Introduction to special volume

Luis Rosa*

Intuitio presents here a special volume on subject matters related to *philosophy of logic*. This is not a popular philosophical topic at PUCRS, where people are mostly occupied with ethics and epistemology, but I decided it was worth giving it a try. I hope the readers of *Intuitio* will find this special volume instructive and thought-provoking.

It is worth noting that this is a *bilingual* publication. We have 2 papers in Portuguese (Rodrigo Cid, Renato Mendes Rocha) and 3 papers in English (Julio Lemos, Wang Jing, Alessio Moretti). The reason for writing in English nowadays is pretty clear: the range of people you can communicate with is wider and richer, and it possibly includes the range of people you can communicate with in another natural language (like Portuguese). Despite the political factors that led to the 'universalization' of English, this is a fact, and if we want to make a critical analysis of such factors our voices will be better echoed in English!

In what follows, I just make some brief introductory points about philosophy of logic.

1.

Logic can be characterized as the formal study of the *consequence* relation between propositions (or sentence-types). From a certain set of assumptions, such-and-such conclusion follows. To know what follows from what, we do logic - but *not only logic*. Logic by itself does not tell us what definitions of logical consequence (or *entailment*) are adequate for this or that purpose and neither does it tell us what is the correct (if any) interpretation of logical consequence. Also, logic by itself does not tell us if our systematic frameworks accurately formalize the domains of discourse they are supposed to formalize. In order to pursue these issues, we have to do philosophy of logic.

* Rutgers University / PUCRS. Contato: fsopho@gmail.com

At the same time, there is no good philosophy without logic[1]. Empirical inadequacy is not as serious a problem for a philosopher as a contradiction in his/her theory - and I mean an *unwanted, unexpected* contradiction (not one about which the philosopher has an explanatory story to tell - I have in mind the *dialetheists* here[2]). A *reductio ad absurdum* is end game - the philosopher has to abandon her previous theory and keep moving to new theses. Or, even when there is no contradiction, if it is shown that a philosopher's theory logically entails a consequence that she has reasons not to adopt, we expect the philosopher to go through theoretical changes. But this is not the only role that logic plays in philosophy. Logical refutation is just the *negative* role. When philosophers are not engaged in induction, they want their arguments to be *valid*. And it is by means of logic that we distinguish the valid arguments from the invalid ones. This is a *positive* role that logic plays in philosophical activity.

So logic needs philosophy and philosophy needs logic. Not quite paraphrasing Kant, we could say that logic without philosophy is empty and philosophy without logic is blind. I will focus on the "logic needs philosophy" side here - not on the "philosophy needs logic" side.

Philosophy of logic is an exciting area of investigation in contemporary philosophy. Since the breathtaking proliferation of logical systems from at least the first half of the 20th century until today, philosophers have been stimulated to ask a whole new set of questions. What is the system (if any) that accurately captures the logical properties of our talk about necessity and possibility? Can we have logical systems in which implication is not interpreted in such a way as to give rise to paradoxical or unwelcome consequences? What is the correct definition of *logical consequence*? Can we include in the domain of logic the study of arguments whose premises do not *guarantee* the truth of the conclusion? Should we still trust in logical systems that cannot be both, complete and consistent? How can we avoid trivialization of inconsistencies?

Then there are questions about the relation between logic and other relevant disciplines. Can first-order logic be used as a foundation for set theory and arithmetic? Is there any way we can build feasible moral systems (or legal systems) using modal logic? Can systems of logic be used as a basis for building rational artificial agents? How is inference, understood as a cognitive process, supposed to be related to argument, understood as an abstract logical structure? Which aspects of natural language (if any) can be understood by means of logical machinery? Corresponding to these questions, we have interface areas of investigation relating logic to mathematics, ethics and the law, AI, epistemology, cognitive science and linguistics. Also, questions in philosophy of logic about *modality*, *quantification*, *existence* and *identity* are intertwined with metaphysical issues, and it is hard to draw the line between philosophy of logic and metaphysics here.

[1] You might want to take a look at Theodore Sider's (2010) *Logic for Philosophy*. New York: Oxford University Press.
[2] See Priest, G. (2006) *In Contradiction*, 2.ed. New York: Oxford University Press.

But I am afraid that, even if we present such a list of questions - the 'pure' questions about logic and its foundations on the one side and the 'interdisciplinary' questions on the other - we may not have a good understanding of how philosophical investigations about logic work, and what are their objects of investigation. In the next section, I will try to illustrate a theoretical setting where important questions in philosophy of logic arise: the structuration of the logical systems themselves (with their syntax and semantics). In the last section I will briefly consider the philosophical problem about the nature of logic.

2.

Suppose you have in front of you a fully developed logical system - an *axiomatics*. Let us forget, for a moment, the *semantics* for this axiomatic system. We can imagine, for example, *Propositional Logic* (*PL*) *without* a model-theoretic structure (where attributions of truth-values to *formulae* are made) and without definitions of validity and entailment based on that structure. What is such a logical system made of? First, there is a basic set of symbols and syntactical rules about how to build well-formed formulas (*wffs*). Some combinations of symbols are *legal*, some are not. In general, we have a set of *atomic formulas* (the most basic kind of well-formed formula) and then we get more complex formulas by concatenating other *wffs* with operators in such a way as to give rise to new *wffs* (the syntactical rules are *recursive*). This is just the *language* part of the system (in order to do the logic, you have to 'speak logiquese').

Second, the system has a set of *axioms* - formulas that are assumed as true no matter what. One can always use an axiom in any part of an argument (more precisely, one will use an *instantiation* of the axiom schema). Here the axioms are just *postulates*. Finally, there is a set of *basic rules of derivation*. These rules are used in arguments to derive new formulas (the 'conclusions' of the arguments) either from the axioms or from new assumptions. The formulas that we derive from the axioms using these rules are the *theorems* of the system. As soon as we have proved a theorem, we can also use it anywhere in any argument. Further, when we show that from a certain type of assumption a certain type of conclusion can be derived, we can also derive rules of derivation. So the whole system is not only amplified by theorems - but also by new rules. Both theorems and derived rules are grounded in the axioms and in the basic rules of derivation.

Summing up, initially a logical system consists of: (1) a *language* with a vocabulary and syntax; (2) a set of *axioms*; (3) a set of *basic rules of derivation*. On the basis of such elements, a system also contains (4) a set of *theorems* and (5) a set of *derived rules of derivation*. At the pure axiomatic and syntactic level, a logical system is just like this - it just appears to be a 'symbol-handler', or maybe an empty game in which we derive formulas. At this point one may start to ask how can we "put meat on these bones". We have a syntactical structure and a derivation engine based on that structure, but the formulas we derive are *meaningless*, and we cannot even say that they are *true*!

Also, given our *axiomatics*, what gets to *validate* its axioms and its consequence relations? Here we need a *semantics* for that axiomatic system. Let me answer the question using the standard type of semantics for formal systems: a *model-theoretic* framework (also called '*truth-theoretic* semantics').

First, we flesh out a *model structure*. Each model M is supposed to 'make true' a certain set of sentences or formulas of a certain language L of a system S. The type of model will be dependent on the language used in the system. In the case of *Propositional Logic*, the models consist of *interpretation functions* that assign truth-values for *atomic formulas*. In the case of *First-Order Logic* (or *Predicate Logic*), the models consist of an ordered pair with a non-empty domain of objects and an interpretation function that assigns to each name of the language an object in the domain and to each predicate a set of individuals (or n-tuples for an n-place predicate) from the domain. In *Modal Logic*, a model is a tuple with three elements: a set of worlds, a set of *relations* between worlds and a valuation function that attributes truth-values *in worlds* to atomic formulas. And so on. In either case, we have a *model structure* such that these models are supposed to *validate* formulas in a language.

Second, we establish truth-conditions for formulas in these models. In general, the truth conditions for formulas are also recursive. We begin by offering the truth-conditions for *atoms* of the language L in a model M. So we state:

$M \vDash_S$ *atom* if and only if...,

where the dots stand for the truth-conditions of the formula *atom* in M of S. The symbol '$\vDash$' is the *semantic consequence* symbol. One can read '$M \vDash_S$ *atom*' as '*atom* is true in model M of system S'. On the basis of the truth-conditions for atoms, we can start fleshing out the truth-conditions for complex formulas of language L (the ones built according to the recursive rules of the used language). Once all truth-conditions are formulated, we have the *criteria* for deciding if any formula of language L is true in model M.

Third, we define validity and consequence relation. A formula is said to be *valid* in a certain model-theoretic framework if it is true in *every model*. So take a formula *formula* of the language L (the language used in system S). If we represent the claim that *formula* is valid in system S with '$\vDash_S$ *formula*', we have that:

$\vDash_S$ *formula* if and only if $M \vDash_S$ *formula* for every model M.

At this point, one could ask what it is for a model M *to be a model for* a system S. Part of what this relation means is understood in the following way: when a model M is a model for a system S, every axiom and every theorem of S is true in M. That is: a model M is a model for S when it makes all axioms and theorems of S true (notice that the *same language* must be used in system S and in the formulas made true by a model M). But, of course, this is not everything there is to say about the relation between M and S - we did not define the consequence relation yet.

We represent the claim that *formula* is a consequence in S of the set of formulas Γ with '$\Gamma \vDash_S$ *formula*'. It is important to realize that the form '$M \vDash_S$ *formula*', where M stands for a model and *formula* stands for a formula, does not represent the same thing as '$\Gamma \vDash_S$ *formula*', where Γ is a set of formulas. The former represent the fact that *formula* is true *in a model M of S*. The latter represent the fact that the truth of *formula* is a (semantic) consequence of the truth of all formulas in Γ (where *truth* must be understood as *truth in a model*). So we define:

$\Gamma \vDash_S$ *formula* if and only if there is no model M in which all sentences in Γ are true and *formula* is false.

Other way to put it: $\Gamma \vDash_S$ *formula* if and only if, for every model M, if all sentences in Γ are true then *formula* is true. We said before that part of what it means for a model M to be a model for S is for M to make true all axioms and theorems of S. Now we can complement this explanation by pointing out that, when M is a model of S, $\Gamma \vDash_S$ *formula* whenever *formula* can be derived from the set Γ in system S (notice that the *making true* of axioms and theorems referred above is just a special case of logical consequence with the empty set of premises). This property of logical systems with respect to a model (or family of models) is sometimes called "soundess".

If we represent the claim that *formula* is derivable from Γ in S by '$\Gamma \vdash_S$ *formula*', to say that S is *sound* is to say that:

(S) Whenever it is the case that $\Gamma \vdash_S$ *formula* it is also the case that $\Gamma \vDash_S$ *formula*.

Going the other way around, if

(C) Whenever it is the case that $\Gamma \vDash_S$ *formula* it is also the case that $\Gamma \vdash_S$ *formula*,

we have *completeness* - S is said to be *complete* when every logical consequence is derivable on it.

Summing up, a standard semantics for a certain system S consists of a model-theoretic structure, a set of truth-conditions for formulas in a language L (the one used in system S) and definitions of validity and consequence. Alternatively, one can work with a *proof-theoretic semantics* for a logical system, but it will be easier for us to consider issues in philosophy of logic by having in mind a *model-theoretic semantics*.

Given this basic picture - an axiomatic logic system on the one side and a semantics for the system on the other - we are now in a better position to point out some of the questions that are pursued in philosophy of logic. (Of course, this is not an attempt to exhaustively list the questions pursued in this area, as it is not an attempt to give a criterion for deciding if some kind of investigation can be properly classified as 'philosophy of logic'). So let us begin from where we started above: the language *L* for a certain system *S*.

3.

First, there are the representational questions about *L*: if it is expressive enough to represent the sentences in the domain of discourse it was supposed to capture (notice that here philosophy of logic becomes intertwined with philosophy of language). If, for example, we want to formalize arguments such as

Facebook is a social network
No social network has only one member
Therefore, Facebook do not have only one member,

we better use something else than the language for *Propositional Logic*. In such a language, the argument would be represented as follows:

p (*premise*)
q (*premise*)
r (*conclusion*)

But the above argument is not valid: the conclusion is unrelated to the premises (and it is not a tautology either). In the language of *Propositional Logic*, each different sentence must be represented by a different sentential letter, so that we have no resources to formally represent the original argument in such a way as to make it valid. And we are assuming that the original argument, formulated in natural language, *is valid* - only it is an informal argument. In this case, it is not difficult to realize that the language of *Predicate Logic* would do the work.

But questions about the language of the intended formal system are not restricted to simple questions about *levels of representation* with respect to natural language. There are also questions about which *logical constants* - *operators* and *connectives* - should be included in a satisfactory formal language. As an example, consider sentences in *intensional contexts*, such as:

Lois Lane believes that Superman can fly.

Would *First-Order Logic* (*FOL*) capture the logical properties of this sentence? Suppose one tries to formalize it with the language of *FOL* and, accordingly, to model the logical relations that this sentence is supposed to maintain with other sentences. One could then represent the sentence above with

$B(l, s)$,

where 'B' represents the relation expressed by '... believes that...', 'l' is a constant for the name 'Lois Lane' and 's' is a constant that stands for the proposition expressed by the sentence '*Superman can fly*'. The semantics for *FOL,* though, is such that if two constants have the same value in the domain of objects, then two sentences in which they occur interchangeably will also have the same (truth) value. Let us (*controversially*) *assume* that the proposition expressed by the sentence '*Superman can fly*' *is the same as* the proposition expressed by '*Clark Kent can fly*'. Further, let us use 'c' as a constant that stands for the proposition expressed by the sentence '*Clark Kent can fly*'. Now, if $B(l, s)$ is true in a model M whose interpretation function is such that it maps both s and c to the same object in the domain, then $B(l, c)$ is also true in M. But this is not right! We do not want to say that *Lois Lane believes that Clark Kent can fly* - only that she believes that *Superman can fly* (it is assumed that Lois Lane does not know that *Clark Kent is Superman*).

Can we use this line of argument to reject the assumption that the sentences '*Superman can fly*' and '*Clark Kent can fly*' express the same proposition? Maybe this assumption is wrong, maybe not - the point now is that we picked the wrong language for representing the original sentence. An argument to the conclusion that '*Superman can fly*' and '*Clark Kent can fly*' do not express the same proposition based on this representational choice would be flawed. The original sentence is in a higher-order context, where some person is related to a *semantic object* (a proposition), and *First-Order Logic* does not have the appropriate resources for dealing with such contexts. We need another language, and a new language will also require a different model-theoretic framework, a different semantics. This is a kind of problem that raises philosophical questions about interpretation of language and types of objects denoted by parts of natural language. Both, philosophy of language and ontology play a role here.

How to represent sentences in natural language, and to what level, constitute one kind of philosophical problem related to formal languages for logic systems. But there is more: given a certain set of symbols for our language L, it is natural for us to ask if the logical constants in such set 'behave' as we expect them to. What does that mean? Presumably, we are trying to capture some relevant properties of sentences in natural language, and these properties should be reflected both in the axiomatic system and in the semantics. Is the semantics for the negation '~' adequate for the natural-language word '*not*'? And, if there is more than one semantics for the negation, which one should we prefer? And to what purpose? Similar questions can be made with respect to other connectives such as the *material conditional* '⊃', for example. It is noteworthy, however, that the question about if a certain semantics for a certain formal language is 'adequate' needs philosophical treatment itself. Maybe no system with the operator '~' and such-and-such semantics can render true *all* the formal representations of sentences in natural language in which '*not*' is adequately used and such that we regard as true. Perhaps we need a restriction as to what class of uses of '*not*' are amenable to the formalization we are working with. But we not only want our semantics to 'make true' formal representations of sentences that we already regard as true. We also want the formulas with the appropriate connectives and operators to have the adequate *inferential relations* with other formulas.

A difficult issue arise here. In general, we can ask about an operator *O* of *L* and a connective *C* of *L* if

O(*formula*),

and,

(*formula1*) *C* (*formula2*),

have the *expected* properties. For *O* and *C* to have the expected properties is for them to be adequate representations of sentences in natural language, in that a particular set of formulas in which they occur are true/false (just as the original sentences) and a particular set of formulas contains all semantic/axiomatic consequences of them.

Now suppose our axiomatic system *S* is such that from *O*(*formula1*) it follows that *formula2*. But the natural language analogue of *formula2* is not supposed to be a (semantic) consequence of the natural language analogue of *O*(*formula1*). Or suppose that from (*formula1*) *C* (*formula2*) it follows that *O*(*formula2*) *C* *O*(*formula1*) and, again, this was not expected. As it happens, though, the semantics for system *S* is such that *S* is both complete and consistent, the system is otherwise adequate in capturing inferential relations we anticipated, it has the desired generality, etc. The virtues of *S* (and its semantics) create a problem for us: should we embrace the system and regard natural language as *wrong* in some respect, or should we use the disanalogy with natural language as a *criterion* for judging that the system is inadequate? Presumably, in such a case, there is a crucial difference between the semantics for the formal language and the semantics for the formalized language, and it is a philosophical issue if the former is adequate or not.

These are some types of questions in philosophy of logic related to both, the language and the semantics, of formal systems. Accordingly, we have a variety of languages and semantics involving different kinds of intensional operators, quantifiers and connectives, and it is clear that philosophical investigations give rise to new systems in this way.

4.

Consider now what one could take to be the most important question in philosophy of logic: the one about the *nature* of logic. We saw, roughly, how a logical system is constituted and surely there is more than one logical system. We can have different systems with exactly the same syntax but with a different semantics, or *extensions* of other systems, or even systems that are completely different of each other. What is common between all the systems such that we can say of each one that it is 'a logic'? Or, more crudely, what is logic?

It is not unusual for authors to explain what logic is by claiming that it is the systematic study of the correct principles of *reasoning*. The concept of *reasoning*, however, is quite ambiguous. There are at least two main meanings that can be attached to this concept. In the first one, reasoning is a type of *cognitive* process, also called 'inference'. Reasoning in this sense is a psychological or cognitive phenomenon whereby beliefs (i.e. *inferential beliefs*) are held by an agent *S because S* holds other beliefs (i.e. *pre-inferential beliefs*). In the second one, reasoning is identified with *argument*.

When it comes to the first reading, we are immediately led to ask what it is for a process of reasoning to be correct. Presumably, we are talking about reasoning being *epistemically* correct. Minimally, a type of reasoning is correct in this sense when it (conditionally) optimizes the epistemic goal of having true beliefs and avoiding false ones[3]. If it is not something like this that one has in mind when one talks about 'correct reasoning' in the first sense, then it is difficult to understand the claim that logic is the study of correct reasoning. Presumably, 'correct reasoning' would include *deductive* reasoning, *inductive* reasoning and *probabilistic* reasoning. If logic is the study of correct reasoning in the first sense we need a system of inductive logic and a system of probability logic taking into account both, epistemically correct ways of inductive reasoning and epistemically correct ways of probabilistic reasoning.

One might be skeptical, though, about the claim that logic is the study of correct reasoning, understood as a psychological process. There may be some kinds of reasoning that can plausibly be regarded as *epistemically rational* and, still, their dynamics do not correspond to a certain logical requirement. As an example, consider *preface* situations: a certain author rationally believes every single proposition expressed in the book she just wrote, but knowing that humans are fallible beings,

[3] When we say that a type of reasoning *conditionally* optimizes the epistemic goal we are saying that, in this type of reasoning, if the content of the pre-inferential belief is true then the content of the inferential belief is true (in the case of deductive reasoning) or probably true (in the case of non-deductive reasoning).

she also believes that her book is not a 100% error-free book. In this situation, a rational agent believes each of a certain set of propositions but she also rationally believes a proposition that is equivalent to the *negation* of the conjunction of all these propositions. More formally, the agent individually believes each proposition from the set:

$\{P_1, Q_1, R_1, ..., P_n, Q_n, R_n\}$,

but she also believes a proposition that is *equivalent* to the negation of the big conjunction:

$\sim(P_1 \ \& \ Q_1 \ \& \ R_1 \ \& ... \& \ P_n \ \& \ Q_n \ \& \ R_n)$.

Here we have a cognitive behavior that we can regard as rational despite the fact that it violates logical, deductive, constraints. Of course, maybe while it violates some deductive constraint (classical deductive consistency) it can be in accordance with a probabilistic constraint. But what is the criterion that we should use to decide the type of logical system (for example, the *Probability Calculus*) that must be applied to this or that epistemic situation? The proponent of logic *as a science of correct reasoning* (in the cognitive reading of *reasoning*) must answer to this question.

Consider again the (normative) interpretation of logic as a study of *correct* reasoning, where reasoning is understood as a cognitive process of inference. A certain theory can be said to be normative both, when it expresses what agents *must do* and when it expresses what agents *are allowed to do*. When someone says that logic is the science of correct reasoning in the sense that it establishes what rational agents *must* believe (given certain conditions), we can have very implausible consequences for epistemology. More to the point, consider what would happen if for every belief held by an agent we were requiring that the agent also believes the logical consequences of its content. For example, given that *P or Q* is a logical consequence of *P*, whenever an agent believes that *P* he/she would also be required to believe that *P or Q*. But this is absurd! Reiterating the requirement would result in a requirement for the agent to believe (*P or Q*) *or R* and then ((*P or Q*) *or R*) *or S*) and so on. So maybe the constraints that logic puts on reasoning are normative in the sense that they establish what one *is allowed* to believe - not what one *must* believe. But, still, is this correct? Imagine I believe a certain mathematical proposition, *P*, and that from *P* a very complicated theorem follows - one that not even the smartest mathematician (or computer) has derived from *P*. I have no clue as to if the theorem follows from *P* and I have no inferential abilities to do the derivation. Does the fact that the complicated theorem is a logical consequence of *P* guarantee that I am allowed to believe the theorem (when I believe that *P*)? It seems it does not, for if I were to form a belief in the complicated theorem, given my belief that *P*, I would do so in an epistemically irresponsible manner.

So maybe we should not hope logic to deliver a normative framework for reasoning. Here we take a completely different route and advance the idea that logic has nothing to do with cognitive processes of inference: its main object of study is constituted by *types of arguments* (abstract structures) and it evaluates which types of arguments are truth-conducive and which are not. Notice that types of arguments are not the same thing as types of reasoning (understood as a cognitive process), for their instantiations differ in nature. Maybe, then, it would be better to abandon the idea that logic is the study of correct reasoning altogether, in order to avoid psychological and epistemic interpretations. Here, logic is a formal science that studies properties of arguments (validity) and sets of propositions (consistency), and if one can derive any epistemic normativity from such study is a contingent and controversial matter. Or, as we pointed out at the beginning, we can take logic to be the formal study of the *consequence* relation between propositions (or sentence-types): it deals with "what follows from what". Either way, we have a characterization of logic in which it is not conceived as a science about inferential (cognitive) processes.

But we can also grant that logic *per se* is not a formal science about the cognitive process of reasoning - that it is rather a formal science about the validity of arguments -, and still try to derive epistemic rules of reasoning from logic. So this second interpretation about the nature of logic *need not* deny the intuition that may lurk behind the first one - the intuition that logic has something to do with reasoning, normatively speaking. In trying to conciliate the more abstract view about the nature of logic with the relevant intuition, though, one faces a big challenge: (i) one needs to show *how* is logic supposed to be used as a basis for epistemic norms; (ii) one needs to explain and justify the use of a certain particular system instead of any other while trying to ground rules about reasoning; (iii) finally, one also needs to show that the epistemic norms based on the relevant logical systems are 'feasible' for standard human cognition - if these norms are supposed to be followed by human reasoners, they better be capable of doing so!

Be it as it may, it is a job for philosophers of logic to say what the nature of logic is and, consequently, to deal with the problems that arise as a result of the preferred view about this issue.

***

Have a good read!

Luis Rosa,

New Brunswick, July 2013