What do ‘evidence’ and ‘truth’ mean in the logics of evidence and truth

Abilio Rodrigues

Federal University of Minas Gerais, Brazil

abilio.rodrigues@gmail.com

Semantical and Computational Aspects of Non-Classical Logics
SCAN 2023

June, 2023

Logics of evidence and truth

Logics of evidence and truth (*LET*s) are **paracomplete** and **paraconsistent** logics that extend the **logic of first-degree entailment** (*FDE*) with a **classicality operator** $\circ$ that recovers classical negation for sentences in its scope.

$A, \neg A \nvdash B$, while $\circ A, A, \neg A \vdash B$

$\nvdash A \vee \neg A$, while $\circ A \vdash A \vee \neg A$.

*LET*s express **preservation of evidence** and **preservation of truth** (or conclusive evidence) in the same formal system.

*LET*s can also be interpreted as **preservation of information**, which can be **reliable or unreliable**.

Belnap-Dunn 4-valued logic

- Belnap (1977) introduced a four-valued semantics for the logic of first-degree entailment (FDE) to represent information stored in a possibly inconsistent and incomplete database.
- Positive information A: the information conveyed by a sentence A; negative information A: the information conveyed by $\neg A$.

  $v(A) = T$: A holds and $\neg A$ does not hold (only positive information);
  $v(A) = F$: $\neg A$ holds and A does not hold (only negative information);
  $v(A) = B$: both A and $\neg A$ hold (contradictory information);
  $v(A) = N$: neither A nor $\neg A$ holds (no information at all);
- The semantic values T and F are not 'ontological'.
- T (resp. F) is a 'told true' (resp. 'told false') sign, in the sense that the computer 'has been told' (only) that a sentence A is true (resp. false).

From da Costa to logics of evidence and truth

- Logics of evidence and truth (*LET*s) developed from logics of formal inconsistency (*LFI*s) (e.g. Carnielli, Coniglio & Marcos 2007).
- Both *LET*s and *LFI*s are part of an evolutionary line that starts with da Costa's seminal work on paraconsistency (da Costa 1963).
- da Costa's work was based on two main ideas:

  (i) to express the metatheoretical notion of consistency at the object language level, and

  (ii) to divide the sentences of the language into two groups, one subjected to classical logic and the other subjected to a subclassical non-explosive logic.

Combining Belnap, Dunn, and da Costa

- Logics of evidence and truth combine da Costa's ideas with Belnap and Dunn proposal of a 4-valued logic.
- More precisely, *LET*s are **logics of formal inconsistency and undeterminedness** (*LFIU*s, see e.g. Marcos 2005) **extending** *FDE* (i.e. with De Morgan and full double negation).

  $A, \neg A \nvdash B$, while $\circ A, A, \neg A \vdash B$ *logics of formal inconsistency*,

  $\nvdash A \vee \neg A$, while $\circ A \vdash A \vee \neg A$ *logics of formal undeterminedness*.

⇒ The intuitive interpretation of $\circ A$ is that the evidence (positive or negative) available for A is conclusive, or that the information A, positive or negative, is reliable.

⇒ It is assumed that conclusive evidence and reliable information behave classically.

Six scenarios, instead of four

LETs add two more scenarios to the four scenarios expressed by FDE.

When $\circ A$ does not hold:

(i) A holds, $\neg A$ does not hold: only positive information A;

(ii) A does not hold, $\neg A$ holds: only negative information A;

(iii) Both A and $\neg A$ hold: contradictory information about A;

(iv) Neither A nor $\neg A$ hold: no information at all about A.

When $\circ A$ holds:

(v) A holds: **reliable positive information** A;

(vi) $\neg A$ holds: **reliable negative information** A.

LETs establish a distinction between the scenarios (i)/(v), and (ii)/(vi) that cannot be established within FDE.

The talk: Part I – the main technical results

- The sentential logic LET_J, which extends Nelson's logic *N4* with a classicality operator $\circ$ and the rules that recover classical negation for sentences in the scope of $\circ$. Recall that *N4* is an extension of *FDE* with a **constructive implication**.

  A sound and complete Kripke semantics for LET_J.

- The first-order logic $Q_v LET_F^-$ that is an extension of *FDE* with quantifiers and with a classicality operator $\circ$ and the respective rules.

  A sound and complete Kripke semantics with **variable domains** for $Q_v LET_F^-$.

- The sentential logic LET_F^+ together with a sound, complete and decidable **deterministic six-valued semantics**.

The talk: Part II – conceptual aspects

- The notion of evidence that underlies the intuitive interpretation of *LET*s in terms of conclusive and non-conclusive evidence.

- The notion of information as *meaningful data* that underlies the interpretation of *LET*s in terms of reliable and unreliable information.

- The connections between information as meaningful data and evidence, as well as an attempt to define the latter based on the former.

Technical aspects

The logics LET_F^- and LET_F

- The logic LET_F^- is a minimal LET that extends FDE with the classicality operator $\circ$ and the rules

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

- LET_J extends LET_F^- with implication and the rules

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

Natural deduction rules for LET_F^- and LET_J

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

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

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

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

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

$FDE + \to = N4 \qquad N4 + EXP^{\circ} + PEM^{\circ} = LET_J \qquad FDE + EXP^{\circ} + PEM^{\circ} = LET_F^-$

Valuation semantics for LET_F^-

A *valuation* for LET_F^- is a function $\rho : \mathcal{L} \to \{0,1\}$ satisfying the following properties:

(v1) $\rho(A \wedge B) = 1$ iff $\rho(A) = 1$ and $\rho(B) = 1$;

(v2) $\rho(A \vee B) = 1$ iff $\rho(A) = 1$ or $\rho(B) = 1$;

(v3) $\rho(\neg(A \wedge B)) = 1$ iff $\rho(\neg A) = 1$ or $\rho(\neg B) = 1$;

(v4) $\rho(\neg(A \vee B)) = 1$ iff $\rho(\neg A) = 1$ and $\rho(\neg B) = 1$;

(v5) $\rho(\neg\neg A) = 1$ iff $\rho(A) = 1$;

(v6) $\rho(A \to B) = 1$ iff $\rho(A) = 0$ or $\rho(B) = 1$;

(v7) $\rho(\neg(A \to B)) = 1$ iff $\rho(A) = 1$ and $\rho(\neg B) = 1$;

(v8) if $\rho(\circ A) = 1$, then: $\rho(\neg A) = 1$ iff $\rho(A) = 0$;

This semantics is sound, complete, and decidable, but non-deterministic.

The semantic value of $\neg A$ is not determined by the semantic value of A, nor the value of $\circ A$ is determined by the values of A and $\neg A$

Kripke semantics for LET_J

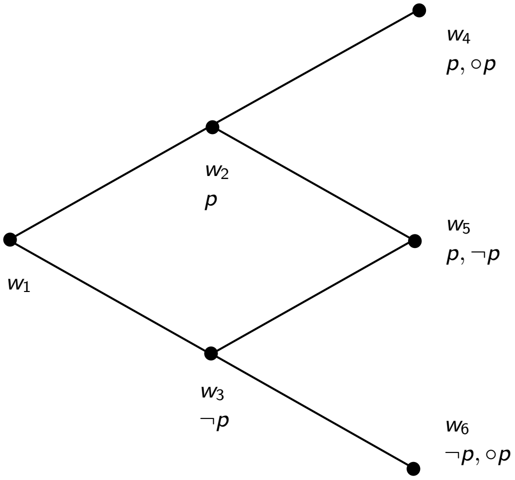
A *Kripke model* $\mathcal{M}$ for LET_J is a structure $\langle W, \leq, v \rangle$ such that W is a non-empty set of *stages*, the *accessibility relation* $\leq$ is a partial order on W, and $v : \mathcal{L}_J \times W \longrightarrow \{0, 1\}$ is a *valuation function* satisfying the following conditions, for every $w \in W$:

1. $v(A \wedge B, w) = 1$ iff $v(A, w) = 1$ and $v(B, w) = 1$;
2. $v(A \vee B, w) = 1$ iff $v(A, w) = 1$ or $v(B, w) = 1$;
3. $v(\neg\neg A, w) = 1$ iff $v(A, w) = 1$;
4. $v(\neg(A \wedge B), w) = 1$ iff $v(\neg A, w) = 1$ or $v(\neg B, w) = 1$;
5. $v(\neg(A \vee B), w) = 1$ iff $v(\neg A, w) = 1$ and $v(\neg B, w) = 1$;
7. $v(A \rightarrow B, w) = 1$ iff for every $w' \geq w$, if $v(A, w') = 1$, then $v(B, w') = 1$;
8. $v(\neg(A \rightarrow B), w) = 1$ iff $v(A, w) = 1$ and $v(\neg B, w) = 1$;
6. $v(\circ A, w) = 1$ only if exactly one of the following conditions obtains:
   For every $w' \geq w$, $v(A, w') = 1$ and $v(\neg A, w') = 0$;
   For every $w' \geq w$, $v(A, w') = 0$ and $v(\neg A, w') = 1$;

P. If $v(A, w) = 1$, then for every $w' \geq w$, $v(A, w') = 1$, for every $A \in \mathcal{L}_J$.

Kripke models for LET_J

A **database over time that** receives **information from different sources** that may be either **reliable or unreliable**. The six scenarios of *LET*s w.r.t. a sentence p are illustrated by the diagram:



The logic LET_F^-

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

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

$$\frac{A}{\neg\neg A}\ DN\ \frac{\neg\neg A}{A}$$

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

The logic $Q_v LET_F^-$

$$\frac{A(c/x)}{\forall x A}\ \forall I \qquad \frac{\forall x A}{A(c/x)}\ \forall E \qquad \frac{A(c/x)}{\exists x A}\ \exists I \qquad \frac{\exists x A \qquad \begin{array}{c}[A(c/x)]\\ \vdots \\ C\end{array}}{C}\ \exists E$$

$$\frac{\neg A(c/x)}{\neg\forall x A}\ \neg\forall I \qquad \frac{\neg\forall x A \qquad \begin{array}{c}[\neg A(c/x)]\\ \vdots \\ C\end{array}}{C}\ \neg\forall E \qquad \frac{\neg A(c/x)}{\neg\exists x A}\ \neg\exists I \qquad \frac{\neg\exists x A}{\neg A(c/x)}\ \neg\exists E$$

$$\frac{A}{A'}\ AV$$

The rules $\forall I$, $\exists E$, $\neg\forall E$ and $\neg\exists I$ are subjected to the usual restrictions.
In AV, A' is any *alphabetic variant of* A, that is any sentence that differs from A only in some of its bound variables.

Kripke models for $Q_v LET_F^-$

A *Kripke model* $\mathcal{K}$ for $Q_v LET_F^-$ is a structure $\langle W, \leq, d, I \rangle$ such that W is a non-empty set; $\leq$ is a partial order on W; d is a function that assigns a non-empty set to each element of W such that if $w \leq w'$, then $d(w) \subseteq d(w')$; and I is a function on $(\mathcal{C} \cup \mathcal{P}) \times W$ such that:

(i) For every individual constant $c \in \mathcal{C}$, $I(c, w) \in d(w)$;

(ii) For every individual constant $c \in \mathcal{C}$ and for every $w, w' \in W$, $I(c, w) = I(c, w')$;

(iii) For every n-ary predicate letter $P \in \mathcal{P}$, $I(P, w)$ is the pair $\langle P_+^w, P_-^w \rangle$ such that $P_+^w, P_-^w \subseteq d(w)^n$.

- P_+ is the *extension* and P_- the *anti-extension* of P.
- $P_+ \cup P_-$ is a subset of $\mathcal{D}^n$.
- However, it may be that $P_+ \cup P_- \neq \mathcal{D}^n$ and $P_+ \cap P_- \neq \emptyset$.

$Q_v LET_F^-$ Kripke models

Let $\mathcal{S}$ be a first-order language and let $\mathcal{K}$ be a Kripke model for $\mathcal{S}$. A mapping $v : Sent(\mathcal{S}_\mathcal{K}) \times W \longrightarrow \{0,1\}$ is a *valuation function* satisfying the following conditions, for every $w \in W$:

1. If $v(A, w) = 1$, and $w' \geq w$, then $v(A, w') = 1$;
2. $v(P(c_1, \ldots, c_n), w) = 1$ if $\widehat{I}(c_i, w)$ is defined for every $1 \leq i \leq n$ and $\langle \widehat{I}(c_1, w), \ldots, \widehat{I}(c_n, w)\rangle \in P_+^w$, for $c_1, \ldots, c_n \in \mathcal{C}_\mathcal{K}$; $= 0$, otherwise.
3. $v(\neg P(c_1, \ldots, c_n), w) = 1$ if $\widehat{I}(c_i, w)$ is defined for every $1 \leq i \leq n$ and $\langle \widehat{I}(c_1, w), \ldots, \widehat{I}(c_n, w)\rangle \in P_-^w$, for $c_1, \ldots, c_n \in \mathcal{C}_\mathcal{K}$; $= 0$, otherwise.

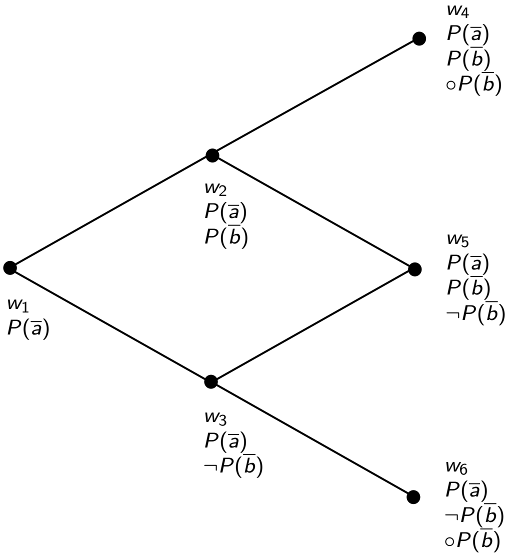
   [. . .]

9. If $v(\circ A, w) = 1$, then $v(A, w) = 1$ iff $v(\neg A, w) = 0$;
10. $v(\forall x A, w) = 1$ iff for every $w' \geq w$, $v(A(\bar{a}/x), w') = 1$, for every $a \in d(w')$;
11. $v(\exists x A, w) = 1$ iff $v(A(\bar{a}/x), w) = 1$, for some $a \in d(w)$;
12. $v(\neg\forall x A, w) = 1$ iff $v(\neg A(\bar{a}/x), w) = 1$, for some $a \in d(w)$;
13. $v(\neg\exists x A, w) = 1$ iff for every $w' \geq w$, $v(\neg A(\bar{a}/x), w') = 1$, for every $a \in d(w')$.

   [. . .]

A $Q_v LET_F^-$*-interpretation* $\mathcal{I}$ for $\mathcal{S}$ is a pair $\langle \mathcal{K}, v\rangle$ such that $\mathcal{K}$ is a Kripke model and v is a valuation function.

The diagram below illustrates the intended interpretation of $Q_v LET_F^-$. It represents **possible stages of a database** over time, where each stage w displays the information that has been stored at w.



Adding propagation rules: the logic LET_F^+

- Propagation of classicality is how classical behavior propagates from less complex to more complex sentences, and vice-versa.
- The logic LET_F^- (as well the LETs investigated so far) enjoy the following property:

  Suppose $\circ\neg^{n_1}A_1, \ldots, \circ\neg^{n_m}A_m$ hold in LET_F, for $n_i \geq 0$ (where $\neg^{n_i}$, $n_i \geq 0$, represents n_i occurrences of negations before the formula A_i).

  Then, for any LET_F-formula B formed with $A_1, \ldots, A_m$ over $\{\neg, \wedge, \vee\}$, B behaves classically, that is, $B \vee \neg B$ and $B, \neg B \vdash C$ hold.

- Example: if $\circ p$ and $\circ\neg q$ hold, $p \wedge q$, $\neg(p \wedge \neg q)$, $\neg\neg q$, etc. behave classically.
- However, the classicality operator $\circ$ *is not transmitted* from less to more complex sentences – that is, the inferences

  $\circ p \vdash \circ\neg p$ and $\circ p, \circ q \vdash \circ(p * q)$ $(* \in \{\vee, \wedge\})$,

  for example, do not hold.

Adding propagation rules: the logic LET_F^+

- Recall that $\circ A \wedge A$ and $\circ A \wedge \neg A$ are intended to mean that the information conveyed, respectively, by A and by $\neg A$, is considered reliable.
- Positive and negative reliable information behaves like truth and falsity in classical logic.
- $\circ A \wedge A$ implies that $A \vee B$ is also reliable, for any B, no matter whether $\circ B$ holds or not, and so $\circ(A \vee B) \wedge (A \vee B)$ holds.
- But $\circ A \wedge \neg A$ does not imply that $\neg(A \vee B)$ is reliable, because to conclude $\neg(A \vee B)$ both $\neg A$ and $\neg B$ are required. $\circ(A \vee B)$ cannot be inferred. This suggests the validity of the following inferences:

$$\circ A, A \vdash \circ(A \vee B) \wedge (A \vee B),$$
$$\circ A, \neg A, \circ B, \neg B \vdash \circ(A \vee B) \wedge \neg(A \vee B).$$
$$\circ A, \neg A \vdash \circ(A \wedge B) \wedge \neg(A \wedge B),$$
$$\circ A, A, \circ B, B \vdash \circ(A \wedge B) \wedge (A \wedge B),$$

Recall the logic LET_F^-

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

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

$$\frac{A}{\neg\neg A}\ DN\ \frac{\neg\neg A}{A}$$

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

$$FDE + EXP^\circ + PEM^\circ = LET_F^-$$

Adding propagation rules: the logic LET_F^+

$A^T \stackrel{\text{def}}{=} \circ A \wedge A$ $\qquad$ $A^F \stackrel{\text{def}}{=} \circ A \wedge \neg A$ $\qquad$ $\dfrac{}{\circ\circ A}\,[I\circ]$ $\qquad$ $\dfrac{\circ A}{\circ\neg A}\,[I\neg\circ]$ $\qquad$ $\dfrac{\circ\neg A}{\circ A}\,[E\neg\circ]$

$$\frac{A^T \quad B^T}{(A \wedge B)^T}\,[I\wedge T] \qquad \frac{A^F}{(A \wedge B)^F}\,[I\wedge F]\,\frac{B^F}{(A \wedge B)^F}$$

$$\frac{A^T}{(A \vee B)^T}\,[I\vee T]\,\frac{B^T}{(A \vee B)^T} \qquad \frac{A^F \quad B^F}{(A \vee B)^F}\,[I\vee F]$$

$$\frac{(A \wedge B)^T}{A^T}\,[E\wedge T]\,\frac{(A \wedge B)^T}{B^T} \qquad \frac{(A \wedge B)^F \quad \begin{matrix}[A^F]\\ \vdots \\ C\end{matrix} \quad \begin{matrix}[B^F]\\ \vdots \\ C\end{matrix}}{C}\,[E\wedge F]$$

$$\frac{(A \vee B)^T \quad \begin{matrix}[A^T]\\ \vdots \\ C\end{matrix} \quad \begin{matrix}[B^T]\\ \vdots \\ C\end{matrix}}{C}\,[E\vee T] \qquad \frac{(A \vee B)^F}{A^F}\,[E\vee F]\,\frac{(A \vee B)^F}{B^F}$$

The six scenarios of LETs

Recall the valuation semantics for LET_F^-: ρ is a function from the sentences of the language to $\{0,1\}$; the value 1 assigned to A can be read as 'A holds', and the value 0 as 'A does not hold',

When $\circ A$ does not hold, $\rho(\circ A) = 0$: $(\rho(A), \rho(\neg A), \rho(\circ A))$

i. **T_0:** $\rho(A) = 1$, $\rho(\neg A) = 0$: only positive information A; (1,0,0)

ii. **F_0:** $\rho(A) = 0$, $\rho(\neg A) = 1$: only negative information A; (0,1,0)

iii. **B:** $\rho(A) = 1$, $\rho(\neg A) = 1$: contradictory information about A; (1,1,0)

iv. **N:** $\rho(A) = 0$, $\rho(\neg A) = 0$: no information at all about A. (0,0,0)

When $\circ A$ holds, $\rho(\circ A) = 1$:

v. **T:** $\rho(A) = 1$, $\rho(\neg A) = 0$: reliable positive information A; (1,0,1)

vi. **F:** $\rho(A) = 0$, $\rho(\neg A) = 1$: reliable negative information A. (0,1,1)

A six-valued deterministic semantics for LET_F^+

$\tilde{\wedge}$	T	T_0	B	N	F_0	F
T	T	T_0	B	N	F_0	F
T_0	T_0	T_0	B	N	F_0	F
B	B	B	B	F_0	F_0	F
N	N	N	F_0	N	F_0	F
F_0	F_0	F_0	F_0	F_0	F_0	F
F	F	F	F	F	F	F

	$\tilde{\neg}$
T	F
T_0	F_0
B	B
N	N
F_0	T_0
F	T

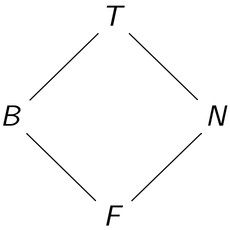
$\tilde{\vee}$	T	T_0	B	N	F_0	F
T	T	T	T	T	T	T
T_0	T	T_0	T_0	T_0	T_0	T_0
B	T	T_0	B	T_0	B	B
N	T	T_0	T_0	N	N	N
F_0	T	T_0	B	N	F_0	F_0
F	T	T_0	B	N	F_0	F

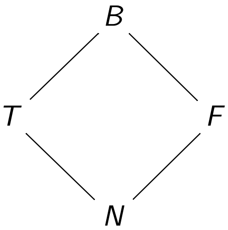
	$\tilde{\circ}$
T	T
T_0	F
B	F
N	F
F_0	F
F	T

M.E. Coniglio & A. Rodrigues. On six-valued logics of evidence and truth expanding Belnap-Dunn four-valued logic. https://arxiv.org/abs/2209.12337

On the lattice structure of LET_F^+

Belnap (1977): two lattice-orderings defined by the four semantic values of *FDE*, **L4** and **A4**.

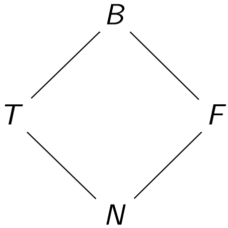




L4 Logical lattice

A4 Approximation lattice

The approximation lattice A4

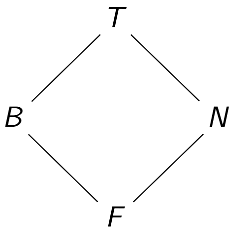


Order is given by inclusion: $a \leq b$ iff $a \subseteq b$.

$T = \{1\}, F = \{0\}, B = \{1, 0\}, N = \emptyset$.

The *amount of information* grows from bottom to top.

The logical lattice L4



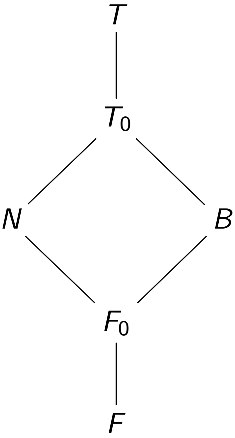
T, F, B, N are pairs (a_1, a_2), where $a_1 = \rho(A)$ and $a_2 = \rho(\neg A)$ in a given bivaluation ρ.

$T = (1, 0), F = (0, 1), B = (1, 1)$, and $N = (0, 0)$.

$(a_1, a_2) \leq (b_1, b_2)$ iff $a_1 \leq b_1$ and $a_2 \geq b_2$.

[T]he worst thing is to be told something is false, simpliciter. You are better off ... in either being told nothing about it, or in being told both that it is true and also that it is false; while of course best of all is to be told it is true, with no muddying of the waters (Belnap 1977).

The logical lattice L6



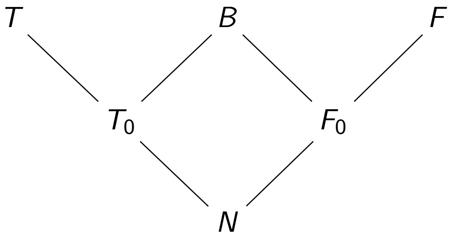
T, T_0, N, B, F_0, F are triples (a_1, a_2, a_3) s.t.
$a_1 = \rho(A)$, $a_2 = \rho(\neg A)$, $a_2 = \rho(\circ A)$.

$T = (1, 0, 1), T_0 = (1, 0, 0), N = (0, 0, 0),$
$B = (1, 1, 0), F_0 = (0, 1, 0), F = (0, 1, 1).$

$(a_1, a_2, a_3) \leq (b_1, b_2, b_3)$ iff $a_1 \leq b_1$, $a_2 \geq b_2$,
$a_3 = (b_1 \sqcap b_3 \sqcap b_1 \sqcap b_3) \sqcup (b_2 \sqcap b_3) \sqcup (b_2 \sqcap b_3)$,
$(\vec{a} \leq \vec{b}$ iff $\vec{a} = \vec{a} \sqcap \vec{b})$.

Still worse than be told that A is false is to be told that A is conclusively false, that is, $\circ A \wedge \neg A$.
Still better than be told that A is true is to be told that A is conclusively true, that is, $\circ A \wedge A$.

The meet semilattice A6



$\mathbf{U} = \{0, 1, \mathrm{c}\}$, c means that the respective information is reliable.

$N = \emptyset,\ T_0 = \{1\},\ F_0 = \{0\},\ T = \{1, \mathrm{c}\},\ B = \{1, 1\},\ F = \{0, c\}$.

$a \leq b$ iff $a \subseteq b$.

The *amount of information* grows from bottom to top.

Conceptual aspects

Evidence and justifications

- The intended intuitive interpretation of *LET*s is based on the notions of positive and negative evidence and on the assumption that people reason classically in the face of conclusive evidence.
- Positive and negative evidence are, respectively, evidence for the truth and for the falsity of a sentence A, and the presence of conclusive evidence for A is expressed by $\circ A$.
- Non-conclusive evidence requires a paracomplete and paraconsistent logic, since there may be circumstances in which there is simultaneous positive and negative evidence for a sentence A, and circumstances in which there is no evidence for A at all.

On the notion of evidence

- The word ‘evidence’ does not have a unified meaning, either in philosophy or in natural language.
- The notion of evidence that underlies the intended interpretation of *LET*s fits a **well-established use of this word, both in philosophical discussions and in the ordinary usage of language**.
- Note that the aim is not to formalize *the* notion of evidence, but rather *a* notion of evidence that is indeed used in philosophy and ordinary language.

On the notion of evidence

(i) 'Evidence for a sentence A' are 'reasons for believing in or accepting A'.

(ii) These reasons for accepting may be non-conclusive (or *non-factive*), therefore evidence for A does not imply the truth of A;

(iii) Acceptance here *is not* rationally accepting A but merely admitting A in a certain context of reasoning – in particular, whenever it is said that a contradiction has been accepted, this does not mean that it has been rationally accepted;

(iv) There may be conflicting evidence for a pair of sentences A and $\neg A$, as well as no evidence at all for either A and $\neg A$;

(v) Evidence for A is objective, i.e. it is independent from the belief of an agent in A.

(vi) *LET*s do not intend to represent cognitive relations between agents and propositions. It is not an account of any kind of propositional attitude. (Carnielli & Rodrigues 2017)

Evidence, reasons, and justifications

The following ideas can be found in a number of philosophical texts:
See Rodrigues and Carnielli, 2022, Section 2.

- The idea of evidence as *reasons for believing*;
- The connection between evidence and justification;
- The notions of positive and negative evidence, which may occur together;
- The distinction between non-conclusive (or potential) evidence and evidence that supposedly implies truth;
- The notion of non-factive evidence, which is a non-conclusive evidence that does not imply truth;
- We may then conclude that the expressions 'x is evidence for A' and 'x justifies A' are synonymous in several contexts – as well as 'non-conclusive evidence' and 'non-factive justification'.

Evidence in natural language

- We find that in a number of circumstances evidence is considered to be non-conclusive, objective, and independent of belief and truth.
- *The Concise Oxford English Dictionary* describes evidence as "**information indicating** whether a belief or proposition is true or valid" (our emphasis).
- *The Cambridge Dictionary Online* explains evidence as "one or more **reasons for believing** that something is or is not true" (our emphasis).
- Google search for 'conflicting evidence', 'lack of evidence', 'conclusive evidence', 'non-conclusive evidence', 'inconclusive evidence', 'partial evidence', 'false evidence', or 'misleading evidence', restricted to reliable sources of English language usage like 'nytimes.com', 'bbc.com', and '.edu', we find thousands of collocations with 'evidence' in line with the notion of evidence as characterized by us.

I'd like to make clear that...

⇒ We are not interested in defining a general notion of evidence, nor in a epistemological investigation of the concept of evidence.

⇒ The proposal is to formalize the deductive behavior of a notion (not *the* notion) of evidence that fits with a usage (not *the* usage) of the word 'evidence' in philosophy and natural language.

The criticisms of the notion of evidence stick to the idea that people should not believe in what is not justified by strong evidence.
In this way, the notion of evidence is read with a bias that assumes that evidence justifies a doxastic attitude toward a proposition.
Such doxastic attitude has to comply with the standards put forth by epistemological discussions about knowledge, justification, etc.
But the logics of evidence and truth have a different purpose: to provide a formal system capable of formalizing positive and negative, conclusive and non-conclusive evidence. (Rodrigues & Carnielli 2022)

LETs as information-based logics

- In a wide sense, an information-based logic is any logic suitable for processing information in the sense of taking a database as a set of premises and drawing conclusions from these premises in a sensible way.
- Since databases often contain contradictions, explosion cannot be valid in an information-based logic.
- *LET*s can also be interpreted as information-based logics.
- Instead of positive (negative) evidence, think of positive (negative) information, and $\circ A$ means that the information A, positive or negative, is reliable.

Evidence vs. information

- The notions of evidence and information are closely related
- The notion of evidence can be made more precise based on the definition of information as *meaningful data* and its connection with the definition of *knowledge as justified true belief*.

  [Information is] what is left from knowledge when you subtract justification, truth, belief, and any other ingredients such as reliability that relate to justification (J.M. Dunn 2008 p. 589).

- This idea of information is a propositional version of the 'standard account' of information, as we read in Fetzer (2004):

  [T]he "standard account" of information would define it simply as meaningful data, which might or might not be true. (Fetzer 2004 p. 224)

A definition of evidence

- Information as meaningful data may appear in both linguistic and non-linguistic forms.
- It may be conveyed by sentences as well as by things like blood spots, details in a photograph, fossil records, fingerprints on a gun, documents, etc.
- These 'pieces of information' can be considered *justifications* for certain sentences, and such justifications may be **non-factive**, i.e., fallible, partial, or wrong – or even in some cases conclusive.
- A definition of *evidence for a sentence A* would be a pair $\langle \Theta, A \rangle$ where Θ contains pieces of information, linguistic or otherwise, that are considered justifications (possibly non-factive) for A.
- Information = knowledge − justification − truth − belief
  Evidence = information + a justification that may be non-factive.

  The terminology 'non-factive justification/evidence' was borrowed from Fitting (2016), where he proposes a justification logic conceived to formalize the notion of evidence of LET_J.

Some additional remarks

- The relation between Θ and A is not to be thought of as a relation of logical consequence, that is, A does not follow logically from Θ through some notion of consequence. In fact, it is not clear that a general account of the relation between Θ and A could even be spelled out.
- The proposed definition of evidence not only makes clear the connections between the notions of evidence and information but also fits with both interpretations. If Θ is empty, and so nothing is being presented as a supposed justification of A, what remains is just A, which is nothing but a piece of linguistic information in the sense of Dunn (2008) mentioned above.
- Science denialism illustrates what it is meant by conflicting evidence based on non-factive justifications. Consider, for example, the creationist claim that God created the Earth less than ten thousand years ago. The scientific consensus estimates the age of the Earth at about 4.5 billion years. On the Web one finds both claims, together with a number of justifications. The same applies to other examples of denialism.

Logics of evidence and truth – main papers

- Carnielli and Rodrigues. An epistemic approach to paraconsistency: a logic of evidence and truth. *Synthese*, 2017.

  The basic ideas, the sentential logic LET_J, an extension of *N4*.
- Rodrigues, Bueno-Soler and Carnielli. Measuring evidence: a probabilistic approach to an extension of Belnap-Dunn Logic. *Synthese*, 2020.

  Probabilistic semantics, interpretation in terms of information.
- Antunes, Rodrigues, Carnielli, Coniglio. Valuation semantics for first-order logics of evidence and truth, *Journal of Philosophical Logic*, 2022.

  First-order with identity, constant domains.
- Rodrigues and Antunes. First-order Logics of Evidence and Truth with Constant and Variable Domains. *Logica Universalis*, 2022.

  Kripke models with variable domains.
- Rodrigues and Carnielli. On Barrio, Lo Guercio, and Szmuc on logics of evidence and truth. *Logic and Logical Philosophy*, 2022.

  Refinement of the notions of evidence and information.
- Coniglio and Rodrigues. On six-valued logics of evidence and truth expanding Belnap-Dunn four-valued logic. *Submitted*.

  Finitely valued sentential *LET*s.

Thanks!

`https://shorturl.ae/hw6ff`

abilio.rodrigues@gmail.com