
Surviving Abduction

WALTER CARNIELLI, *Centre for Logic, Epistemology and the History of Science and Department of Philosophy. UNICAMP. Campinas (Brazil).*
E-mail: carniell@cle.unicamp.br

Abstract

Abduction or retrodution, as introduced by C.S. Peirce in the double sense of searching for explanatory instances and providing an explanation (i.e., involving the procedure of searching, and the function of providing an *explanans* to the *explanandum*) is a kind of complement for usual argumentation. There is, however, an inferential step from the *explanandum* to the (one or more) abductive *explanans* (that is, to the facts that will explain it). Whether this inferential step can be captured by logical machinery depends upon a number of assumptions, but in any case it suffers in principle from the triviality objection: any time a singular contradictory *explanans* occurs, the system collapses and stops working. The traditional remedies for such collapsing are the expensive (indeed, NP-complete) mechanisms of consistency maintenance, or complicated theories of non-monotonic derivation that keep the system running at a higher cost. I intend to show that the robust logics of formal inconsistency, a particular category of paraconsistent logics which permit the internalization of the concepts of consistency and inconsistency inside the object language, provide simple yet powerful techniques for automatic abduction. Moreover, the whole procedure is capable of automatization by means of the tableau proof-procedures available for such logics. Some motivating examples are discussed in detail.

Keywords: abduction, logics of formal (in)consistency, tableaux.

1 Abduction and Contradiction: How to Survive

The abduction problem was originally formulated by Charles S. Peirce (cf. [4]) as the process of forming hypotheses with explanatory purposes (thus a kind of reversed explanation). In a manuscript of 1901, Peirce defined abduction as “accepting or creating a minor premise as a hypothetical solution for a syllogism whose major premise is known and whose conclusion we discover to be a fact”. It was then in the context of his reading of Aristotle that the concept of abduction was conceived, but it is worth noting that the disjunctive act of “accepting or creating ...” something before using it as an explanation poses a second problem: how is it possible to create abductive *explanans*? First, we have to recognize that characterizing the concept of explanation is one of the greatest challenges in the philosophy of science; this problem, I believe, is even harder in Logic and Mathematics, where explanations are sometimes confused with proofs (this particular subtle point is discussed in [21]; a good reference with updated bibliography is [23]).

Although I am not suggesting that “explaining” can be reduced to “deducing”, it is certainly acceptable that the idea of explanation in deductive sciences *includes* the query for missing hypotheses; it is in this context that the general abductive process can be formulated as the process of *generating* new hypotheses within arbitrary deductive systems, and afterwards *using them in deductive terms*. The former task became known in the literature as *creative abduction*, in contrast to *explicative abduction*. The term “explicative” is here understood under the following proviso: a missing link in a deduction certainly does not exhaust the need for an explanation, but does constitute the first necessary step towards

explaining an anomalous (i.e., not yet deducible) fact. Two natural assumptions about explanations that can be posed are the following: first, there can be various explanations for the same anomalous fact, and second, there can be explanations of various degrees for the same anomalous fact. For example, looking for the ultimate scientific or philosophical explanation on why the grass of your garden is wet is one thing, while discovering that the sprinkler was left on all night long is another thing. Both explain the fact, but responding to different needs. The question is analogous to the one in automatic theorem proving: finding *any* proof is one thing, while finding a philosophically-minded proof is another. In the same manner as automatic theorem proving is satisfied with the first level of proofs, so automatic abduction will be satisfied with a first level explanation. I will treat here the question of automatic abduction when no “obvious” explanations are possible, that is, when contradictions are involved.

Let \vdash be a deductive relation; if $\Gamma \not\vdash \alpha$, the creative abductive step consists in finding an appropriate Δ so that $\Gamma \cup \Delta \vdash \alpha$. In this case, the discovered Δ does the explicative abductive step. Obviously, there must be some constraints, otherwise $\Delta = \{\perp\}$ (for \perp the bottom particle) would be a trivial solution for the abduction problem in most deduction relations. Usually another constraint is that $\Gamma \not\vdash \neg\alpha$, for this would (classically) imply that any explicative Δ would be a trivial explanation. This restriction, however, will not be necessary in our case, as the following developments and the examples will make clear.

From the point of view of general argumentation (and not only deduction), abduction concerns the search for hypotheses or the search for explanatory instances that help reasoning. In this sense, it can be seen as a complement for argumentation, in the same manner that in the philosophy of science the context of discovery is a complement for the context of justification. And moreover, keeping the analogy, the question of the logical possibility of creative abduction lies on the same side of the famous question of the logical possibility of scientific discovery.

A renewed interest on abduction acquired impetus due to the factual treatment of data and the question of virtual causality in the information society. The enormous amount of data stored in the world wide web and in complex systems, as well as the virtual relationship among such data, continuously demands new tools for automatic reasoning. These tools should incorporate general logical methods which are at the same time machine-understandable, and sufficiently close to human semantics as to perform sensible automated reasoning.

An example where abductive inference is highly relevant is the *model-based diagnosis* in engineering and AI. Suppose that a complex system, as an aircraft, is being tested before a transatlantic flight. The electronic circuitry permits the testers to predict certain outputs based on specific input tests; if the instruments show something distinct from the expected, it is a task of model-based diagnosis to discover an explanation for the anomaly and use it to separate the components responsible for the problem, instead of disassembling the whole aircraft.

Another example occurs in the process of updating in the so-called *datalog databases*. Suppose we have a logic programming composed of the following clauses, where $desc(x, y)$ means “ x is a descendent of y ”, $parent(y, x)$ means “ y is a parent of x ” and $\beta \leftarrow \alpha$ means that the database contents plus α produce (or answer) β :

$$desc(x, y) \leftarrow parent(y, x)$$

$$desc(x, y) \leftarrow parent(z, x), desc(z, y)$$

There is a subtle difference between inserting information in the database in an explicit

versus an implicit manner: information of the form “ y is a parent of x ” is a basic fact, and can be inserted explicitly, while information of the form “ x is a descendent of y ” is either factual knowledge or is a consequence of the machine reasoning (as simple as it can be). If one wishes to insert an implicit information, it is necessary to modify the set of facts stored in the database in such a way that this information can be deduced: this is an example of creative abduction and explicative abduction at the same time. For instance, if we have stored $desc(Zeus, Uranus)$ and $parent(Uranus, Cronus)$, for implicitly inserting $desc(Aphrodite, Uranus)$ there are two different ways: we may either insert $parent(Zeus, Aphrodite)$ or alternatively insert $desc(Aphrodite, Cronus)$. These two alternative additions are examples of abductive explanations for $desc(Zeus, Uranus)$. In fact, logic programming uses this abductive mechanism for answering queries, in the form: “Is the fact $desc(Aphrodite, Uranus)$ compatible with the program clauses and data?” or “Is there any x such that $desc(x, Uranus)$?”; the whole procedure is creative as much as it can be automatized. Therefore it is evident that a useful abductive mechanism for databases should be based on first-order logic, and not merely on propositional logic.

Moreover, abductive approaches are also used to integrate different ontologies and database schemes, or for integrating distinct data sources under the same ontology, as for example [13], where an abductive-based application for database integration is developed. Suppose that, while a query is being processed by a user, another data source had inserted $desc(Uranus, Aphrodite)$ in our database, plus a constraint of the form: “For no x and y , simultaneously $desc(x, y)$ and $desc(y, x)$ can be maintained in the database”.

If $parent(Zeus, Aphrodite)$ had been inserted for one data source, the insertion of $desc(Uranus, Aphrodite)$ by the second data source would cause a collapse in view of the constraint. What can be done? The alternative of deleting all data seems inconceivable, and the one of having all queries be answered positively (since a database established on classical logic grounds would deduce anything from a contradiction) is of course intolerable. Thus a legitimate logic to ground this process would have to be a first-order logic that sanctions useful reasoning in the presence of contradictions. Proposals from this perspective have been investigated in [7] and [5].

In this paper I argue that simple yet powerful techniques for automatic abduction can be usefully implemented by means of tableau proof-procedures for certain *logics of formal inconsistency* (**LFI**s). I concentrate the attention here on the propositional basis only, commenting on the extensions to the first-order case and briefly discussing the technical difficulties involved. Although a wide-scoped study of tableau and abduction was offered by [12] in 1997, the quite natural idea of using the backward mechanism of tableaux for gaining automatic explanations occurred earlier: [20] in 1992 already proposed a fully detailed treatment for the question of “completing” a database in a way as to deduce (in classical propositional logic) a previously undeducible formula. In the same year a tableau proof system for the paraconsistent system C_1 was proposed in [17], and several examples of using such tableau systems for automatic solving dilemmatic situations was extensively discussed. Even though neither of these references mentions explicitly the concept of abduction, these papers undoubtedly proposed ways to solve the abductive problem, respectively for classical propositional logic and for the propositional paraconsistent logic C_1 . In [16] tableau systems for **LFI**s were proposed, and this logical formalism was used as a method to build database repairs in [7]. Such methods are based upon many-valued semantics, or upon non-truth-functional bivalued semantics ¹.

¹In many cases, however, the same logic can be at the same time characterized by truth-functional many-valued semantics

The question of abduction thus involves two independent, but complementary problems: 1) finding a method to automatically perform abduction (and, if possible, to automatically generate abductive data), and 2) doing this within a robust reasoning environment, in a such a way as to keep running and providing reasonable output even in the presence of the possible contradictions that this search would engender. Any contradictions found in the process of producing a lucid output are traditionally seen as a condemnation of the whole process: contradictions (regarded traditionally, at least) entail triviality, and triviality is but the opposite of lucidness; the abductive experience thus can sometimes appear to be lethal. I want to argue that this is first of all a mistaken view, and second, that very simple and natural logical models can be designed for surviving abduction, by means of defining them in terms of refutation procedures based on **LFI**s.

2 Abduction and Induction *Versus* Deduction

Philosophers after Renaissance already maintained that deduction plays a too modest role in the development of knowledge: Francis Bacon (1561-1626) is known to have criticized Aristotelian logic as useless for the discovery of laws. Bacon called for a new method for developing philosophy; he wrote that philosophers should, instead of staying attached to deductive syllogism to study nature, proceed through (today called inductive) reasoning from facts to laws. In a certain sense, his theory could be seen as a precursor to abduction, if we take into account that there is a difference between multiple-events induction (in the sense of producing a law, starting from a series of events, which by its turn will provide predictions of future events) and single-event abduction, which goes from an event to its explanation (which may be another event, or a law).

This distinction seems to be coherent to the use Peirce made of the concepts of deduction, induction and abduction; although he has refined and revised his account on such concepts through times (even changing terminology), Peirce neatly put deduction on the one side, and induction and abduction on the other (see for example [11]). There are many uses and abuses of the term “abduction” in all areas of knowledge, to a point that for example [12] proposes a careful ‘taxonomy for abduction’ . I will consider here the process of abduction as resulting in a sentence, whether particular or existentially quantified (representing a single event) or universally quantified (representing a law). Even if in this respect we may be departing from Peirce, we will be in line with recent usage, as in e.g. [12]): according to this conception, induction does not differ from abduction on what it produces (an explanatory event or a law), but from where it departs (a single event or a multiple event).

It is well accepted that abduction does not go in the forward direction of deduction. It is not difficult to accept, either, that abduction cannot coincide with any backward form of classical deduction, but it does not follow that *another* form of backward deduction would not work. In this sense, the **LFI**s are good candidates: they are infraclassical logics (in the sense that they do not prove anything that classical logic would not), are resistive to contradictions, but nonetheless can encode the whole of classical reasoning. Backwards proof procedures for **LFI**s indeed constitute a suitable approach for abduction, and I intend to show how this approach can be programmed and treated in a natural basis departing from a very simple formalism.

It is typical that cognitive and epistemological situations can enter into a situation of

and by non-truth-functional bivalued semantics; a thorough investigation on this interesting interplay between bivalence and truth-functionality is discussed in [19], where two-signed tableaux are also provided.

(presumably temporary) contradictory state as the well-known case of scientific theories facing anomalies, as recognized by L. Magnani in [3]. Of course, in a situation where we have serious theories competing around a contradiction, there is little sense in rejecting one of them *a priori* just to save the Principle of Explosion or the Principle of Pseudo-Scotus (which states that logical deduction is *contradiction-explosive* in the sense that from a contradiction everything is derivable); it seems to be out of question that it is more convenient to tame the logic, than to sacrifice a precious (and possibly correct) theory.

This is not only the case for sophisticated cosmological theories: a single digit in a database can of course be extremely valuable, and it is already widely recognized that no automated reasoning is possible without a way to control logical explosion. What is not clear is whether the act of guessing involved in the discovery context of abduction, and furthermore under such conditions, can be the subject of logic. Although it seems that Peirce maintained the negative, the present paper intends to show that in many interesting cases the process of guessing can be solved semi-automatically by means of a careful manipulation of the concept of consistency, viewed as a primitive notion independent from the concept of contradiction, as granted by **LFIs**. This way we can obtain a reasonably efficient and conceptually simple method for discovering new *logical* hypotheses that will serve as *explanans* for a given *explanandum*).

3 Abducting with LFIs: A Safeguard

Defenses of the relevance of reasoning under contradiction for the benefit of epistemology have already been advanced: for example, Magnani argues (cf. [3]) that:

As noted above, when we create or produce a new concept or belief that competes with another one, we are compelled to maintain the derived inconsistency until the possibility of rejecting one of the two becomes feasible. We cannot simply eliminate a hypothesis and then substitute it with one inconsistent with it, because until the new hypothesis comes in competition with the old one, there is no reason to eliminate the old one. Other cognitive and epistemological situations present a sort of paraconsistent behavior: a typical kind of inconsistency maintenance is the well-known case of scientific theories that face anomalies. As noted above, explanations are usually not complete but only furnish partial accounts of the pertinent evidence: not everything has to be explained.

He continues by exemplifying how Newtonian mechanics not only had to survive facing the anomaly of the perihelion of Mercury until the theory of relativity came in, but it also had to accept its own false prediction about the motion of Uranus.

This is an evidence, according to the same paper, of similarities between reasoning in the presence of inconsistencies and reasoning with incomplete information. In order to overcome those contradictions, scientists sometimes adopt the practice of generating auxiliary hypotheses. Such auxiliary hypotheses are not *ad hoc* logical tricks performed to avoid contradictions, but should, if possible, lead to the prediction of something new: the famous case of the auxiliary hypothesis of the existence of an additional planet, later called Neptune, is an example of a very successful case.

The logics of formal inconsistency (**LFIs**) are a wide class of paraconsistent logics which permit the internalization of the concepts of consistency or inconsistency inside the object language; the reader interested in a full treatment of **LFIs** including history, conceptual-

ization, axiomatics and semantics should consult [6]. For the purpose of the arguments developed here, it is sufficient to stress that such a treatment of consistency as a primitive notion, and the resulting dissolution of the purported connections between consistency and contradiction, is accomplished via the introduction of new operators to express the concepts of consistency (and inconsistency) as a primitive notion, thus separating, in principle, the notions of contradictoriness and of inconsistency.

The **LFIs** represent the majority of all extant paraconsistent logics, and express inconsistency as a purely linguistic notion, independent from any ontological presuppositions. A relevant feature of **LFIs** is that the Principle of Explosion, which is an important tool in mathematical reasoning (necessary for example in all forms of proof by absurd) is not lost, but finely controlled. As a consequence of this careful control of explosiveness, most **LFIs** based on classical logic faithfully reproduce all classical inferences by means of suitable translations (despite being themselves only fragments of classical logic). From the semantical viewpoint the **LFIs** are sound and complete with respect to interpretations in valuation semantics and with respect to the new possible-translations semantics (see [1]). This semantics and the proof machinery of these logics testify the robustness of paraconsistent logics (and **LFIs** in particular): (i) they do not validate any non-classical reasoning ; (ii) they do not prove any contradictory theorem, and (iii) they do retain full power of classical reasoning, including forms of *reduction ad absurdum*.

As a by-product, **LFIs** permit to clarify certain misconceptions perpetrated by some authors, concerning the confusion between the validity of the *Principle of Pseudo-Scotus* (or *Principle of Ex Contradictione Sequitur Quodlibet*), which is responsible for the explosive character of the consequence relation \vdash_L in the presence of contradictions in a logic L , and the *Principle of Non-Contradiction*. The first principle states that, for every Γ and α :

$\Gamma, \alpha, \neg\alpha \vdash_L \beta$, and

while the *Principle* (or *Law*) of *Non-Contradiction* states that L must have at least a non-contradictory theory Γ , that is, there must exist Γ such that for every α :

$\Gamma \not\vdash_L \alpha$ or $\Gamma \not\vdash_L \neg\alpha$.

Obviously, contradictory theories exist even in realm of classical logic: just consider the “improper” theory formed by all formulas, or even “proper” theories made up by pairs of contradictory formulas.

An instructive example of a deduction featuring the Principle of *Ex Contradictione Sequitur Quodlibet* was offered by Augustinus Niphus, a 16th century Italian logician. He showed that anything follows from an impossible proposition (and thus that classical deduction is explosive), by proving that “Socrates is and Socrates is not” entails “Man is a horse”; he then shows that anything follows from a false proposition by proving that “Man is a horse; therefore you are at Rome” (cf. [9]). His arguments follow the same pattern in both cases (where here L is classical logic):

1. $\alpha \wedge \neg\alpha \vdash_L \alpha$
2. $\alpha \vdash_L \alpha \vee \beta$
3. $\alpha \wedge \neg\alpha \vdash_L \alpha \vee \beta$
4. $\alpha \wedge \neg\alpha \vdash_L \neg\alpha$
5. $(\alpha \vee \beta) \wedge \neg\alpha \vdash_L \beta$
therefore
6. $\alpha \wedge \neg\alpha \vdash_L \beta$.

As pointed by [10], a similar argument is found not only in John Duns Scotus (Pseudo-Scotus), but in Ockham, Buridan, and Albert of Saxony, among others. Although credited to John Duns Scotus (14th century), this principle is considered to be probably due to Peter Abelard (early 12th century), reputed to be the first medieval logician to develop an early version of propositional logic.

Now, as several authors have pointed, it is not possible to retrieve this kind of argument within the Aristotelian syllogistic: the closest Aristotelian law, as discussed by the famous review of Aristotle by J. Lukasiewicz, is (at least in one of several formulations) the following:

“The same attribute cannot at the same time belong and not belong to the same subject and in the same respect”.

To see that this coincides with the above *Principle of Non-Contradiction* and to understand the roots of the confusion with the *Ex Contradictione Sequitur Quodlibet* just note that this aged law simply says that one should not deduce contradictory statements, but not even Aristotle can forbid anyone from considering contradictory hypotheses! The point is what logic should do with hypotheses of this kind— it should be clear thus that the *Principle of Non-Contradiction* is normative, while the Principle of *Ex Contradictione Sequitur Quodlibet* is operative.

Paraconsistent logics must necessarily refute the latter, while possibly respecting the first. Indeed, with the exception of a few paraconsistent dialectical logics that intend to formalize some dialectical principles, and which permit to deduce contradictions, the vast majority of paraconsistent logics, and by that matter all of **LFI**s, do not *derive* contradictions, but just enable us to reason about them.

A closer attention to the medieval argument shows that the crucial step in the argument of explosive triviality is:

$$(\alpha \vee \beta) \wedge \neg\alpha \vdash_L \beta;$$

This, however, has nothing to do with the *Principle of Non-Contradiction*, but with the fact that, if we accept that the hypothesis $\alpha \vee \beta$ and $\neg\alpha$ as simultaneously true, β has to be true, since we presuppose that the hypothetical α and its negation cannot be simultaneously true. But nothing forbids us (not even the *Principle of Non-Contradiction*) to use as hypothesis sentences α and $\neg\alpha$ that could be taken in some sense as being both true (or at least both non-false) *a priori*. This may be the case if α and $\neg\alpha$ are being evaluated by two observers that disagree due to lack of information, or on account of α and $\neg\alpha$ being defectively defined, or because we have been so lucky as to discover a real contradiction in nature...

Reasoning in such terms is but habitual: that is what scientists (and lawyers, and judges) do in many occasions. If we depart from the uncertain or inconclusive pair α and $\neg\alpha$ the inference to β has to be blocked. Even so, the hypothesis could again entail β , had we any further guarantee that one of them, say α , after all is certain and conclusive; in that case, we would say that α is *consistent* (analogously for $\neg\alpha$). This, as a notion of extra-logical nature, requires a representation in the object language, and there is where the **LFI**s come in.

One of the primarily interesting subclasses of **LFI**s are the **C**-systems, characterized as those **LFI**s where consistency can be expressed as a unary connective. This permits us to formally distinguish between consistency and non-contradiction, and to clearly recognize which are the assumptions hidden within formal logic that force consistency and non-contradiction, or also inconsistency and contradiction, to succumb into indistinguishable concepts. Not only most of the logics produced by the Brazilian school fall within the definition of **LFI**s, but the well-known Jaśkowski’s discussive logic D2 and even normal modal logics in general can

be recast as **dC**-systems, a further subclass of **C**-systems in which the unary connective for consistency may be explicitly defined in terms of other connectives in the language.

4 A Simple and Powerful Logic to Reason Under the Pressure of Contradictions

One of the most natural paraconsistent logics in the ample class of **LFI**'s, if one considers that attributes for a logic to be natural are simplicity and easiness in defining it, is the three-valued logic **LFI1**: it is simple because it has a functional three-valued semantics, but it also has a non-truth-functional bivalued semantics (a *dyadic semantics*, as studied in [19]); it is also natural, in the sense that it was rediscovered (and redefined) several times, by different authors departing from totally distinct perspectives. As we shall see, **LFI1** is indeed very expressive, and has also been extended to a first-order version, cf. [5].

In purely propositional terms, **LFI1** was introduced in [5] as a semantical solution for a problem concerning reasoning with contradictory databases. It coincides, nevertheless, with other logical solutions for totally distinct logic problems since 1960 (see [18] for a discussion). Not only **LFI1**, in different formulations and interpreted differently, appeared as a spontaneous solution in distinct historical moments, but it also plays a central role in interpreting other paraconsistent logics: essentially, another alternative formulation of **LFI1** can be used to interpret and explain other more complex paraconsistent logics. Indeed, despite the fact that the famed da Costa's paraconsistent systems C_n cannot be characterized by finite-valued semantics, it is possible to give a semantical interpretation for C_n as a suitable combination of three-valued logics equivalent to **LFI1** by means of the possible-translations semantics. This was shown in [1] (cf. also [22]).

LFI1 is the logic defined semantically by the following matrices, in the connectives \wedge (for conjunction), \vee (for disjunction), \rightarrow (for implication), \neg (for weak negation) and \bullet (for inconsistency), where 1 and $1/2$ are distinguished truth-values :

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

\vee	1	$1/2$	0
1	1	1	1
$1/2$	1	$1/2$	$1/2$
0	1	$1/2$	0

\rightarrow	1	$1/2$	0
1	1	$1/2$	0
$1/2$	1	$1/2$	0
0	1	1	1

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

As it is usual in many-valued logics, we say, for $\Gamma \cup \{\alpha\}$ a collection of formulas in **LFI1**, that α is a *3-valued consequence* of Γ , denoted by $\Gamma \models_3 \alpha$, if for any three-valued valuation v defined by the above tables, if $v(\Gamma) \in \{1, 1/2\}$ then $v(\alpha) \in \{1, 1/2\}$.

A kind of ‘possibility connective ∇ can be defined in **LFI1** as: $\nabla\alpha \stackrel{\text{def}}{=} \neg\bullet\alpha \rightarrow \alpha$ with the table:

	∇
1	1
$1/2$	1
0	0

LFI1 has some close connections with Lukasiewicz's three-valued logic L_3 : It is possible to define within **LFI1** the table \rightarrow^L for the implication of **LFI1**:

$$\alpha \rightarrow^L \beta \stackrel{\text{def}}{=} ((\nabla(\neg\alpha)) \vee \beta) \wedge ((\nabla\beta) \vee (\neg\alpha)).$$

giving the matrix:

\rightarrow^L	1	$1/2$	0
1	1	$1/2$	0
$1/2$	1	1	$1/2$
0	1	1	1

As the negation \neg^L of L_3 is the same as \neg , we can define the basic connectives of **LF11** from \rightarrow^L and \neg^L of Lukasiewicz's L_3 :

$$\alpha \vee \beta \stackrel{\text{def}}{=} (\alpha \rightarrow^L \beta) \rightarrow^L \beta;$$

$$\nabla \alpha \stackrel{\text{def}}{=} (\neg^L \alpha) \rightarrow^L \alpha.$$

In this way, **LF11** and L_3 define the same matrices in the same language, though they are not equivalent, since **LF11** has only 1 as its distinguished truth-value; this fact has as consequences, for example, that $\vdash_{\mathbf{LF11}} (\neg p \rightarrow p) \rightarrow p$ while $\not\vdash_{L_3} (\neg p \rightarrow p) \rightarrow p$.

Nonetheless, **LF11** and Lukasiewicz's three-valued logic L_3 are algebraizable by the same equivalent algebraic semantics, namely, by means of three-valued MV-algebras; see [14]. for a careful discussion of this unexpected relationship between such logics.

In [18] and [2] it is also shown how **LF11** or its sister formulations extend several other elementary paraconsistent logics in the literature; still another important reason why **LF11** is so natural is that it contains the most part of the classical schemas and rules that do not interfere with paraconsistency. In other words, **LF11** is a maximal subsystem of classical logic, in the sense that there is no other extension of **LF11** which would not explode in the presence of contradictions.

For the sake of completeness, I also show here the hilbertian axiomatization for **LF11**, detailing some relevant subsystems and starting from a kind of basic logic (cf. [18]):

The logic **mbC** is defined, over the signature containing connectives $\wedge, \vee, \rightarrow, \neg$ and \circ by the axioms below. While connectives \wedge, \vee and \rightarrow keep their usual meaning, \neg has a meaning that is weaker than the classical as contradictory sentences do not cause deductive explosion; the connective \circ means "certain", "conclusive", or, as we prefer to call, "consistent" (cf. [18]). The most important point to keep in mind with respect to the logics we are dealing here is that if α and $\neg\alpha$ are contradictory sentences and α is consistent (i.e., $\circ\alpha$ holds) then a contradiction deductively explodes.

Axiom schemas:

$$\mathbf{(Ax1)} \quad \alpha \rightarrow (\beta \rightarrow \alpha)$$

$$\mathbf{(Ax2)} \quad (\alpha \rightarrow \beta) \rightarrow ((\alpha \rightarrow (\beta \rightarrow \gamma)) \rightarrow (\alpha \rightarrow \gamma))$$

$$\mathbf{(Ax3)} \quad \alpha \rightarrow (\beta \rightarrow (\alpha \wedge \beta))$$

$$\mathbf{(Ax4)} \quad (\alpha \wedge \beta) \rightarrow \alpha$$

$$\mathbf{(Ax5)} \quad (\alpha \wedge \beta) \rightarrow \beta$$

$$\mathbf{(Ax6)} \quad \alpha \rightarrow (\alpha \vee \beta)$$

$$\mathbf{(Ax7)} \quad \beta \rightarrow (\alpha \vee \beta)$$

$$\mathbf{(Ax8)} \quad (\alpha \rightarrow \gamma) \rightarrow ((\beta \rightarrow \gamma) \rightarrow ((\alpha \vee \beta) \rightarrow \gamma))$$

$$\mathbf{(Ax9)} \quad \alpha \vee (\alpha \rightarrow \beta)$$

$$\mathbf{(Ax10)} \quad \alpha \vee \neg\alpha$$

$$\mathbf{(bc1)} \quad \circ\alpha \rightarrow (\alpha \rightarrow (\neg\alpha \rightarrow \beta))$$

Inference rule:

$$(\text{MP}) \frac{\alpha, \alpha \rightarrow \beta}{\beta}$$

The logic **mCi** is obtained from **mbC** by the addition of the following axiom schemas:

$$\begin{cases} (\text{ci}) & \neg \circ \alpha \rightarrow (\alpha \wedge \neg \alpha) \\ (\text{cc})_n & \circ \neg^n \circ \alpha \end{cases} (n \geq 0)$$

The logic **Cie** is obtained from **mCi** plus the following axiom schemas:

$$\begin{cases} (\text{cf}) & \neg \neg \alpha \rightarrow \alpha \\ (\text{ce}) & \alpha \rightarrow \neg \neg \alpha \end{cases}$$

Finally, the logic **LF11** is axiomatized by adding to **Cie** the following axiom schemas:

$$\begin{cases} (\text{c}j1) & \bullet(\alpha \wedge \beta) \leftrightarrow ((\bullet\alpha \wedge \beta) \vee (\bullet\beta \wedge \alpha)); \\ (\text{c}j2) & \bullet(\alpha \vee \beta) \leftrightarrow ((\bullet\alpha \wedge \neg\beta) \vee (\bullet\beta \wedge \neg\alpha)); \\ (\text{c}j3) & \bullet(\alpha \rightarrow \beta) \leftrightarrow (\alpha \wedge \bullet\beta). \end{cases}$$

where the inconsistency connective \bullet , meaning “uncertain”, “inconclusive”, etc., is defined as $\bullet\alpha =_{def} \neg \circ \alpha$. Because of axioms (cf) and (ce), we can replace unrestrictedly $\neg \bullet \alpha$ by $\circ \alpha$, though it is simpler to state semantic conditions and tableau rules in terms of \bullet .

For $\Gamma \cup \{\alpha\}$ a collection of formulas in **LF11**, $\Gamma \vdash_{LF11} \alpha$ denotes the fact that α is obtained from Γ by means of the above axioms.

Making explicit the subsystems **mbC**, **mCi** and **Cie** of **LF11** has a purpose: instead of employing **LF11**, one could employ any of those subsystems to deal with abductive reasoning, as in the discussion below. **LF11** has been chosen because of its extra features: it is maximal, and it is characterizable by a three-valued semantics, besides a two-valued one, while the above mentioned subsystems are not characterizable by finite-valued semantics (see [18]).

Following such motivations, here is a short description of other syntax and semantic features of **LF11**, concentrating in the propositional level (the first-order case is discussed in Section 6). Besides the three-valued semantics described above, **LF11** can be also characterized by a two-valued semantics in the special *dyadic format*: these are two-valued (non necessarily truth-functional) semantics described by “and-or” clauses that generalize the subformula property (cf. detailed treatment in [19]).

A dyadic valuation for **LF11** is a function b from the formulas of **LF11** into $\{T, F\}$ defined as follows:

- 1) $b(\neg\alpha) = T \Rightarrow b(\alpha) = F$ or $b(\bullet\alpha) = T$;
- 2) $b(\neg\alpha) = F \Rightarrow b(\alpha) = T$ and $b(\bullet\alpha) = F$;
- 3) $b(\bullet\alpha) = T \Rightarrow b(\alpha) = T$;
- 4) $b(\bullet\bullet\alpha) = T \Rightarrow b(\bullet\alpha) = F$;
- 5) $b(\bullet\bullet\neg\alpha) = T \Rightarrow b(\bullet\alpha) = T$;
- 6) $b(\bullet\bullet\neg\alpha) = F \Rightarrow b(\neg\alpha) = F$ or $b(\alpha) = F$;
- 7) $b(\alpha \wedge \beta) = T \Rightarrow b(\alpha) = T$ and $b(\beta) = T$;
- 8) $b(\alpha \wedge \beta) = F \Rightarrow b(\alpha) = F$ or $b(\beta) = F$;
- 9) $b(\alpha \vee \beta) = T \Rightarrow b(\alpha) = T$ or $b(\beta) = T$;
- 10) $b(\alpha \vee \beta) = F \Rightarrow b(\alpha) = F$ and $b(\beta) = F$;
- 11) $b(\alpha \Rightarrow \beta) = T \Rightarrow b(\alpha) = F$ or $b(\beta) = T$;
- 12) $b(\alpha \Rightarrow \beta) = F \Rightarrow b(\alpha) = T$ and $b(\beta) = F$;
- 13) $b(\bullet(\alpha \wedge \beta)) = T \Rightarrow$
 $b(\alpha) = T$ and $b(\bullet\beta) = T$ or $b(\beta) = T$ and $b(\bullet\alpha) = T$;

- 14) $b(\bullet(\alpha \wedge \beta)) = F \Rightarrow$
 $b(\alpha) = F$ or $b(\beta) = F$ or $b(\alpha) = T$ and $b(\bullet\alpha) = F$, $b(\beta) = T$, $b(\bullet\beta) = F$;
- 15) $b(\bullet(\alpha \vee \beta)) = T \Rightarrow$
 $b(\alpha) = F$ and $b(\bullet\beta) = T$ or $b(\beta) = F$ and $b(\bullet\alpha) = T$ or $b(\bullet\alpha) = T$ and
 $b(\bullet\beta) = T$;
- 16) $b(\bullet(\alpha \vee \beta)) = F \Rightarrow$
 $b(\alpha) = F$ and $b(\beta) = F$ or $b(\alpha) = T$ and $b(\bullet\alpha) = F$ or $b(\beta) = T$, $b(\bullet\beta) = F$;
- 17) $b(\bullet(\alpha \Rightarrow \beta)) = T \Rightarrow b(\alpha) = T$ and $b(\bullet\beta) = T$;
- 18) $b(\bullet(\alpha \rightarrow \beta)) = F \Rightarrow b(\alpha) = F$ or $b(\bullet\beta) = F$.²

We say, for $\Gamma \cup \{\alpha\}$ a collection of formulas in **LF11**, that α is a *dyadic consequence* of Γ , denoted by $\Gamma \models_2 \alpha$, if for any dyadic valuation b defined by the above clauses, if $b(\Gamma) = T$ then $v(\alpha) = T$.

More detailed presentations of **LF11** (but using slightly different, non-gentzenian bivaluations) can be found in the mentioned papers [5] and [7].

For the sake of abductive procedures, it is convenient to formulate the proof machinery of **LF11** in terms of tableaux. Definitions of tableau proof systems abound in the literature; I just ask the reader to recall some relevant points in the usual definitions for tableaux that will be also used here: (1) A *signed* formula of **LF11** is an expression of the form $T(\alpha)$ or $F(\alpha)$, where α is a formula of **LF11**; (2) Tableaux are finitely generated trees, and a branch is said to be *closed* (indicated by $*$) if it contains contradictory formulas (in our case, signed formulas of the form $T(\alpha)$ and $F(\alpha)$). Otherwise the branch is said to be *open* (indicated by \ominus); (3) Tableau rules are branching (indicated by “|”, as in rule (R.1) below) or consecutive (indicated by “,”, as in rule (R.4) below).

An important advantage of dyadic valuations is that tableau systems for logics characterized by these valuations can be almost immediately defined, as shown below.

Based on the dyadic valuations for **LF11**, consider tableau rules as follows: Translate $b(\beta) = T$ as a signed formula $T(\beta)$, and $b(\beta) = F$ as a signed formula $F(\beta)$. As a consequence, the conditional clauses of dyadic valuations induce tableau rules in such a way that tableau completeness is automatically obtained (see [19]).

A tableau system for **LF11** is given below, based on the corresponding dyadic semantics:

$$\begin{array}{lll}
 (R.1) \frac{T(\neg\alpha)}{F(\alpha) \mid T(\bullet\alpha)} & (R.2) \frac{F(\neg\alpha)}{T(\alpha), F(\bullet\alpha)} & (R.3) \frac{T(\bullet\alpha)}{T(\alpha)} \\
 (R.4) \frac{T(\bullet\bullet\alpha)}{F(\bullet\alpha)} & (R.5) \frac{T(\bullet\neg\alpha)}{T(\bullet\alpha)} & (R.6) \frac{F(\bullet\neg\alpha)}{F(\neg\alpha) \mid F(\alpha)} \\
 (R.7) \frac{T(\alpha \wedge \beta)}{T(\alpha), T(\beta)} & (R.8) \frac{F(\alpha \wedge \beta)}{F(\alpha) \mid F(\beta)} & \\
 (R.9) \frac{T(\alpha \vee \beta)}{T(\alpha) \mid T(\beta)} & (R.10) \frac{F(\alpha \vee \beta)}{F(\alpha), F(\beta)} & \\
 (R.11) \frac{T(\alpha \Rightarrow \beta)}{F(\alpha) \mid T(\beta)} & (R.12) \frac{F(\alpha \Rightarrow \beta)}{T(\alpha), F(\beta)} &
 \end{array}$$

²To be completely precise, a dyadic semantics for **LF11** requires two additional clauses, referred to as (C1)–(C2) in [19]; in the present case, however, this is assured by the fact that we are assuming dyadic valuations to be *total functions* with images in $\{T, F\}$.

$$(R.13) \frac{T(\bullet(\alpha \wedge \beta))}{T(\alpha), \mid T(\beta), \mid T(\bullet\beta) \mid T(\bullet\alpha)}$$

$$(R.14) \frac{F(\bullet(\alpha \wedge \beta))}{F(\alpha) \mid F(\beta) \mid T(\alpha), T(\beta), \mid F(\bullet\alpha), F(\bullet\beta)}$$

$$(R.15) \frac{T(\bullet(\alpha \vee \beta))}{F(\alpha), \mid F(\beta), \mid T(\bullet\alpha), \mid T(\bullet\beta) \mid T(\bullet\alpha) \mid T(\bullet\beta)}$$

$$(R.16) \frac{F(\bullet(\alpha \vee \beta))}{F(\alpha), \mid T(\alpha), \mid T(\beta), \mid F(\beta) \mid F(\bullet\alpha) \mid F(\bullet\beta)}$$

$$(R.17) \frac{T(\bullet(\alpha \Rightarrow \beta))}{T(\alpha), T(\bullet\beta)}$$

$$(R.18) \frac{F(\bullet(\alpha \Rightarrow \beta))}{F(\alpha) \mid F(\bullet\beta)}$$

Another tableau system for **LF11** was presented in [16]; that one is based on a non-gentzenian semantics, and the tableau rules, though decidable, may have loops. The equivalence between those distinct tableau systems for the same logic can be checked by means of the notion of derived tableau rules introduced in [17].

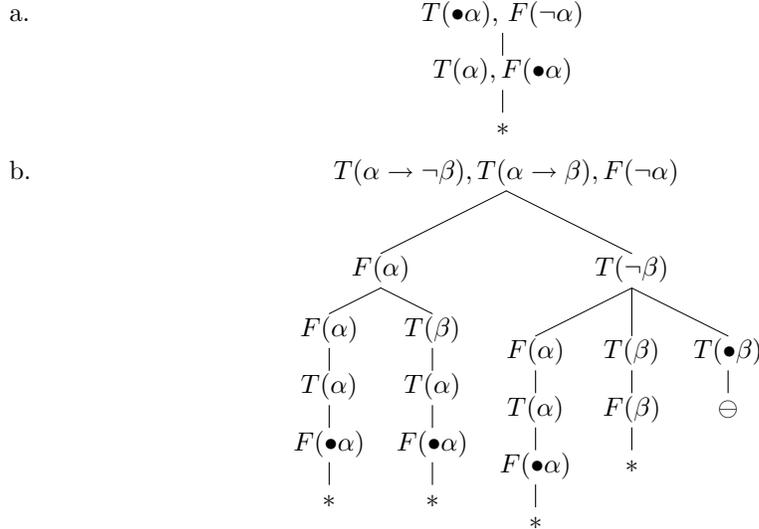
For $\Gamma \cup \{\alpha\}$ a collection of formulas in **LF11**, $\Gamma \vdash_{LF11} \alpha$ denotes the fact that α is obtained from Γ via the above tableau rules (i.e., there exists a closed tableau for $\Gamma_{\top} \cup F(\alpha)$).

EXAMPLE 4.1

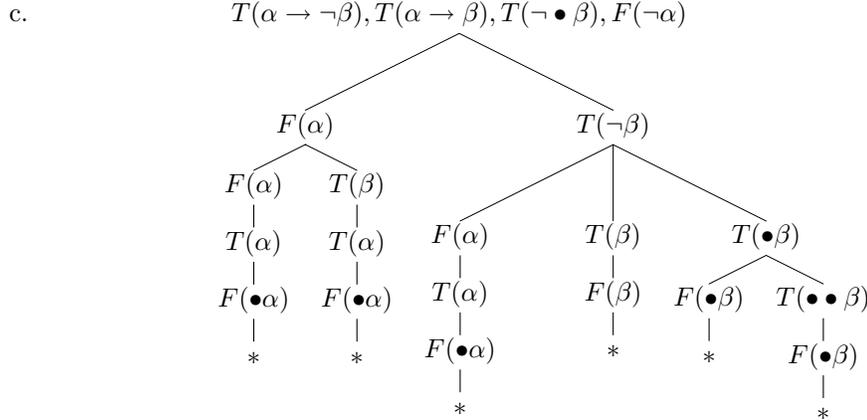
(a) $\bullet\alpha \vdash_{LF11} \neg\alpha$

(b) $\alpha \rightarrow \beta, \alpha \rightarrow \neg\beta \not\vdash_{LF11} \neg\alpha$

(c) $\alpha \rightarrow \beta, \alpha \rightarrow \neg\beta, \neg\bullet\beta \vdash_{LF11} \neg\alpha$. Indeed:



and the resulting tableau is open: the condition that β is certain or consistent is missing here.



and the resulting tableau is now closed, using (R.1) and (R.4) to close the previously open node. This is so because we have added the condition that β is certain or consistent, represented by the formula $\neg \bullet \beta$ (equivalent to $\circ\beta$).

The following result guarantees the soundness and completeness of \mathcal{T} with respect to the logic **LF11**:

THEOREM 4.2

Let Γ be a set of **LF11** formulas and α be a **LF11** formula. Then, $\Gamma \vdash \alpha$ iff $\Gamma \models_3 \alpha$ iff $\Gamma \models_2 \alpha$ iff $\Gamma \vdash_{LF11} \alpha$.

PROOF. By combining all the results about **LF11** in [18] and in [19]. ■

5 LF11 Tableaux as Abductive Procedures

As argued in the preceding section, an important **LF1** is the logic **LF11**. A relevant feature of **LF11** (and of several other **LF1**s), as we have seen, is that they can be defined by means of flexible refutative (tableau type) proof procedures. Such backward proof procedures are very convenient for formalizing abductive routines. This has been recognized in the literature for the case of classical logic, as explained in Section 1. In our case, however, the intention is convince the reader that tableaux for **LF11** are good candidates (in any case, much better than tableaux based on classical logic) to compute abductions.

Part of my purpose is to provide some criticisms, identifying certain problems that could be solved by using more adequate underlying logics in the context of abduction and, as far as possible, to suggest directions for further work. I am especially concerned with abducting in the presence of incomplete as well as overcomplete information. For this reason, through convenient modification of a general trend in the literature (see particularly [12]), the notion of an abductive explanation is so defined:

DEFINITION 5.1

Let Γ, Δ be finite sets of sentences and α be a sentence in the language of a given logic **L**. Γ and α form an *abductive problem* and Δ is an *abductive explanation* for the abductive problem if:

1. (Abductive problem): The context Γ is not sufficient to entail α , that is, $\Gamma \not\vdash_L \alpha$;
2. (Abductive solution): The enriched context Γ plus Δ is sufficient to entail α , that is, $\Gamma, \Delta \vdash_L \alpha$;

3. (Non-triviality of solution): The enriched context Γ plus Δ is non-trivial, that is, there exists γ such that $\Gamma, \Delta \not\vdash_L \gamma$;
4. (Vocabulary restriction of solution): $Var(\Delta) \subset Var(\Gamma) \cup Var(\alpha)$.
5. (Minimality of solution): Δ is minimal (in the sense, for example that it is composed of a set with minimal cardinality and with formulas of minimal length, but there are several ways to implement such minimality). Minimal Δ are called *good explanations*.

While conditions (1) and (2) just define what is an abductive problem and what is a solution, conditions (3) and (4) impose restrictions for a solution to be considered relevant: condition (3) avoids, for instance, that Δ be taken as the collection of all formulas, or as a single bottom particle (which would entail any other formulas). For practical reasons, in the direction of condition (5) it may also be required that the formulas in Δ are restricted to certain special types (usually atomic formulas, or conjunction of atomic formulas).

Since the compactness theorem holds for our logic (and indeed, for all **LFI**s), Γ and Δ can always be taken as finite sets; moreover, since our logic is adjunctive (that is, a finite collection of formulas deduces its conjunction and vice-versa), we may further reduce Γ and Δ to single formulas.

Now, making use of the tableau system for **LFI1**, we can define an *abduction mechanism* by means of the steps below.

We first define the set-theoretical operation of *crossing* of sets: given finite sets $\Delta_1, \Delta_2, \dots, \Delta_n$, their crossing is defined as follows:

- i) $\oplus_{1 \leq i \leq n} \Delta_i = \{X : X = \{y_1, \dots, y_n\}, y_i \in \Delta_i\}$;
 - ii) $cross(\Delta_1, \Delta_2, \dots, \Delta_n) = \{X : X \in \oplus_{1 \leq i \leq n} \Delta_i \text{ and for no } X' \in \oplus_{1 \leq i \leq n} \Delta_i, X' \subseteq X\}$.
- In other words, the elements in the crossing are the minimal elements in $\oplus_{1 \leq i \leq n} \Delta_i$ with respect to the partial ordering \subseteq . For example, if $\Delta_1 = \{a, b\}$ and $\Delta_2 = \{a, c, d\}$ then $cross(\Delta_1, \Delta_2) = \{\{a\}, \{b, c\}, \{b, d\}\}$.

Intuitively, if Δ_i is the set of alternative explanations for the fact ϕ_i , for $1 \leq n$, then $cross(\Delta_1, \Delta_2, \dots, \Delta_n)$ is the minimal set of alternative explanations for the conjunctive fact $\phi_1 \wedge \dots \wedge \phi_n$.

Some immediate and useful properties of crossing are the following:

THEOREM 5.2

Let Δ, Γ and Σ be any sets;

1. If $\Delta \subseteq \Sigma$ then $cross(\Delta, \Sigma) = \Delta$.
2. Crossing is a commutative operation, that is, $cross(\Delta, \Sigma) = cross(\Sigma, \Delta)$
3. Crossing is an associative operation, that is:
 $cross(\Delta, cross(\Sigma, \Gamma)) = cross(cross(\Delta, \Sigma), \Gamma)$

PROOF. By usual set-theoretical methods. ■

Let \mathcal{T} be a tableau with branches $\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_n$. The *closure* of a branch \mathcal{B} is the smallest set $cl(\mathcal{B})$ of signed formulas that, when added to the branch, would close it: $cl(\mathcal{B}) = \{\bar{Q}x : Qx \in \mathcal{B}\}$.

The *closure of the tableau* \mathcal{T} is the crossing of all branch closures: $cl(\mathcal{T}) = cross(cl(\mathcal{B}_1, \dots, cl(\mathcal{B}_n))$.

5.1 *Examples: the Advantages of Using LFI1-Tableaux*

Consider, as an example from the folklore, a theory Γ containing the following sentences:

$\Gamma = \{\alpha \rightarrow \gamma, \beta \rightarrow \gamma\}$ where α means “It rained last night”, β means “the sprinkler was left on”, and γ means “the grass is wet”.

If we observe that the grass is wet, and we want to explain why this is so, “It rained last night” is an explanation, but “the sprinkler was left on” is another competitive (though not excludent) explanation. Tableaux permit to automatically compute “good explanations”, that is, minimal non-trivial explanations, and this is the main task of abduction. After we have such explanations we may employ some form of preference reasoning in order to discard some explanations or to choose the best one among competitors.

For instance, the hypothetical explanation that the sprinkler was left on may be true, but cancelled by the fact that the main water register was known to be off: in any case, some choices may be necessary in order to implement a preference policy for ranking multiple explanations – facts may have precedence over hypothetical explanations, and likelihood may be used to classify explanations. Although this is an important part of the whole question that will affect the usefulness of the automatic explanations produced, it is not part of the abduction problem as originally posed.

Nor, to make it as clear as possible, the question of non-monotonicity has anything to do with the logic that is used to produce the abductive output: **LFII**, the logic we are using there, is plainly monotonic (as much as classical logic, used in [12]). Non-monotonic reasoning, if necessary, will be used in further steps, and is irrelevant here.

We first have to be sure that there exist solutions for the abductive problem under the logic **LFII** satisfying the above conditions (1) to (5) (taking into consideration that Γ and Δ are now reduced to single formulas, as previously discussed).

THEOREM 5.3

Let α and β be sentences such that $\alpha \not\vdash_{LFII} \beta$ (thus constituting an abductive problem). Then there exists an abductive solution γ if and only if $\alpha \not\vdash_{LFII} \neg\beta \wedge \neg\bullet\beta$ and $\neg\alpha, \neg\bullet\alpha \not\vdash_{LFII} \beta$.

PROOF. The result can be established by inspecting all dyadic valuations in the (finite) universe of all propositional variables involved. Alternatively, a constructive argument can be obtained by adapting the proof of Theorem 5 in [20]. ■

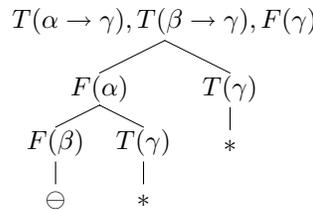
We call *compatible* an abduction problem under the above conditions. Theorem 5.3 shows that compatible abduction problems have at least one solution. Applying this to our definition we have:

EXAMPLE 5.4

A case where **LFII**-tableaux and classical tableaux give the same result:

Let $\Gamma = \{\alpha \rightarrow \gamma, \beta \rightarrow \gamma\}$; of course $\Gamma \not\vdash_{LFII} \gamma$;

- Running an **LFII**-tableau for $T(\Gamma) \cup \{F(\gamma)\}$ produces:

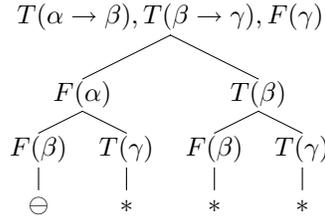


- By collecting all formulas that (when appropriately signed) would close the open branches, that is, by defining sets Σ_i , for $1 \leq i \leq m$, in such a way that $\alpha \in \Sigma_i$ iff $F(\alpha)$ occurs in the open branch, and $\neg\alpha \in \Sigma_i$ iff $T(\alpha)$ in the open branch, we obtain in this case:
 $\Delta_1 = \{\alpha\}$ and $\Delta_2 = \{\beta\}$
- Clearly $\Gamma \cup \Delta_1 \vdash_{LFI1} \gamma$ and $\Gamma \cup \Delta_2 \vdash_{LFI1} \gamma$, and thus Δ_1 and Δ_2 are solutions for the abduction problem;
- Both are good explanations, and is now the user policy to chose between them. The result in this case would be the same using classical tableaux or **LFI1**-tableaux.

Of course, when there are several open branches and several sets Σ_i occur, the crossing $cross(\Sigma_1, \Sigma_2, \dots, \Sigma_n)$ gives the minimal joint explanation.

EXAMPLE 5.5

- Let $\Gamma = \{\alpha \rightarrow \beta, \beta \rightarrow \gamma\}$; clearly $\Gamma \not\vdash_{LFI1} \gamma$.
- Running an **LFI1**-tableau for $T(\Gamma) \cup \{F(\gamma)\}$ produces:



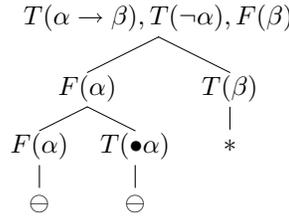
- $\Delta_1 = \{\alpha\}$ and $\Delta_2 = \{\beta\}$ are solutions for the abduction problem.

EXAMPLE 5.6 (“Impossible” explanations explained.)

Suppose that we know that, if it rained last night, then the grass is wet; we know that the grass is wet, but we also know (or have been informed, or have independent evidences for it) that it did not rain. How to explain that the grass is wet?

Let the situation be represented as $\Gamma = \{\alpha \rightarrow \beta, \neg\alpha\}$; here $\Gamma \not\vdash_{LFI1} \beta$, but no classical tableau is able to find an explanation, since the only possible candidate, α , has to be ruled out by clause (3) of Definition 5.1, as it entails deductive triviality. However, **LFI1**-tableaux will be able to provide a solution, based on their intrinsic construction: simply, in situations like this, common sense suggests that raining may be held as an explanation, if the information that it did not rain is uncertain or dubious.

- A **LFI1**-tableau for $T(\Gamma) \cup \{F(\beta)\}$ produces:



- $\Delta_1 = \{\alpha\}, \Delta_2 = \{\alpha, \neg\bullet\alpha\}$ are alternative explanations (as they would close the open branches), and $cross(\Sigma_1, \Sigma_2) = \{\alpha\}$ gives the minimal joint explanation.

This explanation assumes that $\neg\alpha$ is uncertain or dubious, and therefore *does not* violate Definition 5.1: just check that $\Gamma, \alpha \vdash_{LFI1} \beta$ but $\Gamma, \alpha \not\vdash_{LFI1} \gamma$ for γ distinct from β . It is important to note that this situation is considered to be problematic for classical tableaux (cf. [12], pp. 108 and 109).

EXAMPLE 5.7 (Explanations that avoid hasty conclusions.)

We know that taking certain drugs has good consequences for the health, but also the same drugs, under certain conditions, will produce undesirable effects on the health: represent this situation as $\alpha \rightarrow \beta$ and $\alpha \rightarrow \neg\beta$. Under classical reasoning (using classical tableaux, or any other classical inference mechanism) an immediate conclusion would be $\neg\alpha$, that is, we should not take this drug— but this is contrary to common sense, as the contradictory effects could be explained by inappropriate doses, or by different health conditions in different people, and so on. Using **LFI1**-tableaux, however, this case turns out to be an interesting abduction problem: $\alpha \rightarrow \beta, \alpha \rightarrow \neg\beta \not\vdash \neg\alpha$, as shown in item (b) of Example-4.1.

We are thus invited to look for an abductive explanation: this explanation, automatically produced by the **LFI1**-tableau, is that the drug is to be banned only if the contradictory effects are undebatable, that is, if $\neg \bullet \beta$ (or equivalently, $\circ\beta$) holds; item (c) of Example-4.1 then shows that $\Delta = \{\neg \bullet \beta\}$ is an explanation: the resulting **LFI1**-tableau for $\alpha \rightarrow \beta, \alpha \rightarrow \neg\beta, \neg \bullet \beta \vdash \neg\alpha$ is closed.

EXAMPLE 5.8 (Whodunit ?)

A diamond was stolen, and only Alice and Bob were present. Since there are no proofs (but only evidences) against them, the police initially consider that they are not guilty, but certainly one of them is guilty, that is, the evidence basis contains $\Gamma = \{\neg\alpha, \neg\beta, \alpha \vee \beta\}$ where α and β stand, respectively, for “Alice is guilty” and “Bob is guilty”. At this point, $\Gamma \not\vdash_{LFI1} \alpha$ and $\Gamma \not\vdash_{LFI1} \beta$, so we have two abductive problems.

Now, by running **LFI1**-tableaux for $T(\Gamma) \cup \{F(\alpha)\}$ and for $T(\Gamma) \cup \{F(\beta)\}$, we easily see that either $\circ(\neg\alpha)$ (meaning that the initial supposition about Alice’s innocence was indeed consistent) or $\circ(\neg\beta)$ (meaning, alternatively, that the initial supposition about Bob’s innocence was indeed consistent) would decide the question: indeed,

$$\{\neg\alpha, \neg\beta, \alpha \vee \beta, \circ(\neg\alpha)\} \vdash_{LFI1} \beta$$

and

$$\{\neg\alpha, \neg\beta, \alpha \vee \beta, \circ(\neg\beta)\} \vdash_{LFI1} \alpha$$

This coincides with the common sense rule: defending the innocence of one of them amounts to the culpability of the other,

These examples illustrate the fact that employing logics of formal inconsistency in the general problem of abduction has interesting consequences, automatically producing meaningful explanations that would be imperceptible within the classical environment. As mentioned in Section 4, **LFI1** is not the only choice, and several other **LFI**s would play a similar role.

6 Abduction Procedures Involving Quantification

As discussed in Section 1, the extension to the first-order case of the ideas about obtaining abductive explanations by means of tableaux is not only quite natural, but expected in real applications. Although there are some slight technical complications, from the tableau-proof-theoretical standpoint all the grounding constructions are already at our disposal: the

logic **LFII***, first-order extension of the propositional **LFII**, has been studied in details in [5] and [7]. I briefly recall the main ideas about **LFII*** and show how the underlying tableau procedure can be used in abductive problems. Let Σ be the signature of **LFII** enriched with \forall and \exists , (without functional symbols), and **Var** be a set of variables. The formulas of **LFII*** are defined as in classical first-order logic, with the addition of the extra symbol \circ (to be read, as already explained, as “it is consistent that...”). Formulas are inductively defined starting from arbitrary k -ary predicate symbols $R(x_1, \dots, x_k)$ and equality $x_1 = x_2$, considered to be atomic. In general, if α, β are formulas and x is a variable, then $\alpha \vee \beta$, $\neg\alpha$, $\forall x\alpha$, $\exists x\alpha$ and $\circ\alpha$ are formulas,

All the familiar syntactic notions of free and bound variables, closed formulas (sentences), substitution, etc., are defined as usual.

From the semantical side, sentences of **LFII*** are interpreted using either truth-functional three-valued valuations (cf. [5]) or alternatively by (non-truth-functional dyadic (bi-valued) valuations.

From the syntactical side, what interests us here for the sake of abduction, a tableau system for **LFII*** is obtained by adding to the tableau rules of **LFII** the following rules for the quantifiers:

$$\begin{array}{ll}
 (R.19) \frac{T(\forall x(Ax))}{T(\alpha(t))} & (R.20) \frac{T(\exists x\alpha(x))}{T(\alpha(s))} \\
 (R.21) \frac{T(\neg\forall x\alpha(x))}{T(\exists x\neg\alpha(x))} & (R.22) \frac{T(\neg\exists x\alpha(x))}{T(\forall x\neg\alpha(x))} \\
 (R.23) \frac{F(\forall x(Ax))}{T(\alpha(s))} & (R.24) \frac{F(\exists x\alpha(x))}{T(\alpha(t))} \\
 (R.25) \frac{F(\neg\forall x\alpha(x))}{F(\exists x\neg\alpha(x))} & (R.26) \frac{F(\neg\exists x\alpha(x))}{F(\forall x\neg\alpha(x))} \\
 (R.27) \frac{T(\bullet(\forall x\alpha(x)))}{T(\exists x \bullet \alpha(x)), T(\forall x(\alpha(x)))} & (R.28) \frac{T(\bullet(\alpha\exists x\alpha(x)))}{T(\exists x \bullet \alpha(x)), T(\forall x\neg(\alpha(x)))} \\
 (R.29) \frac{F(\bullet(\forall x\alpha(x)))}{F(\exists x \bullet \alpha(x)), F(\forall x(\alpha(x)))} & (R.30) \frac{F(\bullet(\alpha\exists x\alpha(x)))}{F(\exists x \bullet \alpha(x)), F(\forall x\neg(\alpha(x)))}
 \end{array}$$

where t is an arbitrary term and s is a new term with respect to $\forall x\alpha(x)$, i.e., it does not appear in any branch containing $\forall x\alpha(x)$ (respectively for $\exists x\alpha(x)$).

The method introduced here for obtaining automatic explanations can thus be extended to first-order theories; this involves some additional complications, because **LFII***-tableaux, as much as their classical counterparts, are not adequate for showing finite satisfiability. However, **LFII***-tableaux can be modified to yield a procedure that will express finite satisfiability, by adapting the results of [24].

7 Conclusions, and Perspectives

A mechanism for performing automatic abduction based upon tableaux for logics of formal inconsistency has been presented; such logics are infraclassical logics that permit a fine

control of reasoning under contradiction in the presence of malformed or vague hypotheses, while maintaining the full power of logic when the information is consistent. This is a major feature of our systems, as much as profiting from conflicting or contradictory information is a major problem in information systems, given that this information is too valuable to be thrown away (actually, it may be specially rich in content). The mechanism presented here is thus capable of satisfactorily solving an extensive class of abductive problems in propositional reasoning; it can also be upgraded to encompass quantified reasoning as well. This point has been illustrate with examples, and some related issues have been discussed. The points raised here have much in common with belief revision, default reasoning, the “closed world assumption” and “negation as failure” of logic programming, and databases with evolutionary constraints, thus making our proposal valuable for applications. Abduction, however, can also be regarded, from a much more abstract standpoint, as a companion for argumentation; from this perspective, any attempt to make abduction congenerous of deduction is positive. My proposals go in this direction.

References

- [1] W. A. Carnielli. Possible-translations semantics for paraconsistent logics. In “Frontiers in Paraconsistent Logic: Proceedings of the I World Congress on Paraconsistency” ,Ghent, 1998,(editors D. Batens, C. Mortensen, G. Priest, and J.-P. van Bendegem). Baldock: Research Studies Press, King’s College Publications, London, pp.149-163, 2000.
- [2] W. A. Carnielli and J. Marcos. A taxonomy of C-systems. In “Paraconsistency: The Logical Way to the Inconsistent - Proceedings of WCP’2000” (editors W.A. Carnielli, M. E. Coniglio and I. M. L. D’Ottaviano) Marcel Dekker, New York, pp. 1-94, 2002. Pre-print available at CLE e-Prints, vol, 1(5), 2001: URL = http://www.cle.unicamp.br/e-prints/abstract_5.htm.
- [3] L. Magnani. Inconsistencies and creative abduction in science. In “AI and scientific creativity. Proceedings of the AISB99 Symposium on Scientific Creativity”, Society for the Study of Artificial Intelligence and Simulation of Behaviour, Edinburgh College of Art and Division of Informatics, University of Edinburgh, Edinburgh, pp. 1-8, 1999.
- [4] C. S. Peirce. The Collected Papers of Charles Sanders Peirce. Edited by Charles Hartshorne and Paul (vol 1-6) and Arthur Burks (volumes 7-8). Harvard University Press, Cambridge, Massachusetts, 1931-1958.
- [5] W. A. Carnielli, J. Marcos and S. de Amo. Formal inconsistency and evolutionary databases *Logic and Logical Philosophy* 16, pp. 115-152, 2000.
- [6] W. A. Carnielli, M. E. Coniglio and J. Marcos. Logics of Formal Inconsistency. *Handbook of Philosophical Logic* (editors D. Gabbay and F. Guentner), Kluwer Academic Publishers, 2005. In print.
- [7] S. de Amo, W. A. Carnielli and J. Marcos. A logical framework for integrating inconsistent information in multiple databases. *Lecture Notes in Computer Science* 2284, Springer-Verlag, Berlin, 2002, Proceedings of the II International Symposium on Foundations of Information and Knowledge Systems- FoIKS 2002. Thomas Eiter and Klaus-Dieter Schewe, editors, pp.67-84.
- [8] I. M. Bocheński, A History of Formal Logic. Ttranslated and edited by Ivo Thomas. Notre Dame, Indiana, 1961.
- [9] E. J. Ashworth. Propositional logic in the sixteenth and early seventeenth Centuries. *Notre Dame Journal of Formal Logic* Volume IX, Number 2, April 1968.
- [10] M. Tweedale. Abelard and the culmination of the old logic. In N. Kretzmann, A. Kenny, J. Pinborg (eds.), *The Cambridge History of Later Medieval Philosophy*. Cambridge, Cambridge University Press, 1982, pp. 143-157.
- [11] D. R. Anderson. *Creativity and the Philosophy of C.S. Peirce*. Martinus Nijhoff, Dordrecht, 1987.
- [12] A. Aliseda. Seeking Explanations: Abduction in Logic, Philosophy of Science and Artificial Intelligence. Ph.D. Dissertation. Stanford University. 1997. Institute for Logic, Language and Information (ILLC), Dissertation Series, Universiteit van Amsterdam. 1997. Available from: URL = <http://www.wins.uva.nl/research/illc/Dissertations>
- [13] O.Arieli, M.Denecker, B.Van Nuffelen, M.Bruynooghe. Coherent integration of databases by abductivelogic programming. *Journal of Artificial Intelligence Research* 21, pp. 245-286, 2004.

- [14] J. Bueno-Soler and W.A. Carnielli. Possible-translations algebraization for paraconsistent logics, *Bulletin of The Section of Logic*,34(2), pp.77–92, 2005.
- [15] W. A. Carnielli. Systematization of the finite many-valued logics through the method of tableaux. *The Journal of Symbolic Logic*, 52:473–493, 1987.
- [16] W. A. Carnielli and J. Marcos. Tableau systems for logics of formal inconsistency. In Proc. Int. Conf. on Artificial Intelligence (IC-AI'01), Vol. II, 2001pp. 848-852. CSREA Press.
- [17] W. A. Carnielli and M. Lima-Marques. Reasoning under Inconsistent Knowledge. *Journal of Applied Non-Classical Logics* 2(1), 1992, pp. 49-79.
- [18] W. A. Carnielli, M.E. Coniglio and J. Marcos. Logics of Formal Inconsistency, Handbook of Philosophical Logic (editors D. Gabbay and F. Guenther),Kluwer Academic Publishers, volume 14, 2005. Preliminary version available at *CLE e-Prints*, Vol. 5(1), 2005. URL = <http://www.cle.unicamp.br/e-prints/articles.html>
- [19] C. Caleiro, W. A. Carnielli, M. E. Coniglio, and J. Marcos. Two's company: "The humbug of many logical values". In: J.-Y. Bziau, editor, *Logica Universalis*, pp.169-189. Birkhuser Verlag, 2005.
- [20] M. E. Coniglio. Obtención de respuestas en bases de conocimiento a partir de completaciones. In: Query Procedures in Knowledge Bases by means of Completions- Proceedings of XXI JAIIO (Argentinian Symposium on Informatics and Operational Research), SADIO, pp. 1.31–1.63, 1992.
- [21] R. L. Epstein, On mathematics. manuscript, October 2005.
- [22] J. Marcos. Semânticas de Traduções Possíveis (Possible Translations Semantics, in Portuguese). Master Thesis, IFCH-UNICAMP,1999, Campinas, Brazil.
- [23] P. Mancosu. Mathematical explanation: problems and prospects. *Topoi* 20, pp. 97–117, 2001.
- [24] G. Boolos. Trees and finite satisfiability: proof of a conjecture of Burgess. *Notre Dame J. Formal Logic* 25, no. 3 , 193–197, 1984.

Received January 31, 2006