The Reasoner

Volume 16, Number 7
October 2022

thereasoner.org
ISSN 1757-0522

Contents

Guest Editorial

Traditionally, logic has been thought to be finite. In Griffiths and Paseau (2022: *One True Logic*, OUP), we argue that the tradition is wrong. When we consider the nature and purposes of logic, we see that it is in fact infinite, maximally so. The focus of this issue is infinitary reasoning, understood broadly to include infinitary logic as well. The nine articles collected here explore themes from our book, the history of infinitary logic, the mathematical study of infinite systems and the possibility of infinite reasoning in humans. We are hugely grateful to all of our contributors.

Owen Griffiths
University College London
A.C. Paseau
University of Oxford

Features: Focus on infinitary reasoning

Ancestral links

The word 'ancestor' has interesting links to other words. It seems equivalent to an infinite disjunction: you are my ancestor if you are my parent or grandparent or great-grandparent or ..., where the list continues *ad infinitum*. As such, the word 'ancestor' features in apparently valid arguments essentially containing infinitely many premises. Here is a variant of one on p. 102 of Griffiths and Paseau (2022: *One True Logic*, Oxford University Press):

Bob is not my parent.

Bob is not my grandparent.

Bob is not my great-grandparent.

⋮

———————

Bob is not my ancestor.

What are we to make of this argument? It looks valid, though not formally. To see its non-formal nature, change the word 'parent' to 'child' throughout and change no other words. The modified argument's premises would be true and its conclusion false if Bob were, say, my grandfather, so an ancestor of mine, but neither my child nor my grandchild nor my great-grandchild nor.... Hence the original, displayed argument's validity is conceptual rather than formal: it is owed not entirely to form but also to the specific meanings of the words contained.

The second feature of the argument is that its premise set is infinite, essentially so if the conclusion is to follow from it. That Bob is my ancestor cannot be ruled out by any finite subset of the premises: if you just include the first three, for example, you haven't ruled out Bob's being my ancestor because he is my great-great-grandparent. One might object that humans have only a finite number of ancestors. Our actual family trees are finite, whether or not you trace them back beyond homo sapiens (to include previous hominids and even beyond). But the range of possibilities the argument is supposed to take in is not limited in this way. We can imagine a world in which humans have existed forever. Such a world can yield a counterexample to the validity of the argument with any finite subset of the original premise set.

So: the argument before us is valid in a non-formal sense and its validity is owed to infinitely many premises. How should we capture its validity? Here is a natural thought: employ an infinitary logic. The logic $\mathcal{L}_{\omega_1\omega}$ is just like first-order logic except that it allows for countably infinite disjunctions and conjunctions. In $\mathcal{L}_{\omega_1\omega}$, we can easily capture the validity of the original argument augmented with the meaning premise that my ancestor is anyone who is my parent or my grandparent or my great-grandparent or .... Let the predicate P_1 formalise 'is my parent', P_2 'is my grandparent', and so on, let the predicate A formalise 'is my ancestor' and let the constant b formalise 'Bob'. Then the argument's formalisation in $\mathcal{L}_{\omega_1\omega}$ is:

$$\neg P_1 b$$

$$\neg P_2 b$$

$$\neg P_3 b$$

$$\vdots$$

$$\forall x(Ax \leftrightarrow \bigvee_{i\in\omega} P_i x)$$

$$\overline{\qquad\qquad\qquad\qquad}$$

$$\neg Ab$$

This argument is $\mathcal{L}_{\omega_1\omega}$-valid. Taking our foundational logic to be at least as strong as $\mathcal{L}_{\omega_1\omega}$, we can therefore explain the original argument's conceptual validity by supplementing it with a meaning premise and showing the formalisation of the resulting argument to be $\mathcal{L}_{\omega_1\omega}$-valid, as displayed. In short, the logic $\mathcal{L}_{\omega_1\omega}$ does a good job of explaining the original argument's conceptual validity.

What about the competition? Can a finitary logic match $\mathcal{L}_{\omega_1\omega}$? A finitary logic, at the very least, should not allow for infinitary disjunctions and conjunctions. The most standard logic, predicate or first-order logic, is finitary. It is also compact: any valid first-order argument has a valid finite sub-argument (i.e. a sub-argument with a finite premise set). Can first-order logic explain the conceptual validity of the original argument? No. And this for a general reason: the argument's validity cannot be explained by any compact logic.

To sketch why, let $\mathcal{L}$ be a compact logic. Consider what an $\mathcal{L}$-based explanation of the argument's validity presumably looks like. Let's call the original argument (the one exhibited in premise-conclusion form at the start of this article) $\mathcal{B}$. The $\mathcal{L}$-based explanation of $\mathcal{B}$'s conceptual validity will add a meaning premise to $\mathcal{B}$, turning it into the formally valid $\mathcal{B}^+$. It will then formalise $\mathcal{B}^+$ in $\mathcal{L}$, to yield an $\mathcal{L}$-valid argument that we may call $\mathsf{Form}(\mathcal{B}^+)$. Now notice that no finite sub-argument of $\mathsf{Form}(\mathcal{B}^+)$ is $\mathcal{L}$-valid, since it corresponds to a finite sub-argument of $\mathcal{B}^+$. And no finite sub-argument of $\mathcal{B}^+$ can be valid, because Bob's not being reachable by moving n links up my family tree, for any finite n, does not preclude Bob from being my ancestor. But $\mathcal{L}$ is by assumption compact, which contradicts the hypothesis that it can explain $\mathcal{B}$'s validity.

This is not a watertight argument, but a very plausible one nonetheless. We might be sufficiently moved by it to accept the following conclusion. To explain the conceptual validity of $\mathcal{B}$, we must *either* adopt an infinitary logic *or* a non-compact finitary one, such as second-order logic.

These sorts of arguments are developed in much greater detail in Griffiths and Paseau (2022). We called them bottom-up arguments in that book, because they rest on relatively light theoretical principles. They differ from top-down arguments, typically arguments about the nature of logical consequence and the logical constants, as discussed in Griffiths' contribution to this volume. And in Part II of our book, we supplemented the sort of bottom-up argument sketched here with further ones that support the first disjunct at the end of the previous paragraph: only infinitary logics will do the required job, i.e. underwrite the validity of arguments like $\mathcal{B}$ and its generalisations. Frege's definition of an ancestral relation famously used second-order logic, but it will not generalise in the required way. An infinitary logic is needed.

Mathematical logic has not given a central place to infinitary logics. Barwise and Feferman (1985: *Handbook of Model Theoretic Logics*, Springer-Verlag), which summarised the state of knowledge at the time, is a magnificent achievement, but logicians have not built on it in the way they might have. The author of the more recent textbook Marker (2016: *Lectures on Infinitary Model Theory*, Cambridge University Press) cites two reasons to be interested in the model theory of infinitary languages: 'One reason is that we get new insights about first order model theory. But the simplest answer [reason] is that there are many natural classes that are axiomatized by $\mathcal{L}_{\omega_1\omega}$-sentences' (p. 9). Marker's reasons are manifestly mathematical.

As well as mathematical reasons, there are also more philosophical ones to be interested in infinitary logics. These logics are required to capture the validity, be it formal or conceptual, of various arguments, including $\mathcal{B}$. That is one, but by no means the only, reason to accept infinitary logic. Others are offered in *One True Logic*. The one true logic, it appears, is highly infinitary.

A.C. Paseau
University of Oxford

Logical nature and infinity

'You are a reader of the *The Reasoner*, all readers of *The Reasoner* will enjoy *One True Logic*; *so* you will enjoy *One True Logic*.' This is a valid argument in virtue of its logical form. It has the form '*a* is *F*, all *F*s are *G*s; *so a* is *G*' and every instance of this form is valid.

A crucial aspect of logical form is an account of the logical constants. When I specified the form of the above argument, I replaced some expressions—'you', 'reads *The Reasoner*', 'will enjoy *One True Logic*'—with variables. But I kept 'all' fixed, since it is distinctively logical.

Alfred Tarski thought that the logical constants are terms of a general character, which don't have any particular subject matter of their own but which are required for correct reasoning about any subject matter. In modern terminology, they are *topic-neutral*.

Tarski, and others, have made this thought more rigorous by considering permutations. Consider set $\{A, B, C\}$ and a permutation mapping A to B, B to C and C to A. Some sets, such as $\{A, B, C\}$, are mapped to themselves by this permutation. They are *invariant* under it. Others, such as $\{A, B\}$, are not: it is mapped to the distinct set $\{B, C\}$. They are *variant* under the permutation.

We can think of relations as sets and define what it is for a relation to be permutation invariant. Consider a set containing the previous British Prime Minister Boris Johnson, the new British Prime Minister Liz Truss and her logician father John Truss. And consider the relation *being a logician*. The extension of this relation of the set in question is the singleton of John Truss. The image of this set under any permutation that maps him to his daughter is the singleton of Liz Truss. These sets are distinct so the relation fails to be permutation invariant. As a test for topic neutrality, this seems correct: *being a logician* is a relation with a particular subject matter.

This cannot be the end of the story, however. Consider the relation *being human* over the same domain. It is invariant under permutation, since all members of the set are humans. But, like *being a logician*, *being human* is not topic-neutral. To remove all sensitivity to subject matter, as Gila Sher showed, we must consider not just permutations of a domain but bijections to others of the same size. Consider a three-membered domain none of whose members are human. If we consider a bijection from the original domain to this one, the extension of *being human* is not preserved.

Call the resulting test one for *isomorphism* invariance. If a relation is isomorphism-invariant, it has a good claim to being topic-neutral. The result is an attractive account of topic-neutrality: philosophically motivated by traditional thoughts about the formality of logic and capable of rigorous treatment.

What has all of this got to do with infinitary reasoning? In his contribution to this volume, Paseau offers a bottom-up argument for infinitary logic, starting from particular cases. These thoughts about isomorphism invariance can be used to offer a top-down argument for the same conclusion, starting with general theoretical considerations about logic.

The first step in this top-down argument is provided by adapting a theorem from Vann McGee (1996: Logical Operations, *Journal of Philosophical Logic* 25, 567–80). Consider the logic $\mathcal{L}_{\infty\infty}$, which extends first-order logic by allowing, for any cardinal κ, conjunctions and disjunctions of κ-many conjuncts or disjuncts, respectively, and allowing existential and universal quantification over κ-many argument places. In essence, it is the most highly infinitary extension of first-order logic.

What McGee proved, roughly, is that a relation of the right type is isomorphism-invariant just when it is expressible by a formula of $\mathcal{L}_{\infty\infty}$. What is isomorphism-invariant, in other words, is just what is definable in terms of infinitary versions of operations like conjunction and universal quantification. McGee's theorem, therefore, allows us to make the crucial leap from isomorphism invariance to infinitary logic.

So far, we've discussed the logicality of *relations*, which are worldly entities individuated by their extensions. But the logical constants we invoked at the start to explain validity are linguistic, not worldly. The next step in our top-down argument, then, is how we can use isomorphism invariance to deliver verdicts about which expressions are logical constants.

We can treat the relevant expressions as having relations as their semantic values. Isomorphism-invariant expressions are then those that have isomorphism-invariant relations as their semantic values. How does this relate to logical constanthood? Clearly, we need some principles linking isomorphism invariance and logical constanthood. The most simple-minded would be: an expression is a logical constant iff it is isomorphism-invariant.

For various reasons, this seems implausible. One problem is that we've said nothing about *meaning*. Concerned, as we have been, with extension alone, we will judge anything coextensive with a logical constant to be a logical constant. Consider McGee's example of *unicorn negation*:

$$\mathcal{U}\phi =_{Def} (\text{not-}\phi \text{ and there are no unicorns})$$

There are no unicorns, so this is coextensive with ordinary negation and hence a logical constant, by the crude principle. But you might think that an expression which invokes unicorns in its meaning is a poor candidate for being a logical constant. The relationship between logical constanthood and isomorphism invariance is clearly rather subtle.

Fortunately, the top-down argument presented here doesn't rely on such a controversial principle. Rather we need: if a relation is isomorphism-invariant, then logic—the one true logic—should include a logical constant with that relation as its semantic value. This condition avoids the unicorn negation worry above. We require that if a relation is isomorphism-invariant, then there must be a logical constant to name it. So the isomorphism-invariant operation of negation needs to be named by at least one logical constant, e.g. '$\neg$'. If it happens to be named by several, e.g. '$\mathcal{U}$' as well as '$\neg$', that's just fine.

From McGee's theorem, we know that infinite resources are required for all isomorphism-invariant operations to be named in this way; first-order logic, for example, is not up to the task. But a logic at least as strong as $\mathcal{L}_{\infty\infty}$ is. So logic had better be infinitary, otherwise it fails to talk about some logical relation. And that's the top-down argument.

We began with thoughts about the formal nature of validity, developed this in terms of topic-neutrality, which we captured with isomorphism invariance. And that leads us to the one true logic's being highly infinitary. For more, I can't do better than refer you back to the argument with which we started.

Owen Griffiths
University College London

Variation on a theme by Griffiths and Paseau

Consider the following structure, called argument $\mathcal{A}$ in Griffiths and Paseau (2022: *One True Logic*, Oxford University Press, p. 92; all page references will be to this book):

There is at least one planet.

There are at least two planets.

$\vdots$

There are at least n planets.

$\vdots$

There are infinitely many planets.

Is the above structure a logically valid argument? Let us join the authors in thinking that it is. Now, consider what they say later:

> Relatedly, one could try to argue, in the same vein as in Chapter 5, that the argument with premises 'There are at least κ planets' for every cardinal κ and conclusion 'There is an absolute infinity of planets' is logically valid. (p. 132)

Call this argument $\mathcal{A}^*$. First, notice that the formal structure of the two arguments, $\mathcal{A}$ and $\mathcal{A}^*$, is exactly the same. The sequence indexing the premisses of $\mathcal{A}$ is the sequence of natural numbers, and the limit of that sequence is the first aleph. So, if the authors are right that $\mathcal{A}$ is valid, and its premisses are true, we must conclude that there is an infinite number of planets. Similarly, since the sequence indexing the premisses of $\mathcal{A}^*$ is the sequence of the cardinals κ, and if, of course, we believe in absolute infinity being their limit (more on this later), then by the same rationale we get that there is an absolute infinity of planets. Notice that the crucial terms in the two arguments—and those that differ between them—are names of numbers. So to suggest that $\mathcal{A}$ is logically valid while $\mathcal{A}^*$ is not must be to assume that there is a certain threshold in the hierachy of numbers, such that one and the same principle is 'logical' below it, but stops being logical above it. How could this be?

Griffiths and Paseau do not question the validity of $\mathcal{A}^*$. They do have some worries about it, though, mostly for reasons to do with their general thesis in the book, given that they are concerned there with arguments in extended forms of English (see below). As a consequence, they take pains to address the reasonable worry that no argument of English can have an infinite number of premisses (pp. 94–96). They argue— convincingly, I believe—that natural languages are structures that might be consistently thought of as having a denumerably infinite number of terms/sentences, and, so, structures that can deliver arguments such as $\mathcal{A}$. Moreover, and in between the presentation of argument $\mathcal{A}$ and the presentation of argument $\mathcal{A}^*$, they take even more pains to convince the reader that:

> To cut off possible extensions of English at some particular ordinal and declare that beyond this point there can be no others would be arbitrary (p. 104).

They need to disallow any such restriction for they intend to generalize $\mathcal{A}$ to $\mathcal{A}^\kappa$. Argument $\mathcal{A}^\kappa$ has premisses 'There are at least λ planets' for all $\lambda < \kappa$ and conclusion 'There are at least κ planets', for κ any infinite limit cardinal.

The authors are undecided about $\mathcal{A}^*$'s validity. Their reason is that the totality of $\mathcal{A}^*$'s premisses is class-size, and, so, they now need a class-size language in order to express the argument. But can any extended form of English be a class-size language, they wonder.

I find this worry exaggerated. After all, we need a leap of faith to come to terms with the view that natural languages are structures extendable to such a degree that, for any cardinal κ, there is a language with κ-many terms/sentences. If so, we cannot proclaim that it is the absence of any intuition regarding class-size languages that makes us reluctant here. We had no intuition regarding the infinite extensions of English that we have previously come to endorse; but although most of them are not even recursive, we ended up endorsing them.

Language is no reason for scaring us away from $\mathcal{A}^*$. Other things might be, as I now explain.

Consider the sequence of arguments $\mathcal{A}, \cdots, \mathcal{A}^\kappa, \cdots, \mathcal{A}^*$. First, notice that all of its elements follow the same pattern. In the premisses of each such argument, a multiplicity is put into an ascending order, and, by the validity of the argument, the multiplicity mentioned in the conclusion is also affirmed.

Second, notice that the arguments' respective conclusion contains a quantifier of the form 'There are (at least) . . .'. These quantifiers pick up, besides aleph-0, an infinite sequence of limit cardinals, plus absolute infinity, which, if we accept that $\mathcal{A}^*$ is valid, is these cardinals' limit. Now, crucially, notice that these quantifiers, as such, do not discriminate between sets, proper classes, absolute infinities, etc. So, one has to find a means to make them formally sensitive to 'absolute infinity' in a consistent way; otherwise, the sequence will keep going beyond $\mathcal{A}^*$, thereby generating a form of Cantor's paradox. Paradoxes will follow because absolute infinity, unlike the transfinite, is no limit, and for Cantor, at least, everything that reaches a limit can be propagated even further. Treating $\mathcal{A}^*$ as a proper class will not do, for not only do we need to make it impossible for it to belong to anything, but we also need to find some corresponding quantity for it. This is exactly what is implied by the validity and the conclusion in $\mathcal{A}^*$. What this conclusion implies is that there is a quantity of planets (an 'absolute infinity'), that cannot be measured by any of the numbers that appear in the sequence. But these were supposed to all be numbers.

The upshot is that we need some novel post-Cantorian semantics and formalism for 'absolute infinity' in order to make the sequence stop for good, so to speak. Another viable option would be not to shiver in front of the 'inconsistent multiplicity or absolute infinity', as Cantor called it (Letter to Dedekind, 28.7.1899), and embrace the contradiction instead. In fact, there are paraconsistent logicians who genuinely believe that this represents Cantor's own attitude on the matter.

Other possibilities are available too. Among them is that of abandoning the idea of absolute infinity altogether, and going back to the perpetual iteration leitmotif, so much cherished among set theorists of the past century. But if we do so, $\mathcal{A}^*$ will no longer be valid. And this will mean that the validity of the arguments within the sequence $\mathcal{A}, \mathcal{A}^\kappa, \ldots, \mathcal{A}^*$ does not rest entirely on logic, but also on the existence of the collections mentioned in their conclusions as separate, actual, well-defined entities.

Summing up, I think these are open questions that might deserve to be addressed in a sensible follow-up project to *One True Logic*. I end with a Cantorian jest for those who have read this magnificent book already: it takes a super$^\kappa$-human for the full experience of the validity of every argument $\mathcal{A}^\kappa$, but only God has full experience of the validity of $\mathcal{A}^*$.

Doukas Kapantais
Academy of Athens, Research Centre for Greek Philosophy

Löwenheim numbers as a measure of logicality

Current work in logic, at least in a very prominent strand, crucially relies on a theory of models that is, in turn, couched in set theory. Since Tarski formulated his model-theoretic definitions in the 1930s, set theory has progressed into a rich, complex, and, importantly, philosophically perplexing field. Its richness has become part of its self-definition, guiding mathematicians working in the field through principles such as *maximize* and *inexhaustibility*, as well as *richness*; see Maddy (1988: Believing the Axioms I-II, *Journal of Symbolic Logic* 53, 481–511). Logic, by stark contrast, traditionally comes with restrictions, limitations and boundaries. Logic is positioned by philosophers at an extreme: as the stringent underlying foundation of the mathematical edifice; or as the most general discipline; the most widely applicable; the least metaphysically involved; the least controversial, etc.

Criteria for logicality serve the purpose of keeping logic within its purported bounds. The criterion for logicality of invariance under isomorphisms, known as the Tarski-Sher thesis, is articulated in set-theoretic terms. It states that logical constants denote operations that are invariant under isomorphic underlying domains. (Here I slide over issues on the relation between logical constants and operations and on isomorphism invariance's necessity and logicality.) While accepted as a go-to criterion in various contexts (e.g. in linguistics), this criterion has aroused fierce opposition by philosophers of logic. The objections are varied and nuanced, but the majority can be presented as resistant to letting logic succumb to the unbridled extravagance of set theory. Yet, most of the critics of isomorphism invariance aren't prepared to give up *all* use of set-theoretic tools in logical semantics; the explication of logical consequence in model-theoretic terms is recognised as an important advances in the topic, and no one shall drive us out of this Tarskian paradise. The question is whether and how this criterion should be modified and restricted.

The Tarski-Sher Thesis is fully endorsed by Griffiths and Paseau in *One True Logic*. Griffiths and Paseau go further than either Tarski or Sher, and require that the One True Logic should include a constant for each isomorphism-invariant operation (normally, the criterion is used for determining which logical constants are available and permissible for use in systems for logic). This means, for example, that for each cardinality κ, the quantifier *there are at least* κ has to be included in the language. Set theorists debate which cardinals in the higher infinite exist, or are consistent with accepted axioms—but whichever those may be, they will be assigned a constant in the One True Logic. To be sure, the infinite-cardinality quantifiers are invariant under isomorphisms, and as such, are general: they do not distinguish between members of the domain. But they carry with them all the metaphysical weight of contemporary set theory, and have understandably been deemed problematic; see Bonnay (2008: Logicality and Invariance, *Bulletin of Symbolic Logic*, 14, 29–68).

On the other hand, if a set-theoretic background is employed, some assumptions regarding sets will invariably be made. Perhaps some assumptions will be more metaphysically objectionable than others, but it is debatable whether a strict line can be drawn. In *One True Logic*, Griffiths and Paseau write that from the point of view of ontology, "numbers and sets are as good as each other or as bad as each other, or so it seems" (p. 180). Logic, on this point of view, relies on the full set-theoretic hierarchy, whatever that may be. They submit that there's a firm boundary between logic and mathematics (the former is general and topic neutral and the latter is not), but this poses no limitation on cardinality quantifiers.

However, the choice between drawing a strict boundary between acceptable and unacceptable set-theoretic entities and rejecting any such relevant difference between them is a false one. A *graded* notion of logicality can account for the difference between quantifiers. On any occasion in which a system for logic is called for (e.g. assessing arguments' validity), one fixes some terms as logical according to the level of metaphysical or set-theoretic commitment one is willing to take on. The idea would be that the more set-theoretic structure is required in order to fix a term as logical, involving more metaphysical assumptions, the less logical it is.

Contemporary set theory provides us with a possible measure, using the notion of the *Löwenheim number* of a logic; see Sagi (2018: Logicality and Meaning, *Review of Symbolic Logic* 11, 133–59).

Definition. Let L be a logic. The *Löwenheim number* of L, $\ell(\mathrm{L})$, is the least cardinal μ such that any satisfiable sentence in L has a model of cardinality less or equal to μ if such exists. Otherwise, $\ell(\mathrm{L})=\infty$.

For example, by the Downward Löwenheim-Skolem Theorem, the Löwenheim number of first-order logic is $\aleph_0$. This means that facts about validity and logical truth in first-order logic, while defined on the full range of models, require only models of size up to $\aleph_0$ to be determined. Let $\mathcal{L}$ be first-order logic, and $\mathcal{L}(Q)$ be first-order logic with the quantifier Q added to its logical vocabulary. We measure the logicality of Q by the Löwenheim number of $\mathcal{L}(Q)$. The higher the Löwenheim number of a quantifier, a higher infinity of models is required for determining logical facts involving Q—more set-theoretic structure is needed and stronger metaphysical assumptions—thus, the less logical it is. This measure takes $\mathcal{L}$ as a baseline, and builds on it. Some examples:

- Let Q_α be the unary monadic quantifier "there are at least $\aleph_\alpha$ many". We have $\ell(\mathcal{L}(Q_\alpha))=\aleph_\alpha$ for each ordinal α.

- Let Q^W be the unary polyadic quantifier over binary relations such that $M \models Q^W xy\varphi(x,y)$ iff $\varphi(x,y)^M$ is a well-order. We have $\ell(\mathcal{L}(Q^W))=\aleph_0$.

- Let the Härtig quantifier I be the binary monadic quantifier stating equal cardinality of sets: I is a binary monadic quantifier such that $M \models Ix(\varphi x, \psi x)$ iff $|(\varphi x)^M| = |(\psi x)^M|$. $\ell(\mathcal{L}(I))$ is very high, and is independent of ZFC. $\ell(\mathcal{L}(I))$ is a fixed point of the function $\alpha \mapsto \aleph_\alpha$, and further, Magidor and Väänänen showed that it is consistent with ZFC both that $\ell(\mathcal{L}(I))$ is under the first weakly inaccessible cardinal and that it is above the measurable cardinal.
- Let Most be the binary monadic quantifier such that $M \models$ Most $x(\varphi x, \psi x)$ iff $|(\varphi x)^M \backslash (\psi x)^M| < |(\varphi x)^M \cap (\psi x)^M|$. Then we have: $\ell(\mathcal{L}(\text{Most})) = \ell(\mathcal{L}(I))$.
- The logic Griffiths and Paseau argue the one true logic must contain, $L_{\infty\infty}$, is at the very (infinite) end of our spectrum, as $\ell(L_{\infty\infty}) = \infty$.

(See Sagi 2018 for references and more examples).

We see that some everyday quantifiers (e.g. "Most") have complex set-theoretic ramifications in the higher infinite, which may tell against fixing them as logical. To be sure, the proposed measure is not the only consideration to appeal to when setting a logical system—but it does provide a nuanced approach to the problem of logicality in the set-theoretic setting.

Gil Sagi
University of Haifa

Early infinitary logic

Infinitary devices were used by some of the most important pioneers of modern logic. This often goes unrecognised, since *Begriffschrift* (1879) and *Principia Mathematica* (1910) involve only finitary expressions, and are the best known of the early systems of modern logic. But a broader view on the origins of modern logic can serve as a corrective.

Following Boole, there emerged a tradition of algebraic logic which included the school of logicians surrounding C. S. Pierce. Previous algebraists provided algebraic treatments of propositional logic and syllogistic, but with Pierce and his followers we see algebraic treatments of quantificational logic as we would recognise it today. Key innovations attributed to Frege were achieved independently here—in particular, the isolation of the quantifier and with it the capacity for multiple generality. The notation devised by Pierce's school is strikingly familiar. Predications are written x_i, with lowercase letters being predicates and indices denoting individuals. Existential quantifications are rendered $\Sigma_i x_i$, with Σ serving as an existential quantifier (and Π as universal quantifier). As the symbols indicate, quantification is understood in analogy to taking a disjunction (logical sum) or a conjunction (logical product) of claims: $\Sigma_i x_i$ is like an infinite sum $x_a + x_b + x_c + ...$ where $a, b, c, ...$ are all individuals in the domain (Pierce 1885: Algebra of Logic, *American Journal of Mathematics* 7, 195). Where the domain of discourse is infinite, Pierce stresses that quantification is not identical with infinitary disjunction or conjunction; his scruples were not shared by other algebraists. Schröder, the next major figure in this tradition, is happy to use quantification interchangeably with infinitary truth-functions (e.g. Schröder 1890–1905. *Vorlesungen über die Alebra der Logik (3. vols).* Teubner. vol. 3, 10). Moreover Schröeder, Löwenheim, and the early Skolem all explicitly countenanced infinitary expressions (Moore 1990: Proof and the Infinite, *Interchange* 21(2), 49-51). The use of such expressions was not limited to the algebraists. Zermelo and the early Hilbert are two other figures who use of infinitary expressions in some capacity (cf. Moore 1990: 51–52). But early amicability towards infinitary devices did not last. Hilbert later banned infinitary devices from his metamathematics (Moore 1990: 51). Skolem came in 1922 to prohibit infinite expressions in trying to make precise Zermelo's notion of a "definite proposition" used in his axiom of separation (Skolem 1977: Some remarks on axiomatised set theory. In van Heijenoort (ed.) *From Frege to Gödel.* Harvard.) Even Tarski, who would later go on to be a great (re)habilitator of infinitary devices, seems in the 1930s to have had little patience for them (Moore 1990: 53–54).

Why the change? The start of an explanation might go as follows. Two conceptions of logic are seen in the writings of modern logic's early innovators. One is the picture of logic as an algebra or calculus, the other is of logic as a language. These are not contradictory, both owing something to Leibniz. However in the algebraic tradition, where the picture of logic as a calculus was emphasised, there likely did not seem anything so objectionable about infinitary conjunctions and disjunctions. After all, infinite sums and products are indispensable in other areas of algebra; why object to their use in the algebra of logic?

Plausibly, the move away from infinitary expressions did not arise from this algebraic current of thought, but from the current that conceptualised logic primarily as language. This current was delayed in its influence; Frege was its first definite figure, and the unfavourable reception of *Begriffsschrift*, though perhaps sometimes overstated (cf. Vilkko 1998: The Reception of Frege's Begriffsschrift, *Historia Mathematica* 25, 412–422) is part of folklore. Perhaps less known is that Schröder wrote a review of the book (Schröder 1880: Review of Frege's Conceptual Notation. *Zeitschrift für Mathematik und Physik* 25, 81–94). He argues that Frege's system is no more powerful than algebraic systems of logic, and notationally marks a step backwards. Schröder was mistaken on the specifics; he does not address Frege's treatment of multiple generality, which contemporary systems of algebraic logic lacked. But his argument is not wrong in principle. As mentioned, in 1879 Pierce's school was within a few years of publishing algebraic systems of quantificational logic that also handle multiple generality, with the same expressive power as Frege's system. Nevertheless, a distinguishing feature of *Begriffsschrift* was its presentation as a formal language, rather than as an algebra. It is this innovation which doubly ensures the importance of Frege's system in the history of logic. As the foundational crisis in mathematics took shape towards the end of the 19th century, so too did the need for a systematic investigation of metamathematics. It is little surprise that within this turn, a conception of logic as just another calculus would be far less appealing than a more distinctive conception of logic as a universal language in which mathematical theories could be precisely formulated. Thus this conception of logic came to eclipse the one adopted by the algebraic tradition.

Infinitary logic makes less sense under a linguistic conception of logic than it does under an algebraic conception. Infinitely long sums are common enough; infinitely long sentences are not. So it makes sense that, prior to the development of rigorous treatments of infinitary formal languages, logicians outside the algebraic tradition would be sceptical in temperament. We see this above in the younger Tarski. We see it also in Gödel who in one place criticises "the fiction that one can form propositions of infinite (and even non-denumerable) length", going on to say:

> [W]hat else is such an infinite truth-function but a special kind of an infinite extension (or structure) [...] with a hypothetical meaning, which can be understood only by an infinite mind? (Gödel 1944: Russell's Mathematical Logic in Schilpp (ed.) *The Philosophy of Bertrand Russell*, Northwestern University Press, 142).

This attitude was only overcome via the formal investigation of infinitary languages, which are now increasigly well understood. One wonders to what extent the marginalisation of infinitary logic in the first half of the previous century has delayed this understanding.

Michael Bevan
University of Colorado, Boulder

On *One True Logic*

If one takes the surface grammar of mathematics at face value, setting aside the entirety of natural language, one sees a mixture of logical idioms in play: first and second order quantifiers, quantifiers such as "there are infinitely many"; the language is both relational and functional; categorical and set-theoretical terminology can often lurk in the background; a rejection of classical disjunction occasionally emerges. Added to this is perhaps the most frequently used logical idiom of all, namely "..." as in "$A_1(x) \wedge A_2(x) \wedge A_3(x) \ldots$" formalised in infinitary logic by means of (here) infinite conjunction or disjunction.

This seemingly harmless observation has the consequence that any single logic will necessarily fail to be an adequate reconstruction of the discourse, on account of the logic's being committed to specific syntactic features, e.g. order of quantification, etc. On the level of surface grammar, in other words, an easy case can be made in favour of pluralism.

The ideology of "One True Logic," presumably (this author has not yet read *One True Logic* so this note is not to be read as a commentary on that work) is that this linguistic motley of the mathematician is, at best, illusory; that there is a unique logic $\mathcal{L}$ which governs mathematical discourse and/or gives rise to a logical formalism into which that discourse can be reformulated. The motivation behind such formal reductions is well known and goes back to the therapeutic foundational programs of early 20th century, due, for example, to Hilbert.

This is the *prescriptive* approach. In this note I ask: what if we take the surface grammar of the mathematician seriously, and simply investigate the logical territory presupposed by the mathematician's natural language?

The territory is complex. If we hold up the mathematician's natural language against the various formalisations on the market, one sees immediately the great sensitivity of these formal systems, to even small perturbations on the side of syntax. Moving from relational to functional languages causes $0-1$ laws to fail; moving from second to first order quantification in the formulation of arithmetic, induces a failure of categoricity, and so forth. As a rule, formal systems tend to be in this sense *unstable.*

It is an interesting if rather underappreciated feature of the natural language discourse (of mathematics), that while to the logician these perturbations of syntax are highly significant, the discourse of the mathematician is unaffected by framework decisions of this kind, i.e. it is stable with respect to perturbations of syntax, or as this author calls it, *formalism free*. In terms of the prescriptive approach, one would think that the One True Logic would have to exhibit a similar stability in order for it to count as capturing adequately our mathematical discourse—never mind capturing human reasoning *tout court*.

There would be much to say here, impossible in the space allotted. As it turns out though, luckily, there is something to say beyond ideology: one can develop calculi to study syntax sensitivity, and these calculi are of independent interest. One such calculus is the following: a semantically presented mathematical object usually has a detectable underlying logic $\mathcal{L}$, often first order logic. One can change $\mathcal{L}$ for another logic, and ask if there is a change in the object. This tests the sensitivity of the object to the syntactic elements of $\mathcal{L}$. As an example, consider Gödel's constructible hierarchy L, which is built over first order logic. By a result of Myhill and Scott, if one builds L over second order logic, one obtains *HOD*, the hereditarily ordinal definable sets, and these are (consistently) a different inner model from the original L. Investigating L from this point of view has proved revealing; see Kennedy, Magidor and Väänänen (2021: Inner models from extended logics: Part 1, *Journal of Mathematical Logic* 21, Paper No 2150012) and (2022: Inner models from extended logics: Part 2, *Journal of Mathematical Logic*, to appear). For example, Lindström's characterisation of first order logic displays anomalies in this setting: logics close to first order are "misread" by L in the sense of yielding a different inner model, whereas logics far from first order according to the Lindström characterisation, simply return L back, i.e. L (mis)reads them as being first order; see Kennedy (2020: *Gödel, Tarski and the lure of natural language: Logical entanglement, formalism, freeness*, Cambridge University Press).

A second calculus, based on the concept of *symbiosis*, aims at studying the set-theoretical entanglements of a logic, in particular the entanglement of a logic with a concept of set theory such as "x is countable," "x is finite," "x is a cardinal number," "x is the power-set of y"; see Väänänen (1979: Abstract logic and set theory. I. Definability, in *Logic Colloquium '78 (Mons, 1978)*, vol. 97 of *Studies in Logic and the Foundations of Mathematics*, North-Holland, Amsterdam-New York, pp. 391–421). With symbiosis one is able to detect whether a logic "sees" the invariant content of a given set-theoretic predicate—recognises, one might even say, its meaning. And on the other hand the absoluteness of the logic is pinned to the absoluteness of the predicate—whence the name "symbiosis."

Recent debates concerning the comparative virtues of second order logic vs. set theory, for example, decry the entanglement of set theory with second order logic—insofar as it is admitted to exist at all; see Shapiro (1991: *Foundations without foundationalism*, Oxford University Press). Whereas from the symbiosis point of view, one can prove that second order logic is actually symbiotic with the power set operation. Baldwin (2018: *Model theory and the philosophy of mathematical practice*, Cambridge University Press) gives a thorough analysis of the entanglement of infinitary logics with set theoretic assumptions. Or to put it another way: it is useless to try to separate second order logic from set theory.

This, together with the constructibility case, demonstrates the difficulties involved in advocating a single formalism for mathematical discourse. The natural language discourse of the mathematician is *entangled* with various logics, and these entanglements are hard to pull apart.

Juliette Kennedy
University of Helsinki

Infinitary logic in the finite

The study of the expressive power of logics on finite structures is an area of research in the interface between mathematical logic and computer science. A finite structure is an object $\mathbf{A} = (A, R_1, \ldots, R_n)$, where A is a finite set and each R_i is a relation on A. Every relational database can be viewed as a finite structure. For concreteness, the focus here will be on finite graphs, that is, finite structures of the form $\mathbf{G} = (V, E)$, where E is a binary relation consisting of the pairs of nodes in V connected via an edge. The expressive power of a logic is gauged by its ability to express properties of finite structures, where a property of finite structures, such as "connectedness" of a graph (i.e., "every two nodes are connected via a path"), is identified with the collection of all finite structures possessing that property. Thus, given a logic L and a property P, the question is whether or not there is a sentence ψ of L, such that a finite structure $\mathbf{A}$ possesses the property P if and only if $\mathbf{A}$ satisfies the sentence ψ (in symbols, $\mathbf{A} \models \psi$).

On the collection of all finite graphs, first-order logic FO can express local properties, such as "every node has exactly five neighbors" and "every two nodes are connected via a path of length at most three", but FO *cannot* express "connectedness" or any other such property that requires some form of recursion to be computed. This limitation in expressive power can be overcome by enhancing the syntax of FO with infinitary connectives. Indeed, if $\varphi_n(x, y)$ is a FO-formula expressing the property "there is a path of length n from x to y", where $n \geq 1$ is a fixed natural number, then "connectedness" is defined by the expression $\forall x \forall y (\bigvee_{n \geq 1} \varphi_n(x, y))$. Note that $\bigvee_{n \geq 1} \varphi_n(x, y)$ asserts that there is a path from x to y; the ancestor relation discussed in Paseau's article in this issue is an instance of this. Now, the expression $\forall x \forall y (\bigvee_{n \geq 1} \varphi_n(x, y))$ is a formula of the infinitary logic $\mathcal{L}_{\infty\omega}$, the extension of FO obtained by augmenting the syntax with disjunctions and conjunctions that may range over sets of arbitrary cardinality. While $\mathcal{L}_{\infty\omega}$ can make interesting distinctions between infinite structures, it is too powerful to be relevant in finite model theory because *every* collection of finite structures closed under isomorphisms is definable by a sentence of $\mathcal{L}_{\infty\omega}$. Indeed, if C is such a collection, then $C = \{\mathbf{B} : \mathbf{B} \models \bigvee_{\mathbf{A} \in C} \psi_{\mathbf{A}}\}$, where $\psi_{\mathbf{A}}$ is a FO-sentence that defines $\mathbf{A}$ up to isomorphism (i.e., for every structure $\mathbf{B}$ we have that $\mathbf{B} \models \psi_{\mathbf{A}}$ if and only if $\mathbf{B}$ is isomorphic to $\mathbf{A}$ – it is an exercise in logic that such a FO-sentence $\psi_{\mathbf{A}}$ exists, if $\mathbf{A}$ is finite). Note that this property of $\mathcal{L}_{\infty\omega}$ is a version of McGee's theorem in the finite – McGee's theorem is discussed in Griffiths' article in this issue. Barwise (1977: On Moschovakis Closure Ordinals, *J. of Symbolic Logic* 42(2), 292-296) introduced the family $\mathcal{L}^{\omega}_{\infty\omega}$ of the *finite-variable infinitary logics* logics $\mathcal{L}^{k}_{\infty\omega}$, $k \geq 1$, consisting of all formulas of $\mathcal{L}_{\infty\omega}$ with at most k distinct variables (each of these k variables, however, may have infinitely many occurrences in a $\mathcal{L}^{k}_{\infty\omega}$-formula). The finite-variable infinitary logics were introduced to solve an open problem about inductive definability on infinite structures, yet they turned out to have numerous uses in finite model theory.

The first important feature of $\mathcal{L}^{\omega}_{\infty\omega}$ is that it can express "connectedness"; more broadly, $\mathcal{L}^{\omega}_{\infty\omega}$ can express every property of finite structures definable in least fixed-point logic LFP, a powerful extension of FO with a recursion mechanism; see Barwise (1977: 292-296). To illustrate this feature, consider again the property "there is a path of length n from x to y", for some fixed $n \geq 1$. The FO-formula that immediately comes to mind for expressing this property uses $n+1$ distinct variables; for example, "there is a path of length 4 from x to y" is expressed by the FO-formula $\exists z_1 \exists z_2 \exists z_3 (E(x, z_1) \wedge E(z_1, z_2) \wedge E(z_2, z_3) \wedge E(z_3, y))$. Yet, "there is a path of length n from x to y" can be expressed by a FO-formula $\theta_n(x, y)$ that uses just 3 distinct variables. The key idea is to systematically reuse variables to represent nodes in the path from x to y; for example, "there is a path of length 4 from x to y" is expressed by the formula $\theta_4(x, y) =: \exists z (E(x, z) \wedge \exists x (E(z, x) \wedge \exists z (E(x, z) \wedge E(z, y))))$. It follows that "connectedness" is expressed by the $\mathcal{L}^{3}_{\infty\omega}$-formula $\forall x \forall y (\bigvee_{n \geq 1} \theta_n(x, y))$.

The second important feature of $\mathcal{L}^{\omega}_{\infty\omega}$ is that its expressive power can be analyzed using combinatorial games; see Barwise (1977: 292-296) and Immerman (1982: Upper and Lower Bounds for Expressibility, *J. Comput. Syst. Sci.* 25(1), 76-98). The k-pebble game, $k \geq 1$, is played on two structures $\mathbf{A}$ and $\mathbf{B}$ by two players, called Spoiler and Duplicator. Each player has k-pebbles labeled $1, \ldots, k$. The Duplicator picks one of the two structures and places on or removes from an element of the structure one of their pebbles; the Duplicator then responds by a similar move on the other structure using their pebble with the same label. The Spoiler wins if at some point the mapping $a_i \mapsto b_i$ is a not a partial isomorphism, where a_i and b_i are the elements of $\mathbf{A}$ and $\mathbf{B}$ pebbled by the pebbles labeled i, $1 \leq i \leq k$. The Duplicator wins if they can maintain a partial isomorphism in perpetuity. For example, if K_k is the k-clique (i.e., the complete graph with k nodes), then it is easy to see that, for every $k \geq 2$, the Duplicator wins the k-pebble game on K_k and K_{k+1}, while the Spoiler wins the $(k+1)$-pebble game on K_k and K_{k+1}. The pebble games characterize definability in the finite-variable infinitary logics. Specifically, for every $k \geq 1$, a collection C of structures is definable by a sentence of $\mathcal{L}^{k}_{\infty\omega}$ if and only if for all structures $\mathbf{A}$ and $\mathbf{B}$, whenever $\mathbf{A}$ is in C and the Duplicator wins the k-pebble game on $\mathbf{A}$ and $\mathbf{B}$, we have that $\mathbf{B}$ is also in C. As an immediate consequence, the collection of all graphs containing a clique of size $k+1$ is not definable in the logic $\mathcal{L}^{k}_{\infty\omega}$, while this collection is clearly definable in $\mathcal{L}^{k+1}_{\infty\omega}$ and, in fact, in first-order logic with $(k+1)$ variables. Furthermore, since least fixed-point logic LFP is subsumed by $\mathcal{L}^{\omega}_{\infty\omega}$, the pebble games provide a tool for showing that certain properties are not expressible in LFP.

The last important feature of $\mathcal{L}^{\omega}_{\infty\omega}$ discussed here is that a 0-1 law holds for $\mathcal{L}^{\omega}_{\infty\omega}$ under the uniform measure; see Kolaitis and Vardi (1992: *Information and Computation* 98(2), 258-294). This means that for every sentence ψ of $\mathcal{L}^{\omega}_{\infty\omega}$, the asymptotic probability $\lim_{n\to\infty}\mu_n(\psi)$ exists and either 0 or 1, where $\mu_n(\psi)$ is the fraction of finite structures with n elements in their universe that satisfy ψ. For example, $\mu(\text{connectedness}) = 1$, which intuitively means that almost all finite graphs are connected. The 0-1 law for $\mathcal{L}^{\omega}_{\infty\omega}$ subsumes 0-1 laws for FO and LFP that had been established earlier. Moreover, it delineates the boundary of 0-1 laws for infinitary logics since the "even cardinality" property (i.e., "the universe has an even number of elements") has no asymptotic probability, but it is expressible in $\mathcal{L}_{\infty\omega}$ by the sentence $\bigvee_{n=1}^{\infty}\sigma_{2n}$, where σ_{2n} is the FO-sentence asserting that the universe has $2n$ elements.

In conclusion, infinitary logic in the finite may appear at first to be an oxymoron, yet a rich body of work can be produced by focusing on the "right" fragment, namely the finite-variable infinitary logics.

Phokion G. Kolaitis
University of California Santa Cruz and IBM Research

Frege's infinite hierarchy of senses

'Anna believes that Bob Dylan won a Nobel Prize' and 'Anna believes that Robert Zimmerman won a Nobel Prize' may differ in truth-value. How's that possible given that 'Bob Dylan' and 'Robert Zimmerman' refer to just the same person? This is (a version of) Frege's puzzle, one of the most important problems in the philosophy of language. Frege's own solution crucially involves ascribing to an expression not only a *reference* but also a *sense*, a way in which the entity referred to is being presented. In the belief reports above, we are, according to Frege, not so much asserting a relationship between Anna and the unique musician who goes by the names 'Bob Dylan' and 'Robert Zimmerman', but between Anna and the two distinct senses these names express. This solution, when generalized, has been taken to lead to an infinite hierarchy of senses. This infinite hierarchy of senses, in turn, has been taken to render a Fregean language unlearnable. Learning just a single expression, so the objection goes, would require a thinker to apply infinite cognitive resources. Luckily, for Frege, this objection can be resisted. His account, when properly developed, does not require an infinite hierarchy of senses.

Where exactly is the infinite hierarchy supposed to come from? Here's the standard story, endorsed by most commentators from Carnap to Kripke. It has three steps. First, Frege postulates a reference shift for expressions occurring in the 'that'-clause following an attitude verb. This is what allows him to treat such contexts as extensional, a major selling point of his account. In the example above, 'Bob Dylan' no longer refers to Bob Dylan, but to the sense it ordinarily expresses. Accordingly, since 'Bob Dylan' and 'Robert Zimmerman' differ in sense when occurring in ordinary ('direct' as Frege puts it) contexts, they differ in reference when occurring in attitude ('indirect') contexts. That they cannot be substituted for one another *salva veritate* in indirect contexts then no longer counts against the extensionality of these contexts. Second, if expressions shift their *reference* in indirect contexts, they must also shift their *sense*. After all, sense determines reference. So, in the above example sentence, 'Bob Dylan' not only assumes an indirect referent, distinct from its direct referent, but also an indirect sense, distinct from its direct sense. Third, note that we can iterate attitude operators, as in 'Berta believes that Anna believes that Bob Dylan won a Nobel Prize'. If the business of such operators is to induce a shift in sense and reference, the reasoning continues, then each additional operator must lead to an additional such shift. In a doubly indirect context, then, 'Bob Dylan' must refer to the sense it expresses in a singly indirect context. Accordingly, it must express yet another sense (its doubly indirect sense). Further attitude operators will lead to further shifts. Since we can always form a new sentence by prefixing yet another operator, 'Bob Dylan' (and, of course, any other expression) ends up associated with infinitely many indirect senses.

Is this a fair price to pay for a simple, extensional account of attitude ascriptions? Some have thought it prohibitively high. If each expression comes with infinitely many senses, then fully mastering even a single expression would take up infinite cognitive resources, so that a language that works along the lines described would be unlearnable (see Davidson (1965: Theories of Meaning and Learnable Languages, in *Inquiries into Truth and Interpretation*, 1984, Clarendon Press, 3-15)). However, this argument assumes that the infinitely many indirect senses associated with, e.g., 'Bob Dylan' are entirely independent of one another and need to be learned piece by piece. But why should this be the case? Burge (2005: *Truth, Thought, Reason*, Clarendon Press, Ch. 4) and Kripke (2008: Frege's Theory of Sense and Reference: Some Exegetical Notes, *Theoria*, 181-218) have suggested ways of rendering the hierarchy learnable. What their approaches share is the assumption that the n-ly indirect reference of an expression determines its $n + 1$-ly indirect reference: each direct sense (= singly indirect referent) is presented by exactly one singly indirect sense (= doubly indirect referent), which is presented by exactly one doubly indirect sense (= triply indirect referent), and so on. The hierarchy of indirect senses associated with, e.g., 'Bob Dylan' thus doesn't branch out as we move upwards but forms a single, straight column. On the Burge-Kripke view, there is then only two things we need to grasp in order to understand 'Bob Dylan' in any context it may occur in, no matter how indirect. First, we need to grasp the expression's direct sense, the sense at the foot of the infinite column. Second, we need a (perhaps implicit) grasp of a function or rule that gets us from a given sense in the column to the unique sense one level further up.

I think this approach succeeds in taming the infinite hierarchy. In doing so, it also renders the hierarchy entirely useless though. For recall the theoretical pay-off of the original reference shift. The original shift allows us to say that two expressions which co-refer in direct contexts ('Bob Dylan', 'Robert Zimmerman') no longer co-refer in (singly) indirect contexts. This works because the direct reference of an expression does *not* determine its (singly) indirect reference. The two names have the same direct reference, but *not* the same direct sense and thus, given the shift, *not* the same (singly) indirect reference. Now, in assuming that the n-ly indirect reference of an expression determines its $n+1$-ly indirect reference, the Burge-Kripke approach ensures, by its very design, that no further theoretical pay-off is to be had by any of the further shifts. On their approach, any two expressions which co-refer in singly indirect contexts also co-refer in doubly indirect contexts, and in triply indirect contexts, and so on. While the initial shift from direct to singly indirect reference allows us to draw more fine-grained semantic distinctions, the additional shifts, e.g. from singly indirect to doubly indirect reference, have no such effect. But then why even postulate these shifts at all? Why not simply say that each expression is subject to a one-off reference shift when placed in the scope of an attitude operator while prefixing further such operators has no additional semantic effect on the expression? (See Skiba (2015: On Indirect Sense and Reference, *Theoria*, 48-81) for an extended defence of this proposal).

If there is no use for the infinite hierarchy, we should only accept it if, for some reason, we must. Some have thought there to be such reasons. Burge (*ibid.*) takes a rejection of the hierarchy to conflict both with certain principles governing sense composition as well as with the possibility of providing a recursive truth theory with certain desirable features for a Fregean language. But both conflicts can be resolved (see Skiba (2015: 63-75)). The most convincing rational reconstruction of Frege's theory of attitude ascriptions thus only requires one or two layers of sense, depending on whether one takes the single, one-off reference shift which actually pays a theoretical dividend to be accompanied by a single, one-off sense shift. Either way, the infinite hierarchy of indirect senses can and should be avoided.

Lukas Skiba
University of Hamburg

Infinite reasoning and arithmetical undecidability

Gödel's first incompleteness theorem shows that any suitably axiomatized theory of arithmetic, if it is sound (proving only truths), fails to prove all the truths of arithmetic. For such a theory T, we can code syntactic information about the theory itself into the language of arithmetic, and construct T's *Gödel sentence*, G_{T}. This is true if, and only if, it is not provable from the axioms of T. Hence, if there were a T-proof of G_{T}, T would prove a falsehood. So if T is sound, it does not prove G_{T}, and G_{T} is true. The second incompleteness theorem and related results establish further limits on what such theories can prove, including Con_{T}, which is true if and only if T is consistent.

We must distinguish propositions which are unprovable merely in T from those which are *absolutely* undecidable. According to Gödel, 'the epithet "absolutely" means that they would be undecidable, not just within some particular axiomatic system, but by *any* mathematical proof the human mind can conceive' (1951: Some basic theorems on the foundations of mathematics and their implications, in Feferman *et al.* 1995: *Kurt Gödel: Collected Works Vol. III*: 304–323). The existence of absolutely undecidable arithmetical propositions would overturn the deeply-held convictions of some of the greatest mathematicians of recent history (including Hilbert and Gödel himself), and establish a limit on our ability to answer mathematical questions that is far closer to home than the distant reaches of higher set theory. In 'Gödel's Disjunctive Argument' (forthcoming in *Philosophia Mathematica*), I argue that Gödel's theorems, together with some other mathematics, imply the existence of such propositions.

The implication is not immediate. According to Gödel, undecidable sentences such as G_{T} and Con_{T} are *exactly as evident* as the axioms from which they are constructed (193?: Undecidable diophantine propositions, in Feferman *et al.* 1995: 164–175). Consider, for example, PA (standard first-order arithmetic), which we (presumably) know to be sound. By Gödel's second theorem, there is no PA-proof of Con_{PA}. But we know that if PA is *sound*, it is also consistent. So any evidence we have for the soundness of PA is *also* evidence for Con_{PA}. Hence there is some theory which we know to be sound, namely PA + Con_{PA}, which *does* prove Con_{PA}.

The incompleteness theorems, however, are really *incompletability* theorems; PA + Con_{PA} has its own consistency sentence, which is true and unprovable in the strengthened theory. But if we know that PA + Con_{PA} is sound, we also know that it is consistent, and the whole process starts up again. One might wonder—in connection with the theme of this issue—what happens if we repeat this process of reasoning, from truth to consistency, *infinitely* many times? We would keep building stronger theories, but would they remain forever incomplete? And if so, would the consistency sentences for these theories (and related propositions) really be exactly as evident as the humble axioms of PA?

A *reflection principle* is a procedure for iteratively adding new axioms to a theory, such that the soundness of the stronger theories which result is an evident consequence of the soundness of the initial theory. The Gödel sentence construction is an example of one such principle: PA is sound, so PA + G_{PA} is sound, so PA + G_{PA} + $G_{\mathsf{PA}+G_{\mathsf{PA}}}$ is sound, and so on *ad infinitum*. Interestingly, the result of *infinite reflection* for this principle still results in a well-behaved axiomatic theory, PA_ω, obtained from PA by ω-many iterated additions of the Gödel sentence construction. PA_ω is subject to the incompleteness theorems; its Gödel sentence, though true, is PA_ω-unprovable. But how do we even formulate the Gödel sentence for a theory like PA_ω, when the language of arithmetic does not include the symbol 'ω'?

The question is an important one. Suppose ϕ is some sentence the truth of which follows from the soundness of T. If you don't know that the truth of ϕ follows from the soundness of T, then ϕ might fail to be exactly as evident as T for you. In some cases, you might have independent evidence for ϕ which makes it just as evident as the soundness of T. But if not, then you could have a wealth of evidence for the soundness of T, but no means of bringing it to bear in order to demonstrate ϕ. Since the evidence that we have for the Gödel sentence of a theory obtained by iterated reflection on PA piggybacks on our ability to recognize *that* the theory is an extension of PA by iterated reflection, there is a real risk that the Gödel sentence of a theory like PA_ω might turn out to be absolutely undecidable, if we have no means of recognizing that sentence for what it is.

In order to formulate Gödel sentences and consistency sentences for theories which result from infinite reflection, we need a mechanism to code information about the ordinals in the language of arithmetic. The standard system for this is Kleene's $\mathcal{O}$, which assigns natural numbers as "notations" for every computable ordinal. Given the considerations above, the results of infinite reflection will only be as evident as PA in the cases where we can recognize that the numerical notation coding the axioms of the theory does in fact stand for an ordinal in $\mathcal{O}$.

In 'Gödel's Disjunctive Argument', I identify a particular hypothetical infinitary reasoning ability, relating to our ability to enumerate $\mathcal{O}$, or a part thereof. Using Feferman's completeness theorem, I show that if we have this ability, then there are no absolutely undecidable arithmetical propositions. At some point in a particular process of infinite reasoning, not only do all the Gödel sentences and consistency sentences become provable, but *every* arithmetical truth does. However, $\mathcal{O}$ is an exceedingly complex subset of $\mathbb{N}$ (it falls far short of being computably enumerable) and it would be, I argue, nothing short of miraculous if we had the ability to recognize the right notations.

Indeed, the entire issue of absolutely undecidable arithmetical propositions turns on the delicate issue of which notations we can "recognize". I show that if all arithmetical truths are provable, then we do indeed possess the infinitary reasoning ability in question, because the provability of all arithmetical truths includes the provability of those which encode the information that certain numbers *do* stand for ordinals in $\mathcal{O}$.

So, the evidence overwhelmingly favours the existence of absolutely undecidable arithmetical propositions. Not only does one strategy for establishing the absolute decidability of all arithmetical truths fail, because it relies on the unjustified positing of a miracle. *All* routes to establishing the absolute decidability of all arithmetical propositions are complicit. In the case of arithmetic, infinite reasoning promises the greatest possible reward: a complete number theory. Unfortunately, such infinite reasoning goes beyond what is possible for creatures like us.

Wesley Wrigley
London School of Economics and Political Science