PPGFil
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA
perspectiva filosófica
Universidade Federal de Pernambuco

v. 52 n. 2 (2024): Perspectivas Contemporâneas em Filosofia da Lógica

Submetido em: 29 jul. 2024
Aceito em: 21 fev. 2025

https://doi.org/10.51359/2357-9986.2025.263787

Paraconsistent Paradoxes: the price of consistency

Paradoxos Paraconsistentes: o preço da consistência

Rafael Ongaratto[1]
Universidade Estadual de Campinas (UNICAMP)

ABSTRACT

Logics of Formal Inconsistency (LFIs) are usually considered a philosophically neutral logic with respect to paraconsistency, in the sense that they provide a good basis to think in terms of dialetheias, i.e. true contradictions, or in terms of conflicting information, a notion weaker than truth. In this article, I will show how this claim of neutrality fails in the face of the Liar Paradox: in classical logic, the sentence A: "A is false." implies a contradiction. By adopting a paraconsistent logic, we avoid that contradictions lead to triviality. LFIs, however, add a consistency operator that recovers classical logic to propositions to which it applies. In this context, the Liar Paradox arises once again by a type of Liar's Revenge: now, we can construct the sentence B: "B is only false.", and that gives us triviality once again. With this result, we conclude that LFIs are not suitable logics to deal with semantic paradoxes.

Keywords: paraconsistent logics. the liar paradox. logics of formal inconsistencies. dialetheism.

RESUMO

Lógicas de Inconsistência Formal (LFIs) são usualmente consideradas filosoficamente neutras com respeito à paraconsistência, no sentido que elas fornecem uma boa base para pensar em termos de dialetheias, i.e. contradições verdadeiras, ou em termos de informações conflitantes, uma noção mais fraca do que a verdade. Neste artigo, mostrarei como essa alegação de neutralidade falha frente ao paradoxo do mentiroso: na lógica clássica, a afirmação A: "A é falsa." deriva uma contradição. Ao adotar uma lógica paraconsistente, nós evitamos que contradições levem à trivialidade. LFIs, no entanto, adicionam um operador de consistência que recupera a lógica clássica para as afirmações às quais o operador se aplica. Neste contexto, o paradoxo do mentiroso surge mais uma vez por um tipo de vingança do mentiroso: agora,

[1] E-mail: ongarattorafa@gmail.com. ORCID: https://orcid.org/0000-0001-5303-9054. This study was financed, in part, by the São Paulo Research Foundation (FAPESP), Brasil. Process Number 2024/00807-6, and Process Number 2020/16353-3.

podemos construir uma afirmação B: “B é somente falsa.”, e ela leva à trivialidade novamente. Com esse resultado, concluímos que LFIs não são lógicas adequadas para lidar com paradoxos semânticos.

Palavras-chave: lógicas paraconsistentes. paradoxo do mentiroso. lógicas de inconsistência formal. dialeteísmo.

1 Introduction

Paraconsistent logics are logics in which the Principle of Explosion (PE) fails to hold. PE is a principle in classical logic that says that the presence of a sentence α and its negation $\neg\alpha$ imply triviality.

$$\text{(PE) } \alpha, \neg\alpha \vdash \beta, \text{ for all sentences } \alpha, \beta$$

One of the advantages of adopting a paraconsistent position is to avoid undesirable results from semantical paradoxes such as the Liar Paradox: the Liar sentence, in this case, is both true and false, and this does not mean that every sentence is true. In fact, this is one of the main motivations for dialetheism, a philosophical position that holds that there are true contradictions. Priest, in a seminal paper presenting the Logic of Paradox (the name already points out its objective), makes it explicit that the paraconsistent logic he develops is a logic that accepts paradoxes, that is, a logic that accepts (some) contradictions as true, and that this is the *raison d’être* of LP:

> The purpose of the present paper is to suggest a new way of handling the logical paradoxes. Instead of trying to dissolve them, or explain what has gone wrong, we should accept them and learn to come to live with them. [...] For obvious reasons, this will require the abandonment, or at least modification, of classical logic. (Priest, 1979, p. 219)

The Logic of Paradox (LP) is a three-valued logic with the values 1 (“only truth”), 0 (“only false”), and 1/2 (“both true and false”). The idea is that dialetheias are considered acceptable because they receive the truth value 1/2 (Priest, 1979, p.

226).[2] Nevertheless, this does not mean that all contradictions are now allowed to be accepted as *dialetheias*. For example, if someone says "John is tall" and someone else says "John is not tall", this may not be a true contradiction: by measuring John, the two people in conflict conclude John is 1.75m tall. After this measurement, they both agree that John is not tall.

Thus, contradictions that can be solved into the affirmation or the negation of the sentence by further analysis are not true contradictions. That is why the Liar Paradox is a strong candidate for a true contradiction: it has resisted millennia as an *insolubilia*. Before introducing the Liar Paradox, however, let us make a few remarks about how to represent the Liar Paradox in a formal language.

The Liar Paradox is a self-referential sentence that asserts its own falsity: "This sentence is false". The first step towards expressing the Liar sentence, thus, is to adopt a first-order language, that is, a language that has names of individuals, and not only labels for sentences as is the case of propositional logic.

Before formalizing such notions, it's also worth mentioning another paradox that appears in this context of LP, which is Curry's Paradox. Curry's paradox is also about a self-referential sentence that says "This sentence implies its own triviality", and it also trivializes classical logic due to the behavior of the conditional.

In the next section, we will look into both paradoxes formally in more detail. In the third section, we will analyze LP and its first-order extension. In the fourth section, we discuss the results presented in the previous sections philosophically. In the fifth section, we will analyze weaker paraconsistent logics in the face of the Liar Paradox. Finally, section 6 concludes by considering the relation between LFIs and LETs and the Liar. It's worth mentioning the seminal papers that have been the foundation for thinking about LFIs: (Marcos, 2005; Carnielli, Coniglio e Marcos, 2007).[3]

---

[2] In Priest's presentation, he prefers to call those 3 values as t, f, and p. However, we will adopt numerical notation here, which is more used in the contexts of LFIs.

[3] I thank Mahan Vaz for his detailed revision and discussion of the present paper.

2 Preliminaries

In this section, we will consider definitions and notations that will be useful to represent paradoxes in a first-order paraconsistent logic. Let $\mathcal{L}$ be a language formed by a signature $\Sigma = \{\wedge, \vee, \rightarrow, \neg, \forall, \exists\}$, and a set $\mathcal{U}$ of propositional letters. The set of formulae of $\mathcal{L}$ is defined as usual. The logic L is the ordered pair ($\mathcal{L}$, $\models_L$), in which $\models_L$ is a relation of logical consequence.

Following the construction of Barrio, Pailos and Szmuc (2017, p. 4), the Liar Paradox requires two ingredients: we need a truth predicate in the language, which will be the unary predicate T; furthermore, we have to extend our language to include names for all sentences.

Let $\mathcal{A}$ be a structure (A, I) for $\mathcal{L}$, in which A is a nonempty set, and I is an interpretation function of $\mathcal{L}$ over A. The extension of $\mathcal{L}$ with names for each sentence of the language will be denoted $\mathcal{L}'$, and each new name c_ϕ will be interpreted thus in the structure $\mathcal{A}'$:

$$I'(c_\phi) = \phi, \text{ for all } \phi \in A$$

The other semantical notions of consequence and tautology are defined as usual. Now, we have the means necessary to express a sentence such as "ϕ is true". We have the predicate T and the name of ϕ, c_ϕ, therefore, this sentence is formalized as Tc_ϕ.

Furthermore, the T predicate also has to obey some conditions. As we have briefly seen, the paraconsistent solution to the Liar Paradox involves accepting the paradox as a true contradiction, which means that we do not need to change our underlying theory of truth to avoid the Liar. That is, by accepting a paraconsistent solution to the liar, we are in an advantageous position in which we can still defend a naive theory of truth (Barrio, Pailos e Szmuc, 2017, p. 1).

(TRANSPARENCY) ϕ and Tc_ϕ are intersubstitutable "in every non-opaque context"

Furthermore, we also have to commit to the T-schema: ϕ and Tc_{ϕ} are equivalent, that is, $\phi \leftrightarrow Tc_{\phi}$. Finally, we need to include self-reference in the language to produce the Liar. There are various ways of achieving self-reference, as presented in (Smullyan, 1994). In the present case, however, we do not need to worry about which method we use to achieve self-reference, because the solution to the Liar Paradox already supposes that the ingredients to the Liar Paradox are within the expressive power of the language: semantical closeness, self-reference, universality, all ingredients which are already found in the natural language.

Def. 1 *Let K be the closure of L' over $\vDash_{L'}$. K is self-referential if for every ϕ(x), there is a term c that is identical to the name of ϕ(c).*

This definition of self-referentiality is very strong, as Barrio, Pailos and Szmuc (2017, p. 5) point out, by distinguishing between strong and weak self-referential procedures. Barrio et al. choose to pursue a weak procedure that is satisfied by paraconsistent theories. In the present paper, however, we will push to the limit paraconsistent theories by considering the stronger notion of self-referentiality as presented in definition 1.[4] Now, it is possible to construct the Liar sentence. Consider $\phi(x) := \neg T(x)$. Suppose also that K, the theory under consideration, is self-referential. By def. 1, there is a c such that c is identical to the name of $\neg T(c)$. Therefore, c says that "c is not true".

3 LP, LP°, and QLP°

The Logic of Paradox is a logic designed to solve antinomies by incorporating a third truth value. The idea is that propositions can be only true, only false, or both true and false. The propositions that are true and false are the *dialetheias*. As we

[4] We are not pursuing the weaker definition of self-reference, because many paraconsistent systems such as mbC and Cio are too weak to define a congruence relation (Carnielli e Coniglio, 2016, p. 132).

have emphasized before, not all contradictions are true contradictions, only those that are unsolvable conflicts. Let M be the set of values {1, 1/2, 0}, in which 1/2 is the paradoxical value, and D = {1, 1/2} be the set of designated values. The matrix $\mathcal{M}_{\mathcal{LP}}$ = (M, D) of LP is defined as below, presented in the table for the signature with conjunction and paraconsistent negation (Carnielli e Coniglio, 2016, p. 150):

$\wedge$	1	1/2	0
1	1	1/2	0
1/2	1/2	1/2	0
0	0	0	0

	$\neg$
1	0
1/2	1/2
0	1

The extension of LP to its first-order version, QLP, can be done by extending the signature of LP to Σ,[5] and defining the interpretation I to the set $\mathcal{R}$ of atomic formulae $R^n(t_1, t_2, \ldots, t_n)$ as I: $\mathcal{R} \mapsto \{1, 1/2, 0\}$ (Priest, 1979, p. 229). The rules for $\forall$, $\exists$ and the other connectives are defined as usual by recursion using the truth tables above.

Finally, the Liar Sentence can receive a formulation in QLP. Consider the sentence ϕ(x) := $\neg$T(x), and suppose that $\mathcal{L}_{\mathcal{LP}}$ is a self-referential language. By def. 1, there is a c such that c is identical to the name of $\neg$T(c). We will show that $\neg$T(c) is satisfiable.

Suppose that $\vartheta(\neg T(c)) = 1/2$. Given that c is identical to this sentence and that TRANSPARENCY holds for T, $\vartheta(T(c)) = 1/2$. This is a possible valuation of the liar sentence that is guaranteed by the valuation matrix of $\neg$: $\vartheta(\alpha) = 1/2$ iff $\vartheta(\neg\alpha) = 1/2$. Thus, the standard version of the Liar Paradox is tamed in QLP.

Curry's Paradox is also avoided in QLP,[6] due to its nondetachable conditional (Tajer, 2020, p. 8). As $\alpha \rightarrow_{LP} \beta$ is defined as $\neg(\alpha \wedge \neg\beta)$, the conditional in LP behaves as follows:

We can see that $\rightarrow_{LP}$ does not validate *modus ponens*: take, for example, the

[5] The symbols $\vee$ and $\rightarrow$, $\forall$ can be defined from $\neg$, $\wedge$, and $\exists$ as usual.
[6] I thank the anonymous reviewers for the suggestion to include the Curry Paradox and to present the Liar in a first-order theory.

$\to_{LP}$	1	1/2	0
1	1	1/2	0
1/2	1	1/2	1/2
0	1	1	1

valuation $\vartheta(\alpha)$ = 1/2 and $\vartheta(\alpha \to_{LP} \beta)$ = 1/2. In that case, $\vartheta(\beta)$ = 0, which means that α, $\alpha \to_{LP} \beta \not\models_{LP} \beta$. Now, consider a version of Curry's Paradox which says "This sentence implies that I am Superman".

In classical logic, we would reason as follows: if this sentence is true, then, by modus ponens, it implies that I am Superman. If this sentence is false, then the whole conditional is true, which means that the sentence implies that I am Superman.

However, in QLP, we would represent this sentence formally as c$_{\phi(c)}$ such that c$_{\phi(c)}$ is identical to the name of Tc$_{\phi(c)} \to_{LP} \beta$. By TRANSPARENCY, Tc$_{\phi(c)}$ should have the same value as $\phi(c)$, that is, Tc$_{\phi(c)}$ should have the same value as Tc$_{\phi(c)} \to_{LP} \beta$. Is there a situation in which they both hold the same the same value without making β also designated? Take $\vartheta(\phi(c))$ = 1/2. Then, ϑ(Tc$_{\phi(c)}$) = 1/2. Suppose that $\vartheta(\beta)$ = 0. Then ϑ(Tc$_{\phi(c)} \to_{LP} \beta$) = 1/2. That is, TRANSPARENCY is satisfied with β being false. Therefore, the paradox is avoided in LP.

Although QLP avoids the embarrassment of classical logic in the face of the Liar Paradox and Curry's Paradox, it lacks expressiveness: QLP separates the propositions into 3 values: only true, only false, and paradoxical. Can we assert, for example, that "α is only true" in QLP? If we tried to formalize "α is only true" as Tc$_{\alpha}$, then there would be ϑ such that ϑ(Tc$_{\alpha}$) = 1/2: take $\vartheta(\alpha)$ = 1/2. However, "α is only true" should be false if α is paradoxical. Therefore, we cannot express "α is only true" simply by taking the truth predicate in QLP.

In general, QLP cannot express the fact that a proposition is consistent, that is, the fact that it is not paradoxical. Following Carnielli and Coniglio (2016), we define the unary consistency operator as $\circ$.

Theorem 1 *Let $\phi(\alpha)$ be a formula that depends only on the formula α. LP does not define $\phi(\alpha) := \circ\alpha$.*

	$\circ$
1	1
1/2	0
0	1

The proof is a standard proof by induction on the complexity of the formula, and it is left to the reader. Notice that defining such a formula would require that $\vartheta(\alpha)$ = 1/2, and $\vartheta(\phi(\alpha))$ = 0, for some valuation ϑ. However, $\vartheta(\neg\alpha)$ = 1/2 and $\vartheta(\alpha \wedge \alpha)$ = 1/2. Any iteration of the connectives to α, therefore, will result in the value 1/2. The idea is that in LP, we cannot go from the value 1/2 to 0 using the connectives at our disposal.

Therefore, consider an extension of QLP, which we call QLP°, as QLP extended with the connective $\circ$. Now, in QLP°, we can state that a sentence is only true and that a sentence is only false, as $\phi(x_p)$:= T(x_p) $\wedge$ $\circ$p, and $\psi(x_p)$:= $\neg$T(x_p) $\wedge$ $\circ$p, respectively.

Now, consider the self-referential sentence "This sentence is only false". We know that we cannot express this in QLP from Theorem 1,[7]. However, the introduction of the consistency operator allows us to formalize this expression. Let QLP be self-referential. Then, for some c, $\psi(x_\phi)$:= $\neg$T(x_ϕ) $\wedge$ $\circ\phi$ is such that $c_{\psi(c)}$ is identical to the name of $\psi(c_{\psi(c)})$, that is, $c_{\psi(c)}$ is identical to the name of $\neg$T($c_{\psi(c)}$) $\wedge$ $\circ\psi$(c).

Therefore, by TRANSPARENCY, T($c_{\psi(c)}$) must hold the same truth-value as $\neg$T($c_{\psi(c)}$) $\wedge$ $\circ\psi$(c). As we can verify in the truth table below, there is no valuation that satisfies this requirement. Therefore, the Strong Liar sentence cannot receive a valuation in QLP°.

ψ(c)	$\circ\psi$(c)	T($c_{\psi(c)}$)	$\neg$T($c_{\psi(c)}$)	$\neg$T($c_{\psi(c)}$) $\wedge$ $\circ\psi(c)$
1	1	1	0	0
1/2	0	1/2	1/2	0
0	1	0	1	1

[7] Analogously, we can't express that a proposition is not-true, only that the proposition is false. (Goodship, 1996, p. 153) for example, makes this mistake when talking about the Liar Paradox in relation to LP: she constructs the liar sentence as asserting α: "α is not true."

LP, when extended with the consistency operator, is an LFI (Carnielli e Coniglio, 2016, p.32). The importance of the consistency operator is related to the use of *Logics of Formal Inconsistency* (LFIs), introduced in (Marcos, 2005). The idea behind LFIs is that consistency, introduced as a primitive operator in a paraconsistent logic, serves to separate the logical space into two regions: classical and paraconsistent propositions. For those propositions α such that $\circ\alpha$ holds, α becomes explosive in the presence of its negation. The basic axiom that introduces consistency is (bc1) (Carnielli e Coniglio, 2016, p. 34):

$$\text{(bc1) } \circ\alpha \to (\alpha \to (\neg\alpha \to \beta))$$

In summary, by extending QLP with a unary operator of consistency $\circ$, we can characterize an LFI. However, expanding the expressive power of this logic has a cost: we get a new Liar sentence, one kind of Liar revenge. Although we were able to deal with the original Liar sentence, by attributing the value p to it, we form a new sentence α: "α is only false", and it is not possible to attribute any value to it in the extension of QLP.

Curry's Paradox also becomes troublesome in the case of QLP$^\circ$: with the consistency operator, we can define other conditionals that are detachable. Here we consider three conditionals definable in LP$^\circ$:

$$\alpha \Rightarrow \beta := (\alpha \to_{LP} \beta) \wedge (\circ\alpha \vee \beta)$$

$$\alpha \to^\circ \beta := \alpha \to_{LP} \beta \wedge \circ(\alpha \to_{LP} \beta)$$

$$\alpha \to_{LFI1} \beta := (\neg\alpha \wedge \circ\alpha) \vee \beta$$

It can be straightforwardly checked that these three conditionals validate *modus ponens*. That is, if $\vartheta(\alpha) \in$ D and $\vartheta(\alpha \to \beta) \in$ D, then $\vartheta(\beta) \in$ D. In the truth table below, we can check that no conditional of these types will satisfy TRANSPARENCY condition for the case of Curry's Paradox that says "This sentence implies its absurdity". With the consistency operator, we can also define the bottom formula $\perp_\alpha$ as $\alpha \wedge \neg\alpha \wedge \circ\alpha$ (Carnielli e Coniglio, 2016, p. 47). We will drop out α and only write $\perp$.

ϕ(c)	Tc$_{\phi(c)}$	$\perp$	Tc$_{\phi(c)} \Rightarrow \perp$	Tc$_{\phi(c)} \rightarrow^{\circ} \perp$	Tc$_{\phi(c)} \rightarrow^{LFI1} \perp$
1	1	0	0	0	0
1/2	1/2	0	0	0	0
0	0	0	1	1	1

These results on paradoxes and expressiveness are known in the case of LP (Rosenblatt, 2021, p. 2). One could think that these paradoxes are due to the consistency operator being introduced in the language.

However, it is also known that the consistency operator in the LP matrix can also be defined in terms of the unaries operators of J3 (D'Ottaviano, 1985) or the strong negation operator $\sim$. α says that α receives a designated value, while $\sim\alpha$ says that α is not designated (α is equivalent to $\neg\sim\alpha$):

		$\sim$
1	1	0
1/2	1	0
0	0	1

Consistency can be expressed by these operators in the following way:

$$\circ\alpha := \neg\alpha \vee \neg\neg\alpha$$

$$\circ\alpha := \sim\alpha \vee \sim\neg\alpha$$

In this sense, we can also expand QLP with the unary operators , or $\sim$, resulting in QLP and QLP$^{\sim}$. Given that we can define the consistency operator in those cases, the same result about the Liar Revenge and Curry's Paradox follows once again. As we will show, the consistency operator is not the only way to bring paradoxes back. In general, if we allow a unary operator to return only classical values, we can define consistency. The truth tables below show the equivalence between the formulae above.

α	$\neg\alpha$	$\sim\alpha$	$\sim\neg\alpha$	$\sim\alpha \vee \sim\neg\alpha$	$\circ\alpha$
1	0	0	1	1	1
1/2	1/2	0	0	0	0
0	1	1	0	1	1

α	$\neg\alpha$	α	$\neg\alpha$	$\neg\alpha$	$\neg\neg\alpha$	$\neg\alpha \vee \neg\neg\alpha$
1	0	1	0	0	1	1
1/2	1/2	1	1	0	0	0
0	1	0	1	1	0	1

α	$\sim\alpha$	α	$\circ\alpha$	$\bullet\alpha$	$\rhd\alpha$	$\lhd\alpha$
1	0	1	1	0	1	0
1/2	0	1	0	1	0	1
0	1	0	1	0	0	1

Def. 2 *Let M = {1, 1/2, 0}, # be a unary operator. We say that # is a hyperclassical operator if for all formulae $\alpha \in \mathcal{L}_{\mathcal{QLP}}$, $\vartheta(\#\alpha) \in \{1,0\}$.*

What are the hyperclassical unary operators of the three-valued matrix M? We have to consider all combinations of 0's and 1's, except the $\bot$ and $\top$ combinations, which are 0-ary operators. The first three cases we already know: strong negation, designated value, and the consistency operators. $\bullet\alpha$ is simply the paraconsistent negation of $\circ\alpha$, that is, $\bullet\alpha$ says that α is inconsistent. We introduce here the symbols $\rhd$ and $\lhd$, which assert α is only true and α is not only true, respectively. It is easy to check that $\circ\alpha$ can also be defined in terms of $\bullet\alpha$, $\rhd$ and $\lhd$. The proof is straightforward by truth tables and it is left as an exercise to the reader:

$$\circ\alpha := \neg\bullet\alpha$$

$$\circ\alpha := \rhd\alpha \vee \rhd\neg\alpha$$

$$\circ\alpha := \neg\lhd\alpha \vee \neg\lhd\neg\alpha$$

Theorem 2 *Let $QLP^{\#}$ be QLP extended with a unary hyperclassical operator #. $QLP^{\#}$ can define the consistency operator $\circ$.*

As demonstrated above, any unary hyperclassical operator $\# \in \{\sim, , \circ, \bullet, \rhd, \lhd\}$ can express consistency in $QLP^{\#}$. Thus, the Strong Liar sentence is not only a problem when we consider extensions of QLP that introduce a consistency operator. Any hyperclassical operator will also result in this problem. Now, we will discuss the philosophical significance of this result to LFIs and its relation to dialetheism.

4 The "Neutrality" of LFIs

The existence of a Liar sentence in the extension of QLP does not seem to be of much importance at first glance. However, it becomes a problem when we consider the range of interpretations for LFIs: these logics want to express both paraconsistent and classical phenomena in its language, and, allegedly, it serves to represent either an ontological interpretation, in terms of truth preservation, or an epistemic interpretation, in terms of information preservation.

The ontological interpretation is usually called dialetheism, whereas the epistemic interpretation has been exposed, for example, in (Marcos, 2005; Carnielli, Marcos e De Amo, 2000) as a model for inconsistent databases. That is, LFIs are considered a formal apparatus that is independent of these philosophical considerations of how to interpret the paraconsistent system in terms of *dialetheias* or conflicting information:

> ... LFIs do not have any relationship with dialetheism, a philosophical view in which there are true contradictions... If dialetheists turn out to be correct, and some contradictions are actually true, the LFIs will be on their side. If not, the LFIs will continue to be of value, independently of this debate. (Carnielli e Coniglio, 2016, p. 3)

I am not saying that the authors are neutral about the philosophical interpretation of paraconsistent logics. As it is pointed out in (Barrio, 2018), there are issues with regarding an epistemic interpretation as canonical. However, I will not enter such discussions here.[8] Carnielli and Coniglio, along with Rodrigues, extensively defend an epistemic interpretation of paraconsistency (Carnielli e Rodrigues, 2019; Carnielli e Rodrigues, 2015a; Carnielli e Rodrigues, 2021; Rodrigues e Carnielli, 2022). What I am atributing to them is that they think the technical apparatus of LFIs, by itself, should be able to provide a good framework for both the dialetheist and the epistemic interpreter of paraconsistency, which is clear from the quote above.

In this sense, LFIs should be able to provide a logic for dialetheism, even though

[8] I thank an anonymous reviewer for pointing this out to me.

the logic itself does not commit to an ontological interpretation. Let us suppose, for the sake of argument, that we want to use the QLP extension to think in terms of truth preservation. In that case, if $\circ\alpha$ holds, then α can't be a dialetheia, that is, α is only true or α is only false. However, if $\circ\alpha$ does not hold, then α is paradoxical, that is, α is both true and false. In this context, the Liar Paradox is one of the strongest indicators of a true contradiction due to its resistance of solution through time:

> The logical paradoxes... have been around for a long time now. Yet no solution has been found. Admittedly, the Liar Paradox, which has been known for over 2000 years, has often been ignored as a triviality unworthy of serious consideration... However, in the last three quarters of a century, there has probably been more intensive work done in trying to find a solution for the logical paradoxes than on any other topic in the history of logic. The roll call of those who have tried to find a solution reads like a logician's honours list. Yet no widely accepted solution has been found. (Priest, 1979, p. 219)

Thus, a satisfactory logic for dialetheism should at the very minimum provide an acceptance of the Liar Paradox. However, LFIs, by introducing the consistency operator, try to make classical logic coexist with paraconsistent logic. Consequently, as exposed in the previous section, we get a Liar's Revenge, a new Liar sentence that is unacceptable in this logic. Therefore, LFIs do not provide a good basis to tame the Liar Paradox in the case of LP. In fact, due to the Liar's revenge, LFIs are not neutral with respect to their interpretations, because a dialetheist interpretation of LFI is self-defeating, as it does not give a satisfactory account of the Liar Paradox. Therefore, in this sense, LFIs are anti-dialetheist: they do not provide a good model for truth preservation in paraconsistent scenarios where we consider paradoxes of self-reference.

Of course, this does not mean that LFIs are unable to provide a good basis for dialetheism *in general*. Dialetheism can also be defended by looking at other cases of contradictions. For example, in the analysis of time and motion (Priest, 2006, p. 213). We could also try to tackle the Liar sentence in paraconsistent logics different from LP, such as mbC.

The problem posed by the Liar in QLP$^\circ$ also arises in the context of QLFI1, the first-order version of LFI1, and J3: these logics share a three-valued matrix in the case of $\neg$, $\wedge$ and $\circ$ (Carnielli e Coniglio, 2016, p. 159), or they can be inter-definable in the case of LFI1 and J3.

5 Consistency and contradiction

So far, we are faced with the following problem: the first-order versions of LP, LFI1 and J3 that incorporate any hyperclassical unary operator are not able to avoid some type of Liar Revenge. As (Rosenblatt, 2021, p. 35) summarizes, we can either (i) drop the T-schema, or some other principles from naive truth theory, or (ii) we can expurge consistency from our system - as is the case in QLP -, or (iii) we can treat consistency as a notion that can be itself inconsistent.

We will not pursue options (i) and (ii); what we want to investigate is whether paraconsistency can mantain all underlying motivation behind the reasoning of the Liar Paradox, and still not trivialize. For this, we will push to the limit the paraconsistent theories available, and we will choose the weakest LFI that Carnielli and Coniglio start their construction (Carnielli e Coniglio, 2016, p. 33): mbC, the minimal logic of formal inconsistency.

Unlike LP, mbC is a logic that is not strong enough to be characterizable by many-valued matrices (Carnielli e Coniglio, 2016, p. 126). Let $\mathcal{L}$ be defined over a signature $\Sigma = \{\wedge, \vee, \neg, \rightarrow\}$. mbC is characterized by $(\mathcal{L}, \vdash_{mbC})$ such that $\vdash_{mbC}$ validates Modus Ponens and the following axioms:

Ax1 $\alpha \rightarrow (\beta \rightarrow \alpha)$

Ax2 $(\alpha \rightarrow (\beta \rightarrow \gamma)) \rightarrow ((\alpha \rightarrow \beta) \rightarrow (\alpha \rightarrow \gamma))$

Ax3 $\alpha \rightarrow (\beta \rightarrow (\alpha \wedge \beta))$

Ax4 $(\alpha \wedge \beta) \rightarrow \alpha$

Ax5 $(\alpha \wedge \beta) \rightarrow \beta$

Ax6 $\alpha \rightarrow (\alpha \vee \beta)$

Ax7 $\beta \rightarrow (\alpha \vee \beta)$

Ax8 $(\alpha \rightarrow \gamma) \rightarrow ((\beta \rightarrow \gamma) \rightarrow ((\alpha \vee \beta) \rightarrow \gamma))$

Ax9 $((\alpha \rightarrow \beta) \rightarrow \alpha) \rightarrow \alpha)$

Ax10 $\alpha \vee \neg\alpha$

Ax11 $\circ\alpha \rightarrow (\alpha \rightarrow (\neg\alpha \rightarrow \beta))$

This axiom system can be characterized by a bivaluation semantics (Carnielli e Coniglio, 2016, p. 35), which is basically a valuation based on {0, 1} that is non-deterministic for some connectives. Consider QmbC as the first-order extension of mbC analogously to QLP, and let QmbC be self-referential and extendable to a theory with truth predicates and names for its own sentences as defined above. What happens with the Liar sentence in QmbC? Surprisingly, now we have stable valuations for the Liar Revenge.

ϕ(c)	$\neg\phi$(c)	$\circ\phi$(c)	Tc$_{\phi(c)}$	$\neg$Tc$_{\phi(c)}$	Tc$_{\neg\phi(c)}$	$\circ\phi$(c) $\wedge$ $\neg$Tc$_{\phi(c)}$	
6*1	2*1	2*0	2*1	1	1	0	ϑ_1
				0	1	0	ϑ_2
	4*0	2*1	2*1	1	0	1	ϑ_3
				0	0	0	ϑ_4
		2*0	2*1	1	0	0	ϑ_5
				0	0	0	ϑ_6
2*0	2*1	0	0	1	1	1	ϑ_7
		1	0	1	1	0	ϑ_8

For example, in valuation ϑ_3 below, the Liar Revenge can be stably valuated as true: $\vartheta(\phi$(c)) = 1 = $\vartheta(\circ\phi$(c) $\wedge$ $\neg$Tc$_{\phi(c)}$). One could argue that this valuation should be excluded because $\neg$Tc$_{\phi(c)}$ is not equivalent to Tc$_{\neg\phi(c)}$. Even so, we still get the valuation ϑ_8 that stably falsifies the Liar sentence: in this case, this happens because consistency is not assumed to be defined in terms of non-contradiction.

In summary, in QmbC, consistency is weak enough to avoid triviality in the face of the Liar sentence. We can also say that mbC represents consistency inconsistently, because there are valuations for some formulae α such that $\vartheta(\circ\alpha) = 1$ and $\vartheta(\neg\circ\alpha) = 1$, that is, the attribution of consistency is contradictory.

We showed here that taking the $\circ$ operator to be inconsistently expressible avoids the Liar Revenge, as suggested in (Rosenblatt, 2021, p. 35). Unfortunately, the story does not end here. As a finale to this section, we will see that even in QmbC we cannot avoid further revenges of the Liar in two ways: via strong negation and strong consistency.

In mbC, we can define a $\perp$ operator in terms of any formula α. After introducing this operator, given that mbC has a classical conditional, we can recover the strong negation in terms of α by defining it with $\perp$, which behaves as a usual Boolean negation (Carnielli e Coniglio, 2016, p. 47).

$$\perp_\alpha := \alpha \wedge \neg\alpha \wedge \circ\alpha$$
$$\sim\alpha := \alpha \rightarrow \perp_\alpha$$

Of course, we can also define these notions in QmbC. Consider a self-referential sentence $c_{\phi(c)}$ such that $c_{\phi(c)}$ is identical to the name of $\sim Tc_{\phi(c)}$. That is, we are considering the sentence “This sentence is not true”. It is straightforward to check that there is no stable valuation for ϕ(c): suppose that $\vartheta(\phi(c)) = 1$. Then, $\vartheta(Tc_{\phi(c)}) = 1$, which means that $\vartheta(\sim Tc_{\phi(c)}) = 0$. Suppose that $\vartheta(\phi(c)) = 0$. Then $\vartheta(Tc_{\phi(c)}) = 0$, which means that $\vartheta(\sim Tc_{\phi(c)}) = 1$.

This is not the only problem posed by the Liar in QmbC. In fact, we can define consistency consistently in QmbC using a strong negation (Carnielli e Coniglio, 2016, p. 69), which we call $\circ^c\alpha$:

$$\circ^c\alpha := \sim(\alpha \wedge \neg\alpha)$$

We can check from the bivaluation below that this operator behaves just like the consistent consistency operator introduced before in the context of LP. That is, $\vartheta(\circ^c\alpha) = 0$ only iff $\vartheta(\alpha) = \vartheta(\neg\alpha) = 1$.

α	$\neg\alpha$	$\alpha \wedge \neg\alpha$	$\sim(\alpha \wedge \neg\alpha)$ [:= $\circ^c\alpha$]
6*1	2*1	2*1	2*0
	4*0	2*0	2*1
		2*0	2*1
2*0	2*1	0	1
		0	1

Now, consider the Liar Revenge sentence $c_{\phi(c)}$ such that $c_{\phi(c)}$ is identical to the name of $\neg Tc_{\phi(c)} \wedge \circ^c\phi(c) \wedge \circ^c Tc_{\phi(c)}$. This sentence says that "This sentence is only false and it is consistently so", consistency here being understood in the classical way as non-contradictoriness, that is, in terms of $\circ^c$.

We can show that no valuation can give a stable value to this sentence.

ϕ(c)	$\neg\phi$(c)	$Tc_{\phi(c)}$	$\neg Tc_{\phi(c)}$	$\circ^c Tc_{\phi(c)}$	$\circ^c\phi$(c)	$\neg Tc_{\phi(c)} \wedge \circ^c Tc_{\phi(c)} \wedge \circ^c\phi(c)$
4*1	2*1	2*1	1	0	0	0
			0	1	0	0
	2*0	2*1	1	0	1	0
			0	1	1	0
0	1	0	1	1	1	1

In summary, in QmbC we start with an *inconsistent* consistent notion of $\circ$, because judgements of consistency can themselves be contradictory. However, it is possible to introduce classical negation and a *consistent* notion of consistency defined in terms of $\circ$. This result means that we can have revenge results of the Liar even in QmbC, the weakest first-order LFI based on positive classical logic.[9]

Therefore, the option of making the notion of consistency behave in an inconsistent way to avoid the Liar Revenge is not open to LFIs that are based on classical

[9] LFIs can also be based on fuzzy logics, intuitionistic logics, and etc (Carnielli e Coniglio, 2016, p. 171). The relation of the Liar to these other LFIs will not be explored here.

logic. The problem of combining LFIs with naive truth theory is much more deeply rooted in LFIs than was originally described: the behavior of $\circ$ is not important to the proliferation of paradoxes. All we need is that it recovers explosiveness. Consider, for example, the diagnosis of Rosenblatt:

> There are a number of ways to react to the revenge paradox just posed...Third, one could argue that consistency is indeed expressible in the object language, but not consistently. That is, even if one can say that some statement is consistent, and this assertion turns out to be true, there is no guarantee that the assertion will not be false as well. To be sure, this third option involves, once again, a certain compromise on the part of the paraconsistent theorist... This third option takes the revenge paradox above to show that consistency is not consistently expressible. Yet, this does not mean that one cannot inconsistently express consistency. That is, it is still open to the paraconsistent logician to maintain that, even if an operator like $\circ$ is not expressible, perhaps similar and more thoroughly nonclassical operators are. (Rosenblatt, 2021, p. 35)

In his diagnosis, treating consistency inconsistently, even though it has some costs, is a feasible route to avoid the Liar in LFIs. However, we showed above that this is not the case for LFIs based on positive classical logic.

6 Paradefinite Logics

A similar remark of the inadequacy to deal with logical paradoxes was pointed out by (Arenhart e Melo, 2022), in which they discuss the *Logics of Evidence and Truth* (LETs), logics that are both paracomplete and paraconsistent (i.e., paradefinite), and that also have a recovery operator $\circ$ that classicalizes the sentences to which it applies. LETs are extensions of LFIs, because they are Logics of Formal Inconsistencies and Logics of Formal Undeterminedness (LFUs):

> The basic ideas of LFIs and LFUs may be combined in an LFIU – a Logic of Formal Inconsistency and Undeterminedness. In a context that is at the same time paraconsistent and paracomplete, we may recover at once explosion and excluded middle with respect to a given formula A, and hence the properties of classical negation with respect to A (Carnielli e Rodrigues, 2015b, p. 21)

The point Arenhart and Melo make is that the Liar Paradox is not satisfactorily dealt with either in terms of the epistemic interpretation or in terms of truth preservation: for the former, it is difficult to understand how the Liar sentence and its negation have nonconclusive evidence for them both:

> … a naive theory of the very concept of truth that delivers the Liar is a bad theory for this concept, and as such, does not reach reality, as true theories should do, but rather only delivers some kind of non-conclusive evidence. That is, this kind of answer benefits from the fact that the epistemic approach sees inconsistent theories as defective theories, not reaching reality, and that this may apply even to a naive theory of truth that delivers the Liar. The naive theory of truth delivering the Liar, then, would be no different than other defective theories such as Bohr's model of the atom; the only difference is that it deals with the notion of truth. (Arenhart e Melo, 2022, p. 299-300)

In that sense, the naive theory of truth would be a linguistic system that is contradictory, as any scientific theory that is under construction can be. However, as they argue, the epistemic approach is in a worse position than the semantic dialetheist, which defends that there are linguistic contradictions, but not real ones. The Liar Paradox stated in terms of truth, as the semantic dialetheist exposes, would be a better description of our reasoning of the paradox, compared to the epistemic description of conflicting evidence (Arenhart e Melo, 2022, p. 301).

Furthermore, Arenhart and Melo also point out that the Liar sentence in terms of truth is not explained in LETs. Within the epistemic approach, true propositions are those of which we have conclusive evidence, and in that case, we reason with classical inferences. However, as the naive theory of truth contains a contradiction in it, it is a defective theory that is subject to paraconsistent reasoning of non-conclusive evidence. Thus, it seems that the epistemic approach presupposes a theory of truth to be consistent and at the same time to be defective.

> There is still a contradiction at the naive stage that needs to be accounted for and is in terms of the naive notion of truth. We still need to be told how to logically account for that contradiction. The epistemic approach does not have the possibility open to the classical approaches to the truth, which recommends a substitution of the naive concept by a revised consistent one; rather, the epistemic approach proposes that a paraconsistent logic may deal

> with contradictions in theories that are known not to be the final theory... So, even if the concept of truth can be revised successfully, there is still a contradiction, involving truth, that is present at the informal naive level and cannot be accounted for in terms of non-conclusive evidence. (ARENHART; MELO, p. 306-307, 2022)

In that way, the coexistence of classical and paraconsistent reasoning becomes troublesome when we try to think in terms of non-conclusive evidence and truth side-by-side.

But what if we try a full dialetheist position with regard to LFIs and LETs, interpreting the Liar Paradox in terms of truth only? In this context, a bigger problem arises. As we have seen in the previous sections, in LFIs we can strengthen the Liar sentence by a variation that asserts A: “A is only false.”. In that case, we end up with a paradox that is unacceptable in the extension of LP. Suppose that we avoid the problem of the Liar’s revenge sentence by adopting a Tarskian hierarchy of languages for the predicate “only false”. This would technically avoid the paradox, but would we be better off in that case than simply assuming classical logic? At least in LETs based on classical logic, the same embarassment would occur, because their recovery operator recovers classical logic (i.e., it blocks paraconsistency and paracompleteness).

Dialetheism with regards to naive truth theory starts with the motivation of accepting paradoxes such as the Liar as inevitable, in such a way that we have to accept them as true contradictions. By proposing a Tarskian solution to deal with the Liar’s revenge, on the contrary, we are not accepting these paradoxes but avoiding them, in the same way as classical logic does. On the other hand, in LFIs we cannot simply accept the Liar Revenge: it is paradoxical to assert “This sentence is only false”. Even if we weaken the logics to the minimal first-order logic of formal inconsistency, QmbC, paradoxes still arise. Therefore, when thinking in terms of paraconsistency and naive truth theory, LFIs are not a suitable logic: it fails to succeed in its initial purpose of accepting paradoxes within the language as true contradictions.

Conclusion

By adopting a first-order extension of the Logic of Paradox which adds the unary operator $\circ$, self-referentiality, and names for sentences in the language, we were able to produce a variant of the Liar Paradox, which is the sentence "This sentence is only false". We were able to see how this engenders a non-acceptable conclusion in the extension of QLP. Furthermore, we showed that $QLP^{\#}$, # being a hyperclassical unary operator, also produces the same result. Thus, the problem at this level is not only about consistency, but about recovery of classicality. Furthermore, we showed that weakening the notion of consistency cannot avoid all types of Liar Revenges: considering QmbC, it is also possible to define classical negation and classical consistency and produce other paradoxes.

Therefore, it becomes clear that standard LFIs are not adequate to deal with naive theory of truth satisfactorily. However, there are a few routes from here, which are open ideas that the author considers for future research. Following Barrio, Pailos and Szmuc (2017), we could define self-reference in some weak manner and drop the classical conditional to obtain a nondetachable one. It is also possible to consider LFIs that are not based on classical logic (Carnielli e Coniglio, 2016, p. 171). However, these routes are questionable if our goal is to think of truth theory and self-reference naively.

Referências

ARENHART, J. B.; MELO, E. S. The Liar Paradox: between evidence and truth. **Logic and Logical Philosophy**, Nicolaus Copernicus University, v. 31, n. 2, p. 289–311, 2022.

BARRIO, E. A. Models & proofs: LFIs without a canonical interpretations. **Principia: an international journal of epistemology**, v. 22, n. 1, p. 87–112, 2018.

BARRIO, E. A.; PAILOS, F. M.; SZMUC, D. E. A paraconsistent route to semantic closure. **Logic Journal of the IGPL**, Oxford University Press, v. 25, n. 4, p. 387–407, 2017.

CARNIELLI, W.; CONIGLIO, M. E.; MARCOS, J. Logics of Formal Inconsistency. In: **Handbook of Philosophical Logic**. Edição: D.M. Gabbay e F. Guenthner. Dordrecht: Springer Netherlands, 2007. P. 1–93.

CARNIELLI, W.; MARCOS, J.; DE AMO, S. Formal inconsistency and evolutionary databases. **Logic and logical philosophy**, n. 8, p. 115–152, 2000.

CARNIELLI, W.; RODRIGUES, A. Towards a philosophical understanding of the logics of formal inconsistency. **Manuscrito**, SciELO Brasil, v. 38, n. 2, p. 155–184, 2015.

CARNIELLI, W.; RODRIGUES, A. A logic for evidence and truth. **CLE e-Prints**, v. 15, n. 5, p. 1–13, 2015.

CARNIELLI, W.; RODRIGUES, A. An epistemic approach to paraconsistency: a logic of evidence and truth. **Synthese**, Springer, v. 196, n. 9, p. 3789–3813, 2019.

CARNIELLI, W.; RODRIGUES, A. On epistemic and ontological interpretations of intuitionistic and paraconsistent paradigms. **Logic Journal of the IGPL**, Oxford University Press, v. 29, n. 4, p. 569–584, 2021.

CARNIELLI, W. A.; CONIGLIO, M. E. **Paraconsistent Logic: Consistency, Contradiction and Negation**. Berlin: Springer, 2016. v. 40.

D'OTTAVIANO, I. M. The completeness and compactness of a three-valued first-order logic. **Revista Colombiana de Matematicas**, Universidad Nacional de Colombia, v. 19, n. 1-2, p. 77–94, 1985.

GOODSHIP, L. On Dialethism. **Australasian Journal of Philosophy**, Taylor & Francis, v. 74, n. 1, p. 153–161, 1996.

MARCOS, J. **Logics of formal inconsistency**. Brazil: Fundação Biblioteca Nacional, 2005.

PRIEST, G. **In Contradiction**. USA: Oxford University Press, 2006.

PRIEST, G. The Logic of Paradox. **Journal of Philosophical logic**, Springer, p. 219–241, 1979.

RODRIGUES, A.; CARNIELLI, W. On Barrio, Lo Guercio, and Szmuc on logics of evidence and truth. **Logic and Logical Philosophy**, v. 31, n. 2, p. 313–338, 2022.

ROSENBLATT, L. Expressing consistency consistently. **Thought: A Journal of Philosophy**, Wiley Online Library, v. 10, n. 1, p. 33–41, 2021.

SMULLYAN, R. M. **Diagonalization and self-reference**. USA: Oxford University Press, 1994.

TAJER, D. LFIs and methods of classical recapture. **Logic Journal of the IGPL**, Oxford University Press, v. 28, n. 5, p. 807–816, 2020.