

**Nils Kürbis: *Proof and Falsity: A Logical Investigation*, Cambridge University Press, Cambridge, 2019, viii+308 pp.. ISBN 978-1-108-48130-4**

Jaroslav Peregrin, Czech Academy of Sciences & University of Hradec Králové

The key message of Kürbis' book is that the enterprise of proof-theoretic semantics, as projected by Dummett and Prawitz, is bound to founder, because, as it currently stands, it offers no adequate accommodation for negation. Kürbis' diagnosis is that proof-theoretic semantics builds exclusively on "positive" notions, whereas negation is incurably "negative". In his own words (p. 122):

As it stands, Dummett's and Prawitz's theory appeals only to what might be called *positive* primitive notions: affirmation, assertion, truth. They avoid *negative* primitives, such as negation, denial, falsity. The latter are supposed to be *defined* in terms of the primitive notions of the theory. Put this way it may not be too surprising that the project fails: it is a natural thought that nothing negative can be got from the purely positive. The philosophical insight I will tentatively draw from this chapter is that *negative* as well as *positive* notions are needed in the theory of meaning.

This sounds clear and plausible: indeed, "nothing negative can be got from the purely positive". However, it is vital to clarify from the bigger picture what is being meant by the terms "positive" and "negative".

Let me first offer an explication of my own. As I argued elsewhere (Peregrin, 2010; Peregrin, 2014), giving a semantics for a language may be thought of as delimiting the *space of acceptable truth-valuations* of the language. This delimiting must be done by finite means, typically in terms of constraints of the form

if  $|s_1| = t_1$  and ... and  $|s_n| = t_n$ , then  $|s| = t$ ,

where  $s_1, \dots, s_n, s$  are sentences (usually  $s_1, \dots, s_n$  are parts of  $s$ ),  $|s|$  is the truth value of  $s$ , and  $t_1, \dots, t_n, t$  are truth values. (Think about the "if  $|s_1| = 1$  and  $|s_2| = 0$ , then  $|s_1 \rightarrow s_2| = 0$ " of classical logic.) Now if we consider the introduction or elimination rules proof-theoretic semantics works with, we can see that in the typical case they can be seen in the same manner, only with  $t_1 = \dots = t_n = t = 1$ . (Thus, e.g. " $s_1, s_2 \vdash s_1 \wedge s_2$ " can be seen as "if  $|s_1| = 1$  and  $|s_2| = 1$ , then  $|s_1 \wedge s_2| = 1$ .) In this sense, proof-theoretic semantics is restricted to the positive part of the delimitation of acceptable truth-valuations and arguably cannot capture some of those that are not purely positive. (Note, however, that, for example the above constraint "if  $|s_1| = 1$  and  $|s_2| = 0$ , then  $|s_1 \rightarrow s_2| = 0$ " can be turned into a positive one, namely "if  $|s_1| = 1$  and  $|s_1 \rightarrow s_2| = 1$  then  $|s_2| = 1$ ". Nevertheless, this is not possible for constraints such as "if  $|s_1| = 0$  and  $|s_2| = 0$  then  $|s_1 \vee s_2| = 0$ ".)

However, the question is whether such an explication is acceptable, for it pictures proof-theoretic semantics on the background of truth-theoretic (or possibly model-theoretic) semantics, which thus seems more fundamental. This does not seem to be something that a proof-theoretic semanticist would accept - Dummett or Prawitz would certainly insist that the very notion of truth is secondary to inference, rather than underlying it. As Dummett (1991, pp. 166–167) puts it, "without doubt, the source of the concept [of truth] lies in our general conception of the linguistic practice of assertion".

I am afraid that Kürbis takes it for granted that truth is something independent of proofs and inference, something self-standing at least in the minimal sense that it can serve as an anchor for our proof-theoretic enterprises. This is documented by some passages of the book. Thus on page 76 the author claims

There is a difference between rules of inference and rules of chess. Rules of inference are applied to sentences, which have meanings, whereas the pieces to which the rules of chess are applied do not.

I think this is plainly wrong, and it betrays that Kürbis is not really able to put himself into the shoes of a diehard proof-theoretic semanticist. From her viewpoint, meanings are sets of proofs, so a sentence is unable to have a meaning *before* such rules apply to it – it acquires meaning *in that* the rules are applied to it. Hence, just like pieces of wood become pawns, bishops or rooks *in that* they are subordinated to rules of chess, kinds of sounds or inscriptions become meaningful sentences *in that* they are subordinated to rules of inference. (I explained this in greater detail elsewhere – Peregrin, 2017.)

Kürbis continues (*ibid.*):

I have argued earlier that meaning and truth are closely related, as the sense of a sentence, or the thought expressed, can be understood as its truth conditions (...). That much truth enters proof-theoretic semantics, and it is that notion of truth that is preserved by the rules of inference justified by proof-theoretic semantics. This minimal amount of truth does not prejudice issues between realists and anti-realists, but it guarantees that the rules of logic are decisively unlike the rules of chess.

Certainly, a proof-theoretic semanticist can say that the rules of inference preserve truth, but not because there would be some truth outside of the enterprise of proving, which we would discover to be preserved by inference (or make inference preserve it), but rather because truth, *ex definitio*, is what is preserved by inference.

A similar viewpoint of the author is demonstrated on p. 111:

$\perp$  is intended to be a logical falsehood, a sentence that is false under any circumstances. Reading off its meaning from the rules governing it, the result should be that  $\perp$  must be false. Although this characterization of  $\perp$  appeals to semantics, it does not violate the intended semantic neutrality of the proof-theoretic justification of deduction. It is legitimate to appeal to semantic knowledge in order to see whether it has been reconstructed correctly in a theory of meaning. Looking from the outside, as it were, at speakers using  $\perp$  according to the rule *ex falso quodlibet*, is one bound to say that they cannot mean anything but a necessarily false sentence with it? The requirements that no semantic assumptions enter the theory is fulfilled, as no such assumptions enter the rule *ex falso quodlibet*.

Again, there is, according to Kürbis, "the outside, as it were" of proof theory, something that can be used as a touchstone. But according to proof-theoretic semantics, there is no such outside. Even truth is "inside".

Let me add that I think that we can make sense of the boundary between "positive" and "negative" notions without calling for any "outside" of proof theory. Such a boundary is provided by the so called logical expressivism (Brandt, 2000; Peregrin, 2014; Arazim, 2016), which says that logical constants are tools of making explicit the (material) inferential links that exist in language prior to their introduction. And while conjunction and a version of implication can be seen as such tools quite straightforwardly, what negation appears to help make explicit is not directly inferability, but rather non-inferability (see Peregrin, 2008).

However, these foundational qualms aside, Kürbis' book is interesting and thought-provoking. Once we agree about the boundary between "positive" and "negative" notions (however we want to explicate

it in detail), Kürbis' book provides a meticulous survey of the possibilities of including enough of the negative notion into proof-theory to make it viable. Also he argues that as there is no possibility of including negation within the pure proof-theoretic program, the arguments of proof-theoreticians against classical negation and classical logic fail.

Kürbis argues that the standard, Dummettian accommodation of negation within proof-theoretic semantics, where negation reduces to the implication of  $\perp$ , is unsatisfactory, as there is no way to adequately define  $\perp$ . His reasons were already sketched in the above quote, but let me add one more (*ibid.*):

What is it that makes a sentence from which everything follows unacceptable? The answer can only be that it is assumed that some sentences are false. If 'anything' ranges only over true sentences, there is nothing absurd about a sentence that entails that one might as well say anything. But it is a contingent feature of a language that some sentences are false.

Here again, Kürbis uses truth as a touchstone that comes from outside of language – but I am afraid that here, as Wittgenstein (1953), §102, would put it, "there is no outside; outside you cannot breathe". However, to be fair, Kürbis documents that Dummett himself is dissatisfied with his own treatment of negation and of  $\perp$ , and the problems he sees are not dissimilar to those pointed out by Kürbis. Hence the pursuit for some new "negative" element to be incorporated into the proof-theoretic framework to enable a more satisfying treatment of negation is certainly not unsubstantiated.

Kürbis goes on to argue that we cannot but enrich standard proof theory with some "negative" element; and he considers in turn various possibilities, namely, adding incompatibility, adding negation as a primitive notion, adding denial, and adding falsity. He rejects the first three as unsatisfactory and then embraces the last.

I am not quite comfortable with his explanation for his rejection of incompatibility as an additional primitive concept, because his reasons seem to me to miss their target. The Brandomian approach, which is one of those he discusses and with which I am familiar, has it that incompatibility (as well as inference) exists primarily in the material form, that is, among logically simple statements (while the logical vocabulary makes these links explicit). Given this, in a language without logical vocabulary, some statements may have their minimal incompatibles, others perhaps not. The introduction of the explicit negation helps produce a minimal incompatible for every statement – hence a statement that is logically equivalent to any already extant minimal incompatible.

Kürbis' "diagnosis" is the following (p. 173):

Incompatibility is supposed to be an undefined, primitive relation that holds between atomic sentences. Thus, whether it holds or not is a matter of the specific content of atomic sentences. To remedy the problem of the last sections, what would be called for is a general characterization of the notion of primitive incompatibility. But how are we supposed to generalize about when the relation of incompatibility holds, if the holding of the relation is essentially tied to the particular content of atomic sentences?

This, I think, is off the mark. We may assume that some atomic sentences are incompatible with other ones, just like we assume that some atomic sentences are inferable from others. But this is not what logic is interested in: what logic is after is how to extend these primitive relations to logically complex sentences. It tells us, e.g., that if  $X \vdash A$  and  $X \vdash B$ , then also  $X \vdash A \wedge B$  and similarly it tells us that if  $X$  is

incompatible with A, then  $X \vdash \neg A$ . We do not need to assume anything about the underlying relation of incompatibility, and certainly we do not need to assume that every atomic sentence has its minimal incompatible among the atomic sentences (just like we do not assume that for every two atomic sentences there must be an atomic sentence that is their conjunction). It is only after we constitute the operator of negation or conjunction in terms of the corresponding inferential rules (and, as the case may be, rules involving incompatibility) that we can assume this.

The possibility of adding negation as a primitive notion, discussed in Kürbis' next chapter does not make much sense to me. If the role of logic is to capture logical consequence or logical inference in terms of an explication of the inferential behavior of logical constants, then taking negation as primitive would seem to amount not to a step in this enterprise, but to giving it up. Why, if we were to take negation as primitive, could we not take also the other constants as primitive? Kürbis is less directly dismissive of such an approach, but in the end he dismisses it, too.

Another, relatively popular possibility is to accept, aside of the "positive" notion of assertion, also its "negative" counterpart, denial, which is now usually called bilateralism. Here Kürbis' conclusion is that bilateralism leads only to costs and no benefits, that it "fails to serve the purpose of justifying classical logic as the one correct logic". This is slightly puzzling: insofar as the issue is to provide for an adequate definition of negation, it is not obvious why it should involve "justifying classical logic". This would be entailed by the presupposition that the adequate negation is the classical one - but it does not seem that the author would endorse this presupposition explicitly.

Having rejected all these possibilities, Kürbis turns to his last, which he sees as the only viable possibility: adding falsity. He is convinced that proof theory rests, at least implicitly, on the concept of truth, and that we are at liberty both to add falsity, and also to embrace it without committing ourselves to embracing the explicit semantics of the truth functions. His point is that as the way in which proof theory rests on truth is that it stipulates truth-preserving inferences, the appropriate way for it to come to rest on falsity is to add falsity-preserving ones. This procedure, he argues, would allow us to reach, in effect, classical logic, and hence classical negation, without embracing the fully-fledged semantics with its basis of truth-tables.

Despite all the things I find disputable in this book, it remains a remarkably stimulating piece of work and readers interested in the perplexities of proof theory and proof-theoretic semantics will benefit from studying it. Kürbis succeeds, I think, in throwing fresh light on many of these perplexities and thus offers us some nourishing food for thought.

#### References.

- Arazim, P. (2016): 'Logical Expressivist's Logical Constants'. *Organon F* 23, 2–20.  
Brandom, R. (2000): *Articulating reasons*, Cambridge (Mass.): Harvard University Press.  
Peregrin, J. (2008): 'What is the logic of inference?'. *Studia Logica* 88, 263–294.  
Peregrin, J. (2010): 'Inferentializing Semantics'. *Journal of Philosophical Logic* 39, 255–274.  
Peregrin, J. (2014): *Inferentialism: why rules matter*, Basingstoke: Palgrave.  
Peregrin, J. (2017): 'Is inferentialism circular?'. *Analysis* 78, 450–454.  
Wittgenstein, L. (1953): *Philosophische Untersuchungen*, Oxford: Blackwell; English translation *Philosophical Investigations*, Oxford: Blackwell, 1953.