Organon F 23 (2) 2016

Contents

ARTICLES

BOOK REVIEWS

Organon F 23 (2) 2016: 142-161

What Is (Modern) Logic Taken to Be About and What It Is About

JAROSLAV PEREGRIN

Department of Logic. Institute of Philosophy. Academy of Sciences of the Czech Republic
Jilská 1. 110 00 Prague 1
Department of Philosophy and Social Sciences. Faculty of Philosophy
University of Hradec Králové. Náměstí Svobody 331. 500 02 Hradec Králové. Czech Republic
jarda.peregrin.cz

RECEIVED: 23-11-2015 ▪ ACCEPTED: 17-02-2016

ABSTRACT: Since Antiquity, logic has always enjoyed a status of something crucially important, because it shows us how to reason, if we are to reason correctly. Yet the twentieth century fostered an unprecedented boost in logical studies and delivered a wealth of results, most of which are not only not understandable by non-specialists, but their very connection with the original agenda of logic is far from clear. In this paper, I survey how the achievements of modern logic are construed by non-specialists and subject their construals to critical scrutiny. I argue that logic cannot be taken as a theory of the limits of our world and that its *prima facie* most plausible construal as a theory of reasoning is too unclear to be taken at face value. I argue that the viable construal of logic takes it to be explicative of the constitutive (rather than strategic) rules of reasoning, not of the rules that tell us *how* to reason, but rather of rules that make up the tools *with which* (or *in terms of which*) we reason.

KEYWORDS: Logic – mathematical logic – philosophical logic – reasoning.

1. The word "logic"

Logic, in the traditional sense of the word, is taken to be something general and universal, and though there have been disputes over some logical principles (e.g. the principle of the excluded middle, which states that either *A* or not-*A* must be the case; see Church 1928), and sometimes it even seems that these disputes augur a split of logic into different kinds (such as, in the context of the twentieth century, classical and intuitionist logic; see Mancosu et al. 2009), it still seems a contradiction in terms to say that everybody can help themselves to their own logics or that logic can be changed according to the subject matter it is applied to. Our traditional notion remains that the bulk of logic must be general, universal and topic-neutral.

However, if we examine how the word "logic" is actually applied in practice, we soon see that its usage does not really conform to this notion. When we search the British National Corpus (http://www.natcorp.ox.ac.uk/) – a representative set of electronic texts ("of written and spoken language from a wide range of sources, designed to represent a wide cross-section of current British English, both spoken and written"), we discover that (leaving aside the idiosyncratically technical contexts, and those conforming to the above delimitation of logic) there are many examples of usages of the following kind:

> ***AK6 323*** *If sport carries on combining with showbiz at the present rate, this dingy logic will eventually be hard to resist.*
>
> ***CEP 3274*** *Apart from the effect on the playing side, Strudwick sees another benefit with his typically straightforward brand of Aussie logic.*
>
> ***CGF 963*** *And the logic they sought was the logic of sexual difference and male superiority.*
>
> ***CR9 160*** *If not, the logic of the threats made so far is that bombing must follow; threats may have been unwise, but it would be even less wise to make them and fail to carry them out.*
>
> ***CRY 647*** *The particular towns on the list thus had no obvious logic.*
>
> ***CTY 371*** *To characterize only recent French thought as 'the logic of disintegration', as Peter Dews has recently done, masks over the fact that such a logic is fundamental to Marxism itself, the unassimilable dark other to its 'primacy of the category of totality'.*
>
> ***ED6 435*** *Everything is worked out, every detail of orchestration and balance, and there always seems to be an inner logic to the composition – there isn't a bar that isn't absolutely essential.*

EDA 1188 *Why this should have been thought evidence of scurrility was known only to Joyce's peculiar logic.*

FA0 569 *That theory's notion of a world economy is, in its simplest form, based on the view that, since the inception of capitalism in Europe, every part of the globe is linked together through a world market and, thereafter, all that happens obeys the logic of that world market so as to generate profits for enterprises in the advanced capitalist countries.*

FP2 473 *What is the logic that dictates that the shareholders should be entitled to the corporate surplus, instead for instance of the employees or management, with the entitlement of the shareholders reduced to a fixed return on capital?*

G0D 120 *Environmentalism in its first phase had advanced its hegemony through a grandiose moral and scientific logic.*

G12 945 *It was strange; everything he had done on the programme had seemed at the time to be imbued with an exact sense of logic and purposiveness, but now that he looked back on it, all the logical connections had disappeared, like secret writing when the special lamp is taken away.*

G13 739 *I stole a look at Conchis as he gazed up at the picture; he had, by no other logic than that of cultural snobbery, gained a whole new dimension of respectability for me, and I began to feel much less sure of his eccentricity and his phoniness, of my own superiority in the matter of what life was really about.*

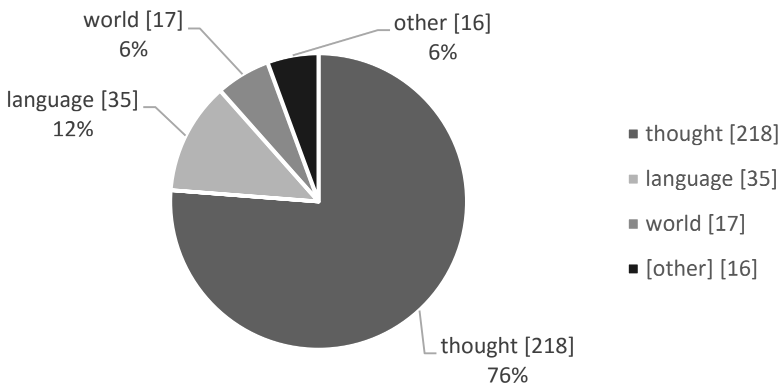
Such quotations suggest that the term "logic" is actually used in multifarious ways. It would seem that one of the most frequent uses of the term is as denoting "the bulk of organizing principles behind an institution, an activity or a domain", or "a set of reasons for a standpoint", or perhaps "a way of conceiving of a matter". This usage would entail that different things may indeed have different logics, and in fact it sometimes seems that it is precisely what is called a "logic" that captures the peculiar "essence" of a thing.

Of course, if we concentrate on logic as a doctrine, as that which is taught in schools and universities, the variability of understandings narrows down significantly. But does this narrowing down lead to something consistent? Is there a general agreement on what logic is and how it should help us? I am afraid not.

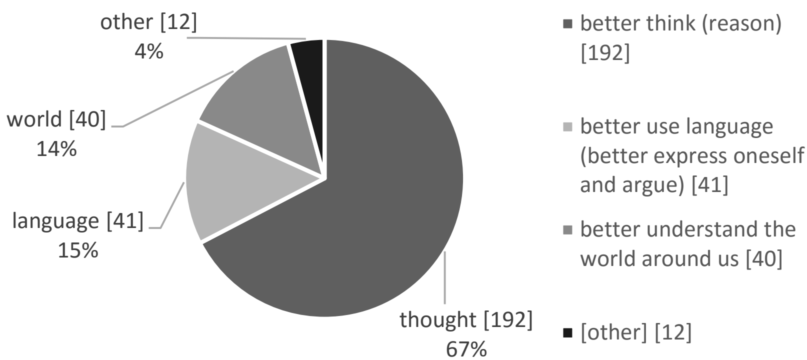
2. Logic as a tool

The following survey was undertaken at the Philosophy Faculty of the University of Hradec Králové (Czech Republic). 286 students took part; of which 114 had taken a course in elementary logic; none of them had studied logic at an advanced level. The participants were asked to express their personal opinions, uncontaminated by what they might think would be "correct" in any other sense.

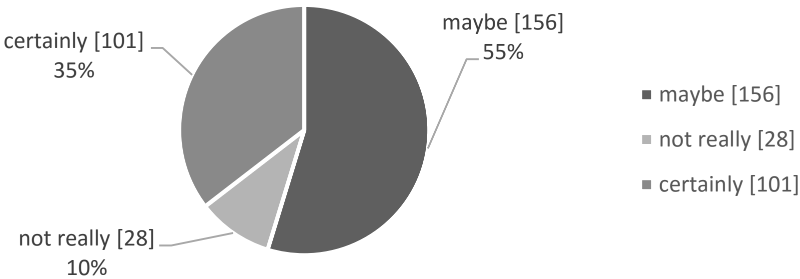
Logic is primarily interested in:



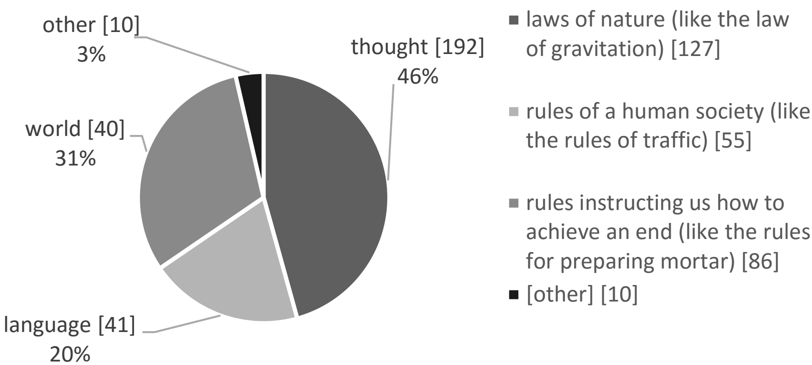
Logic should primarily help us:



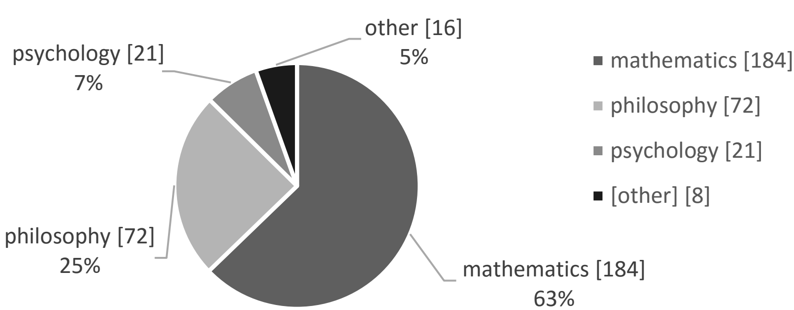
Do you think that logic really helps us in this way?



The laws of logic are, according to their nature, most similar to:



Which of the disciplines below is closest to logic?



In some respects, these results are unsurprising. The first two questions document that the prevailing notion of logic is very traditional: according to this, logic is some kind of a theory of correct thinking, or correct reasoning. But the results for the fourth question are more interesting: the majority of students are convinced that the laws of logic are best compared with the laws of nature, and hence perhaps they are those laws of nature which govern how we think; or they are those which get reflected by our minds in some peculiar way. Now the results of the fifth question indicate that the great majority of students assume that logic falls into the province of mathematics. Taken together, this seems to indicate that it is mathematics that is competent to study those natural laws that govern human thought – and this is a result which would call for some elucidation (for though to elucidate reasoning, as well as almost any other phenomenon, we may conceivably make use of mathematics, it is not quite clear how mathematics alone could be its theory).

In general, it seems that the most popular understanding of logic is its construal as the science (art?) of correct reasoning. This much seems to be clear and this, it would seem, is what was directly imprinted into the answers to the first two questions. What is no longer so clear is what the sense of "correct" is in the "correct reasoning". Does it amount to simply "successful" or "effective"? The majority of respondents appear to deny this – they probably feel that there must be some more substantial notion of correctness, something more akin to the correctness expressed by natural laws.

Now it would seem that if logic is to teach us to reason correctly, then we can also reason incorrectly – indeed logical training should make us abandon any possible habits of incorrect reasoning in favor of reasoning following the canons of logic. Hence, in what sense could it be that the laws of logic, that spell out the ways how we *should* reason *correctly*, are akin to natural laws, which specify how things happen *inevitably*?

Moreover, what is the nature of the close connection between logic and mathematics? True, many kinds of science use mathematical tools to build various models of their domains, compute their parameters etc.; but the relationship between logic (especially modern logic) and mathematics appears to have become more intimate. But if the task of logic is to instruct us which mental processes open to us warrant our engagement, why is it more intimately connected with mathematics than with psychology, the discipline devoted to studying mental processes?

3. Logic and mathematics

In the late nineteenth and early twentieth centuries, Aristotelian logic, which had been the paradigm of logic for more than two millennia, was superseded by a wholly new paradigm: a paradigm that was created by thinkers such as George Boole, Gottlob Frege, Giuseppe Peano and Bertrand Russell (see Grattan-Guiness 2000; or Haaparanta 2009), and which soon led logic into the embrace of mathematics. The idea was that just as physics was able to move to a wholly new level once it was learned how to formulate its problems so that they were amenable to mathematical treatments, so logic might expect a similar acceleration by opening itself to mathematics. However, although the mathematical logic of the twentieth century undoubtedly supplied logicians with a huge number of problems of kinds never dreamed of by previous generations of logicians, the relevance of these problems for logic, in the original sense of the word, is debatable.

Take, by way of comparison, modern physics. Here, too, we can encounter lots of highly complicated mathematics. But here the mathematics is never an end in itself – the results of any computations must be "translated" into the language of physics and tell us something about the physical world. If the role of mathematics in logic is to be similar, then the results reached in the books on mathematical logic also require "translating" so that they tell us something about the correctness of our reasoning.[1]

There is an explanation for this lack of the final "translation" of the mathematical results of logic into a theory of human reasoning: the investigations have changed their nature and become pure mathematics. This may happen in any field of theory – trying to build mathematical models leads us to new mathematics, which, apart from throwing light on the original domain, may become interesting in its own right, and some theoreticians begin pursuing it not for the sake of modeling the original problem, but simply for the sake of

---

[1] To avoid misunderstanding, of course that *any* mathematical theory can be seen as telling us how to use the concepts it is based on – and especially how to reason with them. It must be based on *definitions* of the concepts and the definitions can be seen as instructions for use. But this is the usage of specific concepts on which the theory is based – not of reasoning in general, which is supposed to be the subject matter of logic. If the point of the new concepts is to help us reason better, or at least to better understand how we reason, then if the result is just to master the new concepts, it is radically unsatisfactory.

investigating the new mathematical structures. (And, indeed, a lot of the mathematics inspired by physics has become the ultimate subject matter of purely mathematical studies; but who would consider this as *replacing* physics?)

There is no doubt that much of what is now being done under the heading of "logic", and especially "mathematical logic" (cf. Barwise 1977; Crossley 2011), is a branch of mathematics, with no intention of analyzing, elucidating or improving how we actually reason. And doing mathematics is, of course, a respectable business. The only problem is that there is still a need for theories holding to the traditional agenda of logic.

Not all of contemporary logic declares itself as mathematical. Some protagonists of logical investigations are keen to distance themselves from wholly embedding logic into mathematics; they want to do "non-mathematical" logic (which does not, of course, preclude them from using some mathematics as a tool!). The term "philosophical logic" is occasionally used to distinguish their kind of logic from the mathematical one. The problem, however, is that the term "philosophical logic" is so ambiguous that it is not very useful.

In some cases, the adjective "philosophical" is to be taken at face value, in that the term "philosophical logic" is employed to refer to "philosophy of logic" (see Haack 1978; Grayling 1998) or "philosophy done with the help of logic" (cf. Goble 2001; Jacquette 2002). In other contexts, it has acquired a rather technical sense, in which "philosophical logic" refers to investigations (including purely mathematical) of systems of the so-called "non-classical logic" (cf. Burgess 2009). And it is only in the remaining cases that it is used simply to mark the logic that is referred to as not a pure mathematics, but rather working towards the traditional aims of logic.

The fact that logicians do not seem to be univocally interested in the agenda of correct reasoning and argumentation has also invoked a large movement of those who want to take up this task straightforwardly at face value. The so-called "theories of argumentation" (van Eemeren & Grootendorst 2003; Walton et al. 2008), "informal logics" (Walton 1989; Copi & Burgess-Jackson 1996) or "critical thinking" (Paul & Elder 2002; Bowell & Kemp 2002) are usually very practical enterprises on the boundary between logic and a kind of "technology" of reasoning and argumentation, which does not distinguish between the domain that has been traditionally assigned to logic, and other domains traditionally covered by rhetoric etc.

4. Logical laws

As asking about the nature of logic is mostly asking about the nature of logical laws, it is good to clarify what exactly is meant by these. One of the most traditional claims that has come to be called *logical law* is the principle of non-contradiction, stating that no *a* can be both *P* and not-*P*, or, more generally, that nothing can both hold and not hold at the same time. Using the symbolism of modern logic, we can record it as

(NC) $\neg(A \wedge \neg A)$.

The complementary law stating that anything must either hold or not hold, the so-called law of excluded middle (which is accepted by far not as univocally as (NC)), is then

(EM) $A \vee \neg A$.[2]

Aside of laws of this kind, there are laws which do not have a form of a statement, but rather of an inference, a transition from statements (premises) to statements (conclusion). A typical example is *modus ponens*, the inference from *if A, then B* and *A* to *B*, symbolically

(MP) $$\frac{A \qquad A \rightarrow B}{B}$$

We can consider other laws of this kind, such as

(∧I) $$\frac{A \qquad B}{A \wedge B}$$

[2] It is a peculiar fact that within classical propositional logic, the two laws formulated in this way turn out to be equivalent. (Indeed, as $\neg(A \wedge B) \leftrightarrow (\neg A \vee \neg B)$, $\neg(A \wedge \neg A)$ is equivalent with $\neg A \vee \neg\neg A$, and as $\neg\neg A \leftrightarrow A$, this is further equivalent with $\neg A \vee A$.) But this should not be construed as the proof that the two laws say the same, but rather as a demonstration of the restricted expressive power of this logical system.

or

$$(\vee I) \quad \frac{A}{A \vee B}$$

All these laws contain *logical constants*: ¬, ∧, →, … (or, perhaps, in their informal articulation, "logical" words of natural language, such as *not*, *and*, *if-then*, …). So the elucidation of the nature of logical laws clearly involves an elucidation of the nature of these constants.

There are several ways to view logical constants. One possibility is to consider them as primarily elements of the "furniture of the universe", as potential constituents of facts making up our world. (Our signs then being their – better or worse – representations; the natural language words being not very faithful, whereas the signs of our formal languages being much better.) The statement of the form "$A \rightarrow B$", for example, can be seen as expressing a worldly fact, perhaps concerning some kind of (causal?) dependence of B on A. If we view logical constants in this way, the laws of logic will be something akin to natural laws.

Another possibility is to consider logical constants as primarily constituents of thoughts. (Then again, our signs are their representations.) Given this, logical laws are some kind of laws of thought (perhaps directives how to think so as to arrive at the truth?) On this construal, "$A \rightarrow B$" expresses a specific kind of thought.

Then there is the possibility to locate logical constants in our language, the "logical words" of natural languages being their tentative versions, while those of the artificial languages the versions that has been regimented and stabilized (see Peregrin & Svoboda forthcoming). Given this construal, logical constants are not entities *represented* by these signs; rather they are these very signs (perhaps the regimented kind). However, the fact that logical constants do not represent anything should not be read as saying that they are meaningless – they have meanings in that they have certain uses within our language games, especially they have certain roles with respect to the inferential rules governing the sentences containing the constants. Construed thus, "$A \rightarrow B$" is an element of an artificial language, characterized by its inferential role (e.g. that we can infer from $A \rightarrow B$ and A to B).

5. Logic and the world

Let us now turn our attention to the viability of various notions of logic based on the various understanding of logical constants sketched above. We will try to indicate that the *prima facie* most plausible construals of the nature of logic are in fact *not* viable and that the question of what exactly modern logic is useful for is not quite easy to answer.

Let us start from the relationship between the laws of logic and the laws of nature. Could the former be akin to the latter, as the survey showed they are often taken to be? (Adherents of this view can appeal to the patronage of no lesser person than a key founding father of modern logic, Bertrand Russell, who in Russell (1919, 169-170) famously claimed that logic "is concerned with the real world just as truly as zoology, though with its more abstract and general features".) Natural laws deal with a certain kind of necessity or impossibility (which is clearly the other side of necessity): they tell us, we can say, that some things that might appear to be possible are in fact impossible. So let us consider the concept of impossibility.

It is, for example, impossible for me to speak Portuguese right now, because I have not learned the language. This is an impossibility stemming from the contingencies of my life; should it have run slightly otherwise, it might have been that I would have learned Portuguese. Then there are impossibilities that appear to be somewhat more categorical. I cannot, for example, fly like a bird. This has nothing to do with the course of my life, it is a necessity which we can call *physiological*. My organism is simply not made to make this possible. But then there is an even more categorical kind of impossibility: I cannot move faster than light. This has nothing to do with my particular physiology, it is a rather of the physical law discovered by Einstein. We can call it *physical* necessity.

Now logic might perhaps be seen as dealing with an even more categorical kind of necessity: *logical* necessity. I cannot, for example, run and at the same time not run. This might be said to be excluded by (NC). However, what is it that this law excludes? The previous kinds of necessity tell me that I cannot do something that I could have imagined I would be able to do. In principle, I could imagine experiments in which I tried to check whether I really cannot do what I am told I cannot. But what would it be to run and not to run? How could I try to do so to check whether it is really impossible?

Maybe, then, the necessity logic spells out is a matter of conception or imagination. I cannot imagine myself running and at the same time not running.

But here we must be careful. I clearly can imagine myself running and also I can imagine myself not running, and perhaps I can do this at the same time (e.g. superimposing two pictures before the mind's eye). Hence, what I cannot do is not *imagine that I-am-running* and, and at the same time, *imagine that I-am-not-running*, but rather *imagine that I-am-running-and-at-the-same-time-not-running*. But again, what is it that I cannot imagine? What is it *to-be-running-and-not-running*?

I can say that I cannot imagine myself *himbajsing*. Indeed, I cannot, for I simply do not know what it is. But surely this is just a dull fact, pointing out no limitations of my imagination. And the fact that I cannot imagine myself running-and-not-running does not seem to be significantly different in this respect: again, I cannot imagine something simply because I do not know what it is (the difference being only that while *himbajsing* is utterly nonsensical, *running-and-not-running* is composed of meaningful parts, though in a way which makes the whole meaningless).[3]

Hence, perhaps it is not *imagination*, but rather *belief* the boundaries of which are drawn out by logic? Can I believe that I am running-and-not-running? Again, it seems that I *can* believe that I am running and at the same time believe that I am not running – this would, to be sure, not be a very frequent situation, but people are known to be able to harbor contradictory beliefs, so though improbable, it may seem not utterly impossible. Now, what about the belief that I am running-and-not-running? Again the same situation as before: I cannot believe something that does not make sense.

The upshot seems to be that, unlike the laws of nature, laws of logic cannot be straightforwardly and transparently derived from the world. To say that they are akin to the laws of nature brings about many more questions than it can answer.

---

[3] This was pointed out already by Wittgenstein (1922), who took logical necessities and impossibilities as pathologic by-products of logical vocabulary, which otherwise helps enhance the representational capacities of our language. See also Coffa (1991, chap. 8). In his later writings, Wittgenstein was even more explicit about this. Note, for example, what Wittgenstein (1956, I.§132) has to say about the law of identity: "Frege calls it 'a law about what men take for true' that 'It is impossible for human beings … to recognize an object as different from itself'. – When I think of this as impossible for me, then I think of *trying* to do it. So I look at my lamp and say: 'This lamp is different from itself'. (But nothing stirs.) It is not that I see it is false, I can't do anything with it at all."

6. Logic and acquiring true beliefs

Hence the notion that the laws of logic are kind of natural laws (perhaps the most general one?), that the laws tell us what is impossible despite being seemingly possible (be it in the world or in our thought) does not seem to stand up to scrutiny. We must, it seems, discard the notion that the laws of logic are of the kind of natural laws, despite its popularity. And as the view that logical laws are akin to social norms is neither appealing, nor popular, the outcome appears to be that we should accept that the laws of logic are best viewed as *instrumental* rules, telling us how we should reason in an efficient way.

This may seem to be a welcome happy-end: we have already seen that this view of logic, *viz*. the view of logic as a theory of correct reasoning, is almost generally accepted and insofar as the view of laws of logic as natural laws is not compatible with it, then the popularity of this latter view must be a matter of some delusion, and we should be happy to relinquish it. However, to see this as a happy-end would be premature. The notion of logic as a theory of reasoning, as the pursuit of reasoning correct in an instrumental sense, raises some very awkward questions.

First, if the kind of correctness involved here is to be instrumental correctness, then it must be derived from a goal at which the whole enterprise of reasoning is aiming – indeed the instrumental "correct" is nothing other than "effectively helping us achieve a goal". Hence, according to this view, logic tells us how to reason to reach a goal. But what is the goal?

Answering this question may not be as easy as it might look at first sight. What comes to mind immediately is something like acquiring true beliefs. But does this mean that logic is to help us acquire *only* beliefs that are true, or rather that it helps us acquire *as many* true beliefs *as possible*? Neither response holds water. If the task of logic were to prevent us from acquiring beliefs that are not true, then it could accomplish it simply by instructing us to acquire *no* beliefs at all; whereas if the task were acquiring as many true beliefs as possible, then it could well instruct us to acquire an arbitrary trivial true belief, conjoin it with itself and continue further conjunctions with the original belief *ad infinitum*.

In reality it would seem that what we need is some *reasonable* collection of *relevant* true beliefs. We certainly do not need all true beliefs, and we probably should not despair about acquiring some exceptional non-true beliefs. What is the precise sense of the "reasonable" and "relevant" in the "*reasonable* collection of *relevant* true beliefs"? Well, it seems to be highly context-dependent – what is useful to know may differ quite radically from one context to another. Anyway, to find out what is thus useful is a vital part of our art of steering clear of the perils of our world; and logic does not seem to offer us any clear instructions how to do this.

All of this is not to say that construing logic as a theory of correct reasoning is utterly misguided, but it does indicate that if we propose this construal, it is very blurry what exactly it is we are proposing. It is obvious that in some cases logic may help us acquire new and useful beliefs; but much of the reasoning we perform is either stimulated or influenced by factors which are beyond the boundaries of logic; and, conversely, a lot of the reasoning logic sanctions as correct is of no use for us.

It is of no help to claim that logic equips us with a useful *tool*, and the question of what we do with the tool, whether we are able to achieve something valuable with its help, is not its business. The point is that we cannot say that a tool is useful until we know what useful end it can serve as a means. Hence again, the question is *a tool for what*? And if the answer is as simple as *for acquiring new true beliefs*, then it does not seem quite satisfactory.

7. Logic as sanctioning selected inferences

Given this, we may try to reduce the role of logic to an acceptable minimum. Perhaps logic does not tell us how to acquire the reasonable collection of relevant true beliefs we need to cope with the world successfully; perhaps it only helps us with some partial aspect of this process. Perhaps the only thing that logic is able to do is to tell us that certain ways of going from beliefs already had to a belief to be acquired are impeccable in the sense that if the former beliefs are true, then the latter is true too.

In comparison to the role of logic considered in the previous section, this role is truly minimalistic. Logic tells us nothing about which beliefs to acquire or how to acquire them; it only tells us that if we choose certain ways of acquiring them, we will not fail. (But perhaps there are ways of acquiring beliefs that are better than those sanctioned by logic, perhaps ways that are not quite impeccable, but much more effective?)

It may seem that though assigning this task to logic may look, to many, as its denigration, at least this is what logic really does. But I am afraid that even this is not as straightforward as it might *prima facie* seem. For consider the inferences we usually find in logical textbooks. Take, for example, such laws as (∧I) or (∨I) above, i.e. the rule that $A \wedge B$ is inferable form A and B; or that $A \vee B$ is inferable form A. What is the use of such inferences? Do we ever, when we reason, consider them so that we could use the fact that logic tell us that we can use them safely?

The answer is not quite clear. At least *prima facie*, if we have the beliefs A and B, we do not need the extra belief that $A \wedge B$. If I have the belief that *it rains* and that *it is dark*, why would I need the extra belief that *It rains and it is dark*? I know that when I get out it may be useful to take an umbrella and a flashlight already on the basis of the former beliefs, I do not need the latter. More generally, whatever I can infer from the latter one I can infer already from the former.

The situation is even more problematic in case of the inference from A to $A \vee B$. Here I not only do not gain anything, but I do loose something. It is hard to imagine that it would ever be useful to make such a step in reasoning. Well, of course we cannot see through all possible twists and turns our reasoning may take and we cannot exclude that even such inference, in combination with other ones, might be useful; however, the picture that logical laws are useful and impeccable ways of acquiring new true beliefs, in general, does not seem to stand to scrutiny.

8. So what is logic about?

The brief discussions presented in the previous sections are not to be taken as a substitute for a thorough analysis of the relationship between logic and reasoning.[4] My point was to indicate that the *prima facie* plausible view of logic as a theory of reasoning, on closer inspection, is by far not as satisfactory as it might seem. We surveyed some popular approaches and indicated that some ways of understanding the nature and agenda of logic lead us up alleys that, if not quite blind, are nevertheless tortuous and shady. It is not clear

---

4 See, e.g., Harman (1986); Perkins (2002); Milne (2009); Stenning & Van Lambalgen (2008); and see Peregrin (2014, especially chap. 10 and 11) for my fuller discussion of this topic.

whether the *prima facie* most plausible ways are ultimately viable. Now the question that may strike us is whether we have a viable way left at all.

To indicate that we do, let us consider the third of the possibilities of construing the nature of logical laws. Aside of taking them as akin to laws of nature and as seeing them as instrumental rules of efficient reasoning, we can look at them as at a kind of social rules. This, to be sure, seems *prima facie* quite absurd: is the law of contradiction akin to the rules telling me whom to greet, or those telling that I should help the poor? Despite this, I believe that even this option is worth being investigated.

The point is that some of our social rules are constitutive of useful social tools and institutions. Certain social rules, for example, constitute the institution of police, which is, needless to say, immensely helpful when it comes to human interaction and its pathological aspects. Other rules constitute, for example, loans, which may be an immensely useful tool for everybody short of money. Could it not be that the laws of logic constitute something useful in this way? And I think that we might consider a positive answer to this question: namely that logic constructs and provides us with certain "cognitive tools", which open up for us new modes of thinking. We usually do not register this, for as a species we have become too accustomed to these new modes.

What does the rule that we can reason from A to $A \vee B$ actually tell us? How to effectively manage our systems of beliefs in that we should extend it by $A \vee B$ whenever it contains A? We have already noted that except perhaps in some extenuating circumstances, we never reason to a disjunction from its disjunct. (Why adopt a belief that is a mere dilution of what we already know?) So what is the use of this rule? The answer I propose as worth considering is that the rule is *not* an instruction how to reason or how to manage our system of beliefs, but rather a rule constitutive of a tool *by means of which* we reason – in this case, disjunction. It allows us to say things and think thoughts of a usefully unspecific nature. ("Either we mend this, or we are doomed.")

The most instructive case is that of implication, a tool which represents an entering wedge into the hypothetical ways of argumentation and reasoning, and which, needless to say, offers a formidable upgrade to our thinking powers. Conditional statements, formed with the help of implication, allow us to say not only what is the case, but what is the case *if something else is the case*, which paves the way to counterfactuals (although the classical, material implication is not strictly speaking a counterfactual, there is a close connection to genuine counterfactuals).

In this case, we must keep in mind that on this construal, logical constants, and consequently logical laws, are not any entities beyond the words of natural language or the sign of an artificial ones – they are directly the linguistic items (understood as governed by certain laws). And it is vital to see that this conception of logic differs from the conception discussed in previous sections more fundamentally than it might *prima facie* seem. This point can be illustrated by comparing logic with chess. With respect to chess, we can consider two kinds of rules: the "constitutive" ones that delimit the moves that are legal, and the "strategic" ones that indicate moves that are good in the sense of being likely to lead to victory. The latter, strategic rules presuppose that we already have the chess pieces (that are, as such, constituted by the constitutive rules) and we are consequently in a position to put them to an effective use. Now, the notion of logic as the science of reasoning discussed in previous sections, took the laws of logic as the strategic kind of rules, hence it presupposed that we already had the logical concepts and it was telling us how to make an efficient use of them. By contrast, the notion considered at in this section takes the laws of logic as constitutive, rather than tactical rules: they produce the basic logical tools that open up new spaces for our argumentation and reasoning, without "strategic" advice about how to steer through them.

But are ways of thinking "social tools" or "institutions"? Are they not a matter of individual psychologies? Part and parcel of this way of looking at logic is the conviction that they are not; that they are imprints of social practices – that the kinds of tools exemplified by that logical constants are forged in social mold. This might seem strange, but the view that human covert reasoning is based on inter-human overt argumentation (rather than vice versa) can be backed by arguments of both philosophers and social scientists.[5]

I think that if we accept this, we could embrace the notion of logic as a theory of reasoning only on a very specific reading. According to this view, logic does not tell us how to reason in the sense of instructing us how to weave our webs of our beliefs effectively. It makes explicit the constitution of the

---

5 As for the philosophical arguments, I think that the clearest ones were formulated by Davidson (1991), who argues that that the very fact of propositional thought (including reasoning) presupposes communication. See also Dutilh Novaes (2015); and for a more empirically grounded account see, e.g., Mercier & Sperber (2011). For a more detailed exposition see Peregrin (2014, chap. 11).

most powerful, and most general, cognitive tools we have – disjunction, implication, negation etc. – and in this way it makes for the very possibility of having beliefs and weaving their web. In one sense this is not very much – we need a lot of additional rules or experiences to put these tools to effective use. In another sense, it is quite a lot – without these basic tools, there would be no space within which such instructions would help us steer (see Peregrin 2014, chap. 10 for a further elaboration).

9. Conclusion

Modern logic is in a paradoxical situation. On the one hand, it has inherited at least part of the prestige logic has always enjoyed, and, moreover, it has introduced logical investigations into unprecedented mathematical intricacies; on the other hand, it has become unclear what exactly it thereby brings us and how exactly it is useful to us. This situation can only be overcome by in-depth analyses of the nature of logic and its achievements. We can no longer make do with received wisdoms such as "logic tells us how to reason", not because these are completely false, but because they are misleading and often engender a mere illusion of explanation. We must clarify which results of modern logic are to be seen as part of mathematics and which bring us something more – and what exactly this "something more" is.

In this paper, I have proposed that the most adequate way of viewing logical laws, and consequently logic, is to view them not as strategic rules telling us how to manage our beliefs effectively, but rather as certain constitutive rules, rules that open up, for us, certain modes of reasoning by constituting the "cognitive tools" on which these modes rest. I have argued that to think otherwise is akin to mistaking the rules constitutive of chess, rules telling us which moves are allowed for which piece, for strategic rules advising us how to play so as to win.

References

BARWISE, J. (ed.) (1977): *Handbook of Mathematical Logic*. Amsterdam: North-Holland.

BOWELL, T. & KEMP, G. (2002): *Critical Thinking: A Concise Guide*. London: Routledge.

BURGESS, J. P. (2009): *Philosophical Logic*. Princeton: Princeton University Press.
COFFA, A. (1991): *The Semantic Tradition from Kant to Carnap*. Cambridge: Cambridge University Press.
COPI, I. M. & BURGESS-JACKSON, K. (1996): *Informal logic*. Upper Saddle River: Prentice Hall.
CROSSLEY, J. N. (2011): What Is Mathematical Logic? A Survey. In: van Benthem, J., Gupta, A. & Parikh, R. (eds.): *Proof, Computation and Agency*. London: Springer, 3-17.
DAVIDSON, D. (1991): Three Varieties of Knowledge. In: Griffiths, A. (ed.): *A. J. Ayer: Memorial Essays*. (Royal Institute of Philosophy Supplement 30), Cambridge: Cambridge University Press, 153-166; reprinted in: Davidson, D. (2001): *Subjective, Intersubjective, Objective*. New York: Oxford University Press, 205-220.
DUTILH NOVAES, C. (2015): A Dialogical, Multi-Agent Account of the Normativity of Logic. *Dialectica* 69, No. 4, 587-609.
GOBLE, L. (2001): *The Blackwell Guide to Philosophical Logic*. Oxford: Blackwell.
GRATTAN-GUINESS, I. (2000): *The Search for Mathematical Roots 1870-1940*. Princeton: Princeton University Press.
GRAYLING, A. (1998): *An Introduction to Philosophical Logic*. Oxford: Blackwell.
HAACK, S. (1978): *Philosophy of Logics*. Cambridge: Cambridge University Press.
HAAPARANTA, L. (ed.) (2009): *The Development of Modern Logic*. Oxford: Oxford University Press.
HARMAN, G. (1986): *Change in View (Principles of Reasoning)*. Cambridge (Mass.): MIT Press.
CHURCH, A. (1928): On the Law of Excluded Middle. *Bulletin of the American Mathematical Society* 34, No. 1, 75-78.
JACQUETTE, D. (ed.) (2002): *A Companion of Philosophical Logic*. Oxford: Blackwell.
MANCOSU, P., ZACH, R. & BADESA, C. (2009): The Development of Mathematical Logic from Russell to Tarski: 1900-1935. In: Haaparanta, L. (ed.): *The History of Modern Logic*. New York: Oxford University Press, 318-470.
MERCIER, H. & SPERBER, D. (2011): Why Do Humans Reason? Arguments for an Argumentative Theory. *Behavioral and Brain Sciences* 34, No. 2, 57-111.
MILNE, P. (2009): What Is the Normative Role of Logic? *Proceedings of the Aristotelian Society*, Suppl. Vol. 83, 269-298.
PAUL, R. & ELDER, L. (2002): *Critical Thinking*. Upper Saddle River: FT Press.
PEREGRIN, J. (2014): *Inferentialism: Why Rules Matter*. Basingstoke: Palgrave.
PEREGRIN, J. & SVOBODA, V. (forthcoming): Logical Formalization and the Formation of Logic(s). *Logique et Analyse*.
PERKINS, D. N. (2002): Standard Logic as a Model of Reasoning: The Empirical Critique. In: Gabbay, D. M. et. al. (eds.): *Handbook of the Logic of Argument and Inference*. Amsterdam: Elsevier, 186-223.
RUSSELL, B. (1919): *Introduction to Mathematical Philosophy*. London: Allen & Unwin.

STENNING, K. & VAN LAMBALGEN, M. (2008): *Human Reasoning and Cognitive Science*. Cambridge (Mass.): MIT Press.

VAN EEMEREN, F. H. & GROOTENDORST, R. (2003): *A Systematic Theory of Argumentation: The Pragma-Dialectical Approach*. Cambridge: Cambridge University Press.

WALTON, D. N. (1989): *Informal Logic: A Handbook for Critical Argumentation*. Cambridge: Cambridge University Press.

WALTON, D., REED, C. & MACAGNO, F. (2008): *Argumentation Schemes*. Cambridge: Cambridge University Press.

WITTGENSTEIN, L. (1922): *Tractatus Logico-Philosophicus*. London: Routledge. English translation: London: Routledge, 1961.

WITTGENSTEIN, L. (1956): *Bemerkungen über die Grundlagen der Mathematik*. Oxford: Blackwell. English translation: *Remarks on the Foundations of Mathematics*. Oxford: Blackwell, 1956.

Organon F 23 (2) 2016: 162-176

Between the Actual and the Trivial World

MACIEJ SENDŁAK

Institute of Philosophy. University of Szczecin
Ul. Krakowska 71-79. 71-017 Szczecin. Poland
maciej.sendlak@gmail.com

RECEIVED: 11-08-2015 ▪ ACCEPTED: 25-02-2016

ABSTRACT: The subject of this paper is the notion of similarity between the actual and impossible worlds. Many believe that this notion is governed by two rules. According to the first rule, every non-trivial world is more similar to the actual world than the trivial world is. The second rule states that every possible world is more similar to the actual world than any impossible world is. The aim of this paper is to challenge both of these rules. We argue that acceptance of the first rule leads to the claim that the rule *ex contradictione sequitur quodlibet* is invalid in classical logic. The second rule does not recognize the fact that objects might be similar to one another due to various features.

KEYWORDS: Counterfactuals – counterpossibles – impossible worlds – possible worlds – trivial world.

1. Introduction

It is significant that we make some inferences which are based on what is impossible. Consider the following examples:

(1) If Hobbes had squared the circle, then mathematicians would be impressed.

(2) If Hobbes had squared the circle, then mathematicians would not be impressed.

(3) If it were the case that $2+2=5$, then it would not be the case that $2+3=5$.

(4) If it were the case that $2+2=5$, then it would be the case that $2+3=5$.

Common intuition and practice show that we tend to take (1) and (3) to be true and (2) and (4) to be false. Since all of these claims are in the form of conditionals, it is reasonable to expect that their truth or fallacy can be explained in terms of theories of counterfactuals. Unfortunately, according to the well-known analysis of worlds semantics, all of them are taken to be vacuously true.

Because of this, many contemporary philosophers of modality have been arguing that a standard analysis of counterfactuals in the framework of possible worlds semantics is insufficient when it comes to counterpossibles, i.e., counterfactuals with impossible antecedents.[1] As an alternative to the traditional approach, they have proposed an extended account that is based on worlds semantics which commits to possible as well as impossible worlds. One of the main aims of this extension was to satisfy the need for an explanation of reasoning about what is taken to be impossible (see Yagisawa 1988; Mares 1997; Nolan 1997; Restall 1997; Vander Laan 1997; 2004). Introducing impossible worlds raises many philosophical questions, and even though one can find various analyses of the logical structure and ontological status of impossible worlds and their application, few of these analyses discuss the important notion of similarity between worlds.[2] The importance of this notion lies in its role, which is to determine whether a given counterfactual (with a possible or impossible antecedent) is true or false.

Although "the discussion developed so far should show that the issue of the structure, closeness and ordering of impossible worlds is quite open" (Berto 2013), there are two claims which are in some sense the core of the standard understanding of the notion of similarity. The first one is commonly shared among the advocates of impossible worlds; the second one raises some doubts. According to the first claim, the trivial world, i.e., the world where everything

---

1 As 'standard analysis' we mean theories delivered by Robert Stalnaker (see Stalnaker 1968) and David Lewis (see Lewis 1973).

2 For a comprehensive analysis of the ontological status of impossible worlds, see Berto (2013), Nolan (2013).

is the case, is the most dissimilar to the actual world (@). In other words, every non-trivial world (possible or impossible) is closer (more similar) to the actual world than the trivial world is. We will call this claim the Dissimilarity of the Trivial World (*DTW*). The second assumption about similarity and impossible worlds is the Strangeness of Impossible Condition (*SIC*), according to which every possible world is closer to the actual world than any impossible world is. Both of these claims were formulated by Daniel Nolan (see Nolan 1997). Though *prima facie* both are compelling, we will show the reasons for believing that a proper analysis of counterfactuals requires that they be rejected.

The result of our investigation should be as general as possible, and because of this we will not discuss any particular account of the metaphysics of an impossible world. The reason for this is that the notion of similarity, which is our main concern, is taken to be a part of semantics, and not of metaphysics of impossible worlds. Moreover, *DTW* and *SIC* have their advocates among philosophers who take impossible worlds to be concrete, spatiotemporal objects (cf. Yagisawa 1988), as well as among those who believe that impossible worlds are abstract entities (cf. Nolan 1997). As such, our investigations are in an important respect independent of what the metaphysical nature of impossible worlds is. Nevertheless, we will base our analysis on two heuristic assumptions. According to the first one, the actual world is ruled by classical logic. The second assumption is that postulating impossible worlds should not lead to changes in the logic of the actual world.[3] These assumptions will help us point to the main concern about *DTW*. Even though the acceptance of *DTW* has particular consequences for advocates of the two above-mentioned assumptions, we shall see that philosophers who believe that the actual world is ruled by one of the non-classical logics are in no better position.

2. Counterpossibles

Counterpossibles can be represented as sentences of the form: "If it were the case that *A*, then it would be the case that *C*" ($A > C$), in which it is stated that the truth of an impossible antecedent (*A*) leads to a given consequent (*C*). Examples were already provided at the very beginning of the text:

---

[3] This view is shared by Daniel Nolan (see Nolan 1997) and David Vander Laan (see Vander Laan 1997), among others.

(1) If Hobbes had squared the circle, then mathematicians would be impressed.
(2) If Hobbes had squared the circle, then mathematicians would not be impressed.
(3) If it were the case that $2+2=5$, then it would not be the case that $2+3=5$.
(4) If it were the case that $2+2=5$, then it would be the case that $2+3=5$.

Each of the above counterfactuals contains impossible (necessarily false) antecedents. This means that there are no possible worlds in which these antecedents are true. After all, it is impossible to square the circle, and it is impossible that $2+2=5$. According to the standard analysis of counterfactuals:

(CF) "$A>C$" is true in @ iff either (a) there is no world where *A* is true or (b) every world *w* where *A* and *C* are true is more similar to the actual world than any world *w'*, where *A* is true but *C* is false.

In virtue of CF, sentences (1)-(4) are true since all of them satisfy condition (a). On the contrary, we would rather like to consider only *some* of them to be true and others to be false; for that reason, a more sensitive analysis of their truth is required.

To solve this problem, many philosophers have argued that one needs to invoke impossible worlds, i.e., worlds where what is impossible is true. They claim that just as for every possibility there is a possible world which represents it, then for every impossibility there is an impossible world which represents what is impossible from the actual world's point of view (e.g., Yagisawa 1988). As a consequence, the advocates of impossible worlds postulate worlds where, for example, a round square exists, 10 is a prime number, $2+2=5$, it is raining and not raining at the same time, etc.

To avoid the trivial consequences of postulating worlds where necessarily false claims are true, one should assume that impossible worlds are elements of other logical spaces than the space of possible worlds. It is worth noting that because of this, modal terms should be taken as indexical with respect to given logical spaces: What is impossible in our logical space (i.e., in all worlds which are ruled by classical logic) is possible in some other logical spaces (e.g., paraconsistent spaces). In this sense, every impossibility is true in some world, but that world has to be outside the set of possible worlds.

Of course, there is no "equality" between different impossible worlds. Some of them are closer (more similar) to the actual world than others. As we have already mentioned, there are issues that pertain to determining how to measure similarity between worlds. This was not easy in the case of standard analysis, and now, when one introduces a plenitude of impossible worlds, it is even more puzzling. Nevertheless, it seems that we can point to a claim which at least tells us what the most dissimilar world is:

> First, it is intuitive to claim that some impossible worlds are more similar to the actual world @ than others. For instance, the explosion world (call it *e*) at which everything is the case, that is, at which every sentence is true, seems to be as far from @ as one can imagine, provided one can actually imagine or conceive such an extremely absurd situation. Now, pick the impossible world, *t*, at which everything is as in @, except that I wear an impossible t-shirt which is white all over and black all over. Intuitively, *t* is closer to @ than *e*. (Berto 2013)[4]

Regardless of the detailed account of the similarity, the existence of a plenitude of possible as well as impossible worlds and their sets allows us to avoid the vacuous truth of counterfactuals with necessarily false antecedents. Thanks to these, one can easily extend the standard analysis by claiming that since every impossibility is true in some of impossible worlds, we can add these kinds of worlds to the original analysis:

(CF*) "*A* > *C*" is true in @ iff every (possible or impossible) world *w* where *A* and *C* are true is more similar to the actual world than any world *w'* where *A* is true but *C* is false.

This extension should keep the analysis of counterfactuals from being insensitive to the problem of counterpossibles. Sentence (1) is considered to be true because there is an impossible world in which the antecedent and the consequent of this counterfactual are both true, and this world is more similar to the actual world than any world where the antecedent and consequent of (2)

---

4 See also Nolan (1997).

are true.[5] Thanks to this, one can present non-vacuously true reasoning that is based on necessarily false claims.

3. The trivial world

The above extension works well for most examples of counterpossibles, but it seems that when it comes to the trivial world, troubles arise. Although it might be bizarre enough, postulating the existence of this world is the simple consequence of the claim that for every impossibility there is a world where it is true. If we agree that it is impossible that everything is true, then there is an impossible world where everything is true – the trivial world. Since we assumed that the actual world is ruled by classical logic, when considering the trivial world, it is worth assembling it with one of the fundamental rules of this logic, i.e., the so-called Rule of Explosion, also known as *ex contradictione sequitur quodlibet* (*ECQ*). It is usually expressed as an implication $[A \wedge \neg A] \rightarrow B$ and states that from contradiction everything follows. The reason we mention it here is that there is only one world where B as mentioned above is true, and this is the trivial world.

Analysis of the relationship between implication and counterfactuals has a rich history in the philosophical literature (cf. Bennett 2003, 20-44), but besides the many differences in the various approaches to this issue, lately one claim seems to be commonly accepted. It can be expressed as $A \rightarrow B \vdash A > B$, and it states that “any logical truth of the form $A \rightarrow B$ gives rise to the true conditional $A > B$” (Priest 2009, 331).[6] This connection between implication and conditionals allows us to consider the following sentences:

(5) If there were a true contradiction, then everything would be the case.
(6) If there were a true contradiction, then (still) not everything would be the case.

Let us assume that the antecedent and consequent of (5) are true in w_1, while those of (6) are true in w_2. From classical logic’s (i.e., the actual world’s)

---

5 Similarly in the case of (3) and (4).

6 See also Gibbard (1981); and Kratzer (2012, 87-9). It should be stressed that this does not mean that any true conditional results in a true implication.

point of view, the antecedents of both of these counterfactuals express impossibility, so in order to evaluate their truth we should assume that both w_1 and w_2 are impossible worlds. The important difference between them is that w_1 is the trivial world, whereas w_2 is a non-trivial one. Assuming that the actual world is ruled by classical logic, we would rather like to admit the truth of (5) than of (6). This is so because the first one is just a counterfactually expressed *ECQ*, and as such it tells us what would be the consequence of a true contradiction in the actual world, and in any other world in which classical logic is valid. If this is truly so, then (according to CF*) we have to admit that w_1 is more similar to the actual world than w_2. But as was stressed above, one of the basic assumptions in the theories of impossible worlds is that the trivial world is the most dissimilar from the actual world. If we assume *DTW* and admit that the trivial world (w_1) is the most dissimilar to the actual world, then w_2 is more similar than w_1. As a result, (6) becomes a true counterfactual and (5) should be taken to be false. If (5) is false, then *ECQ* is false (invalid) as well. In consequence, the analysis of counterpossibles leads to a rejection of one of the fundamental rules of classical logic, which means that classical logic is invalid in the actual world.

This conclusion might lead to at least two consequences. On the one hand, we can claim that since *DTW* leads to falseness of classical logic in the actual world, we should reject *DTW*. In our opinion this is a correct way of addressing the above issue. Nevertheless, what for us is a *modus tollens*, some philosophers might take to be a *modus ponens* and argue that it is the case that *ECQ* is false in the actual world. This result is consistent with those theories of impossible worlds which are based on paraconsistent logic (see Mares 1997; Priest 1997; Restall 1997). Although it is one of the interpretations of the concept of an impossible world, it leads to a controversial conclusion: that true contradictions are possible. After all, if the actual world is an element of space of paraconsistent logic, and every world of this space is a possible one, then it is possible that there are true contradictions. Because of this consequence, many theorists of impossible worlds would like to avoid changing the logic of the actual world in order to deal with impossibilities.

Moreover, the problem is more complicated than deciding what the logic of the actual world is. As we will see, the question of validity of *DTW* is in a way independent of the question about the logic of the actual world. One may argue that taking (5) to be true undermines the impossible worlds analysis of counterpossibles in general. After all, this entire framework was meant to show some non-vacuously true reasoning based on what is impossible, and (5) shows us that from contradiction everything follows. In this sense, every sentence that is both true and false should imply everything, and it seems to contradict the basic motivations of introducing impossible worlds in the first place. Now the question is: how can one believe in the truth of (5) and at the same time make some non-vacuous inferences based on paraconsistent logic?

To answer this question, we should notice that there is an important difference between assuming true contradiction in classical logic, on the one hand, and contradictions which are true in one of the worlds in the space of paraconsistent logic, on the other. When we are thinking about such a contradiction which does not lead to the truth of everything, we are considering the last option. In this sense, every non-vacuously true counterpossible with a contradiction as an antecedent is (implicitly or explicitly) assigned as true in the world of paraconsistent logic. Consider the two examples:

(7) If it were raining and not raining at the same time, then everything would be the case.

(8) If it were raining and not raining at the same time, then not everything would be the case.

Both of these contain impossible antecedents, and it seems that we can find two different contexts in which they have different truth values. If we try to analyze them with the assumption that classical logic is valid, then (7) would be true and (8) would be false, just as in the case above of (5) and (6). On the other hand, if the counterfactuals above were preceded by a claim such as "Assuming the validity of paraconsistent logic, …" then obviously we would say that (8) is true and (7) is false. After all, that is what the advocates of paraconsistent logic would like to claim. In other words, one can find a reason to believe that there is a context in which (7) is true and others where it is false. In this sense, just because we take (5) to be true does not mean we treat *every* contradiction in the same way; especially not those which are true in a world ruled by paraconsistent logic.

In virtue of the above, if one either hesitates to admit the truth of (5) and the falseness of (6), or one *does* believe that a change of the logic of the actual world would help to save the validity of *DTW*, one can easily change examples (5) and (6) to:

(5*) *If classical logic were valid*, and if there were a true contradiction, then everything would be the case.

(6*) *If classical logic were valid*, and if there were a true contradiction, then (still) not everything would be the case.

Similarly as our previous examples, both counterfactuals have an impossible antecedent. Moreover, (5*) corresponds to the trivial world w_1, and (6*) corresponds to the non-trivial world w_2. What differentiates our examples is that the antecedent of (5*) and (6*) is impossible *regardless* of what the logic of the actual world is. After all, no matter what the logic of the actual world is, the conjunction "classical logic is valid and there are true contradictions" is necessarily false. This shows that the consequence of accepting *DTW* is not merely that *ECQ* is invalid in the actual world, but rather that *ECQ* is invalid in classical logic in general. After all, (5*) expresses one of the basic views held by the advocates of classical logic. Obviously, if one believes that the actual world is ruled by classical logic, then this implies that *ECQ* is not valid in the actual world. Nevertheless, the problem that we are trying to point to does not affect only classical logicians. As (5*) and (6*) show, this problem is in an important aspect irrelevant to what the true logic of the actual world is. What is important is that according to classical logic, *ECQ* is a valid principle and that the consequence of *DTW* contradicts this.

4. Diagnosis

It seems that the problem with *ECQ* and impossible worlds as presented above is based on acceptance of the following assumptions:

(i) For every impossibility there is a world that represents this impossibility.

(ii) The valid implication $A \rightarrow C$ entails the true counterfactual $A > C$.

(iii) The trivial world is the most dissimilar to the actual world (*DTW*).

Because of this, if one would like to save the validity of *ECQ* in classical logic and give an interesting analysis of counterpossibles, one should reject one of the above assumptions. Let us consider the reasons for and the consequences of rejecting each of them.

The first assumption expresses the fundamental claim of the advocates of impossible worlds. Of course, it may be controversial, and leads to the "incredulous stare", mostly because it is difficult to conceptualize a world where everything is true (w_1). It is even more complicated to conceptualize worlds where classical logic is true, where contradiction is true and where it is not the case that everything is true (w_2). After all: what does it even mean to say that classical logic is true in a world where contradiction is true? The truth of one of these claims implies the falseness of the other. In this sense, one could say that we are in fact neither talking about classical logic nor about contradiction.

This objection seems to be the standard reaction to postulating the worlds w_1 and w_2. Someone might say that it is impossible for there to be a world where classical logic is true and where contradiction is true as well. Fair enough, but let us remember that we are dealing with *im*possible worlds, and a world where classical logic is true and contradiction is true is one of them. Because of this, if one would like to exclude the above-mentioned world from the modal universe, then there is no reason not to also exclude worlds where a round square exists or where 10 is a prime number.[7] It is difficult to find a reason for which one should accept the existence of a world where a round square exists and at the same time reject the existence of a world where classical logic is true and contradiction is true as well. Just as our understanding of a notion of being round excludes being square, our understanding of the notion of contradiction excludes the possibility of classical logic being true. As long as we accept that there are worlds where round squares exist or where 10 is a prime number, there is no reason to exclude worlds such as w_1 and w_2 from our analysis of impossibilities. After all, they represent impossibilities.

A possible justification for rejecting (ii) might be that *ECQ* is a logical law, and as such it remains valid in every possible world regardless of the truth value of (5) or (6). In this way the falsehood of (5) (or (5*)) would not result in the invalidity of *ECQ* in classical logic. Nevertheless, it seems that (5) expresses exactly the same claim that is expressed in *ECQ*, so it is difficult to imagine what could be a better way of expressing *ECQ* in the natural language than (5) is. As Graham Priest pointed out: "Conditionals may not express laws

---

7 Naturally one may take this as a reason for rejecting the view that there are impossible worlds. Although most philosophers do not believe in this kind of objects, the problem that we are dealing with is addressed to those who believe in a theoretical value of impossible worlds' analyses.

of logic; but which conditional holds may certainly depend on logical laws. Thus, $[A \wedge B] > A$ since $[A \wedge B]$ entails A" (Priest 2009, 330). Although the rejection of (ii) may allow one to avoid the problem that we have presented above, it may be considered to be misleading. The only way of taking *ECQ* to be true in the actual world and (5) to be false (and consequently (6) to be true) is if we consider the antecedent to be true in a world of paraconsistent logic. But as we have seen above, this is clearly *not* the trivial world, and what we are interested in is a world where classical logic is true, contradiction is true and where everything is the case, i.e., the trivial world.

It seems thus that what is left is to reject (iii). Among (i)-(iii), it is the least supported assumption of the analysis of counterpossibles in terms of impossible worlds. Compared to (i) and (ii), (iii) looks merely like a pre-theoretical intuition that is not so well supported by argument. As we know, some of intuitions are simply deceptive. Therefore, it is worth considering an analysis towards such a notion of similarity between worlds which will be consistent with rejecting the last assumption. Otherwise, we should conclude not only that the actual world is a world where classical logic is false, but also that *ECQ* is invalid in classical logic.

Surely one could argue that our investigation shows that, actually, (ii) is false. It might be argued that since validity is taken to be the truth in every possible world, then, when taking into consideration impossible worlds, (ii) has no applications anymore. This might be an interesting way of dealing with the problem that we are analyzing here; especially for classical logicians who would like to deliver an analysis of non-vacuously true counterpossibles and save the validity of *ECQ* at the same time. This may allow one to keep *DTW* as one of the guides for an interpretation of the notion of similarity. Nevertheless, what might be a justification for the rejection of *DTW* is that its problematic consequence is in some sense independent of what the correct logic of the actual world is (as (5*) and (6*) show). As such, if the dismissal of *DTW* would help to avoid it, it is worth considering such an interpretation of similarity which does not rely on this assumption.

5. The Strangeness of Impossibility Condition

The second of the rules that we are going to challenge is the Strangeness of Impossibility Condition (*SIC*). According to this condition, "any possible world is more similar (nearer) to the actual world than any impossible world" (Nolan 1997, 550). In this sense, a world where there is no woodpecker (which is a possible world) is more similar to the actual world than a world where a round square exists. Contrary to the claim of the dissimilarity of the trivial world, *SIC* is not very widely accepted, and some philosophers doubt its validity. We will join them here and argue that *SIC* should not be taken to be a guide for understanding the notion of similarity.

Let us start with an analogy. Consider three objects: a ball, a tomato and a ladder. If one asks "What is more similar to the ball? A tomato or a ladder?", most of us would probably answer "a tomato". When asked why, we can say that both have the same shape. This will make our answer correct, but only if we understood the question as "What is more similar to the ball as far as having the same shape?" But if one presents the question in a different way, e.g., "What is more similar to the ball as far as having the same nature?", the answer would be different. In this case we should say that the ladder is more similar. After all, a ladder and a ball are artifacts, while the tomato is not. This shows that it is very difficult to think about similarity *per se*. Usually, our understanding of similarity between objects depends on a chosen feature that we take to be the most important. In this sense, each time we compare objects we (either in an explicit or implicit way) focus on a given feature. Without this restriction the result of such a comparison might be misleading. Similarity understood in this way is in fact a ternary relation, S $\langle a, b, F\rangle$, i.e., object a is similar to object b because of factor (property) F. In this sense, two objects are similar with respect to property F iff they both have F. A ladder is similar to a ball because they are both artifacts, and a tomato is similar to a ball because they are both round. By analogy, being more similar (MS) is a quaternary relation MS $\langle a, b, c, F\rangle$, which states that because of factor F, object a is more similar to object b than object c is.

Consider the possible world as mentioned above where there are no woodpeckers (but where no circle is a square) and impossible worlds where a round square exists (but where woodpeckers also exist). When it comes to a lack of round squares (and presumably being possible) we can say that the former is more similar to the actual world than the latter. Nevertheless, we can also say that, when considering the number of woodpeckers, the last one is more similar to the actual world than the first one is. In this sense, the similarity between worlds depends on a chosen aspect. If the most important feature of a world is to have an adequate number of woodpeckers, and one does not care about geometrical impossibilities, then one can say that there is an impossible world that is more similar to the actual world than one of the possible worlds is.

Someone who would like to save the validity of *SIC* might argue that the most important feature of a world is whether it is possible or impossible. After all, we should consider worlds in their fundamental aspects, and logical or metaphysical possibility is one of them. Surely these are important features of a world, especially when we are dealing with an analysis of modality. By accepting this assumption, *SIC* might easily be taken to be true. Even more, it would be obviously true since it would state that when considering the feature of being possible, every possible world is more similar to the actual world than any impossible world is. Though it is difficult to argue against this claim, it is presupposed that the only important feature of a world is being either possible or impossible and, as we have seen, we do not have to compare worlds (neither any other object) only because of this feature. As such *SIC* should not be used as a guide for a proper understanding of the notion of similarity.

6. Conclusion

We believe that the above considerations give good reasons to claim that the trivial world should be taken to be more similar to the actual world than some non-trivial worlds are, and that there are impossible worlds which are (in some respects) more similar to the actual world than some possible worlds are. Because of this, both *DTW* and *SIC* should not be considered to be good guides for understanding the notion of similarity between worlds.

This conclusion raises two important questions – is it possible to deliver such an interpretation of the notion of similarity which does not rely on *DTW* and *SIC*? And if this is so, is the refutation of *SIC* necessary in order to save the validity of *ECQ* in the actual world (resp. in classical logic)? We believe that there is a positive answer to the first question. A project of such an account of the notion of similarity was delivered in Sendłak (2016). Although the interpretation that was presented in this work gives further reasons to dismiss *SIC*, we believe that *SIC* is independent of *ECQ* and *DTW*. After all, both w_1 and w_2 are impossible worlds, and as such *SIC* has no important application to determine which of these is closer to the actual world; it applies to them in exactly the same way.

Nevertheless, as we argued in Sendłak (2016), regardless of the problem of the validity of *ECQ,* one can indicate the reasons for a refutation of *SIC*. We believe that this modification in the interpretation of the notion of similarity (i.e. refutation of both *SIC* and *DTW*) helps us better understand the use of counterpossibles in general.[8]

References

BENNETT, J. (2003): *A Philosophical Guide to Conditionals*. Oxford: Oxford University Press.

BERTO, F. (2013): Impossible Worlds. In: Zalta E. N. (ed.): *The Stanford Encyclopedia of Philosophy*. Available at: http://plato.stanford.edu/entries/impossible-worlds.

GIBBARD, A. (1981): Two Recent Theories of Conditionals. In: Harper W., Stalnaker R. & Pearce G. (eds.): *Ifs*. Reidel, 211-247.

KRATZER, A. (2012): *Modals and Conditionals: New and Revised Perspectives*. Cambridge: Cambridge University Press.

LEWIS, D. (1973): *Counterfactuals*. Oxford: Blackwell.

MARES, E. D. (1997): Who's Afraid of Impossible Worlds? *Notre Dame Journal of Formal Logic* 38, No. 4, 516-527.

NOLAN, D. (1997): Impossible Worlds: Modest Approach. *Notre Dame Journal of Formal Logic* 38, No. 4, 535-572.

NOLAN, D. (2013): Impossible Worlds. *Philosophy Compass* 8, 360-372.

PRIEST, G. (1997): Sylvan's Box: A Short Story and Ten Morals. *Notre Dame Journal of Formal Logic* 38, No. 4, 573-582.

---

[8] Earlier versions of this material were presented in Bratislava (Slovak Academy of Science) at "Issues on the Impossible Worlds" in May 2014, in Warsaw (University of Warsaw) at "Philosopher's Rally" in July 2014, in Ghent (Centre of Logic and Philosophy of Science) at "Entia et Nomina" in July 2014, in New York at Graham Priest's graduate student seminar in November 2014 (CUNY Graduate Center)), and in Budapest at "Ontology and Metaontology" in July 2015 (Central European University). I am grateful to the participants of these meetings for their helpful comments and discussions. I would like to thank to the anonymous reviewers for this volume for their comments concerning the earlier versions of the paper.

This material is based on work supported by the Polish National Center of Science under Grant No. 2012/05/N/HS1/02794. Thanks to the Polish-U.S. Fulbright Commission I had the opportunity to develop the ideas presented here during my stay at CUNY Graduate Center.

PRIEST, G. (2009): Conditionals: A Debate with Jackson. In: Ravenscroft, I. (ed.): *Minds, Worlds and Conditionals: Themes from the Philosophy of Frank Jackson*. Oxford: Oxford University Press, 311-336.

RESTALL, G. (1997): Ways Thing's Can't Be. *Notre Dame Journal of Formal Logic* 38, No. 4, 583-597.

SENDŁAK, M. (2016): Counterpossibles, Similarity, and Abduction. In: Payette G. & Urbaniak R. (eds.): *Applications of Formal Logic. The Road Less Traveled*. Springer.

STALNAKER, R. (1968): A Theory of Conditionals. In: Rescher N. (ed.): *Studies in Logical Theory*: Oxford: Blackwell.

VANDER LAAN, D. (1997): The Ontology of Impossible Worlds. *Notre Dame Journal of Formal Logic* 38, No. 4, 597-620.

VANDER LAAN, D. (2004): Counterpossibles and Similarity. In: Jackson F. & Priest G. (eds.): *Lewisian Themes: The Philosophy of David K. Lewis*. Oxford: Oxford University Press, 258-274.

YAGISAWA, T. (1988): Beyond Possible Worlds. *Philosophical Studies* 53, No. 2, 175-204.

Organon F 23 (2) 2016: 177-203

The Narrowness of Wide-Scope Principles

DAVID BOTTING

Instituto de Filosofia da Nova. FCSH. Universidade Nova de Lisboa
4 Piso, 26-C Av. de Berna. Lisbon 1069-061. Portugal
davidbotting33@yahoo.co.uk

RECEIVED: 05-09-2015 ▪ ACCEPTED: 04-01-2016

ABSTRACT: In this paper I will propose that the unpalatable consequences of narrow-scope principles are not avoided by altering the scope of the principle but by changing the kind of conditional. I argue that a counterfactual conditional should do the trick and that the rational requirement of modus ponens can be understood as something like a "Ramsey test" on this conditional.

KEYWORDS: Coherence – counterfactual conditionals – modus ponens – Ramsey test – rationality – wide and narrow scope conditionals.

1. Introduction

What does rationality demand of us?

One thing demanded of us, many philosophers think, is to avoid having attitudes that are inconsistent or incompatible. So, we will often see *modus ponens* made into the rational principle that our beliefs must be deductively closed, that is to say that if you believe p and you believe $p \rightarrow q$, then believing q complies with what rationality requires and one is rationally criticizable if one does not have this belief while having the others.[1] Obviously these principles must be conditionals. But what parts of the conditional should be inside

[1] Deductive closure is a stronger condition than is actually required to avoid logical inconsistency, sufficient for which is the weaker condition that if you believe p and you

the scope of rationality's demands? And what kind of conditional should they be?

2. The Scope of the Conditional

2.1. Modus ponens as a narrow-scope material conditional

A straightforward formalization of *modus ponens* gives a rational principle something like this:

$B(p) \wedge B(p \rightarrow q) \rightarrow$ rationality requires $B(q)$

This is called "narrow-scope" because the scope of the propositional operator "rationality requires" is the consequent. What the conditional says, in words, is that if you believe p and you believe $p \rightarrow q$ then the consequent can be detached that says you are rationally required to believe q. Because it tells you that on the basis of the antecedent you should have a particular belief, or that you should draw a particular inference, this is called in the literature a *process-condition*. A process-condition tells you to reason in a certain, determinate way.

However, this seems to have counter-intuitive consequences in cases where it is not rational to believe p or to believe $p \rightarrow q$. Suppose that p is 'The moon is made of cheese' and q is 'The moon is made of a dairy product'. According to the principle, believing that the moon is made of a dairy product complies with the rational requirement and not believing that the moon is made of a dairy product violates the rational requirement, yet this is counter-intuitive

---

believe $p \rightarrow q$, then you comply with what rationality requires as long as you do not believe $\neg q$. Sometimes the principle is modified so that believing q complies with what rationality requires only when you care whether q or the necessity of q given p is sufficiently obvious. I do not intend to take any position on whether rationality demands something as strong as deductive closure of belief or whether something slightly weaker is required. Nor do I wish to engage the vexed issue of the normativity of rational requirements. The issue I wish to take up is over the formulation of these principles of rationality, and for this purpose the more straightforward principle will serve as illustration, and the approach I propose can be adapted to whatever principle of modus ponens is correct. Ultimately, my approach should extend to other principles as well, including those of practical rationality.

in three ways: 1) It is not rational (even if we changed the example so that *q* was actually true); 2) Surely we cannot somehow make it rational just by having beliefs that satisfy the antecedent of this conditional;[2] 3) It seems consistent that it could be believed by the subject herself that the beliefs that satisfy the antecedent of this conditional are not rational, and yet believing the consequent will still come out as complying with what is rationally required.

Perhaps it might be argued that (1) is actually false, that subjectively, if she is aiming at complying with what rationality requires of her, then assuming that pursuing this aim provides her a reason to modify her beliefs in order to attain that aim (i.e., logically consistent beliefs), she should believe that the moon is made of a dairy product, rationality being a constraint only on attitudes' consistency and coherence with each other and not on how they fit the world.

I think I can afford to be agnostic on this since it would not affect the problem I want to focus on which is (3), which is precisely that in some cases complying with the principle involves being *consciously* irrational. Aiming at logically consistent beliefs provides the reasoner with a reason not only to believe a proposition that is irrational and unjustified, but a reason to believe a proposition *she knows to be* irrational and unjustified. Paradoxically, rationality seems to require us to have beliefs we know to be irrational, and says that we are being irrational if we do not have such beliefs. Suppose that our reasoner's belief that the moon is made of cheese is one she simply cannot shake, despite the fact that she knows that she has no good reason for it and that she is not justified in having this belief. Since she knows this, she knows also that the proposition 'The moon is made of cheese' is not safe for use as a premise and hence that any consequences she infers on its basis will be likewise unjustified (although they may conceivably be true). Yet the principle implies that she complies with what rationality requires in believing these unjustified consequences even while knowing them to be unjustified, that to comply with what rationality requires involves consciously putting herself into a worse epistemic position than she is in now, and she is rationally criticizable if she does not.

---

2 As Broome (1999, 402-403) has pointed out, $B(p) \wedge B(p \rightarrow p) \rightarrow$ rationality requires $B(p)$ also fits the pattern of this principle, yet its implausible consequence is that we are rationally required to have whatever beliefs we actually have, or equivalently that our actual beliefs are infallibly those that we are rationally required to have. This objection is usually made in terms of reasons; having something as a belief cannot give you a reason to believe that it is true unless there is already a reason to believe it is true.

This is the main problem I wish to solve, and I do not see how the scope of the conditional affects the matter, having formulated the problem in such a way that it is actually neutral with respect to the scope of the rational requirement (as will become clear in a moment).

Also, with regards to (2) it is not simply a matter of whether the belief in the antecedent is rational or not. Let us suppose that this belief is rational and that there are objective reasons for it, but that the belief is not based on those reasons. To simplify, suppose that the belief is innate, imprinted at birth, and although the reasoner may come to learn of reasons that justify her having this belief, she would have it anyway, and if she came to learn of reasons that justify her dropping this belief, she would not do so. Should a principle like *modus ponens* make believing the consequences of such a belief the only way of complying with what is rationally required?

Until she grasps the relation between the belief and the reasons that justify it, I think the intuition is that she would not be rational in this case – the risk of propagating false beliefs is too great, even though in the particular case the innate belief is true and so would be all of its consequences.

Intuition is less clear after she does grasp the relation between the belief and the reasons that justify it, although her belief does not actually depend on the reasons in any way; that is to say, she is justified and knows herself to be justified in having the innate belief. Here, she improves her epistemic situation by following the principle, and that seems a good reason to say that the principle should be formulated in a way so that it does apply in this situation, in spite of the belief's questionable historical credentials. The moral is that even if we have the right attitudes we are not being rational if we fail to be aware of the normative relations between them.[3] This is a fairly weak historical condition,

---

[3] See Brunero (2005, 8) for discussion of the claim that coherence conditions, because they concern only combinations of attitudes, wrongly ignore whether any particular attitude was formed in a rational manner. In suggesting that attitudes not formed in a rational manner are not themselves rational, I am not necessarily saying that they are irrational; I am saying that coming to have these attitudes was not an exercise in rationality. They are non-rational.

I am not sure that this scenario is accurately described, since it assumes that once we have a belief with a particular propositional content then any reasons we may have or acquire for believing that content to be true must be linked to that particular belief-token or not be linked at all. I am inclined to think that in grasping the relation between the belief and the reasons that justify it one does *ipso facto* have a belief-token with that because being aware of normative relations is not to say that our belief-formation processes are responsive to these normative relations.

On the other hand, the following claim also seems intuitive:

> No Rationality Without Autonomy: S must be *autonomous* towards her attitudes in order for them to count as being held rationally.

In other words, if I were going to believe *q* anyway irrespective of other things I believed, then my believing *q* is not an exercise of rationality – something I have come to believe by trying to comply with what rationality demands of me or by a process of reasoning that ensures this compliance (whether one is reflectively aware of this or not) – even if it is correctly supported and coherent with other things I rationally believe. Perhaps we could even say that it violates something like the Principle of Conditional Non-Contradiction,[4] since it seems

---

content based/depending on those reasons, although this will be a different token from (in our hypothetical case) the innate belief. In short, we can have more than one token with the same propositional content.

4 Aristotle actually argues for something like this (sometimes called "Aristotle's Thesis") as a principle of logic on the grounds that he finds its consequence $p \rightarrow \neg p$ absurd. However, unlike $p \wedge \neg p$, $p \rightarrow \neg p$ does not actually violate the Law of Non-Contradiction, and some logical proofs actually use it. Modern logic, therefore, rejects "Aristotle's Thesis" as a principle of logic. It is possible, though it must be investigated further, that it may be resurrected as a principle of rationality, and if so, this might be a way of formally capturing the idea of being non-autonomous with respect to a proposition. However, we must be careful because logically true propositions and any theorem of the logical language will satisfy the condition of being true both when any other proposition is true and false, and our believing it as a logical truth seems to be captured by a condition like this; i.e, our believing a logical truth when knowing it to be a logical truth should always be rational whatever else we believe, so satisfying the conditional in this circumstance implies rationality rather than irrationality. On the other hand, it may not be the case that we would believe a logical truth whatever else we believe, even if we believe it to be a logical truth. For instance, we may believe a logical truth because we believe that we have a proof of it, and would not believe it had we no such proof (or belief that there was such a proof); furthermore, if we did believe it in the absence of belief in a proof, this belief would, I think, be irrational. The proof of the theorem is not the kind of thing whose truth or falsity is immaterial to the truth of the logical truth or the rationality of believing the logical truth. These are complications that would have to be worked out; theorems do not, I think, get their being rational for free, despite the impossibility of their being false.

to imply that $B(\Gamma) \rightarrow B(p)$ and $B(\neg\Gamma) \rightarrow B(p)$, since $B(p)$ would be true whatever else I believe, even if the Γ that I actually base my belief that p on does actually support p and I am quite unaware of the fact that I would believe that p even if I believed that Γ was false.[5] Since I do not say that $B(\neg\Gamma)$ is true – that I actually have this belief, or for that matter that I believe either of the conditionals – the attitudes need not be actually inconsistent, but the mere danger that they could be seems reason enough for caution.

There are various accounts of autonomy available, but the one I wish to appeal to here is from Mele (1995), where being autonomous toward an attitude is for it to be *sheddable*, where an attitude is sheddable provided that it results from our psychological processes operating in the normal way and could in principle be shed by their continued operation in the normal way. An attitude that we have been psychologically compelled to have by the intervention of something exogeneous to the normal operation of our belief-forming processes, or more simply because of their temporarily abnormal operation, and that we have no control over (the hackneyed examples being hypnosis and brain-washing) will be practically unsheddable. Practically unsheddable attitudes are held non-rationally.

Does this mean that rational requirements should be formulated in such a way that believing the consequences of such an unsheddable belief should not count as complying with them? If so, this is a stronger historical condition than that described above, because it says that the normative relations must not simply be grasped in the particular case but must actually guide our belief-

5 Suppose that I believe that I was immaculately conceived, and base this on the belief that I do not have a biological father; that is to say, I believe the conditional 'If I do not have a biological father, then I was immaculately conceived.' These beliefs would be consistent, and my belief that I was immaculately conceived would be justified (both propositionally and doxastically) by my belief that I do not have a biological father. But let us suppose, contrary-to-fact, that even if I believed that I do have a biological father, I would still believe that I was immaculately conceived (I am just that kind of guy). In this possible world, I would be prepared to countenance both conditional beliefs 'If I do not have a biological father, then I was immaculately conceived' and 'If I do have a biological father, then I was immaculately conceived' rather than give up the belief that I was immaculately conceived. From this I conclude that the belief that I was immaculately conceived is not rational in the actual world either; although this belief is justified it is not rational because it is not responsive to the belief it was justified by in relevant counterfactual situations.

formation processes. The intuition is unclear, because there is a sense in which the attitude is rational and a sense in which it is not.

2.2. Modus ponens as a wide-scope material conditional

Objections (1) and (2) are well-known and have led many to a wide-scope formulation of the rational principles that contrasts with the narrow-scope formulation given above. Thus, the rational principle should be something like this:

Rationality requires $[B(p) \wedge B(p \rightarrow q)] \rightarrow B(q)$

This is called "wide-scope" because the scope of the propositional operator "rationality requires" is the conditional as a whole. What the conditional says, in words, is that what is rationally required is to make the conditional itself true, and this can be done in two ways: by ceasing to have one of the beliefs referred to in the antecedent or by having the belief referred to in the consequent. It does not tell you that you should or should not have a particular belief, or that you should draw a particular inference, or that you should reason in a certain, determinate way, because although it is still true that you have a reason to make the conditional true, reasoning itself is not given a determinate direction in that the principle does not tell the reasoner *how* to make the conditional true. It is not a *process-condition* but a *state-condition*: it proscribes being in a state where there is a certain combination of attitudes which would constitute a counter-interpretation of the logical principle, which in this case is a state where I believe that p and $p \rightarrow q$ are true while also believing that q is false. What rationality requires through these principles is avoiding such combinations, i.e., attitudes (in this case, beliefs) that are incoherent. Satisfying the conditional avoids any such combination, but rationality judges symmetrically with regards to how the conditional is satisfied and is thus agnostic towards how the particular combination is best avoided.

The cause of the implausible consequences of the narrow-scope formulation is held to be the fact that the scope of what 'rationality requires' is the consequent of the conditional. By changing the scope to the conditional as a whole we avoid these consequences. For example, the wide-scope conditional does not have the consequence that rationality requires us to believe that the moon is made of a dairy product because we can obey this principle by dropping the irrational belief that the moon is made of cheese. Unlike the narrow-scope formulation, the consequent does not detach, so it is not the case that rationality requires that I believe *q*, though it is true that in believing *q* I would be complying with what rationality requires. Thus, for the wide-scoper it is strictly speaking incorrect to say that rationality requires us to have any particular belief – this is why I have tended to use the rather tortuous expression of a belief's complying with what rationality requires rather than simply that rationality requires having that belief, and although when discussing the narrow-scope formulation I could have used the simpler expression, my reason for not doing so was in order to formulate the problem in a way that did not depend on a wide or narrow-scope reading of the conditional. For the wide-scoper, there is more than one way of complying with what rationality requires; the point of having rationality require us to be such that the conditional is true is that there is more than one way of making the conditional true – we can make the consequent true (which is what the narrow-scope principle endorses exclusively) or we can make the antecedent false by dropping the antecedent beliefs.

Superficially this solution is attractive and the wide-scope view seems to avoid the consequences of the narrow-scope view. But does it really? Are there not at least some circumstances under which there is after all only one way to comply with what rationality requires, that is to say, only one way psychologically and/or physically open to us to make the conditional true? There are a number of ways we might imagine this happening that have turned up in the literature. In fact, even our earlier example seems to say that the only rational way to satisfy the conditional is to drop the irrational belief that *p* given that the belief that *q* is irrational. To make the case stronger we may suppose that we believe that believing *q* would be irrational. It seems then that the only way to be rational is by not having the belief that *p*.

In the last case, dropping the belief that *p* was the right way to comply with what rationality requires, so perhaps it might be thought that the complaint that it is the only way open to us to be rational is of little consequence. Of course, the problem is exacerbated when not believing that *p* is the *wrong* way to be rational. Suppose that we are doxastically akratic and believe that believing *q* is the right way to comply with what rationality requires, but for some reason cannot bring ourselves to believe that *q*; belief that *q* is not a psychologically open possibility for us. Again, the only way that we can actually comply with what rationality requires (at least, on purpose – this is an important qualification that will be discussed later) in these circumstances is by making the antecedent false, that is to say, by dropping beliefs that we may rationally have and quite likely believe ourselves rationally to have, but this is to make akrasia a rationally principled response to the situation we find ourselves in, yet surely akrasia is a paradigm case of irrationality. At the very least, principles of rationality should not provide the akratic with reasons to behave akratically. Supposing that it is psychologically open for the akratic simply not to comply with the rational principle (e.g., by just staying in the state he is now), then it seems that this is what he should rationally prefer, despite the principle's being violated.

Analogous arguments could be made for circumstances where we are simply unable to believe that *q*. These are all cases where making the antecedent false seems to be the wrong thing to do, yet it is what we must do to avoid finding ourselves with an incoherent combination of attitudes, and if we have a reason to avoid such combinations (and *ex hypothesi* we do have a reason to make the wide-scope conditional itself true) then we have a reason to do the wrong thing; despite the fact that rational requirements are not themselves reasons, taking coherence (as expressed by rational requirements) as a norm does provide reasons.

Cases where the only way to comply with what rationality requires is to make the consequent true are even more common, and in these cases the wide-scope principle works out the same as the narrow-scope principle after all. In fact, if we make time a factor, this is universally the case, since until we actually make the antecedent false by dropping the irrational belief (supposing now that this is psychologically and physically open to us) we are still in a situation where the only way to make the conditional true is to make the consequent true.[6] As before, if dropping the irrationally or non-rationally held belief is not psychologically and physically open to us (e.g., if they are unsheddable) then the only way (purposely) to comply with what rationality requires is by making

---

6 This way of pressing the general objection comes from Schroeder (2009, 227) who remarks that 'it follows that people are in general infallible about what they ought to do, as long as they do not try to change their minds'. Note also that the reconstruction of Broome's first-order model of practical reasoning in Bratman (2009, 14) has as a crucial premise: 'If you also believe that E only if M, and if these beliefs do not change, BC requires that you believe M; and that is where your reasoning can lead.' Even here it seems to be the mere fact of beliefs' not changing that seems to lead the reasoning one way rather than another, but if it did not do this it is questionable whether we could achieve any of our cognitive ends through reasoning alone.

the consequent true.[7] We have not solved the problem that the wide-scope principle was introduced to solve.

Perhaps it might be objected that the only thing that counts is that it is logically possible to make the conditional true in the right way, irrespective of whether it is psychologically or physically possible; the fact that we cannot do something does not alter the fact that it is what it would be rationally and epistemically best to do. But this still leaves the reasoner in a dilemma: there is an option he can take that would result in his complying with what rationality requires insofar as it would result in his attitudes going from a state where they are incoherent to a state where they are coherent.[8] His only other option is to leave his attitudes in a state of incoherence and thereby be irrational. Which should he rationally prefer? When coherence can only be purchased at the price of believing further falsehoods or the logical consequences of beliefs that are irrational and quite possibly believed to be irrational, I think it is plausible to think that he should leave his attitudes as they are, inconsistency notwithstanding. On the other hand, whoever consciously holds inconsistent attitudes seems to be rationally criticizable in a distinctive way, and it seems distinctly odd to give as an excuse: 'Well, what I actually wanted to do was to stop believing that the moon was made of cheese, but I couldn't.'

So far I have argued that wide-scope principles do not avoid the consequence that only one way of purposely complying with it is rational, and further that this way will often be the wrong way of complying with the principle

---

[7] Even the wide-scope formulation is subject to the kind of detachment that is called necessary detachment. Necessary detachment says that if *p* then *q*, and necessarily *p*, then if rationally required *p* then rationally required *q*. If we treat unsheddable beliefs as being necessary in the relevant sense, then we will detach as if the requirements were narrow scope. I owe this observation to an anonymous reviewer of *Organon F*. However, I am not convinced that unsheddable beliefs are necessary in the relevant sense. We must distinguish between detaching B(*q*) and detaching rationally required B(*q*). Narrow-scope formulations and necessary detachment detach rationally required B(*q*), whereas in the scenario described it is only B(*q*) that is detached. However, this is enough to create the problem. See the later discussion of Hussain's view.

[8] To make the incoherence more marked we may suppose that the subject actually believes the negation of the consequence, e.g., that the moon is not made of a dairy product. This makes the subject's belief set logically inconsistent and both the strong (deductive closure) version of modus ponens and its weaker version will apply to it. And unfortunately the result on both versions will be to retract the true and rationally held belief that the moon is not made of a dairy product.

and may even be known by the subject to be the wrong way and to be irrational, and that the reason for both of these things is beliefs that the reasoner actually has simply because he actually has them, irrespective of their rationality.[9] The wide-scope view does not seem to avoid the consequences it was expressly introduced to avoid, then, at least in certain circumstances. I do not wish to reject it completely, however, because I think that its central insight that the principles of rationality are principles prohibiting certain combinations of attitudes is correct and worth preserving, which is to say that it is still the truth of the conditional that rationality requires. Only I deny that this conditional is a material implication.

3. The type of conditional

The situation so far is that I have described a kind of scenario where reasoners would have to consciously put themselves into a worse epistemic situation than they are already in to comply with rational requirements, irrespective of whether those requirements are formulated as material conditionals with wide or narrow scope. This problem, at least, is not solved by altering the scope of the material conditional. The aim in this section is to investigate whether the problem is the conditional's being a material conditional rather than some other type.

In the first part I will note that originally Broome did not actually use a purely material conditional but what I will call, for want of a better name, a "quasi-material" conditional. This is defended in the work of Hussain (2007),

---

[9] Way (2010, 224-225) argues further that wide-scope principles do not avoid the consequence that we have *reasons* to obey the principle in a particular way. Although the wide-scope view avoids the consequence that any particular way of making the conditional true is required, making the conditional true is itself something that is required and plausibly something that we have a reason to do. It is plausible to suppose that if we have a reason to do something then we have at least some reason to do whatever is necessary and/or sufficient for doing it, or on a slightly different principle, whatever is a means to doing it. Since we have a reason to make the conditional true and, trivially, each way of making it true is sufficient for making it true, it follows that there is a reason to make oneself rational in *both* ways, and this is so simply because of attitudes we actually have regardless of whether they are rational or whether we have reasons to have them. Way calls this the *transmission problem*. I will not be dealing with this problem in this paper.

whose views will be discussed in some detail. It will be shown that he does not avoid the problem described above.

In the second part I will make tentative inroads into formulating the principles of rationality as counterfactual conditionals. Probably, this should be qualified as "quasi" too, for the conditional is not quite the classical counterfactual conditional, and the Ramsey test that I propose to use to evaluate the counterfactual is not quite the classical Ramsey test. All that I wish to present here is a basic approach to the problem.

3.1. Modus ponens as a wide-scope quasi-material conditional

It is interesting to note Broome's own attitude towards the conditional. In an early paper Broome (1999, 401-402) says that the relation he calls a rational requirement is not a material conditional but something like a material conditional with determination added where this determination is 'roughly analogous to causation' (Broome 1999, 401). This contrasts with a presentation of the wide-scope principle as something like 'It is rationally required that (you do not believe the antecedent or you do believe the consequent)' where the disjuncts are offered as the rational options for making the conditional true. Yet in Broome (2007) he says:

> When a wide-scope requirement holds, what is required of you is a material conditional proposition $p \varepsilon q$. We must be able to substitute logical equivalents within the scope of a requirement. So rationality also requires of you the contrapositive $\neg q \varepsilon \neg p$. Wide-scope requirements have this sort of symmetry.
>
> But sometimes this symmetry seems wrong. Look at the wide-scope formulation of the anti-akratic requirement …
>
>> Rationality requires of you that (You believe you ought to F ε You intend to F).
>
> Contraposing gives:
>
>> Rationality requires of you that (You do not intend to F ε You do not believe you ought to F).
>
> But the relation between believing you ought to F and intending to F is not symmetric
> … [It would be] irrational to disbelieve you ought to F because you do not intend to F. (Broome 2007, 35-36)

Here Broome seems to be saying that the wide-scope principle in question (admittedly not *modus ponens*) is a material conditional after all and as such it contraposes, and explains away the apparent irrationality of, for instance, being rational by akratically modifying one's beliefs in line with one's intentions, as the failure of a material conditional to capture a relation of explanation or grounding. Broome changed his mind in the intervening years: his later position seems to be that it is not the role of the principle to capture this relation, whether we call it determination, explanation or grounding. To capture this relation is not a matter of scope but a matter of the type of conditional (cf. Broome 2007, 36). This goes equally for non-instrumental principles: *modus ponens* contraposes into a principle where from *not* having a belief one reasons to *not* having other beliefs,[10] and it is not obvious that it is even possible to reason from the lack of a belief. It is nonetheless true that one would be complying with what rationality requires if one satisfied this conditional, so Broome could maintain that this is in fact a rational requirement even if it was one that could not be satisfied by reasoning. But, because the material conditional does not capture the relation of grounding, one could not say that one would be in this situation *because* one does not believe that *q*, or that it is rational to intend to F *because* one ought to F, irrespective of whether the principle is wide-scope or narrow-scope. A principle that resolved this asymmetry properly could not be a material conditional, and Broome (2007, 37) briefly suggests as an alternative a conditional that does not contrapose.

We are here considering three possibilities: two alternative readings of the wide-scope principle as an ordinary material conditional and as a material conditional with determination added, and as an as yet unspecified conditional that does not contrapose. I will later be suggesting a way of making good on this third alternative that Broome (2007) cursorily passes over, but first I want to

---

[10] At first glance contraposition seems reasonable for *modus ponens* and not to have the kind of problematic asymmetry that arise for instrumental principles; after all, if the consequent is false then the antecedent must be false. However, it should be remembered that the antecedent and consequent in this case are the beliefs and not the propositions believed; what you get when you contrapose the rational principle *modus ponens* is not the rational principle *modus tollens* which is:

Rationality requires $(B(\neg q) \rightarrow [B(\neg p) \vee B(\neg(p \rightarrow q))])$

but the much more peculiar:

Rationality requires $(\neg B(q) \rightarrow [\neg B(p) \vee \neg B(p \rightarrow q)])$.

note that even when a condition 'roughly analogous to causation' is appended to the material conditional, we still face the counter-intuitive result that in order to comply with what is rationally required a reasoner may need to consciously put himself into a worse epistemic situation that he is already in. This will motivate serious consideration of the third alternative, for which I will make some preliminary proposals.

The early way of reading the conditional as a material conditional with determination added is defended in Hussain (2007, 42-45) who denies that adding determination amounts to a narrowing of the wide-scope principle on the grounds that the important point is that the conditional nonetheless does not license detachment, and it is detachment that causes the problem, not the direction of reasoning as such. That you cannot, for instance, drop the belief that is held irrationally or believed to be held irrationally, does not, Hussain seems to say, imply that you are rationally required to believe its equally irrational consequence, even though this is in fact the only way open to you in which you can comply with what rationality requires, for this is a matter of detaching the consequent of a wide-scope principle as you would of a narrow-scope principle, and this detachment is not valid.

On this subject (in a context other than *modus ponens*) he makes a number of interesting comments:

> Consider the matter from the third person perspective of someone assessing *S*'s rationality and assessing what mental states *S* ought to have. To keep things simple, consider the case where not only is *S*'s belief [that *S* lacks sufficient reason to *X*] false, but in fact there is conclusive reason to *X* and so *S* ought to intend to *X*. The assessor can still think that rationality requires someone in *S*'s situation to get rid of the intention to *X*. Of course, the assessor doesn't think *S* should be in that situation and so doesn't think that *S* ought not to intend to *X*. Wide-scope is precisely what allows the assessor to think these thoughts without contradicting herself. (Hussain 2007, 43)

S seems to have got herself into a situation where the only way open to her to comply with what rationality requires is to make the consequent true despite the fact that in some sense what she should do is not be in that situation in the first place; her being in that situation indicates that somewhere in the past she formed a belief irrationally (or non-rationally). Hussain seems to be saying here that the wide-scope principle's result that complying with what rationality requires by making the consequent true holds even in cases where the antecedent cannot be made false.

He goes on:

> But if there is only one way of responding by reasoning to what is wrong with me, then does that not mean that I am rationally required to take that way? Well, given the situation I find myself in … there is only [one] way for me to proceed by reasoning. Rationality requires me to reason in a certain way, but rationality doesn't require me to be in the situation I am in: it does not require me to have the belief that I lack sufficient reason to X. That belief is not, so to speak, rationality's responsibility and neither, therefore, is the result that the only way for me to proceed by reasoning is not to have the intention to X. Things could have been otherwise without violating the rational requirements and they could still be otherwise.
>
> … [T]here are two ways of rationally resolving the conflict. The agent cannot change the belief by reasoning and thus things couldn't be otherwise by reasoning, but it still does not follow that what rationality requires is not to intend to *X*; i.e., that we can detach the conclusion that rationally requires that I not intend to *X*. What rationality requires is a specific process of reasoning in certain circumstances. One can engage in that process of reasoning or one can change the circumstances, though, sometimes, not by reasoning. It does not follow that rationality requires the particular outcome to the process that would result if the circumstances were not changed. This becomes clear when we see that if one were to change the circumstances – again not necessarily by reasoning – one would not be violating the requirement, indeed, one would now make it the case that one was living up to the requirement. Imagine … that I just, somehow, forget the content of the belief – I no longer have that belief. Now I would be back in conformity with the requirement, but not by reasoning. This is a way of ending up in accord with the requirement though not a way of coming into accord that I could manage by reasoning. The requirement directs me to reason in certain ways in certain circumstances. But removing those circumstances is one way for me to have the requirement no longer apply to me. (Hussain 2007, 44-45).

Principles of rationality are for Hussain principles for reasoning in the correct way, and if we cannot by reasoning make the antecedent of the conditional false then we are rationally required to engage in reasoning that, as it happens, makes the consequent true, and if we cannot by reasoning make the antecedent of the conditional false then this amounts to saying that this is not psychologically open for us. Our beliefs regain coherence if the antecedent becomes false of its own accord, but that is nothing to do with us or with rationality as such.

Hussain is trying to have his cake and eat it. In saying that "[w]hat rationality requires is a specific process of reasoning" he is trying to use the wide-scope quasi-material conditional as a process-condition in that, although it does not tell you that some belief is rationally required, it does provide a direction of reasoning. Narrow-scope principles are process-conditions because, by virtue of detaching their consequents, they tell you what to think or at least what it would be rational to think, but wide-scope conditionals are not generally thought of as process-conditions but as state-conditions. How exactly does Hussain make it into a process-condition, then?

Some background is necessary to answer this question. To a large extent, the dispute between the wide-scopers and the narrow-scopers is a dispute over whether the principles of rationality are state-conditions or process-conditions. This is a main bone of contention between Kolodny and Broome. Kolodny (2005) argues that "for any rational requirement on you, there must be a process of reasoning through which you can bring yourself to satisfy that requirement." Broome (2006, 2) quite explicitly rejects Kolodny's arguments for this view, expressing agnosticism towards its conclusion, and in Broome (2009, 18) we see why, for he says that reasoning cannot always bring you to satisfy a certain putative principle and that in that situation we are in a dilemma of concluding either that the principle "is not a genuine requirement of rationality, or alternatively that it is a genuine requirement but not one that reasoning can always bring you to satisfy." Ultimately Broome seems to prefer the latter horn of this dilemma.

Hussain seems to want to steer a middle course. He says that "[r]ationality requires me to reason in a certain way," apparently in agreement with Kolodny, and where "there is only [one] way for me to proceed by reasoning" Hussain says that rationality requires me to reason in that way. He also seems to accept something similar to Broome's point against Kolodny that not every way of satisfying a wide-scope principle (i.e., each way of making the conditional true) is such that it can be satisfied by reasoning. Where there is one way of satisfying the principle that can be reached by reasoning and one that cannot be so reached Hussain is clear that what rationality requires is to satisfy the principle by reasoning. This rationally required reasoning has a definite direction as indicated by the "added determination" in the conditional, and the result of this reasoning is the consequent of the wide-scope conditional. On Hussain's view of the quasi-material conditional, then, by complying with what rationality requires we will end up with the consequent; in effect (if not formally), we may detach the consequent. It is, then, a process-condition in this sense. This does not, however, make it equivalent to the narrow-scope principle. Detaching the consequent is rationally required and the destination to which rationally required reasoning leads us, but actually believing the detached belief is not rationally required; what is detached is the belief itself, whereas in the narrow-scope view[11] (and this is how Hussain's view differs from the narrow-scope view) what is detached is not the belief itself but the belief's being rationally required. So, we avoid the problem of being rationally required to have beliefs when the beliefs in the antecedent of the conditional are not themselves rational; for example, we are not rationally required to believe that the moon is made of a dairy product. But we are, it is implied, rationally required to draw this belief as an inference, at least if this is the only way to proceed by reasoning (and probably also if it is not, for the conditional's "added determination" already seems to give it that direction).

Apply this to our case. Despite Hussain's acceptance that an assessor could judge that the subject ought not to have the irrational or unsheddable belief that is the cause of the problem and in that sense the subject ought not to believe what follows from it, the subject is rationally required to engage in reasoning (and this is what makes it a process-condition) which results in believing whatever follows from it. Hussain (2007, 44) says also that there is no inconsistency for the subject: believing $\{p, p \rightarrow q\}$ and assuming that these beliefs do not change, then she must believe q. Note that the way I formulated the problem, by saying that complying with what rationality requires can lead one to have irrational beliefs, applies equally to Hussain's view. Believing q, it is true, is not rationally required, just as it was not rationally required in the ordinary wide-scope view, yet it is still the only way by reasoning and on purpose that we may comply with the rational principle and make our attitudes coherent. The belief that q is sufficient to cause the problem, so drawing a distinction between detaching the belief and detaching the belief's being rationally required does no good here.

---

[11] Similarly for the wide-scope view when there is necessary detachment. This is why I think there is a disanalogy between necessary detachment and the problem described here.

I will argue that it is implausible that a reasoner should be rationally required to reason as Hussain claims, at least in the case where the subject not only has a belief that is held non-rationally but is conscious of this fact and of the fact that this realizes a disvalue for her, as we supposed in objection (3) above and the ensuing discussion. Assuming that these beliefs do not change, a subject in this situation is incoherent whatever she does, whether she believes *q* or not. The subject's belief that she has an irrational belief is a second-order belief about her first-order belief with evaluative content. It does not seem too much of a stretch to suppose that she also has some second-order beliefs about various combinations (whether particular combinations or general patterns of combination) of the irrational belief with other beliefs. Some combinations (e.g., believing the logical consequences of the irrational belief) will increase the number of irrational beliefs, and so presumably the evaluation here will also be negative. However, some of these combinations seem to be rationally required, and it does not seem impossible that she has second-order beliefs about this as well. By complying with the rational requirement the subject consciously puts herself in an even worse position than before; in fact, the combinations that put her into these positions are precisely those that avoid incoherence. If it is nevertheless the case that she believes that she should do as she is rationally required, it must be because she believes that in this particular case coherence has a value of its own that outweighs the disvalue of holding and propagating irrational beliefs. It will not do to simply say that avoiding incoherence is a valuable disposition to have for this will not explain her decision or why the rational requirement applies in this particular case. I find this belief about the value of coherence somewhat implausible and would question whether this belief is itself a rational one; if satisfying the consequent is the only way you can satisfy the conditional by reasoning, then there is no reason or value in proceeding by reasoning in this particular case, and the subject will know this. So, I do not agree with Hussain's claim that rationality should require us to engage in the kind of reasoning he is talking about. Reasoning may be a useful disposition to have, but it is not one we ought to manifest in cases when we know that by doing so we are only making our epistemic situation worse. What is more, it seems that this is where reasoning itself can take us; we can reason (on the basis of evaluative beliefs about our first-order beliefs) that we would do better by not reasoning (on the basis of first-order beliefs). The correct thing to do is to leave oneself in the incoherent state after all, contrary to what Hussain says.

3.2. Modus ponens as a counterfactual conditional

Since the wide-scope conditional seems to endorse the view that in these cases rationality requires one to consciously form beliefs that one knows will increase the overall irrationality of one's set of beliefs, it cannot be correct. We need another kind of conditional. The counterfactual conditional does not contrapose or detach and seems to instantiate the kind of causal and grounding relation that Broome talks about. I will now sketch an account of a principle of rationality using counterfactual conditionals.

According to the Ramsey test for the truth of counterfactual conditionals, the open counterfactual conditional $p > q$ is true if adding p to a body of knowledge would, after minimal adjustments to preserve consistency, result in q belonging to a body of knowledge. That is to say, to add $p > q$ to the body of knowledge is to commit oneself to adding q to the body of knowledge should it come about that p is in the body of knowledge. This means that if $p > q$ already belongs to the body of knowledge, the preservation of consistency dictates that if p came to belong to the body of knowledge then so also would q, or else $p > q$ would be false. This derives something similar to the logical (not rational) principle of *modus ponens* for counterfactual conditionals from the Ramsey test something like

$$B(p > q) \to [B(p) \to B(q)]$$

The antecedent is the knowledge base. It is important to understand that what this is telling us is what must be the case when the counterfactual conditional is true; it is not telling us what beliefs to have – all references to beliefs are to be understood subjunctively as a test on the truth of the conditional, including the reference to belief in the counterfactual conditional.

What we want is a principle where it is a material conditional that one has a prior belief in and a counterfactual conditional that one is required to make true, e.g.,

$$B(p \to q) \to O[B(p) > B(q)]$$

This is a kind of intermediate-scope conditional[12] and seems to put the grounding in the right place – we want to be able to say that we ought to be such that

[12] See Way (2011) for another example of intermediate-scope conditionals.

our belief that q is grounded on our belief that p. Unfortunately, however, the principle does not say this – note that since the truth of the counterfactual conditional $p > q$ meant that $B(p) \rightarrow B(q)$, the truth of the counterfactual relation between *belief* that p and the *belief* that q that is here in the scope of the ought-operator (i.e., $B(p) > B(q)$) means a relation of material implication between the belief that the subject believes that p and the belief that the subject believes that q (i.e., $B(B(p)) \rightarrow B(B(q))$; the body of knowledge after the revision would be $\{p \rightarrow q, B(p), B(q)\}$, or in other words $B(p \rightarrow q)$, $B(B(p))$, and $B(B(q))$.

So, this principle does not give us what we want, which is a relation between $B(p)$ and $B(q)$ rather than a relation between $B(B(p))$ and $B(B(q))$. More importantly, the principle is not valid: $B(B(q))$ does not follow from $B(p \rightarrow q)$ and $B(B(p))$ because the former is a conditional concerning the contents of belief and the latter a second-order belief. We can cope with this in two stages.

First, we must be able to convert the second-order belief that p to a first-order belief, giving a principle more like

$$\{B(B(p)) \rightarrow B(p), p \rightarrow q\} \rightarrow (B(p) > B(q))$$

If we add $B(p)$ to the body of knowledge now we will get $B(q)$ because we can get $B(B(p))$ as the first step of the Ramsey Test, $B(p)$ from $[B(B(p)) \rightarrow B(p)] \wedge B(B(p))$, and then $B(q)$ from $B(p) \wedge B(p \rightarrow q)$.

This still does not give us what we want. Although it is having $B(q)$ and not $B(B(q))$ that we ultimately want to say is rationally required in the given situation, the truth of the counterfactual $B(p) > B(q)$ requires still that $B(B(q))$ would be in the body of knowledge after the hypothetical addition of $B(p)$, and from the principle above all we know is that $B(q)$ would be in the body of knowledge.

The second stage, then, is to amend the original body of knowledge further to convert the first-order belief that q to a second-order belief that q, giving the principle

$$\{[B(q) \rightarrow B(B(q))], [B(B(p)) \rightarrow B(p)], p \rightarrow q\} \rightarrow (B(p) > B(q))$$

If we add $B(B(p))$ to the body of knowledge now we will get $B(B(q))$ because we can get $B(q)$ as previously described and then $B(B(q))$ from $B(q)$ and $B(q) \rightarrow B(B(q))$. Given these three conditionals (the knowledge base), we make the counterfactual true either by believing both p and q or believing neither. However, later I will propose that when we also have a negative evaluative belief about our belief that *p* (such as that it is irrational or unsheddable), the counterfactual can be evaluated in such a way that we can make it true by not having the belief that *q*.

Plausibly, the only reason we should be able to convert from first-order beliefs to second-order beliefs and back again for these particular propositions is if this were so for all propositions, that is to say if $\forall p\,.\,B(B(p)) \to B(p)$ and $\forall p\,.\,B(p) \to B(B(p))$. Smullyan (1986) calls those for whom these are true *stable reasoners* and *normal reasoners* respectively. Those for whom these are not true are called *unstable reasoners* and *peculiar reasoners* respectively, and although such reasoners find themselves in a strange psychological position they are not necessarily inconsistent. I propose, then, that the following principle be restricted to *stable* and *normal reasoners*:

MODUS PONENS*: $B(p \to q) \to$ rationality requires $(B(p) > B(q))$

This says that stable and normal reasoners are rationally required to have beliefs such that believing *q* counterfactually depends on believing *p* if they believe that $p \to q$.[13] If you comply with MODUS PONENS*, then if you believe that *p* then you will believe that *q*, and if it had not been the case that you believe that *p* then it would not have been the case that you will believe that *q*. However, this seems only to be true if you did not already believe that *q*. To avoid this, the revised principle is:

MODUS PONENS**: $(B(p \to q) \wedge \neg B(q)) \to$ rationality requires $(B(p) > B(q))$

[13] It simplifies the Ramsey test if these universal conditionals are part of the body of knowledge, and if we insist that they must be then this restricts the principle further to those who know themselves to be *stable* and *normal reasoners*. Or it may be that the requirement is actually a requirement for everyone, but only normative for *stable* and *normal reasoners* or for those who know themselves to be so. But the ability to make these inferences in the cause of making minimal adjustments to preserve consistency in the knowledge base would perhaps be sufficient, or even it may be part of the notion of consistency in use that the knowledge base should be such that it satisfies the constraints of stability and normality.

This seems to guarantee that believing *p* is the cause or grounds of believing *q* when these states are the result of complying with the principle.

Does it solve our problems? Suppose again that our reasoner believes *p* and is not able not to believe *p*. Adding the belief that *p* to a knowledge base including the belief that $p \rightarrow q$, a belief that *q* will result for stable and normal reasoners. This is because believing that *q* appears to be the minimal adjustment that preserves consistency with what we believe when updated with what we have hypothetically added.

However, consider the case where we know that believing *p* is unsheddable and/or irrational. We will not, in this case, have a coherent knowledge base to begin with, and whatever we do – whether we believe *q* or not – we will not get a coherent knowledge base, as indicated earlier; although the contents of our beliefs may be logically compatible, there are certainly some among them that it is irrational to hold together, e.g., a belief B and the higher-order belief that one ought not to have belief B, and in general to have those attitudes held akratically or non-autonomously. Here we must settle for maximizing consistency, and in this case I will stipulate that it is permissible *not* to preserve in our considerations what we have hypothetically added.

In fact, this is what we are rationally required to do in this case, on the grounds that we are aware that this belief is the result of a belief-forming mechanism that is functioning abnormally, and therefore the closest possible world W in which the knowledge base is fully consistent is one where this mechanism is functioning normally. In W, the belief that *q* is not part of the knowledge base either. So, the belief that *q* still depends on the belief that *p* – the conditional still passes this version of the Ramsey test – because in W we have neither of these beliefs.[14] Let me put it this way: normally, for B(*q*) to depend counterfactually on B(*p*) it must be the case that if B(*p*) is present then B(*q*) is present and if B(*p*) is absent then B(*q*) is absent. I am proposing instead that B(*q*) can be absent in the actual world even though B(*p*) is present, if in the

---

[14] This may be a little hasty, because the world where the belief-forming process functions normally is not necessarily one where it does not result in the belief that *p*. As I said before, having bad credentials does not necessarily mean that *p* is false or is a belief one ought not to have if everything were working perfectly. It seems enough to say here that it is not the case that one ought to believe that *p*, i.e., that this belief is not one that needs to be preserved; we do not need to say that the belief that *p* is one that one ought not to have. That is to say, $\neg O(B(p))$ rather than $O(B(\neg p))$ or even, perhaps, $O(\neg B(p))$.

actual world the knowledge base is incoherent and in the nearest possible world in which the knowledge base is coherent B(*q*) is absent, and this is because in that world B(*p*) is also absent; the presence or absence of B(*q*) in this world counterfactually depends on the presence or absence of B(*p*) in W, since by grounding it on B(*p*) in W one would minimize avoidable irrationality. Counterfactual dependencies between our attitudes then should depend on what is true in W – in the closest maximally coherent world – rather than in the actual world. Although we cannot shed the unsheddable belief in fact (in the actual world), it is still the case that the knowledge base where the belief's consequences are not added is more coherent than the knowledge base where they are; we cope with the unsheddable belief simply by making the truth of the conditional depend on counterfactual situations where the belief is not unsheddable and everything functions normally. Although we cannot get to W in this way (because this requires us to shed the unsheddable), we can get as close as possible to it, as coherent and consistent as it is possible for us to be, by complying with MODUS PONENS**. I will continue to talk of this as 'making a counterfactual conditional true' though I acknowledge that it is a non-standard interpretation of the phrase 'counterfactual conditional' and a non-standard version of the Ramsey test that I use to evaluate it.

This gives the right result in the case where we are aware that believing *p* is unsheddable and/or irrational. But what if we do not know this? There would not appear to be any inconsistency in this case, or therefore any obstacle to believing *q*. I admit that I am not really sure how to answer this, but one might bite the bullet: we actually should rationally prefer to believe *q* after all, and the reason that we have intuitions to the contrary and for accepting the No Rationality Without Autonomy claim is because when considering this scenario we are *ipso facto* in the position of knowing that the belief is unsheddable and/or irrational, and take it in the same way that we take the case above where the subject also knows this and we should not rationally prefer to believe *q*.

This seems to work out like a version of Schroeder's account of weak subjective reasons; the belief that *p* does provide a subjective reason for believing that *q* even if *p* is irrational, as long, I would add, as we are not aware of that fact. Schroeder gives the following example: you see Tom Grabit leave the library, pull out a book from under his shirt and run away. On this evidence you form what appears to be a fairly safe conclusion that Tom has stolen a book. Suppose now that you learn that Tom is indistinguishable from his identical twin Tim. Now your conclusion is less safe – a 50-50 bet, in fact. On learning further that there is a third identical sibling Tam it has now become more likely that your original belief that Tom stole a book is false than true. Objectively, Tom always had identical siblings and there was no good objective reason for the conclusion. Yet your belief-forming processes have functioned correctly and your behaviour is rational, and this seems to imply that your beliefs generated reasons on their own. So, an unsheddable belief that you do not believe to be unsheddable generates a reason to believe its consequences. Only knowledge that the original information was unsound, e.g., if after Tom leaves the library you hear a director saying 'Cut!,' are there no subjective reasons at all for believing that Tom has stolen a book. Schroeder (2004, 358) calls this *complete undermining*. Maybe knowledge of lack of autonomy could be considered as *completely undermining* – when an unsheddable belief is believed to be unsheddable, it no longer generates any reasons. However, remember that unsheddable beliefs are not always false or irrational since even malfunctioning processes will get things right sometimes. Throughout I have been assuming that unsheddability is an epistemic vice, but the reality may well be far more nuanced than this; our ordinary perceptual apparatus gives us beliefs that are in many respects unsheddable, but this is a feature of their proper functioning, rather than their malfunctioning.

4. Conclusion

In this paper I considered several objections in the literature that the wide-scope strategy did not avoid the consequence it was designed to avoid, namely that to comply with what rationality requires sometimes means believing the consequences of beliefs that are irrational or otherwise defective, whether you are aware of this defectiveness or not. This is typically because the belief named in the antecedent of the conditional is one that one cannot avoid having, for one reason or another (although there is no reason in principle why it is not a consequent that one cannot bring oneself to have that is not the source of the problem). Following a suggestion of Broome (2007, 37) I suggested that the problem is not with the scope of the conditional but the type of the conditional, and proposed that instead of a material conditional we should put in a conditional that does not contrapose or detach, like a counterfactual conditional. Noting a similarity between the rational requirement of modus ponens and the Ramsey test, I suggested that we are rationally required to have those beliefs that passed the Ramsey test. Then, I worked through some complications so that the conditional properly linked the beliefs themselves rather than their contents, coming up with:

> **For all *stable* and *normal* reasoners, $(B(p \rightarrow q) \wedge \neg B(q)) \rightarrow$ rationality requires $(B(p) > B(q))$**

An interesting question I intend to leave open is whether rationality (or anything else, for that matter) requires one to be a stable and normal reasoner; there is reason to think that what distinguishes the kind of rationality human cognizers enjoy in contrast to lower animals and that is necessary for autonomy is that we are able to reflect on our own practices and thought processes and that this is mentally represented by higher-order attitudes, but this is a far more general requirement than having these specific inferential dispositions to convert between first- and second-order beliefs.

Another question that I intend to leave open is cashing out the clause, essential to the Ramsey test, of minimal adjustment to maintain consistency. This goes beyond logical consistency and I envisage it as excluding certain paradoxical cases like Moore-sentences, violations of the Law of Conditional Non-Contradiction, and in particular using as a basis for belief revision a belief you believe to be not properly connected to reasons, for even if this belief is true and you take it or believe it to be such, any kind of closure, whether under deductive entailment or some more limited principle, will result in a network of beliefs that is fundamentally unsafe. It is an essential part of my analysis that the rational requirement should be formulated in such a way as to make believing the consequences of such a belief rationally impermissible, for even though by having this belief one would maintain consistency between the beliefs named in the conditional, because of the bad history of the antecedent belief it is still more consistent over all (because of the comparative closeness of the counterfactual world in which there is no bad history) not to believe its consequences. However, consistency cannot be construed so widely as to smuggle in all the requirements of rationality that it is being used to explain, or to reduce to the claim that the beliefs we are rationally required to have are those that are best all things considered, for although this is undoubtedly true there do seem to be local requirements that are no less strict for the fact that sometimes global requirements may require us to violate them. Perhaps the way to do this would be to exclude normative beliefs from the body of knowledge, although evaluative beliefs will be necessary, for it is these that would comprise the subject's being conscious of the fact that she would be making her own epistemic situation worse in the scenarios in question.

One final thing to note is that the rational requirement of modus ponens involves belief in the material conditional $p \rightarrow q$. I have not been much concerned with the credentials of this belief or how it has emerged. However, the Ramsey test for the conditional $p > q$ seems to follow equally as for the material conditional, and suggests that the rational requirement can be further generalized to say:

For all *stable* and *normal* reasoners, $(B(p > q) \wedge \neg B(q)) \rightarrow$ rationality requires $(B(p) > B(q))$

This seems to successfully rule out subjects consciously putting themselves in a worse epistemic situation than that from which they started.

Acknowledgments

This work was supported by the Portuguese Foundation for Science and Technology (FCT) under grants SFRH/BPD/77687/2011 and PTDC/FIL-FIL/110117/2009.

References

BRATMAN, M. (2009): Intention, Belief, and Instrumental Rationality. In: Sobel, D. & Wall, S (eds.): *Reasons for Action*. Cambridge: Cambridge University Press, 13-36.

BROOME, J. (1999): Normative Requirements. *Ratio* 12, No. 4, 398-419.

BROOME, J. (2006): Reasoning with Preferences. In: Olsaretti, S. (ed.): *Preferences and Well-being*. Cambridge: Cambridge University Press.

BROOME, J. (2007): Requirements. In: Rønnow-Rasmussen, T., Petersson, B., Josefsson, J. & Egonsson, D. (eds.): *Hommage à Wlodek. Philosophical Papers Dedicated to Wlodek Rabinowicz*. Sweden: Lund University. Available online at: http://lup.lub.lu.se/record/753262.

BROOME, J. (2009): The Unity of Reasoning. In: Robertson, S. (ed.): *Spheres of Reason*. New York: Oxford University Press, 62-93.

BRUNERO, J. (2005): Two Approaches to Instrumental Rationality and Belief Consistency. *Journal of Ethics and Social Philosophy* 1, No. 1, 1-20.

HUSSAIN, N. (2007): The Requirements of Rationality. Unpublished manuscript. Available at: http://web.stanford.edu/~hussainn/StanfordPersonal/Online_Papers_files/HussainRequirementsv24.pdf

KOLODNY, N. (2005): Why Be Rational? *Mind* 114, No. 455, 509-563.

MELE, A. (1995): *Autonomous Agents*. Oxford: Oxford University Press.

SCHROEDER, M. (2009): Means-End Coherence, Stringency, and Subjective Reasons. *Philosophical Studies* 143, No. 2, 223-248.

SCHROEDER, M. (2004): The Scope of Instrumental Reason. *Philosophical Perspectives* 18, Ethics, 337-364.

SMULLYAN, R. M. (1986): Logicians Who Reason about Themselves. *Proceedings of the 1986 Conference on Theoretical Aspects of Reasoning about Knowledge*. San Francisco (CA) – Monterey (CA): Morgan Kaufmann Publishers Inc., 341-352.

WAY, J. (2010): Defending the Wide-Scope Approach to Instrumental Reason. *Philosophical Studies* 147, No. 2, 213-233.

WAY, J. (2011): Explaining the Instrumental Principle. *Australasian Journal of Philosophy* 90, No. 3, 487-506.

Organon F 23 (2) 2016: 204-221

Δ-TIL and Normative Systems[1]

DANIELA GLAVANIČOVÁ

Faculty of Arts. Comenius University in Bratislava
Gondova 2. 811 02 Bratislava. Slovakia
dada.baudelaire@gmail.com

RECEIVED: 22-12-2015 ▪ ACCEPTED: 12-03-2016

ABSTRACT: According to a widespread view, deontic modalities are relative to normative systems. Four arguments in favour of this suggestion will be presented in this paper. Nevertheless, I have proposed and defended an analysis of deontic modalities in terms of Transparent Intensional Logic (TIL) that is non-relativistic (with respect to normative systems) and accommodates minimal semantics of TIL. This leads to a question whether one can do justice to arguments for deontic relativism and put forward a *relativistic* analysis of deontic modalities in TIL. The main aim of this paper is to amend the former analysis of deontic modalities in terms of TIL to incorporate both the standard (relativistic) view and the minimal semantics of TIL.

KEYWORDS: Circumstances of evaluation – conflicts of obligations – deontic relativism – minimal semantics – normative systems – Transparent Intensional Logic.

0. Introduction

Deontic operators such as “it is obligatory”, “it is forbidden” and “it is permitted” are of a particular interest to descriptive deontic logic. These operators are sentential operators, i.e. as Dretske (1970, 1007) puts it, “when affixed to

[1] I would like to thank M. Zouhar for his brilliant criticism, useful remarks and interesting discussions. I am also indebted to L. Bielik, F. Gahér, I. Sedlár and the anonymous referees of *Organon F*.

a sentence or statement, they operate on it to generate another sentence or statement." The pre-theoretical meanings of these operators are called *deontic modalities*. I have proposed and defended an analysis of deontic modalities in terms of Transparent Intensional Logic (TIL);[2] see Glavaničová (2015a; 2015b). I will use the term "Δ-TIL" to refer to it. In sum, Δ-TIL makes a (semantically based) distinction between *implicit* and *explicit* deontic modalities. The former are analysed as properties of propositions and the latter as properties of propositional constructions. The distinction proves to be useful in resolving deontic paradoxes, but also in analysing strong and weak permissions.[3] These are the main motivations for employing Δ-TIL. The analysis is non-relativistic[4] and is in perfect line with the spirit of minimal semantics of TIL. On the other hand, there are substantial arguments in favour of deontic relativism. Deontic relativists argue that deontic expressions (and their meanings) are relative to various authorities (I will refer to them as "normative systems"). The main aim of this paper is to show that Δ-TIL can accommodate deontic relativism without violating its minimal semantics.

Section 1 presents four arguments in favour of deontic relativism. Section 2 contains a brief summary of Δ-TIL. Section 3 introduces the problem of implementing deontic relativism to Δ-TIL and section 4 suggests two possible solutions of this problem. Section 5 concludes the results and the final section examines some possible objections to the proposed analysis.

1. A case for deontic relativism

We can state deontic relativism as follows:

(DR) Normative systems enter into the truth-conditions of some descriptive deontic sentences.

---

[2] TIL was comprehensively introduced in Tichý (1988). See also recent works on TIL, most notably Raclavský (2009), Duží, Jespersen & Materna (2010) and Duží & Materna (2012). I will briefly explain some basic notions of TIL in Section 2 of this paper.

[3] Hansson (2013) argues for the usefulness of the distinction between implicit and explicit permissions in a similar vein.

[4] Whenever I refer to relativism in the present paper, I have in mind deontic relativism, i.e. relativism with respect to normative systems.

Without loss of generality, we may confine our attention to deontic sentences of the form $O\varphi$ (i.e. *φ is obligatory*), since parallel arguments can be made for descriptive deontic sentences of the form $F\varphi$ and $P\varphi$. Consequently, we may replace (DR) with:

(DRO) Normative systems enter into the truth-conditions of some $O\varphi$-sentences.

The first argument in favour of deontic relativism then goes as follows: Let us consider a situation, where we talk about $O\varphi$-sentences without mentioning any normative system. A person *A* asks a person *B*, whether some $O\varphi$-sentence is true or not. It may happen that (i) *B* assigns a truth-value to a given $O\varphi$-sentence with respect to *relevant* normative system or (ii) *B hesitates to answer* the question and asks for further information.

Let *P* represent the sentence "Some men have more than one wife at a time" and let us look at examples of cases (i) and (ii):

The case (i): Imagine that Mr. Fiable, an inhabitant of France, is in Saudi Arabia, asking one of its inhabitants, Mr. Amin, whether $O\neg P$ is true or not. Suppose that Mr. Amin is a reliable source of information about the legal system of Saudi Arabia. He supposes that Mr. Fiable's question concerns the legal system of Saudi Arabia and replies that $O\neg P$ is false. From now on, Mr. Fiable will (truly) think that the legal system of Saudi Arabia permits polygamy.

The case (ii): Imagine that case (i) has never happened. Mr. Amin and Mr. Fiable are visiting Tilburg. Neither of them knows the Dutch legal system. Again, Mr. Fiable asks Mr. Amin, whether $O\neg P$ is true or not. In this case, Mr. Amin is not likely to make similar supposition as in the case (i). He would hesitate and ask which normative system the question concerns.

However, in both cases, Mr. Amin was not able to assign a truth-value to $O\neg P$ without relativizing it to normative systems. Thus, both cases support (DRO).

The second argument in favour of deontic relativism has the following form: It is quite reasonable to demand that normative systems be internally consistent. The commonly employed system of deontic logic – *Standard Deontic Logic* (SDL), has an axiom that accommodates this requirement; cf. McNamara (2006, 207-208):

(A1) All tautologous wffs of the language.

(A2) $O(\varphi \to \psi) \to (O\varphi \to O\psi)$

(A3) $O\varphi \to \neg O\neg\varphi$

(R1) If $\vdash \varphi$ and $\vdash \varphi \to \psi$ then $\vdash \psi$

(R2) If $\vdash \varphi$ then $\vdash O\varphi$

In particular, it has an axiom A3, which tells us that if φ is obligatory, then its negation is not. As Goble (2000, 113) puts it, this principle "explicitly precludes conflicts of obligation". However, different normative systems can give rise to conflicts of obligations (sometimes called normative contradictions or moral dilemmas). The conflict of obligation is a statement of the form $O\varphi \wedge O\neg\varphi$. Besides entailing normative conflicts, different normative systems can be explicitly contradicting each other. This happens when one normative system permits some φ (i.e. $\neg O\neg\varphi$ holds for such system), whilst the other does not (i.e. $O\neg\varphi$ holds for such system).[5]

Let Q represent the sentence "Antigone buries her brother Polynices". Consider the following story:

> **The case of Antigone:** Polynices is a (dead) traitor to the city. Creon is a king. The burial of Polynices is forbidden by Creon's proclamation. Therefore, it ought to be the case that $\neg Q$ (under *human law*). However, the soul of Polynices needs the proper burial of his body to proceed to the underworld. Polynices should go to the underworld, so the gods demand his burial. Antigone is the only one who is willing to bury Polynices. Therefore, it ought to be the case that Q (under *divine law*).

Therefore, we have both $O\neg Q$ and OQ. Consequently, we can derive a contradiction by deriving $\neg O\neg Q$ or $\neg OQ$ (by A3 and R1). SDL as it stands thus cannot consistently allow for conflicts of obligation even across different normative systems.

A possible solution is deontic relativism. Let O_xA represent schematically the formula OA relativized to normative system x. We should amend A3 in such a way that if φ is obligatory (under a *certain* normative system x) then $\neg\varphi$ is not (under *that* normative system); schematically:

---

[5] Recall our first example concerning polygamy.

(A3*) $O_x\varphi \rightarrow \neg O_x\neg\varphi$

The reasonable requirement of internal consistency is preserved, whilst the unreasonable requirement of consistency across various systems is dismissed: "Each set of norms or regulations is presumed to be internally consistent, and conflicts only emerge as a result of rivalry between sets of norms" (Goble 2000, 117). Furthermore, remaining axioms and rules have to be decorated with subscripts too. Otherwise A3* would be useless in proofs.

The third argument is similar to the second one. It goes as follows: certain English text (namely well-known *Contrary-to-Duty Paradox*) is apparently consistent. However, its (most plausible) formalisation in SDL immediately leads to contradiction. Deontic relativism enables us to account for this problem in a simple and straightforward way.

Roderick Chisholm introduced so-called *contrary-to-duty (CTD)* imperatives as "imperatives telling us what we ought to do if we neglect certain of our duties" (Chisholm 1963, 33).[6] The problem with CTD obligations can be set forth as an argument of the following form:

(P1) Sophie shall not kill.
(P2) It ought to be that if Sophie does not kill, she is not punished for killing.
(P3) If Sophie kills, she ought to be punished for killing.
(P4) Sophie kills.

The text consisting of (P1)-(P4) is obviously consistent. However, its most plausible formalisation in SDL is inconsistent:

(P1') $O\neg A$

(P2') $O(\neg A \rightarrow \neg B)$

(P3') $A \rightarrow OB$

(P4') A

6 Throughout this paper, I will ignore the difference between descriptive and declarative (modes of) deontic sentences. While the distinction constitutes an interesting and widely discussed problem for deontic logic, it does not affect my arguments.

1. OB R1, P3', P4'
2. $O(\neg A \rightarrow \neg B) \rightarrow (O\neg A \rightarrow O\neg B)$ A2
3. $O\neg A \rightarrow O\neg B$ R1, P2', 2
4. $O\neg B$ R1, P1', 3
5. $OB \rightarrow \neg O\neg B$ A3
6. $\neg O\neg B$ R1, 1, 5
7. $\bot$ A1, 4, 6

We can solve CTD problem via deontic relativism treating primary and secondary subsystems of certain normative systems as different normative systems. Subsequently, we acquire relativistic version of our argument:

(P1*) $O_n\neg A$

(P2*) $O_n(\neg A \rightarrow \neg B)$

(P3*) $A \rightarrow O_mB$

(P4*) A

1. O_mB R1*, P3*, P4*
2. $O_n (\neg A \rightarrow \neg B) \rightarrow (O_n\neg A \rightarrow O_n\neg B)$ A2*
3. $O_n\neg A \rightarrow O_n\neg B$ R1*, P2*, 2.
4. $O_n\neg B$ R1*, P1*, 3.
5. $O_mB \rightarrow \neg O_m\neg B$ A3*
6. $\neg O_m\neg B$ R1*, 1., 5.
7. $O_n\neg B \rightarrow \neg O_nB$ A3*
8. $\neg O_nB$ R1*, 4., 7.

Inconsistency is thus avoided, for the set $\{O_n\neg B, \neg O_m\neg B, \neg O_nB, O_mB\}$ is consistent. Therefore, deontic relativism can solve the CTD paradox. However, this is clearly not the only possible solution to the CTD paradox (cf. Goble 2013). Nevertheless, it illustrates the usefulness of deontic relativism.

Finally, let us consider the fourth argument in favour of deontic relativism. This argument takes its inspiration from Lou Goble, though his aims are different from ours. Obviously, the opponent of deontic relativism can still reject the axiom scheme A3 $O\varphi \rightarrow \neg O\neg\varphi$. He can thus avoid the derivation of explicit contradiction from the conflict of obligation. Yet, he has another problem, namely *deontic explosion*: the formula $(O\varphi \wedge O\neg\varphi) \rightarrow O\psi$ is still valid. Therefore, as Goble (2000, 114) puts it, "if there is any conflict of obligation, then everything is obligatory." We can give an axiomatic proof of that proposition (in SDL):

1.	$(\varphi \wedge \neg\varphi) \rightarrow \psi$	A1
2.	$O((\varphi \wedge \neg\varphi) \rightarrow \psi)$	R2, 1
3.	$O(\varphi \wedge \neg\varphi) \rightarrow O\psi$	A2, 2, R1
4.	$\varphi \rightarrow (\neg\varphi \rightarrow (\varphi \wedge \neg\varphi))$	A1
5.	$O(\varphi \rightarrow (\neg\varphi \rightarrow (\varphi \wedge \neg\varphi)))$	R2, 4
6.	$O\varphi \rightarrow O(\neg\varphi \rightarrow (\varphi \wedge \neg\varphi))$	A2, R1, 5
7.	$O(\neg\varphi \rightarrow (\varphi \wedge \neg\varphi)) \rightarrow (O\neg\varphi \rightarrow O(\varphi \wedge \neg\varphi))$	A2
8.	$O\varphi \rightarrow (O\neg\varphi \rightarrow O(\varphi \wedge \neg\varphi))$	A1, R1, 6, 7
9.	$(O\varphi \wedge O\neg\varphi) \rightarrow O(\varphi \wedge \neg\varphi)$	A1, R1, 8
10.	$(O\varphi \wedge O\neg\varphi) \rightarrow O\psi$	A1, R1, 3, 9

Goble (2000, 113) claims that any logic, which contains all of

(a) $\vdash (\varphi \wedge \neg\varphi) \rightarrow \psi$,

(b) if $\vdash \varphi \rightarrow \psi$, then $\vdash O\varphi \rightarrow O\psi$ and

(c) $\vdash (O\varphi \wedge O\psi) \rightarrow O(\varphi \wedge \psi)$

will necessarily contain

(d) $(O\varphi \wedge O\neg\varphi) \rightarrow O\psi$.

Suppose that we are in a situation where a conflict of obligation comes to play: the case of Antigone, the CTD paradox or some real-world moral dilemma. Furthermore, suppose we reject deontic relativism as well as the axiom scheme A3. The derivation of explicit contradiction from conflict of obligation is thus avoided. Yet we derive $O\psi$ for any formula ψ whatsoever. This result is obviously counterintuitive and poses a problem for the opponent of deontic relativism. One possible solution is to repudiate one of (a)-(c). Another one is to adopt deontic relativism, since this does not pose a problem for deontic relativist: The theorem (d) is still valid. Yet we cannot use it, since all we have is a formula of the form $O_n\varphi \land O_m\neg\varphi$, which does not constitute a genuine conflict of obligation (i.e. it is not the formula of the form $O_x\varphi \land O_x\neg\varphi$).

2. Δ-TIL and its minimal semantics

Δ-TIL is a part of the system of TIL. For this reason, there is a need to introduce TIL briefly. Furthermore, there is a need to explain semantic minimalism of TIL.

The first comprehensive account of TIL was provided by Pavel Tichý in *The Foundations of Frege's Logic*. TIL is a hyperintensional partial lambda calculus with types. It is the logic of *constructions*. Construction is a hyperintensional, structured entity, a theoretical explicate of the notion of *meaning*. TIL employs six different kinds of constructions, most important among them are variables, trivialisation, composition, and closure. Tichý devised an objectual analysis of *variables* (so variables are understood as full-fledged objects). A variable is a construction that constructs an object with respect to some valuation; notation $w, t, x, y\ldots$ (possibly with subscripts). *Trivialisation* is a simple construction which picks out an object and returns the very same object; notation ^0X. *Composition* is a construction that applies a function to some arguments and returns the value of this function on the given arguments (if there is such a value); notation $[X\ Y_1\ldots Y_n]$. Composition has its syntactic surrogate in lambda calculus, namely application. *Closure* is a construction that construes a function by abstraction; notation $[\lambda x_1\ldots x_n\ Y]$. Closure has its syntactic surrogate in lambda calculus too, namely lambda abstraction.

It is quite common in TIL to use four basic types: ο for two truth-values, ι for individuals, τ for moments of times (or real numbers), and ω for possible worlds. Constructions have higher order atomic types $*_n$ ($n \in \mathbb{N}$). These atomic types are the building blocks, and all mathematically possible functions are built upon them (as is quite common in lambda calculus). For instance, proposition is a function from world courses to truth-values, i.e. it has a type ((οτ)ω), in an abbreviated form $ο_{τω}$; property of individuals has a type $οι_{τω}$; set of propositions has a type $(ο(ο_{τω}))$ and so on. Constructions of propositions (propositional constructions) are theoretical explicates of philosophically important notion of *structured proposition*.

The semantics of TIL is in accordance with semantic minimalism. According to semantic minimalism, as Borg (2009) puts it, "syntax provides the sole route to semantic content." Yet there are two characteristic versions of semantic minimalism (see Zouhar 2012, 708-713).[7] According to the first version of semantic minimalism, "[t]he semantic content of a sentence S is the content that all utterances of S share. It is the content that all the utterances of S express no matter how different their contexts of use are" (Cappelen & Lepore 2005, 143). According to the second version of semantic minimalism,

> literal meaning is held to be entirely context-invariant – a sentence, individuated in terms of its syntax, possesses the very same meaning no matter when, where, or by whom it is produced. (…) [A]s far as semantics is concerned, we should (…) concentrate just on the meaning of sentence-types as formal objects of study. (Borg 2004, 215)

This second version of semantic minimalism is closer to the semantics of TIL. However, these issues have not been extensively discussed yet. Sufficient examination of background syntactic theory of TIL is needed for a sufficient examination of minimal semantics of TIL. In any case, background syntactic theory of TIL needs to distinguish between surface structure and deep structure (or logical form) of expressions. This is so because the constructions of TIL involve modal and temporal variables, despite the fact that such variables are not present in surface structure of (empirical) expressions, but only in deep structure. It is this deep structure what constitutes the relevant basis for semantic analysis. Moreover, the semantic analysis is context invariant. Surely, one needs context to find the *intended* meaning, yet one does not need context to find the *literal* meaning. The apparent semantic function of context is explained away by *ambiguity*. Finally, since TIL has minimal semantics, Δ-TIL (as a part of TIL) has to adopt semantic minimalism too.

Let me now briefly introduce Δ-TIL. Deontic operators *O*, *P* and *F* stand for *implicit deontic modalities*. Implicit deontic modality is a function from world courses (i.e. a function from possible worlds to function from moments of time) to sets of propositions. Deontic operators O^*, P^* and F^* stand for *explicit deontic modalities*. Explicit deontic modality is a function from world courses to sets of constructions.

---

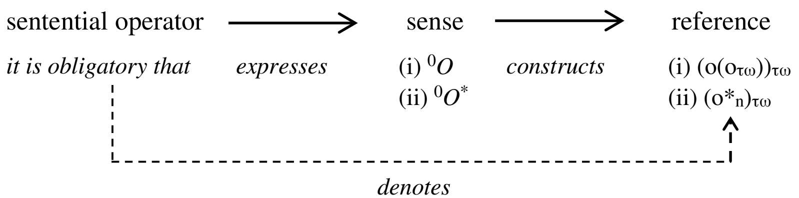
7 For further discussion of semantic minimalism (and its competitors), cf. Zouhar (2011).

Δ-TIL assumes that a sentence of the form "*A* is obligatory" is true *simpliciter*, as one can see from the truth-conditions stated below. Let ${}^0T$ constructs the truth-value True, ${}^0F$ constructs the truth-value False and let *C* be a construction of a proposition. We write α : β if and only if (iff) α construes the same object as β (with respect to some valuation).[8] The truth-conditions of formulas involving *O* and O^* are then as follows:[9]

${}^0T : [{}^0O_{wt}\ C]$ iff $C \in O_{wt}$
${}^0F : [{}^0O_{wt}\ C]$ otherwise.

${}^0T : [{}^0O^*_{wt}\ {}^0C]$ iff ${}^0C \in O^*_{wt}$
${}^0F : [{}^0O^*_{wt}\ {}^0C]$ otherwise.

Similarly, for *P* and P^*. The truth-conditions of formulas involving *F* and F^* are defined in the standard way via *O* and O^* (i.e., something is forbidden iff its negation is obligatory). The following schema represents the analysis of the expression "it is obligatory, that":



Note that when someone asserts that *It is obligatory that A*, it is ambiguous. The analysis (i), but also the analysis (ii) is correct. Surely, one can consequently ask which one is the *preferred* analysis. There is a way to answer such a question: namely, by answering additional question, whether the individual in question is talking about explicitly formulated obligations of about implicit consequences of some explicitly formulated obligations. Yet this is a step beyond the realm of semantics.

---

8 This definition employs the notion of *match*, introduced by Pavel Tichý; see Tichý (1982, 64-65).

9 We will use an arbitrary construction *X* with lower-case „*wt*" in a standard way as an abbreviation for [[*X w*] *t*].

3. The problem

The problem can be stated this way: How to amend the analysis to be both in line with deontic relativism and semantic minimalism? There are at least two possible solutions:

> we might seek to complicate the syntax of natural language sentences, positing a range of 'hidden indexicals' which provide the syntactic triggers for the additional context-sensitivity (…) [or] introduce additional complexity into the way in which sentences map to truth-conditions, holding that the context-sensitivity (…) lies within the circumstances of evaluation, not in a truly indexical content for sentences. (Borg 2009, 424)

The former is characteristic of indexicalism. However, we want the analysis to accommodate semantic minimalism (recall that Δ-TIL is a part of TIL, so it should be consistent with TIL). Hence, we will consider *minimal indexicalism*[10] rather than mere indexicalism. What does it mean for Δ-TIL? A free variable ranging over normative systems would occur in deontic construction. Therefore, this construction would be open and we would need the process of completion (saturation – see Bach 1994) for obtaining a closed construction. Normative systems would thus belong to the *context of use*[11] and expressions denoting them would function just like indexicals. If normative systems belong to the context of use, one needs to specify them to determine *what has been said*.

The latter option is characteristic of (non-indexical) relativism. In this case, a lambda-bound variable ranging over normative systems would occur in deontic construction. Therefore, we gain closed construction, but it would not be a propositional construction anymore, since additional parameter for normative systems would be present in its type. Normative systems would thus belong to the *circumstances of evaluation* and would function just like possible worlds (and moments of time). If normative systems belong to the

---

[10] Minimal indexicalism was introduced by Marián Zouhar. For the most comprehensive account, see Zouhar (2011).

[11] The distinction between context of use and circumstances of evaluation was introduced in Kaplan (1989) (written already in 1977). For further discussion, see Zouhar (2013a).

circumstances of evaluation, one needs to specify them to determine *truth-values* of deontic sentences in question.

4. The two possible solutions

To begin with, let us look at a more detailed version of the previous analysis. The deontic sentence "C is obligatory" was supposed to represent either (i) the implicit deontic construction $[\lambda w \lambda t\ [^0O_{wt}\ C]]$ or (ii) the explicit deontic construction $[\lambda w \lambda t\ [^0O^*_{wt}\ ^0C]]$. The construction $[\lambda w \lambda t\ [^0O_{wt}\ C]]$ is an abbreviation for $[\lambda w \lambda t\ [[[^0O\ w]\ t]\ C]]$; the construction $[\lambda w \lambda t\ [^0O^*_{wt}\ ^0C]]$ abbreviates $[\lambda w \lambda t\ [[[^0O^*\ w]\ t]\ ^0C]]$.

We need to add variables for normative systems. This leads to a problem: What is the proper type of normative systems? This remains, however, an open question. One option is to add a further atomic type to the basis. However, for the purposes of this paper, it will suffice to analyse them simply as individuals (note that even individual authorities such as parents, teachers, emperors etc. can be integrated into this framework). Therefore, variables for normative systems will be *individual variables*, so they will construe individuals, in technical notation $n \rightarrow_v \iota$ (we read this as "*n v*-construes an individual"). The operator O will represent a function from individuals to properties of propositions, in technical notation $O/(o o_{\tau\omega})_{\tau\omega\iota}$. The operator O^* will represent a function from individuals to properties of propositional constructions, in technical notation $O^*/(o^*_n)_{\tau\omega\iota}$. Remaining types are $w \rightarrow_v \omega$ (*w v*-constructs a possible world, i.e. w is a possible-world variable) and $t \rightarrow_v \tau$ (*t v*-constructs a moment of time, i.e. t is a time-moment variable).

4.1. Minimal indexicalism

Δ-TIL combined with minimal indexicalism offers the first possible solution. The analysis of some $O\varphi$-sentence will be

(1.1) implicit deontic construction $[\lambda w \lambda t\ [[[[^0O\ n]\ w]\ t]\ C]]$ or

(1.2) explicit deontic construction $[\lambda w \lambda t\ [[[[^0O^*\ n]\ w]\ t]\ ^0C]]$.

As we can see, the constructions in (1.1) and (1.2) are open, since they contain a free variable *n*. The evaluation (saturation) of this variable is needed to acquire a propositional construction.

4.2. Non-indexical relativism

Δ-TIL combined with non-indexical relativism offers the second possible solution. The analysis of some $O\varphi$-sentence will be

(2.1) implicit deontic construction $[\lambda n \lambda w \lambda t[[[[{}^0O\ n]\ w]\ t]\ C]]$ or

(2.2) explicit deontic construction $[\lambda n \lambda w \lambda t[[[[{}^0O^*\ n]\ w]\ t]\ {}^0C]]$.

As we can see, constructions in (2.1) and (2.2) are closed (because all variables are bound by lambda abstractors). Note that as a straightforward consequence of this analysis, constructions in (2.1) and (2.2) are no longer propositional constructions, and subsequently, deontic sentences do not denote propositions (in the standard sense) anymore.

5. Concluding remarks

Either way, the evaluation of *n* is needed for a truth-evaluation of certain $O\varphi$-sentence. Is it plausible? Let us recall our examples from the first section. We might (reasonably) hesitate to answer a question such as "Is φ obligatory?", since for a truth-evaluation of the sentence of the form $O\varphi$, we need to check the normative system in question. If no normative system is given at all, we do not know what to check. This result is in perfect accordance with the above presented analysis.

Finally, we may ask which of the competing options is better. Since in deontic sentences there is no explicit reference to normative systems (exactly as no explicit reference to possible worlds and moments of time), the second option seems more plausible. Yet the first option is feasible too. Further research is needed to examine them.

6. Response to possible objections

This section will anticipate some possible objections to the analysis and respond to them. To begin with, one can accept deontic relativism proposed in section 1 without thereby accepting the version of deontic relativism proposed in section 4. Certainly, there are alternative theories designed to account for the problems described in section 1: namely contextualism, ambiguity theory, subjectivism and objectivism; cf. MacFarlane (2014, 280-285).

To put it simply: Contextualists claim that there is just one word "obligatory", but since this word is context-sensitive, different contexts assign different meanings to this word. Ambiguity theory claims that there are many words "obligatory_1", "obligatory_2", …, with different meanings corresponding to them. Subjectivism claims that "obligatory" is relative to the normative system the speaker has in mind. Objectivism claims that "obligatory" is relative to the most general (common, important…) system of norms.

Yet none of them is able to explain disagreements.[12] Let *M* represent the sentence "The Maori children learn the names of their ancestors". Sophie asserts that *OM* is true, whilst Pavel asserts that *OM* is false. As regards the sentence *OM*, they are in a disagreement. How is it possible? Easily: Sophie thinks of the tribal laws Maoris have and Pavel thinks of the official law in New Zealand, however, they are talking about the same sentence with the same meaning. It is this sentence (and its meaning) what is the subject matter of their disagreement.

According to contextualism, they use the same word "obligatory", but contexts assign different meanings to this single word. They are both right, they assert the same sentence, but with different meanings. There is thus no disagreement. The same holds for subjectivism. According to ambiguity theory, they use different words (with different meanings). Again, they are both right. Yet they assert different sentences (with different meanings). Hence, there is again no disagreement. According to objectivism, they use the same word "obligatory", which is relative to the "universal law" (whatever it is). Yet, it is problematic to say what this so-called universal law is supposed to be.

Can deontic relativism solve the problem of disagreements? That is beyond doubt, since Sophie and Pavel are in a disagreement about a certain sentence with certain meaning. However, the meaning of this sentence needs evaluation (or saturation) of deontic variables for a truth-evaluation of the sentence. In our case, such evaluation (or saturation) will reveal the fact that Sophie and Pavel were thinking about different normative systems. Strictly speaking, they can be both right, because (free or lambda-bound) variables for normative systems will be evaluated differently. This does not necessarily mean the end of disagreement: Sophie and Pavel can still disagree about the preferred normative system.

---

[12] The argument takes its inspiration from MacFarlane (2014, 280-285) and Kratzer (1977, 338).

Moreover, there are further disadvantages of alternative theories. Firstly, contextualism is worse than minimal indexicalism for methodological reasons (see Zouhar 2013b). Secondly, ambiguity theory causes annoying profusion of "oughts"; (see Jackson 1991, 471; and MacFarlane 2014, 284). Similarly, contextualism causes annoying profusion of oughts. Finally, the objective sense of "obligatory" is too general. We usually use this word in talking about different legal systems, moral codes or tribal laws and so on. We can thus conclude that "obligatory" we are using in natural language is not the objective one.

Furthermore, one can accept deontic relativism without thereby accepting semantic minimalism. It is not the purpose of the present paper to criticize all the alternatives to semantic minimalism. Rather we outline some positive reasons, inspired by Cappelen & Lepore (2005, 151-154). Firstly, semantic minimalism does not end up requiring that semanticists do metaphysics. Fortunately, minimalism does not require of semantics to answer the question "What is obligation?" which is far beyond the borders of semantics. Secondly,

> [it] can account for how the same content can be expressed, claimed, asserted, questioned, investigated, etc. in radically different contexts. It is this content that enables audiences who find themselves in radically different contexts to understand each other, to agree or disagree, to question and debate with each other. (Cappelen & Lepore 2005, 152)

Moreover, it "can account for how Inter-Contextual Disquotational Indirect Reports can be true where the reporter and the reportee find themselves in radically different context…" (Cappelen & Lepore 2005, 152). Suppose the speaker S utters the sentence "A is obligatory". We can (truly) utter the sentence "S said that A is obligatory." Semantic minimalism can explain this fact, since it admits certain common content – in particular, the (minimal) semantic content of the sentence "A is obligatory".

Finally, one can accept deontic relativism without thereby accepting that deontic operator is relative to *a particular* (explicitly unspecified) normative system. We claim that a formula of the form $O\varphi$ means that φ is obligatory under *a particular* normative system. However, it seems that many other quantifiers can be employed. Let us discuss at least the applicability of existential and universal quantifiers. Hence, $O\varphi$ can mean (E) φ is obligatory under *some* normative systems or (A) φ is obligatory under *all* normative systems.

We can make use of the example discussed in section 1. Again, let *P* represent the sentence “Some men have more than one wife at a time”. Consider the following situations:

> **The case of polygamous Saudi Arabia:** Mr. Fiable, an inhabitant of France, is in Saudi Arabia, asking one of its inhabitants, Mr. Amin, whether *O*¬*P* is true or not. Mr. Amin is a reliable source of information about the legal system of Saudi Arabia. He supposes that Mr. Fiable’s question concerns the legal system of Saudi Arabia and replies that *O*¬*P* is false. From now on, Mr. Fiable will (truly) think that the legal system of Saudi Arabia permits polygamy. However, the option (E) claims that *O*¬*P* is true because there is at least one normative system, which forbids *P* (e.g. the legal system of France). Yet it seems that Mr. Amin was right in claiming the opposite.
>
> **The case of monogamous France:** Mr. Amin is in France, asking Mr. Fiable whether *O*¬*P* is true or not. Mr. Fiable is a reliable source of information about the French legal system. He supposes that Mr. Amin’s question concerns the French legal system and replies that *O*¬*P* is true. From now on, Mr. Amin will (truly) think that the French legal system forbids polygamy. However, an option (A) claims that *O*¬*P* is false, because there is at least one normative system, which permits *P* (e.g. the legal system of Saudi Arabia). Yet it seems that Mr. Fiable was right in claiming the opposite.

The answer “*O*¬*P* is false” was expected in the case of polygamous Saudi Arabia. Yet according to (E), *O*¬*P* is true, since there is at least one normative system, which forbids *P*. Hence, the option (E) gives a wrong prediction. Moreover, the answer “*O*¬*P* is true” was expected in the case of monogamous France. Yet according to (A), *O*¬*P* is false, since there is at least one normative system, which permits *P*. Hence, the option (A) gives a wrong prediction.

Certainly the argument demonstrates only the *insufficiency* of analysing *Oφ*-sentences in a fashion suggested by (E) or (A). This result is sufficient for the present purposes. Note, however, that it does not demonstrate their uselessness. We can employ such quantifiers when needed. For instance, we can use them to analyse sentences such as “It ought to be the case that φ under some system of norms” and “It ought to be the case that φ under any system of norms”. Furthermore, we can use restricted quantifiers.

References

BACH, K. (1994): Conversational Impliciture. *Mind and Language* 9, No. 2, 124-162.

BORG, E. (2004): Formal Semantics and Intentional States. *Analysis* 64, 215-223.

BORG, E. (2009): Semantic Minimalism. In: Cummings, L. (ed.): *The Pragmatics Encyclopedia*. Routledge, 423-425.

CAPPELEN, H. & LEPORE, E. (2005): *Insensitive Semantics: A Defence of Semantic Minimalism and Speech Act Pluralism*. Oxford: Blackwell Publishing.

DUŽÍ, M., JESPERSEN, B. & MATERNA, P. (2010): *Procedural Semantics for Hyperintensional Logic. Foundations and Applications of Transparent Intensional Logic*. Berlin: Springer.

DUŽÍ, M. & MATERNA, P. (2012): *TIL jako procedurální logika (Průvodce zvídavého čtenáře Transparentní intensionální logikou).* [*TIL as a Procedural Logic.*] Bratislava: aleph.

DRETSKE, F. (1970): Epistemic Operators. *Journal of Philosophy* 67, No. 24, 1007-1023.

GLAVANIČOVÁ, D. (2015a): K analýze deontických modalít v Transparentnej intenzionálnej logike. [On the Analysis of Deontic Modalities in Transparent Intensional Logic.] *Organon F* 22, No. 2, 211-228.

GLAVANIČOVÁ, D. (2015b): Implicitné a explicitné deontické modality. [Implicit and Explicit Deontic Modalities.] In: Brzobohatá, K. (ed.): *Logika pro právní praxi*. Brno: Masarykova univerzita, 36-45.

GOBLE, L. (2000): Multiplex Semantics for Deontic Logic. *Nordic Journal of Philosophical Logic 5*, No. 2, 113-134.

GOBLE, L. (2013): Prima Facie Norms, Normative Conflicts, and Dilemmas. In: Gabbay, D. M. et al. (eds.): *Handbook of Deontic Logic and Normative Systems*. College Publications.

CHISHOLM, R. M. (1963): Contrary-to-Duty Imperatives and Deontic Logic. *Analysis* 23, 33-36.

JACKSON, F. (1991): Decision-Theoretic Consequentialism and the Nearest and Dearest Objection. *Ethics* 101, No. 3, 461-482.

KAPLAN, D. (1989): Demonstratives. In: Almog, J., Perry, J. & Wettstein, H. (eds.): *Themes from Kaplan*. Oxford: Oxford University Press, 565-614.

KRATZER, A. (1977): What 'Must' and 'Can' Must and Can Mean. *Linguistics and Philosophy* 1, No. 3, 337-355.

MACFARLANE, J. (2014). *Assessment Sensitivity: Relative Truth and its Applications*. Oxford: Clarendon Press.

MCNAMARA, P. (2006): Deontic Logic. In: Gabbay, D. M. & Woods, J. (eds.): *Handbook of the History of Logic*: *Logic and the Modalities in the Twentieth Century*. Vol. 7. Amsterdam: Elsevier, 197-288.

RACLAVSKÝ, J. (2009): *Jména a deskripce: logicko-sémantická zkoumání*. [*Names and Descriptions: Logico-Semantical Investigations*.] Olomouc: Nakladatelství Olomouc.

TICHÝ, P. (1982): Foundations of Partial Type Theory. *Reports on Mathematical Logic* 14, 59-72.

TICHÝ, P. (1988): *The Foundations of Frege's Logic*. Berlin – New York: de Gruyter.

ZOUHAR, M. (2011): *Význam v kontexte*. [*Meaning in Context.*] Bratislava: aleph.

ZOUHAR, M. (2012): Indexikalizmus, sémantický minimalizmus a pravdivostný obsah. [Indexicalism, Semantic Minimalism, and Truth-Conditional Content.] *Filozofia* 67, No. 9, 705-717.

ZOUHAR, M. (2013a): Epistemický kontextualizmus a jeho motivácia. [Epistemic Contextualism and Its Motivation.] *Organon F* 20, Suppl. issue 2, 171-186.

ZOUHAR, M. (2013b): Sémanticky relevantné pragmatické procesy (2): Testy a metodologické aspekty. [Semantically Relevant Pragmatic Processes (2): Tests and Methodological Aspects.] *Filozofia* 68, No. 4, 296-308.

Organon F 23 (2) 2016: 222-242

Common Source of the Paradoxes of Inference and Analysis

KAREL ŠEBELA

Department of Philosophy. Faculty of Arts. Palacký University in Olomouc
Křížkovského 12. 771 80 Olomouc. Czech Republic
karel.sebela@upol.cz

RECEIVED: 03-08-2015 ▪ ACCEPTED: 07-12-2015

ABSTRACT: The paper deals with the paradoxes of inference and analysis. It attempts to show what is specific about these paradoxes. They have got a lot in common. Often, they are not considered paradoxes in the strict sense at all. Moreover, they both raise the same problem: How can the requirements of correctness and informativeness be both met for inference and for conceptual analysis? The strategies developed to address the problem are similar for both cases. In the paper, I claim that the paradoxes have common origins. This claim is supported by comparing different strategies adopted to resolve the problem. Regarding their origins, both paradoxes share the epistemological framework that is grounded in Aristotle's theory of science. This is related to the problem of implicit knowledge, which is a variation on a dilemma formulated by Plato in his *Meno*. Aristotle's solution to the dilemma of *Meno* is discussed and considered as another plausible strategy for dealing with the paradoxes of inference and analysis.

KEYWORDS: Aristotle – pre-knowledge – the paradox of analysis – the paradox of inference.

0. Introduction

This paper focuses on two remarkable paradoxes related to the subject of rational cognition, namely the paradoxes of inference and analysis. The objective is to investigate the nature of these paradoxes and to somehow resolve them. The structure of the paper is as follows: first, I briefly introduce both paradoxes and, by comparing them, I come to their common source, pre-existent knowledge. After that, I outline a very interesting conception of this subject as offered by Aristotle in *Analytics*. Then I apply this concept to both paradoxes and, thus, offer it as their common solution. To keep it simple, I try to present all paradoxes in the simplest form possible, *i.e.* in the form of a simple syllogism.

1. Introducing the paradoxes

1.1. Paradox of inference[1]

To keep it simple, this paradox can be presented in the form of the following argument:

Valid inference from true premises is a good tool
for expanding knowledge.
Valid inference does not provide any new knowledge.

Some good tools for expanding knowledge do not
provide any new knowledge.

Commentary on individual parts of the argument:

The first premise ("Valid inference from true premises is a good tool for expanding knowledge"): Many philosophers believe that the tools or means for how we get to know the world around us are our senses and reason. For now, we can leave the manner, competence, and mutual relationship between both means of knowledge aside; the only important matter is that inference definitely belongs among the methods that reason – a tool of knowledge – "works with". Thus, what we assign to rational knowledge as a whole is also relevant for inference. This opinion is also for many people the reason why philosophers (or scientists) should – to some extent – master the discipline which, above all, applies inference, *i.e.* logic. The first premise is, therefore, a compact

---

1 The origin of the paradox definitely goes back to Classical Antiquity, its modern form and name were coined by Cohen & Nagel (1934, 173).

expression of this conviction (the word “good” should emphasize that it is a functioning, not a damaged tool).

The second premise (“Valid inference does not provide any new knowledge”): It contains a finding that many philosophers and logicians come to when facing the question of what generally justifies a specific conclusion from given premises. For instance, let us consider the popular syllogism “all men are mortal; Socrates is a man; thus, he is also mortal”. We often face the opinion that a conclusion is based on premises because it is in them somehow – implicitly – contained. In our case, Socrates’ mortality is given by the fact that as a man he belongs to creatures that the first premise mentions. If the conclusion is contained in the premises before its actual inference, then it seems that the explicit statement of the conclusion cannot provide any new information which would not have already been in the premises, *ergo*, inference does not provide any new knowledge. In this sense, old sceptics – or more recently, J. S. Mill (see Mill 1882, 228) – criticised inference. We can come to the same conclusion by considering the famous Deduction Theorem. It states that every deductively valid argument can be transformed into tautology (in the form of implication where the conjunction of premises forms an antecedent and the conclusion a consequent of implication). In a widely accepted understanding of tautology, it does not provide any new information about the world. If it is therefore possible to equivalently transform deductively valid inferences to tautologies and tautologies do not provide any new knowledge, then also the deductively valid arguments do not provide any new knowledge; *ergo*, inference does not provide any new knowledge.

The conclusion (“Some good tools for expanding knowledge do not provide any new knowledge”): It is correctly inferred from the given premises, specifically it is the syllogistic mood Felapton. It is also seemingly strange; analogically we could, for instance, say that a good tool to hammer nails does not hammer nails. It either is not a good tool for hammering nails and then we should not take it for one or it makes sense and then it can crack down on those nails, despite our opinion – and yet, both simultaneously are impossible.

We, consequently, have premises here which are at least in some philosophical and logical communities quite commonly used, but together they come to an absurd conclusion. A false conclusion, correctly inferred from premises, clearly indicates that (at least) one premise is false; at least one of the given, rather widely accepted opinions is, thus, false.

Given that the second premise seems to contain hardly questionable results of logic, the more common strategy of how to contest the inference paradox is by questioning the truth of the first premise (see the aforementioned ancient sceptics or J. S. Mill). Consequently, we have to revise the opinion regarding the role of inference in human cognition.

The less common strategy is questioning the second premise, *i.e.* questioning the belief that a conclusion does not provide new information. That the conclusion is "somehow" contained in the premises cannot be challenged, therefore they mostly use redefinition or distinction of the terms novelty and/or information.[2]

It is worth remembering that for modern logics the *locus classicus* of the inference paradox is the publication by Cohen and Nagel, *An Introduction to Logic and Scientific Method* from 1934 (see Cohen & Nagel 1934). It defines the inference paradox in the following way:

> If in an inference the conclusion is not contained in the premises, it cannot be valid; and if the conclusion is not different from the premises, it is useless; but the conclusion cannot be contained in the premises and also possess novelty; hence inferences cannot be both valid and useful. (Cohen & Nagel 1934, 173)

Cohen and Nagel thus understand it as a dilemma between validity and usefulness. The concept of usefulness is worth considering – a valid inference is seen useless because it does not provide any new information. Thus, the criterion for usefulness of an inference is that it provides new information. We will see later that Aristotle approached it differently.

1.2. Paradox of analysis

This paradox is most often formulated as a dilemma when analysis is said to be either correct or informative, but never both. Thus, dilemma is surprising because the possibility of a correct and at the same time informative analysis is often considered unproblematic. If converted into a simple argument, we could present it in the following way:

---

[2] See, e.g. Duží (2006), or Duží (2010). Needless to say that the redefinition or distinction is not the whole solution, but only a part of the more complex argumentation.

Correct analysis is a good tool for expanding knowledge.
Correct analysis does not provide any new knowledge.

Some good tools for expanding knowledge do not provide any new knowledge.

Commentary on the individual parts of the argument:

The first premise ("Correct analysis is a good tool for expanding knowledge"): Here almost the same applies as in the commentary on the first premise of the inference paradox: analysis, a mental decomposition or a break-down of a given compound into its constituent components, is understood as one way of how reason – a tool of knowledge – "works". If it holds true that reason is really a good tool for getting to know the world, then the same holds for analysis.

The second premise ("Correct analysis does not provide any new knowledge"): The specific problem here was established by the British philosopher, G. E. Moore (originally in Moore 1903, 7) and it is, let us say, of a semantic nature. It states that if we have, for instance, an analysed concept (an analysandum) and an analysing concept(s) (or analysans), then the basic requirement of a correct analysis is that all that the analysandum contains must be also contained in the analysans and vice versa, thus, analysandum = analysans. If the analysans contained something that was not in the analysandum, it would have been an incorrect analysis. Taking an analogical example from the analysis of physical things, a traditional component of military training was to disassemble a soldier's machine gun down to its components and to reassemble it again. If the soldier conducted the disassembly correctly, then he dismantled the machine gun into and only into the parts that it was composed of – if any additional component appeared, it could not have come from the dismantled machine gun and the disassembly had not been performed correctly (or alternatively, it was a disassembly of the machine gun and something else). If the analysans cannot contain what was not in the analysandum to begin with, then no correct analysis can come up with something new, *ergo*, analysis does not provide any new knowledge.

The conclusion ("Some good tools for expanding knowledge do not provide any new knowledge"): It is completely analogical to the paradox of inference with the modification that the "some" in each paradox targets a different tool for expanding knowledge. Solution strategies are also similar:

The more common strategy is to oppose the first premise, *i.e.,* in some sense revise the view on the role of analysis in human knowledge. This revision – seemingly absurdly – appears in analytical philosophy, that is, a philosophical stream which has analysis in its very title and which focuses its very philosophical work on rigorous analysis. However, the absurdity is only illusory. Analytical philosophers, who reject the first premise of the paradox of analysis, do not wish to claim that the result of their analytical work is new information about the world; they do not wish to compete with sciences. In other words, philosophy – in their understanding – is not a theory but an activity.[3]

Another used strategy is to question the second premise. In this case, it usually means specification of what exactly the equal sign between the analysandum and the analysans relates to: whether to language expressions, meanings, or (non-language) objects depicted by these expressions.

1.3. The two paradoxes compared

The preceding text should anticipate the similarity of both paradoxes. For better illustration, let us present them again, but this time together:

Valid inference from true premises is a good tool
for expanding knowledge.
Valid inference does not provide any new knowledge.

Some good tools for expanding knowledge do not
provide any new knowledge.

Correct analysis is a good tool for expanding knowledge.
Correct analysis does not provide any new knowledge.

Some good tools for expanding knowledge do not provide
any new knowledge.

The similarity of both paradoxes should definitely not be only about the possibility to convert them into almost identical syllogisms. On the contrary, that we can present them in this way is but one indication of their similarity. These similarities are more numerous and deeper:

[3] See Wittgenstein (2001, 4.112): “Philosophy is not a body of doctrine but an activity”.

Above all – both are called paradoxes even though they are clearly different from typical and more famous paradoxes – such as the liar paradox, Russell's paradox, etc. In my view, the main difference lies in the fact that the aforementioned, more typical paradoxes are based on premises or methods, where we cannot easily determine where and if there is any problem at all. They are paradoxes for the very reason that the drawing the absurd consequences from seemingly unproblematic premises is striking; it is not clear what is wrong. Our paradoxes often constitute some logical-noetic intuitions shared by a community but always rejected by many other experts. These statements are not unproblematic and their rejection does not seem that shocking or fatal. They are paradoxes almost literally – the term "paradox" refers to a situation, when various *doxai*, *i.e.*, opinions or intuitions, go "against each other". In our case it is the contradiction of various intuitions related to the role of rational knowledge, thus, being incompatible, together forming an absurd result. Therefore, it concerns the divisions within the community of logicians and philosophers.

Furthermore – both paradoxes' first premise is always a general statement on the role of the relevant component of reason for knowledge, while the second premises capture a finding of a more logical or semantic nature.

Thirdly – the second premise always captures a logical-semantic piece of knowledge, *i.e.*, from a sphere that enjoys a relatively high level of authority. This fact gives rise to the attempts to solve the paradoxes by criticizing the first, more of a philosophical (specifically noetic) premise.

Fourthly – paradoxes form relatively simple arguments, where it is difficult to question the fact that the conclusions really follow from these premises, thus, the attempts to question the second premise most commonly use the method of concept distinction, specifically concepts of novelty, information or equality.

Fifthly – the most significant similarity is related to a problem which is difficult and is the proper subject matter of both paradoxes. I believe it stands behind the reason why so many contradictory beliefs can arise about the same thing – it is the question of implicit knowledge or pre-existent knowledge. For both paradoxes, this problem is hidden in the second premise. In the case of the paradox of inference as stated above, the conclusion follows from the premises for the very reason that it is somehow – implicitly – contained in them. Yet, what does this really mean? If the conclusion is "in some way" in the premises, then this could mean that every validly inferred argument incorporates a mistake of the so-called reasoning in a circle, which is a situation when the conclusion is one of the premises. Circular reasoning is a deductively valid argument, but is considered as faulty because it does not contain any new information compared to the premises, *i.e.*, is trivial. If the fact that the conclusion of every deductively valid argument is somehow contained in the premises means that the conclusion appears in the premises then every deductively valid argument is at the same time circular reasoning, trivial and uninformative. On the other hand, we have rich experience, especially with more complex arguments, arguments with more and/or complicated premises where the conclusion is really surprising and we are inclined to believe that the conclusion is new information for us. Thus, it seems that the situation when the conclusion is implicitly contained in the premises is not simply a situation where the conclusion is one of the premises as would be the case of circular reasoning. Information contained in the conclusion is somehow already present for us in the premises, but at the same time we know it in a somewhat different way than we know the premises.

In the case of the paradox of analysis, we can understand analysis generally as decomposition, as breakdown of a whole into its components. When decomposing a whole, we have to have this whole available, it has to be somehow given to us – as a whole with all its components. Analysis can be correct for the very reason that if we are decomposing a given whole and not something else, then what we get are parts of this whole and not something else. This should, however, mean that a correct analysis cannot provide anything new because the parts obtained by analysis had already been somewhat available when we had the whole thing in front of us prior to commencing the analysis. It is as with the soldier who disassembles his machine gun – before he starts the dismantling, he has a machine gun in front of him, *i.e.*, the whole with all its components. When he starts to disassemble the machine gun, then everything he gets, for instance lock frame or striker, he had already somehow at his disposal at the moment when he started the dismantling. On the other hand, analysis of more complex concepts or statements can really strike us, it might surprise us that some parts of the analysans belong to this very analysandum.

The situation is, therefore, similar to that of the paradox of inference – *on one hand the "result" (conclusion or analysans) is already somehow present in the given (premises or analysandum) and this very presence is the guarantee of correctness (of inference or analysis); on the other hand, the presence is such that its (re)emergence "in the result" can surprise us and is thus present only indirectly*. We somehow know the conclusion and analysans in advance, this *pre-existent knowledge* is a special type of knowledge which deserves special attention. An insight into the nature of this pre-existent knowledge can simultaneously provide a way how to better understand the challenges of both paradoxes or even more provide a solution.

2. Aristotle's theory of pre-existent knowledge

In the next part, I would like to offer a very interesting – and, as far as I can see, also plausible – theory of pre-existing knowledge that Aristotle presented in some of his logical treatises. I would like to defend the plausibility of Aristotle's concept by its very ability to solve the paradoxes of inference and analysis which, based on what was said above, I consider derived from the problem of pre-existing knowledge.

2.1. Plato

Aristotle "inherited" this problem like many others from Plato. Given that Plato's answer to this problem is his famous doctrine of ideas, we can claim that it could not have been a marginal issue for him. Thus, it is apt to further investigate how he handled it. It allows us to find out what, specifically, Aristotle responded to and in what form he took it over from Plato.

We can find a very concise introduction into the problem in Plato's dialogue *Meno*. Socrates discusses with Meno the nature of virtue and demonstrates to him the unsustainability of the definition of virtue as presented by Meno. Meno is at a loss as to how, and if at all, to continue. Socrates encourages him and maintains that we should continue to pursue the definition of virtue. Yet, Meno objects:

> Why, on what lines will you look, Socrates, for a thing of whose nature you know nothing at all? Pray, what sort of thing, amongst those that you know not, will you treat us to as the object of your search? Or even supposing, at the best, that you hit upon it, how will you know it is the thing you did not know?
>
> I understand the point you would make, Meno. Do you see what a captious argument you are introducing—that, forsooth, a man cannot inquire either about what he knows or about what he does not know? For he cannot inquire about what he knows, because he knows it, and in that case is in no

> need of inquiry; nor again can lie inquire about what he does not know, since he does not know about what he is to inquire. (Plato, *Meno* 80d-e)

All in all – it states the following dilemma: if we want to know something, then we have two options – we know or we don't know the thing which we are discovering. If we know the thing we are discovering, then it makes no sense to commence with the discovery because we know it already. If we don't know the thing we are discovering then even if we would come across it during our discovery, we would not know that it was the thing we sought. Thus, even in this case, it makes no sense to commence with discovery.

The answer to this "captious argument" is in this dialogue the doctrine of ideas – Plato reinterprets cognition so that putative knowledge (compromised by Meno's dilemma) is in fact re-cognition because in its quest the soul starts to remember the idea of the object which it is looking for.

2.2. Aristotle's solution

It is well known that Aristotle rejected Plato's doctrine of ideas which meant that he, besides other things, had to readdress the Meno paradox. He explicitly handled this problem mostly in chapter 21 of *Prior Analytics*, *Book II*, and in the first chapters of *Posterior Analytics*.[4] Like Plato, Aristotle reinterprets the common term cognition whose limits were rightly accentuated by Meno's paradox. He, however, did not claim that cognition is in reality re-cognition and distinguished between several meanings of the term "knowledge". The starting point was for him one of the basic dichotomies of his philosophy, potentiality and actuality. His *Posterior Analytics* start with the sentence: "All instruction given or received by way of argument proceeds from pre-existent knowledge". Thus, Aristotle – like his teacher – accepted some knowledge preceding one's own cognition.

Instead of identifying this pre-existent knowledge with the recollection of ideas, he understood this *pre-existent cognition as potential knowledge*, or knowledge in possibility, which only in the process of cognizance becomes true actual knowledge. In *Prior Analytics*, Aristotle even distinguished between three meanings of knowledge: knowledge of what is general (further called universal knowledge), what is particular (particular knowledge) and

---

[4] The presented interpretation of Aristotle's solution was taken mostly from Mráz (2000) and Barnes (1978).

knowledge of what is actual (cf. An.Prior. II, 67a33-b5). For a better illustration and deeper understanding of this concept, we use Aristotle's own examples. Aristotle considered (in the chapter mentioned) these arguments:

Every triangle has two right angles.[5]
This is a triangle.

This has two right angles.

Every mule is infertile.
This is a mule.

This is infertile.

Both have the same logical form:

$$\frac{\forall(x)(\mathrm{P}(x) \rightarrow \mathrm{Q}(x)) \quad \mathrm{P}(a)}{\mathrm{Q}(a)}$$

In both, every judgment represents one type of knowledge. The first one is universal knowledge, the second is particular knowledge and the conclusion is actual knowledge. We will step-by-step discuss these individual types of knowledge:

Universal knowledge. Question is how we obtain this kind of knowledge. In Aristotle's case, we can talk about, for example, proof or incomplete induction. These topics are very delicate from the point of modern epistemology but we do not need to further follow them for our purposes.

What matters is that in terms of a given argument we do not know the "general" premise in the same way that we know the conclusion, *i.e.*, in terms of the given argument it is not justified knowledge (or in other words – it is not knowledge justified by the given argument). Simultaneously, we do not have at our disposal all the subjects of universal knowledge. Aristotle demonstrated it in chapter 1 of *Posterior Analytics* in the example of a sophisticated "trick" which had probably been known in his times. The point was to present the

---

5 Aristotle wanted to say that all interior angles of a triangle add up to two right ones.

opponents a sentence "every couple is even" and ask whether they knew it. If they said no, then they would be easily shown that if they understood the term couple well, then they would have had to admit that a couple was even. If they said yes, then they would be presented with a couple that they did not know it was a couple and therefore they also did not know that this couple was even. As Aristotle stated, one suggested solution was to say that if someone claimed that every couple is even, then they actually only said that every couple they know is even. Aristotle rejected this *ad hoc* constriction and insisted that the statement is really relevant to all couples, not only those that the speaker know. Here comes the point – related to the couples that the speaker did not know, the statement "every couple is even" is potential knowledge. That means that he did not actually know it about these couples, but if he realized it was a couple then he would have known it was even. Thus, potential knowledge becomes actual knowledge. In relation to these subjects it is potential knowledge, pre-existent knowledge! *Pre-existent knowledge is potential knowledge*, which is in some way at our disposal, but we always need something else to turn it into actual knowledge.

Particular knowledge. Unlike the aforementioned, this knowledge regards only one thing. What we know about it is that it falls under some universal, or that we recognised this individual as falling under some general determination. Using modern logics, it is a finding that an individual *a* lies in the domain of predicate *P*, while according to the first premise, the domain of *P* is a subset of domain *Q*.

Actual knowledge. It is knowledge which unites the preceding universal and particular knowledge. It is based on *Aristotle's notion that actual knowledge is knowledge of causes or reasons,* thus, actual knowledge is (ideally) the conclusion of an argument. The given premises are the reasons why we actually know what the conclusion states. Actual knowledge is, therefore, justified knowledge and for Aristotle also knowledge of what is necessary.

In order to know what the conclusion says we must know the premises, *i.e.*, think them simultaneously and in mutual correlation. If not, then the conclusion would simply be an accidental statement. Aristotle had this in mind when mentioning both arguments in *Prior Analytics*. If we, for instance, knew that every triangle had two right angles, it would not justify the statement that this had two right angles – we would lack the knowledge that it was a triangle. Similarly, if we only knew it was a triangle, then we could not justify that this had two right angles. Even if we knew that every triangle had two right angles and simultaneously that this was a triangle, it would not mean that we knew that this had two right angles. We have to actually think both findings simultaneously and in mutual correlation. Only this correlation establishes a justifying relationship between the premises and the conclusion.

2.3. Summary

It is a good idea to sum up here what has been said so far. When investigating the paradoxes of inference and analysis, we discovered many common points and most importantly we uncover the problem of pre-existent knowledge as a common source of both paradoxes. I outlined as a plausible concept of pre-existent knowledge Aristotle's theory, which rests on distinguishing three types of knowledge: universal, that is, only potential knowledge, particular knowledge and actual knowledge, which is justified knowledge connecting both previous types, thus, (ideally) the conclusion of the argument. Pre-existent, universal knowledge is therefore (a kind of) potential knowledge. It is neither unknowing nor actual knowledge (thus, none of the Meno's paradox possibilities), but something in the middle: knowledge in possibility.

We should also mention in connection with Meno's paradox that it emerges in relation to universal knowledge. If, for instance, the argument's premises represent particular knowledge (and conclusion actual knowledge), then this dilemma will not arise.

I will now try to apply this Aristotle's concept to both paradoxes.

3. Paradoxes revisited

3.1. Paradox of inference revisited

To remind the reader, the paradox of inference was presented in the form of this argument:

Valid inference from true premises is a good tool for expanding *knowledge*.
Valid inference does not provide any new *knowledge*.

Some good tools for expanding knowledge do not provide any new *knowledge*.

The argument repeats the concept of knowledge several times. After listing Aristotle's classification of types of knowledge, it is worth mentioning which types of knowledge refer to which place.

The second premise ("Valid inference does not provide any new knowledge"): As mentioned in section 1.1., inference does not provide any new knowledge because the conclusion of the sound argument is somehow (implicitly) contained in the premises. It is crucial to grasp its implicit presence in the premises. This is quite difficult to do and so we will use examples. Possibly the best way to demonstrate how the conclusion is contained in the premises is to use arguments of modus ponens form. Let's take argument (a): If this animal is a mule, then this animal is infertile. This animal is a mule. Thus, this animal is infertile. Conclusion of (a) – this animal is infertile – is even to a naked eye contained in the premises. It is not included, though, as an independently standing premise making such a statement. That would make (a) reasoning in a circle and I surely do not wish to claim that every valid argument is reasoning in a circle. Conclusion of (a) is contained in the premises as a consequent of implication (premise one). That means that conclusion of (a) is stated only conditionally, *i.e.*, in order to say that, we would have to know whether the respective condition was met. The first premise does not say that though and we, therefore, cannot find out from the first premise only whether the animal is infertile. This judgment is somehow available to us, but is not stated; we do not know whether it is true. In Aristotle's terms, it is not actual knowledge.

Somewhat less visible is the conclusion of an argument contained in the premises in case of the argument of modus tollens form. Let's take argument (b): If this animal is a mule then this animal is infertile. This animal is not infertile. Thus, this animal is not a mule. Conclusion of (b) – this animal is not a mule – is not already contained in the premises in such an obvious way as was the case of conclusion of (a). This judgement is in the premise as a condition, thus, is not itself stated, it is in a similar "situation", as in case (a), the same judgement is in the conclusion even negated. Not only we have this judgement somehow available to us in the premises but is not stated; we don't know if it is true, but it is also in the conclusion negated as untrue.

Even less obvious is the presence of the conclusion in the premises of the famous syllogism (c): All men are mortal. Socrates is a man. Therefore, Socrates is mortal. Conclusion of (c) – Socrates is mortal – is to the naked eye not at all contained in the premises! Only its "parts" are included, in the second premise subject, in the first predicate, the matter of its affirmation or negation, thus, does not arise at all. What the first premise of (c) says is that the predicate of mortality belongs to all men, but it neither mentions a specific man, nor talks about any specific man, thus, neither mentions Socrates. This premise states that the predicate of mortality is common to men, thus, any man. If someone agrees with this premise, then he/she does not claim that the predicate of mortality belongs only to men he/she knows. It is the same situation as "every couple is even" mentioned above (cf. section 2.2.). If someone agrees with the statement that all men are mortal, then he/she agrees that this judgment relates also to entities he/she does not know, that is, to entities he/she does not know that they are also people. If then he/she does not know that, for instance, someone called Socrates is a man then regarding this Socrates the knowledge that he is mortal is only knowledge in possibility. It is not complete unknowing because the knowledge of first premise *allows* us to gain knowledge that Socrates is also mortal. If the first premise was replaced with for instance a judgment "every mule is infertile", then the first premise would not allow us to gain knowledge that Socrates is also a man.

All in all – the implicit presence of a conclusion in a premise means that this premise offers potential knowledge of (future) conclusion, thus, its preexistent knowledge. That a given judgment appears in the conclusion of a valid argument is not extension of potential knowledge. This potential knowledge was available already in one of the premises. We can rephrase the second premise of paradox of inference as: Inference does not provide any new potential knowledge.

The first premise ("Valid inference from true premises is a good tool for expanding knowledge"): If we recognize that inference does not provide any new potential knowledge, then how to avoid the conclusion that inference is epistemically useless and does not provide any new knowledge? We hinted the answer in the preceding analysis of the first premise of inference paradox – we know the argument's conclusion differently than as shown/hidden in the premises.

To be more rigorous: to know that the conclusion is contained in the premises actually means that we know how to infer it from the premises. Such (explicit) knowledge, thus, means that we thought of both premises at the same time and in mutual correlation which Aristotle stated as a condition for actual, thus justified, knowledge. We saw above that the potential knowledge of the conclusion in the premises is something that requires addition – with modus ponens this is the knowledge of the antecedent's truth, with modus tollens the knowledge of the consequent's falsity, with our syllogism the knowledge that someone called Socrates is also a man. After this addition, knowing the conclusion is knowledge of different kind because the interconnected knowledge contained in the premises (thought simultaneously and in mutual correlation) provides a reason for affirming the conclusion. As late as now the conclusion becomes a real conclusion and the justifying knowledge become premises. Strictly speaking, Aristotle's actual knowledge is therefore not new knowledge because potential knowledge must always precede this knowledge, but it is knowledge of different kind. If we revisit Cohen and Nagel's formulation of the paradox of inference, valid argument is not useful according to them because it does not provide any new knowledge. In line with Aristotle's reasoning, we can reply that we know the conclusion of the argument differently than how we know the premises, *i.e.*, in a justified manner. The conclusion does not provide any completely new knowledge but provides this knowledge in a different way. The utility of the valid argument, thus, rests on how it is given to us – it is different to have some knowledge and to have some knowledge together with its justification. Expansion of actual (not potential) knowledge, hence, means expanding knowledge, which is given in a justified manner.[6]

We can therefore reword the first premise of the paradox of inference as: inference provides new actual knowledge.

The original paradox of inference will, accordingly modified, look like this:

Valid inference from true premises is a good tool
for expanding actual knowledge.
Valid inference does not provide any new potential knowledge.

Some good tools for expanding actual knowledge
do not provide any new potential knowledge.

The difference from the original wording is fundamental and it says that the premises that are true lead to a correctly inferred conclusion that is also true. Thus, paradox is eliminated.

---

6 In modern epistemological logics, this distinction corresponds with the difference between explicit and implicit knowledge (I would like to thank an anonymous reviewer of this paper for mentioning this).

3.2. Paradox of analysis revisited

As a reminder, the paradox of analysis was presented in the following argument:

Correct analysis is a good tool for expanding *knowledge*.
Correct analysis does not provide any new *knowledge*.

Some good tools for expanding knowledge do not provide any new *knowledge*.

Similarly to the paradox of inference, the concept of knowledge also appears here several times. After listing Aristotle's categorization of knowledge types, it will be interesting to see which types of knowledge belong where.

The second premise ("Analysis does not provide any new knowledge"): As mentioned in section 1.2., equality between the analysandum and the analysans means that correct analysis cannot provide anything new. Correctness of the analysis is given by the fact that analysans is already in some way (implicitly) contained in the analysandum. Thus, it is crucial to somehow more thoroughly grasp its implicit presence.

Let us again begin with examples – we mentioned above the analogy of disassembling a machine gun. In the case of a machine gun, I have in front of me a specific whole which I can take apart into pieces (and let us assume that I have never done it before and did not receive any theoretical instruction about the components of a machine gun). When I had the machine gun in front of me, I somehow also had all its components in front of me. If I dismantled it correctly and dismantled only this machine gun, I really have in front of me all its parts. I have in front of me parts after the dismantling of the machine gun which I had somehow in front of me when I had the machine gun in front of me. What is the difference here, why do I use in the first case the vague expression "somehow"? Simply because in the first case, I could see only some parts of the machine gun, for instance, trigger, but others such as the striker were hidden from my view. Even before dismantling I considered the machine gun as a whole and due to its function expected it to contain other parts than those I could see when having a machine gun in front of me as a whole. Thus, I knew that it contained also parts which I could not see before dismantling it. Furthermore, I knew that those parts invisible to me would be in some relation to those which I could see and to each other. However, only executing the dismantlement clarified what the hidden parts were and how they were located in relation to each other, for example when I found a striker and learnt about its relation to the trigger. Thus, I knew already before the dismantlement that there were parts in the machine gun which somehow caused the bullet to leave the gun barrel, but I knew neither what these parts were nor what their relationship to each other parts were. The knowledge of these parts before the dismantlement was, thus, only potential knowledge, pre-existent knowledge, actualised by executing the dismantlement.

If we leave the analogy with a physical analysis, we can use the examples of sentences and their analysis. Let us take the sentence (a) "Every man is mortal"; according to classical modern logics, we would find out after analysis that (a) could be rephrased as: "For every individual, if it is a man, then it is mortal." It is important here that between "man" and "mortal" is according to this analysis a relationship of implication which could be expressed in a phrase such as "if …, then …". Before the analysis, I had in front of me a whole (a) composed of certain parts. These parts also appeared in the analysans of (a). What the analysandum did not explicitly contain was a phrase "if…, then…", in other words, it was not obvious to a naked eye that according to this sentence between man and mortality was a relationship of implication. Even before the analysis, after learning the analysandum, I knew that there was some relationship between man and mortality in the analysandum though expressed by a somewhat ambiguous "is". Only the correct analysis of the statement found out that this relationship is one of sufficient condition when being a man is a sufficient condition for being mortal. Who learns about the analysandum can, thanks to this acquaintance, know that the relation in question is a relationship of sufficient condition. Only after the analysis process, however, has explicit knowledge of structure (a), thus, has actual knowledge. Knowing analysandum without knowing the analysis is, hence, potential knowledge of the sentence's structure. It is not complete unknowing, let us say it is a rather necessary (but not sufficient) condition for the explicit knowledge of implication structure (a).

Even less evident is the analysans' presence in the analysandum in Russell's famous analysis of the sentence (b) "The present king of France is bald". According to Russell, (b) actually contains three sentences connected by conjunction. We could simplify it a lot and put the analysans (b) in the following way: "There is an individual who is presently the king of France and there is only one such individual and this individual is bald". If Russell's analysis (b) is correct, then (b) implicitly contains a conjunction of these three sentences. This knowledge is only implicit, or potential, because (b) contains neither connective "and" which is a sign of a conjunction nor other components of the analysans. Yet, it is not simple unknowing as indicated also by the fact that in case of non-existence of the king of France or in the matter of proper negation (b), the person not having the knowledge of correct analysis (b) can be at a loss. The analysandum, without knowing its analysis, does not in these cases offer clear answers because some parts of the whole (b) and their mutual relations (conjunction here) are hidden in the analysandum. It is hidden analogically to how the striker was hidden in case of the machine gun. Modern logic talks about the grammatical form of the sentence not corresponding with its logical form.

In both examples provided, the implicit knowledge of the whole is a different type of knowledge than explicit knowledge of its parts and their mutual relations. That the analysandum and the analysans contain "the same" (and analysis then does not provide any new knowledge) needs to be corrected. The second premise of the paradox of analysis could be rephrased as: Analysis does not provide any new potential knowledge.

The first premise ("Analysis is a good tool for expanding knowledge"): If we speak about analysis as expanding knowledge, it is clear that we do not mean potential knowledge but justified knowledge, or actual knowledge. If we say, for example, that part of a concept (analysandum) is a part *X*, then it is so to say a snapshot. If we, however, say it after executing an analysis then it is (by a process of correct analysis) justified knowledge. The statement that part of a given concept is component X will be the same in both cases (thus, it is not completely new knowledge), yet in the second case it is justified knowledge. Most importantly though, we emphasised above that many parts of the analysed wholes cannot be simply detected from these wholes, even more so in case of the parts' mutual relations. The first premise of the paradox of analysis can be therefore rephrased as: Analysis provides new actual knowledge.

The original paradox of analysis will after modification look like this:

Analysis is a good tool for expanding actual knowledge.
Analysis does not provide any new potential knowledge.

Some good tools for expanding actual knowledge
do not provide any new potential knowledge.

The difference from the original wording is substantial. It states that a correct conclusion is inferred from premises which are also correct. The paradox is eliminated.

4. Conclusion

The paradox of inference and paradox of analysis have both been understood in our research as grounded in the problematics of pre-existent knowledge. The clarified concept of pre-existent knowledge as presented by Aristotle offers the opportunity to understand both paradoxes as based on a vacillating concept of knowledge which leads some authors to opposite opinions. Further specifications of this concept leads in contrast to a discovery that these opposing opinions refer to different types of knowledge and as such do not actually need to be contradictory and do not lead to a paradox.[7]

References

BARNES, J. (1975): *Aristotle's Posterior Analytics*. Oxford: Clarendon Press.

BARNES, J. (1984): *Complete Works of Aristotle. Vol 1*. Princeton: Princeton University Press.

BEANEY, M. (2015): Analysis. In: Zalta, E. (eds.): *Stanford Encyclopedia of Philosophy*. Available at: http://plato.stanford.edu/entries/analysis/.

COHEN, M. R. & NAGEL, E. (1934): *An Introduction to Logic and Scientific Method.* London: Routledge & Kegan Paul.

DUŽÍ, M. (2006): Informativnost matematických či analyticky pravdivých tvrzení a paradox inference. [The informativeness of mathematical or analytically true statements and the paradox of inference.] *Filosofický časopis* 54, No. 4, 501-522.

DUŽÍ, M. (2010): The Paradox of Inference and the Non-triviality of Analytic Information. *Journal of Philosophical Logic* 39, No. 5, 473-510.

MILL, J. S. (1882): *A System of Logic, Ratiocinative and Inductive*. New York: Harper & Brothers Publishers.

MOORE, G. E. (1903): *Principia Ethica*. Cambridge: Cambridge University Press.

MRÁZ, M. (2000): Aristotelova kritika úvah o anamnési v Platónově dialogu Menón. [Aristotle's Critique of Anamnesis Considerations in Plato's Meno.] In: Havlíček, A. (ed.): *Platónův dialog Menón*. Praha: Oikoymenh.

---

7 I would like to thank an anonymous reviewer for the very inspiring comments.

PLATO (1967): *Plato in Twelve Volumes. Vol. 3.* Lamb, W. R. M. (trans.); Cambridge (Mass.): Harvard University Press – London: William Heinemann Ltd.

WITTGENSTEIN, L. (2001): *Tractatus Logico-Philosophicus*. Pears, D. F. & McGuinness, B. F. (trans.); London: Routledge.

Organon F 23 (2) 2016: 243-262

The Reception of Stanisław Leśniewski's Ontology in Arthur Prior's Logic

ZUZANA RYBAŘÍKOVÁ

Department of Philosophy. Palacký University Olomouc
Křížkovského 12. 771 80 Olomouc. Czech Republic
zuzka.rybarikova@gmail.com

RECEIVED: 01-10-2015 ▪ ACCEPTED: 11-01-2016

ABSTRACT: Arthur Prior's logic was influenced, among others, by logicians from the Lvov-Warsaw school. This paper introduces the impact Leśniewski's Ontology had on Prior's logical system. The paper describes the main characteristics of Leśniewski's Ontology, Prior's logical system and the manner in which Prior became acquainted with Leśniewski's logical system. Since Leśniewski was no longer alive when Prior began to develop his logical system and Leśniewski's papers were not easily available to Prior, this paper also includes Prior's interpretation of Leśniewski's logical system which did not always correspond to Leśniewski's original ideas.

KEYWORDS: A. N. Prior – Leśniewski's names – Leśniewski's Ontology – Stanisław Leśniewski.

1. Introduction

Arthur Prior's ontological position was in many cases unique as he combined intensional logic and nominalism. The aim of this paper is to demonstrate that Prior's distinctive ontological position was also made possible through his adoption of certain features of Leśniewski's Ontology.[1] This paper conse-

[1] Ontology is often written with a capital O in this paper. This indicates that we are

quently discusses the impact Leśniewski's Ontology has on Arthur Prior's ontological position and the ontological commitment of his logic.[2] One of Prior's (1971) texts even has Leśniewski's name in its title and Prior wrote several reviews of texts by students of Leśniewski which discussed Leśniewski's logical systems. The reasons Leśniewski's Ontology was interesting for Prior will also be mentioned.

Prior was acquainted with Leśniewski's works despite the fact that Leśniewski's papers were not easily available when Prior developed his logical systems.[3] Prior knew this logical system from works of Leśniewski's students and colleagues and from personal communication with them (see Sobociński 1953, Lejewski 1956). Leśniewski's ideas could have, however, been misinterpreted by Prior because his knowledge of Leśniewski's logical system was primarily based on the work of his students and colleagues not on Leśniewski's own ideas. In addition, this paper discusses to what extent Prior departed from Leśniewski's original ideas when he incorporated his theory into his logical systems.

The form of Ontology in Prior's logical system is primarily examined in the works that both authors wrote at the end of their lives as their logical and ontological positions changed a great deal over the course of their lives. Leśniewski's *Foundation of Mathematics* and *On the Foundation of Ontology* and Prior's *Time and Modality*, *Object of Thought* and *Existence in Russell and Leśniewski* were thereby chosen for the analysis. In light of the fact that Prior primarily knew the works of Leśniewski's students, these works are also discussed, in particular Lejewski's *Logic and Existence* and Słupecki's *Leśniewski's Calculus of Names*. Prior was also aware of Sobociński's works but Sobociński chiefly deals with issues which are not deeply investigated in this pa-

---

not speaking of ontology as it is understood in most philosophical debates but specifically as in Leśniewski's system of logic which is in some cases similar to ontology but which in many ways also differs.

2 Although Leśniewski's impact on Prior is well known among logicians which handle with Lvov-Warsaw School (see e.g. Woleński 1989, 155; Simons 1982, 191; Urbaniak 2014a, 104 and 192), it is not discussed among logicians who focus on Prior.

3 Storrs McCall's book *Polish Logic* was published in 1967 and included two of Leśniewski's papers. Prior would have known this book since he wrote a review on it. Prior never quoted one of the papers as far as I am aware.

per. From the Ontological point of view, Sobociński's letters are of most interest. These are deposited in the Bodleian Library and in them Sobociński attempts to explain to Prior the main aims and procedures of Ontology and Protothetic.

It is worth emphasizing here that Leśniewski's and Prior's philosophies shared a common thread even though they came from different logical traditions. Prior began studying logic in traditions which were referred to as 'orthodox logical systems'[4] by Simons (1982, 165). The orthodox systems of logic are systems created on the foundations laid by Peano, Frege and Russell. It is these systems which are the most widespread in modern logic at present. Leśniewski, in contrast, formulated his own logical system which differed from the orthodox systems in a number of aspects. It seems unusual that Prior, a logician from New Zealand who was primarily familiar with the Anglo-Saxon logical tradition, found common ground with a logician from Poland whose logical system is unusual in many features. As Uckelmann (2012, 352) points out, however, Prior discovered Łukasiewicz's work on the history of logic and his innovation in modern logic during his teaching at Canterbury University and became interested in his logical systems. Prior began to be introduced to the concepts of Lvov-Warsaw School through Łukasiewicz and his student Bocheński.

In light of the teachers of both Prior and Leśniewski, there is a common thread leading to the same person. This person was Franz Brentano who was the teacher of Meinong and Twardowski. Twardowski was the philosopher who established the Lvov-Warsaw School in Lvov before World War I and was more (e.g. in Łukasiewicz's case) or less (e.g. in Leśniewski's case) the teacher of nearly all the members of the Lvov-Warsaw School (see Woleński 1989, 3-7). John N. Findlay, who was A. N. Prior's teacher, studied for several years in Europe and published an influential book which discussed Meinong's Objects (see Copeland 2008). As a result, Leśniewski's logical systems were not as unfamiliar to Prior as they might have otherwise been.

When Prior discussed Leśniewski's logical system he nevertheless tried to adapt it to the orthodox logical systems. This approach was not without sacrifices on both sides and certain authors have doubts as to whether it was actually

4 This title is used throughout the paper.

successful (cf. Sagal 1973, 259-262; Simons 1982, 177). Their remarks will be introduced in the further part of the paper.

2. Leśniewski's System of Logic

Stanisław Leśniewski was one of the most renowned members of the Lvov-Warsaw School. He was born in 1886 and died in 1939. Leśniewski began to develop his logical system in 1916. He tried to invent a logical system which mathematics could be based on as Russell did in his *Principia Mathematica*. There are certain differences between Leśniewski and Russell. Leśniewski (1992a, 74-75, 126) was dissatisfied by Russell's solution of Russell's antinomy. Namely, as a nominalist he did not approved the existence of classes and sets. Hence, wanted to devise a system which would not contain antinomies and any of the other ambiguities which appeared in Russell's system and which at the same time would not presuppose existence of classes and sets (see Luschei 1962, 25-33; Urbaniak 2014b, 290-292; and Urbaniak 2015, 127-131).

Although Leśniewski was convinced that his system could solve previously mentioned problems which occurred in Russell's *Principia Mathematica*, this system is not widely used. Simons (2011) asserts that it might have been caused by the fact that Leśniewski's papers were primarily written in Polish and to a lesser extent in German. Leśniewski's perfectionism could have been another reason why his work was not well known in Prior's day. Since Leśniewski (1992a, 174-176) did not allow the publication of his texts until they were perfect, only a fragment of his work was published while he lived. After his death all his works were prepared for publication by his students. World War II began, however, shortly after Leśniewski's death and brought a stop to the publication of the texts. Leśniewski's texts were deposited in Warsaw which burned down when the Warsaw Uprising was defeated. Leśniewski's students and colleagues reconstructed Leśniewski's logic after the War (see Luschei 1962, 25-26) but following Leśniewski's death and the destruction of his works, it proved impossible to entirely reconstruct his work.

Leśniewski (1992a, 176-177) built his logical system on three theories: Protothetic, Ontology and Mereology. They are usually presented in this order because it represents a hierarchy. Protothetic together with Ontology are the theories which demonstrated Leśniewski's logical position. Mereology is an extra-logical theory which deals with parts and wholes. These three theories are, according to Leśniewski, the basis for the foundation of mathematics. The division of logical theories into Protothetic and Ontology corresponds more or less to the division of two fields of logic, the logic of propositions and the logic of terms. Protothetic is also sometimes known as the calculus of propositions and Ontology is called the calculus of names (see Słupecki 1984; Paśniczek 1996).

Since each of Leśniewski's theories has been discussed in numerous papers, only Ontology, which had the greatest influence on Prior when he formulated his ontological position, is introduced. Protothetic is also dealt with in some of Prior's works and as Sobociński (1953) demonstrated in his letter, Protothetic and Ontology are strongly connected. In order to keep the paper within limits, however, I will focus exclusively on Ontology.

3. Leśniewski's Ontology

Leśniewski (1992a, 373-374) named the system Ontology, based on the Greek "ὄντος", which means "being" in English. He was aware that "ontology" was the name of a discipline which deals with "the general principles of existence" and that this description does not correspond with his concept of ontology. He also pointed out that his theory had certain similarities with the ontology defined by Aristotle and was part of a philosophical tradition spanning back centuries. Leśniewski presumed that if Aristotle's theory was described as the "the general theory of objects", it is not far from his own Ontology. Ontology introduces "some principles of existence" but in an extremely narrow sense. It describes Leśniewski's linguistic intuitions, the language and its usage but does not deal with beings themselves.

The way this works is contained in Leśniewski's concept of quantifiers, or more precisely the concept of a quantifier, because Leśniewski's Ontology includes only the universal quantifier in Leśniewski's original concept. Sobociński (1953) claims that the existential quantifier should not occur in Leśniewski's Ontology, even though, Leśniewski's students used it in his papers in order to simplify explanations.

3.1. The functor ε

The most important and the only primitive functor in Ontology is *ε*. According to Leśniewski, the best translation of this functor is the verb "is", nevertheless, it is "is" with the meaning it has in Polish or Latin. Leśniewski (1992b, 608-609), as well as Russell (1919, 172), were aware that in English "is" could have more than one meaning. This is due to the difference between the definite and indefinite article which occurs in English and some other languages where it is not present in Latin, Polish and other Slavic languages. Although there is such a difference, Leśniewski, whose logical systems were influenced by his linguistic intuitions (see Miéville 2009, 4-5), expressed the functor *ε* as the colloquial Polish word "jest". Leśniewski (1992a, 376-382) was aware that there are also differences in the use of colloquial Polish but as his followers (cf. Słupecki 1984, 65; Rickey 1998, 31-32; Woleński 1999, 18-19) have demonstrated, the main difference in the usage of "is" lies between the languages which contain the definite and indefinite article and those languages which do not contain them.

The three meanings of the word "is" can be demonstrated by three statements which were also used by Leśniewski's student Słupecki (1984). In his article entitled *Leśniewski's Calculus of Names*, Słupecki introduces three examples of statements in which the word "is" has a different meaning in English and in Latin (Słupecki 1984, 65):

Socrates is a man.	*Socrates est homo.*
The dog is an animal.	*Canis est animal.*
Socrates is the husband of Xantippe.	*Socrates est coniunx Xantippae.*

Słupecki claims that the three statements in the first column have different meanings. Furthermore, if the statements in the second column are considered correct translations of the first column, their meanings have also to differ. Słupecki points out, however, that Leśniewski worked with the form in which they all have the same meaning. Moreover, this meaning differs from the meaning "is" has in English statements. This meaning can be demonstrated by the description of the functor *ε* which occurs in Luschei's book *The Logical System of Leśniewski*. According to Luschei (1962), the definition of the formula $A \varepsilon b$ is:

> *Singular predication* or *inclusion ("relation of being")*: A is b; (the sole) A is (a or the sole) b; (individual) A is (one of the one or more) b; A is (an individual that is) b; A is one of the one or more individuals that are b; being b characterizes (individual) A; there is exactly (i.e., at least and at most) one A, and (any) A is b. (Luschei 1962, 10)

Słupecki (1984, 65-68) argues that Polish and Latin statements can be found in which "is" does not correspond with Leśniewski's description. Leśniewski (1992a, 376-382) problematizes them in his *Foundation of Mathematics*. Although the functor ε is equivalent to the word "is" in an ordinary article-free language, there are certain exceptions. The word "is" is not the equivalent of the word "exist" and also does not have the meaning "is now". Apart from the statement being meaningful, if the subject of the statement combined with ε is a common noun or an empty name where that statement is always false. The statement '*The dog is an animal*' consequently has to be rewritten. The correct form of this statement is '*Whatever is a dog is an animal*'. There is no such aid, however, for the statements which contain an empty name. As Słupecki (1984, 68) discusses, statements which have an empty name as a subject or predicate, such as '*Hamlet is the king of the Danes*' or '*Barack Obama is a vampire*', are false and there is no way to change it.

The difference in the meaning of the word "is" which exists between English and Latin led Słupecki to the conclusion that Leśniewski's functor ε cannot be translated into English. Rickey (1998, 31-32) and Woleński (1999, 18-19) disagree with his findings. Rickey suggests that English-speaking authors should use ε in a specific technical definition as it is used in Ontology. Woleński points out that the correct usage of the functor ε is not a case of linguistic intuition but requires a detailed analysis.

3.2. Nouns and names

A description of the Leśniewskian names is necessary since not every noun can serve as a value for the formula $A\, \varepsilon\, b$ if one intends to create a true statement. As was mentioned before, the statement '*Charlotte is a fairy*' or '*The giraffe is a mammal*' are grammatically correct and meaningful but nevertheless false in Leśniewski's Ontology.

When Sobociński (1953) describes Leśniewski's system of logic in his letter to Prior, he asserts that there are two semantic categories in this system, the category of names and the category of propositional functors. The former category is the point of interest for this chapter. It will be demonstrated that there is a difference between the concept of names, as is well known in Russell's logical system and the concept of names in Leśniewski's ontology.

As Zuber (1998, 219) points out, this is also based on dissimilarities between Polish and English. Zuber (1998, 230-233) demonstrates that Polish is an inflected language and hence the statements do not have a strict form. It is grammatically correct in Polish to form the sentence '*Jacek jest przewodnikiem'* (Jacek is a guide) and the sentence '*Przewodnikiem jest Jacek*' is also correct. The subject is consequently not defined by the position of the term in a statement.

In addition, if the common noun in the sentence is connected with the determiner as in the sentences '*Każdy żołnierz jest odważny*' (Every soldier is brave), '*Ten żołnierz jest odważny*' (This soldier is brave), or '*Nasz żołnierz jest odważny*' (Our soldier is brave) then it belongs to the same semantic category as proper names in Polish. The bare noun without a determiner, in contrast, has no proper sense in Polish. The sentence '*Żołnierz jest odważny*' (A soldier is brave) is only tolerable when interpreted very broadly. Hence certain, but not every, common noun can be the term of a true statement in Leśniewski's logic. The problem lies in the fact that Polish, as well as other article-free languages, lack articles which play the role of determiners in other languages.

Apart from this distinction, which occurs between the Russellian and the Leśniewskian names, there is one more important feature of Leśniewski's system of logic from the ontological point of view, namely quantification. There is no doubt that the concept of quantification is one of the core concepts of Leśniewski's ontology. Namely, Leśniewski's quantification is not as linked with existence as Russell's. Leśniewski introduced an operator "ex" in order to formalize the verb "exist". The statement "Unicorn does not exist" is formalized as $[\exists a].\sim ex(a)$ in his system of logic, which means "Some unicorn does not exist" (see Urbaniak 2008, 120).

As Urbaniak (2014a, 189-191) claims, several questions arise which Leśniewski did not address. Firstly, he points out that Leśniewski did not postulate which entities are values of variables which are bound by quantifiers. Secondly, there is no consensus among authors as to whether Leśniewski's quantifiers required ontological commitment as Quine's do. In contrast, there is mostly agreement among them that there is a difference between Leśniewski's and Quine's theories of quantification.

Several authors suggested solutions to these queries. Prior was inspired in this case primarily by Lejewski. Hence the interpretation, which Lejewski presented in his paper *Logic and Existence*, will be discussed. Lejewski demonstrates the diverse ontological concepts by a thought experiment:

> To have a still simpler though fictitious example let us think of the universe as limited to two objects **a** and **b**. Then the corresponding expansions would be: F**a** ∨ F**b** and F**a** ∧ F**b**. Our language, which for reasons of simplicity needs not synonyms, may leave room for noun-expressions other than the singular names "a" and "b". We may wish to have a noun-expression "c" which would designate neither of the two objects, in other words which would be empty, and also a noun-expression "d" which would designate either. (Lejewski 1954, 109)

If the predicate F can be truly asserted to *a* and *b*, then the formula $\exists x(Fx)$ is true in Ontology but the formula $\forall x(Fx)$ is false, although both formulas are true in Quine's interpretation. This is caused by the fact that in Lejewski's interpretation the variable *x* in both formulas stands for all the noun-expressions. The formula $\exists x(Fx)$ in Lejewski's interpretation means either *a* or *b* or *c* or *d* have this property which is true since *F* can be ascribed to *a*, *b* and also *d*. The formula $\forall x(Fx)$ means that *a* and *b* and *c* and *d* have this property. The latter formula has to be upheld for all noun-expressions to be true and it is not since the noun-expression *c* has no reference. The formula $\forall x(Fx)$ consequently has to be false.

Although this seems to be the disadvantage of the system, other differences occur, which was later used by Prior, if it is analysed deeper. Namely, as Lejewski (1954, 109-110) pointed out that *d* behaves like a noun. It has to be reformulated as $D(x)$ in Quine's interpretation but not in Leśniewski's. It does not cause any harm that the constant *d* refers to two individuals in Ontology, unlike Russell's and Quine's system of logic where constants stand for precisely one individual.

Lejewski was of the opinion that this experiment also expresses the differences between quantification in Quine's and Leśniewski's logic. Based on Quine's famous theory of ontological commitment, variables which are bound by existential quantifiers have to signify something existent. In contrast, the Leśniewskian quantification in Lejewski's (and also Prior's interpretation) is different. Lejewski (1954, 113-114) therefore suggested that the designation "existential quantifier", which could be misleading in the Leśniewskian interpretation, should be replaced by the designation "particular quantifier". This replacement is in accordance with the Leśniewskian interpretation of quantifiers. Lejewski was more likely to interpret more formulas with existential quantifiers in an Aristotelian way. A formula such as $\exists x(Fx)$ is not translated as "There exists *x*, such that *Sx*." but "For some *x*, *Sx*".

The variables in Lejewski's interpretation represent noun-expressions which refer to a concrete object or objects in the case of the noun-expression *d*. In addition, objects which can be unproblematically bound by quantifiers in Lejewski's interpretation of Ontology include such dubious entities as numbers and colours. It therefore seems that objects are values of variables in Lejewski's interpretation, even though in a quite wide sense of the word "object" and the variables refer to them indirectly. In the following chapters, the way in which Prior adopted these ideas will be presented.

4. Arthur Prior's approach to logic

Arthur Prior is considered one of the founders of modern temporal logic and also created new systems of modal logic. Although Prior was an intensional logician, as Hugly & Sayward (1996, 47-48) point out, he did not postulate the existence of such entities as intensional objects because as a nominalist he did not acknowledge the existence of all abstract entities. The intensionality of his system consequently meant that he admitted intensional functions.

When Prior (1957) formulated his temporal logic, he intended to enclose it in natural language. This is the reason he also assumes the medieval concept of propositions, which differs from Frege's. Based on this concept, a proposition can be true at one time and false at another time. The proposition '*The head of my Department is a logician*' was therefore true when I wrote the first version of my paper and when it referred to the Department of Philosophy at which I work. The same proposition is currently false and would also be false when referring to a different department, because, according to the medieval concept of propositions, it is still the same proposition.

Although a logical system that includes this concept of propositions is closer to natural language, it has to manage the problem of entities which do not exist permanently. This is particularly the case when Prior developed temporal logic, where past, present but also future figure. The following chapters discuss to what extent Ontology played a crucial role in solving this problem.

5. Prior discovers Leśniewski

Prior did not discover Leśniewski's logical system directly. When Prior began his correspondence with Polish logicians, Leśniewski had already been dead for several years. His archive had been destroyed and all Poles were far from their homeland. Thus, their access to Leśniewski's papers and papers of his other students and colleagues were limited (see Sobociński 1953, 5). Łukasiewicz was recognised by Prior (1955-1956, 199) as the man who introduced him to Leśniewski's logic. Łukasiewicz was not the only one of Leśniewski's colleagues, however, whose work Prior knew. Prior also mentions Sobociński's introduction to Protothetic.

Prior describes his first impression of Leśniewski's logical system in his paper *Definition, Rules and Axioms*. In this paper, Prior also discusses Protothetic logic and Leśniewski's theory of definition, not only Ontology. The article also demonstrates that Prior had several comments on Leśniewski's logical system. He firstly criticizes multiplying axioms which occur in Leśniewskian systems of logic and which were introduced to him by Leśniewski's students. He secondly has an aversion to Leśniewski's concept of names in which empty-names also occur. Despite his criticism, he later uses this specific concept of names in his nominalism. Even in this paper he appreciates certain features of Leśniewski's system.

Although Prior was initially critical of Leśniewski's logical system, it influenced a great deal of his own logic. This can be illustrated through a comparison of two of Prior's books. Prior was interested in the history of logic and was preparing the publication of an exhaustive book about this topic. Its title would have been *The Craft of Formal Logic*[5] but due to its length the publishing house recommended that Prior shorten it. Prior instead wrote a new book entitled *Formal Logic* (see Copeland 2008). After Prior's death, certain fragments of *The Craft* were published by P. T. Geach and A. J. P. Kenny. This book was entitled *The Doctrine of Propositions and Terms*.

[5] The title of this book is consequently shortened as *The Craft*.

Prior demonstrated here his brilliant knowledge of the history of logic. Although the Polish logicians Łukasiewicz and Bocheński are mentioned in this book, his main discussion is on ancient and medieval logic along with the logical theories of the 19th century and the beginning of the 20th century. In contrast, *Formal Logic* focuses considerably on the logical systems of Polish logicians.

Leśniewski's system is often discussed in Prior's later works. It seems that Prior appreciated Leśniewski's work more when he developed his own temporal logic. When Prior (1957, 63-75) formulated his ontological position in his *Time and Modality*, he criticised Russell's concept of names as inappropriate to his systems of logic. Hence, in his system of logic ΣT_2, he combined tense logic and Ontology, primarily Leśniewski's concept of names. Prior emphasizes that the difference between this system of logic and the ΣT_1 system, where Russell's calculus is used instead of Ontology, is that proper names are replaced by common nouns in ΣT_2. Prior also uses Leśniewski's functor[6] when he describes Ontology here. Ontology helps Prior to solve the problem of entities which do not have an actual existence.

In contrast, Prior was aware that Leśniewski's concept of logic differs considerably from his own. Leśniewski considered propositions as timeless and has a preference for extensional logic. Hence ΣT_2 could not completely replace ΣT_1, but Prior incorporated some parts of ΣT_2 to ΣT_1 to utilize the advantages of both systems. He also pointed out that the ΣT_1 system had to be enriched by special propositional and predicate variables.

Since Prior first worked on improving his logical systems of temporal and modal logic, he postponed addressing questions which arose in his own ontology. Consequently, his most important book, *Past, Present and Future*, which was published after the publication of *Time and Modality*, does not contain any satisfactory improvement of his ontology. The concept of names which Prior prefers is clearly formulated:

> …we just have no Russellian individual name-variables at all, bound *or* free, but only devices for referring to individuals obliquely, as in Leśniewski's 'ontology'. (Prior 1967, 162)

6 However, Prior used in *Time and Modality* the symbol "∈" instead of "ε" which could be misleading as will be discussed further.

It is still doubtful, however, whether Prior was actually able to interpret Ontology correctly. In spite of the differences that exist between orthodox logic and Leśniewski's system, Prior was more acquainted with the works of Leśniewski's students than Leśniewski's own papers. The differences which arise between Prior's interpretation of Leśniewski's system and Leśniewski's original system are consequently discussed in the following chapter.

6. The reception of Ontology in Prior's logic

6.1. Prior's interpretation of the functor ε

Although *ε* was described as "is" in article-free languages such as Polish and Latin, it was shown in a previous part of this paper that English researchers are able to use it properly. It only requires precision in the use of this specific term. This chapter will therefore investigate whether Prior used the functor *ε* correctly and what his interpretation of this functor was.

Prior's concept of the functor *ε* was influenced by Lejewski's understanding of it, since they discussed it in their letters (see Lejewski 1956). Prior was also acquainted with Słupecki's paper *S. Leśniewski's Calculus of Names* in which Słupecki introduced Ontology. Prior adopted this concept in his own paper *Existence in Leśniewski and in Russell.*

> Leśniewski meets this difficulty by introducing an undefined constant expressing a relation between classes – it can be, but does not need to be, the functor "ε" previously mentioned. This functor, as I have also previously said, has arguments of the same logical type, so that what it express is *not* Russellian class-membership. It express rather the *inclusion* of a unit class in another class. (Prior 1971, 163)

This is not Leśniewski's original interpretation of the functor *ε*, however, and Prior is aware of this. He continues in the very next part of his paper: "…and although Leśniewski himself did not like it, no other interpretation of the symbol seems to me intelligible" (Prior 1971, 151); and Prior clearly admits in his paper that the interpretation of the functor *ε*, which he has chosen, is not Leśniewski's.

Additional reasons for why Prior rejected Leśniewski's interpretation of the functor *ε* can also be found. The most plausible explanation seems to be, however, the one offered by Simons in his paper *On Understanding Leśniewski*. Simons (1982, 165) examines ways of understanding Leśniewski's logical system by logicians which came from a tradition that Simons calls the orthodox systems of logic. Since Prior came to Ontology from this position, he must have perceived Ontology by means of the tools of the orthodox logical systems.

When Prior (1957, 63-75) uses Leśniewski's names, he does so without the ambition of reconstructing Ontology. He attempts to implement some of Leśniewski's inventions in his own logical system. His system is consequently closer to orthodox logic than the Leśniewski system, as he interpreted the functor ε in a way in which it is more translatable in orthodox logic. The meaning of the functor ε in Prior's interpretation lies somewhere between the Leśniewskian ε and the Russellian $\in$. This could be problematic. As Słupecki (1984, 69-72) stresses that and Russell's $\in$ cannot replace Leśniewski's ε, and vice versa, since Leśniewski's functor binds two words which belong to the same semantic categories while Russell's binds a name and a class.

There is still one more distinction between Leśniewski's and Prior's concept of the functor ε. Since in Prior's logic propositions can have different truth values at different times, and individuals are postulated as temporal, Prior distinguishes three possible meanings of the functor ε. Prior (1957, 76-83) emphasizes this in his *Time and Modality* when he discusses the two meanings of the article "the", the weak "the" and the strong "the". As Leśniewski did not hold this concept of propositions, nothing similar occurs in his logic system.

The weak "the" is an article in the formula "The *a* is a *b*" where the specification depends on the time frame, as in the statement "The president of Russia is the owner of a dog". This statement can only be true when there is only one individual which the predicate can be assigned to at the time of utterance. The statement was consequently true when this paper was written but was not true several years ago when Dimitri Medvedev was president of Russia and it might not be true after Vladimir Putin finishes his career. Therefore, the weak "the" has only a temporary significance. This sense holds the functor $\in$ in the ΣT_2 system.

In contrast, the strong "the" fixes its signification regardless of time. If there is the strong "the" in the statement "The *a* is a *b*", the *a*, which is a *b*, is the only one individual that ever was, is or will be the *a*. Examples of such a statement could be "The best known pupil of Plato was a clever man". When the strong "the" is used in the logical system, the functor ∈ is replaced by the functor ∈'. Prior additionally defines a new logical system, ΣT_3, where the functor ∈' applies.

Prior also introduces the functor ∈". The functor ∈" is derived from the functor ∈' and describes the situation when, in the statement "The *a* is *b*", the subject is characterised by the strong "the" and the predicate is an identifiable individual.[7] Prior called the logical system in which the functor ∈" occurs the naïve object-existent system. The functor ∈" is the most useful functor among the ∈-functors because it enables Prior to create statements which deal with non-existent entities. He does not need to postulate either their existence or their properties. Prior does not want to postulate *possibilia* in this way, but in his concept of individuals, he has to deal with entities which do not actually exist, but which existed or will exist.

Prior's treatment of individuals of Ontology is distant, however, from Leśniewski's own interpretation. Simons (1982, 177-182) in his article demonstrates that the functor *ε* can be interpreted in accordance with Leśniewski's definition, but that interpretation does not suit the requirements of Prior's logical systems. The functor *ε* and the terms which are bound with it have an existential import in Simons' interpretation. Prior (1971, 161) requires, for the applicability of the system, a different concept of terms bound by the functor *ε* (or ∈ in some of Prior's works (Prior 1957, 63-75)). In addition, it is obvious that there are more differences between Prior's and Leśniewski's positions. They will be introduced in the following chapter, where Prior's concept of Leśniewskian names is discussed.

6.2. The difference between the concept of nouns in Ontology and Prior's logical systems

Since Leśniewski's Ontology is also described as the calculus of names, Prior's concept of names can demonstrate to what extent Prior actually associated Ontology with his own logical system. As was shown in previous chapters, Prior rejected the Russellian names because he did not want to postulate the existence of actual non-existent entities, although he had to work with them

---

7 The Identifiable individual is an individual which has a contingent existence but is determined by its past. Its future is open but it cannot act otherwise than it acted in the past. Moreover, events that happened to it also cannot change (see Prior 1968, 66-77).

in his temporal logic. He consequently incorporated Leśniewski's names into his system instead of those of Russell.

Simons (1982, 177-182) emphasizes that Prior also did not fulfil all the stipulations that are identified in Ontology. Although Prior was aware that his interpretation of the functor ε differed from Leśniewski's, he in all probability did not possess any doubts about his interpretation of the Leśniewskian names which he included in his own logical system. Simons observes, however, two different interpretations of Ontology which can be found in Prior's work. He finds that Prior construed Leśniewski's names as class names or as common names. This does not mean, however, that Prior had two different understanding of Leśniewskian names. These two concepts are primarily connected in Prior's work.

An example of such a connection can be found in Prior's paper *Existence in Leśniewski and in Russell*. Firstly, Leśniewski's names are described as class names. Prior claims:

> Ontology's so called "names", in other words, are not individual names in the Russellian sense, but *class* names. This immediately explains the first two of the peculiarities I have mentioned. For while it makes nonsense to divide up individual names in this way, class-names *are* divisible into those which apply to no individuals, those which apply to exactly one, and those which apply to several. It makes sense also to say that some classes "exist", either in the sense of having at least one member or in the sense of having exactly one member, and some classes do "exist" in these senses and some do not. (Prior 1971, 162)

Prior's replacement of the symbol ε with the symbol $\in$ also affirms that he considered that Leśniewski's names behave like classes.[8]

Simons (1982, 177-178) emphasizes that Leśniewski as a nominalist cannot agree with the postulation of classes. Prior does not agree, however, with the postulation of classes either. He also considered himself a nominalist. In spite of the fact that he uses class to approximate Leśniewskian names to the orthodox logical systems, he rejects them having some means of existence. He claims:

> It may seem from what I have said that ontology, on my interpretation of it, is committed to the existence of classes as nameable entities, though in

---

8 This feature of Prior's paper was particularly criticised by Sagal (1973, 259-262).

> fact Leśniewski was notoriously nominalistic. But this is a misunderstanding, arising from the use of the perhaps unfortunate term "class-name". What we have to do deal with here are *common nouns*, and these are not strictly speaking *names of objects* at all. (Prior 1971, 165)

When Prior postulates Leśniewskian names in his logical systems, however, both descriptions can be used. They are defined as class names, and the hierarchy of classes can describe precisely how they operate in a logical system. From the ontological point of view, however, they are treated as common nouns. As Słupecki (1984, 71) emphasizes, nouns bound by the functor ε should be of the same semantic category. Hence the concept in which Leśniewskian names are described as common nouns complies better with Leśniewski's requirements.

In contrast, as Urbaniak (2014a, 189) points out, there is no consensus among authors as to which entity is represented by the bound variables in Leśniewski's Ontology. It consequently cannot be claimed that it was actually Leśniewski's concept of names which played such an important role in Prior's ontological ideas. It was instead several of Leśniewski's ideas primarily surmised from the works of Leśniewski's pupils. In addition, these ideas were occasionally misunderstood by Prior and when mixed with orthodox logic resulted in the formulation of names that Prior calls Leśniewskian in his logical systems.

7. Conclusion

To sum up, although Prior adopted Leśniewski's concept of names, not everything that he attributed to Leśniewski was actually compatible with Leśniewski's concepts. There are common features in both Prior's and Leśniewski's systems of logic. They both tried to create systems of logic which can be combined with nominalism and both also had a preference for natural language to the formal system. In contrast, Prior developed some ideas which he found in the papers of Leśniewski's students so radically that even he had to admit that they differed from Leśniewski's thoughts. This can be demonstrated by the functor ε, in which the change in usage was caused by the difference between Prior's and Leśniewski's concept of propositions. Finally, the concept of names, which is not the same in Ontology and Prior's logical systems, can be representative of the third way in which Prior adopted concepts of Ontology. Although Prior thought that he introduced the Leśniewskian names into his logical system, they did not fulfil all of Leśniewski's requirements.

The adaptation of some features of Leśniewski's system of logic, even though misinterpreted in certain ways, nevertheless enabled Prior to formulate his ontological position. The contribution of Leśniewski's Ontology is remarkable primarily in Prior's concept of names in which he had to combine nominalism with the intensional context and medieval concept of propositions.

Acknowledgments

I am grateful to professor Jan Štěpán and anonymous reviewers for their comments on the previous version of this paper. This work was supported by the student project "Informal Logic and Argumentation Theory" No. FF_2013_050 of Palacky University.

References

COPELAND, B. J. (2008): Arthur Prior. In: Zalta, E. N. (ed.): *The Stanford Encyclopedia of Philosophy*. Available at: http://plato.stanford.edu/archives/fall2008/entries/prior/.

HUGLY, P. & SAYWARD, C. (1996): *Intensionality and Truth: An Essay on the Philosophy of A. N. Prior*. Dordrecht: Kluwer.

LEJEWSKI, C. (1954): Logic and Existence. *The British Journal for the Philosophy of Science* 5, No. 18, 104-119.

LEJEWSKI, C. (1956): *Letter from 23. 5. 1956 to A. N. Prior*. Unpublished manuscript stored in Bodleian Library.

LEŚNIEWSKI, S. (1992a): On the Foundation of Mathematics. In: Leśniewski S.: *Collected works I.* Barnett, D. I. (trans.), Surma, S. J., Srzednicki, J. T. J., Barnett, J. D & Rickey, V. F. (eds.), Dordrecht – Warszawa: Kluwer/Polish Scientific Publishers, 174-382.

LEŚNIEWSKI, S. (1992b): On the Foundation of Ontology. In: Leśniewski, S.: *Collected works II.* O'Neil, M. P. (trans.), Surma, S. J., Srzednicki, J. T. J., Barnett, J. D & Rickey, V. F. (eds.), Dordrecht – Warszawa: Kluwer/Polish Scientific Publishers, 606-628.

LUSCHEI, E. (1962): *The Logical Systems of Leśniewski*. Amsterdam: North-Holland.

MCCALL, S. (ed.) (1967): *Polish Logic 1920–1939*. Oxford: Clarendon Press.

MIÉVILLE, D. (2009): *Introduction à l'oeuvre de S. Lesniewski: Fascicule 2: L'ontologie*. Neuchâtel: Centre de Recherches Sémiologiques Université de Neuchâtel.
PAŚNICZEK, J. (1996): Meinong's Ontology vs. Leśniewski's Ontology: Toward a Meinongian Calculus of Names. *Axiomathes* 7, No. 2, 279-286.
PRIOR, A. N. (1955-1956): Definitions, Rules and Axioms. *Proceedings of the Aristotelian Society* 56, 199-216.
PRIOR, A. N. (1957): *Time and Modality.* Oxford: Clarendon Press.
PRIOR, A. N. (1967): *Past, Present and Future.* Oxford: Clarendon Press.
PRIOR, A. N. (1968): *Papers on Time and Tense*. Oxford: Clarendon Press.
PRIOR, A. N. (1971): *Object of Thought*. Geach, P. T & Kenny, A. J. P. (eds.), Oxford: Clarendon Press.
PRIOR, A. N. (1976): *The Doctrine of Propositions and Terms.* Geach, P. T. & Kenny, A. J. P. (eds.), London: Duckworth.
RICKEY, F. (1998): A Survey of Leśniewski's Logic. In: Srzednicki, J. T. J. & Stachniak, Z. (eds.): *Leśniewski's Systems: Protothetic*. Dordrecht: Kluwer, 23-42.
RUSSELL, B. (1919): *Introduction to Mathematical Philosophy*. London: George Allen and Unwin.
SAGAL, P. T. (1973): On How Best to Make Sense of Lesniewski's Ontology. *Notre Dame Journal of Formal Logic* 14, No. 2, 259-262.
SIMONS, P. (1982): On Understanding Leśniewski. *History and Philosophy of Logic* 3, No. 2, 165-191.
SIMONS, P. (2011): Stanisław Leśniewski. In: Zalta, E. N. (ed.): *The Stanford Encyclopedia of Philosophy*. Available at: http://plato.stanford.edu/archives/win2011/entries/lesniewski/.
SŁUPECKI, J. (1984): S. Leśniewski's Calculus of Names. In: Srzednicki, J. T. J. & Rickey, V. F. (eds.): *Leśniewski's Systems: Ontology and Mereology*. The Hague: Nijhoff, 59-122.
SOBOCIŃSKI, B. (1953): *Letter from 16. 9. 1953 to A. N. Prior*. Unpublished manuscript stored in Bodleian Library.
SOBOCIŃSKI, B. (1984): Leśniewski's Analysis of Russell's Paradox. In: Srzednicki, J. T. J. & Rickey, V. F. (eds.): *Leśniewski's Systems: Ontology and Mereology*. The Hague: Nijhoff, 11-44.
UCKELMAN, S. L. (2012): Arthur Prior and Medieval Logic. *Synthese* 188, No. 3, 349-366.
URBANIAK, R, (2008): Leśniewski and Russell's Paradox: Some Problems. *History and Philosophy of Logic* 29, No. 2, 115-146.
URBANIAK, R. (2014a): *Leśniewski's System of Logic and Foundations of Mathematics*. Cham: Springer.
URBANIAK, R. (2014b): Słupecki's Generalized Mereology and Its Flaws. *History and Philosophy of Logic* 35, No. 3, 289-300.
URBANIAK, R. (2015): Stanisław Leśniewski: Rethinking the Philosophy of Mathematics. *European Review* 23, No. 1, 125-138.

WOLEŃSKI, J. (1989): *Logic and Philosophy in the Lvov-Warsaw School*. Dordrecht: Kluwer.

WOLEŃSKI, J. (1999): Reism and Leśniewski's Ontology. In: Woleński, J.: *Essays in the History of Logic and Logical Philosophy*. Cracow: Jagiellonian University Press, 13-21.

ZUBER, R. (1998): Polish Logic, Language and Philosophy. In: Kijania Placek, K. & Woleński, J. (eds.): *Lvov Warsaw School and Contemporary Philosophy*. Dordrecht: Kluwer, 219-238.

Organon F 23 (2) 2016: 263-274

Manuel García-Carpintero & Genoveva Martí (eds.): *Empty Representations: Reference and Non-existence* Oxford University Press, Oxford, 2014, 368 pages

Proper names (such as 'Obama'), indexicals (such as 'I') and demonstratives (such as 'that dagger') are *singular terms*. The job of singular terms is to pick out objects, and according to current orthodoxy not only the truth but also the meaning of – or *proposition expressed by* – a sentence containing a singular term depends on what object is, in fact, picked out. Insofar as the identity of the singular proposition depends on the identity of the referent, the identity of a singular thought corresponding to (or involving) that proposition depends on the identity of the object thought about as well. An obvious problem with this picture is that some singular terms, such as 'Santa Claus' and 'Vulcan', do not refer to anything but are nevertheless able to occur in sentences that seem perfectly meaningful and sometimes even true ('Vulcan does not exist,' for instance). Moreover, we entertain thoughts about Santa Claus and Vulcan that seem to be genuinely singular, and even in the absence of referents they seem to succeed in fulfilling their representational task. To put the core problem as succinctly as possible: When we characterize thoughts expressed by sentences containing singular terms, we need to involve the objects referred to by those singular terms themselves; yet the fact that these sentences are meaningful even if the terms (apparently) do not refer to anything suggests that the referents themselves are extrinsic to the nature of those thoughts.

The amount of philosophical work done on fiction and on mental and linguistic representation of things that do not exist (if anything) is staggering, and an impressive number of strategies have been tried out. Virtually all of the claims in the above paragraph have been rejected by someone at some point, for instance. *Empty Representations: Reference and Non-existence*, edited by Manuel García-Carpintero and Genoveva Martí, can hardly claim to be comprehensive – there is, beyond the useful introduction, no detailed discussion of Meinongian theories, for instance – but it does cover a number of important contemporary questions, views and strategies deployed to account for such representations. As such, it will be valuable to anyone who wishes to get up to date on the state-of-the-art research on these issues, and indispensable to anyone working on empty names or singular thought. It should be noted, however, that apart from the lucid overview by García-Carpintero in the introduction, the book is perhaps less helpful as an introduction to the issues. Some of the contributions are rather technical, and several are applications or modifications of frameworks that the authors have developed in detail elsewhere – a comprehensive evaluation of which will probably require consulting those other works.

The book is divided into four parts, of which the first, *Foundational Matters: Singular Thoughts and Their Attribution*, concerns, somewhat loosely, general theoretical issues surrounding singular thought, in particular the association between singular thought and *acquaintance* – whether the ability to entertain singular thoughts about x requires that one stands in some special epistemic relationship to x – and how these concerns play out when those thoughts turn out not to refer. In "Transparency and the Context-Sensitivity of Attitude Reports" Cian Dorr mounts a powerful and interesting defense of the idea that although singular beliefs and reports of singular beliefs are *transparent* (that if $x = n$, then one believes that $\varphi(x)$ if one believes that $\varphi(n)$), the intuition that substitution of singular terms does not generally hold can be explained by the context-sensitivity of belief ascriptions. He provides a careful explanation of how context-sensitivity may be invoked to explain such intuitions, and an interesting Kaplan-style semantics for context-sensitive sentences as well as a thoughtful discussion about what the source of the context-sensitivity of belief reports may be. What is perhaps missing is a precise formulation of what the semantic content of a typical belief report actually is when given what he calls a *non-uniform* interpretation, or what an ascription whose embedded clause contains a non-referring term could convey to an audience, but the framework nevertheless provides an interesting starting point for further exploration.

Robin Jeshion, in "Two Dogmas of Russellianism", argues that contemporary Russellians have often failed to address a tension between three theses Russell himself ascribed to: That noun phrases are either directly referential or quantificational, that the only objects to which we can directly refer are those with which we are acquainted, and that there are two ways of thinking about particular objects: singularly or descriptively. Those theses constitute what she calls 'the Russellian trinity' – an agent A is acquainted with *o* iff A can directly refer to *o* iff A can think singularly about *o* – and she endeavors to show that contemporary Russellians who accept that *all* proper names are directly referential, should reject the other two theses. In particular, she argues, first, that perceptual and testimonial information usually considered sufficient to establish reference is often insufficient to make agents acquainted with that reference in any meaningful sense (Russell himself, of course, maintained that we are only acquainted with sense-data, thus avoiding the possibility of error through misidentification, but most contemporary philosophers adopt a more liberal view of acquaintance). Second, some of our referring expressions are empty and so, *pace* realists about mythical characters, resist any sort of acquaintance relation, as do referential terms introduced by purely descriptive stipulation; so-called 'descriptive names'. Of course, the second point assumes (Jeshion offers no argument here) that empty and descriptively used names are genuinely referential names, which is certainly a controversial assumption. The main problem with maintaining that descriptively introduced names are referential, for instance, is not that it threatens a close association between *acquaintance* and reference, but that it threatens to commit us to *a priori* contingent beliefs that should not be *a priori* – it would entail that the true proposition *Whitcomb Judson invented the zip* can be known *a priori* by anyone who knows that *someone* did. If descriptive names were not directly referential they would be no counterexample to an acquaintance condition on reference (though, as Hawthorne & Manley (2012) have argued in detail, an acquaintance condition may not help circumvent the problem either). If descriptive names are not ordinary, referential proper names it is a short step to concluding that empty ones aren't either.

Jeshion goes on to argue against the acquaintance requirement in favor of a view she calls 'cognitivism' (defended in more detail in Jeshion 2010), according to which we think about individuals through object files, where the fact that a mental file is an object file is a matter of the normative and functional role of the file – the *significance* of the object itself to a thought – rather than whether the agent is *acquainted* with the referent of the file. Ken Taylor and François Recanati defend similar conceptions of singular thought in their contributions to this volume; Recanati even retains an acquaintance condition, though his understanding of acquaintance is, at least as compared to Russell's original idea, so diluted that any disagreement with Jeshion on that score may be little more than terminological.

Jeshion also rejects Russell's claim that there are two mutually exclusive and jointly exhaustive ways of thinking about objects in the world – singularly and descriptively – and the arguments for this claim are less developed. The idea that singular thoughts are *object-involving* is dismissed in a paragraph, and the most intriguing version is rejected simply by appeal to incredulity: 'It cannot be literally true that the object itself is a constituent of one's thought itself, the mental particular,' says Jeshion (p. 85), yet that is precisely what defenders of *de re* intentionality are committed to (see McDowell 1986, for instance). Other properties associated with singular thought, *non-descriptiveness* and *directness* are also dismissed without much – or only gestures toward – argument. This is a bit puzzling, since once we get rid of the acquaintance requirement (and perhaps object-involvement) maintaining a clear distinction between singular and descriptive thought would seem all the more natural as a distinction between tracking an object of thought 'directly' or as individuated by the role it plays (individuated by a description), a difference that becomes significant for instance when tracking the objects through counterfactual situations or over time, and perhaps to mark a distinction between whether the identity of the object of thought is determined by factors external to the thinker's mind. Without the acquaintance requirement, whether a thought is a singular thought can be defined purely in terms of its normative and functional role, which would also be straightforwardly reflected in the semantic and syntactic properties of sentences expressing those thought. Counterexamples to a sharp, formally defined distinction usually rely precisely on the lack of a proper epistemic grounding for certain thoughts (and sentences) with an apparently singular form. (Jeshion mentions referentially used description as a worry for maintaining the distinction, but it is at least *prima facie* plausible to think that a referentially used description is a means for handily naming an object *individuated* by description (or demonstration) but *thought about* non-descriptively.) Moreover, rejecting a sharp distinction between descriptive and singular thought would tend to weaken her own arguments against the acquaintance requirement – thoughts expressed by sentences containing descriptive names, for instance, will be evidence against the acquaintance requirement on singular thought only on the assumption that they are, indeed, genuinely singular thoughts. Of course, Jeshion has written about these issues elsewhere, and although the discussion in the current essay is interesting and thought-provoking, one would need to consult those other works to get a full sense of the force of her arguments.

In "Intersubjective Intentional Identity" Peter Pagin discusses the problem of intentional identity in the context of Geach sentences (intentional identity is also discussed by Stacie Friend). His solution draws on his and Kathrin Glüer's development of *relational modality* (see Glüer & Pagin 2006; 2008), which is also the topic of their joint contribution "Vulcan Might Have Existed, and Neptune Not: On the Semantics of Empty Names", which opens part II of the book, *Accounts of Empty Representations*. The latter article develops their own *switcher semantics* for proper names to deal with empty names in a truth-theoretic framework. The basic idea is that proper names are associated with two different intensions, a standard *possibilist* intension that picks out in each world the unique satisfier of descriptive information associated with the name, and an *actualist* intension that picks out the satisfier of that information in the actual world. Modal operators then *switch* the evaluation from the possibilist to the actualist intension. The result is an interesting, essentially two-dimensionalist solution to the problem of empty names flexible enough to respect a variety of *prima facie* conflicting intuitions about such cases. (Note that few two-dimensional systems have dealt with empty names in detail.) That said, the article relies heavily on their previous work, and to those unfamiliar with that work certain moves, such as the introduction of different truth-predicates, may seem *ad hoc*. They are also committed to the controversial idea of treating empty names as picking out possible individuals (see for instance Briggs Wright's article in this volume).

In "Content Relativism and the Problem of Empty Names" Frederick Kroon mounts a powerful defense of the descriptive-proxy account of empty names – that is, a *causal-descriptivist* view adapted to empty names – in particular against the so-called variation problem (that there may be considerable variation in descriptions agents associate with the name) by invoking *content relativism*, (roughly) that what is said by an utterance *u* in a context may vary between the circumstances from which *u* is interpreted. François Recanati's contribution ("Empty Singular Terms in the Mental File Framework") is based on his earlier work on mental files (cf. Recanati 2012) but introduces some more nuance to deal with empty names, in particular a meta-representational function that some mental files have that allow their owners to represent how other subjects think about objects rather than the objects themselves; the result bears some affinities with Frege's reference-shifting account of indirect discourse. Ken Taylor ("The Things we Do with Empty Names: Objectual Representations, Non-Veridical Language Games, and Truth Similitude") discusses three distinctions – between objective and objectual representations (which is similar to Recanati's distinction between thought-vehicles and thought contents), between veridical and non-veridical language games, and between truth and truth-similitude – that together allow him to offer a sophisticated and interesting account of how sentences that fail to express determinate propositions may still carry cognitive significance and be 'correctly assertable'. That I do not discuss the details of the articles by Kroon, Recanati and Taylor should not be taken to indicate that I found them anything but rich, compelling and interesting; they are, however, based on ideas that these authors have defended in detail elsewhere, and it is sometimes difficult to evaluate them independently of this background.

One of the most original and thought-provoking contributions to the collection is Imogen Dickie's "A Practical Solution to the Problem of Empty Singular Thought". Assuming that singular thoughts are genuinely singular (not descriptive) there is a tension between the claims that i) empty singular beliefs are not about objects; ii) for there to be a fact of the matter about what it would take for the singular thought to be true, there must be an object for it to be about; iii) empty beliefs are (often) justified; and iv) there is justification only if there is truth-conditional content. Dickie defends the rather novel idea of rejecting iv). Although she provides some arguments against rejecting ii), she does not discuss in detail the more obvious option of denying iii). Many Russellians would presumably reject iii) (indeed, many of them, such as Taylor above, deny that there even *is* a singular belief involved in empty cases) but argue that, in empty cases, the agent may entertain related descriptive beliefs that are, in fact, justified and genuinely truth-conditional. However, Dickie's suggestion is worth taking seriously; in the article it is only developed in detail for demonstrative thoughts guided by the representational needs of the agent, and further work is needed to extend it to other singular thoughts (those expressed by proper names or other uses of demonstratives).

To develop her argument, Dickie draws on Anscombe's distinction between speculative and practical knowledge (see Anscombe 1957). In particular, she rejects the classical assumption that the sort of relation with objects that is required for singular thoughts is *theoretically oriented* rather than *practically oriented*. She suggests, very roughly, that perceptual demonstrative thoughts fill a basic cognitive (practical) need and are justified iff formed in a way that tends toward securing that the demonstrative element refers to a thing outside the mind (I will have to refer the reader to the article itself for the empirically informed details). Empty singular thoughts, then, are justified insofar as they are formed through processes that reliably secure reference to a mind-independent reference, even if the particular demonstrative thought happens to fail to do so because the world is uncooperative in that particular instance, and even if the resulting thought lacks genuinely truth-conditional content.

Two related worries that Dickie does not address should be mentioned, however. Dickie explicitly assumes that empty, singular, demonstrative beliefs *are* justified, and attempts to develop an account of *how* they can be justified. It is unclear to me, however, whether such beliefs really are the kinds of things that *can* be justified. Empty, singular, demonstrative beliefs may be formed by processes that reliably result in true beliefs, but a belief being justified is a matter of it being *likely to be true* given the available evidence or reliability of processes by which is was formed, which an empty singular belief with no truth-conditional content cannot be (Dickie rather explicitly commits herself to this conclusion on p. 237). Now, this worry may perhaps assume a notion of *truth-likelihood* that is too externalist; perhaps a response could be that these beliefs *seem*, to the agent, to be likely to be true and that this is enough to make them justified. But surely, from an internalist point of view, the belief that the agent thinks is likely to be true is not the empty singular belief but some other, truth-conditional and perhaps descriptive belief also formed on the basis of the perceptual experience (since the empty belief isn't truth-conditional, the agent cannot be thinking that *it* is likely to be true). To bring that point home, consider another, related worry with Dickie's account: She provides no criteria for *distinguishing* empty, singular thoughts. Indeed, one suspects that, on pain of equipping them with descriptive content, there can be no such distinction. But then all empty, singular thoughts are identical, and if one empty, singular belief is justified by being formed in a reliable manner, then all are. Surely, though, the agent *herself* can think that the empty singular beliefs she expresses by 'that$_x$ is *F*' and 'that$_y$ is *F*', respectively, where the corresponding demonstrative beliefs are formed in different circumstances, are different beliefs. And then it seems that the beliefs that are available to the agent for her own cognitive processes or networks of justification are not the singular beliefs themselves, but (perhaps) some descriptively enhanced proxies. And if the singular beliefs do not play any role in the agent's own cognitive economy, why does it matter whether they are *justified*? These worries may, however, merely show that there is work left to do; Dickie's account is intriguing and deserves further attention.

The third section, *Existence and Non-Existence*, concerns the question of what we mean when we assert or deny that something exists. In "What is Existence?" Nathan Salmon offers a typically lucid explanation for why 'existence' should be considered an ordinary predicate, contrary to the view famously espoused by Kant and Russell. Greg Ray, in "The Problem of Negative Existentials Inadvertently Solved", provides a neat solution to the problem of negative existentials like 'Pegasus does not exist' in a truth-theoretic framework. Indeed, the solution falls out of his axioms with such apparent inevitability that Ray is led to wonder why it has been overlooked. The answer is surely that he offers what is essentially a form of wide-scope descriptivist interpretation of names rather than a Millian or traditional descriptivist account, but the straightforward solution to the problem of negative existentials that follows is a potential argument for treating names this way. Ray also shows why wide-scope descriptivism is a natural position in a truth-theoretic framework given natural formulations of the axioms associated with names.

The final section, *Fiction*, deals with an assortment of issues concerning fictional discourse, including the notion of *truth in fiction* and the ontological status of fictional characters: *Realists* claim that fictional characters exist – they are part of the 'furniture of the world', though they are, of course, abstract objects rather than real wizards or hobbits or detectives – whereas *irrealists* claim that the universe contains no such entities. Both views claim some basis in common sense: The realist can make sense of typical literary criticism by deeming for instance 'some fictional characters are better known than others' to be straightforwardly and literally true, but arguably struggles to get 'Sherlock Holmes does not exist' come out true, as it intuitively should. In "Fictional Realism and Negative Existentials", Tatjana von Solodkoff endeavors to show how denials of existence in ordinary discourse should be analyzed to ensure that such claims receive the correct truth-value. To do so, she elaborates on a suggestion by Amie Thomasson (1999) to the effect that '*a* does not exist' should be interpreted as conveying the claim that *a* is not a *K*, where i) *K* is a conversationally salient kind and ii) *a* is fictionally characterized as being a *K* in conversationally salient fiction (p. 337). The task, then, is to provide an apposite value for '*K*', and von Solodkoff provides a thoughtful and compelling argument for interpreting '*K*' as (roughly) *concrete thing*.

In "Fictional Worlds and Fiction Operators" Mark Sainsbury argues against a particular kind of realist alternative, David Lewis's *possibilist* account of fiction (see Lewis 1978). Instead of focusing on the familiar indeterminacy problems that arise when fictional characters are treated as possible objects, Sainsbury argues that Lewis's account is beset by a variety of other problems, perhaps in particular the worry that in order to identify the possible worlds that give us the truth-conditions for the fiction one will first have to determine what the content is, which threatens to give rise to a vicious circularity – at least if Lewis's account is intended to have a particular kind of explanatory power that, I might add, I am not sure Lewis really intended it to have. It is worth noting, though Sainsbury does not, that it seems possible to raise parallel concerns for possible-world semantics for ordinary discourse in general.

Kripke's (2011) objection to treating fictional characters as possible objects has been influential: There is at most one Sherlock Holmes (*uniqueness*), but on a possibilist interpretation there will be a multitude of different possible objects that satisfy the descriptions associated with Holmes in the stories and no satisfactory means for specifying which of these is the referent of 'Holmes' (*multiplicity* and *arbitrariness*). (Notice that non-possibilist realists may face similar challenges.) One may, of course, wonder how forceful this objection is if one follows Jeshion (this volume) and denies that reference presupposes a special epistemic access to the referent: Why not say that the referent of 'Holmes' is the possible object *Holmes*, who is not, and could not be, identical to any *actual* object – even if we do not have any non-arbitrary means for distinguishing between the different candidates that satisfy the descriptions associated with him – and that 'truth according to the Holmes stories' is defined in terms of worlds in which those stories are told as known facts *about him*? Perhaps it is a worry that we aren't appropriately causally linked to such referents, but those who reject the acquaintance requirement will presumably have to adopt a fairly liberal view of causal links as a source of reference fixing to begin with. Now, we would not be able to identify them by distinguishing them from worlds where otherwise identical stories are told as known fact about persons that are *not* Holmes, of course, and would accordingly need to *stipulate* the worlds we are interested in when evaluating sentences for truth-according-to-the-Holmes-stories. But that is presumably how we should think about modal talk about individuals in general: To determine whether 'Aristotle could have failed to teach Alexander' is true, we don't search the space of possible worlds to discover one where an identifiable Aristotle fails to have that property. Rather, we stipulate that we are talking about Aristotle and (roughly) determine whether he can be part of a world where he doesn't have that property; similarly with Holmes.

That response does, in fairness, assume a rather radical anti-acquaintance view. Briggs Wright, in "Many, But Almost, Holmes?", considers a less radical response by noting similarities between Kripke's argument and Peter Unger's *problem of the many* (cf. Unger 1980), which can roughly be illustrated as follows (cf. Lewis 1993): A cloud is an aggregate of droplets. At the outskirts the density of the droplets gradually falls off, with the consequence that it is impossible to tell where the boundaries of the cloud actually are. As a consequence, many different aggregates of droplets are equally good candidates to be the cloud, and we seem unable to say that the cloud is one particular aggregate rather than another. But if all these aggregates count as clouds we have many clouds; and there is just one. Wright explores whether standard solutions to the problem of the many can be used to salvage possibilism from the multiplicity and arbitrariness problems, ultimately concluding that such strategies fail.

There is one response I wonder whether Wright is a bit too quick to dismiss, however. According to *supervaluationism*, 'there is but one cloud' is *super*-true since, despite the multitude of potential candidate clouds, the sentence is 'true under all ways of making the unmade semantic decisions' (Lewis 1993, 31). Similarly, one might suggest, 'there is just one Sherlock Holmes' is super-true, since no matter what we decide on matters not specified in the Holmes-stories, the sentence will ostensibly come out true. Wright's worry is that although the move will superficially circumvent the tension between *uniqueness*, *multiplicity* and *arbitrariness*, it does so at the cost of what a rather paradoxical-looking result: Although the object-language sentence 'there is just one Sherlock Holmes' is true, 'when we examine the meta-theory for that language, we find, paradoxically, that there are *many* things, each of which qualifies as [Sherlock Holmes] on some interpretation of the language' (p. 299). The result may not *formally* be a paradox, but seems to be in tension with ordinary conceptions of uniqueness since assertions of uniqueness will be 'ultimately made true by the existence of *many* things.' And although the worry may perhaps be circumvented in the case of the problem of the many, the common strategies, such as invoking Lewis's notion of almost-identity, are unavailable in the fiction case.

I agree that the result may seem strange. But then, there is something strange about *uniqueness* intuitions about fictional characters in the first place. In her interesting contribution on intentional identity in an irrealist framework, “Notions of Nothing”, Stacie Friend discusses and evaluates various ways to account for the feeling that different thoughts expressed by sentences involving, say, ‘Santa’ may be about the same thing even though ‘Santa’ does not refer, different agents may associate different descriptions with Santa, and the name may have gained traction in a linguistic community without the associated information having any clearly discernible, single origin. One of her points is that in some contexts, such as when a child says that ‘Santa will come tonight’ and a different child that ‘Father Christmas will come tonight’, we feel compelled to say that they are talking about the same thing; yet in other contexts, such as when a historian talks about how Father Christmas came to be associated with Santa, they seem to be represented as distinct due to the distinct origins of the myths. Her irrealist account manages to make sense of these variations in judgment; a realist account, however, including Lewis’s possibilism, may have more trouble resolving these apparently conflicting intuitions.

But what do these considerations do to the uniqueness intuition that drives Kripke’s challenge to Lewis? Consider the unmade semantic decisions regarding Sherlock Holmes. Now, the Sherlock Holmes fictional universe did not end its expansion with Conan Doyle. Holmes has made numerous later appearances – think for instance of the recent TV series set in modern-day London – where many decisions left unmade by Conan Doyle have been made. Different later expansions (and overlook the popular but rather artificial notion of a *canon* often guarded with some fervor by fans of the original stories) take the Holmes stories in different directions, and will often make incompatible semantic decisions regarding elements left unspecified in the original stories. What happens with our intuition that we are still talking about *Holmes*, or that these more recent contributions concern the same character as Conan Doyle’s original stories? It seems that our intuitions about identity or uniqueness start to become shaky and, even more obviously, to context dependent.

In a possibilist framework these incompatible semantic decisions must be reflected by different worlds where different possible people play the role of Holmes. And these possible people must then have been different Holmes candidates all along, even when Conan Doyle wrote his original stories (it’s not like Lewis could be a *creationist* about fictional characters). So the possibilist must deny that Kripke’s uniqueness intuition should be accounted for by there being a single referent for Holmes in the first place, and our (and Friend’s) discussion suggests that such intuitions *ought to* be somewhat shaky, no matter one’s stance on the ontology of fictional characters. The supervaluationist move gives Lewis (and possibly other realists) a means to explain why our judgment that 'there is a unique Holmes' may nonetheless count as true in some contexts: Even if there are several truth-makers for the sentence, no matter how the semantic decisions that remain unmade *relative to some salient class of properties* (which is, of course, context dependent) are decided, the sentence 'there is a unique Holmes' may still be *super-true* and hence assertable. (In other contexts, with a different class of properties, it may not be.) Whether this line of response will ultimately be successful is a matter of debate, of course, but it seems to me a potentially promising explanation for why we seem to harbor uniqueness intuitions even though there obviously (and not only in a possibilist framework) *isn't* one, unique thing that is Sherlock Holmes.

As should be clear the contributions to *Empty Representations* provide ample food for thought, and given that the contributions offer a multitude of intuitively compelling arguments going in very different directions it should be unnecessary to say that it hardly provides the final word on any of the issues discussed. Yet for anyone with an interest in fiction, non-existence, the semantics of empty names, or mental representation in general, it should remain an indispensable reference book in a rapidly developing field for some time to come.

Fredrik Haraldsen
fredrikhar@yahoo.no

References

ANSCOMBE, G. E. M. (1957): *Intention*. Oxford: Blackwell.

GLÜER, K. & PAGIN, P. (2006): Proper Names and Relational Modality. *Linguistics and Philosophy* 29, No. 5, 507-535.

GLÜER, K.& PAGIN, P. (2008): Relational Modality. *Journal of Logic, Language, and Information* 17, No. 3, 307-322.

HAWTHORNE, J.& MANLEY, D. (2012): *The Reference Book*. Oxford: Oxford University Press.

JESHION, R. (2010): Singular Thought: Acquaintance, Semantic Instrumentalism, and Cognitivism. In: Jeshion, R. (ed.): *New Essays on Singular Thought*. Oxford: Oxford University Press, 105-140.

KRIPKE, S. (2011): Vacuous Names and Fictional Entities. In: *Philosophical Troubles: Collected Papers. Vol 1*. Oxford: Oxford University Press, 52-75.

LEWIS, D. (1978): Truth in Fiction. *American Philosophical Quarterly* 15, No. 1, 37-46.

LEWIS, D. (1993): Many, But Almost One. In: Bacon, J., Campbell, K. & Reinhardt, L. (eds.): *Ontology, Causality and Mind*. Cambridge: Cambridge University Press, 23-42.

MCDOWELL, J. (1986): Singular Thought and the Extent of Inner Space. In: Pettit, P. & McDowell, J. (eds.): *Subject, Thought, and Context*. Oxford: Clarendon Press.
RECANATI, F. (2012): *Mental Files*. Oxford: Clarendon Press.
THOMASSON, A. (1999): *Fiction and Metaphysics*. Cambridge: Cambridge University Press.
UNGER, P. (1980): The Problem of the Many. *Midwest Studies in Philosophy* 5, No. 1, 411-467.

CALL FOR PAPERS

The Emergence of Structuralism and Formalism

organized by Catholic Theological Faculty, Charles University
and Institute of Philosophy, Czech Academy of Sciences, v.v.i.

June 24.-26., 2016

(Prague, Czech Republic)

Keynote speakers

S. SHAPIRO
M. DETLEFSEN
M. RESNIK
L. HORSTEN

To submit a proposal, please send a proposal of your paper to the following email conference@ktf.cuni.cz

Proposals for papers should be prepared for anonymous review. Proposals should include: Title and abstract of the paper (maximum 500 words).

If you have inquiries about the conference or about the submission process, please write to conference@ktf.cuni.cz

Deadline for submission: **April 30, 2016**
Notification of acceptance: **May 10, 2016**

The scheduled length of lectures is 30 minutes including approx. 10 minutes for discussion. Selected contributions will be published.

Department of Analytic Philosophy, Institute of Philosophy
of the Czech Academy of Sciences Prague, Czech Republic

cordially invites you

to take part in the 5th Ernst Mach Workshop/Werkstatt Ernst Mach.

The title of this year's workshop is

Contemporary Perspectives on Reductionism

Keynote speakers

ALYSSA NEY
(University of California, Davis)

MARCIN MIŁKOWSKI
(Institute of Philosophy and Sociology of the Polish Academy of Sciences)

It seemed, not a long while ago, that nonreductive approaches in philosophy of science had prevailed. Recently, however, a number of philosophers have re-examined the standard objections against reductionism, and have come to regard it again as viable. We invite the submission of papers contributing to the current discussion on reductionism, including the criticisms of the standard objections and the new arguments based on recent developments in science, from physics and biology to neuroscience.

Please send abstracts of up to 500 words to **emw@flu.cas.cz**
Deadline for abstracts: **May 15, 2015**

Workshop will be held on **June 30 – July 1, 2016**
at the Academic Conference Center of the Institute of Philosophy, Prague.

For more information on both this workshop and the whole Ernst Mach Workshop series, see **www.flu.cas.cz/emw**.

Institute of Philosophy, Slovak Academy of Sciences
and
Slovak Philosophical Association

Modal Metaphysics: Issues on the (Im)Possible IV

August 30-31, 2016
(Bratislava, Slovakia)

Keynote speakers

IRA KIOURTI
GIDEON ROSEN

We invite submissions for a 30 minute presentation followed by 10 minute comments and a 15 minute discussion. Areas of interest might include any aspect of analytic metaphysics, epistemology and logic of modality.

A paper of approximately 3000 words should be prepared for blind review and include a cover page with the full name, title, institution and contact information.

Papers can be submitted in pdf or doc(x) and should be sent to **modalmetaphysics@gmail.com.**

Deadline for submission: **March 30, 2016**
Notification of acceptance: **May 31, 2016**

If you wish to submit a paper, comment on an accepted paper or would need any further details, please, email us to the above address, or visit the conference website **www.metaphysics.sk**

Institute of Philosophy, Slovak Academy of Sciences

Philosophy of Language (I): Semantics of Fictional Discourse

October 20-21, 2016
(Bratislava, Slovakia)

Keynote speakers

STACIE FRIEND (Birkbeck College)
PETR KOŤÁTKO (Czech Academy of Sciences)
ALBERTO VOLTOLINI (University of Turin)

We invite submission for a 30 minute presentation (followed by 15 minute discussion). Topics of interest might include (but are not limited to):

semantics of fictional, mythical and empty names; reference to non-existents; truth in fiction; existence in fiction; fictional entities; fiction and ontology

An extended abstract of approximately 1000 words should be prepared for blind review and include a cover page with full name, institution and contact information. Files can be submitted in pdf or doc(x) and should be sent to ***philolangba@gmail.com***. If you wish to submit a paper, or need further information, please, email us to the above e-mail address.

Deadline for submission: **April 30, 2016**
Notification of acceptance: **June 30, 2016**