PRINCIPIA **22**(1): 189–202 (2018) doi: 10.5007/1808-1711.2018v22n1p189
Published by NEL — Epistemology and Logic Research Group, Federal University of Santa Catarina (UFSC), Brazil.

AGAINST A METAPHYSICAL UNDERSTANDING OF REJECTION

MARIELA RUBIN
BA-Logic Group, UN, ARGENTINA
marurubin@gmail.com

ARIEL ROFFÉ
CEFHIC-UNQ-CONICET, UNTREF, ARGENTINA
ariroffe@hotmail.com

Abstract. In this article, we defend that incorporating a rejection operator into a paraconsistent language involves fully specifying its inferential characteristics *within the logic*. To do this, we examine a recent proposal by Berto (2014) for a paraconsistent rejection, which — according to him — avoids paradox, even when introduced into a language that contains self-reference and a transparent truth predicate. We will show that this proposal is inadequate because it is too incomplete. We argue that the reason it avoids trouble is that the inferential characteristics of the new operator are left (mostly) unspecified, exporting the task of specifying them to metaphysicians. Additionally, we show that when completing this proposal with some plausible rules for the rejection operator, paradoxes do arise. Finally, we draw some more general implications from the study of this example.

Keywords: Rejection • paraconsistent logic • revenge paradoxes

RECEIVED: 16/07/2017 ACCEPTED: 30/01/2018

1. Introduction

One of the richest and most interesting expressions in our language is negation. It would be a bit hyperbolic to state that negation is the cornerstone of arguing. Yet, the idea of discussing whether some proposition (or hypothesis) should be accepted or rejected seems to capture the essence of this activity. We engage in this activity constantly, as philosophers, scientists or even in our daily life. And to reject a claim seems to mean, in some way, to negate its truth.

As logicians, one of the tasks we have is to clarify this expression of rejection or strong negation, and the way it is to be understood or used. We can think of it in an inferential way by showing under what circumstances a negation-statement can be concluded, by specifying what follows from a negation-statement, etc. Or we can understand it model-theoretically by showing the truth conditions for these sorts of claims.

As it is known, different logics contain different notions of negation. For instance, in classical logic, there can be no valuation that assigns value 1 (true) to both a sentence A and its negation $\sim A$. In contrast, many paraconsistent logics allow for a sentence to be both true and false at the same time (this is most clear, for instance, in relational presentations, where it might be the case that both A and $\sim A$ relate to both truth and falsity). In inferential terms, we may see that $A, \sim A \vdash_{CL} B$, while $A, \sim A \nvdash_{PL} B$ (where PL is some paraconsistent logic).

One might ask which of these two negations is a better explication of the strong idea of negation (associated, as we saw, with rejection). However, a prior question is if both negations are *intend* to explicate that notion. Perhaps when one changes logic, one also "changes subject", as in the old Quinean debate. However, as paraconsistent logicians themselves recognize (see for example Priest 1987; 2006), when they want to express a strong idea of negation or rejection, they utilize other kinds of technical devices, such as the expression $\neg A =_{\text{df}} A \rightarrow \bot$.

A different way of expressing a strong idea of rejection in a paraconsistent setting can be found in the Logics of Formal Inconsistency (LFI's; see Carnielli and Coniglio 2016; Carnielli and Marcos 2000; Da Costa 1974). For example, some LFI's include a "classicality" operator $\circ$, with the following truth table:

A	$\circ A$
1	1
$\frac{1}{2}$	0
0	1

Once this classicality (or consistency) operator is introduced into a paraconsistent language, the strong negation of A can be expressed as ($\circ A$ & $\sim A$) (for more on this, the reader may also see Barrio, Pailos and Szmuc, Forthcoming; Omori and Sano 2014).

However, in certain contexts, these new expressions can turn out to be problematic. More specifically, they are problematic when the base language contains a transparent truth predicate and is capable of self-reference (as, for example, the language of arithmetic is). There are, of course, many independent reasons for wanting to have these two features. For example, a transparent truth predicate can function as a generalizing device. Having it inside the language, one is able to state, for instance, that every theorem of arithmetic is true (thus "endorsing" them all at the same time). And one can do this without actually having to state each and every one of these theorems. As is known, Tarski (1936) proved that, if the language into which this new predicate is introduced contains self-reference, and the underlying logic is classical, then paradox ensues. This can be seen, for instance, in the instance of the diagonalization theorem that gives us the liar sentence: $\vdash_{PA} L \leftrightarrow \sim Tr(< L >)$. Sentence L can receive no *classically* stable truth value, and is thus, paradoxical. Many logicians attempt to deal with these issues by changing the underlying logic to a paraconsistent one, in which sentences like L can be both true and false, without trivializing the logic. In fact, paradox-related issues are frequently cited as one of the reasons for adopting paraconsistent logics (see for example, Priest 2006, Beall 2009).

It is in a context such as this, that expressions like $\neg$ or $\circ$ (as defined above) can become problematic. Particularly, some "revenge" paradoxes ensue; for instance, a paradox that looks like a combination of the liar and the Curry paradoxes can be formed with the first of those expressions, as $L' \leftrightarrow \neg Tr(< L' >)$, which by definition of $\neg$ is equivalent to $L' \leftrightarrow (Tr(< L' >) \rightarrow \perp)$. The second expression gives us the paradoxical sentence $L'' \leftrightarrow (Tr(< \circ L'' \; \& \sim L'' >))$.

The lesson seems to be that any attempt to formally characterize a strong idea of negation as a logical operator inside a language that contains self-reference and a transparent truth predicate will run into trouble (at least as long as structural rules are held constant). Particularly, this means that the paraconsistents among us, despite their weakened concept of negation, have the same problems that anybody else has to express rejection in certain contexts. In Berto's words:

> The problem generalizes: *no* sentential operator, \$, that applied to A outputs a $\$A$ which has a designated value only if A doesn't, can be a dialetheic exclusion-expressing device. For then we have the explosive logical consequence: $\{A, \$A\} \models \perp$. We can then always build the relevant revenge sentence, which gives triviality, using \$. (Berto 2014, p.9)

Paraconsistent logicians have attempted to devise a variety of different solutions this problem (see Priest 2006, Beall 2009, and others). One of the latest attempts comes from Berto (2014), where he claims that he can introduce into a language a logical expression that successfully explicates the notion of rejection which does not lead to paradox, even when introduced into a language that satisfies the conditions stated above. His idea, in short, consists of introducing a rejection operator as a defined term, in terms of a primitive second-order operator $\bullet$ (which is considered as already interpreted in the language), intuitively expressing the metaphysical incompatibility between two properties. The goal of this paper is to criticize attempts such as this. We will claim that Berto's solution is inadequate, because it merely exports the problem to metaphysics. That is, until he has a fully developed (and perhaps even operational) notion of metaphysical incompatibility between properties, the paraconsistent logician is in no better position to express rejection than he was before. We believe this discussion has far-reaching consequences not limited to the adequacy of Berto's particular proposal, as it functions as an argument against any kind of attempt to export logical problems to metaphysics, or any other field of philosophy.

From here on, we proceed as follows. The next section reconstructs Berto's proposal. Section 3 formulates some critiques to it. In section 4 we draw some general consequences from this critique. Finally, section 5 contains some conclusions.

2. Berto's proposal

In a recent paper, Berto (2014) claims that there is a way to express a strong idea of rejection within a paraconsistent setting, that avoids revenge. In order to do so, he defines:

(D1) $\underline{A}(x) =_{\mathrm{df}} \exists Q(Q(x) \wedge Q \bullet A)$

Where the intended interpretation of $\underline{A}(x)$ is the rejection of A, and $Q \bullet A$ is supposed to mean that "Q is metaphysically incompatible with A" (he takes this as a term already interpreted in the language). In other words, given an ascription of a (simple or complex) property P to an individual x, what $\underline{P}(x)$ states is that there is another property Q, which x has, that is metaphysically incompatible with P. So, according to Berto, if we read $|P(x)|$ as the set of the objects x that satisfy $P(x)$, then it should be the case that:

(P1) $|P(x)| \cap |\underline{P}(x)| = \emptyset$

There is, however, an ambiguity here. The underscore can be taken to be either an operator (or a predicate) inside the object language, with $D1$ being its *theoretical* definition, or it can be taken to be part of the vocabulary of the metalanguage, but not of the object language, with $D1$ being a *metatheoretical* postulate. Let us dig deeper into this.

If the second path is chosen, then the underscore would be a shorthand way of talking about a *property*. That is, $\underline{P}(x)$ would not be a sentence of the object language. What $D1$ tells us (from the metatheory) is that is that writing $\underline{P}(x)$ is nothing but a shorthand way of writing $\exists Q(Q(x) \wedge Q \bullet P)$. Thus, to express the rejection of a sentence like $P(a)$, inside the language, we would have to say $\exists Q(Q(a) \wedge Q \bullet P)$.

This is inconvenient for several reasons. First off, sometimes we want to reject statements that are not atomic. For instance, suppose we want to reject the statement $A(a) \wedge B(a)$. As Berto defines it, it makes no sense to express this as $\underline{A(a) \wedge B(a)}$, since the underscore is a metatheoretical device that applies to predicates, not sentences. It would also be incorrect to state this rejection as $\exists Q(Q(a) \wedge Q \bullet (A \wedge B))$, since $(A \wedge B)$ is not grammatically well formed. Perhaps the best way of expressing the rejection of $A(a) \wedge B(a)$ would be $\underline{A}(a) \vee \underline{B}(a)$ (provided that De Morgan's laws hold for the logic into which the operator is being introduced?); that is, $\exists Q(Q(a) \wedge Q \bullet A) \vee \exists Q(Q(a) \wedge Q \bullet B)$. But this is somewhat uncomfortable because the sentence has now been turned into a disjunctive one. The same goes if one wishes to reject a quantified statement, such as $\forall x P(x)$. Again, this is not correctly expressed as either $\underline{\forall x P(x)}$, because this is not metatheoretically well formed, or $\forall x \underline{P}(x)$, because we do not want to reject *every* instance of the universal, but only some instance. Thus, we would have to claim something like $\exists x \underline{P}(x)$, again, entirely changing the logical form of the sentence. A different option could be to define two new predicates, $C(x) =_{\text{df}} A(x) \wedge B(x)$, and $R(x) =_{\text{df}} \forall x P(x)$, and then express both rejections as $\underline{C}(x)$ and $\underline{R}(x)$. But this is also uncomfortable because we would have to introduce new defined predicates into the language every time we wish to reject a non-atomic sentence.

On the other hand, if one chooses the first path from above, and decides to incorporate the underscore as part of the vocabulary of the object language, one can do this in two ways. Firstly, the underscore can be treated as an *operator* that applies to sentences (or open formulas). Thus. the rejection of $A(a) \wedge B(a)$ could be written as $\underline{A(a) \wedge B(a)}$ (in prefix notation, this would be $REJ(A(a) \wedge B(a))$). Secondly, the underscore could be treated as a special *predicate*, which would then apply to names of sentences. Thus, in prefix notation, $\underline{A(a) \wedge B(a)}$ should be understood as $REJ(< A(a) \wedge B(a) >)$.

Although we prefer this last path, because it does not have the inconveniences of the other one, we leave open the issue of how to treat the underscore. Our critiques to Berto in section 3 (especially the one in section 3.2) hold independently of the choice one makes. From hereafter, to express the rejection of a sentence $\Phi(x)$ we underscore the entire sentence (i.e. $\underline{\Phi(x)}$) simply because this notation seems clearer to us. Thus, depending on the choice one makes above, a sentence like $\underline{A(a) \wedge B(a)}$ should be understood (in object language terms) as either $(\exists Q(Q(a) \wedge Q \bullet A) \vee \exists Q(Q(a) \wedge Q \bullet B))$, $REJ(A(a) \wedge B(a))$ or $REJ(< A(a) \wedge B(a) >)$.

Moving on, as Berto notes, when this apparatus is introduced into a system with an LP kind of negation $\sim$,[1] P1 entails the following:

(Ent1) For every $x, \underline{P(x)} \models {\sim}P(x)$

This can be seen as follows. Suppose that, for an object o, $\underline{P(o)}$ is true. Then, object o cannot be in the extension of P (because it would then be the case that $P(x)$ and $\underline{P(x)}$). Additionally, because the system under consideration is not paracomplete, it cannot be the case that o is not in the anti-extension of P (that is, in the extension of ${\sim}P(x)$).

However, the following does not hold generally:

(Ent2) For every $x, {\sim}P(x) \models \underline{P(x)}$

Both of these facts make sense at a first glance. Consider an arbitrary object o. If $\underline{P(o)}$ holds (i.e. o possesses some property incompatible with having P), then $P(o)$

must be false, and thus $\sim P(o)$ must be true. However, $P(o)$ might be false without positively having a property that rules out that happening (i.e. without $\underline{P(o)}$ being true).

With this apparatus, according to Berto, it would also be possible to formulate absolute contradictions (AC), which would have the following form:

(AC) $P(x)$ & $\underline{P(x)}$

Since, as we saw, $|P(x)| \cap |\underline{P(x)}| = \emptyset$, then it cannot be the case that AC holds for some x. This way of formulating rejection would also block paradox (or at least the usual derivations of them) in the following way. Consider a possible strengthened liar, defined in the following way:[2]

(UL) $L^* = \underline{Tr(< L^* >)}$

By $(D1)$, what L^* claims then is $\exists Q(Q(< L^* >) \wedge Tr \bullet Q)$. That is, sentence L^* claims that it possesses a property that is incompatible with truth. The existence of this sentence would not lead to paradox, because it can be given a stable truth assignment. If L^* is true or a dialetheia, then we can assert L^*, and we get the AC:

1. L^*
2. $Tr(< L^* >)$ $\quad T$-in, 1
3. $\underline{Tr(< L^* >)}$ $\quad UL$, 1

However, if L^* is just false (and thus we can assert only $\sim L^*$) we do not get AC, since neither $Tr(< \sim L^* >)$ nor $\sim Tr(< L^* >)$ imply $\underline{Tr(< L^* >)}$ (because Ent2 fails). So, there seems to be a consistent assignment of truth values to L^*, one in which it is simply false. This would not entail a contradiction because something could in principle be false, but not positively possessing a property that is incompatible with truth.

The problem that we see with in this solution is that the reason it avoids paradox is simply that the proposal is *too incomplete*. That is, the author introduces into the language an operator $\bullet$, but he does not tell us the ways of legitimately reasoning with it. All that he tells us is that $|P(x)| \cap |\underline{P(x)}| = \emptyset$, *just enough* for the concepts he introduces to express some form of rejection. However, any reasonable account of either rejection or metaphysical incompatibility should be more robust. In the next section, we argue this point further, and show that diverse ways of actually making it more robust do lead to various paradoxes.

3. A critique of Berto

As a first approximation, let us draw an analogy to show more clearly why this proposal is inadequate. If we are allowed (as Berto does), to introduce already interpreted terms into the language, then an obvious question arises: if we want to express the idea of 'rejection', why complicate ourselves with a defined term? Let's just introduce into the language a predicate $R(x)$, "metaphysically interpreted" to mean "the rejection of sentence x". Since we consider it interpreted, we don't give any rules whatsoever for operating with it. Then, it seems pretty much obvious that we won't get a contradiction from the sentence $L^{**} \equiv Tr(< R(< L^{**} >) >)$. If we suppose that L^{**}, by transparency we get that $R(< L^{**} >)$, and there is absolutely nothing relevant we can infer from that because we haven't given rules for operating with R. We assume that no one would find this satisfactory, but it is in essence what Berto does.[3] The only thing that he does claim is that $|P(x)| \cap |\underline{P(x)}| = \emptyset$, but this is a clearly incomplete characterization.

Let us dig deeper into this. When he defends his project, he does say some things about this sort of issue:

> Indeed, I submit that exclusion had better not be defined at all. Exclusion should be taken as a primitive concept with a general metaphysical import. There are reasons for so taking it. First, that there must be primitive notions is uncontroversial: were all notions definable in terms of others, we would face either a bad infinite regress, or a (large) circulus in *definiendo* (on this, see Williamson (2007), pp.50–1). Definitions have to come to an end. There being primitive concepts, no fool-proof decision procedure for them is likely to be available. Many take the concept of set, for instance, as a candidate primitive. We say that a set is an aggregate or collection of objects, but that is no definition. To elucidate the concept, we give examples and hope for the best. (...) There is something in exclusion being characterized as a primitive completely general and, in this sense, metaphysical (contrast logical) feature of our experience the world. For this makes plausible the view that the holding of worldly exclusion relations is (only) ascertainable fallibly and a *posteriori*. (Berto 2014, pp.10–11)

However, there seems to be a slippage here. That a term is primitive within a theory (i.e. that it is not explicitly defined within it), does not mean that it is not regulated by axioms (or rules, sequents, truth-clauses, etc.) within that theory. The idea of presenting an axiomatization is, precisely, to regulate the behavior of terms that are not explicitly defined. We say that the axiomatization is sound when it includes the intended interpretation among its models (i.e. among the acceptable interpretations for the terms, according to the restrictions set by the axioms, rules, sequents, clauses, etc.). On the other hand, it is complete when it is able to pin down that intended interpretation (i.e. other, non-intended interpretations are not models). A theory that puts too few restrictions in place will almost always be sound, but at the cost of being too incomplete, since it will leave almost nothing aside. In the same way, a term characterized by a theory (though not defined in it) will plausibly express its intended interpretation if and only if the theory itself puts enough restrictions on the interpretation of that term. What we hold is that the underscore operation does not successfully express denial because the only "axiom" set in place is too weak, it leaves too many non-intended interpretations in.

On the other hand, incompatibility may indeed be a relation to which we have epistemic access only empirically, as Berto claims. What this would mean, is that in order to *apply* the theory (or "logic") of deniability being formulated here we must take empirical considerations into account. In the same way, in standard propositional logic, determining if an argument is solid requires that one knows if the premises are true, and that may be an empirical matter. However, we can formulate a "logic" (we prefer the term theory) of rejection, which deals with the way in which we should *reason* with rejection claims, regardless of the "content" of those claims. Therefore, postulating that rejection is an empirically determinable concept does not excuse us from formulating more robust requisites for reasoning with rejection claims, especially if our project is to elaborate this "logic of rejection".

In the following subsections we look at some restrictions that should intuitively hold for a rejection operator as defined by Berto, and show a number of ways in which they lead back to paradox.

3.1. Axiomatizing rejection

In this subsection, we try to give some rules directly for Berto's underscore operator. We show that some rules that should plausibly hold involve the reintroduction of paradoxes. First, a plausible introduction rule could go something like this: "If assuming that an object o has property P leads to an Absolute Contradiction, then we can infer that A has the property of being incompatible with P, namely $\underline{P(o)}$".

(REJ-In) If $P(o) \vdash AC$, then $\vdash \underline{P(o)}$ (for every P, and for every o)

That is, if supposing that o has property P leads to affirm that some object possesses a property Q and some other property incompatible with having Q, then o has some property incompatible with having P (we may think of that property being something like "leading to an absolute contradiction"). A nice thing about this rule is that Ent2 remains generally invalid. Now, with this rule in mind, sentence L^* (defined above) does lead to paradox. The demonstration would go as follows:

1. $Tr(< L^* >)$	Supposition
2. L^*	2, T-Out
3. $\underline{Tr(< L^* >)}$	2, UL
4. $Tr(< L^* >) \,\&\, \underline{Tr(< L^* >)}$	1,3, &-in

And we now have that $Tr(< L^* >) \vdash AC$. Then, by the REJ-in rule, we get that:

5. $Tr(< L^* >)$	1-4. REJ-in
6. L^*	5, UL
7. $Tr(< L^* >)$	6, T-In
8. $Tr(< L^* >) \,\&\, \underline{Tr(< L^* >)}$	5,7, &-in

Since supposing that $Tr(< L^* >)$ is true leads to an AC, we get, by &-in, that $Tr(< L^* >)$ is strongly false (i.e. $Tr(< L^* >)$ must have a property that makes it incompatible with being true). But this leads again to an AC (the demonstration would be identical to the one above, from step 5 onwards). Notice that this demonstration is only possible with sentences such as L^*, not with any ordinary statement such as $P(a)$ & $\sim P(a)$.

A possible reply to this would be that the REJ-In rule we proposed is not really that intuitive (i.e. it doesn't preserve truth that clearly). Indeed, it might be held that what actually follows from $P(o) \vdash AC$ is not that $\vdash \underline{P(o)}$ (that o has some property that is incompatible with having P), but only that $\vdash \sim P(o)$ (i.e. that $P(o)$ is false). But this would be a strange reply for a paraconsistent logician, since what $\sim P(o)$ expresses for him is that $P(o)$ is either false *or a dialetheia*. But if it were a dialetheia, then there should be no reason for it to lead to an absurd claim (an AC). In other words, $P(o) \vdash AC$ must mean that $P(o)$ is *just* false, that is, $\underline{P(o)}$.

3.2. Incompatibility

On the other hand, a supporter of Berto's proposal may not be entirely convinced by the arguments in 3.1. As said before, he may not take the underscore as part of the theoretical vocabulary, and thus the may claim that giving axioms for it is not necessary. Unfortunately for this supporter, there seems to be an even more direct way of obtaining a paradox directly from Berto's binary relation $\bullet$. Let the liar be, as before, the sentence $L^* = \underline{Tr(< L^* >)}$; that is,

$$L^* = \exists Q(Q(< L^* >) \wedge Q \bullet Tr)$$

Supposing that L^* is either true or a dialetheia leads to an absolute contradiction in the same way as before. So, exactly as before, L^* must be just false. This means that $\exists Q(Q(< L^* >) \wedge Q \bullet Tr)$ must be just false, and therefore, that there is no property Q such that Q applies to $< L^* >$ and is incompatible with the property of being true. Now, let P be the property of "$\underline{\text{being true}}$ iff being L^*" (i.e. the property of being the Liar*). That is, let

$$P(x) =_{\text{df}} (L^* = \underline{Tr(x)})$$

Then, $P(< L^* >)$ is true, because the instance $P(< L^* >)$ is just $L^* = \underline{Tr(< L^* >)}$, and that is true by definition of L^*. We also know, from the prior demonstrations, that being L^* is *logically* incompatible with being true; and since logical incompatibility is stronger than metaphysical incompatibility, then being L^* must also be *metaphysically* incompatible with being true (whatever this means). We have then that $P(< L^* >)$ and $P \bullet Tr$. Thus, by definition of the underscore operation, this means that $\underline{Tr(< L^* >)}$ is true. And again, by definition of L^*, this means that L^* is true, but this is absurd since we claimed that L^* must be just false.

Another way to see this is the simple derivation:

1. $P(< L^* >)$ & $P \bullet Tr$	By the reasoning above
2. $\exists Q(Q(< L^* >) \wedge Q \bullet Tr)$	$\exists$-in, 1
3. $\underline{Tr(< L^* >)}$	By definition of the underscore, 2
4. $\underline{L^*}$	3, UL
5. $Tr(< L^* >)$	4, Transparency
6. $\underline{Tr(< L^* >)}$ & $Tr(< L^* >)$	3,5, &-In

More intuitively, if L^* says "I have a property that is metaphysically incompatible with truth", and L^* is just false, then there should be no such property. But the property in question is just "being the liar* sentence". The liar* sentence does have the property of being the liar* sentence, and — as we saw — being the liar* sentence is (logically, it seems, and therefore also metaphysically) incompatible with truth.

What we are doing here is to implicitly utilize a rule governing the binary operator $\bullet$. This rule can be reconstructed as follows:

(L-M) If for every $x, P(x)$ & $Q(x) \vdash \bot$, then $\vdash P \bullet Q$

What it states is simply that logical incompatibility implies metaphysical incompatibility, which also seems like a plausible restriction to put in place regarding incompatibility. Utilizing this rule more explicitly, our argument from above can be reconstructed in the following way:

1. $P(x) =_{\text{df}} (L^* = \underline{Tr(x)})$	Definition
2. $P(< L^* >)$	1, UL
3. $P(< L^* >)$ & $Tr(< L^* >) \vdash \bot$	Demonstrated in section 3.1.
4. for every $x, P(x)$ & $Tr(x) \vdash \bot$	
5. $P \bullet Tr$	4, L-M
6. $P(< L^* >)$ & $P \bullet Tr$	2, 5, &-In
7. $\exists Q(Q(< L^* >)$ & $Q \bullet Tr)$	6, $\exists$-In
8. $\underline{Tr(< L^* >)}$	7, UL

And, from here, the demonstration follows in the same way as in the one above.[4] Notice that step 4 follows because of 3, and the fact that for every $x \neq < L^* >$, $P(x)$ is false (and the premise in the inference is false, thus, not constituting a counterexample to the validity of the inference).

There are other ways of completing Berto's apparatus, which also lead to problems. For instance, Omori and De (2017) recently developed a modal account of this issue. They interpret "reject *A*" to be a kind of negation which is true at a world *w* "just in case all worlds where the sentence is true are incompatible with *w*" (De and Omori 2017, p.1). By doing this, the paraconsistent logic ends up in several crossroads. First, they show how this negation satisfies contraposition, which forces the paraconsistent logician to drop weakening in order to avoid EFQS. The problem with dropping weakening is that one should have independent reasons to take structural rules out of the logic others than just avoiding some other rule. To be fair, Berto does recommend dropping weakening for independent reasons, yet the weakening free logic shouldn't be motivated by our understanding of negation, and therefore, one restriction shouldn't impose the other one.[5]

We now turn our attention to some more general lessons that can be extracted from this discussion.

4. Some general lessons

The general lesson that we believe can be extracted from this discussion can be introduced by considering a possible reply to our objections to Berto. This reply consists of arguing that having a property is not equivalent to satisfying a sentence with a free variable, because not every sentence expresses a "metaphysically substantial" property. It seems that Berto would not actually be willing to give this reply, since he asserts that:

> Talk of properties should not be taken as metaphysically too committing (we could in fact rephrase the view in a strictly nominalistic but more cumbersome fashion). We mean by "property" what Field has called conceptual properties, and have as our background a naïve property theory: "there is a conceptual property corresponding to every intelligible predicate" (Berto 2014, p.13)

In the above quote, he seems to indicate precisely that satisfying any sentence with a free variable should be taken as satisfying a property (in the metaphysically deflated sense of 'property' that he speaks of). But if this is the case, then our objections (especially the one presented in 3.2) do mean trouble for him.

However, let us think of an imaginary logician who is willing to give the reply introduced above. What would be missing, then, is an account of what properties count as "metaphysically substantive", in order to restrict the *Q*'s with which the $\exists Q$ can be instantiated. Until one has that, it cannot be claimed that the proposal is free from trouble, since there is no way to know if properties such as "being the liar* sentence" are legitimate or not. Of course, our rival may simply assert that "being the liar* sentence" is not a "metaphysically substantive" property, but that would just be an *ad hoc* reply. There needs to be some reason for rejecting it, as well as for accepting any other.

In other words, the general lesson seems to be this: one can pretty much always avoid paradoxes, simply by recurring to expressions that are considered as already interpreted within the language. Using an analogy, the same would happen with classical logic + arithmetic and the Tr predicate. If one merely states that Tr is an already interpreted predicate, which means "true in the standard model", and does not give any rules, or clauses, etc. for operating with it, then of course paradox will not ensue.

Even other kinds of problems could potentially be "solved" by operating in this way. For instance, consider the inability of standard, first order arithmetic (take, for instance, PA) to capture its intended model (see the discussion by Barrio 2014, Da Re 2014, and Roffe 2014). It is widely known that, because of metatheorems concerning isomorphisms, and others like the Löwenheim–Skolem and Gödel theorems, the models of PA include much more than its single intended model. If one simply considered the expressions of PA to be already interpreted (to mean those of the intended model), then the problem disappears. But this has a cost. The reason for doing formal arithmetic is, precisely, to be able specify the legitimate ways of reasoning using those terms, instead of doing naïve arithmetic (with already interpreted natural language terms). The same goes here. To elaborate a "logic" (or a theory) of rejection requires that we go beyond already interpreted naïve terms. Otherwise, despite the fact that paradoxes do not accrue, the project simply does not have the same value.

To be sure, not every problem will be tractable in a formal-logical manner; for instance, there are some notions whose application is irreducibly pragmatic, not subject to a codification by rules (such as the notion of an intended application for empirical scientific theories, see for example Moulines 2002). But the moment that we postulate that, our job as logicians (although perhaps not as philosophers) is over. In any case, one has to distinguish between irreducibly pragmatic expressions (and for which we have independent reasons to think this is the case), and expressions which are clearly governed by some "logic", but that are problematic in some contexts (for instance, rejection).

5. Conclusions

In this article, we have considered one of the latest attempts to express a strong idea of rejection in a paraconsistent setting that, supposedly, does not lead to paradox. We have shown that this proposal is inadequate on various fronts. First, we showed that if one attempts to "complete" the proposal with some plausible rules (such as the

REJ-In rule), then paradox does arise again. In the following subsection we argued that, if one allows properties such as "being the liar* sentence" to count as acceptable instances of the existential in the definition of the underscore operator then, again, paradox ensues (this time, without adding any new rule). Finally, in section 4 we drew some general conclusions about attempts of solving a problem within logic by exporting it to a different field of philosophy (for instance, metaphysics). We claimed that logical problems (concerning the legitimate ways to *reason* with some expressions) must be dealt with in a logical fashion, by showing the rules, or clauses, etc. that govern the inferential use of such expressions. Merely claiming that rules need not be given because the expressions in question are already interpreted amounts to removing all the value in elaborating a "logical theory" about those expressions.

References

Barrio, E. 2014. Lógica de segundo orden y el modelo estándar de la aritmética. *Cuadernos de Filosofía* **62**: 43–54.

Barrio E.; Pailos F. M.; Szmuc D. 2017. A Paraconsistent Route to Semantic Closure. *Logic Journal of the IGPL* **25**(4): 387–407.

Beall, J. C. 2009. *Spandrels of Truth*. Oxford University Press.

Berto, F. 2014. Absolute Contradiction, Dialetheism, and Revenge. *Review of Symbolic Logic* **7**(2): 193–207.

Carnielli, W. A.; Coniglio, M. E. 2016. *Paraconsistent Logic: Consistency, contradiction and negation*. Springer.

Carnielli, W. A.; Marcos, J.; de Amo, S. 2000. Formal Inconsistency and Evolutionary Databases. *Logic and Logical Philosophy* **8**: 115–152.

da Costa, N. C. A. 1974. On the Theory of Inconsistent Formal Systems. *Notre Dame Journal of Formal Logic* **15**: 497–510.

Da Re, B. 2014. El modelo estándar de la aritmética: recursividad y lógica de primer orden. *Cuadernos de Filosofía* **62**: 55–64.

De, M.; Omori, H. 2017. There is more to negation than modality. *Journal of Philosophical logic*. DOI: 10.1007/s10992-017-9427-0

Moulines, C. U. 2002. ¿Dónde se agazapa la pragmática en la representación estructural de las teorías? In: J. Diez; P. Lorenzano (eds.) *Desarrollos actuales de la metateoría estructuralista*. Universidad Nacional de Quilmes.

Omori, H.; Sano, K. 2014. da Costa meets Belnap and Nelson. In: R. Ciuni, R.; H. Wansing; C. Willkommen (eds.) *Recent Trends in Philosophical Logic*. Springer.

Priest, G. 2006. *In Contradiction: A Study of the Transconsistent* (2nd ed.). Oxford University Press.

———. 2006a. *Doubt Truth to Be a Liar*. Oxford: Oxford University Press.

Roffe, A. 2014. Sobre la capturabilidad de teorías informales en sistemas axiomáticos formales. *Cuadernos de Filosofía* **62**: 65–75.

Tarski, A. 1936. Über den Begriff der logischen Folgerung. In: *Actes du Congrès International de Philosophie Scientifique*, fasc. 7. Hermann et Cie.

Notes

[1] Remember the three facts hold for ∼: If *A* is true, ∼*A* is false; if ∼*A* is false, *A* is true; and if *A* is a dialetheia, ∼*A* is also a dialetheia.
[2] Again, the formulation is slightly different than it appears in Berto's text, because he uses $\underline{L}$ as a name, not as a means of applying _ to a sentence L, which is a bit confusing.
[3] To be fair, Berto does not only want to be able to *express* rejection inside a formal language, he also wishes to *explicate* rejection in terms that do not involve rejection itself. Introducing an already interpreted rejection term into the language would accomplish the first, but not the second of these goals. The analogy is intended only to show that introducing already interpreted terms into the language is not a satisfying way of dealing with the first problem.
[4] All the rules we considered in this section are introduction rules. However, we do not claim that the behavior of Berto's operator(s) should *necessarily* be characterized by introduction rules. All we show is that, if one adds some rules which seem plausible (and which happen to be introduction rules) to the system, paradox arises again. We are open to the possibility that the operator be characterized via other kinds of rules. In any case, the burden of the proof is on Berto's side to provide those rules, not on ours.
[5] For a fuller understanding of the issue and an interesting proposal for a paraconsistent negation, see De and Omori (2017).