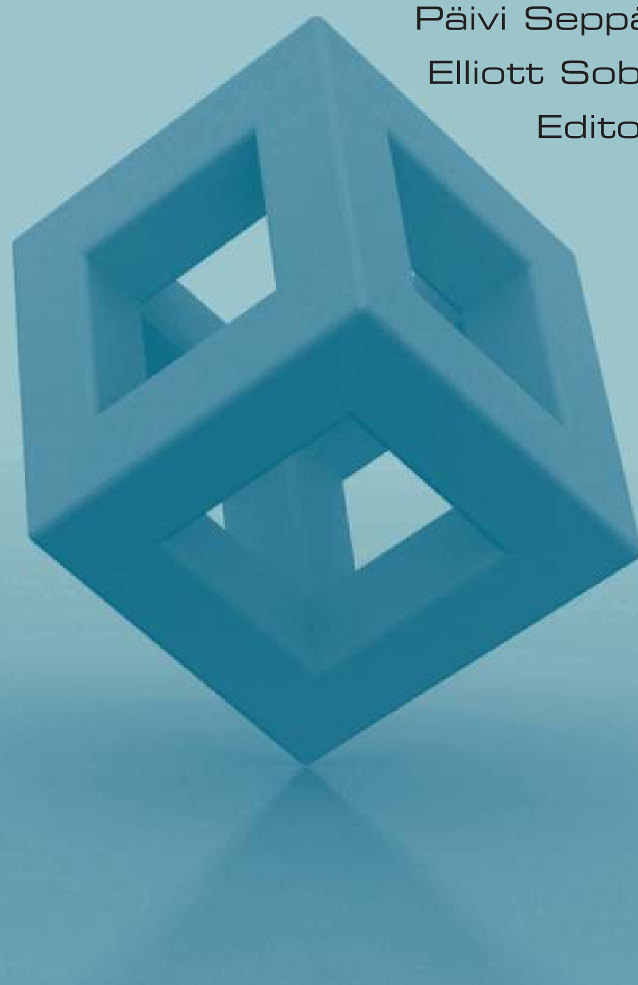
Logic, Methodology and Philosophy of Science

Proceedings of the Fifteenth International Congress

Hannes Leitgeb
Ilkka Niiniluoto
Päivi Seppälä
Elliott Sober
Editors



3 Squeezing arguments and strong logics

JULIETTE KENNEDY AND JOUKO VÄÄNÄNEN

Abstract. G. Kreisel has suggested that squeezing arguments, originally formulated for the informal concept of first order validity, should be extendable to second order logic, although he points out obvious obstacles. We develop this idea in the light of more recent advances and delineate the difficulties across the spectrum of extensions of first order logics by generalised quantifiers and infinitary logics. In particular we argue that if the relevant informal concept is read as informal in the precise sense of being untethered to a particular semantics, then the squeezing argument goes through in the second order case. Consideration of weak forms of Kreisel's squeezing argument leads naturally to reflection principles of set theory.

Keywords: generalized quantifiers, logical validity, reflection principles, second-order logic, squeezing principles.

1 Introduction

The foundational project is driven by the idea of modeling mathematical discourse, more or less globally, by giving an *adequate* formal reconstruction of it. Adequacy here, as in the phrase "adequate formal system," is delivered by the following: the formalism should be sound, complete, and at least to the degree possible, effective and syntactically complete. The formalism should also be *meaning-preserving*, relative to a given semantics (ideally). The idea of foundational formalism, as we called it in (Kennedy, 2013), is that with such a system in hand one could reasonably claim that the formalism has "captured" the informal discourse—whichever way one wishes to express this idea of "capturing."

At the same time the idea of considering just the informal mathematical discourse on its own, so to say *in situ*, has also attracted interest. This is implicit in so-called "practice-based" philosophy of mathematics—that practice is situated, after all, in natural language—while other philosophers take a more direct interest in natural language. M. Glanzberg, for example, in his (Glanzberg, 2015), argues that the notion of, e.g., consequence at work in natural language is to be distinguished from a genuinely logical consequence relation:

> The success of applying logical methods to natural language has led some to see the connection between the two as extremely close. To put the idea somewhat roughly, logic studies various languages, and the only special feature of the study of natural language is its focus on the languages humans happen to speak.
>
> This idea, I shall argue, is too much of a good thing.

Glanzberg is not only alerting us to the pitfalls of conceiving of natural language, at least in its logical aspects, as a kind of thinly disguised formal language, a matter of cleaning up the relevant definitions, concepts and so on. For Glanzberg, natural language and formal discourse are instead, with respect to their logical scaffolding anyway, two autonomous domains—"though the processes of identification, abstraction, and idealization can forge some connections between them." "Natural language has no logic" is the paper's central claim.[1]

What does it mean to say that natural language has no logic? For Glanzberg this is to deny *the logic in natural language thesis* (LNL thesis henceforth), namely the thesis that:

> A natural language, as a structure with a syntax and a semantics, thereby determines a logical consequence relation.

This is as opposed to the *logics in formal languages thesis* (LFL thesis henceforth):

> Logical consequence relations are determined by formal languages, with syntactic and semantic structures appropriate to isolate those relations.... Thus, the logics in formal languages thesis holds that consequence relations are *in* formal languages, in the sense that they are definable from them.

The LFL thesis is uncontroversial, if not trivial. The argument against the LNL thesis is subtle and turns partly on a critique of the model-theoretic account of logical consequence.[2] We will not consider Glanzberg's critique here, but will do so below, as the

[1] Glanzberg construes the term "logic" rather narrowly in the paper. The quote in the central claim is a deliberate reference to Strawson's (Strawson, 1950):

> Neither Aristotelian nor Russellian rules give the exact logic of any expression of ordinary language; for ordinary language has no exact logic.

[2] Glanzberg's arguments against the LNL thesis generally read "logic" as logical consequence.

issue becomes relevant to Kreisel's so-called squeezing arguments, to which we now turn.

2 Squeezing arguments and their critics

Squeezing arguments may be thought of as falling on the other side of the spectrum of belief in the LNL thesis (albeit tacitly). Introduced by Kreisel in his 1967 "Informal Rigour and Completeness Proofs" (Kreisel, 1967), and since taken up by W. Dean (Dean, 2016), H. Field (Field, 2008), V. Halbach (Halbach, 2016), P. Smith (Smith, 2011) and others, the arguments go as follows:

> Consider an informally defined mathematical concept I. Formally define two concepts A and B such that falling under the concept of A is a sufficient condition for falling under the concept of I, and falling under the concept of I suffices for falling under the concept of B. Thus $A \subseteq I \subseteq B$, where the inclusions are understood as applying to the extensions of the concepts A, B, I.
>
> Now suppose the formal notions A and B have the same extension. Then by the inclusions $A \subseteq I \subseteq B$ the informal concept I must coincide, again extensionally, with that of A and B.

For the informal concept I Kreisel took intuitive validity, denoted Val, understood as truth in all possible structures. This includes set and class-sized structures, as well as, in principle at least, structures that have no set-theoretical definition. Taking formal first order provability, denoted D_F, on the left, and taking truth in all set-theoretical structures,[3] denoted V, on the right, Kreisel argued as follows: By soundness, $\mathrm{D}_F \subseteq$ Val. By the fact that truth in all structures entails truth in all set-theoretical structures, Val $\subseteq$ V. Thus

$$\mathrm{D}_F \subseteq \mathrm{Val} \subseteq \mathrm{V}. \tag{3.1}$$

Invoking the completeness theorem for first order logic Kreisel concludes the following *theorem*, as he calls it, for α any first order statement:

$$\mathrm{Val}\,\alpha \leftrightarrow \mathrm{V}\alpha \text{ and } \mathrm{Val}\,\alpha \leftrightarrow \mathrm{D}_F\,\alpha.$$

Kreisel's presentation of the argument has been criticised in the literature. Smith (Smith, 2011) objects that Kreisel's somewhat model-theoretic construal of Val does not obviously capture the pre-theoretic notion in question, validity-in-virtue-of-form, as Smith prefers to think of Val.

[3] A set-theoretical structure is one whose domain, relations and functions are sets in the usual sense.

Field's criticism of the argument in (Field, 2008) involves the soundness claim, namely the first inclusion $\forall\alpha(\mathrm{D}_F\alpha \to \mathrm{Val}\,\alpha)$:

> In chapter 2 I argued that ... there is no way to prove the soundness of classical logic within classical set theory (even by a rule-circular proof): we can only prove a weak surrogate. This is in large part because we cannot even *state* a genuine soundness claim: doing so would require a truth predicate. And a definable truth predicate, or a stratified truth predicate, is inadequate for stating the soundness of classical logic, and even less adequate for proving it.[4]

Field goes on to prove soundness—or a weak surrogate of soundness—by means of a formal truth predicate, applying in restricted cases. Taken as a repair of Kreisel's argument one might argue that it ignores the methodology of the paper, which is heavily semantic (see below). Kreisel's argument does not depend on a proof of soundness in classical set theory. Kreisel is asking us to take soundness for granted, on the basis of *historical experience*—or as Kreisel puts it, *intuitive notions standing the test of time*. Instead, Field takes D_F as "primary", in Kreisel's terminology.[5]

As for Kreisel's own "proof" of soundness, extending to α^i (interpreting α as an *i*th order sentence) for all *i*, it amounts to arguing that the universal recognition of the validity of Frege's rules (D_F) at the time, together with the "facts of actual intellectual experience" accumulated subsequently, should amount to no more and no less than the acceptance of

$$\forall i \forall\alpha(\mathrm{D}_F\,\alpha^i \to \mathrm{Val}\,\alpha^i)$$

for us. And though a century of logical history has taught the logician nothing if not to be extremely suspicious of inclusions such as (3.1)—suspicions that Kreisel himself airs at the end of the paper—surely what Kreisel has in mind here is the idea that D_F was formulated *ex post facto*, that is precisely so as to *guarantee* soundness. D_F stood for Frege and his contemporaries, Kreisel claims, and stands for the contemporary mathematician too, as a completely adequate formalisation of Val.

Returning to our discussion of (Kreisel, 1967)'s critics, Halbach (Halbach, 2016) offers a repair of squeezing arguments in the form of a formal, syntactic substitutional notion of logical validity, to be substituted in for Kreisel's somewhat model-theoretic notion.[6] Such a concept of validity is, in Halbach's view, "closer to rough and less rig-

[4](Field, 2008, p. 191)

[5]From Kreisel (1967, p. 153):

> First (e.g. Bourbaki) 'ultimately' inference is nothing else but following formal rules, in other words D is primary (though now D must not be regarded as defined set-theoretically, but combinatorially). This is a specially peculiar idea, because 99 per cent of the readers, and 90 per cent of the writers of Bourbaki, don't have the rules in their heads at all!

[6]Halbach's conceptual analysis applies more widely, that is, it is an analysis of the "natural" concept of logical consequence in terms of substitutional validity überhaupt, i.e. not just in connection with squeezing

orous definitions of validity as they are given in introductory logic courses":

> I put forward the substitutional analysis as a direct, explicit, formal, and rigorous analysis of logical consequence. The substitutional definition of logical validity, if correctly spelled out, slots directly into the place of 'intuitive validity' in Kreisel's squeezing argument, as will be shown below.[7]

This is as opposed to the model-theoretic account of consequence, with its many (in Halbach's view) drawbacks:

> ...on a substitutional account it is obvious why logical truth implies truth *simpliciter* and why logical consequence is truth preserving. On the model-theoretic account, valid arguments preserve truth in a given (set-sized) model. But it's not clear why it should also preserve simple ('absolute') truth or truth in the elusive 'intended model'. Truth-preservation is at the heart of logical validity. Any analysis of logical consequence that doesn't capture this feature in a direct way can hardly count as an adequate analysis.[8]

Under a substitutional account, the connection with set theory is severed, or such is the claim; and interpretations of logical formulae are now syntactic objects:

> On the model-theoretic account, interpretations are specific sets; on the substitutional account they are merely syntactic and (under certain natural assumptions) computable functions replacing expressions.[9]

We do not wish to address the question here whether Halbach's is a reasonable conceptual analysis of the intuitive notion of logical consequence, the notion of consequence in itself. What seems clear to us is that the *informal* concept of consequence at work in natural mathematical languages is often plainly semantic, and moreover model-theoretic. That when the mathematician draws inferences in natural language, s/he imagines a situation in which the hypothesis is true—i.e. one has a model for the hypothesis in view—then s/he argues that the conclusion must hold in that model.

Kreisel states the point rather colorfully—"they don't have the rules in their heads at all!"—but what he means is that, e.g. group theorists do not derive theorems directly from the group axioms *in practice*, rather they employ the semantic method, i.e. they imagine a group and then show that the group has the property claimed for it. Analysts do not derive formal theorems from the axioms of real numbers, the real numbers are

arguments.

[7](Halbach, 2016)

[8](Halbach, 2016)

[9](Halbach, 2016). It is on account of this passage that we characterized Halbach's solution as syntactic. On the other hand, Halbach takes truth as a primitive in the paper, so in that sense his account is at the same time semantic.

taken as a *structure* which satisfies the axioms, and then theorems are proved about that structure. In fact this is what mathematicians are trained to do, as a cursory look at most standard introductory texts demonstrates. For example, Walter Rudin begins his classic text *Principles of Mathematical Analysis* (Rudin, 1976) with the existence of the field of reals:

> Theorem 1.19. There exists an ordered field $\mathbb{R}$ which has the least-upper-bound property.

The rest of the book is the investigation of this field $\mathbb{R}$. The proof of the existence of $\mathbb{R}$, i.e. the construction of the reals from the rationals, is relegated to an appendix.

For another example, Halsey Royden introduces the real numbers in his *Real Analysis* (Royden, 1988) with the following statement:

> We thus assume as given the set $\mathbb{R}$ of real numbers, the set P of positive real numbers and the functions "$+$" and "$\cdot$" on $\mathbb{R} \times \mathbb{R}$ to $\mathbb{R}$ and assume that these satisfy the following axioms, which we list in three groups.

As for the reasons behind the mathematicians' semantic mode of thought, this is in some sense the moral of Gödel's speed-up theorem:

> Thus, passing to the logic of the next higher order has the effect, not only of making provable certain propositions that were not provable before, but also of making it possible to shorten, by an extraordinary amount, infinitely many of the proofs already available.[10]

This can be interpreted as saying that the semantic method, the method of establishing logical consequence by considering models, enjoys a so-called "speed-up" over the method of formal proofs. This may explain why the model-theoretic notion of logical consequence seemed natural to Tarski and others.

The above-mentioned objections to (Kreisel, 1967) are certainly appropriate. Kreisel's notion of intuitive validity is clearly overly theorised, in the sense that the intuitive notion considered is not sufficiently intuitive or pre-theoretic, per Smith; and under theorised, per Field and Halbach respectively. At the same time though one might ask, if the intelligibility of Kreisel's squeezing argument depends on replacing the pre-theoretic notion of intuitive consequence by some formal, syntactic counterpart notion (as in (Field, 2008) or (Halbach, 2016)), what is the point of squeezing arguments at all? Why not simply analyse formal consequence directly, as logicians have always done?

For it seems to us that the interest of squeezing arguments lies in their being carried through in such a way as to fulfill what was originally claimed for them in (Kreisel, 1967), namely to capture an informal, natural language mathematical notion by "squeezing" it between two formal ones. It can be argued whether Kreisel himself

[10](Gödel, 1986, p. 397)

succeeded in this. The point is that if the aim was to capture (or "squeeze") informal notions used in practice—to provide a conceptual analysis of the *informal* notion of validity, as it were—then as we have argued above, Kreisel's model-theoretic construal of intuitive consequence was the correct and natural one. This is not to question the validity of the conceptual analyses of logical consequence which have been pursued so vigorously, especially in the period since Etchemendy's (Etchemendy, 1990), but rather to ask whether the concepts emerging from such analyses ought to replace the informal notions which serve as the object of Kreisel's analysis here, i.e. in the original squeezing argument. We will return to the issue of "genuine informality" below.

A final objection may concern what is surely the very unnatural restriction to the first order case, that is to propositions of the form α^1. The issue is addressed by Kreisel in the paper, who points to a partial result in this direction, an analogue of the squeezing argument derivable in the case of extensions of D_F to the ω-rule.[11]

Of interest to us here, and one of the topics of this paper, is the possible development of squeezing arguments in the direction of infinitary and second order logics. As it turns out somewhat more is known about completeness theorems for extensions of first order logic than was known in 1967. Going beyond the ω-rule, which Kreisel mentions, completeness theorems have been obtained for a number of infinitary logics, as well as for logics intermediate between first and second order. The second order logic perspective has also been developed a great deal since the publication of Kreisel's (Kreisel, 1967).

As a result of this logical work squeezing arguments of the kind Kreisel seems to be asking for in (Kreisel, 1967) may now be available. It is an oddity of the paper that while Kreisel mentions both the completeness of ω-logic and infinitary logics, he doesn't mention completeness theorems that would have been available already during the writing of (Kreisel, 1967), namely the Henkin Completeness Theorem (Henkin, 1950) for second order logic with the so-called Henkin semantics, and the completeness theorem for $L(Q_1)$, the extension of first order logic by the quantifier "there are uncountably many" due to Vaught (Vaught, 1964) and published in 1964. We will return to these logics below.

3 Squeezing arguments and the logic in natural language thesis

Before we consider the possibility of expanding squeezing arguments in the direction of second order and other strong logics, we ask, are the relevant natural language concepts available for this analysis at all, i.e. even in the first order case? Keeping

[11] The completeness of the ω-rule is due to Orey (1956).

Glanzberg's rejection of the LNL thesis in mind, can one simply extract what one thinks of as the notion of informal validity at work in the natural language mathematicians use, and devise a squeezing argument for that notion? In other words, do squeezing arguments require the LNL thesis?

Many researchers in the foundations of logic and mathematics may share a tacit belief in what one might call the "logic in natural *mathematical* language thesis". At the more formalist end of the spectrum of foundationalist views which have been pursued traditionally, one might even attribute a belief in the identity of the notion of logical consequence at work in natural mathematical languages with the notion defined by the relevant formal language. This is just the thought that the logical consequence relation defined by a suitable, maximally adequate formal language is the correct version of the logical consequence relation at work in natural language mathematics—what had been *meant* by the natural language concept all along. Tarski, though no formalist, seemed to argue for this or a similar view in his conversations on nominalism with Quine and Carnap at Harvard in 1940-41, when he remarked that "the difference between logic and mathematics" was that "Mathematics = logic + ϵ".[12]

Others, with Glanzberg, might see the notion of consequence at work in the mathematician's natural language as exact but fundamentally different from the formal notion. In that case one might ask, what separates formal entailment from its counterpart in (mathematical) natural language? We remarked above on Kreisel's observation that Frege's rules gained acceptance among mathematicians over time. This is to say, presumably, that if the relevant part of the discourse is formalized, then Frege's rules would be the mathematician's obvious choice of logical rules. Formalization also involves a choice of a semantics, but in contrast to the rules (D_F) it is not clear that a choice of semantics is determined by the informal practice in the second order case, which is our interest here. The view taken in this paper is that in the case of *intuitive*, informal second order validity, a choice of semantics is entirely irrelevant to the conceptual analysis of the notion of informal consequence, or so we will argue below. To be informal in the second order case necessitates prescinding from a choice of semantics.

We first consider the case of strong logics in general.

4 Squeezing arguments with completeness theorems

Kreisel's paper is entitled "Informal Rigour and Completeness Proofs," and indeed an apparently implicit assumption in (Kreisel, 1967) is that any time one has a completeness theorem in hand for a given logic, the associated squeezing argument should go through. More precisely, let $\mathscr{L}$ be a logic, and let $V_{\mathscr{L}}$ denote $\mathscr{L}$-validity understood

[12]Tarski is quoted in Carnap's notebooks. See Mancosu, (Mancosu, 2005).

in the standard semantic sense, that is set-theoretically. Let $\mathrm{Val}_{\mathscr{L}}$ denote informal validity (via $\mathscr{L}$), understood in Kreisel's sense, that is as referring to truth in all structures. Finally, let $D_{\mathscr{L}}$ denote derivability in the formal system introduced for $\mathscr{L}$.[13] Then one would expect that a completeness theorem for the logic $\mathscr{L}$ together with the inclusions:

$$D_{\mathscr{L}} \subseteq \mathrm{Val}_{\mathscr{L}} \subseteq \mathrm{V}_{\mathscr{L}}$$

should underwrite the extensional equivalence of the concepts $\mathscr{L}$-provability, informal $\mathscr{L}$-validity and $\mathscr{L}$-validity construed semantically.

Is this plausible? That is, can new squeezing arguments be obtained from completeness theorems for strong logics? The following completeness theorems for strong logics are known: in *ZFC*, completeness theorems have been obtained for $L(Q_1)$, the extension of first order logic by the quantifier "there are uncountably many", due to Vaught (Vaught, 1964), as was mentioned; for $L_{\omega_1\omega}$, the logic which is otherwise first order, but allowing conjunctions and disjunctions of countably many formulae, due to Karp (Karp, 1964); for so-called cofinality logic, the extension of first order logic by the quantifier denoted $Q^{cf}_{\kappa}xy\varphi(x,y)$, meaning "$\varphi$ defines a linear order of cofinality κ" for κ a regular cardinal, due to Shelah (Shelah, 1975); and for so-called stationary logic, the extension of first order logic by the quantifier denoted $\mathtt{aa}s\varphi(s)$, meaning "a club of countable sets s satisfies $\varphi(s)$", due to Barwise, Kaufmann and Makkai (Barwisw, Kaufmann, & Makkai, 1978). The last two require the Axiom of Choice for their completeness theorems.

Going beyond *ZFC* there is the logic $L(Q_2)$, the extension of first order logic by the quantifier "there are at least $\aleph_2$ many" proved complete by C.C. Chang (Chang, 1965), as pointed out by G. Fuhrken (Fuhrken, 1965), using the continuum hypothesis *CH*.

Another interesting case going beyond *ZFC* is the extension of first order logic by the Magidor-Malitz quantifier (Magidor & Malitz, 1977), defined as follows: $\mathscr{M} \models Q^{MM,n}_{\alpha}x_1 \ldots x_n\varphi(x_1,\ldots,x_n) \iff \exists X \subseteq M(|X| \geq \alpha \wedge \forall a_1,\ldots,a_n \in X \mathscr{M} \models \varphi(a_1,\ldots a_n))$. The completeness theorem for this logic uses the set-theoretical principle $\diamond$, which is stronger than *CH*, and it is consistent that completeness fails in the absence of $\diamond$ (Abraham & Shelah, 1993) .

The squeezing argument for the logic $L(Q_1)$, for example, would look like this: Let $D_{L(Q_1)}$ denote the concept of formal provability relative to this logic. Let $\mathrm{V}_{L(Q_1)}$ denote the truth of $L(Q_1)$-statements in all set-theoretical structures. Finally let $\mathrm{Val}_{L(Q_1)}$ stand for the validity of $L(Q_1)$-statements relative to all possible structures. Then if the inclusions

$$D_{L(Q_1)} \subseteq \mathrm{Val}_{L(Q_1)} \subseteq \mathrm{V}_{L(Q_1)}$$

[13] For many $\mathscr{L}$ it is obvious what $D_{\mathscr{L}}$ should be, but this is not always so.

hold, the squeezing argument relative to the logic $L(Q_1)$ must also hold, by the completeness theorem for $L(Q_1)$.

Strengthening the logic escalates one's set-theoretic commitments, clearly. The completeness theorem for first order logic is actually equivalent to Weak König's Lemma (WKL), which is also required to prove the completeness both of ω-logic and of $L_{\omega_1\omega}$. The Axiom of Choice is used for proving completeness theorems for the cofinality and stationary logics, corresponding to the generalised quantifiers $Q^{cf}_{\kappa} xy\varphi(x,y)$ and $\mathtt{aa}s\varphi(s)$. The *CH* is used for proving the completeness theorem for the logic $L(Q_2)$, and finally $\diamond$ is required for proving the completeness of the extension of first order logic obtained from the Magidor-Malitz quantifier. A comprehensive study of the exact nature of these commitments would seem to be in order, but is not our concern here.

Countenancing such a hierarchy of commitments is acceptable in some quarters and unacceptable in others—a matter of deciding whether the relevant completeness theorems "speak for themselves," to quote Kreisel.[14] What about the soundness claim in this advanced setting? In the first order case we claimed, in the spirit of informal rigour, that D_F was formulated so as to guarantee soundness—in fact Kreisel's soundness claim $\forall i \forall \alpha (\mathrm{D}_F\, \alpha^i \rightarrow \mathrm{Val}\, \alpha^i)$ extends to all orders, as we saw. There is no obvious reason why Kreisel's argument could not be extended to strong and infinitary logics. In the case of ω-logic, the order is somewhat reversed. That is, formal validity (V) is considered with respect to ω-models, in which the positive integer part is standard. Thus in the case of ω-logic the *semantics* is designed so as to underwrite the soundness of the omega-rule.

The case of second order logic in this regard is also striking, in that the Henkin semantics is formulated specifically so as to guarantee not soundness but completeness. We will now take up the question of whether squeezing arguments can be obtained in the second order case.

5 Squeezing for second order logic

The mathematician's informal discourse very naturally includes second order concepts—quantifying over functions and relations and so forth—so it is reasonable to ask for a squeezing argument for informal second order validity. But if a logic has a completeness theorem, then if the proof system of the logic is effective in the sense that the set

[14] Kreisel used this phrase in discussing the independence of the *CH* in the paper (p. 140):

> The present conference showed beyond a shadow of doubt that several recent results in logic, particularly the independence results for set theory, have left logicians bewildered about what to do next: in other words, these results do not 'speak for themselves' (to these logicians).

of axioms and rules are recursive and proofs are finite, then the set of valid sentences is recursively enumerable. By (Väänänen, 2001, Theorem 1) the set of valid sentences of second order logic is actually Π_2-complete in the Levy hierarchy. Thus on simple grounds of complexity no reasonable completeness theorem can exist for second order logic.

Does this mean one shouldn't pursue a squeezing argument for second order α? Kreisel himself took the view that "For higher order formulae we do not have a convincing proof of $\forall\alpha^2(\mathrm{V}\alpha^2 \leftrightarrow \mathrm{Val}\alpha^2)$ though one would expect one." We will now argue that a squeezing argument for second order α is available, once one incorporates the concept of validity with respect to general models.

Before addressing this point, recall that the Henkin semantics is defined simply so that in the so-called general models, the second order variables of a given formula are thought of as ranging over a fixed subset of the power-set of the domain. This subset of the power set may be a proper subset but it has to satisfy the axioms of second order logic, including the full comprehension axioms. In case the domain of quantification is actually the full power set, one refers to the model as "full" or "standard", and the associated semantics as full or standard semantics.

Taking the definition of Henkin semantics into account, a squeezing argument for informal validity of second order α would be the following:

Let D_Γ denote the usual axiom system of second order logic, already given in (Hilbert & Ackermann, 1928). As above, let $\mathrm{Val}\,\alpha^2$ mean that α^2 is informally true in all structures, including class-sized structures and including, in principle at least, structures that have no set-theoretical definition. Now let $\mathrm{V}\alpha^2$ mean that α^2 is true in all set-theoretical structures. The unproblematic implications are:

$$D_\Gamma\alpha^2 \to \mathrm{Val}\,\alpha^2 \to \mathrm{V}\alpha^2.$$

Note that if the completeness theorem held for second order logic, we could conclude straightway that $\mathrm{Val}\,\alpha^2 \leftrightarrow \mathrm{V}\alpha^2$ and $\mathrm{Val}\,\alpha^2 \leftrightarrow D_\Gamma\alpha^2$, as before. Now denote by $\mathrm{V}'\alpha^2$ the statement that α^2 is valid with respect to set-theoretically defined general models (satisfying—as we have assumed—the full comprehension axioms). Consider the following implications:

$$D_\Gamma\alpha^2 \to \mathrm{Val}\,\alpha^2 \to \mathrm{V}'\alpha^2 \to \mathrm{V}\alpha^2. \tag{3.2}$$

Suppose we assume (3.2). Then by Henkin's (Henkin, 1950) proof of

$$\mathrm{V}'\alpha^2 \to D_\Gamma\alpha^2$$

together with (3.2) we would obtain:

$$\mathrm{Val}\,\alpha^2 \leftrightarrow D_\Gamma\alpha^2 \leftrightarrow \mathrm{V}'\alpha^2.$$

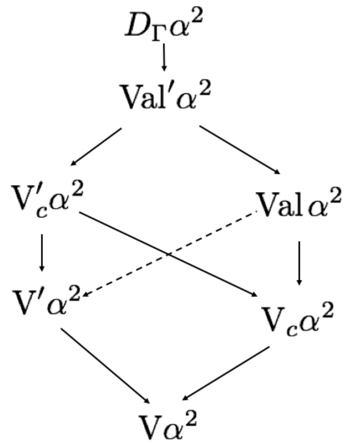


Figure 3.1: Varieties of validity in second order logic.

The first implication of (3.2) is clear, modulo the soundness claim, and so is the last, for trivial reasons. What about the middle implication Val $\alpha^2 \rightarrow \mathrm{V}'\alpha^2$? If α^2 is informally valid in all structures, why is it that general models should count as such structures? If the second order variables of α^2 are thought of as ranging over the full power set of the domain in question, why is it the case that these second order variables can be regarded as ranging over the subsets in a general model? Is there a principled way to distinguish non-standard general models from standard (full) structures, from the "informal practice" point of view?

To analyse the situation in more detail, let us write $\mathrm{V}_c \models \alpha^2$ if α^2 is valid in both class-sized structures and set structures. Let us also write $\mathrm{V}'_c \models \alpha^2$ if α^2 is valid in both class-sized general models and set-sized general models. Finally, let us write $\mathrm{Val}'_c\, \alpha^2$ if α^2 is informally true in all structures, full or general, including class-sized structures. Figure 3.1 depicts the trivial implications. The possible implication Val $\alpha^2 \rightarrow \mathrm{V}'\alpha^2$ is the problematic one.

So what distinguishes Val α^2 from $\mathrm{Val}'\alpha^2$? We claim that on the informal level it is impossible to see a difference between a standard model and a general model. It is true that if we consider a general model in isolation, from outside, so to speak, it is easy to imagine that something is missing from the model, in order for it to count as a standard model. For example, if we consider an infinite general model with a countable set of relations as the range of second order variables, we know that the model is not standard. There may be other ways of seeing the non-standardness from the outside. We may, for example "see" that a general model of arithmetic has an element with infinitely many predecessors. The position taken here is that it is contrary to the idea of informal validity that one should be able to look at the situation from outside.

One might still think[15] that there really is an *informal* concept of a general model, encapsulated by the thought: "All the sets I need are there and if some are missing, they do not change anything”. This would seem to be different from the informal concept of a standard model, encapsulated by the thought: "All the sets are there and no set, whether I need it or not, is missing”. If this is the case it is conceivable that for some α^2 we make the judgement that it is informally valid only in standard models, not in all general models. However, while it is crystal clear what the difference is between standard models and general models in the technical, logical sense, it is a different matter to see the difference on the informal level.

We go further and claim that on the informal level, the difference is not discernable. The reason for this is (essentially) that the general models "know" all the definable sets and relations (by the Comprehension Axioms) and they are the ones we refer to in mathematical practice.[16]

A similar line is articulated in (Väänänen, 2012), in which the second author has argued that from the point of view of mathematical practice, when we actually use second order logic we do not and in fact cannot see a difference between ordinary ("full”, or "standard”) models and general models.

> I will argue in this paper that if second-order logic is used in formalizing or axiomatizing mathematics, the choice of semantics is irrelevant: it cannot meaningfully be asked whether one should use Henkin semantics or full semantics. This question arises only if we formalize second-order logic after we have formalized basic mathematical concepts needed for semantics. *A choice between the Henkin second-order logic and the full second-order logic as a primary formalization of mathematics cannot be made*; they both come out the same.[17]

For example, let us consider Bolzano’s Theorem:

Theorem 1. *(Bolzano) Every continuous real function on* $[0,1]$ *which has a negative value at* 0 *and a positive value at* 1 *assumes the value* 0 *at some point of* $(0,1)$.

For the proof, by the second order comprehension axiom one can instantiate a universal second order quantifier at $X = \{x|f(x) < 0\}$. The set X is even first order definable, with f as parameter. This is a paradigm example: we operate on sets definable from existing sets. Of course, principles such as the Axiom of Choice force us to introduce also non-definable sets, but they do not exist because "all" sets exists but because we assume—and the general models are assumed to satisfy—the Axiom of

[15] We are indebted to A. Blass for suggesting this line of thought.

[16] Note that the definable sets taken on their own are not sufficient as they do not satisfy the Comprehension Axiom. One needs a little "blurring" around the edges, otherwise one can diagonalise out of the class.

[17] (Väänänen, 2012, p. 505, emphasis ours.)

Choice.

We now turn to the issue of set vs class-sized models. Consider the weaker claim that

$$\mathrm{V}\alpha^2 \rightarrow \mathrm{V}_c\,\alpha^2 \tag{3.3}$$

that is, the claim that second order formulas valid relative to set-theoretical structures are also valid relative to class-sized structures. In other words, we ask, is it true that if a second order sentence has a class-sized model, it also has a set-sized model? This cannot be proved from the axioms of von Neumann-Gödel-Bernays class theory (NGB), as the following "Zermelian" argument shows: Let α^2 be the second order sentence which says that the universe of the model is an inaccessible cardinal. Let κ be the least inaccessible and let M_κ denote the cumulative hierarchy up to κ. Then $\langle M_\kappa, \mathscr{P}(M_\kappa)\rangle$ is a model of NGB satisfying α^2. But no set-sized model, in the sense of $\langle M_\kappa, \mathscr{P}(M_\kappa)\rangle$, satisfies α^2.

Paul Bernays (Bernays, 1961) formulated more or less exactly (3.3), albeit in dual form, as a reflection principle, and observed that it implies the existence of inaccessible cardinals.[18] In fact, (3.3) implies a parameter-free version of so-called Levy's Schema (Lévy, 1960), which says that every definable closed unbounded class C of ordinals contains a regular cardinal. In the original Levy's Schema the definition of C is allowed to have parameters. Since the class of all cardinals is definable without parameters, we obtain from (3.3) a proper class of inaccessible cardinals. Bernays goes on to formulate (3.3) with second order parameters and arrives at what became later to be known as indescribable cardinals.[19] L. Tharp (Tharp, 1967) showed that the parametrized principle implies that for every n, the class of Π^1_n-indescribable cardinals is a proper class. This gave immediately a proper class of e.g. weakly compact cardinals. For an analysis and discussion of the situation we refer to Tait (Tait, 2005, Ch. 6).

Thus we cannot expect a proof of $\mathrm{V}\alpha^2 \rightarrow \mathrm{V}_c\,\alpha^2$, at least without additional axioms. On the other hand, the assumption (3.3) formulated in a reasonable class theory (such as NGB) seems plausible. By a result of Scott (Scott, 1961), it is true in the above $\langle M_\kappa, \mathscr{P}(M_\kappa)\rangle$, assuming that κ is not only weakly compact, but even measurable. In fact, it suffices to assume that κ is $\Pi^1_{<\omega}$-indescribable, hence (3.3) is consistent with $V = L$, assuming the consistency of a $\Pi^1_{<\omega}$-indescribable cardinal.

What about (3.3) for sentences α in other extensions of first order logic than second order logic? For first order logic this is an immediate consequence of the Levy Reflection Principle. For extended logics of the form $L(Q_\alpha)$ we can use translation to first order set theory and get the analogue of (3.3) as for first order logic. The same is true for $L(Q^{MM,n}_\alpha)$, $L(Q^{cf}_\kappa)$, and the extension of first order logic by the Härtig-quantifier

[18] We are indebted to A. Blass for pointing this out.

[19] In our model theoretic context second order parameters would correspond to adding generalised quantifiers to second order logic.

$Ixy\varphi(x)\psi(y)$, meaning: the cardinality of the set of elements x satisfying $\varphi(x)$ is the same as the cardinality of the set of elements y satisfying $\psi(y)$. For these powerful logics, unlike for second order logic, the analogue of the small part of the squeezing argument represented by (3.3) is simply provable in *ZFC*. The situation with stationary logic is more complicated. We leave the status of (3.3) open, if α^2 is taken to be a formula of so-called stationary logic rather than second order logic. The Open Question is, whether it is provable in *ZFC* or not.

Attempts to formulate higher order reflection with higher order parameters leading to larger large cardinals than $\kappa(\omega)$ have failed (see Koellner, 2009). However, a different approach, due to P. Welch, to a very strong reflection principle with second order parameters, called the *Global Reflection Principle*, gives a proper class of Woodin cardinals (Welch & Horsten, 2016; Welch, 2012).

6 Löwenheim-Skolem Theorems

Kreisel asks for a convincing proof of $\forall\alpha^2(\mathrm{V}\alpha^2 \leftrightarrow \mathrm{Val}\,\alpha^2)$, on its face impossible as we saw. Short of such a proof, Kreisel then asks a more specific question, which *can* be answered. Stating the Löwenheim-Skolem Theorem for first order logic in the form $\forall\alpha\forall\sigma > \omega(\mathrm{V}^{\omega+1}\alpha^1 \leftrightarrow \mathrm{V}^{\sigma}\alpha^1)$, what is the analogue to ω for second order formulae?[20]

First we recall some definitions. Given a logic $\mathscr{L}$, we say that $\mathscr{L}$ has Löwenheim-Skolem number κ if κ is the least cardinal such that for all vocabularies τ such that the cardinality of τ is $\leq \kappa$, if a sentence φ in the vocabulary τ of the logic has a model $\mathfrak{M}$, then it has a model $\mathfrak{N}$ of size $\leq \kappa$. In case $\mathfrak{N}$ can be taken to be a submodel of $\mathfrak{M}$ then κ is called the Löwenheim-Skolem-Tarski (*LST*) number of the logic.

Let $\mathscr{L}^2$ denote second order logic. We can now state Magidor's result (Magidor, 1971), which answers Kreisel's question: κ is the the least supercompact cardinal if and only if $\kappa = LST(\mathscr{L}^2)$.

In fact there is now a whole range of logics calibrated by large cardinals, in the sense that the assumption of the cardinal is equivalent to or implies a Löwenheim-Skolem-Tarski theorem for the logic. For the cases already mentioned the results are as follows: for cofinality logic, corresponding to the generalised quantifier $Q^{cf}_{\kappa}xy\varphi(x,y)$, the *LST* number is $\aleph_1$.[21] For stationary logic, corresponding to the quantifier $\mathtt{aa}s\varphi(s)$, the *LST* number is consistently $\aleph_1$, assuming the consistency of a supercompact cardinal,[22] but the *LST* number of stationary logic can also be the first supercompact

[20]In Kreisel's notation $\mathrm{V}^{\sigma}\alpha^1$ denotes the assertion "α^1 is true is the cumulative hierarchy up to σ".

[21](Shelah, 1975).

[22](See Ben-David, 1978).

cardinal.[23]

Finally, the interesting case of the Härtig quantifier: It is now known that if the *LST* number $LST(I)$ of this logic exists, then there is a weakly inaccessible cardinal and $LST(I)$ is at least the least weakly inaccessible cardinal. It is consistent relative to the consistency of a supercompact cardinal that $LST(I)$ is the first weakly inaccessible, and also consistent that it is the first supercompact.[24]

A general approach to strong logics and the reflection principles they give rise to is presented in J. Bagaria et al. (Bagaria & Väänänen, 2016), where a close connection is established between *LST* numbers of strong logics and so-called *structural reflection principles* in set theory.

Just as in the completeness theorems, and the ensuing squeezing arguments, obtaining Löwenheim-Skolem type theorems may require principles that go beyond Weak König's Lemma (WKL), sufficient in the case of first order logic.

7 Squeezing very simple concepts

Consider the concept W of finite words in a given vocabulary X. Intuitively we construct a word by placing letters from X one after another a finite number of times. What does this mean? We can use a squeezing argument to shed light on this question. As an analogue of derivability consider the concept D of starting from the empty word and then adding one letter from X at a time to the end of any word we already have. As an analogue of set theoretic validity we take the concept C of being a member of every closed set, where a set A is called *closed* if the empty word is in A, all one-letter words are in A, and the concatenation ww' of any two words w, w' of A are in A. Clearly,

$$D \subseteq W \subseteq C.$$

The first "$\subseteq$" is intuitively obvious because adding one letter to the end of a word certainly yields another word. The second is less obvious but one can run an informal induction on the length of the word to see that if A is closed, then the word is in A. It is a mathematical fact that

$$C \subseteq D,$$

because D is one of the sets that C is the intersection of. Hence

$$D = W = C,$$

[23] Magidor, unpublished.

[24] (Magidor & Väänänen, 2011).

and the informal concept of a finite word is squeezed between two (extensionally) identical exact concepts. Although everything in this squeezing argument in on a very elementary level, it is noteworthy that strictly speaking the inclusion $C \subseteq D$ is based on an impredicative argument.

Similarly, we may consider the concept F of a finite set. Intuitively we call a set *finite* if we can use some natural number to list the elements of the set. On the other hand, natural numbers can be identified with finite ordinals. Thus there is a certain amount of circularity in the concept of finiteness. So what does "finite" exactly mean? Let us take as D the concept of starting from the empty set and then adding one element at a time to get more sets. Let C be the concept of belonging to every *ideal class* i.e. to every class which contains the empty set, all singletons and is closed under unions of any two elements of the class. Clearly,

$$D \subseteq W \subseteq C.$$

The first "$\subseteq$" is again intuitively obvious because adding one element to a finite set certainly preserves the set finite. For the second one can use informal induction on the finite size of the set to see that if A is an ideal class, then the set is in A. It is a mathematical fact that

$$C \subseteq D,$$

because D is one of the classes that C is the intersection of. Hence

$$D = W = C,$$

and the informal concept of a finite set is squeezed between two (extensionally) identical exact concepts of class theory.

8 Conclusion

Do squeezing arguments capture the mathematician's informal discourse, even as it strays beyond first order talk, quantifying over relations and functions, and making implicit use of infinitary rules? This is difficult enough to argue for in the first order case. Nevertheless, we hope to have reinforced Kreisel's original argument in (Kreisel, 1967) that squeezing arguments have a general role in the conceptual analysis of informal mathematical concepts. Moreover, we have pointed out and given evidence to the claim that the circumstance that the two sides of the squeeze (extensionally) agree is based in general on a non-trivial mathematical fact.

In particular, we hope to have shown for strong logics that if we refashion the relevant informal concepts appropriately (here validity), we can, so to say, filtrate the informal discourse involving those concepts through a hierarchy of set-theoretic commitments ranging from Weak König's Lemma (WKL) up to $\diamondsuit$.

We also saw that various strategies present themselves in the second order case, that go beyond what Kreisel suggests, if intuitive second order validity is understood in the right way.

Acknowledments. This paper was written while the authors were participants in the Intensive Research Program - IRP Large Cardinals and Strong Logics at the Centre de Recerca Matemàtica during the fall semester of 2016. The authors thank the CRM for their support. Jouko Väänänen thanks the Simons Foundation for their support.

Bibliography

Abraham, U., & Shelah, S. (1993). A Δ_2^2 well-order of the reals and incompactness of $L(Q^{MM})$. *Ann. Pure Appl. Logic*, *59*(1), 1–32.

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

Barwisw, J., Kaufmann, M., & Makkai, M. (1978). Stationary logic. *Ann. Math. Logic*, *13*(2), 171–224.

Ben-David, S. (1978). On Shelah's compactness of cardinals. *Israel J. Math*, *31*(1), 34–56.

Bernays, P. (1961). Zur Frage der Unendlichkeitsschemata in der axiomatischen Mengenlehre. In *Essays on the foundations of mathematics* (pp. 3–49). Hebrew Univ., Jerusalem: Magnes Press.

Chang, C. C. (1965). A note on the two cardinal problem. *Proc. Amer. Math. Soc. 16*, 1148–1155.

Dean, W. (2016). Squeezing feasability. In *Pursuit of the universal (lecture notes in computer science)* (pp. 78–88). Springer.

Etchemendy, J. (1990). *The concept of logical consequence*. Cambridge, MA: Harvard University Press.

Field, H. (2008). *Saving truth from paradox*. Oxford: Oxford University Press.

Fuhrken, G. (1965). Languages with added quantifier "there exist at least $\aleph_\alpha$". In *Theory of models (Proc. 1963 Internat. Sympos. Berkeley)* (pp. 121–131). Amsterdam: North-Holland.

Glanzberg, M. (2015). Logical consequence and natural language. In *Foundations of logical consequence* (pp. 71–120). Oxford: Oxford University Press.

Gödel, K. (1986). *Collected works. Vol. I.* New York: The Clarendon Press, Oxford University Press.

Halbach, V. (2016). The substitutional analysis of logical consequence. Retrieved from https://mdetlefsen.nd.edu/assets/201030/halbach.consequence.pdf

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

Hilbert, D., & Ackermann, W. (1928). *Gründzuge der theoretischen Logik*. Berlin: J. Springer (Die Grundlehren der mathematischen Wissenschaften Bd. 27). VIII, 120 S.

Karp, C. R. (1964). *Languages with expressions of infinite length*. Amsterdam: North-Holland Publishing Co.

Kennedy, J. (2013). On formalism freeness: Implementing Gödel's 1946 Princeton bicentennial lecture. *Bull. Symbolic Logic*.

Koellner, P. (2009). On reflection principles. *Ann. Pure Appl. Logic*, *157*(2-3), 206–219.

Kreisel, G. (1967). Informal rigour and completeness proofs. In I. Lakatos (Ed.), *Proceedings of the International Colloquium in the Philosophy of Science, London, 1965, Vol. 1.* (pp. 138–157). Amsterdam: North-Holland Publishing Co.

Lévy. (1960). Axiom schemata of strong infinity in axiomatic set theory. *Pacific J. Math*, *10*, 223–238.

Magidor, M. (1971). On the role of supercompact and extendible cardinals in logic. *Israel J. Math*, *10*, 147–157.

Magidor, M., & Malitz, J. (1977). Compact extensions of $L(Q)$. Ia. *Ann. Math. Logic*, *11*(2), 217–261.

Magidor, M., & Väänänen, J. (2011). On Löwenheim-Skolem-Tarski numbers for extensions of first order logic. *J. Math. Log. 11*(1), 87–113.

Mancosu, P. (2005). Harvard 1940–1941: Tarski, Carnap and Quine on a finitistic language of mathematics for science. *Hist. Philos. Logic*, *26*(4), 327–357.

Orey, S. (1956). On ω-consistency and related properties. *J. Symb. Logic*, *21*, 246–252.

Royden, H. L. (1988). *Real analysis* (3rd ed.). New York: Macmillan Publishing Company.

Rudin, W. (1976). *Principles of mathematical analysis* (3rd ed.). International Series in Pure and Applied Mathematics. New York-Auckland-Düsseldorf: McGraw-Hill Book Co.

Scott, D. (1961). Measurable cardinals and constructible sets. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. 9*, 521–524.

Shelah, S. (1975). Generalized quantifiers and compact logic. *Trans. Amer. Math. Soc. 204*, 342–364.

Smith, P. (2011). Squeezing arguments. *Analysis (Oxford)*, *71*(1), 23–30.

Strawson, P. F. (1950). On referring. *Mind*, *9*, 320–344.

Tait, W. (2005). The provenance of pure reason. In *Logic and computation in philosophy*. Essays in the philosophy of mathematics and its history. New York: Oxford University Press.

Tharp, L. H. (1967). On a set theory of Bernays. *J. Symbolic Logic*, *3*, 319–321.

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

Väänänen, J. (2012). Second order logic or set theory? *Bull. Symbolic Logic*, *18*(1), 91–121.

Vaught, R. L. (1964). The completeness of logic with the added quantifier "there are uncountably many". *Fund. Math. 54*, 303–304.

Welch, P. (2012). Global reflection principles. Isaac Newton Institute preprint series NI12051-SAS.

Welch, P., & Horsten, L. (2016). Reflecting on absolute infinity. *Journal of Philosophy*, *113*(2), 89–111.

Author biography. Juliette Kennedy received her Ph.D. from the Graduate School and University Center of CUNY in 1996 with a thesis in mathematical logic. She is now University Lecturer (eq. Associate Professor) at the Department of Mathematics and Statistics at the University of Helsinki. She works in set-theoretic model theory and in the history and philosophy of mathematics.

Jouko Väänänen is a Professor of Mathematics of the University of Helsinki. He is also a Professor Emeritus of Mathematical Logic and Foundations of Mathematics of the University of Amsterdam. He works in several fields of logic, such as set theory, model theory, computer science logic, and foundations of mathematics. His current work is best described by his two books “Dependence logic”, CUP 2007, and “Models and Games”, CUP 2011.