An epistemic approach to paraconsistency: a logic of evidence and truth

Abilio Rodrigues Filho

Joint work with Walter Carnielli

Federal University of Minas Gerais, Brazil
New College, Oxford, UK

abilio.rodrigues@gmail.com

Logic Colloquium 2018

Udine

Overview

Paraconsistent logics

- The principle of explosion does not hold: $A, \neg A \nvdash B$.
- A paraconsistent logic accepts (some) contradictions without triviality.

What is the nature of contradictions that are accepted in paraconsistent logics?

No true contradictions

- We reject dialetheism, the view according to which there are true contradictions – e.g. Priest and Berto, *Dialetheism*, Stanford.
- We do not endorse a metaphysically neutral position about the nature of contradictions.
- In order to endorse a paraconsistent logic and reject dialetheism it is necessary to attribute a property weaker than truth to pairs of contradictory propositions A and $\neg A$.

Contradictions as conflicting evidence

- A paraconsistent logic may be concerned with a notion **weaker than truth** that allows an intuitive understanding of contradictions.
- A and $\neg A$ may be understood as some kind of 'conflicting information', namely, that there is **conflicting evidence** about A.

A holds $\rightsquigarrow$ **evidence** that A is true $\rightsquigarrow$ **reasons for believing** that A is true.

$\neg A$ holds $\rightsquigarrow$ **evidence** that A is false $\rightsquigarrow$ **reasons for believing** that A is false.

Contradictions as conflicting evidence

- Positive evidence and negative evidence are two primitive, independent and non-complementary notions.
- Four scenarios with respect to the evidence for a proposition A:

1. No evidence at all: both A and $\neg A$ do not hold;
2. Only evidence that A is true: A holds, $\neg A$ does not hold.
3. Only evidence that A is false: A does not hold, $\neg A$ holds.
4. Conflicting evidence: both A and $\neg A$ hold.

Paraconsistency as preservation of evidence

- The *Basic Logic of Evidence* (*BLE*), is a paraconsistent and paracomplete formal system capable of expressing **preservation of evidence**, instead of preservation of truth.
- *BLE* ends up being equivalent to Nelson's logic *N4*, but the motivations are quite different – Nelson was interested in constructive mathematics.
- *BLE*: supposing the availability of evidence for the premises, we ask whether an inference rule yields a conclusion for which evidence is available.

The Basic Logic of Evidence – *BLE*

$$\frac{A \quad B}{A \wedge B}\ \wedge I \qquad \frac{A \wedge B}{A}\ \wedge E \quad \frac{A \wedge B}{B} \qquad \frac{\begin{array}{c}[A]\\ \vdots \\ B\end{array}}{A \to B}\ \to I \qquad \frac{A \to B \quad A}{B}\ \to E$$

$$\frac{A}{A \vee B}\ \vee I \quad \frac{B}{A \vee B} \qquad \frac{A \vee B \quad \begin{array}{c}[A]\\ \vdots \\ C\end{array} \quad \begin{array}{c}[B]\\ \vdots \\ C\end{array}}{C}\ \vee E$$

$$\frac{\neg A \quad \neg B}{\neg(A \vee B)}\ \neg \vee I \qquad \frac{\neg(A \vee B)}{\neg A}\ \neg \vee E \quad \frac{\neg(A \vee B)}{\neg B} \qquad \frac{A \quad \neg B}{\neg(A \to B)}\ \neg \to I \qquad \frac{\neg(A \to B)}{A}\ \neg \to E \quad \frac{\neg(A \to B)}{\neg B}$$

$$\frac{\neg A}{\neg(A \wedge B)}\ \neg \wedge I \quad \frac{\neg B}{\neg(A \wedge B)} \qquad \frac{\neg(A \wedge B) \quad \begin{array}{c}[\neg A]\\ \vdots \\ C\end{array} \quad \begin{array}{c}[\neg B]\\ \vdots \\ C\end{array}}{C}\ \neg \wedge E$$

$$\frac{A}{\neg\neg A}\ DNI \qquad \frac{\neg\neg A}{A}\ DNE$$

The Basic Logic of Evidence – Inversion principle

A little tweak in the inversion principle (Gentzen 1935, and Prawitz 1965):

> *Let* α *be an application of an elimination rule that has* B *as consequence. Then, any* κ *that is evidence for the major premise of* α*, when combined with evidence for the minor premises of* α *(if any), already constitutes evidence for* B*.*
> *The existence of evidence for* B *is thus obtainable directly from the existence of evidence for the premises, without the addition of* α*.*

$$\frac{A \quad \neg B}{\neg(A \to B)} \neg\to I \qquad \frac{\neg(A \to B)}{A} \neg\to E \; \frac{\neg(A \to B)}{\neg B}$$

and so on.

The Basic Logic of Evidence – Inversion principle

$$\frac{A \quad B}{A \wedge B}\ \wedge I \qquad \frac{A \wedge B}{A}\ \wedge E\ \frac{A \wedge B}{B} \qquad \frac{\begin{matrix}[A]\\ \vdots \\ B\end{matrix}}{A \to B}\ \to I \qquad \frac{A \to B \quad A}{B}\ \to E$$

$$\frac{A}{A \vee B}\ \vee I\ \frac{B}{A \vee B} \qquad \frac{A \vee B \quad \begin{matrix}[A]\\ \vdots \\ C\end{matrix} \quad \begin{matrix}[B]\\ \vdots \\ C\end{matrix}}{C}\ \vee E$$

$$\frac{\neg A \quad \neg B}{\neg(A \vee B)}\ \neg\vee I \qquad \frac{\neg(A \vee B)}{\neg A}\ \neg\vee E\ \frac{\neg(A \vee B)}{\neg B} \qquad \frac{A \quad \neg B}{\neg(A \to B)}\ \neg\to I \qquad \frac{\neg(A \to B)}{A}\ \neg\to E\ \frac{\neg(A \to B)}{\neg B}$$

$$\frac{\neg A}{\neg(A \wedge B)}\ \neg\wedge I\ \frac{\neg B}{\neg(A \wedge B)} \qquad \frac{\neg(A \wedge B) \quad \begin{matrix}[\neg A]\\ \vdots \\ C\end{matrix} \quad \begin{matrix}[\neg B]\\ \vdots \\ C\end{matrix}}{C}\ \neg\wedge E$$

$$\frac{A}{\neg\neg A}\ DNI \qquad \frac{\neg\neg A}{A}\ DNE$$

The Basic Logic of Evidence – Symmetry

$$\frac{A}{A \vee B} \vee I$$

Suppose κ is positive evidence for A. Then, κ is also positive evidence for any disjunction $A \vee B$.

$$\frac{\neg A}{\neg(A \wedge B)} \neg \wedge I$$

Suppose κ is negative evidence for A, i.e. positive evidence for $\neg A$. Then κ is also negative evidence for any conjunction $A \wedge B$, i.e. positive evidence for $\neg(A \wedge B)$.

and so on.

The Basic Logic of Evidence – Symmetry

$$\frac{A \quad B}{A \wedge B}\ \wedge I \qquad \frac{A \wedge B}{A}\ \wedge E\ \frac{A \wedge B}{B} \qquad \frac{\begin{matrix}[A]\\ \vdots \\ B\end{matrix}}{A \to B}\ \to I \qquad \frac{A \to B \quad A}{B}\ \to E$$

$$\frac{A}{A \vee B}\ \vee I\ \frac{B}{A \vee B} \qquad \frac{A \vee B \quad \begin{matrix}[A]\\ \vdots \\ C\end{matrix} \quad \begin{matrix}[B]\\ \vdots \\ C\end{matrix}}{C}\ \vee E$$

$$\frac{\neg A \quad \neg B}{\neg(A \vee B)}\ \neg \vee I \qquad \frac{\neg(A \vee B)}{\neg A}\ \neg \vee E\ \frac{\neg(A \vee B)}{\neg B} \qquad \frac{A \quad \neg B}{\neg(A \to B)}\ \neg \to I \qquad \frac{\neg(A \to B)}{A}\ \neg \to E\ \frac{\neg(A \to B)}{\neg B}$$

$$\frac{\neg A}{\neg(A \wedge B)}\ \neg \wedge I\ \frac{\neg B}{\neg(A \wedge B)} \qquad \frac{\neg(A \wedge B) \quad \begin{matrix}[\neg A]\\ \vdots \\ C\end{matrix} \quad \begin{matrix}[\neg B]\\ \vdots \\ C\end{matrix}}{C}\ \neg \wedge E$$

$$\frac{A}{\neg\neg A}\ DNI \qquad \frac{\neg\neg A}{A}\ DNE$$

The logic of evidence and truth – LET_J

The Logic of Evidence and Truth (LET_J) is obtained by extending the language of BLE with a classicality operator $\circ$ and adding the following inference rules:

$$\frac{\circ A \quad A \quad \neg A}{B}\ EXP^{\circ} \qquad \frac{\circ A \quad \begin{matrix}[A]\\ \vdots \\ B\end{matrix} \quad \begin{matrix}[\neg A]\\ \vdots \\ B\end{matrix}}{B}\ PEM^{\circ}$$

- The operator $\circ$ works as a **context switch**: if $\circ A, \circ B, \circ C \ldots$ hold, the *argumentative context* of $A, B, C \ldots$ is classical.
- $\circ A \wedge A$ holds $\rightsquigarrow$ A is true.
- $\circ A \wedge \neg A$ holds $\rightsquigarrow$ A is false.

The intended interpretation of LET_J

- When $\circ A$ does not hold, four scenarios (non-conclusive evidence):
  1. Only evidence that A is true: A holds, $\neg A$ does not hold.
  2. Only evidence that A is false: $\neg A$ holds, A does not hold.
  3. No evidence at all: both A and $\neg A$ do not hold.
  4. Conflicting evidence: both A and $\neg A$ hold.

- When $\circ A$ holds, two scenarios (truth and falsity):
  5. 'A holds' $\rightsquigarrow$ 'there is conclusive evidence that A is true';
  6. '$\neg A$ holds' $\rightsquigarrow$ 'there is conclusive evidence that A is false'.

Valuation semantics for BLE and LET_J

- *Valuation semantics* have been proposed for the logics of da Costa C_n hierarchy (da Costa & Alves 1977, Loparic & Alves 1980, Loparic 1986), intuitionistic logic (Loparic 2010), and several Logics of Formal Inconsistency (*LFI*s) (Carnielli, Coniglio & Marcos 2007, Carnielli & Coniglio 2016).
- Given a language L, valuations are functions from the set of formulas of L to $\{0, 1\}$ according to certain conditions that somehow 'represent' the axioms and/or rules of inference.
- The attribution of the value *0* to a formula A means that A *does not hold*, and the value *1* means that A *holds*.
- Valuation semantics are better seen as mathematical tools that represent the inference rules in such a way that some technical results can be obtained.

Valuation semantics for BLE and LET_J

A semivaluation s for BLE is a function from the set S_1 of formulas to $\{0, 1\}$ such that:

(i) if $s(A) = 1$ and $s(B) = 0$, then $s(A \to B) = 0$,

(ii) if $s(B) = 1$, then $s(A \to B) = 1$,

(iii) $s(A \wedge B) = 1$ iff $s(A) = 1$ and $s(B) = 1$,

(iv) $s(A \vee B) = 1$ iff $s(A) = 1$ or $s(B) = 1$,

(v) $s(A) = 1$ iff $s(\neg\neg A) = 1$,

(vi) $s(\neg(A \wedge B)) = 1$ iff $s(\neg A) = 1$ or $s(\neg B) = 1$,

(vii) $s(\neg(A \vee B)) = 1$ iff $s(\neg A) = 1$ and $s(\neg B) = 1$,

(viii) $s(\neg(A \to B)) = 1$ iff $s(A) = 1$ and $s(\neg B) = 1$.

A semivaluation s for LET_J is a function from the set S_2 of formulas to $\{0, 1\}$ that satisfies the clauses (i)-(viii) above plus the following clause:

(ix) if $s(\circ A) = 1$, then $s(A) = 1$ if and only if $s(\neg A) = 0$.

Valuation semantics for *BLE* and LET_J

A *valuation* for BLE/LET_J is a semivaluation for which the condition below holds:

(Val) For all formulas of the form $A_1 \to (A_2 \to ... \to (A_n \to B)...)$ with B not of the form $C \to D$:
if $s(A_1 \to (A_2 \to ... \to (A_n \to B)...)) = 0$, then there is a semivaluation s' such that for every i, $1 \leq i \leq n$, $s(A_i) = 1$ and $s(B) = 0$.

- Logical consequence is defined as usual: $\Gamma \vDash A$ if and only if for every valuation v, if $v(B) = 1$ for all $B \in \Gamma$, then $v(A) = 1$.
- The semantics above is sound and complete, and provides a decision procedure for *BLE* and LET_J by means of the *quasi-matrices* (see Carnielli & Rodrigues 2017).

Extending inferential (proof-theoretical) semantics

- The plan: to apply to the logic *BLE* the basic idea of inferential (or proof-theoretical) semantics, developed for intuitionistic logic.
- What is at stake in each case is what is expressed by the inference rules:

Intuitionistic logic: preservation of **constructive proof**.

BLE: preservation of **availability of evidence**.

- In both cases, the *meanings* of the connectives are given by the inference rules.

The context of intuitionistic logic

- The idea that the meanings of intuitionistic connectives is to be explained in terms of solutions of problems (Kolmogorov), assertibility conditions (Heyting), and proofs (Troelstra & van Dalen) was based on the claim that intuitionistic logic is not about truth but rather about constructive reasoning in Mathematics.
- So, an assertion of A means that a constructive proof of A is available.

- The **context** is constructive Mathematics.
- The **property of propositions** at stake is availability of a constructive proof.

Inferential semantics for *BLE*

- *BLE* intends to represent **paracomplete and paraconsistent argumentative contexts**.
- The **property of propositions** at stake is availability of **evidence**.
- An inferential semantics for *BLE* needs two primitive notions: positive and negative evidence. Recall that evidence for truth (resp. falsity) is different from absence of evidence for falsity (resp. truth).
- In *BLE*, '$\neg A$ holds' means that there is negative evidence for A. This same evidence is positive evidence for $\neg A$. So,

  κ is negative evidence for A iff κ is positive evidence for $\neg A$.

The Evidence Interpretation for *BLE*

[E1] Positive evidence κ for $A \wedge B$ amounts to positive evidence κ_1 for A and positive evidence κ_2 for B;

[E2] Positive evidence κ for $A \vee B$ amounts to positive evidence κ_1 for A or positive evidence κ_2 for B;

[E3] Positive evidence κ for $A \rightarrow B$ is given when the supposition that there is positive evidence κ_1 for A leads to the conclusion that there is evidence κ_2 for B;

[E4] Negative evidence κ for $A \wedge B$ amounts to negative evidence κ_1 for A or negative evidence κ_2 for B;

[E5] Negative evidence κ for $A \vee B$ amounts to negative evidence κ_1 for A and negative evidence κ_2 for B;

[E6] Negative evidence κ for $A \rightarrow B$ amounts to positive evidence κ_1 for A and negative evidence κ_2 for B;

[E7] Negative evidence κ for $\neg A$ amounts to positive evidence κ for A.

The Evidence Interpretation for *BLE*

[E1] Positive evidence κ for $A \wedge B$ amounts to positive evidence κ_1 for A and positive evidence κ_2 for B;

[E2] Positive evidence κ for $A \vee B$ amounts to positive evidence κ_1 for A or positive evidence κ_2 for B;

[E3] Positive evidence κ for $A \rightarrow B$ is given when the supposition that there is positive evidence κ_1 for A leads to the conclusion that there is evidence κ_2 for B;

[E4] Negative evidence κ for $A \wedge B$ amounts to negative evidence κ_1 for A or negative evidence κ_2 for B;

[E5] Negative evidence κ for $A \vee B$ amounts to negative evidence κ_1 for A and negative evidence κ_2 for B;

[E6] Negative evidence κ for $A \rightarrow B$ amounts to positive evidence κ_1 for A and negative evidence κ_2 for B;

[E7] Negative evidence κ for $\neg A$ amounts to positive evidence κ for A.

The Basic Logic of Evidence – *BLE*

$$\frac{A \quad B}{A \wedge B}\ \wedge I \qquad \frac{A \wedge B}{A}\ \wedge E\ \frac{A \wedge B}{B} \qquad \frac{\begin{array}{c}[A]\\ \vdots \\ B\end{array}}{A \to B}\ \to I \qquad \frac{A \to B \quad A}{B}\ \to E$$

$$\frac{A}{A \vee B}\ \vee I\ \frac{B}{A \vee B} \qquad \frac{A \vee B \quad \begin{array}{c}[A]\\ \vdots \\ C\end{array} \quad \begin{array}{c}[B]\\ \vdots \\ C\end{array}}{C}\ \vee E$$

$$\frac{\neg A \quad \neg B}{\neg(A \vee B)}\ \neg \vee I \qquad \frac{\neg(A \vee B)}{\neg A}\ \neg \vee E\ \frac{\neg(A \vee B)}{\neg B} \qquad \frac{A \quad \neg B}{\neg(A \to B)}\ \neg \to I \qquad \frac{\neg(A \to B)}{A}\ \neg \to E\ \frac{\neg(A \to B)}{\neg B}$$

$$\frac{\neg A}{\neg(A \wedge B)}\ \neg \wedge I\ \frac{\neg B}{\neg(A \wedge B)} \qquad \frac{\neg(A \wedge B) \quad \begin{array}{c}[\neg A]\\ \vdots \\ C\end{array} \quad \begin{array}{c}[\neg B]\\ \vdots \\ C\end{array}}{C}\ \neg \wedge E$$

$$\frac{A}{\neg\neg A}\ DNI \qquad \frac{\neg\neg A}{A}\ DNE$$

Inferential semantics for *BLE*

- The meaning of a composite expression A is given by the meanings of its constituents and the way they are combined.
- The meaning of $A * B$, $* \in \{\wedge, \vee, \to\}$, is given by the meanings of A, B and the meaning of the respective connective, given by the introduction rule $*I$.
- The meaning of $\neg A$ is primitive if A is a propositional letter, and if $A = \neg B$, it depends on the meaning of B and on the rule *DNI*.
- Otherwise, if $A = B * C$, the meaning of $\neg A$ depends on B, on C, and on the respective rule $\neg * I$.

Inferential semantics for LET_J

- Recall that LET_J divides propositions into two contexts: the first governed by BLE, and the second, in the scope of $\circ$, subject to classical logic.
- The classical meanings of the logical connectives $\wedge$, $\vee$, $\rightarrow$ and $\neg$ cannot be given by their introduction rules.
- We suggest that the meaning of a connective $*$ in LET_J, in the contexts governed by classical logic, is given globally by all the truth-preserving inferences in which $*$ occurs.

What we have done?

- We presented an approach to paraconsistency in which contradictions are understood as conflicting evidence.
- Two natural deduction system have been proposed: BLE, conceived to express preservations of evidence instead of preservation of truth, and LET_J, an extension of BLE that recovers classical logic for propositions for which there is conclusive evidence available.
- Adequate valuation semantics and decision procedures have been presented for BLE and LET_J.
- The basic ideas of an inferential semantics that would be able to explain the meanings of the expressions of BLE and LET_J in a compositional manner have been presented.

Now, what comes next?

What comes next?

1. To provide a probabilistic semantics for BLE and LET_J, where let $P(A) = n$ means that n is the measure of evidence available for A. In such a semantics, the amount of evidence available for a proposition could be quantified, and $P(A) + P(\neg A)$ could be less or greater than 0.

2. To develop the proof-theory of BLE and LET_J, in order to see to what extent the ideas here presented will fit the technical results.

- W. Carnielli, and A. Rodrigues. An epistemic approach to paraconsistency: a logic of evidence and truth. *Synthese*. DOI: 10.1007/s11229-017-1621-7, 2017. Preprint available at http://bit.ly/SYNLETJ.
- W. Carnielli, and A. Rodrigues. Inferential Semantics, Paraconsistency and Preservation of Evidence. Forthcoming, 2018. Preprint available at http://bit.ly/INFSMTS.
- W. Carnielli and M. Coniglio. *Paraconsistent Logic: Consistency, Contradiction and Negation*. Springer, 2016.
- W. Carnielli, M. Coniglio, and J. Marcos. Logics of Formal Inconsistency. *Handbook of Philosophical Logic (2nd ed.)*, vol. 14. Springer, 2007.
- N. C. A. da Costa and E. Alves. A semantical analysis of the calculi C_n. *NotreDame Journal of Formal Logic*, 18(4):621-630, 1977.
- A. Loparic. A semantical study of some propositional calculi. *The Journal of Non-Classical Logic*, 3(1):73-95, 1986.
- A. Loparic. Valuation semantics for intuitionistic propositional calculus and some of its subcalculi. *Principia*, 14(1):125-133, 2010.
- A. Loparic and E. Alves. The semantics of the systems C_n of da Costa. *Proceedings of the III Brazilian Conference on Mathematical Logic*, Recife, 1979.

Muito obrigado!

abilio.rodrigues@gmail.com