

Relevant Logic and Paraconsistency

John Slaney

The Australian National University
and National ICT Australia *

Abstract. This is an account of the approach to paraconsistency associated with relevant logic. The logic **fde** of first degree entailments is shown to arise naturally out of the deeper concerns of relevant logic. The relationship between relevant logic and resolution, and especially the disjunctive syllogism, is then examined. The relevant refusal to validate these inferences is defended, and finally it is suggested that more needs to be done towards a satisfactory theory of when they may nonetheless safely be used.

1 Why Paraconsistency?

The core business of logic is to underpin reasoning. The distinction is important: a logic is a *theory*; reasoning is a *process*. The activities of a reasoner are not dictated by logic, but are described by it in the sense that logic issues permissions—assurances that certain forms of inference will never lead into error—and restrictions. The restrictions are not on inferences, for the reasoner may of course reason invalidly if it so wishes, but on the formation of theories not closed under the principles enunciated in the logic. Thus logic is a guide to reasoning as well as a description of it: a “normative science”, as Ramsey succinctly put it.¹ That sets up the central paradox of the philosophy of logic: norms are necessarily prior to the behaviour they circumscribe or guide, but phenomena are prior to the science that describes them, so how is such a normative science possible? The nature of logic will not be settled in this paper, but we may at least investigate an important issue that bears directly upon it. The problem arises in matching logic to the activity of reasoning with inconsistent information. Such reasoning is desirable in practice, and apparently rational, yet accommodating it challenges the most basic principles on which standard formal logic is built. The standard view of contradictions is that they are (necessarily) false, so no inference with contradictory premisses can lead into error: if you start from a contradiction, you are already in error, so it simply does not matter what inference you make. That is, logic validates the argument form:

$$\frac{A \quad \neg A}{B}$$

* National ICT Australia is funded through the Australian Government’s *Backing Australia’s Ability* initiative, in part through the Australian Research Council.

¹ This conversational remark of Ramsey’s, reported by Wittgenstein in the *Philosophical Investigations*, applies equally to other formal sciences such as grammar.

On the other side of the coin, the classical account prohibits theories which are inconsistent in the sense of containing a contradiction but for which it is a non-trivial question what else they contain. This prohibition is not consonant with reasoning practice, in which it is fairly common to encounter bodies of information which are inconsistent in ways which should not inhibit reasoning.

Three examples of such reasoning situations in which giving up in the face of inconsistency is not an appropriate response will serve as illustrations:

1. Database management: data integration. It is common for databases to make available data from many sources, and important for them to be able to do this while imposing strong integrity constraints. Since the sources may have overlapping domains, in which they may conflict, and since they are not always expected to satisfy the integrity constraints, especially when taken in combination with each other, there arises a need for the deductive part of the database to cope with inconsistency. Many techniques for this have been suggested, sometimes based on non-monotonic reasoning because it is expected that consistency should be restored through repeated revision of the data over time. However, the static logical description of inconsistent data also demands to be taken seriously, especially as there may be cases in which the inconsistency of the global database is undetected at the time of query answering and in which data may become obsolete and be replaced faster than the corpus can be checked for consistency and corrected.

2. Software engineering: merging specifications. Software specifications must frequently be put together from many sources, fragments being in different languages and contributed by different experts. There may be indeterminacy as to whether an apparent conflict (one expert says that a transition from state a to state b is possible, while another says that a transition from state α to state β is impossible) is to be resolved by distinguishing between a and α , b and β or whether it should remain in the proto-specification as a genuine disagreement, perhaps to be cleaned up later in some way. A standard move is to represent the amalgamated specification in a logical framework where the “truth values” are tuples of values (true, false, unknown) standing for the opinions of the various experts from whom the fragments are taken. An alternative is to use a four-valued scheme to allow the cases in which we have been told that something is true, told that it is false, both or neither. The logical manipulation of this “useful 4-valued logic” [4] requires para-consistency in order to cope with the truth value gaps and gluts without collapse.

3. Epistemic logic: first order beliefs. It is usual in doxastic and epistemic logic to consider belief sets closed under logical entailment rather than sets of propositions immediately and explicitly available to an agent. This concentration on implicit beliefs is essential to any treatment based on normal [multi-]modal logics, since it is an outcome of the K principles that the belief sets in question are deductively closed—the *theories* of agents rather than their explicit contents. The question of whether

an agent believes p is then one of whether p follows logically from the agent's explicit beliefs, not one of whether p is actually present to the agent. Thus it does not lie open to introspection. In a language as rich as first order logic, indeed, it does not lie open to effective determination at all. It is thus quite possible for an agent to arrive at a belief that p , unaware (explicitly) that $\neg p$ follows from its beliefs, and thus (implicitly) to embrace a contradiction. This situation is not particularly abnormal, and calls for paraconsistent logical treatment, not for dismissal as a case that “cannot happen”.

There are three styles of approach to such inconsistency, very likely each with its appropriate range of applications. Firstly, the inconsistent theory may be regarded as a temporary departure from a previously consistent one and the problem as one of restoring consistency by revision or some other type of nonmonotonic reasoning. Secondly, the formulae whose consequences are (globally) inconsistent may be treated like soft constraints in an overconstrained CSP, and “large” consistent subtheories sought without necessarily changing the inconsistent theory. Of course, these two responses may be combined in various ways. The third approach is to regard the inconsistent theory as logically respectable just as it stands, and therefore to adopt a genuinely paraconsistent logic as the underlying theory of valid inference. It is this third option which is the subject of the present paper.

Paraconsistency requires inconsistent theories to be entertained without collapse into triviality, and hence affects most directly the logic of negation. However, the logic of negation is obtained by fitting an account of denial into the framework provided by positive logic. Accordingly, it is with the negation-free part of logic that we begin the next section.

2 The Relevant Approach

There are many paraconsistent systems on the menu, but one of the oldest and most systematically developed is relevant logic [2, 3, 21, 23] in which paraconsistency is not itself the main motivation but arises naturally out of other concerns. Those other concerns historically included securing relevance—most simply that in the propositional part of the logic, no implication should be accounted valid unless antecedent and consequent share a variable. The classical inference from $p \wedge \neg p$ to q of course violates even this simple relevance requirement. However, the deeper motivation of the relevant family of logics is to formalise a notion of proof in which part of what constitutes a derivation of a conclusion *from* assumptions is that the assumptions be *used* in deriving the conclusion, and to marry this structural condition on derivations with systematic logical properties such as a decent deduction theorem and, importantly, with a very “ordinary” account of the familiar truth-functional connectives.

2.1 Relevant positive logic

Their truth-functional character gives the operations of conjunction and disjunction all of their logical properties from the relevant point of view.

This puts relevant logic in the same family as intuitionist logic and the usual modal logics, rather than that of linear logic and the other substructural systems, as regards its treatment of the extensional connectives. Semantically, these connectives are evaluated at each world using only information local to that world—specifically, the values of the conjuncts or disjuncts at that world—in the standard way. This results in the set of “positive first-degree entailments” which may be characterised using \wedge and \vee as multiary connectives thus:

1. Where p_1, \dots, p_m and q_1, \dots, q_n are all atomic,
 $p_1 \wedge \dots \wedge p_m \vdash q_1 \vee \dots \vee q_n$ iff for some $i \leq m$ and $j \leq n$, $p_i = q_j$.
2. $A_1 \wedge \dots \wedge A_{i-1} \wedge (B \vee C) \wedge A_{i+1} \wedge \dots \wedge A_m \vdash D$ iff
 $A_1 \wedge \dots \wedge A_{i-1} \wedge B \wedge A_{i+1} \wedge \dots \wedge A_m \vdash D$ and
 $A_1 \wedge \dots \wedge A_{i-1} \wedge C \wedge A_{i+1} \wedge \dots \wedge A_m \vdash D$
3. $A \vdash B_1 \vee \dots \vee B_{j-1} \vee (C \wedge D) \vee B_{j+1} \vee \dots \vee B_n$ iff
 $A \vdash B_1 \vee \dots \vee B_{j-1} \vee C \vee B_{j+1} \vee \dots \vee B_n$ and
 $A \vdash B_1 \vee \dots \vee B_{j-1} \vee D \vee B_{j+1} \vee \dots \vee B_n$

Algebraically, the structures modelling this fragment of logic are simply distributive lattices. That is, a “propositional structure” for this logic is a set on which are defined binary operations \wedge (meet) and \vee (join) each of which is idempotent, commutative and associative, such that:²

$$\begin{aligned} a \wedge (a \vee b) &= a \\ a \vee (a \wedge b) &= a \\ a \wedge (b \vee c) &= (a \wedge b) \vee (a \wedge c) \\ a \vee (b \wedge c) &= (a \vee b) \wedge (a \vee c) \end{aligned}$$

A model in such a propositional structure is just a homomorphism from the formula algebra into the lattice in the obvious sense, lattice meet corresponding to conjunction and lattice join to disjunction. There is nothing specific to relevant logic about this: the story for classical logic is exactly the same, as it is for intuitionist logic and many others. This is worth emphasising, since it is a feature of the relevant approach that it does not depend on any strange definition of entailment or the like, but agrees totally with the standard logics as regards the basic truth-functional connectives.

Classically, negation and implication are basic truth-functional connectives as well. Relevantly, although we shall urge below that negation can reasonably be regarded as truth-functional, implication is something else. The relevant theory of implication belongs firmly to the tradition of substructural logic. There is a semantic account, to be sure, but the fundamental intuitions concerning implication are deduction-theoretic in nature.

The essence of the implication operation, \rightarrow , is encapsulated in the deduction equivalence:

$$\Gamma \vdash A \rightarrow B \text{ iff } \Gamma, A \vdash B$$

² As a trivial fact of lattice theory, only one of the first two and one of the last two of these postulates are needed, but all four are given here to emphasise the complete duality of the two operations.

That is, some information Γ suffices for an implication $A \rightarrow B$ iff the assumption of A in the context of Γ suffices for B . The implication records the availability of an inference from A to B . This much, too, is common to classical logic, intuitionist logic, the whole range of substructural logics, several many-valued logics and others. Those logics tend, however, to disagree about the details of which entailments hold among formulae involving the implication operator. Does $(p \rightarrow q) \rightarrow p$ entail p ? Does $(p \rightarrow q) \rightarrow (q \rightarrow p)$ entail $q \rightarrow p$? Does p entail $q \rightarrow q$? Does $p \rightarrow (p \rightarrow q)$ entail $p \rightarrow q$? These questions are not answered by the deduction equivalence alone, but by the underlying theory of how inference is structured, which differs from logic to logic. Differences at that level affect the account of what it is to “assume A in the context of Γ ”, and hence the possible readings of the compound object Γ, A .

The fundamental idea of relevant implication is simple: a derivation of B is not “from” a structure Γ, A as required for the deduction equivalence unless A is in the appropriate sense *used* in reaching B . The paradigm cases of “use” are clear: both premises are used in an application of the rule of detachment in which D is derived from C and $C \rightarrow D$, but the first premise is *not* used in the derivation of D from C and D by the rule of iteration. Use is transitive, so whatever is used in deriving lemmas from axioms is used in the derivation of a theorem from those axioms by means of the lemmas.

The implicational fragment of the standard relevant logic **R** results by making this guiding idea rigorous in a very natural way. Let X, Y , etc be the sets of assumptions used in derivations. Note that the *sets* of assumptions correspond in general to *multisets* of formulae, since nothing prevents two or more distinct assumptions of the same formula. We write $X : A$ to mean that A is derived relevantly from assumptions X . Now the introduction and elimination rules for the implication connective \rightarrow are simple and obvious. For introduction, we have

$$\frac{X : B}{X \setminus \{x\} : A \rightarrow B}$$

where x is an assumption of the formula A and (for relevance) it is required that $x \in X$. The elimination rule is detachment, as is familiar:

$$\frac{X : A \rightarrow B \quad Y : A}{X \cup Y : B}$$

To get the calculus started, it is of course relevantly fine to derive a formula from an assumption of itself, though not in general from a set of assumptions of which it happens to be a member, since there may be no way to *use* all the side assumptions in the derivation. It will be apparent that this logic is the pure implication fragment of the substructural logic BCIW: that which allows collections of formulae to have the associative and commutative character of multisets, and allows the structural rule of contraction, but disallows weakening. Thus, despite its appearance and reputation as something exotic, the relevant logic of implication is actually a logic very much in the mainstream tradition: it is in fact exactly like intuitionist pure implicational logic except for the extra feature that

due attention is paid to which assumptions are used and which are idle. In particular, it does not involve violent departures from logical tradition such as non-transitive entailment relations, restrictions on the nesting of connectives or special treatment for formulae of some distinguished kind such as inconsistent ones.

R is natural, given the motivation, but it is not the only possibility. There is not one relevant logic, just as there is not one modal logic, but a family, since the interpretation of the guiding notion of “relevance” is an equivocal matter. The paradigm cases of “use” may indeed be clear, but more delicate questions soon arise. Does the order of assumptions matter? – that is, is the effect of assuming B in the context of A different from that of assuming A in the context of B ? For the standard relevant logic **R** there is no difference, but the systems **T** and **E** originally preferred by Anderson and Belnap [2] do make a distinction. Again, how are uses to be *counted*? Have we used *all* of the assumptions in deriving A from A, A ? Not according to **R**: if you want to discharge A twice you have to use it twice; in the semi-relevant system **RM** [7] however, repetitions don’t count. Conversely, according to **R**, once an assumption is in play it may be used as often as it takes to reach a conclusion, and then discharged in just one step, whereas in the weaker system **C**, whose pure implication fragment is that of linear logic, once an assumption is used it is consumed so to use it twice you must assume it twice. Different decisions as to the structural rules lead to different logics, although in a straightforward sense the *meaning* of the implication connective is the same in all of them.

It is not part of the present purpose to examine the differences between the many logics in the relevant family. Still less is it to declare one of them the One True Logic. Rather let it be noted that they are all constructed by fitting together the stable and relatively uncontroversial classical theory of conjunction and disjunction with a weakening-free substructural logic of implication. Because the fundamental motivation for the former is semantic, based on the “truth tables” for the dual pair of lattice connectives, while the latter arises from considerations concerning the structure of deduction and is thus essentially proof-theoretic, combining the two is not trivial. Specifically, the principle of distributivity of \wedge over \vee and *vice versa* is proof-theoretically unnatural except in the context of intuitionist logic and the like in which the fine distinctions required for substructurality are obliterated. The relevant logics, however, derive their distinctive character from the way in which they manage to maintain both the “intensional” and the “extensional” subtheories as motivated above, while combining them without restriction in the richer logic.

To specify the relevant positive logic R_+ deductively, we use two ways of combining assumptions: the formulae A_1 through A_n may be collected into a set $\{A_1, \dots, A_n\}$ or a multiset $[A_1, \dots, A_n]$.³ The set represents “pooling” information, with no particular consideration for which sub-

³ The idea of formulating these logics with two operations for combining assumptions goes back at least to Dunn [8] and has since become fairly standard and been elaborated by many authors [5, 6, 12, 20, 21, 25, 26].

set is used in making a deduction, while the multiset represents pieces of information which have been all been used in combination. Bunching of formulae under these two operations may be nested arbitrarily: any formula A is both an S-bunch (“set bunch”) and an M-bunch (“multiset bunch”); any non-singleton set of M-bunches is an S-bunch and any non-singleton multiset of S-bunches is an M-bunch. Our notation follows standard practice (e.g. [25]) whereby the comma-separated list X_1, \dots, X_n stands for $\{X_1, \dots, X_n\}$ while the semicolon-separated list $X_1; \dots; X_n$ stands for $[X_1, \dots, X_n]$. We also write $\Gamma(\Delta)$ in the normal way to indicate a bunch in which Δ occurs in some place as a sub-bunch, so $\Gamma(\Delta')$ differs from it exactly in that Δ' occurs instead of Δ in that particular place. The logic may be presented in natural deduction style as a calculus of sequents with bunches of formulae on the left and single formulae on the right. The axioms are the sequents of the form $A \vdash A$ and the logical (introduction and elimination) rules are:

$$\begin{array}{c}
\frac{\Gamma \vdash A \quad \Delta \vdash B}{\Gamma, \Delta \vdash A \wedge B} (\wedge\text{I}) \qquad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash A} (\wedge\text{E}) \\
\frac{\Gamma \vdash A \wedge B}{\Gamma \vdash B} (\wedge\text{E}) \\
\frac{\Gamma \vdash A \vee B \quad \Delta(A) \vdash C \quad \Delta(B) \vdash C}{\Delta(\Gamma) \vdash C} (\vee\text{E}) \qquad \frac{\Gamma \vdash A}{\Gamma \vdash A \vee B} (\vee\text{I}) \\
\frac{\Gamma \vdash B}{\Gamma \vdash A \vee B} (\vee\text{I}) \\
\frac{\Gamma \vdash A \rightarrow B \quad \Delta \vdash A}{\Gamma; \Delta \vdash B} (\rightarrow\text{E}) \qquad \frac{\Gamma; A \vdash B}{\Gamma \vdash A \rightarrow B} (\rightarrow\text{I})
\end{array}$$

Note that the arrow goes with intensional (multiset) combination of assumptions, while conjunction goes with extensional (set) combination. There are also structural rules marking the difference between sets and multisets. Both operations, symbolised by the comma and semicolon respectively, are associative and commutative, and both satisfy contraction in the form:

$$\frac{\Gamma(\Delta, \Delta) \vdash A}{\Gamma(\Delta) \vdash A} (\text{W-set}) \qquad \frac{\Gamma(\Delta; \Delta) \vdash A}{\Gamma(\Delta) \vdash A} (\text{W-multiset})$$

Set combination, though not multiset combination, also satisfies the standard rule of weakening:

$$\frac{\Gamma(\Delta) \vdash A}{\Gamma(\Delta, \Delta') \vdash A} (\text{K-set})$$

A small but important detail is that the empty set \emptyset and the empty multi-set \mathbf{I} are different objects satisfying the conditions $\emptyset, \Gamma = \Gamma$ and $\mathbf{I}; \Gamma = \Gamma$ respectively. Formula A is a logical theorem iff $\mathbf{I} \vdash A$ is provable.

All of this looks very much as normal, except for the well-motivated distinction between assumptions which have been combined and those which just co-occur. The very ordinariness of the system is what needs to be stressed: this is a logic in the mainstream tradition of logical theory—static, monotonic and with a familiar look. But for a few wrinkles, it could have been proposed by Frege or Tarski.

Semantically, too, it is much what should be expected of a marriage between a substructural treatment of implication and a truth functional account of the lattice connectives. On the substructural side, the fundamental semantic idea is that of combining two bodies of information, and in particular taking one such body to supply the available inferences and applying it to another which supplies the facts available to serve as inputs to those inferences. Thus if the first body contains the information that all tigers are carnivores (an inference ticket) and the second gives us the information that Timmy is a tiger (a fact), then by applying the first to the second we may deduce that Timmy is a carnivore.

Formally [22] a frame for a logic in the relevant family is a set of evaluation points, which may be thought of as bodies of information, or information states. The set is partially ordered by increasing strength—intuitively, by inclusion of the information. We write $x \leq y$ to mean y is stronger than x . More generally, there is a *ternary* relation defined on the set: $Rxyz$ means that z contains everything that can be derived from y by applying an inference warranted by x .⁴ What properties does this relation have? In the basic case, nothing beyond monotonicity. That is, if x or y is weakened, or if z is strengthened, the relation $Rxyz$ still holds. In the case of particular logics such as \mathbf{R} there are more postulates on the relation, just as there are in the semantic stories corresponding to modal logics stronger than \mathbf{K} , but these are best considered as additions to the basic theory, again as in the modal case. There is one more component to a frame: a distinguished point 0 representing the truth, or the real world, or that which is logically correct. Its characteristic property is that if it says that x is included in y then x is really included in y . That is, $R0xy$ iff $x \leq y$.

A model in such a frame is a function assigning to each atomic formula p a set of points, intended to be those points at which p is evaluated to “true”. Naturally, this satisfies a heredity condition that the set assigned to p is closed under \leq . The modelling condition for implications is the obvious one: $A \rightarrow B$ holds at a point x iff x warrants the inference from A to B ; that is, for all points y and z such that $Rxyz$, if A holds at y then B holds at z . The true formulae in a model are those which hold at 0 , and the valid formulae on a frame (or set of frames) are those which are true in all models on that frame (or on all frames in that set).

⁴ We have sometimes [26] written $Rxyz$ as $y \leq_x z$ which is suggestive of the meaning: y is contained in z from the perspective of x . However, the neutral ‘ R ’ notation is standard in the literature and is followed here for that reason.

Accounting for the lattice connectives \wedge and \vee is simple: require the evaluation points to be world-like in that they treat truth-functional operators truth-functionally. That is, $A \wedge B$ holds at x iff A holds at x and B holds at x , and $A \vee B$ holds at x iff A holds at x or B holds at x . To secure the usual relevant logic \mathbf{R} , it is necessary to impose conditions saying that the “application” of points to each other is associative, commutative and square-increasing:

1. $Rabc \& Rcde \implies \exists x(Rbdx \& Raxe)$
2. $Rabc \implies Rbac$
3. $Raaa$

Of course, these postulates are not inescapable: other logics in the family result by modifying them in various ways, just as they result proof-theoretically by modifying the structural rules governing premise combination. For a wide range of such modifications, the first degree logic (with no nested arrows) is invariant, and conversely the positive logic is a conservative extension of the pure implication fragment.⁵

2.2 Negation

Just as there are many choices along the route to relevant positive logic, the motivating considerations being insufficiently precise to determine how those choices are to be made, so there are several more or less natural ways to add negative particles to the logic. One possible addition is an “absurd” constant \perp with its characteristic property that it (relevantly) entails everything, or equivalently that it holds in no world. This constant is in some sense out of the spirit of the relevant view, though its addition is easily seen to be conservative over the positive logic and since it is a connective, not containing variables, it does not break the relevance conditions such as variable-sharing. It allows a sort of negation to be defined as in intuitionist logic:

$$\hat{\neg}A =_{\text{df}} A \rightarrow \perp$$

However, this kind of negation will hardly do for relevant knowledge representation purposes, since what we typically do with negation is to deny things on the grounds that we think they are false, and few of these things, we may suppose, are so false that they *relevantly* imply absolutely everything. Absurd negation, then, is too strong. On the other hand, relevant *minimal* negation is rather weak. This is obtained by introducing the constant \mathbf{f} without any distinguishing properties, and defining another kind of negation:

$$\tilde{\neg}A =_{\text{df}} A \rightarrow \mathbf{f}$$

This is better, but suffers from the usual drawback of minimal negation, that it does not yield much of a theory because there is nothing to mark the constant \mathbf{f} as a *negative* expression.

⁵ Completeness theorems and similar results for logics in the relevant family may be found in many places: in [22] for instance, or more accessibly in [9] or [23].

A more interesting possibility, yielding a much less trivial account of negation, is to add to the language a connective corresponding to the operation of boolean complement [18]. This is easy enough both syntactically and semantically. Deductively, add to the system of positive logic outlined above the new connective, here symbolised by overscoring, with the rules:

$$\frac{\Gamma, A \vdash B \quad \Gamma, \overline{A} \vdash B}{\Gamma \vdash B} \text{SLEM}$$

$$\frac{\Gamma \vdash A \quad \Gamma \vdash \overline{A}}{\Gamma \vdash B} \text{ECQ}$$

Note that in the rule SLEM (strong law of the excluded middle) the boolean negation goes essentially with the comma of set combination, establishing it as a connective in the truth functional group, like \wedge . Note also that both this rule and ECQ (*ex contradictione quodlibet*) are quite out of keeping with the concern for relevance. The addition to relevant positive logic is conservative, however, so there is a sense in which it does not upset relevant insights.

On the semantic side, boolean negation is naturally introduced by giving it the expected classical truth table at each world. To secure the necessary model-theoretic properties, however, the worlds need to be unordered. What that amounts to is that a new “base” world $0'$ be added with the property that $R0'xy$ iff $x = y$. It must then be shown that such an addition does not change the set of valid inferences in the old vocabulary, so that every counter-model to a nontheorem of the positive logic remains so after the addition of $0'$. In the case of a well-behaved propositional logic like \mathbf{R} , this is routine, though in some cases, especially in richer vocabularies, it may be nontrivial.

The question of whether to add a boolean negation to relevant logic has traditionally divided the community of relevant logicians. There are those who, like the founders of the field Anderson and Belnap, wish to have none of it, and others such as Martin and Meyer [17] who wish to embrace it—at least as a kind of recommended optional extra. Certainly the relevant use criterion for valid implication does not sit well with inference principles such as SLEM: the latter allows the conclusion that there is a *relevant* deduction of B from Γ even in the case where all the work was done by the A and \overline{A} and in which Γ may have been introduced by the explicitly irrelevant weakening rule K-set. Certainly also boolean negation does not readily mix with the relevant implication connective. Even the apparently innocent addition of relevant contraposition in the form

$$(A \rightarrow B) \rightarrow (\overline{B} \rightarrow \overline{A})$$

has unfortunate side effects such as the loss of conservative extension results⁶

The standard approach to negation in relevant logics, also to be found in linear logic [13] and Łukasiewicz many-valued logics [15] among others,

⁶ See for example [23], p.379. With strong contraposition as above, we have $A \rightarrow (A \rightarrow C), (A \rightarrow B) \rightarrow C \vdash C$ which is not valid in \mathbf{R} .

is to weaken the boolean theory sufficiently to bring it into line with the positive logic while keeping its most important systematic property of being an involution in the lattice sense—a dual automorphism of period 2. That is, it has the effect of reversing the order of implication, dualising the other connectives and maintaining left-right symmetry in the logical system. Its characteristic properties are equivalences:

$$\begin{aligned}
& \neg\neg A \dashv\vdash A \\
& \neg A \rightarrow \neg B \dashv\vdash B \rightarrow A \\
& A \rightarrow \neg B \dashv\vdash B \rightarrow \neg A \\
& \neg A \rightarrow B \dashv\vdash \neg B \rightarrow A \\
& \neg(A \wedge B) \dashv\vdash \neg A \vee \neg B \\
& \neg(A \vee B) \dashv\vdash \neg A \wedge \neg B
\end{aligned}$$

Of course, there are one-way inference principles involving negation as well. In \mathbf{R} , for instance, as a result of the structural rule of exchange (commutativity of the semicolon in the deductive system given above) we have

$$1. \quad A \dashv\vdash (A \rightarrow A) \rightarrow A$$

from which, on substituting $\neg A$ for A , rewriting $\neg A \rightarrow \neg A$ as $A \rightarrow A$ and contraposing,

$$2. \quad \neg A \dashv\vdash A \rightarrow \neg(A \rightarrow A)$$

hence

$$3. \quad A \rightarrow \neg A \dashv\vdash A \rightarrow (A \rightarrow \neg(A \rightarrow A))$$

so by contraction (W-multiset) and rewriting $A \rightarrow \neg(A \rightarrow A)$ as $\neg A$:

$$4. \quad A \rightarrow \neg A \vdash \neg A$$

This strong *reductio* principle does not hold in all logics in the relevant family; Anderson and Belnap [2] postulated it by fiat for their systems \mathbf{T} and \mathbf{E} , but in general it fails in systems without exchange and contraction. One of its outcomes is the provability of the truth table tautologies in the vocabulary of \wedge , \vee and \neg , in virtue of the theorem $A \vee \neg A$. To prove this, note that $A \wedge B$ entails $A \vee C$, as a special instance of which $A \wedge \neg A$ entails $A \vee \neg A$. But $A \vee \neg A$ is equivalent to $\neg(A \wedge \neg A)$, so by the strong *reductio* principle we have both $\neg(A \wedge \neg A)$ and $A \vee \neg A$ as theorems of \mathbf{R} .

As logical rules sufficing to govern negation in the deductive calculus, we may take this pair:

$$\frac{\Gamma; A \vdash \neg B \quad \Delta \vdash B}{\Gamma; \Delta \vdash \neg A} (\neg\text{I})$$

$$\frac{\Gamma; \neg A \vdash B \quad \Delta \vdash \neg B}{\Gamma; \Delta \vdash A} (\neg\text{E})$$

Note that relevant negation is defined using multiset (use-sensitive) combination of assumptions and that it thus fits well into the substructural theory of implication underlying the intensional side of relevant logic.

It also fits the truth-functional account of the extensional connectives, with the twist that paraconsistency and its converse are allowed. Suppose that truth and falsehood are neither collectively exhaustive nor mutually exclusive, but rather are treated in the formal semantics of logic as two independent properties that propositions may have. That is, at each point in a frame, there are those formulae which are asserted, or accepted, or true according to that point, and there are those which are denied, rejected, false according to that point. A particular formula may be simply accepted, simply rejected, both accepted and rejected (if the evaluation is confused) or neither (if the evaluation is incomplete). Another view of the matter is that each evaluation point presents two theories: one consists of the formulae asserted and the other of the formulae not denied. Both of these theories are required to be closed under logic. The distinctive feature of the relevant semantics for negation is that to each evaluation point a there corresponds another a^* which asserts just what the first fails to deny, denies just what the first fails to assert. The conditions governing evaluation points in frames apply to both equally. Clearly $a^{**} = a$. Equally clearly, if a is contained in b then b^* is contained in a^* . In fact, to secure the full force of relevant reasoning with negation, this last condition holds also under assumptions, meaning that if $Rabc$ then Rac^*b^* . Naturally, $\neg A$ holds at a point a iff A does not hold at a^* . The classical, boolean account of negation results by imposing the condition $a^* = a$ but relevant logic, leaving open the paraconsistent possibilities, does not require such a strong condition.⁷

As in the case of the $\{\wedge, \vee\}$ fragment, the logic of \wedge , \vee and \neg arises directly from the truth functional semantics. Atomic formulae may or may not be true, and orthogonally to that may or may not be false. Compound formulae have truth conditions and falsehood conditions exactly as in the boolean case except that the two are independent. Thus $A \wedge B$ is true iff its two conjuncts are both true (whether or not they are false as well) and false iff at least one of them is false (whether or not it is also true). Dually, $A \vee B$ is true iff at least one disjunct is true and false iff they are both false. $\neg A$ is true iff A is false and false iff A

⁷ There is a large literature on this kind of negation and the semantic postulates governing it. The 'star' operation on worlds is due to Routley and was featured in [22], since when it has been attacked and defended many times. See [3] for an entry point to the literature.

is true. A entails B iff B is true on every valuation on which A is true and A is false on every valuation falsifying B . This scheme gives rise to the logic **fde** of “first degree entailments” [2]. To check whether A entails B in **fde** it suffices to reduce A to disjunctive normal form and B to conjunctive normal form; the logic validates all of the DeMorgan laws, distribution principles and other moves necessary for this reduction. Then A entails B iff each disjunct of $\text{DNF}(A)$ separately entails each conjunct of $\text{CNF}(B)$. A disjunct of $\text{DNF}(A)$ is a conjunction of literals and a conjunct of $\text{CNF}(B)$ is a disjunction of literals; the former entails the latter in **fde** iff they have a literal in common.

fde is a well known paraconsistent logic. The reason for taking so long to come to this point is to emphasise that in the context of relevant logic **fde** is not an arbitrary choice but arises naturally from the background motivations and is of a piece with the larger logical system. In particular, it represents a theory of negation in harmony with the positive logic to which it is added. In the next section **fde** will be defended against some common objections, after which it will be noted that the first degree entailments themselves do not form an adequate logical system but stand in need of extension to something like a logic in the relevant tradition.

3 Disjunctive Syllogism: Baby or Bathwater?

It is easy enough to set up a logic to omit some unwanted principle of inference—here the inference from a contradiction to an arbitrary conclusion—but less easy to do this in such a way as not to lose with it too many other principles which are not so unwanted. A common view is that the weakened logic should remain as close as possible to classical logic while avoiding the “bad” principle, where “as close as possible” is taken to mean that as many inference forms as possible should be retained, and perhaps that outside the disputed area (say, in consistent domains) the logic should be exactly classical. The logic **fde** draws criticism on this point, usually for failing to validate the classical principle of the *disjunctive syllogism* in the form

$$\frac{A \vee B \quad \neg A}{B} (DS)$$

Undeniably, we sometimes reason in this way: ‘Somebody has eaten the last cookie, and it’s not me, so it must be you!’ Simply denying that the disjunctive syllogism is valid reasoning will not do, for such ordinary episodes of inference should be accounted for rather than legislated away. On the other hand, from a paraconsistent point of view the case against DS is strong. Most famously, it is clearly implicated in the derivation of ECQ through the ancient argument re-discovered some eighty years ago by C. I. Lewis:

$$\frac{\frac{A}{A \vee B} (\vee I) \quad \neg A}{B} (DS)$$

In the light of this argument, every paraconsistent logic must either disallow $\vee I$, disallow DS or somehow allow that a two-step argument whereby a conclusion is derived from a lemma which is derived from an assumption is not really a derivation of the conclusion from the assumption. Of the possible suspects here, DS looks by far the most guilty. Consider the premises of DS. What does $(A \vee B) \wedge \neg A$ amount to? By the distributivity of \wedge and \vee , it amounts to $(A \wedge \neg A) \vee (B \wedge \neg A)$. And we have no reason to think *that* entails B if we are not already wedded to the doctrine that its first disjunct does. The classical reasoning to obtain B from $(A \wedge \neg A) \vee (B \wedge \neg A)$ is that the first disjunct $A \wedge \neg A$ just cannot be the case and so can be ignored, leaving the second, from which B obviously follows. But if we are in a reasoning situation in which inconsistency is a serious possibility and in which we do not take $A \wedge \neg A$ to entail B , it simply is not true that the first disjunct can be ignored. In particular, in the Lewis argument, the classical thought is that because of the $\neg A$, the $A \vee B$ can't come from A , so it must come from B —but this is just plain wrong, because the $A \vee B$ did come from A , whatever the other premise says.

What, then, of the desideratum of staying as close as possible to classical logic? Is **fde** not throwing away the baby with the bathwater, retreating from well-motivated and useful logical principles just to be able to entertain theories of a kind nobody really wants to regard as first-class inhabitants of logical space?

There appear to be three reasons for wanting a paraconsistent logic to retain DS in spite of the *prima facie* case against it. The first is that it is essential to keeping the logic “almost classical”. The second is that it is needed: that reasoning would be hamstrung without it. The third, which is perhaps the most persuasive in practice but also the least defensible is that it is *just obviously good reasoning*. This last reason is also puzzling. It rests, presumably, on examples like the one above (‘Someone has eaten the last cookie. . .’). Yet we have seen that in the paraconsistent context, DS is equally obviously *not* good reasoning, so how is this conflict of “obvious” intuitions to be resolved? A first step is to note that too hasty generalisation from a meagre set of examples is the enemy of logical good sense. It may be true that we regard the reasoning in the above example as rational, but it by no means follows that whatever plausible formalisation we make of it may be applied with equal rationality to quite different types of reasoning. We shall return below to this question of what it is to apply locally inferential principles which are globally invalid, and in what sense it may be rational to do so.

The argument that DS is part of keeping as much as possible of classical logic is unconvincing. There is a particular reason why the classical paradigm is not a suitable goal for paraconsistent logic: classically, negation is *the* central connective, whereas in paraconsistent logic it occupies a less important place. The mainspring of classical inference is the absolute intolerance of inconsistency, whereas a guiding principle of (monotonic) paraconsistent logic is that it is possible to come to rest on a contradiction without the collapse of all rationality. In a sense, the paradigm classical inference form is resolution: inference is driven by the existence of a clash (a contradiction) and only by that, and what in-

ference does is to remove the clash leaving whatever parametric literals happen to be around. In paraconsistent logic, where even if there is nothing else around p and $\neg p$ together do not necessarily call for any action, resolution as classically construed⁸ is an unmotivated style of reasoning. If paraconsistent logic is worth anything, it is a decent theory of reasoning in inconsistent theories whereas classical logic furnishes no such thing. And it is not a virtue of a good theory that it stays as close as possible to a bad one.

The remaining plea for DS, and for resolution more generally, is the pragmatic one that reason cannot get by without it. This may be an important consideration, though opponents of the classical paradigm should not concede the point until it has been established. Naturally, whether resolution is needed depends on the alternatives to it, and the best view of that at present is that not enough is known to enable it to be settled.

4 First Degree Entailment and Beyond

The relevant abandonment of DS, and of the classical tenet that inference is driven by the need to avoid contradictions, is not an arbitrary choice of response to the Lewis argument and the like, but is an outcome of a systematic account of logic. This account indeed stays within the orthodox logical tradition, but at a deeper level than just maintaining certain individual theorems. It is like classical logic and the other mainstream systems in that it allows unlimited nesting of connectives, validates cut, is monotonic everywhere and admits semantic treatment in terms of truth-preservation at worlds. However, the fragmentary logic **fde** which emerges from relevant logic is capable of being examined on its own terms as a medium for paraconsistent reasoning, and then of being re-extended to a full system which may eventually fit that application better than the parent relevant logic.

The best explanation of the potential value of **fde** in computer science applications comes by comparing it with a reasonable alternative in the form of a “vector of values” system. The domain is software specification, and in particular the fusion of partial specifications gleaned from several experts or users. On a specific question such as whether a transition is possible from state s_1 to state s_2 , each expert may either assert ‘yes’, assert ‘no’ or be agnostic. This makes it natural to represent the expert’s opinions in a three-valued fashion, where the values stand for ‘true’, ‘false’ and ‘unknown’. It is quite standard to adopt the usual three-valued matrices, associated with Kleene, to extend the trivalent valuation to compound formulae. If there are several experts, then instead of obtaining just one of the three values, we obtain a vector, one value per expert. The logical operations such as conjunction, disjunction and negation extend in a pointwise fashion to operations on such vectors

⁸ There are other construals of resolution, as a form of cut rule, which is of course fine from a relevant perspective as from many others. The complaint is only about the view of it as resolving a clash, which makes sense only in a theory where clashes need to be resolved.

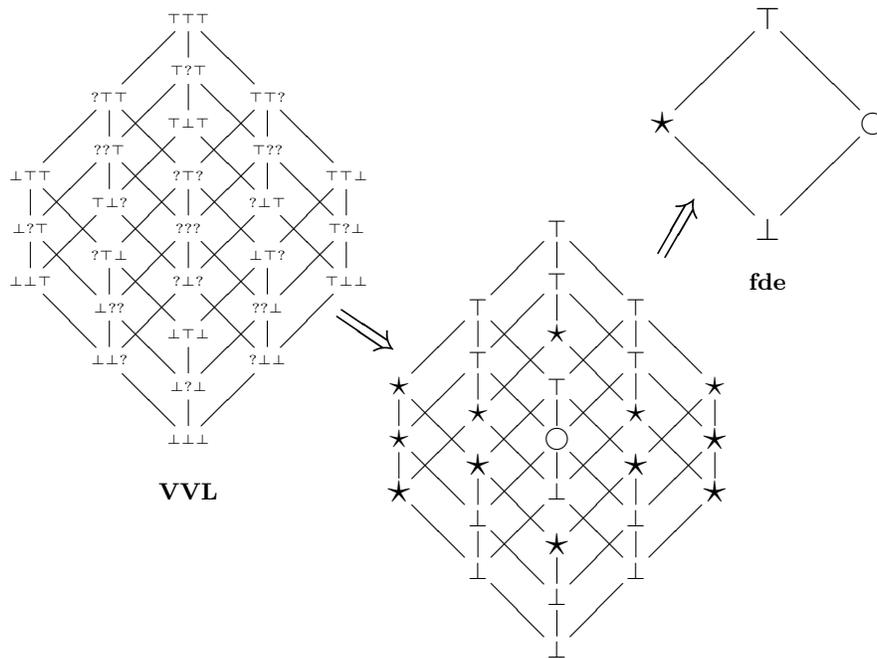


Fig. 1. The four values of **fde** as an abstraction of the twenty-seven vector values of three experts. The star represents ‘confused’ and the circle represents ‘no value’. Note that with the trivial exception of ‘no value’, the preimages of the **fde** values are not sublattices of the **VVL** values.

of values, permitting a semantic representation of the situation in which partial and possibly conflicting pieces of information are combined.

It may be suspected that since epistemic values are in question here, a multi-agent epistemic logic along modal lines could be a better option, but in fact in the attempt to build a theory as to the truth of the matter, rather than one as to the beliefs of the experts, many-valued approaches seem to perform not too badly. Let us therefore consider some advantages and shortcomings of **fde** as opposed to the vector approach. A first striking difference is that **fde** remains four-valued irrespective of the number of experts, whereas the vector-valued logic (henceforth **VVL**) has 3^k values where there are k experts. This extra complexity of **VVL** is more apparent than real, however, because the theories in the vector do not interact: the value (true, false or unknown) in each vector position can be calculated independently of the values in other positions, so the calculation of the value of a compound is no worse than linear in the number of experts (and in the length of the formula, of course). Still, for **fde** it is linear in just the formula length.

In a sense, the four values of **fde** are an abstraction from those of **VVL**. In passing from **VVL** to **fde**, information is lost: the information as

to who said what. All that remains is whether at least one expert said ‘true’, and whether at least one said ‘false’. The notion of the relationship between the logics as one of abstraction must, however, be treated with care, because the function taking a vector to the corresponding “abstracted” value is not a homomorphism between the two algebras (Figure 1). That is, they go on to treat compounds in very different ways.

Consider a simple case in which two experts a and b disagree about proposition p . a says p and b says $\neg p$. Suppose neither of the experts has any opinion about q . What should we conclude about the compounds $p \wedge q$ and $p \vee q$? On the **VVL** approach, $p \wedge q$ is false for expert b because p is, and it is unknown for expert a because for a it stands or falls with the unknown q . So it has value $\langle ?, \perp \rangle$. Similarly, $p \vee q$ gets value $\langle \top, ? \rangle$. In **fde** however, the conjunction and disjunction come out simply false and simply true respectively. I have seen incredulity at this: if we are confused about p , how can disjoining it with something about which we have no information at all remove the confusion and leave pure truth? We shall return to this response below, but for now note that there is a reasonable story to tell in reply: consider $p \vee q$; we have been told that it is true, since we have been told that p by someone who is an expert on p (we have also been told $\neg p$, but never mind). Nobody, however, has told us that the disjunction is false, because nobody has any evidence against q . Therefore, the only truth value we have for $p \vee q$ is ‘true’.⁹

Consider another case, in which **fde** appears closer to the intuitively correct view than **VVL**. Again the experts disagree about p , a saying p and b saying $\neg p$. This time, however, they also disagree about q : a says $\neg q$ and b says q . What now of $p \wedge q$ and $p \vee q$? On the **VVL** approach, the disjunction is true and the conjunction false, because the two experts agree on that much although their reasons for the compound assertions are completely opposite. On the **fde** approach, however, we mark $p \wedge q$ as false, because an expert has said that p is false (and for good measure another expert has said that q is false), but we also mark it as true for the very good reason that we have expert testimony that p is true and also that q is true. Therefore we take the experts in combination to be confused, or in disagreement, about $p \wedge q$, and similarly about $p \vee q$. This is surely right: it is quite possible that a is the dominant expert about p and b about q , in which case it is quite correct to mark $p \wedge q$ as true despite their opinions to the contrary. Each of *them* has a reason to regard the conjunction as simply false and the disjunction as simply true, and of course we have these reasons as well, but we also have, as they do not, sufficient expert testimony to regard the conjunction as true and the disjunction false.

At the level of representation of simple propositions, therefore, **fde** compares reasonably well with certain other natural approaches to paraconsistent theory-building. It is worth pausing to note some more features of the logic. Importantly, there is a sense in which classical resolution-

⁹ We also have ‘unknown’, but this is better seen as the lack of a truth value than a third value in the same sense as the other two. On the **fde** story, at any rate, each proposition has a set of truth values, the values in the sets being just the classical ‘true’ and ‘false’.

based reasoning can be reconstructed in **fde**, despite the invalidity of resolution as such. For this purpose it is convenient to enrich the logic slightly by adding the sentential constants **t** (true) and **f** (false). Intuitively, **t** is the infinite conjunction of all tautologies: we may think of it as $\bigwedge_i (A_i \vee \neg A_i)$. Dually, we may think of **f**, the minimally contradictory proposition, as $\bigvee_i (A_i \wedge \neg A_i)$. Now any resolution inference of the form

$$\frac{A \vee B \qquad \neg B \vee C}{A \vee C}$$

may be reproduced in **fde** with the addition of the false constant:

$$\frac{A \vee B \qquad \neg B \vee C}{A \vee C \vee \mathbf{f}}$$

Hence where there is classically a derivation of the empty clause from a set of clauses, in **fde** there is an analogous resolution derivation of **f** from the same set of clauses. More generally, any classical resolution derivation of any formula A has a corresponding **fde** derivation of $A \vee \mathbf{f}$. Of course, in **fde**, $A \vee \mathbf{f}$ does not imply A , because we cannot generally overlook the possibility that the situation in which we are reasoning is itself inconsistent and contains **f**. However, for showing the inconsistency of a set of clauses by resolution, **fde** lacks nothing in comparison with classical logic, and for deriving an arbitrary conclusion, it is as good provided we are prepared to tack the precautionary "... or I contradict myself" onto the conclusion.

fde is, however, inadequate for all but the most basic knowledge representation purposes. The reason is that it cannot express generality. It lacks quantifiers. Without the means to say that all men are mortal, all tigers are carnivores, all footballers are bipeds and so forth, there is *no hope* of serving the essential purpose of representing lawlike conditions or knowledge about relationships between sorts. Of course, it can easily be equipped with the familiar ' $\forall x$ ' and ' $\exists x$ ' with the obvious semantics, but this hardly helps. The problem is that it is not sufficient that the language contains some particles that *look* like quantifiers: to function as quantifiers, they have to validate the right inference forms. Just as an arrow is not an implication connective unless it satisfies a rule of detachment, so a variable-binding operator is not a universal quantifier unless it features appropriately in the passages of inference:

- (a) Let ABC be a triangle;
 then ... *(some reasoning)* ... ABC has an acute angle;
 therefore every triangle has an acute angle.
- (b) All footballers are bipeds;
 Socrates is a footballer;
 therefore Socrates is a biped.

These principles for the introduction and elimination of universal quantifiers are central to the logic of generality, and have nothing to do with the presence or absence of given structural rules or with attitudes towards inconsistency. In order to formalise such reasoning, a logic must contain, or have a way of securing, quantifiers as binary operators on formulae. It is not enough to be able to express 'Everything is a biped':

there must be a way to say that every *footballer* is a biped. Classically, of course, ‘Everything is either a biped or else not a footballer’ will suffice, but in a weaker logic such as **fde** it will not because it does not validate argument (b) above.

The effect of introducing a binary universal quantifier is to add an implication operator to the logic. If there is an implication \rightarrow in the language, the unary quantifier produces a binary one by the usual move of parsing ‘All A are B ’ as $\forall x(A \rightarrow B)$. Conversely, if there is a suitable binary universal quantifier $(\forall x : A)B$ expressing ‘All A are B ’, it can be used to define $A \rightarrow B$ neatly, if a little artificially, as $(\forall y : A)B$ where y is a variable not occurring free in either A or B . In previous work on this subject [27] it was suggested that if **fde** is to be equipped with universal and existential quantifiers in the most basic way, without stepping outside the truth functional part of the logic, the semantic conditions for these should be as follows. The notion of satisfaction (truth under an assignment to variables) has to be accompanied by a dual notion of dissatisfaction (falseness under assignment to variables) in the obvious way. Then:

1. $(\forall x : A)B$ is satisfied by a valuation v iff B is satisfied by all x -variants of v that satisfy A and A is dissatisfied by all x -variants of v that dissatisfy B .
2. $(\forall x : A)B$ is dissatisfied by a valuation v iff for some x -variant v' of v , A is satisfied by v' and B is dissatisfied by v' .
3. $(\exists x : A)B$ is satisfied by a valuation v iff some x -variant of v satisfies both A and B .
4. $(\exists x : A)B$ is dissatisfied by a valuation v iff every x -variant of v either dissatisfies A or dissatisfies B .

This gives the implication connective the matrix:

\rightarrow	\top	\circ	\star	\perp
\top	\top	\circ	\perp	\perp
\circ	\top	\top	\circ	\circ
\star	\top	\circ	\star	\perp
\perp	\top	\top	\top	\top

The valid formulae are those which always take values \top or \star (if they were required to take only \top then even $A \rightarrow A$ would not be valid). Now an interesting thing has happened, for the logic **BN4** with this implication matrix does not contain the relevant logic **R**. It rejects the structural rule of contraction (W-multiset) since $\circ \rightarrow (\circ \rightarrow \perp)$ evaluates to \top and so does not imply $\circ \rightarrow \perp$ which evaluates to \circ . In fact, the four-valued structure contains the three-valued logic of Łukasiewicz as the subalgebra on $\{\top, \circ, \perp\}$ and also the (unique) three-valued matrix for **R**, as the subalgebra on $\{\top, \star, \perp\}$. We may note that *no* connective definable on the four values of **fde** is an implication in the sense of **R**.

R accommodates paraconsistency without strain, and has **fde** as its truth functional fragment, but it does not fit the interpretation of \circ as the lack of a truth value. As already noted, **R** has the theorem scheme $A \vee \neg A$ which requires that in any **R** model either A or $\neg A$ is true at the base

world 0. Hence although there can be worlds in which A has no truth value, the real world cannot be one of them. This is not necessarily fatal to the advertised use of **fde** and its implicative extensions, to account for fusion of theory fragments, since after all even if the experts' knowledge leaves gaps, we may reasonably suppose that reality does not have gaps. However, it does suggest that we might do well to examine alternatives to the **R** theory of implication in the hope of finding a plausible logic with **fde** as its extensional fragment that can have as a model the partial and inconsistent theory resulting from amalgamating expert opinions. While there are many options for enhancing **fde** with an implication in the relevant family, one particularly attractive suggestion [25] is the paraconsistent version of Nelson's logic [19] of constructible falsity, called **NP** in [14]. A frame for this system is a set of information states, partially ordered by inclusion. As in the models of **fde**, truth and falsehood are independently assigned at states in the frame, subject to the heredity condition that *both* truth and falsehood are preserved under the inclusion order. This gives rise to two modelling relations \models^+ (makes true) and \models^- (does not make false). The semantics of conjunction, disjunction and negation at each state are as in **fde** while implication is evaluated:

$$\begin{aligned}
w \models^+ A \rightarrow B & \quad \text{iff} \quad \text{for every } x \text{ such that } w \subseteq x, \\
& \quad \text{(i) if } x \models^+ A \text{ then } x \models^+ B; \\
& \quad \text{(ii) if } x \models^- A \text{ then } x \models^- B. \\
w \models^- A \rightarrow B & \quad \text{iff} \quad w \not\models^+ A \text{ or } w \models^- B.
\end{aligned}$$

Evidently, the four-valued matrix of **BN4** is the special case in which there is only one information state in the frame, so this logic is a refinement of the "truth functional" implication most naturally associated with **fde**.

NP is a strong logic, an extension of linear logic with a distinctly intuitionist flavour. It does not validate contraction, but comes as close to it as possible, validating the structural rule

$$\frac{\Gamma; \Gamma; \Delta \vdash A \quad \Gamma; \Delta; \Delta \vdash A}{\Gamma; \Delta \vdash A}$$

This is not the place to go into a detailed account of constructible falsity. The interested reader would do well to start with [14] for a readable account and entry to the literature. **NP** has been noted not only half a century ago by Nelson and others, but more recently by a number of writers [1, 10, 11, 30] who see it as useful especially in the context of logic programming. What is worth noting is that it represents, at least arguably, an advance on **R** for the purposes of paraconsistent reasoning such as occurs in merging databases or system specifications. In the first place, it has the four-valued characteristic matrix of **BN4** as a model, and in the second place it is decidable in polynomial space, unlike **R** which is undecidable [28] and whose decidable fragments such as the pure implication fragment tend to have EXPSPACE-hard decision problems [29].

5 Finally: The Disjunctive Syllogism Again

Here we are thinking of the logics of constructible falsity as substructural systems related to the relevant logics, rather than in the more usual way as intuitionist logic with a “strong” negation. They wear both aspects, of course. Like **R** and the other relevant logics, they have **fde** as their fragment of entailments between extensional formulae, and therefore do not validate DS or resolution. The remaining task for the present paper is to revisit DS and consider the status of reasoning in that way in the framework of paraconsistent logics such as **R** or **NP**.

As noted in the opening section, logics which do not validate DS nonetheless do not prohibit its use in reasoning. They offer no guarantee that such reasoning will never go awry, and indeed those of their models which show the disputed principle to be invalid also show *how* it is unreliable. They provide examples of the circumstances in which it fails. In the case of DS, these examples are the obvious ones of inconsistent theories which should be regarded as non-trivial.

This also points to a set of circumstances in which resolution and DS are rational principles to use: those circumstances in which there is no threat of inconsistency (even in counter-factual suppositions) or at least in which no inconsistent state of information can possibly be of any interest. We must note carefully that taking ourselves to be in such circumstances is not a matter of making an assumption of consistency: an extra assumption cannot make an inconsistent theory consistent, even if the extra assumption is “...and this is consistent”. Rather it is a methodological decision to regard any contradiction as rendering the reasoning state absolutely useless. As already observed,¹⁰ resolution derivations can be copied inside paraconsistent logics which contain **fde** provided ‘... $\vee f$ ’ is tacked onto every conclusion. We can choose to disregard the caveat, inferring A from $A \vee f$, if we wish, provided we do not care about the lack of a first-class logical guarantee for the move.

There is a class of reasoning situations in which an invalid rule such as DS may be applied with more logical backing, namely in making deductions in theories in which the rule, though not derivable, is provably admissible. This is general: admissibility is all that is required, though of course derivability is the most direct argument for admissibility no matter what the logic. It is worth rehearsing the commonest technique for showing the admissibility of DS in theories based on relevant logics such as **R** since this is independently interesting and widely applicable (though hardly new, having been around for 35 years or so [2]).

Consider some logic **L** in the relevant family. By a *theory* we mean a set θ of formulae closed under **L** entailment in the sense that where $A_1, \dots, A_n \vdash B$ in **L** if $\{A_1, \dots, A_n\} \subseteq \theta$ then $B \in \theta$. We say that θ is *prime* if wherever it contains a disjunction $A \vee B$ it contains either A or B , that it is *normal* if it is prime and consistent, and that it is *regular* if it contains all the theorems of **L**. The key to showing that θ is closed under resolution is to show that θ is the intersection of its normal

¹⁰ This observation has been made frequently enough before, by R. Meyer and D. Batens among others, so no claims of originality are made for it here.

supertheories. For classical logic, this amounts to Lindenbaum’s lemma (θ is the intersection of its maximal consistent supertheories) but for **L** it must be remembered that maximal theories are not in general normal, and that the lemma is not true of arbitrary θ —it is a rather special feature that must be proved again for each individual case. Clearly, if $A \vee B$ and $\neg A \vee C$ are both in θ , then every normal supertheory of θ contains $B \vee C$, because it either contains A or contains B and either contains $\neg A$ or contains C , and it does not contain both A and $\neg A$, so if θ is the intersection of such theories, it too contains $B \vee C$. Adding unification to resolution for the purposes of this observation is merely a technical detail.

The standard procedure is to construct normal extensions of θ by metavaluation, a technique dating back at least to the 1950s and Harrop’s work on intuitionist logic. Where N is a nontheorem of θ , there is by Zorn’s lemma a maximal θ' extending θ while excluding N (with appropriate machinery to deal with existentials as always in completeness proofs). Now set \mathcal{T} to agree with θ' on atomic formulae, define it to contain $A \wedge B$ iff it contains both A and B , $A \vee B$ if it contains A or contains B , and $\exists xA$ iff it contains A_x^t for some term t free for x in A . $\neg A \in \mathcal{T}$ iff both $A \notin \mathcal{T}$ and $\neg A \in \theta'$. Similarly, $A \rightarrow B \in \mathcal{T}$ iff both $A \rightarrow B \in \theta'$ and if $A \in \mathcal{T}$ then $B \in \mathcal{T}$. Finally, $\forall xA \in \mathcal{T}$ iff $A_x^t \in \mathcal{T}$ for every t free for x in A and $\forall xA \in \theta'$. It remains to show that \mathcal{T} is normal (which is trivial) and that $\theta \subseteq \mathcal{T} \subseteq \theta'$ (which is not trivial). The hard part of the proof is usually to show that all the axioms of θ are in \mathcal{T} .

Many piecemeal results by metavaluation have been established, including crucially the cases in which θ is the set of theorems of a logic such as **R** or **E**. Some special theories have also been shown to admit DS, including the infinitary arithmetic **R**^{##}, and some others have been shown *not* to admit DS, including the relevant Peano arithmetic **R**[#] [16]. The technique has also been elaborated to deal with contraction-free logics in the family such as **C** and **NP** [24] but there remain serious limits to what has been done. Most annoyingly, the technique has been restricted to regular theories—this is natural, since in regular theories the closure conditions are just detachment and adjunction, the logical theorems supplying the rest—but many irregular theories are closed under resolution too, and it would be satisfying to have a routine way of applying metavaluations to them.

6 Conclusion

Relevant logic, considered as a family of systems rather than one specific theory, presents a coherent approach to paraconsistent reasoning while remaining within the mainstream logical tradition in many ways. This is not to deny a place to nonmonotonic reasoning as a way of restoring consistency in flawed theories, nor to soft constraint solving as an approach to overconstrained problems. It is merely to offer a logical point of view from which inconsistent theories may be admitted as first class logical citizens without putting an end to critical rationality.

An examination of one common source of resistance to the relevant approach, the supposed plausibility of resolution and its special case the disjunctive syllogism, suggests that these are in fact poorly motivated in inconsistent contexts, and that there is no particular virtue in trying to maximise preservation of them. It does, however, open the issue of when these inference forms can be used responsibly. To that question there is as yet no completely satisfactory answer.

What is clear, however, is that a thoroughgoing paraconsistency, seen as logical reasoning of the plain deductive variety rather than as a process of reconciling conflicting but individually consistent theories, is an option and is in tune with an account of logic that stands on its own terms, rather than as paraconsistent superstructure on a classical foundation.

References

1. S. Akama. Tableaux for Logic Programming with Strong Negation. *Proceedings of the Conference on Automated Reasoning with Analytic Tableaux and Related Methods (TABLEAUX'97)*, 1997, 31–42.
2. A. R. Anderson and N. D. Belnap. *Entailment: The Logic of Relevance and Necessity, Vol 1*. Princeton University Press, Princeton, 1975.
3. A. R. Anderson, N. D. Belnap and J. M. Dunn. *Entailment: The Logic of Relevance and Necessity, Vol 2*. Princeton University Press, Princeton, 1992.
4. N. D. Belnap. A Useful Four-Valued Logic. Dunn and Epstein (eds), *Modern Uses of Multiple-Valued Logics*, Reidel, Dordrecht, 1977: 8–37.
5. N. D. Belnap. Display Logic. *Journal of Philosophical Logic* 11 (1982): 375–417.
6. R. Brady. *Universal Logic*. Cambridge University Press, Cambridge, 2001.
7. J. M. Dunn. Algebraic Completeness Results for R-mingle and its Extensions. *Journal of Symbolic Logic* 35 (1970): 1–13. Reprinted in [2].
8. J. M. Dunn. A ‘Gentzen’ System for Positive Relevant Implication. *Journal of Symbolic Logic* 38 (1974): 356–357 (abstract). Reprinted in [2].
9. J. M. Dunn. Relevance Logic and Entailment. in D. Gabbay and F. Günthner (eds) *Handbook of Philosophical Logic* Vol. 3, Reidel, Dordrecht, 1986: 117–229.
10. T. Eiter, N. Leone and D. Pearce. Assumption Sets for Extended Logic Programs. *JFAK. Essays dedicated to Johan van Benthem on the occasion of his 50th birthday*. Amsterdam University Press, 1999. <http://www.illc.uva.nl/j50/contribs/pearce/>.
11. M. Gelfond. Representing Knowledge in A-Prolog. *Computational Logic* 2408: ‘Logic Programming and Beyond: Essays in honour of Robert A. Kowalski’ (2002): 413–451.
12. S. Giambrone and A. Urquhart. Proof theories for Semilattice Relevant Logics. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik* 33 (1987: 301–304.

13. J.-Y. Girard. Linear Logic. *Theoretical Computer Science* 50, 1987: 1–101.
14. I. Hasuo and R. Kashima. A Proof-Theoretical Study on Logics with Constructible Falsity. Report C-165, research Reports on Mathematical and Computing Sciences, Tokyo Institute of Technology, 2003, <http://www.is.titech.ac.jp/research/research-report/C/>
15. J. Lukasiewicz. *Selected Works* (ed. L. Borkowski), North-Holland, Amsterdam, 1970.
16. R. K. Meyer and H. Friedman. Whither Relevant Arithmetic? *Journal of Symbolic Logic* 57 (1992): 824–831.
17. R. K. Meyer and E. P. Martin. Logic on the Australian Plan. *Journal of Philosophical Logic* 15 (1986): 305–332.
18. R. K. Meyer and R. Routley, Classical Relevant Logics, I and II. *Studia Logica* 32 (1973): 51–66 and 33 (1973): 183–194.
19. D. Nelson. Constructible Falsity. *Journal of Symbolic Logic* 14 (1949): 16–26.
20. D. Pym. *The Semantics and Proof Theory of the Logic of Bunched Implications*. Kluwer, Dordrecht, 2002.
21. Relevant and Substructural Logics. in D. Gabbay and J. Woods (eds) *Handbook of the History and Philosophy of Logic* forthcoming.
22. R. Routley and R. Meyer. Semantics of Entailment. in H. Leblanc (ed) *Truth, Syntax, Modality*, North Holland, 1973: 194–243.
23. R. Routley, V. Plumwood, R. Meyer and R. Brady. *Relevant Logics and their Rivals*. Ridgeview, Atascadero CA, 1982.
24. J. Slaney. Reduced Models for Relevant Logics Without WI. *Notre Dame Journal of Formal Logic* 28 (1987): 395–407.
25. J. Slaney. A General Logic. *Australasian Journal of Philosophy* 68 (1990): 74–88.
26. J. Slaney and R. Meyer. Logic for Two: The Semantics of Distributive Substructural Logics. *Proceedings of the Conference on Qualitative and Quantitative Practical Reasoning (ECSQARU-FAPR)* (1997): 554–567.
27. J. Slaney. The Implications of Paraconsistency. *Proceedings of the 12th International Joint Conference on Artificial Intelligence* (1991): 1052–1057.
28. A. Urquhart. The Undecidability of Entailment and Relevant Implication. *Journal of Symbolic Logic* 49 (1984): 1059–1073.
29. The Complexity of Decision Procedures in Relevance Logic. J. M. Dunn and A. Gupta (ed), *Truth or Consequences: Essays in Honour of Nuel Belnap*, Kluwer, Dordrecht, 1990: 77–95.
30. H. Wansing. *The Logic of Information Structures (LNAI 681)*. Springer-Verlag, Berlin, 1993.