Book review: CARNIELLI, W., CONIGLIO, M. *Paraconsistent Logic: Consistency, Contradiction and Negation.* Logic, Epistemology, and the Unity of Science Series. (New York: Springer, 2016. ISSN: 2214-9775.)

Henrique Antunes
State University of Campinas
Department of Philosophy
Campinas, SP
Brazil
antunes.henrique@outlook.com

Vincenzo Ciccarelli
State University of Campinas
Department of Philosophy
Campinas, SP
Brazil
ciccarelli.vin@gmail.com

Article info
CDD: 160
Received: 01.12.2017; Accepted: 30.12.2017
DOI: http://dx.doi.org/10.1590/0100-6045.2018.V41N2.HV

Keywords:
Paraconsistent Logic
LFIs

Abstract
Review of the book 'Paraconsistent Logic: Consistency, Contradiction and Negation' (2016), by Walter Carnielli and Marcelo Coniglio

The *principle of explosion* (also known as *ex contradictione sequitur quodlibet*) states that a pair of contradictory formulas entails any formula whatsoever of the relevant language and, accordingly, any theory regimented on the basis of a logic for which this principle holds (such as classical and intuitionistic logic) will turn out to be trivial if it contains a pair of theorems of the form A and $\neg A$ (where $\neg$ is a negation operator). A logic is *paraconsistent* if it rejects the principle of explosion, allowing thus for the possibility of contradictory and yet non-trivial theories.

Among the several paraconsistent logics that have been proposed in the literature, there is a particular family of (propositional and quantified) systems known as *Logics of Formal Inconsistency* (**LFI**s), developed and thoroughly studied within the Brazilian tradition on paraconsistency. A distinguishing feature of the **LFI**s is that although they reject the general validity of the principle of explosion, as all other paraconsistent logics do, they admit a a restrcited version of it known as *principle of gentle explosion*. This principle asserts that a contradiction that concerns a *consistent* formula logically entails any other formula of the language. The expression 'consistent' here is a generic term susceptible to several alternative interpretations (not necessarily coinciding with non-contradiction), depending on the particular **LFI** under consideration. Another (related) feature that distinguishes the **LFI**s from other paraconsistent logics is that they internalize this unspecified notion of consistency inside the object language by means of a unary sentential operator $\circ$ (called 'consistency operator' or simply 'circle'). When prefixed to a formula A, $\circ$ expresses that A is consistent or well behaved, however these expressions are to be interpreted in each particular case.

Paraconsistent Logic: Consistency, Contradiction and Negation, by Walter Carnielli and Marcelo Coniglio, is entirely devoted to the Logics of Formal Inconsistency. The book covers the main achievements in the field in the past 50 years or so, presenting them in a systematic and (to a great extend) self-contained way. Although the book is mostly concerned with particular logical systems, the relations among them, and their corresponding metatheoretical properties, it also sets the basis of a new philosophical interpretation of paraconsistent logics.

The book contains nine chapters, which altogether cover several topics about the **LFI**s. In Chapter 1 the authors explain the rationales behind paraconsistent logics in general and the **LFI**s in particular, and discuss the philosophical problems related to paraconsistency under the light of some general issues in the philosophy of logic (such as the nature of logic and the nature of contradictions). It is argued that since there are some real life situations in which contradictions do actually turn up, paraconsistent logics are justified, no matter how those contradictions are interpreted – whether they are seen as concerning reality or knowledge. The chapter also discusses the relation between paracomplete and paraconsistent logics and analyzes some key notions related to paraconsistency, such as consistency, contradiction (and the *principle of non-contradiction*) and negation.

In Chapter 2 the concept of **LFI** is precisely defined, as well as other basic technical notions employed throughout the book. A minimal propositional **LFI**, called **mbC**, is introduced by means of an axiomatic system. **mbC** results from *positive classical propositional logic* by the inclusion of two additional axioms: the principles of excluded middle and gentle explosion – $A \vee A$ and $\circ A \rightarrow (A \rightarrow (\neg A \rightarrow B)$, respectively. **mbC** is then provided with a *valuation semantics* with respect to which it is proved to be sound and complete. The relations between **mbC** and classical propositional logic are carefully analyzed. The analysis reveals that **mbC** can be viewed both as a *sublogic* and as an *extension* of classical logic, when these terms are suitably qualified.

Chapter 3 presents several extensions of **mbC** and analyzes the relations between the notions of *consistency/inconsistency* and *contradictoriness/non-contradictoriness* – formally expressed by the formulas $\circ A/\neg\circ A$ and $A \wedge \neg A/\neg(A \wedge \neg A)$, respectively. As it turns out, although consistency and non-contradictoriness (and inconsistency and contradictoriness) are partially independent in **mbC**, they may or may not coincide in some of its extensions. In addition, the notion of a *C*-system is introduced. Despite the complexity of the relevant definition, a *C*-system simply amounts to an **LFI** within which the consistency operator is definable in terms of the other connectives of the language. *Da Costa's hierarchy of paraconsistent logics* – a family of paradigm examples of *C*-systems – is briefly presented and explained. The chapter also deals with the important notions of propagation and retro-propagation of the consistency operator.

The first part of Chapter 4 is devoted to the problem of the algebraizability of some **LFI**s, and the second part discusses some many-valued **LFI**-systems. In Section 4.1 some preliminary concepts concerning logical matrices are introduced. Section 4.2 contains a Dugundji-style proof of the uncharacterizability by finite matrices of the **LFI**s presented so far. Section 4.3 contains a proof of the algebraizability of some extension of **mbC** in the broader sense of Block and Pigozzi. The remaining sections deal separately with different many-valued **LFI**s, most of which were proposed several decades before the emergence of the concept of Logic of Formal Inconsistency.

Chapter 5 represents a partial detour from the main exposition, for the systems presented therein are not extensions of positive classical propositional logic. The first case considered by the authors is that of intuitionistic logic: more specifically, it is shown how a consistency operator $\circ$ can be defined within Nelson's logic **N4** in terms of a strong negation $\sim$ operator (i.e., $\circ A \equiv \sim(A \wedge \neg A)$). Another interesting case covered by the chapter is that of modal logic, where the consistency operator is shown to be interpretable as having a sort of "modal flavor". In particular, the definition $\circ A \equiv A \rightarrow \Box A$ can be introduced in normal non-degenerate modal logics. Some systems of fuzzy logic are also analyzed in the chapter. In all of the aforementioned logics, the strategy pursued by the authors consists in defining a consistency operator within the system in question and then showing that it satisfies the general definition of an **LFI**.

Chapter 6 is devoted to the problem of defining non-deterministic semantics for non-algebraizable systems (even in the broader sense of Block and Pigozzi). It presents three main formal semantics – based, respectively, on **F**-*structures*, *non-deterministic logical matrices*, and *possible translations*. Of particular interest, especially from a more philosophical point of view, is the so-called *possible translation semantics*, whose main idea is to translate a given logic into logics whose semantics are well known and deterministic. The relevant notion of translation is that of a mapping preserving logical consequences and the rationale for this approach is the interpretation of a logic as a combination of "possible world views".

Chapter 7 concerns first-order **LFI**s. The chapter is mainly devoted to two systems: **QmbC**, the first-order extension of **mbC**, and **QLFI1**. Due to the non-deterministic nature of **mbC**, a non-standard semantics is defined for its first-order extension: the authors introduce the notion of a *Tarskian paraconsistent structure*, defined as an ordered pair composed of a Tarskian structure (in the classical sense) together with a non-deterministic valuation. Concerning **QLFI1**, the approach is twofold: on one hand, it is shown how the language may be interpreted in a suitable Tarskian paraconsistent structure; on the other hand, a different semantics is proposed, given that the propositional fragment of **QLFI1** can be characterized by a three-valued matrix. The semantics is represented by a *partial structure*, defined in a similar way to a classical Tarskian structure, except for the fact that all predicate symbols are interpreted as *partial relations*. Both **QmbC** and **QLFI1** are proved to be sound and complete with respect to the corresponding semantics. Compactness and Lowenhëim-Skolem theorems are proved for **QmbC**.

Chapter 8 concerns one of the most straightforward applications of paraconsistent logics: set theory. Nevertheless, the authors' approach to the subject is substantially different from what has been traditionally done in the field of paraconsistent set theory – namely, to formulate a non-trivial naïve set theory countenancing the unrestricted comprehension principle for sets. The systems presented in the chapter include all of Zermelo-Fraenkel set theory's axioms (except for the *axiom of foundation*, which is replaced by a weaker version of it) with an **LFI** as the underlying logic. Another distinguishing feature of those systems is that they include a consistency predicate for sets whose behavior is governed by a set of additional axioms. Hence, whereas in a propositional **LFI** the property of consistency applies only to formulas, in the corresponding paraconsistent set theories it applies to both formulas and sets. The main results of the chapter are the *derivability adjustment theorem* (establishing that any derivation in **ZF** can be recovered within its paraconsistent counterpart) and a proof of the non-triviality of the strongest system presented in the chapter.

Chapter 9 discusses the significance of contradictions for science, describing some historical paradigm examples where contradictions seem to have played an important role in the development of scientific theories. It also proposes an interpretation of paraconsistent logics according to which they are better viewed as possessing an epistemological, rather than an ontological, character; in a nutshell, this means that they are not supposed to deal primarily with reality and truth (as in the case of classical logic), but with the epistemic notion of *evidence*. This interpretation is meant to be a more palatable alternative to *dialetheism* (the thesis that there are true contradictions), since it neither affirms the existence of true contradiction nor rejects classical logic as incoherent – adhering thus to logical pluralism.

One of the main virtues of *Paraconsistent Logic: Consistency, Contradiction and Negation* is that it keenly highlights the pervasiveness and generality of the notion of logic of formal inconsistency. Firstly, because it shows through the definition of an **LFI** how several systems of paraconsistent logic proposed in the literature – which at first sight might have appeared to be quite unrelated with one another – can be framed under a single unifying concept. Secondly, because it emphasizes that the definition of an **LFI** is applicable to systems based on logics of various different kinds, such as classical, intuitionistic, fuzzy, and modal logic. The resulting multiplicity of systems allows for various alternative semantic approaches, which are carefully described in several chapters of the book (e.g., valuation semantics, deterministic and non-deterministic matrices, **F**-structures, swap structures, possible translations semantics).

The book is mainly devoted to the taxonomy of **LFI**-systems, leaving little room for a more detailed discussion of the intrinsic properties of each particular system. This is understandable, though, since it is not meant to be a textbook. However, it is possible to use the book as an introductory text on formal paraconsistency by skipping some of the more technical chapters (e.g., a reader merely interested in those **LFI**s based on positive classical propositional logic may well skip chapters 5, 6 and possibly 8).

Concerning the more philosophical chapters of the book (chapters 1 and 9), the reader might think that the issues discussed therein would have deserved a more extended and rigorous analysis, especially when compared to the painstakingness of the other chapters. In particular, she might find the epistemic interpretation of paraconsistent logics wanting, despite its initial plausibility, this view in not sufficiently argued for. Moreover, specific relations between the epistemic interpretation and the particular features of the **LFI**s are missing. Nevertheless, this apparent shallowness is presumably due to the fact the purpose of those chapters is not to thoroughly develop a philosophical theory about paraconsistency, but merely to indicate some conceptual possibilities. After all, *Paraconsistent Logic* is mainly a technical piece of work.

So much for the general considerations. There are two specific points that we think would deserve a more detailed discussion. The first one concerns the cumbersome notation employed in the characterization of the semantics of first-order **LFI**s (Chapter 7): the strategy adopted by the authors in that chapter consists in extending the (non-deterministic) propositional valuations to the first-order case, combining these with a (classical) Tarskian structure – characterized, as usual, by a non-empty domain together with an interpretation function. The resulting first-order valuations apply thus only to sentences and the notion of truth, as in the propositional case, is not defined in terms of assignments, sequences, or any other technical device usually employed in order to interpreted the variables. The absence of any of these devices leads the authors to locally indicate all the relevant substitutions of individual constants for the free variables of a given formula. In the case of **QmbC**, for example, the semantic value of a quantified formula $\forall x A$ (under a structure $\mathfrak{A}$ and a valuation v) is defined by means of the following clause:

$v(\forall x A) = 1$ *iff* $v(A[x / \bar{a}]) = 1$, *for every* a *in the domain of* $\mathfrak{A}$

where $A[x / \bar{a}]$ denotes the result of substituting the constant $\bar{a}$ for all free occurrences of x in A, and where the language is supposed to have at least one individual constant $\bar{a}$ for each elements a of the domain of $\mathfrak{A}$ (that is, the language is supposed to be diagrammatic). At first sight, the use of the notation $[x / \bar{a}]$ (and its generalization $[x_1,\ldots, x_n / \bar{a}_1,\ldots, \bar{a}_n]$ to multiple simultaneous substitutions) does not seem to compromise readability at all – in fact, they are usually employed in the definition of substitutional semantics for first-order logic. However, matters become much more complicated when it comes to the additional clauses introduced in the definition of $v(A)$ in order to guarantee that the *substitution lemma* holds for Tarskian paraconsistent structures. One of these clauses, which concerns the negation operator, is formulated as follows:

(sNeg) *For every contexts* $(\vec{x}; z)$ *and* $(\vec{x}; y)$, *for every sequence* $(\vec{a}; \vec{b})$ *in the domain of* $\mathfrak{A}$ *interpreting* $(\vec{x}; \vec{y})$, *for every* $A \in L(\mathfrak{A})_{\vec{x};z}$ *and every* $t \in T(\mathfrak{A})_{\vec{x};\vec{y}}$ *such that* t *is free for* z *in* A, *if* $A[z/t] \in L(\mathfrak{A})_{\vec{x};\vec{y}}$ *and* $c = (t[\vec{x};\vec{y} / \vec{a}; \vec{b}])^{\mathfrak{A}}$ *then:*

If $v((A[z/t])[\vec{x};\vec{y} / \vec{a}; \vec{b}]) = v(A[\vec{x}; z / \vec{a}; c])$ *then*

$v((\neg A[z/t])[\vec{x};\vec{y} / \vec{a};\vec{b}]) = v(\neg A[\vec{x}; z / \vec{a}; c])$

Without attempting to individually explain every piece of notation above, **(sNeg)** merely expresses that if the substitution lemma holds for a formula A, then it holds for its negation as well (the introduction of this clause, absent in the definition of classical first-order structures, is necessary given the non-deterministic behavior of the negation operator in **mbC**). Now, it is quite clear that the reader would probably take several minutes to read and understand **(sNeg)**. Moreover, this situation is not restricted to **(sNeg)**, but it also happens with the similar clause concerning the consistency operator and the formulation and proof of various semantic theorems enunciated in Chapter 7. The notational cumbersomeness of the chapter is further worsened by the introduction of the notion of *extended valuation*, which assigns a truth value to an arbitrary formula A (not necessarily a sentence) by indicating a sequence of individual constants with respect to which A is to be evaluated. More precisely, if the free variables in A are among $x_1,\ldots, x_n$ (abbreviated by $\vec{x}$) then the truth value of A under the extended valuation $v_{\vec{x}\to\vec{a}}$ is simply $v(A[x_1,\ldots, x_n / \bar{a}_1,\ldots, \bar{a}_n])$. This notion represents a simile of the notion of *satisfaction* and is necessary in order to provide an interpretation for the open formulas.

The notation of Chapter 7 could, however, be greatly simplified in the following way: instead of importing the notion of valuation from the corresponding propositional **LFI**, the authors could well have defined a new notion of valuation which assigns one of the truth values 0 or 1 to each pair (s, A), where s is an assignment of objects of the domain to first-order variables and A is an arbitrary formula (open or closed). All definitions and theorems of the chapter could then be easily adapted according to this strategy, yielding much simpler formulations. In particular, clause **(sNeg)** above would become:

(sNeg') *Let A be a formula with at least one free variable z and let t be a term free for z in A. Let s be an assignment in a structure $\mathfrak{A}$ and let s' be the assignment which is just like s except that is assigns the interpretation of t under s to the variable z. Then:*

If $v(s', A) = v(s, A[z / t])$ *then* $v(s', \neg A) = v(s, \neg A[z / t])$

In addition to the evident simplicity of this new formulation, it is worth mentioning that since the notion of valuation above applies to any formula whatsoever of the language (open or closed), it is unnecessary to introduce extended valuations, resulting in a significant conceptual simplification.

Our second criticism concerns the paraconsistent set theories of Chapter 8. In general, the main motivation for a paraconsistent set theory is to recover the intuitive notion of set codified in the unrestricted principle of comprehension – i.e., the idea that every property *P* determines a set of all and only those objects having *P*. Of course, this can only be achieved by renouncing to classical logic, since that principle classically entails the existence of contradictory sets (e.g., Russell's set, universal set, etc.). On the other hand, classical set theories (such as **ZF**) maintain classical logic at the cost of imposing what seems to be *ad hoc* restrictions to the comprehension principle and countenancing additional principles whose justification seems also *ad hoc*. Hence, paraconsistent and classical set theories are symmetrically opposed to one another: what the former tries to achieve (i.e., preserve the intuitive notion of set) is given up by the latter, and what the latter preserves (i.e., classical logic) the former revises.

Nevertheless, the approach to paraconsistent set theory adopted by the authors diverges significantly from these two trends. Firstly, because the attempt to recover the intuitive notion of set codified in the principle of comprehension is explicitly given up once they opt for **ZF**-like axiomatizations of their theories – ruling out well-known inconsistent collections from the outset. Secondly, given that those theories are variations of **ZF** based on one or another **LFI**, the revision of the underlying logical theory is achieved by extending classical logic, rather than renouncing to it. In fact, each of the set theories of Chapter 8 is equivalent to **ZF** under the assumption that all sets enjoy the property of consistency.

This particular take on paraconsistency may leave the reader wondering what is the point of having a paraconsistent set theory that does not explicitly countenance contradictory collections ('Why not just stick with **ZF**?', she might ask.). The book does not provide an explicit answer to this question, though. However, it would not be difficult to imagine a scenario in which the systems of Chapter 8 would be vindicated: suppose that **ZF** is someday shown to be inconsistent. Under this circumstance, any of those systems could be used to preserve the strength of **ZF** while avoiding its triviality. Even though a paraconsistent set theory of this kind may turn out to be fruitful, its fruitfulness turns on an unlikely possibility, though – namely, that **ZF** could be inconsistent. In view of such a possible application, we suggest that the approach to paraconsistent set theory adopted by the authors is aimed at presenting alternative versions of **ZF** that are more "cautious" in the sense that they would be able to withstand contradictions, should they ever arise within **ZF**. For this reason, we believe that those theories should not be viewed as competitors to classical set theories, but rather as interesting and possibly useful variations of it, whose mathematical properties are nonetheless worth investigating.

Paraconsistent Logic: Consistency, Contradiction and Negation is a comprehensive text on the **LFI**s and fulfills an important gap in the literature on paraconsistency. A huge amount of significant results is presented for the first time in a single text, providing the reader with an extensive survey of the research in the area. Moreover, the content of the book is not limited to the achievements of the so-called *Brazilian* *school of logic*, but also encompasses contributions coming from other areas and research groups. As a result, it is highly recommended for everyone interested in both the formal and the philosophical aspects of paraconsistency, including mathematicians, linguistics, computer scientists, and philosophers of language, mathematics and science.

References

CARNIELLI, W., CONIGLIO, M. Paraconsistent Logic: Consistency, Contradiction and Negation. Logic, Epistemology, and the Unity of Science Series. New York: Springer, 2016.