexity of metalogical concepts of $\mathcal{LF}$. In §4.2 and §4.3, I'll respectively consider the cases in which the domain has been chosen to be some fixed finite or infinite collection. Finally, §4.4 will address the complexity of $\mathcal{D}$-validity and $\mathcal{D}$-consequence considered prior to a choice of $\mathcal{D}$.

***4.1. Measuring definitions.*** Let's start with a simple framework for measuring the complexity of logical notions, which we can then apply to the notions of $\mathcal{D}$-validity and $\mathcal{D}$-consequence for $\mathcal{L}$ and $\mathcal{LF}$. A standard measure of logical complexity derives from the theory of computable functions on the natural numbers. Its application to logic depends on the technique of arithmetization, according to which formulas are "coded" as natural numbers. In the present context, arithmetization is somewhat annoying, since it prejudges nontrivial interpretive questions. First, it applies only if the signature is countable. Since a Russellian signature includes the domain of its Russellian structures, this excludes structures with uncountable domains. Second, arithmetization yields the definability of, e.g., the class of names coded by the even integers. This leads to troubles of the sort indicated in §1.2.

Sundholm (1992, 71) recommends an approach to this problem derived from Barwise (1975, 78ff): code syntactical constructions not arithmetically but set-theoretically, treating logical vocabulary as pure sets, and the terms of the signature as urelements. In a little more detail, suppose $\mathcal{S}$ is a structure. Let HF($|\mathcal{S}|$) be the class of hereditarily finite sets over the domain $|\mathcal{S}|$ of $\mathcal{S}$; thus the elements of HF($|\mathcal{S}|$) are generated, given the elements of $|\mathcal{S}|$ initially, by repeatedly forming all finite sets of what's obtained already. This determines a first-order structure $\mathbb{HF}(\mathcal{S})$ whose domain is HF($|\mathcal{S}|$), together with the natural membership relation on HF($|\mathcal{S}|$), the property of belonging to the domain $|\mathcal{S}|$ of urelements, and each of the relations and functions baked into $\mathcal{S}$ itself. I will write HF and $\mathbb{HF}$ for HF($\emptyset$) and $\mathbb{HF}(\emptyset)$ respectively.

We'll be specially concerned with $\mathbb{HF}(\mathcal{S})$ for $\mathcal{S}$ not any old structure, but a Russellian signature considered as a structure. The class $\mathbb{HF}(\mathcal{S})$ will be regarded as the universe of possible syntactical constructions from the initially given assortment of nonlogical vocabulary. For simplicity of coding, I will just identify the signature with its collection $\mathcal{D}$ of names. This implies no serious loss of generality. For one thing, a signature with countably many predicates can be interpreted as a signature with at most one predicate of each arity: take the $j$th predicate of arity $k$ to be the predicate whose arity is the numerical code of the pair $j, k$. But second, a language with at most one predicate of each arity can take an atomic formula to be simply a sequence of terms, so doesn't need predicates. We will therefore work over $\mathbb{HF}(\mathcal{D})$, where the underlying structure $\mathcal{D}$ is a bare class of urelements. Logical vocabulary should be coded by pure sets. Atomic formulas are coded in $\mathbb{HF}(\mathcal{D})$ as

https://doi.org/10.1017/S1755020316000472 Published online by Cambridge University Press

finite sequences of variables and of elements of $\mathcal{D}$. Nonatomic formulas result from atomic formulas through this or that finitary set-theoretic construction.

The structure $\mathbb{HF}(\mathcal{D})$ is of course itself a structure for a first-order language whose two nonlogical predicates are those of membership and urelementhood, or $\in$ and D. The complexity of a class on $\mathbb{HF}(\mathcal{D})$ can now be measured by the logical complexity of the simplest formulas which define it. I'll just sketch the portion of the framework we'll need; the details can be found in Barwise (1975). A formula is $\Delta_0$ if it is built up from atomic formulas by negation, disjunction, and bounded existential quantification $\exists x \in y \ldots$. A formula is $\Sigma_1$ (or $\Pi_1$) if it's the result of prefixing a $\Delta_0$ formula with a sequence of existential (universal) quantifiers; also the result of prefixing a $\Pi_n$ (or $\Sigma_n$) formula with a string of existential (universal) quantifiers is said to be $\Sigma_{n+1}$ ($\Pi_{n+1}$). Now the complexity of a subclass of $\mathbb{HF}(\mathcal{D})$ is given by the complexity of the simplest formula which defines it. A class is said to be $\Delta_n$ if it is both $\Sigma_n$ and $\Pi_n$.

The definability-theoretic characterization of complexity of classes of hereditarily finite sets generalizes naturally the recursion-theoretic measures of complexity on the natural numbers. Note that each natural number is uniquely represented as a sum of powers of two, and so it can be taken to code the set of whatever is coded by the powers (Ackermann, 1937). This gives a bijection between the naturals and HF. It's now straightforward to verify that a set of integers is recursive iff it is $\Delta_1$ on $\mathbb{HF}$ (Barwise, 1975, 47ff), likewise for r.e. and $\Sigma_1$, and so on. As Kirby (2009) argues, $\mathbb{HF}$ can therefore be seen as a natural home of finitary constructions.

Moreover, the extension of $\mathbb{HF}$ to $\mathbb{HF}(\mathcal{D})$ preserves the alignment of definability-theoretic and computability-theoretic classifications. In particular, the introduction of $\mathcal{D}$ does not affect first-order definability-theoretic complexity of pure classes.

PROPOSITION 4.1. *Suppose $X \subseteq$ HF. Then $X$ is $\Sigma_n$ on $\mathbb{HF}$ iff $X$ is $\Sigma_n$ on $\mathbb{HF}(\mathcal{D})$.*

*Proof.* In one direction, note that the class HF is $\Sigma_1$ on $\mathbb{HF}(\mathcal{D})$. So if $\phi$ is a $\Sigma_n$ definition of $X$ on $\mathbb{HF}$, then the relativization of $\phi$ to HF is a $\Sigma_n$ definition of $X$ on $\mathbb{HF}(\mathcal{D})$.

Conversely, the downward Löwenheim-Skolem theorem implies that $\mathbb{HF}(\mathcal{D})$ has a countable elementary substructure $\mathbb{HF}(\mathcal{D})'$. Let $\mathcal{D}'$ be the least subclass of $\mathcal{D}$ such that each element of the domain of $\mathbb{HF}(\mathcal{D})'$ belongs to HF$(\mathcal{D}')$. Each element of HF$(\mathcal{D}')$ is definable on $\mathbb{HF}(\mathcal{D})$ by a first-order formula with parameters in $\mathcal{D}'$. So in fact, HF$(\mathcal{D})' =$ HF$(\mathcal{D}')$. So any $\mathbb{HF}(\mathcal{D})$ has a countable elementary substructure $\mathbb{HF}(\mathcal{D}')$. We may therefore assume that $\mathcal{D}$ is countable.

Let $p$ be a bijection from $\mathcal{D}$ onto the class of pure sets of the form $(0, a)$. Define $f$ : HF$(\mathcal{D}) \to$ HF so that $f(a) = p(a)$ for $a \in \mathcal{D}$, and otherwise $f(a) = (1, \{f(b) : b \in a\})$. Let HF$(\mathcal{D})^* = \{f(a) : a \in$ HF$(\mathcal{D})\}$; let $\mathcal{D}^* = \{p(a) : a \in \mathcal{D}\}$, and let $a \in^* b$ iff $b = (1, c)$ for some $c$ such that $f(a) \in c$. Then $f$ is an isomorphism of $\mathbb{HF}(\mathcal{D})$ onto (HF$(\mathcal{D})^*, \mathcal{D}^*, \in^*$), while $\mathcal{D}^*, \in^*$ are $\Sigma_1$ on $\mathbb{HF}$. Now, suppose that $\phi$ is a $\Sigma_n$ definition of $X$ on $\mathbb{HF}(\mathcal{D})$. Let $\phi^*$ be the relativization to HF$(\mathcal{D})^*$ of the result of replacing $\in$ and D in $\phi$ with the definitions of $\in^*$ and $\mathcal{D}^*$. Then $\phi^*$ is a $\Sigma_n$ definition of $X^*$ on $\mathbb{HF}(\mathcal{D})$. Moreover, the restriction of $f$ to HF is $\Sigma_1$ on $\mathbb{HF}$. Therefore, $\phi^*[f(x)]$ is a $\Sigma_n$ definition of $X$ on $\mathbb{HF}$. □

***4.2. The finite case.*** Let's now apply this framework to analyze the metatheory of $\mathcal{L}$ and $\mathcal{LF}$. The first couple of results should be reassuring.

PROPOSITION 4.2. *Suppose $\mathcal{D}$ is finite. Then $\mathcal{D}$-validity for $\mathcal{LF}$ is $\Delta_1$ on $\mathbb{HF}(\mathcal{D})$, and $\mathcal{D}$-consequence for $\mathcal{L}$ is $\Sigma_1$.*

https://doi.org/10.1017/S1755020316000472 Published online by Cambridge University Press

*Proof.* Consider the infinitary expansion $\overline{\overline{A}}$ of a formula $A$. Let $B$ be an unquantified infinitary formula which results by successively replacing each existentially quantified subformula of $\overline{\overline{A}}$ with the finite disjunction of its $\mathcal{D}$-instances. Now consider an infinite disjunctive subformula $C$ of $B$; suppose by induction that all subformulas of $C$ are finitary. Then $C$ has the form $\bigvee(D, E[p/D], E[p/E[p/d]], \ldots)$, where $D$ and $E$ are constructed from some $k$ atomic formulas by negation and finite disjunction, though $E$ itself may also contain $p$. By extensionality, Proposition 2.14, $C$ must be equivalent to the disjunction of its first $2^{2^k}$ disjuncts, and we can just drop the rest. Let $A^{\downarrow}$ be the formula, constructed from atomic sentences by negation and finite disjunction, which results by eliminating from $B$ in this way each of its infinitary subformulas. Clearly there is a $\Sigma_1$-definable function on $\mathbb{HF}(\mathcal{D})$ associating $A^{\downarrow}$ to $A$. But validity for finitary formulas is $\Delta_1$.

If $\mathcal{D}$ is finite, then $\mathcal{D}$-consequence is compact. So, $A$ is an $\mathcal{D}$-consequence of $X$ iff there's a conjunction $B$ of elements of $X$ such that $B \to A$ is valid; this can be expressed as a $\Sigma_1$ formula on $\mathbb{HF}(\mathcal{D})$. □

Let's now drop the assumption that $\mathcal{D}$ is finite. The second reassuring result is that, if we drop form-series from the logic, and consider only the notion of $\mathcal{D}$-validity, then this is no more complicated than usual.

PROPOSITION 4.3. *$\mathcal{D}$-validity for $\mathcal{L}$ is $\Sigma_1$ on $\mathbb{HF}(\mathcal{D})$, for all $\mathcal{D}$.*

*Proof.* It suffices to consider the case where $\mathcal{D}$ is infinite. Let $I$ be a collection of $\mathcal{L}$-formulas to the effect that "there is at least one thing, there are at least two things, ...". For $A$ a formula of $\mathcal{L}$, Let $T_A$ be an $\mathcal{L}$-formula $a \neq b \wedge a \neq c \wedge \ldots$ to the effect that no two names in $A$ denote the same thing.

We now argue that $A$ is $\mathcal{D}$-valid iff $A$ is a classical consequence of $I \cup T_A$. In one direction, suppose that $A$ is not $\mathcal{D}$-valid, so that $\mathcal{M} \not\models A$ for some $\mathcal{D}$-structure $\mathcal{M}$. Clearly $\mathcal{M} \models T_A$ since $A$ is a $\mathcal{D}$-structure, and $\mathcal{M} \models I$ since $\mathcal{D}$ is infinite. Hence $A$ is not a classical consequence of $I \cup T_A$. Conversely, suppose that there's a classical structure $\mathcal{M}$ such that $\mathcal{M} \models I \cup T_A$ but $\mathcal{M} \not\models A$, where the signature of $\mathcal{M}$ consists just of the nonlogical vocabulary of $A$. Since $\mathcal{M} \models T_A$, it follows by the Lowenheim-Skolem theorems that $\mathcal{M}$ is elementarily equivalent to a structure $\mathcal{M}'$ whose domain is $\mathcal{D}$. Since $\mathcal{M}' \models I$, therefore $\mathcal{M}'$ is isomorphic to an $\mathcal{M}''$ such that $a^{\mathcal{M}''} = a$ for each constant $a$ which occurs in $A$. In turn, $\mathcal{M}''$ may be expanded to an $\mathcal{D}$-structure $\mathcal{M}'''$ such that $\mathcal{M}''' \not\models A$.

The completeness theorem for first-order logic implies that $A$ is a classical consequence of $I \cup T_A$ iff there is a proof of $A$ from $I \cup T_A$. Formally, however, the property "being a proof of $A$ from $I \cup T_A$" is $\Delta_1$ on $\mathbb{HF}(\mathcal{D})$, so that "having a proof from $T \cup I_A$" is $\Sigma_1$. From the previous paragraph, it follows that the collection of $\mathcal{D}$-valid formulas is $\Sigma_1$ on $\mathbb{HF}(\mathcal{D})$. □

So, $\mathcal{D}$-validity is never more complicated than classical validity. Does this also hold for $\mathcal{D}$-consequence? By the compactness theorem, classical consequence is no more complicated than classical validity. But if $\mathcal{D}$ is infinite, then $\mathcal{D}$-consequence is clearly not compact. For example, a universal generalization is a $\mathcal{D}$-consequence of the set of its instances, but not of any finite subset. In other words, the complexity of the $\mathcal{D}$-consequence relation remains to be determined.

***4.3. Countability and categoricity.*** Let's now fix the domain to be infinite. What, then, is the complexity of $\mathcal{D}$-validity and of $\mathcal{D}$-consequence for $\mathcal{LF}$ in this case? Recall,

https://doi.org/10.1017/S1755020316000472 Published online by Cambridge University Press

from §2.2 and §3.2, that the Russellian constraint and the form-series device each lead to categorical axiomatizations of the standard model of arithmetic. I'll now show that $\mathcal{LF}$ categorically axiomatizes another infinite structure as well: $\mathbb{HF}(\mathcal{D})$ itself. This leads to an amusing application of Tarski's theorem, which shows that validity is not first-order definable on $\mathbb{HF}(\mathcal{D})$. After establishing the definability of satisfaction relative to truth, we then conclude that in fact $\mathcal{LF}$-validity is $\Pi^1_1$-complete. Finally I'll show that $\mathcal{L}$-consequence is $\Pi^1_1$-complete as well.

Let $\mathcal{M}$ be a relational structure without constants. Now, say that $\mathcal{M}$ is $\mathcal{D}$-axiomatized up to isomorphism by some formulas $X$ iff the formulas $X$ are true in some $\mathcal{D}$-structure, and if by an addition of constants, $\mathcal{M}$ can be expanded to a structure which is isomorphic to every $\mathcal{D}$-structure satisfying $X$.

PROPOSITION 4.4. *Suppose $\mathcal{D}$ is infinite. Then $\mathbb{HF}(\mathcal{D})$ is $\mathcal{D}$-axiomatizable up to isomorphism by a single formula of $\mathcal{LF}$.*

*Proof.* We want to find a formula $Z$ such that $\mathcal{M} \models Z$ iff $\mathcal{M}$ is isomorphic to $\mathbb{HF}(\mathcal{D})$, for all $\mathcal{D}$-structures $\mathcal{M}$. Write $D$ and $E$ for a monadic and dyadic predicate of $\mathcal{L}$; these can serve, inside $\mathcal{L}$ or $\mathcal{LF}$, as counterparts of the predicates "$x$ is an urelement" and "$x$ belongs to $y$".

Note that $\mathrm{HF}(\mathcal{D})$ can be seen as the smallest set which contains the empty set and all urelements, and which is closed under the procedure $x, y \mapsto x \cup \{y\}$ of adjoining $y$ to $x$. This suggests that the class of $\mathcal{D}$-structures which are isomorphic to $\mathbb{HF}(\mathcal{D})$ can be defined by a single formula of $\mathcal{LF}$. Using the predicate $E$ for the counterpart of membership, it's straightforward to express the graph of the corresponding procedure of adjunction:

$$Adj = \forall w(Ewz \leftrightarrow Ewx \vee w = y).$$

Let $F$ be a new monadic predicate. Let

$$ClAdj = \exists x \exists y(Fx \wedge Fy \wedge Adj)$$

and let $\Gamma = |ClAdj|^{F:\mathcal{M}}$. Let $M_0 = |\neg\exists y Exy|^{\mathcal{M}}$. By Proposition 3.5, let $Z'$ be a formula in the one free variable $x$ such that $|Z'|^{\mathcal{M}} = \mathrm{I}^{\omega}_{\Gamma}(M_0)$. And let $Z_1 = \forall x Z'$. Thus, $Z_1$ is true in just those $\mathcal{M}$ where everything results from things without elements$^{\mathcal{M}}$ by repeated adjunction$^{\mathcal{M}}$. Furthermore, let

$$\begin{aligned}
Z_2 &= \forall x \forall y \exists! z Adj[x, y, z]) \\
Z_3 &= \exists! x(\neg Dx \wedge \forall y \neg Eyx) \\
Z_4 &= \forall x(Dx \rightarrow \forall y \neg Eyx) \\
Z_5 &= \forall x \exists y(Dy \wedge \neg Eyx).
\end{aligned}$$

And finally let

$$Z = Z_1 \wedge Z_2 \wedge Z_3 \wedge Z_4 \wedge Z_5.$$

It remains to show that $\mathcal{M} \models Z$ iff $\mathcal{M}$ is isomorphic to $\mathbb{HF}(\mathcal{D})$. In one direction, it is clear that any $\mathcal{D}$-structure isomorphic to $\mathbb{HF}(\mathcal{D})$ satisfies $Z$.

Conversely, suppose that $\mathcal{M} \models Z$. We want to construct an isomorphism from $\mathcal{M}$ onto $\mathbb{HF}(\mathcal{D})$. Since $\mathcal{M}$ is a $\mathcal{D}$-structure, its domain is just $\mathcal{D}$.

By $Z_3$ and $Z_4$, the class $M_0$ consists of the stuff in $D^{\mathcal{M}}$ plus exactly one thing $\emptyset^{\mathcal{M}}$ not in $D^{\mathcal{M}}$. Conditions $Z_1$ and $Z_2$ ensure that the cardinality of $\mathcal{D}$ differs by at most $\aleph_0$ from the cardinality of $D^{\mathcal{M}}$. Moreover if $\mathcal{D}$ is countable, then $Z_5$ together with $Z_2$ ensures that $D^{\mathcal{M}}$ is infinite and therefore also countable. So in any case, there is a bijection $f$ from $D^{\mathcal{M}}$ onto $\mathcal{D}$. For $a \in M_0$, let $h(a) = \emptyset$ if $a = \emptyset^{\mathcal{M}}$, and otherwise let $h(a) = f(a)$.

A couple of quick observations before extending $h$ from $M_0$ to all of $\mathcal{D}$. First: it is routine to show by induction on $k$ that if $a \in \mathrm{I}^k_\Gamma(M_0)$, then every $E^{\mathcal{M}}$-chain descending from $a$ has length at most $k$. Since $\mathcal{D} = \mathrm{I}^\omega_\Gamma(M_0)$, therefore $E^{\mathcal{M}}$ is wellfounded.

Second: the truth of $Z_2$ implies that adjunction$^{\mathcal{M}}$ is functional; write $a \fatsemi b$ for the $c$ such that $\mathcal{M} \models Adj[a, b, c]$. By $Z_1$, each $c \in \mathcal{D} - M_0$ can be written in the form $c = a \fatsemi b$ for some $a, b$. Now $Z_2$ says that there is exactly one thing whose elements$^{\mathcal{M}}$ are $b$ together with the elements$^{\mathcal{M}}$ of $a$. So, $c$ is determined by its elements$^{\mathcal{M}}$. But $c$ was arbitrary. So no two sets$^{\mathcal{M}}$ have the same elements$^{\mathcal{M}}$.

For $a \in \mathcal{D} - M_0$, we may now define $h(a) = \{h(b) : E^{\mathcal{M}}ba\}$. Let's argue that $h$ is an isomorphism of $\mathcal{M}$ onto $\mathbb{HF}(\mathcal{D})$. Suppose $s \in \mathrm{HF}(\mathcal{D})$. Clearly if $s = \emptyset$ or $s \in \mathcal{D}$, then $s$ is in the range of $h$. Otherwise $s = \{s_1, \ldots, s_k\}$; by induction on rank we may assume each $s_i = h(a_i)$ for some $a_i$, and so $s = h(\emptyset^{\mathcal{M}} \fatsemi a_1 \fatsemi \cdots \fatsemi a_k)$ where $\fatsemi$ associates to the left. Conversely, suppose that $h(a) = h(b)$ but $a \neq b$. Without loss of generality, we can assume there's a $c$ such that $E^{\mathcal{M}}ca$ but not $E^{\mathcal{M}}cb$ while $h(c) \in h(b)$, so that $h(c) = h(d)$ for some $d \neq c$, contradicting the wellfoundedness of $E^{\mathcal{M}}$. Since $h$ is a bijection, we must therefore also have that $h(b) \in h(a)$ iff $E^{\mathcal{M}}ba$. □

We're exploring the complexity of concepts of the metatheory of $\mathcal{L}$ and $\mathcal{LF}$. So far it has been enough simply to identify formulas of these object-languages with elements of a certain convenient structure, namely $\mathbb{HF}(\mathcal{D})$. At this point we'll need to begin making explicit reference to expressions of the formal metatheory. Indeed, it will also be useful to think of $\mathcal{L}$ as containing copies of the first-order expressions of the metalanguage.

To keep things clear, I will in the rest of this subsection never any longer use unadorned logical notation to refer to expressions of $\mathcal{L}$ or $\mathcal{LF}$. Whereas $\exists, \neg, \vee, (, )$, $x, y, \ldots, \in, \mathrm{D}\ldots$, etc., now refer to expressions of the metalanguage, only $\dot\exists, \dot\neg, \dot\vee, \dot(, \dot)$, $\dot x, \dot y, \ldots, \dot E, \dot D \ldots$ now refer to corresponding expressions of $\mathcal{L}$, which are understood as elements of $\mathbb{HF}(\mathcal{D})$. Where $\phi$ is a formula of the metalanguage, its $\mathcal{L}$-copy is $\dot\phi$: for example if $\phi$ is $\exists x(x \in y \rightarrow x \in z)$, then $\dot\phi$ is $\dot\exists\dot x\dot(\dot E\dot x\dot y\dot\rightarrow\dot x\dot\in\dot z\dot)$. We'll also consider second-order quantification in the metalanguage; the coding scheme must then be extended correspondingly. I'll continue to use uppercase italic letters informally to range over those elements of $\mathbb{HF}(\mathcal{D})$ which code formulas of $\mathcal{L}$ or of $\mathcal{LF}$. Finally, the notions of $\mathcal{D}$-validity and $\mathcal{D}$-consequence have obvious counterparts as relations on codes of $\mathcal{L}$- and $\mathcal{LF}$-formulas; I'll refer to these counterparts as $\textsc{Valid}_{\mathcal{L}}$, $\textsc{Implies}_{\mathcal{L}}$, $\textsc{Valid}_{\mathcal{LF}}$ and $\textsc{Implies}_{\mathcal{LF}}$.

With Proposition 4.4, we saw that the structure $\mathbb{HF}(\mathcal{D})$ can be $\mathcal{D}$-axiomatized up to isomorphism by a single formula of $\mathcal{LF}$. This implies that the concept of first-order truth-in-$\mathbb{HF}(\mathcal{D})$ is embedded in the concept of $\mathcal{LF}$-validity.

PROPOSITION 4.5. *Suppose that $\mathcal{D}$ is infinite. Then, there is a first-order formula $\theta$ in the signature of $\mathbb{HF}(\mathcal{D})$, with one extra monadic predicate* V*, such that*

$$\mathbb{HF}(\mathcal{D}), \textsc{Valid}_{\mathcal{LF}} \models \theta[\dot\phi] \leftrightarrow \phi$$

*for all first-order formulas $\phi$ in the signature of $\mathbb{HF}(\mathcal{D})$.*

*Proof.* Let $Z$ be the categorical $\mathcal{LF}$-axiom for $\mathbb{HF}(\mathcal{D})$, which is given by the proof of Lemma 4.4. Let $\zeta$ be a formula of the metalanguage such that $\mathbb{HF}(\mathcal{D}) \models \zeta[x/a, y/b]$ iff $a$ is an $\mathcal{LF}$-formula $A$ and $b$ is $Z\dot\rightarrow A$. Let $\theta$ be the formula $\exists y(\zeta \wedge Vy)$. Given that $\mathcal{D}$ is infinite, $\mathbb{HF}(\mathcal{D})$ is isomorphic to an $\mathbb{HF}(\mathcal{D})$-coded $\mathcal{D}$-structure. Using this fact, it is routine to verify that $\theta$ satisfies the claim of the lemma. □

https://doi.org/10.1017/S1755020316000472 Published online by Cambridge University Press

We are clearly in the vicinity of Tarski's theorem, which has the following corollary.

LEMMA 4.6. *Suppose that $\theta$ is a first-order formula in the signature of $\mathbb{HF}(\mathcal{D})$ together with a single new predicate $R$, such that*

$$\mathbb{HF}(\mathcal{D}), X \models \theta[\dot{\phi}] \leftrightarrow \phi$$

*for every first-order formula $\phi$ in the signature of $\mathbb{HF}(\mathcal{D})$. Then $X$ is not first-order definable on $\mathbb{HF}(\mathcal{D})$.*

*Proof.* Suppose to the contrary that $X$ were first-order definable, say by a formula $\psi$ whose free variables are just $\vec{x}$. Let $\theta'$ be the result of everywhere replacing $R\vec{x}$ with $\psi$. Then

$$\mathbb{HF}(\mathcal{D}) \models \theta'[\dot{\phi}] \leftrightarrow \phi$$

for all $\phi$, and this contradicts Tarski's theorem. □

PROPOSITION 4.7. *Suppose that $\mathcal{D}$ is infinite. Then the class of $\mathcal{D}$-valid $\mathcal{LF}$-formulas is not first-order definable on $\mathbb{HF}(\mathcal{D})$.*

*Proof.* Immediate from Proposition 4.5 and Lemma 4.6. □

In fact, something stronger is true. A formula is said to be (explicitly) $\Pi^1_1$ if it is the result of prefixing a first-order formula with a string of universal second-order quantifiers. The official definitions of $\mathcal{D}$-validity and of $\mathcal{D}$-consequence have the form "for all $\mathcal{D}$-structures…". Moreover, a $\mathcal{D}$-structure is determined by its diagram. And under the coding of $\mathcal{L}$-formulas with elements of $\mathbb{HF}(\mathcal{D})$, a diagram is just a class on $\mathbb{HF}(\mathcal{D})$. So the official definitions of $\mathcal{D}$-validity and $\mathcal{D}$-consequence are naturally expressed over $\mathbb{HF}(\mathcal{D})$ with $\Pi^1_1$ formulas. Can this rendering be simplified?

A class $P$ is said to be $\Pi^1_1$-complete over some structure if for every $\Pi^1_1$ formula $\phi$, there's a first-order formula $\psi$ such that $\phi[\vec{a}] \leftrightarrow \psi[P, \vec{a}]$ holds for all $\vec{a}$ in the domain. Thus, $P$ is $\Pi^1_1$-complete if every class definable by a second-order universal generalization of a first-order formula is first-order definable relative to $P$. Note that since $\mathbb{HF}(\mathcal{D})$ permits coding finite sequences of classes as classes of finite sequences, it is enough to consider the case of a single universal second-order quantifier.

I'll now show that for formulas of $\mathcal{LF}$, the concept of $\mathcal{D}$-validity is $\Pi^1_1$-complete. To this end, I'll show that in $\mathbb{HF}(\mathcal{D})$, the concept of satisfaction can be reduced to the concept of truth, and then show that the concept of truth for $\Pi^1_1$ formulas is first-order definable relative to $\mathcal{D}$-validity.

Reducing satisfaction to truth is somewhat like trying to replace *de re* mental states with *de dicto* ones (except perhaps in being easier). As is well known, each element of $\mathbb{HF}$ has a parameter-free first-order definition, so in that case, the satisfaction of a formula by an element can be reduced to truth using something like Russell's theory of descriptions. This is not so for elements of $\mathbb{HF}(\mathcal{D})$, since the urelements are indiscernible. But for present purposes, a notion weaker than that of definition will suffice. Say that sets $a, b$ are isomorphic if they are identical modulo swapping of urelements. Now let's say that a formula describes $a$ if it is satisfied precisely by the sets isomorphic to $a$. The crucial fact about descriptions will be this: that no objects satisfying the same description differ over any parameter-free formula. As I'll verify below, every element of $\mathbb{HF}(\mathcal{D})$ is first-order describable. Indeed, a "canonical" description of $a$ can be generated from $a$ in a simple uniform manner. This means that lying within $\mathbb{HF}(\mathcal{D})$, there is a definable "road back" from each object to a bunch of codes of its canonical descriptions.

LEMMA 4.8. *There is a formula $\delta$ in the signature of $\mathbb{HF}(\mathcal{D})$, such that (i) if $\mathbb{HF}(\mathcal{D}) \models \delta[a, b]$ then $b$ codes a canonical description of $a$, and (ii) $\mathbb{HF}(\mathcal{D}) \models \forall x \exists y \delta[x, y]$.*

*Proof.* Let's begin by associating to each element $a$ of $\mathbb{HF}(\mathcal{D})$ a formula $\gamma_a$ without parameters which canonically describes it. First, write $\nu$ for a one-one association of variables $\nu_a, \nu_b, \ldots$ for some finitely many elements $a, b, \ldots$ of $\mathbb{HF}(\mathcal{D})$. Now, let's construct a formula $\beta_{a,\nu}$ which says of the objects assigned to its free variables that they're arranged as the elements of the transitive closure of $a$:

$$\beta_{a,\nu} = \begin{cases} \mathrm{D}\nu_a & \text{if } a \in \mathcal{D}, \text{ and} \\ \forall y \left(y \in \nu_a \leftrightarrow \bigvee_{b \in a} y = \nu_b\right) \wedge \bigwedge_{b \in a} \beta_{b,\nu}, & \text{otherwise.} \end{cases}$$

Conjoin $\beta_{a,\nu}$ with each formula $\nu_b \neq \nu_c$ such that $b, c$ are distinct urelements in the transitive closure of $a$. Finally, let $\gamma_{a,\nu}$ be the result of existentially generalizing this with respect to each variable $\nu_b$ such that $b \neq a$. Then $\mathbb{HF}(\mathcal{D}) \models \gamma_{a,\nu}[b]$ iff $b$ is isomorphic to $a$, and so $\gamma_{a,\nu}$ describes $a$.

Clearly the construction of $\gamma_{a,\nu}$ from $a$ can be first-order definably replicated inside of $\mathbb{HF}(\mathcal{D})$. The formula $\gamma_{a,\nu}$ is given relative to the choice $\nu$ of variables for sets. To wash this out, we'll now take as defined, by induction on rank, the relation that holds between $a$ and $b$ iff $b = \dot{\gamma}_{a,\nu}$ for some $\nu$. It remains only to conjoin the requirement that the one free variable of $\dot{\gamma}_{a,\nu}$ is $\dot{x}$. The result is a first-order formula $\delta$ in the signature of $\mathbb{HF}(\mathcal{D})$ such that

$$\mathbb{HF}(\mathcal{D}) \models \delta[a, b] \text{ iff } b = \dot{\gamma}_{a,\nu} \text{ for some } \nu \text{ such that } \dot{\nu}_a = \dot{x}.$$

Then $\delta$ satisfies

$$\text{if } \mathbb{HF}(\mathcal{D}) \models \delta[a, b], \text{ then there's a } \phi \text{ such that } b = \dot{\phi} \text{ and } \phi \text{ describes } a,$$

for all $a, b$. Thus, $\delta$ defines the graph of a function which takes an element of $\mathbb{HF}(\mathcal{D})$ to a nonempty class of codes of its canonical descriptions. Since every element of $\mathbb{HF}(\mathcal{D})$ has a canonical description, therefore also $\mathbb{HF}(\mathcal{D}) \models \forall x \exists y \delta[x, y]$. □

The definable describability of each element of $\mathbb{HF}(\mathcal{D})$ now yields the desired reduction of $\Pi^1_1$ satisfaction. The reduction is more general, but I'll just state the relevant case. Let $\mathrm{TRUE}_{\Pi^1_1}$ be the class of codes of $\Pi^1_1$ formulas which are true in $\mathbb{HF}(\mathcal{D})$.

LEMMA 4.9. *There is a first-order formula $\theta$ such that*

$$\mathbb{HF}(\mathcal{D}) \models \phi[a] \quad \textit{iff} \quad \mathbb{HF}(\mathcal{D}), \mathrm{TRUE}_{\Pi^1_1} \models \theta[a, \dot{\phi}] \tag{12}$$

*for all $a$ and all $\Pi^1_1$ formulas $\phi$ in the signature of $\mathbb{HF}(\mathcal{D})$.*

*Proof.* Let $\delta$ be the formula provided by Lemma 4.8, and let T be the predicate interpreted by $\mathrm{TRUE}_{\Pi^1_1}$. We can assume that the free variable of $\phi$ is $x$. Let's work out how to say that $\dot{\exists}\dot{x}\dot{(}\dot{\phi}\dot{\wedge}\dot{\gamma}_a\dot{)}$ belongs to $\mathrm{TRUE}_{\Pi^1_1}$ for some description $\gamma_a$ of $a$.

Let $\eta$ be the formula $\exists w(\mathrm{T}w \wedge w = \dot{\exists}\dot{x}\dot{(}y\dot{\wedge}z\dot{)})$. Then $\mathbb{HF}(\mathcal{D}), \mathrm{TRUE}_{\Pi^1_1} \models \eta[b, c]$ iff $b$ and $c$ are codes of $\Pi^1_1$ formulas such that $\dot{\exists}\dot{x}\dot{(}b\dot{\wedge}c\dot{)}$ is in $\mathrm{TRUE}_{\Pi^1_1}$. Let $\theta$ be the formula $\exists z(\delta[x, z] \wedge \eta)$. Noting that any metalinguistic formula $\phi$ is parameter-free, it is routine to verify that (12) holds of $\theta$. □

PROPOSITION 4.10. *If $\mathcal{D}$ is infinite, then $\mathrm{TRUE}_{\Pi^1_1}$ is first-order definable relative to $\mathrm{VALID}_{\mathcal{LF}}$.*

https://doi.org/10.1017/S1755020316000472 Published online by Cambridge University Press

*Proof.* Let $\theta$ be the formula $\mathrm{V}(Z \dot{\to} x)$ where V is a predicate not in the signature of $\mathbb{HF}(\mathcal{D})$, and $Z$ is the $\mathcal{LF}$-formula of Proposition 4.4 which categorically axiomatizes $\mathbb{HF}(\mathcal{D})$. Suppose $\phi$ is a first-order formula in the free second-order variables $\vec{X}$. Then

$$\begin{aligned}\mathbb{HF}(\mathcal{D}) \models \forall \vec{X}\phi \quad &\text{iff} \quad \mathbb{HF}(\mathcal{D}), \vec{P} \models \phi \text{ for all subclasses } \vec{P} \text{ of HF}(\mathcal{D})\\ &\text{iff} \quad \mathcal{M}, \vec{P} \models \phi \text{ for all } \mathcal{D}\text{-structures } \mathcal{M} \simeq \mathbb{HF}(\mathcal{D}) \text{ and all } \vec{P} \subseteq \mathcal{D}\\ &\text{iff} \quad \mathbb{HF}(\mathcal{D}), \mathrm{VALID}_{\mathcal{LF}} \models \theta[\dot{\phi}]. \qquad \square\end{aligned}$$

PROPOSITION 4.11. *If $\mathcal{D}$ is infinite, then the class of $\mathcal{D}$-valid formulas of $\mathcal{LF}$ is $\Pi^1_1$-complete on $\mathbb{HF}(\mathcal{D})$.*

*Proof.* Immediate from Lemma 4.9 and Proposition 4.10. $\square$

Note that *a priori*, the definition of $\mathcal{D}$-consequence is $\Pi^1_1$, since quantification over structures is essentially just second-order quantification over $\mathbb{HF}(\mathcal{D})$. So, Proposition 4.11 is as strong as possible. It shows that if the underlying universe is infinite, then there cannot be a notion of proof whose completeness would, as in the case of first-order logic, secure any reduction in the complexity of class of valid formulas. The fact which underlies this is that if the universe is infinite, then $\mathcal{LF}$ can itself characterize $\mathbb{HF}(\mathcal{D})$ up to isomorphism.

Needless to say, a concept of proof need not serve only to situate validity or consequence in the complexity-theoretic universe. For example, the existence of a sound and complete proof procedure might be held to supply its correlative notion of consequence with some sort of an explanation or analysis. From that point of view, the limitive results of this section don't rule out philosophical importance for some sound and complete notion of proof for $\mathcal{LF}$.

Under the assumption that $\mathcal{D}$ is countable, the languages $\mathcal{L}$ and $\mathcal{LF}$ can be translated into countably infinite truth-functional logic. And so they do have a complete notion of proof (Lopez-Escobar, 1965). Unlike the notion of proof for classical logic, this cannot show the consequence relation to be any simpler than what's given by its *a priori* specification. Nonetheless, it might be claimed to demonstrate that the corresponding consequence relation is uniformly realized through infinitary symbolic patterns, and, therefore, still appropriately grounded in propositional structure. The philosophical development and evaluation of this proposal is a difficult problem, which must be left to further work.

Let's conclude this subsection by evaluating the complexity-theoretic significance of the restriction to countable Russellian structures independently of the form-series device. We saw in §2.2 that under $\mathcal{D}$-consequence, every structure is categorically axiomatized by its diagram. Consequently, a counterpart of Proposition 4.11 holds for the notion of $\mathcal{D}$-consequence over $\mathcal{L}$-formulas. The counterpart is somewhat weaker, because infinite diagrams are not in general definable over $\mathbb{HF}(\mathcal{D})$. So within $\mathbb{HF}(\mathcal{D})$, the concept of $\mathcal{D}$-consequence yields the concept of truth-in-$\mathbb{HF}(\mathcal{D})$ only relative to an enumeration of $\mathcal{D}$. The proof mimics that of Proposition 4.11. Let $(\mathrm{HF}(\mathcal{D}), \mathcal{D}, \in, f)$ be the result of adding to $\mathbb{HF}(\mathcal{D})$ a mapping $f$ of $\omega$ onto $\mathcal{D}$.

PROPOSITION 4.12. *Suppose that $\mathcal{D}$ is countably infinite.* $\mathrm{IMPLIES}_{\mathcal{L}}$ *is $\Pi^1_1$-complete on* $(\mathrm{HF}(\mathcal{D}), \mathcal{D}, \in, f)$.

*Proof.* By Lemma 4.9, $\mathrm{TRUE}_{\Pi^1_1}$ is $\Pi^1_1$-complete on $\mathbb{HF}(\mathcal{D})$. Since $\mathbb{HF}(\mathcal{D})$ contains a first-order definable copy of $(\mathrm{HF}(\mathcal{D}), \mathcal{D}, \in, f)$, therefore $\mathrm{TRUE}_{\Pi^1_1}$ must also be $\Pi^1_1$-complete on $(\mathrm{HF}(\mathcal{D}), \mathcal{D}, \in, f)$. So it suffices to show that $\mathrm{TRUE}_{\Pi^1_1}$ is first-order definable on $(\mathrm{HF}(\mathcal{D}), \mathcal{D}, \in, f, \mathrm{IMPLIES}_{\mathcal{L}})$.

Suppose that $\mathcal{M}_{\mathbb{HF}(\mathcal{D})}$ is a $\mathcal{D}$-structure isomorphic to $\mathbb{HF}(\mathcal{D})$ itself. Let

$$\Delta(\mathcal{M}_{\mathbb{HF}(\mathcal{D})}) = \text{the diagram of } \mathcal{M}_{\mathbb{HF}(\mathcal{D})}.$$

From Proposition 2.2 it follows that $\mathcal{M} \models \Delta(\mathcal{M}_{\mathbb{HF}(\mathcal{D})})$ iff $\mathcal{M} = \mathcal{M}_{\mathbb{HF}(\mathcal{D})}$. So there's a first-order formula $\theta$ such that

$$\mathbb{HF}(\mathcal{D}), \textsc{Implies}_{\mathcal{L}}, \Delta(\mathcal{M}_{\mathbb{HF}(\mathcal{D})}) \models \theta[\dot{\phi}] \text{ iff } \mathbb{HF}(\mathcal{D}) \models \phi$$

for all $\Pi^1_1$ formulas $\phi$. But there is an $\mathcal{M}$, isomorphic to $\mathbb{HF}(\mathcal{D})$, whose diagram is first-order definable on $(\mathrm{HF}(\mathcal{D}), \mathcal{D}, \in, f)$. So, the class $\textsc{True}_{\Pi^1_1}$ is first-order definable on $(\mathrm{HF}(\mathcal{D}), \mathcal{D}, \in, f, \textsc{Implies}_{\mathcal{L}})$, as desired. □

***4.4. Uncountability and nonabsoluteness.*** So far, we've been investigating the complexity of concepts of the metatheory of $\mathcal{L}$ and $\mathcal{LF}$ relative to this or that fixed choice of $\mathcal{D}$. If the universe must be finite, then the notion of tautology is in any case decidable. If the universe is fixed as some infinite set $\mathcal{D}$, then the notion of $\mathcal{LF}$-tautology is $\Pi^1_1$-complete on $\mathbb{HF}(\mathcal{D})$. However, from an epistemological point of view it is not clear that the most significant measure of complexity presumes a determination of $\mathcal{D}$. For as we'll see in §5, Wittgenstein appears to demand, in some sense, that the size of the universe not be prejudged. So, I want now to investigate what can be said about the complexity of $\mathcal{D}$-validity as a relation between $\mathcal{LF}$-formulas and arbitrary choices of $\mathcal{D}$, and analogously for $\mathcal{D}$-consequence with sets of $\mathcal{L}$-formulas.

As we've seen, the concept of $\mathcal{D}$-validity for $\mathcal{LF}$ can be expressed by a formula using second-order quantification over $\mathbb{HF}(\mathcal{D})$. The definition can also be phrased as a formula which uses a single first-order universal quantification ranging over the power set of $\mathbb{HF}(\mathcal{D})$. Can that universal quantifier be replaced with an existential one? By Proposition 4.11, the existential quantifier couldn't just range over $\mathbb{HF}(\mathcal{D})$, but we might hope to find some infinitary notion of "proof", or more broadly of "pattern", to witness the concept of $\mathcal{D}$-validity for $\mathcal{LF}$. Could such a complete, if perhaps quite profligate, notion of proof be established on the basis of the axioms of set theory?

As before, it's natural to avoid assuming that $\mathcal{D}$ has been coded as some pure set. Here, let's adapt ZFC to handle $\mathcal{D}$ as a collection of urelements. Add to the language of set theory a primitive predicate D which corresponds to the property of belonging to the class $\mathcal{D}$ of urelements. Add an axiom to the effect that nothing in $\mathcal{D}$ has an element, and replace the usual axiom of extensionality with an axiom that things outside of $\mathcal{D}$ are the same if they have the same elements. It simplifies matters to assume that the elements of $\mathcal{D}$ form a set, although the result established here can be extended to the case in which the assumption is dropped.

In the resulting set theory ZFCU, the theory of logics $\mathcal{L}$ and $\mathcal{LF}$ easily results by formalizing the construction which earlier took place in $\mathbb{HF}(\mathcal{D})$. In particular, the assumption that $\mathcal{D}$ forms a set means that the definition of validity and consequence become expressible by universal, first-order generalization over subsets of $\mathcal{D}$.

It is clear that the Russellian constraint on the class of all structures implies that the consequence relation discriminates between domains of different cardinality. For example, it certainly discriminates between domains of different finite cardinality. As we'll now see, the Russellian constraint also implies that the consequence relation effects transfinite discriminations as well.

PROPOSITION 4.13. (*i*) *There is an $\mathcal{LF}$-formula Uncountable such that Uncountable is $\mathcal{D}$-valid iff $\mathcal{D}$ is uncountable.* (*ii*) *For any subset $s$ of $\mathcal{D}$, there is a set* $\textsc{Onto}_s$ *of $\mathcal{L}$-formulas such that* $\textsc{Onto}_s$ *is $\mathcal{D}$-satisfiable iff $s$ has the same cardinality as $\mathcal{D}$.*

https://doi.org/10.1017/S1755020316000472 Published online by Cambridge University Press

*Proof.* (i) Call a binary relation successorlike if it stands to $\mathcal{D}$ as the successor relation stands to the finite ordinals. Let *Uncountable* be an $\mathcal{LF}$-formula which says that $R$ is not successorlike. Clearly *Uncountable* works as desired. (ii) Suppose that $\mathcal{D}$ is infinite, and let $s$ be a subset of $\mathcal{D}$. For some monadic predicate $F$, let $F_s$ be the collection of all formulas $Fa$ for $a \in s$, together with all formulas $\neg Fa$ for $a \notin s$. And, for some dyadic predicate $R$, let $B$ be an $\mathcal{L}$-formula to the effect that $R$ maps the $F$s onto $\mathcal{D}$. Let $\textsc{Onto}_s = F_s \cup \{B\}$. Then $\textsc{Onto}_s$ is $\mathcal{D}$-satisfiable iff there is a surjection from $s$ onto $\mathcal{D}$. □

However, the following lemma shows that the condition of uncountability of the class of urelements cannot be expressed by a $\Sigma_1$ formula.

LEMMA 4.14. *There is no $\Sigma_1$ formula $\theta$ such that* $\mathrm{ZFCU} \vdash \mathrm{D} \succ \aleph_0 \leftrightarrow \theta$.

*Proof.* Assume to the contrary that $\theta$ is a $\Sigma_1$ formula such that $\mathrm{ZFCU} \vdash \mathrm{D} \succ \aleph_0 \leftrightarrow \theta$. Let $\mathbb{V}(x)$ formalize "the class of all sets with urelements drawn from $x$". For any formula $\phi$, write $\phi^{\mathbb{V}(x)}$ for the relativization of $\phi$ to $\mathbb{V}(x)$. And let $\mathrm{Tc}(x)$ formalize "the transitive closure of $x$".

Clearly, ZFCU proves that any surjection from $\omega$ to a set $m$ belongs to the class of all sets with urelements drawn from $\mathrm{Tc}(m)$. So,

$$\mathrm{ZFCU} \vdash m \preccurlyeq \aleph_0 \rightarrow (\mathrm{D} \preccurlyeq \aleph_0)^{\mathbb{V}(\mathrm{Tc}(m))}. \tag{13}$$

Since $\mathrm{D} \leq \aleph_0$ and $\theta$ are $\Sigma_1$, therefore $\mathrm{D} \preccurlyeq \aleph_0 \rightarrow \neg\theta$ is $\Pi_1$. Moreover, $\mathbb{V}(\mathrm{Tc}(m))$ is transitive. But, $\Pi_1$ formulas are downward absolute for transitive classes (Jech, 2003, 185). So from $\mathrm{ZFCU} \vdash \mathrm{D} \succ \aleph_0 \leftrightarrow \theta$, it follows that

$$\mathrm{ZFCU} \vdash (\mathrm{D} \preccurlyeq \aleph_0)^{\mathbb{V}(\mathrm{Tc}(m))} \rightarrow \neg\theta^{\mathbb{V}(\mathrm{Tc}(m))}. \tag{14}$$

And (13) and (14) together imply

$$\mathrm{ZFCU} \vdash m \preccurlyeq \aleph_0 \rightarrow \neg\theta^{\mathbb{V}(\mathrm{Tc}(m))}. \tag{15}$$

On the other hand, we assumed that $\mathrm{ZFCU} \vdash \mathrm{D} \succ \aleph_0 \rightarrow \theta$. Write $\phi^m$ for the relativization of $\phi$ to $m$. Using the axiom of choice, a version of the reflection theorem implies that if $\mathrm{ZFCU} + \phi \vdash \theta$ then $\mathrm{ZFCU} + \phi \vdash \exists m(m \preccurlyeq \aleph_0 \wedge m \text{ is transitive} \wedge \theta^m))$ (Jech, 2003, 165ff). So by the deduction theorem,

$$\mathrm{ZFCU} \vdash \mathrm{D} \succ \aleph_0 \rightarrow \exists m(m \preccurlyeq \aleph_0 \wedge m \text{ is transitive} \wedge \theta^m). \tag{16}$$

We also assumed that $\theta$ is $\Sigma_1$, so that $\theta$ is upward absolute for transitive classes. Consequently, (16) implies

$$\mathrm{ZFCU} \vdash \mathrm{D} \succ \aleph_0 \rightarrow \exists m(m \preccurlyeq \aleph_0 \wedge \theta^{\mathbb{V}(\mathrm{Tc}(m))}). \tag{17}$$

From (15) and (17), it follows that

$$\mathrm{ZFCU} \vdash \mathrm{D} \succ \aleph_0 \rightarrow \bot. \tag{18}$$

But $\mathrm{ZFCU} + \mathrm{D} \succ \aleph_0$ is consistent relative to ZF. So (18) contradicts the consistency of ZF. □

We can now put the pieces together, establishing the last technical result of this paper. Let $\mathrm{Valid}_{\mathcal{LF}}$ formalize in ZFCU the property of being a $\mathcal{D}$-tautology; and let $\mathrm{Implies}_{\mathcal{L}}$ formalize the relation of $\mathcal{D}$-consequence which holds between a set of formulas and a formula. Say that a formula is $\Sigma_1$-definable in ZFCU if it is provably equivalent to a $\Sigma_1$ formula. Then

PROPOSITION 4.15. *Neither* $\mathrm{Valid}_{\mathcal{LF}}$ *nor* $\mathrm{Implies}_{\mathcal{L}}$ *is $\Sigma_1$-definable in* ZFCU.

https://doi.org/10.1017/S1755020316000472 Published online by Cambridge University Press

*Proof.* First suppose that $\phi$ is a $\Sigma_1$ formula such that

$$\text{ZFCU} \vdash \phi \leftrightarrow \text{Valid}_{\mathcal{LF}}. \tag{19}$$

Let $\xi$ be a $\Sigma_1$ term for the (coded) $\mathcal{LF}$-formula *Uncountable* of Proposition 4.13. Clearly, it is provable in ZFCU that *Uncountable* is a tautology iff $\mathcal{D}$ is uncountable:

$$\text{ZFCU} \vdash \text{Valid}_{\mathcal{LF}}[\xi] \leftrightarrow \text{D} \succ \aleph_0. \tag{20}$$

So by (19) and (20) there would be a $\Sigma_1$ equivalent of $\mathcal{D} \succ \aleph_0$, and this contradicts Lemma 4.14.

Similarly, suppose that $\psi$ is a $\Sigma_1$ formula such that

$$\text{ZFCU} \vdash \psi \leftrightarrow \text{Implies}_{\mathcal{L}}. \tag{21}$$

The proof of Proposition 4.13(ii) gives a method of constructing, from any subset $s$ of $\mathcal{D}$, a set $\text{ONTO}_s$ of formulas which is $\mathcal{D}$-satisfiable iff $s$ has the same cardinality as $\mathcal{D}$. When formalized in ZFCU, this construction yields a $\Sigma_1$ term $\zeta$ in the free variable $x$ such that

$$\text{ZFCU} \vdash \forall x(\text{Implies}_{\mathcal{L}}[\zeta, \bar{\bot}] \leftrightarrow x \prec \text{D}) \tag{22}$$

where $\bar{\bot}$ is a $\Sigma_1$ term for the code of an $\mathcal{L}$-contradiction. By (21) and (22) it follows that

$$\text{ZFCU} \vdash \exists x(\aleph_0 \preccurlyeq x \wedge \psi[\zeta, \bar{\bot}]) \leftrightarrow \aleph_0 \prec \text{D}. \tag{23}$$

But the left side of (23) is $\Sigma_1$, again contradicting Lemma 4.14. □

By itself, Russellian constraint broadens but does not significantly complicate the class of validities. However, on classical semantics, it's compactness which ensures that if validity is witnessed by finite proofs, then so is consequence. And $\mathcal{D}$-consequence is not compact, because for example a universal generalization is a $\mathcal{D}$-consequence of the class of its instances. So $\mathcal{D}$-consequence might turn out to be more complicated than validity. And indeed, the effects are as strong as possible: $\mathcal{D}$-consequence permits no simplification of the analysis of consequence through universal generalization over structures. On the other hand, in the case of $\mathcal{D}$ countably infinite, the form-series device makes the concept of validity already that bad regardless of the Russellian constraint. But, the form-series device also has the effect of concentrating, into the notion of validity, the complications of the consequence relation which follow from the Russellian constraint. This implies, in the general case, that the concept of tautology, or of logically valid formula, is not even $\Sigma_1$ definable in set theory. So in the general case, it is not just that we cannot replace the search through the collection of structures on some infinite domain with an enumeration of finite proofs. Rather, validity cannot in general be witnessed by any system of mathematical constructions which are identifiable by properties intrinsic to the constructions themselves. It depends essentially on the extrinsic matter of which bijections happen to exist in the mathematical universe.

**§5. Conclusion.** Let me now summarize the results of the previous sections, and then sketch an implication which I take to be of philosophical significance. We've seen that logic as Wittgenstein conceives it in the *Tractatus* differs from classical logic in two ways. The presence of the form-series device contributes to $\mathcal{LF}$ the capacity to express induction, and therefore allows categorical finite axiomatization of rich countable structures. Wittgenstein also relativizes the notion of structure to some fixed domain $\mathcal{D}$ of what has a unique name. This doesn't, by itself, significantly complicate the notion of validity. But it breaks the compactness theorem: indeed, the $\mathcal{D}$-consequence relation embeds the concept of truth.

https://doi.org/10.1017/S1755020316000472 Published online by Cambridge University Press

When the two departures from classical logic are combined, then we obtain the notion of $\mathcal{D}$-validity for $\mathcal{LF}$-formulas. And this notion turns out to depend on the distinction between $\mathcal{D}$ countable and uncountable.

I want now to consider the significance, for the *Tractatus*, of distinctions in the cardinality of domain. In the *Tractatus*, the question of the number of existing objects has a problematic status. Specifically, Wittgenstein distinguishes in the 5.55s between logic and its "application". The application of logic supposedly "decides what elementary propositions there are" (5.557a). At the same time, "logic and its application must not overlap" (5.557e), because logic "cannot anticipate" what lies in its application (5.557b). Now, the totality of objects, and the totality of elementary propositions, are alike in that each constitutes, or makes manifest, some limit to "empirical reality" (5.5561). So, the size of the universe would appear to belong to the "application" of logic, rather than to logic itself; and so it could not be anticipated by logic alone (see also 5.55b, together with 5.551a).

It would seem obvious that Wittgenstein wants to maintain that relations of consequence, contradictoriness, and so on do belong to logic (6.1ff). For example, it might be supposed to belong to logic that the disjunction of some proposition with its negation is a tautology. Among what belongs to logic, then, Wittgenstein would include the consequence relation, but exclude the size of the universe. From what became the classical understanding of the consequence relation, this pair of constraints seems quite natural. Indeed, the classical understanding would exclude the size of the universe from logic: no proposition "there are at least $n$ $F$s" classically entails the proposition "everything is $F$", yet nor is any proposition "there are at least $n$ things" a logical truth. But as we've seen, Wittgenstein restricts the class of all structures to those with some common underlying domain. So from the *Tractatus* point of view, if the number of objects is finite, then the consequence relation would evince this; likewise it could be seen that the number of objects isn't any particular finite $n$. Wittgenstein maintains that if infinitely many objects exist, then this fact would show up in the extent of the consequence relation. Thus, if consequence belongs to logic, then something that would make itself felt in logic is the size of the universe.

We've now reached something of an antinomy. It's natural to understand 5.557 as denying that logic might by itself anticipate the number of objects. But as we've just seen, logic would anticipate the number of objects if the consequence relation belonged to logic.

I'm inclined to conclude that for the early Wittgenstein, the consequence relation does not belong to logic without qualification. Instead, the consequence relation emerges in the application of logic. It is in the application of logic, for example, that "formal-logical properties of language and the world" would get shown through the fact that a given proposition of logic is a tautology: "the fact that a tautology is yielded by *this particular way* of connecting its constituents characterizes the logic of its constituents" (6.12b). The consequence relation belongs to logic only once logic is taken together with its application.

One thing that is fairly clear, though, is that applying logic was not a job Wittgenstein undertook in the *Tractatus* itself: he gives only an illustrative example at 6.3751. So, he cannot coherently have intended the consequence relation itself to be revealed in the *Tractatus*. The question then arises: in what respect could that book have been intended to illuminate logic?

As I suggested toward the end of §1.2.3, Wittgenstein's promised solution to philosophical problems appears to depend on the evolution of the general propositional form: it begins with the gnomic "this is how things are" at 4.5, but becomes the "general form of the truth-function" at 6. Specifically, philosophical progress would, in any particular case, involve rewriting propositions in such a way that their logical interrelations would be made manifest in the resulting propositional signs. For example, rewriting "the mug reflects white light" as a conjunction one of whose conjuncts is "the mug reflects red light" would make manifest that the former entails the latter. Wittgenstein's rather attenuated conception of philosophical progress would seem to turn on the conviction that analysis is possible: that the totality of propositions can be rewritten in such a way that all modal interdependence between propositions would be made manifest merely in the signs for them. Thus it would become clear that "the only necessity that exists is *logical* necessity" (6.37).

This conception of philosophical progress plainly requires some independent standard for a logical relationship's being made manifest. Suppose people recognize some propositions $p$ and $q$ to be such that $p$ entails $q$, and suppose some purported analysis to rewrite them in the signs $A$ and $B$. The adequacy of such an analysis would be determined at least partly by whether the entailment of $q$ by $p$ is shown in the new signs $A$, $B$. Without such a standard, the adequacy of the purported analysis would be indeterminate.

But what is it, exactly, for something to be made manifest, or shown? It is unlikely that Wittgenstein could have meant by this anything like "made evident" in an epistemic sense. He wrote that "it is remarkable that so exact a thinker as Frege should have appealed to the degree of self-evidence as the criterion of a logical proposition" (6.1271). The assimilation of epistemology to psychology tells against evidentness as a criterion for logical consequence. Moreover, the number of objects that exist is supposed to be shown (Wittgenstein, 1979, 19.8.19): but the epistemic status of the question of the number of objects is, to say the least, obscure.

Let's say that a feature of symbols is symbolically realized if it supervenes purely on what makes anything into those symbols. I propose that Wittgenstein took something's being shown to require that it be symbolically realized. The philosophical purpose of the evolution of a general form of the proposition at 4.5 into a general form of the truth-function at 6 can then be understood at least in part as responding to the demand for a univocal account of the symbolic realization of logical consequence.

Wittgenstein deigns to no concerted application of logic himself. So, if he had achieved his goal in the general form of the truth-function, then the symbolic realization of consequence would have to have been identified independently of the application of logic. It is tempting to conclude that this could be achieved. For example, the contradictoriness of any proposition with its negation would be so fixed, no matter what the proposition's ultimate analysis reveals it to be. Another example of something fixed by Wittgenstein's purported logical achievement would be this: that $\forall x A$ would follow from $A[a]$, $A[b]$, ... under the condition that $A[a]$, $A[b]$, ... are all the instances of $\forall x A$. More generally, as I understand it, the *Tractatus* presents an attempt to show how the entirety of the consequence relation might be fixed, conditional on any determination of the totality of objects. Having proposed an account of the symbolic realization of logical consequence, Wittgenstein then could declaim *a priori* that logic, in its application, makes manifest "formal-logical properties of language and the world" (6.12).

> The propositions *show* the logical form of reality.
> They exhibit it. (4.121e–g)

> Thus a proposition $fa$ shows that in its sense the object $a$ occurs, two propositions $fa$ and $ga$ that they are both about the same object.
> If two propositions contradict one another, this is shown by their structure; similarly if one follows from another, etc. (4.1211)

Within this sketch of Wittgenstein's logical aims in the *Tractatus*, the complexity-theoretic analyses of §4 assume a new significance. Compatibly with those analyses, is there some way in which mere propositional structure could be taken to make manifest the $\mathcal{D}$-validity of all $\mathcal{D}$-valid formulas, no matter the size of $\mathcal{D}$? Wittgenstein suggested that axiomatic proofs represent progress toward this goal:

> Whether a proposition belongs to logic can be calculated by calculating the logical properties of the *symbol*.
>
> And this we do when we prove a logical proposition. For without troubling ourselves about a sense and a meaning, we form the logical propositions out of others by mere symbolic rules (6.126a,b).

We've seen that if the number of objects is infinite, then it is completely out of the question that some notion of finite proof could suffice to demonstrate the validity of all valid formulas. But Wittgenstein himself couldn't have rested content with axiomatic proof, because he held that validity, or tautologousness, is only a limiting case of the broader phenomenon of logical consequence.

> If the truth of one proposition follows from the truth of others, this expresses itself in relations in which the forms of these propositions stand to one another, and we do not need to put them in these relations first by connecting them with one another in a proposition; for these relations are internal, and exist as soon as, and by the very fact that, the propositions exist. (5.131)

Toward an account of the consequence relation Wittgenstein gives only this seemingly proto-Tarskian sketch:

> If the truth-grounds which are common to a number of propositions are all also truth-grounds of some one proposition, we say that the truth of this proposition follows from the truth of those propositions. (5.12)

Unfortunately, this apparent analysis falls under a remark that "truth-functions can be arranged in series" (5.1a). So it looks to be lodged in a notational proposal which is at best incomplete. The proposal is that some propositions can be written as tables of agreement and disagreement with the distributions of truth-value over elementary propositions (4.442). Surely, Wittgenstein couldn't say that all propositions can be so written, for *a priori* the number of elementary propositions might be infinite. The details of the construction of a general form of the truth-function through iterations of joint denial are evidently driven by a demand that propositions can be written down even if they have infinite logical ancestry. So, the truth-tabular symbolism does make some logical consequence manifest. But like a proof system for first-order logic, it is only partial.

The question then remains: is there some very loose, even wildly profligate, notion of symbolic pattern, which could serve to make manifest all instances of the $\mathcal{D}$-consequence relation, or even the $\mathcal{D}$-validity of all $\mathcal{D}$-valid formulas, no matter the size of $\mathcal{D}$? The results of §4.4 suggest that the answer to this question is no. We may suppose the relevant "pattern" to be exemplifiable by sets at any level of the cumulative hierarchy, and require only that exemplification of the pattern by a set be discernible by a first-order formula whose quantifiers are restricted to the set's transitive closure. No first-order formula could, in this way, univocally discern a pattern which would be adequate to Wittgenstein's notion of consequence, not without prejudging the number of objects. At least, that's the drift of Proposition 4.15. The underlying problem is that Wittgenstein's concept of consequence detects the difference between countable and uncountable domains: by Proposition 4.13, there is even a single $\mathcal{LF}$-formula which is valid iff the domain is uncountable. Whether some set is countable doesn't depend only on its internal structure, but on the extrinsic matter of which functions happen to exist in the mathematical universe. The sensitivity of consequence to distinctions in the transfinite suggests that there is no reasonable notion of symbolic pattern whose exemplifications could serve to realize all instances of the consequence relation—at least not without some *a priori* constraint on the number of objects.

Thus, the reason Wittgenstein's concept of consequence cannot in general be realized in some univocally specified kind of symbolic pattern is that so conceived, consequence depends on distinctions of transfinite cardinality. It might be wondered whether this sensitivity to the transfinite could be trimmed away without great loss. But its origins were already recorded way back in Proposition 2.3: that Wittgenstein's concept of consequence embeds the concept of truth. In turn, the embedding of truth in consequence is immediate from the book's core thesis, that a proposition is a truth-function of elementary propositions. So the results are deeply rooted.

To conclude, Wittgenstein's early conception of philosophical progress requires some independent standard for a logical relationship's being made manifest in signs. The evolution of the general form of a proposition into the general form of a truth-function looks like a gesture toward such a standard. Yet that attempt proceeds without any concerted application of logic, and hence avowedly without prejudice to the size of the universe. The results of this paper suggest, however, that without prejudging the number of objects, no single concept of manifestation could be adequate to Wittgenstein's concept of consequence.

**§6. Further work.**

1. *Formal procedures.* The implementation in $\mathcal{LF}$ of the form-series device does not accommodate everything Wittgenstein might have accepted as a formal procedure. For example, $\mathcal{LF}$ does not allow operator-signs to contain form-series expressions; nor does it accommodate many-place or multigrade operations. Could a richer operation-scheme maintain a tractable semantics but yield a more expressive logic?

I have also not attempted in this paper to clarify the interaction between the form-series device and the interpretation of objectual variables which is developed by Wehmeier (2004). The example of §2.3 gives an expression of the ancestral under the exclusive interpretation. But Wehmeier's translation schemes do not extend straightforwardly to $\mathcal{LF}$, and I do not know how to avoid the use of equality in proving the adequacy of $\mathcal{LF}$ to finitary induction (see Lemma 3.4 and Propositions 3.5 and 3.10).

2. *The bar notation.* Wittgenstein introduces a bar notation, which is said to convert a propositional variable $\xi$ into an expression $\overline{\xi}$ of the plurality of its values (see especially 5.501). In §2.3, we saw that form-series contexts complicate the interpretation of free variables; I handled the problem somewhat brutally by dividing semantic evaluation into two stages. A more interesting approach is to introduce Wittgenstein's bar notation into the object-language. In a propositional sign, each enclosed propositional variable would have a bar written somewhere above it, with the priority of the expansions indicated by the vertical order of the bars. It is tempting to speculate that Wittgenstein's use of a bar, rather than an "inline" device, represents an attempt to escape the parse tree in some such way as we've seen 5.501 to require.

3. *Philosophy of arithmetic.* In this paper, I've argued that the form-series device permits definitions by induction on ordinary dyadic relations, like adjacency or temporal priority. Any such ordinary relation thereby generates material analogues of the system of finite ordinals, together with addition, multiplication, and indeed all arithmetically definable relations. Within this account, ordinal concepts are generated by an operation of taking the relative product of a relation with itself. So understood, "a number is an exponent of

an operation" (6.021). Frascolla (1997) offers another perspective on 6.021, developing an operational interpretation of the equational fragment of Peano arithmetic; he concludes that the interpretation cannot be extended to quantified arithmetical formulas. Can the two appearances of arithmetic be somehow understood as complementary?

4. *A presumption of countability*? Some commentators do already appear to find Wittgenstein to be committed in the *Tractatus* to the domain's countability (Ricketts, 2012). As I've already noted, there exist complete notions of "proof" for countably infinite truth-functional logic. The $N$-operator lends itself to elegant proof-theoretic analysis, particularly through a transfinite generalization of "bilateral" proof systems (Smiley, 1996 and Rumfitt, 2006). Could some such notion be relevantly found to underwrite a realization of logical consequence by symbolic patterns? And can a presumption of countability be squared with Wittgenstein's insistence that the number of names not be prejudged?

**§7. Acknowledgements.** Thanks to Nate Ackerman, Roberta Ballarin, Juliet Floyd, Warren Goldfarb, Michael Kremer, Rehana Patel, John Schlipf, Sanford Shieh, Ori Simchen, and Kai Wehmeier.

## BIBLIOGRAPHY

Ackermann, W. (1937). Die widerspruchsfreiheit der allgemeinen mengenlehre. *Mathematische Annalen*, **114**, 305–315.

Avron, A. (2003). Transitive closure and the mechanization of mathematics. In Kamareddine, F. D. editor. *Thirty five Years of Automating Mathematics*. Dordrecht: Springer, pp. 149–171.

Barwise, J. (1975). *Admissible Sets and Structures*. Berlin: Springer-Verlag.

Barwise, J. (1977). On Moschovakis closure ordinals. *Journal of Symbolic Logic*, **42**(2), 292–296.

Bays, T. (2001). On Tarski on models. *Journal of Symbolic Logic*, **66**(4), 1701–1726.

Büchi, J. (1960). On a decision method in restricted second-order arithmetic. In Nagel, E., Suppes, P., and Tarski, A., editors. *Proceedings of the International Congress for Logic, Methodology and Philosophy of Science*. Stanford: Stanford University Press, pp. 1–11.

Fine, K. (2010). Towards a theory of part. *Journal of Philosophy*, **107**(11), 559–589.

Floyd, J. (2001). Number and ascriptions of number in Wittgenstein's *Tractatus*. In Floyd, J. and Shieh, S., editors. *Future Pasts*. Cambridge, MA: Harvard University Press, pp. 145–193.

Fogelin, R. (1976). *Wittgenstein*. London: Routledge and Kegan Paul.

Frascolla, P. (1997). The "Tractatus" system of arithmetic. *Synthese*, **112**(3), 353–378.

Frege, G. (1879). *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Halle: Louis Nebert.

Geach, P. (1981). Wittgenstein's operator $N$. *Analysis*, **41**(4), 178–181.

Goldfarb, W. (2012). Wittgenstein against logicism. Unpublished.

Heck, R. (2011). A logic for Frege's theorem. In *Frege's Theorem*. New York: Oxford University Press, pp. 267–296.

Henkin, L. (1950). Completeness in the theory of types. *Journal of Symbolic Logic*, **15**(2), 81–91.

Hintikka, J. (1956). Identity, variables, and impredicative definitions. *Journal of Symbolic Logic*, **21**(3), 225–245.

Jech, T. (2003). *Set Theory* (third edition). Berlin: Springer.

Kirby, L. (2009). Finitary set theory. *Notre Dame Journal of Formal Logic*, **50**(3), 227–244.

Krivine, J.-L. (1998). *Théorie des Ensembles*. Paris: Cassini.

Lopez-Escobar, E. G. K. (1965). An interpolation theorem for denumerably long formulas. *Fundamenta Mathematicae*, **LVII**, 253–272.

Lopez-Escobar, E. G. K. (1966). On defining well-orderings. *Fundamenta Mathematicae*, **59**(1), 13–21.

Mancosu, P. (2010). Fixed- versus variable-domain interpretations of Tarski's account of logical consequence. *Philosophy Compass*, **5**(9), 745–759.

Martin-Löf, P. (1996). On the meanings of the logical constants and the justifications of the logical laws. *Nordic Journal of Formal Logic*, **1**(1), 11–60.

Mirimanoff, D. (1917). Les antinomies des Russell et de Burali-Forti et le problème fondamentale de la théorie des ensembles. *L'Enseignment Mathématique*, **19**, 37–52.

Potter, M. (2009). The logic of the *Tractatus*. In Gabbay, D. M. and Woods, J., editors. *From Russell to Church*, Volume 5 of *Handbook of the History of Logic*. Amsterdam: Elsevier, pp. 255–304.

Ricketts, T. (2012). Logical segmentation and generality in Wittgenstein's *Tractatus*. In Potter, M. and Sullivan, P., editors. *Wittgenstein's* Tractatus*: History and Interpretation*. Oxford: Oxford University Press, pp. 125–142.

Rogers, B. & Wehmeier, K. (2012). Tractarian first-order logic: Identity and the $N$-operator. *Review of Symbolic Logic*, **5**(4), 538–573.

Rumfitt, I. (2006). 'Yes' and 'no'. *Mind*, **109**(436), 781–824.

Russell, B. (1903). *The Principles of Mathematics*. London: Allen and Unwin.

Russell, B. (1905). On denoting. *Mind*, **14**(56), 479–493.

Russell, B. (1918). The philosophy of logical atomism. In Marsh, R. C., editor. *Logic and Knowledge*. London: Allen and Unwin, pp. 177–281.

Shoenfield, J. (1967). *Mathematical Logic*. Natick, MA: A K Peters.

Smiley, T. (1996). Rejection. *Analysis*, **56**, 1–9.

Soames, S. (1983). Generality, truth-functions, and expressive capacity in the *Tractatus*. *Philosophical Review*, **92**(4), 573–589.

Sullivan, P. (2004). The general form of the proposition is a variable. *Mind*, **113**(449), 43–56.

Sundholm, G. (1992). The general form of the operation in Wittgenstein's *Tractatus*. *Grazer Philosophische Studien*, **42**, 57–76.

Tarski, A. (1936). O pojeciu wynikania logicznego. *Przeglad Filozoficzny*, **39**, 58–68.

Väänänen, J. (2001). Second order logic and foundations of mathematics. *Bulletin of Symbolic Logic*, **7**(4), 504–520.

Wehmeier, K. (2004). Wittgensteinian predicate logic. *Notre Dame Journal of Formal Logic*, **45**, 1–11.

Wehmeier, K. (2008). Wittgensteinian tableaux, identity, and codenotation. *Erkenntnis*, **69**(3), 363–376.

Wehmeier, K. (2009). On Ramsey's 'silly delusion' regarding *Tractatus* 5.53. In Primiero, G. and Rahman, S., editors. *Acts of Knowledge – History, Philosophy and Logic*. London: College Publications, pp. 353–368.

Whitehead, A. N. & Russell, B. (1910–1913). *Principia Mathematica*. Cambridge: Cambridge University Press.

Wittgenstein, L. (1921). Logische-philosophische abhandlung. *Annalen der Naturphilosophie*, **XIV**, 168–198.

Wittgenstein, L. (1922). *Tractatus Logico-Philosophicus*. London: Kegan Paul.

Wittgenstein, L. (1961). *Tractatus Logico-Philosophicus*. London: Routledge & Kegan Paul.

Wittgenstein, L. (1971). *Prototractatus*. London: Routledge & Kegan Paul.

Wittgenstein, L. (1979). *Notebooks 1914–1916*. Oxford: Blackwell.
Yablo, S. (1993). Paradox without self-reference. *Analysis*, **53**(4), 252–252.

DEPARTMENT OF PHILOSOPHY
BOSTON UNIVERSITY
BOSTON, MA 02215, USA
*E-mail*: max@whythis.net

https://doi.org/10.1017/S1755020316000472 Published online by Cambridge University Press