Significant Inferences: Preliminary Report

Philippe Besnard
IRIT–CNRS
118 route de Narbonne
F–31062 Toulouse Cedex
besnard@irit.fr

Torsten Schaub
Institut für Informatik
Universität Potsdam
Postfach 60 15 53, D–14415 Potsdam
torsten@cs.uni-potsdam.de

Abstract

We explore the possibility of a logic where a conclusion substantially improves over its premise(s): Specifically, we intend to rule out inference steps such that the premise conveys more information, in a simpler form, than the conclusion does.

In fact, most reasoning formalisms, among them classical logic, come with means for generating disjunctive or conditional information in a fairly arbitrary way.

The basic principle for drawing disjunctive information is disjunctive weakening, which allows for deriving $\varphi \vee \psi$ from φ (for any ψ). Thus, given that "Nancy is married to Ron", disjunctive weakening makes us infer that "Nancy is married to Ron or Monica is married to Bill". Although the latter propositions may still be seemingly related, one should not forget (1) that any arbitrary proposition can serve as the additional disjunct, eg. "Nancy is married to Ron or the Queen of England is bald", and (2) that this can be iterated so that real information is buried in the generated disjunction among irrelevant propositions. What is the point in inferring such disjunctive formulas?

Similar phenomena can be traced back to conditionalization, which allows for deriving $\psi \rightarrow \varphi$ from φ (for any ψ).

As a result, we propose a natural deduction system along the intuitions sketched above.

1 MOTIVATION

Perhaps the most salient feature of reasoning is to make explicit what is only implicit. Accordingly, the less obvious a (correct) conclusion is, the more valuable is any reasoning by which that conclusion is drawn. In a sense, some conclusions may then not be worth inferring. Such a point of view is latent in relevance logics [Anderson and Belnap,1975] because they reject certain conclusions that weaken some premise in a peculiar way, as happens with arbitrary conditionalization for instance, i.e., $p \vdash q \rightarrow p$ (where p and q are propositional symbols). In this respect, relevance logics pave the way to a logic where only conclusions worth expressing are drawn. However, relevance logics stop short of fully achieving this idea. This can be viewed from relevantly valid schemes such as $p \vdash p \vee q$. Indeed, one may regard $p \vee q$ as a dubious weakening of p, as much as $q \rightarrow p$ is. In fact, we regard the scheme $p \vdash p \vee q$ as drawing a conclusion which fails to be significant *in view of* its premise. What do we mean by significant here? We mean that the conclusion is not worth stating when its premise is stated: Once it is clear that "the winning ticket is number 36", there is no point in making the inquiry whether "the winning ticket is number 17 or 36". Much as it would make no sense to investigate iteratively about the fact that "the winning ticket is number n_1 or…or n_k or 36" for k increasing. It does not even matter that any n_i be 36 as well: What is the point of stating that "the winning ticket is number 36 or the winning ticket is number 36" when it has already been stated that "the winning ticket is number 36"? That is, even the restricted form $p \vdash p \vee p$ is a particular form of dubious weakening.

Hence, we want to explore the possibility of a logic where a conclusion substantially improves over its premise(s): Specifically, we intend to rule out inference steps such that the premise conveys more information, in a simpler form, than the conclusion does.

Indeed, we are not aiming at ignoring a conclusion that is less informative than its premises, provided that the conclusion has a simpler form. For instance,

$$p \vee q, p \rightarrow \neg r, s \leftrightarrow r, \neg p \rightarrow (q \rightarrow \neg s) \vdash \neg s$$

is an inference whose conclusion $\neg s$ does not exhaust all the information about p, q, r, s that the premises provide but is much easier to grasp.

2 TOWARDS SIGNIFICANT INFERENCE

In contrast to a proof system such as resolution [Robinson,1965] that has a single inference rule (ignoring factorization), natural deduction [Gentzen,1935] is traditionally a suitable framework for analyzing inferences. So let us consider the matter of *conditionalization* and *disjunctive weakening* from the perspective of natural deduction.

Conditionalization consists of turning an inference of ψ from some premises including φ into a proof of $\varphi \rightarrow \psi$. In extensional logics, the notion of being a premise is fairly liberal so that it is not required that φ actually serves for deriving ψ. Intensional logics such as relevance logics insist on φ being actually used to infer ψ: Relevance is then a necessary condition for $\varphi \rightarrow \psi$ to be derived. We adopt the same criterion because it matches also our idea of significance: on the one hand, $\varphi \rightarrow \psi$ is certainly significant whenever ψ is derived by means of φ, and, on the other hand, if ψ is derived independently of φ, then $\varphi \rightarrow \psi$ is both less simple and less insightful than ψ itself. In natural deduction, generalizing this leads us to requiring that the so-called "auxiliary assumptions" have to take part in the derivation of the conclusion that is declared to rely on them.

Disjunctive weakening consists of concluding $\varphi \vee \psi$ from φ (or similarly from ψ). We have argued above that this is never justified on its own, although it is sometimes useful as a device for inferring intermediate conclusions, as needed for reasoning by cases (whose principle is that, given $\varphi \vee \psi$, if χ is concluded from φ and if χ is concluded from ψ, then χ is inferred). So, weakening is needed for combining the conclusions obtained in each case:

$$p \vee q, p \rightarrow r, q \rightarrow s \vdash r \vee s \ .$$

In fact, we permit disjunctive weakening only for deriving conclusions drawn by reasoning by cases (see below).

Technically, banning disjunctive weakening while restricting conditionalization as just indicated could simply yield a subsystem of a relevance logic. However, we ban disjunctive weakening according to our intuitions about significant conclusions and these intuitions are different from those behind relevance logics, in particular, they make us depart from the relevance view against disjunctive syllogism (viz. $\varphi \vee \psi, \neg\varphi \vdash \psi$).

A well-taken objection by the relevantists is that if $\varphi \vee \psi$ holds because φ does then applying disjunctive syllogism is flawed: There is a contradiction between φ and $\neg\varphi$ but ψ is irrelevant in the matter. We agree that disjunctive syllogism is inappropriate in such a case but we contend that it is a valuable pattern when no contradiction is involved. Indeed, our intuitions about $\varphi \vee \psi$ are that the disjunction is really an alternative between φ and ψ so that it actually is about φ *as well as* about ψ. Accordingly, it is not about just φ or about just ψ and this warrants that if either case is denied then the other must hold.

As appears from the preceding discussion, significant inference is paraconsistent (i.e., the so-called *ex falso* $\varphi, \neg\varphi \vdash \psi$ does not hold). The intuitive notion of a significant conclusion appeals for paraconsistency in at least two different ways. One is that nothing is more informative or substantially simpler than $\varphi \wedge \neg\varphi$. Another is that, should paraconsistency be ruled out, then everything would be concluded from a contradiction (meaning that everything is significant, which is antinomic).

The issue of paraconsistency naturally leads us to that of analyzing the standard deduction of the ex falso:

$$\cfrac{\cfrac{\varphi}{\varphi \vee \psi}\,{}_1 \quad \neg\varphi}{\psi}\,{}_2$$

Relevance logics preclude such a deduction because they rule out disjunctive syllogism (step (2)). A different perspective is to rule out disjunctive weakening (step (1)), as we advocate. The last option is to allow for both inference schemes but to preclude transitivity of inference. This is the path taken by Tennant in his most interesting work on entailment [Tennant,1987]. Unfortunately, failure of transitivity has major drawbacks.

Also, it is worth paying attention to the interaction between disjunction and implication embodied by the formula $(\varphi \vee \psi) \rightarrow \chi$. Let us repeat that, for us, $p \vee q$ really is an alternative between p and q so that $p \vee q$ is actually about p as well as about q. For this reason, $(\varphi \lor \psi) \to \chi$ is specifically meant to deduce χ from $\varphi \lor \psi$ whereas $\varphi \to \chi$ and $\psi \to \chi$ serve the same purpose with respect to φ and ψ. Accordingly, $(\varphi \lor \psi) \to \chi$ together with φ, or similarly together with ψ, does not make χ to be deduced. When thinking of it, all this is the more sensible and only is the prejudice attached to unlimited disjunctive weakening making the usual equivalence of $(\varphi \lor \psi) \to \chi$ with the couple of formulas $\varphi \to \chi$ and $\psi \to \chi$ sound all right.

Of course, the notion of significance is intuitive so not everything is clear-cut about significant conclusions. Nonetheless, the above discussion gives us enough constraints to define a first inference system for significant reasoning.

3 NATURAL DEDUCTION

The reader familiar with natural deduction can skip this section.

Throughout the text, we focus on trees in which each node is labeled with a formula. We call them *trees of formulas* (actually, we rather tend to identify a node with its labeling formula).

It is assumed that every link (in fact, hyper-link) relating a parent node to its children is uniquely determined by a so-called *identifier* (which we take to be a natural number).

Given a tree of formulas, we also say that a formula A is an *hypothesis* for a node N iff A is a leaf in the subtree rooted at N. By abuse of language, we say that A is an hypothesis for B when B is the labeling formula of N.

We assume a notational device called *discharge* such that any leaf (in a tree of formulas) can be marked as "discharged". Such a leaf is said to be a discharged formula, as is usual, even though a discharged formula refers to an occurrence: Not all leaves labeled with the same formula need have the same status regarding discharge.

Intuitively, discharged formulas are auxiliary hypotheses that serve to deduce intermediate conclusions whereas the final conclusion does *not* depend on these formulas. In contrast, the final conclusion depends on all non-discharged hypotheses.

As usual, inference rules are used to form trees of formulas corresponding to a deduction. Here, all the inference rules have the following general form:

$$\frac{A_1 \cdots A_n}{C} \qquad \left\{ \begin{array}{l} \text{given } A_1 \text{ and } \cdots \text{ and } A_n\text{, deduce } C \\ \text{possibly subject to some condition(s)} \end{array} \right.$$

where $A_1, \ldots, A_n$ are the *assumptions* of the rule and C is the *consequence* of the rule.

Definition 1 *A derivation* Π *of a formula* C *from a set of formulas* $\mathcal{P}$ *is a finite tree of formulas such that:*

- *The root node of* Π *is* C*.*
- *Each leaf of* Π *is either a discharged formula or a formula in* $\mathcal{P}$ *(or an axiom, if any).*
- *Each node* B *of* Π *has* $A_1, \ldots, A_n$ *as its child nodes only if there exists an inference rule of which* $A_1, \ldots, A_n$ *are the assumptions and* B *the consequence (all constraints, if any, attached to the rule must be met).*
- *Discharge marks in* Π *only occur as per the stipulations* $(\S_n)$ *stated for* $(I \to)$ *and* $(E \lor)$ *(see below).*

We write $\mathcal{P} \vdash C$ *whenever we have such a derivation* Π*.*

In a derivation, every link is thus to be identified with an instance of an inference rule (which in turn is uniquely determined by the identifier of the link under consideration).

At some point, we will need to consider *normal derivations* [Prawitz,1965] where the intuitive meaning of a normal derivation is that no node in it is both an assumption of the elimination rule yielding its parent node and the conclusion of the corresponding introduction rule yielding one of its child nodes:

Formally, a *maximum segment* in a derivation is a sub-branch $\{A_1, \ldots, A_k\}$ (starting with A_1 and ending with A_k) such that

1. A_1 is the consequence of an introduction rule.
2. For each $i < k$, A_i is a minor assumption of a $(E \lor)$ rule.
3. A_k is a major assumption of an elimination rule.

A_k is a *maximum formula* and k is the *length* of the maximum segment.

A closer look at the following inference system reveals that the introduction and elimination rule referred to in 1. and 3. are necessarily applied to the same connective.

Finally, a derivation is *normal* iff it contains no maximum segment.

4 NATURAL DEDUCTION FOR SIGNIFICANT REASONING

We adopt a restricted set of logical symbols: $\perp$ (absurdity), $\vee$ (disjunction), $\rightarrow$ (implication). First of all, there is no negation. Formulas $\varphi \rightarrow \perp$ are then used to overcome the absence of explicit negation. Second, there is no conjunction. Instead, sets of formulas are to be interpreted conjunctively. Then, the disjunction of two sets of formulas $\Phi = \{\varphi_1, \ldots, \varphi_n\}$ and $\Psi = \{\psi_1, \ldots, \psi_m\}$ can be given the form of the set

$$\{\varphi_i \vee \psi_j \mid 1 \leq i \leq n, 1 \leq j \leq m\} \ .$$

And, $\varphi_1 \rightarrow (\ldots \rightarrow (\varphi_n \rightarrow \psi_i) \ldots)$ (for $i = 1..m$) is the way for a series of formula to state that each ψ_i (for $i = 1..m$) follows from $\varphi_1, \ldots, \varphi_n$ taken together (in any order, actually).

In order to avoid heavy notation, we follow the convention that $\vee$ binds stronger than $\rightarrow$.

We adapt a system due to Prawitz [1965] with its notion of a *minor assumption* for an inference rule (it is either of φ, ψ_1, ψ_2 in the list below) and of a *major assumption* for an inference rule (any other assumption).

Our system consists of the following five inference rules:

Introduction rules

$$(I \rightarrow) \ \frac{\psi_0}{\varphi_0 \rightarrow \psi_0} \ (\S_0)$$

$$(I_l \vee) \ \frac{\varphi_1}{\varphi_1 \vee \varphi_2} \ (\dagger) \qquad (I_r \vee) \ \frac{\varphi_2}{\varphi_1 \vee \varphi_2} \ (\dagger)$$

Elimination rules

$$(E \rightarrow) \ \frac{\varphi \qquad \varphi \rightarrow \psi}{\psi}$$

$$(E \vee) \ \frac{\varphi_1 \vee \varphi_2 \qquad \psi_1 \qquad \psi_2}{\psi} \ (\ddagger) \ (\S_1)$$

Simplification rule

$$(E_l \ \vee\perp) \ \frac{\psi \vee \perp}{\psi} \qquad (E_r \ \vee\perp) \ \frac{\perp \vee \psi}{\psi}$$

The symbols ($\dagger$), ($\ddagger$), and ($\S_n$) indicate that the rules are subject to the following provisos:

Proviso **($\dagger$):** It must be the case that $\varphi_1 \vee \varphi_2$ is a minor assumption of $(E \vee)$.

Proviso **($\ddagger$):** It must be the case that $\psi = \psi_1 = \psi_2$. Moreover, if ψ_1 and ψ_2 are conclusions of $(I\vee)$ rules then one of these must be a $(I_l\vee)$ rule and the other must be a $(I_r\vee)$ rule.

Proviso **($\S_n$):** For $i = n..2n$, it must be the case that φ_i is an hypothesis for ψ_i. Then, any occurrence of φ_i as an hypothesis for ψ_i inherits the identifier of the current instance of the inference rule as its discharge mark.

The application of these provisos is discussed in Section 5.

Axioms for associativity of $\vee$ can be introduced freely:

$$((\varphi_0 \vee \varphi_1) \vee \varphi_2) \rightarrow (\varphi_0 \vee (\varphi_1 \vee \varphi_2))$$

and

$$(\varphi_0 \vee (\varphi_1 \vee \varphi_2)) \rightarrow ((\varphi_0 \vee \varphi_1) \vee \varphi_2) \ .$$

A classical extension for the above intuitionistic version can be obtained by adding the axiom for excluded middle: $\varphi \vee (\varphi \rightarrow \perp)$.

Observe that all introduction and elimination rules are standard except for the provisos. However, the above system is unusual in several respects. First, there is a $\perp$-related simplification rule for $\vee$. Second, as can be seen from the examples in Section 5, subtrees in a derivation, when taken in isolation, need not be derivations themselves.

5 EXAMPLES AND MAIN RESULT

An illustrative derivation involving discharge is the following one:

$$\cfrac{\cfrac{\cfrac{{}^{(1)}\varphi \quad \varphi \rightarrow \psi}{\psi} \ {\scriptstyle (E\rightarrow)} \qquad \cfrac{{}^{(1)}\varphi \quad \varphi \rightarrow (\psi \rightarrow \perp)}{\psi \rightarrow \perp} \ {\scriptstyle (E\rightarrow)}}{\perp} \ {\scriptstyle (E\rightarrow)}}{\varphi \rightarrow \perp} \ {\scriptstyle (1)\ (I\rightarrow)}$$

The derivation "starts" with letting $\varphi \to \psi$ and $\varphi \to (\psi \to \bot)$ as well as two occurrences of φ (at first, ignore the superscript in front of these) be hypotheses (for the time being). Now, the rule $(E \to)$ "is applied" to yield ψ on the one hand and $\psi \to \bot$ on the other. Then, $(E \to)$ is applied once more to yield $\bot$. Next, the rule $(I \to)$ is applied to yield $\varphi \to \bot$. All in all, the final conclusion $\varphi \to (\varphi \to \bot)$ depends on the non-discharged hypotheses, which are $\varphi \to \psi$ and $\varphi \to (\psi \to \bot)$. As for notation, we write $\varphi \to \psi, \varphi \to (\psi \to \bot) \vdash \varphi \to (\varphi \to \bot)$.

The fact that the occurrences of φ are discharged when applying the rule $(I \to)$ is indicated by (1) next to the $(I \to)$ bar and by (1) as a superscript in front of the occurrences of φ. That is, (1) is the identifier for that instance of $(I \to)$ and is also the discharge mark for the occurrences of φ.

Still regarding discharge, the proviso for the $(I \to)$ rule discriminates

$$\dfrac{{}^{(1)}\varphi}{\varphi \to \varphi}\ {\scriptstyle (1)\ (I\to)}$$

which is a derivation in our system, from

$$\dfrac{\psi}{\varphi \to \psi}\ {\scriptstyle (I\to)}$$

which is not derivable in our system.

The proviso for the $(I\vee)$ rule lets

$$\dfrac{\psi \vee \varphi \qquad \dfrac{{}^{(1)}\psi}{\varphi \vee \psi}\ {\scriptstyle (I_r\vee)} \qquad \dfrac{{}^{(1)}\varphi}{\varphi \vee \psi}\ {\scriptstyle (I_l\vee)}}{\varphi \vee \psi}\ {\scriptstyle (1)\ (E\vee)}$$

be a derivation in our system, whereas

$$\dfrac{\varphi}{\varphi \vee \psi}\ {\scriptstyle (I_l\vee)}$$

is not derivable in our system.

For the record, here is a derivation of disjunctive syllogism:

$$\dfrac{\dfrac{\varphi \vee \psi \qquad \dfrac{\dfrac{{}^{(1)}\varphi \quad \varphi \to \bot}{\bot}\ {\scriptstyle (E\to)}}{\psi \vee \bot}\ {\scriptstyle (I_r\vee)} \qquad \dfrac{{}^{(1)}\psi}{\psi \vee \bot}\ {\scriptstyle (I_l\vee)}}{\psi \vee \bot}\ {\scriptstyle (1)\ (E\vee)}}{\psi}\ {\scriptstyle (E_l\vee\bot)}$$

Apart from disjunctive syllogism, the following derivable inferences are of interest:

$$\dfrac{\varphi_1 \vee \varphi_2 \qquad \varphi_1 \to \psi_1 \qquad \varphi_2 \to \psi_2}{\psi_1 \vee \psi_2}$$

$$\dfrac{\varphi \to (\psi \to \chi)}{\psi \to (\varphi \to \chi)}$$

$$\dfrac{\varphi \to \psi \qquad \psi \to \chi}{\varphi \to \chi}$$

$$\dfrac{\varphi \to \psi \qquad \psi \to \bot}{\varphi \to \bot}$$

In fact, the last inference is modus tollens. Similarly, one may derive all other forms of reasoning by contraposition.

As common with natural deduction systems, the most fundamental property is that of normalization (because it corresponds to cut-elimination in sequent calculi):

Theorem 1 (Normalization) *Every derivation can be transformed into a normal derivation (with the same premises and conclusion).*

In our case, normalization yields additionally the following salient features:

Corollary 1 (Transitivity) $\varphi \vdash \psi$ *and* $\psi \vdash \chi$ *implies* $\varphi \vdash \chi$ *for all* φ, ψ, χ.

Corollary 2 (Paraconsistency) *For all* φ *there exists* ψ *such that* $\varphi, \varphi \to \bot \nvdash \psi$.

Corollary 2 is usually expressed as $\varphi, \neg\varphi \nvdash \psi$.

Furthermore, observe that

$$\begin{aligned} \varphi \vee \psi, \varphi \to \bot &\ \vdash \psi, \\ \varphi, \varphi \to \bot &\ \nvdash \psi, \\ \varphi, \varphi \vee \psi, \varphi \to \bot &\ \vdash \psi. \end{aligned}$$

6 DISCUSSION

All rules in our system are either restricted forms of rules for classical logic or rules derivable in classical logic. Thus, our system is a subsystem of natural deduction for classical logic. Let us look at some of the inferences that are no longer derivable.

First, tautologies corresponding to conditionalization and disjunctive weakening are no longer derivable in our system:

$$\nvdash \varphi \to (\varphi \vee \psi),$$
$$\nvdash \varphi \to (\psi \to \varphi).$$

The latter comes together with the non-derivability of the classical tautology

$$\nvdash (\varphi \to \psi) \vee (\psi \to \varphi),$$

for which it has then no reason to be supported, since it relies on arbitrary conditionalization.

As was to be expected from the discussion at the end of Section 2,

$$\frac{\varphi_1 \qquad (\varphi_1 \vee \varphi_2) \to \psi}{\psi}$$

is not derivable. Analogously, the following inference is not valid in our system:

$$\frac{\varphi_1 \qquad (\varphi_2 \to \varphi_1) \to \psi}{\psi}$$

Similarly to the case of disjunction, where $(\varphi \vee \psi) \to \chi$ is not related to $\varphi \to \chi$ and $\psi \to \chi$, we must therefore distinguish between $(\psi \to \varphi) \to \chi$ and $\varphi \to \chi$. In accord with the discussion at the end of Section 2, we differentiate between a formula q and a formula $p \to q$ that links information about q with that about p. Therefore, we also distinguish between inferences drawn from q and $p \to q$.

The most surprising case is presumably the failure of

$$\frac{\varphi \to \psi}{\varphi \to (\varphi \to \psi)}$$

but more relaxed technical conditions may turn it into a valid scheme (this is further discussed in the conclusion).

Adapting traditional natural deduction had us make an implicit decision and that is that the system is monotonic although this property did not arise from our discussion about significant reasoning. Consider a tautology such as $p \to p$. It certainly is significant information to conclude when no premise is given. This is no longer the case, should the premise p be given. Does this mean that $p \to p$ must then be withdrawn? We think it would be too strict a principle. Rather, we prefer to consider as significant any conclusion that could be viewed as such for *some* reason (in particular, in light of *part* of the given premises). The philosophy here would rather be that we can dispense with drawing some inferences but not to rule them out altogether.

Our system seems promising whenever it comes to knowledge representation problems involving (some extent of) relevance. Let us illustrate this by considering a classically valid yet counterintuitive inference that has been identified by Stephen Read in [Read,1989].

> "Roy has claimed that John was in Edinburgh on a certain day, and Crispin has denied it."

Now, consider

1. "If John was in Edinburgh, Roy was right."
2. "It is not the case that if Crispin was right, so was Roy."
3. "If John was in Edinburgh, Crispin was right."

The latter sentence is false but the former two are true. This gives us an invalid argument (all its premises are true and its conclusion is false). However, the argument is classically valid because

$$\varphi \to \psi, \neg(\chi \to \psi) \vdash \varphi \to \chi$$

holds in classical logic. To see this, consider the following derivation.

$$\dfrac{\dfrac{\dfrac{(\chi \to \psi) \to \bot \qquad \dfrac{\dfrac{{}^{(3)}\varphi \quad \varphi \to \psi}{\psi}\,{\scriptstyle (E\to)}}{\chi \to \psi}\,{\scriptstyle (1)\ (I\to)}}{\bot}\,{\scriptstyle (E\to)}}{(\chi \to \bot) \to \bot}\,{\scriptstyle (2)\ (I\to)}}{\dfrac{\chi}{\varphi \to \chi}\,{\scriptstyle (3)\ (I\to)}}\,{\scriptstyle (E\neg\neg)}$$

This derivation is invalid in our systems since it violates proviso $(\S_0)$ at (1) $(I \to)$ and (2) $(I \to)$. In a classical natural deduction system the inference

$\varphi \to \psi, \neg(\chi \to \psi) \vdash \varphi \to \chi$ is always established by applying arbitrary conditionalization, which is disallowed in our system as in relevance logics.

However, relevance logics only overlap with the requirements of a notion of significant inference. On the one hand, disjunctive syllogism is valid in our system but is invalid in relevance logics. On the other hand, disjunctive weakening is invalid in our system but is valid in relevance logics (including first-degree entailment [Anderson and Belnap,1975]).

Our system also bears some connections with relatedness [Epstein,1979; Krajewski,1986] but it is closer to Parry's [1989] and even closer to Tennant's [1987]. However, Tennant is strongly concerned with relevance when implication is involved and he ignores the matter when it comes to disjunction. Also, Tennant's system fails transitivity but ours satisfies it (namely, $\varphi \vdash \psi$ and $\psi \vdash \chi$ yields $\varphi \vdash \chi$).

As already mentioned, negation can be introduced in our system by way of the usual implication to absurdity. In fact, if we take

$$\neg\varphi \stackrel{\text{def}}{=} \varphi \to \bot$$

then the usual inferences hold (except that the degenerate case where $\bot$ does not depend on ${}^{(1)}\varphi$ is not allowed here):

$$\dfrac{\begin{matrix}{}^{(1)}\varphi \\ \vdots \\ \bot\end{matrix}}{\varphi \to \bot}\ {\scriptstyle (1)\ (I\to)} \qquad \Longleftrightarrow \qquad \dfrac{\begin{matrix}{}^{(1)}\varphi \\ \vdots \\ \bot\end{matrix}}{\neg\varphi}\ {\scriptstyle (1)\ (I\neg)}$$

and

$$\dfrac{\varphi \to \bot \quad \varphi}{\bot}\ {\scriptstyle (E\to)} \qquad \Longleftrightarrow \qquad \dfrac{\neg\varphi \quad \varphi}{\bot}\ {\scriptstyle (E\neg)}$$

Turning to conjunction,

$$\varphi \wedge \psi \stackrel{\text{def}}{=} ((\varphi \to \bot) \vee (\psi \to \bot)) \to \bot$$

only the usual intuitionistic inference holds, as shown in Figure 1, and the result is not as meaningful as in the case of negation, since the corresponding elimination rule is not derivable in our system. As with intuitionistic logic [Dummett,1977], there is no way of deriving φ from $\neg(\neg\varphi \vee \neg\psi)$. Among others, this is a reason why we advocate modeling conjunctions by appeal to sets, as described at the start of Section 4.

7 CONCLUSION

With significant reasoning, we have elaborated upon a new notion of reasoning that is distinct from existing approaches, although there are close ties to relevance logic and intuitionistic logic.

Our contribution can thus be looked at from two perspectives: First, we have identified and elaborated upon the notion of significant reasoning. And second, we have defined a version of natural deduction for significant reasoning.

Of course, the notion of significance is intuitive so that not everything is clear-cut about significant conclusions. There are several places where another choice could make sense. In fact, our choices were motivated by the strict realization of our intuitions discussed in Section 2. For instance, we could be less strict about the policy for discharging hypotheses so that $\varphi \to \psi$ yields $\varphi \to (\varphi \to \psi)$. (Although this is characteristic for linear logic [Girard,1987], it should be clear that our system has nothing in common with resource logics.) Also, we could consider to get closer to familiar practice so that φ and $(\varphi \vee \psi) \to \chi$ yield χ (although we still disapprove of such an inference).

As discussed in Section 5, our system seems promising whenever it comes to knowledge representation problems involving (some extent of) relevance. As pointed out by one of the anonymous referees, this is of interest for the definition of reasoning capabilities for agents in a multi agent systems.

An important issue of future research consists of elaborating appropriate semantical underpinnings. A promising starting point could be to adapt the semantics of relatedness logic [Epstein,1979]. That is, models would be equipped with relations reflecting a notion of significance among propositions.

A SOME TECHNICALITIES

Let $\mathcal{F}$ be the set of all formulas of the language.

A *tree of formulas* is a triple $\langle I, T, f\rangle$ where

1. I is a finite initial portion of the numerals $\{\overline{1}, \overline{2}, \overline{3}, \ldots\}$ for the natural numbers
2. T is a subset of the free monoid I^* such that:
   (a) if $uv \in T$ then $u \in T$
   (b) if $u\overline{n} \in T$ then $u\overline{m} \in T$ for $m < n$
3. f is a function from T to $\mathcal{F}$

$$\dfrac{\dfrac{^{(2)}((\varphi\to\bot)\vee(\psi\to\bot)) \quad \dfrac{^{(1)}(\varphi\to\bot)\quad\varphi}{\bot} \quad \dfrac{^{(1)}(\psi\to\bot)\quad\psi}{\bot}}{\bot}\,{\scriptstyle (1)\ (E\vee)}}{((\varphi\to\bot)\vee(\psi\to\bot))\to\bot}\,{\scriptstyle (2)\ (I\to)} \quad\iff\quad \dfrac{\varphi\quad\psi}{\varphi\wedge\psi}\,{\scriptstyle (I\wedge)}$$

Figure 1: Derivation of adjunction.

Each $w \in T$ is a *node* of the tree. The tree is *finite* iff it only has finitely many nodes (i.e., T is finite). The empty tree has no node (i.e., $T = \emptyset$). The empty word ε is the *root* of the tree (on condition that the tree is non-empty, $T \neq \emptyset$). A *leaf* of the tree is a node w such that $w\overline{1} \notin T$. If $w \in T$ and $w\overline{n} \in T$ then $w\overline{n}$ is a *child* node of w and w is the *parent* node of $w\overline{n}$.

A *branch* of the tree $\langle I, T, f\rangle$ is any $B \subseteq T$ such that:

1. $\varepsilon \in B$
2. if $u \in B$ and $u\overline{1} \in T$ then there exists exactly one $\overline{n} \in I$ such that $u\overline{n} \in B$

Given a branch B of the tree $\langle I, T, f\rangle$, a *sub-branch* starting with $u \in B$ and ending with $v \in B$ is any non-empty $B' \subseteq B$ such that $w \in B$ is a member of B' iff w is a (possibly improper) suffix of u and v is a (possibly improper) prefix of w.

Let $\langle I, T, f\rangle$ with $u \in T$. The *subtree rooted at u* is the tree $\langle I, T', f'\rangle$ where $T' = \{v \mid uv \in T\}$ and $f'(v) = f(uv)$ for all $v \in T'$.

B PROOF OF NORMALIZATION, PARACONSISTENCY, AND TRANSITIVITY

Theorem 1 (Normalization) *Every derivation can be transformed into a normal derivation (with the same premises and conclusion).*

Corollary 1 (Transitivity) $\varphi \vdash \psi$ *and* $\psi \vdash \chi$ *implies* $\varphi \vdash \chi$ *for all* φ, ψ, χ.

Corollary 2 (Paraconsistency) *For all* φ *there exists* ψ *such that* $\varphi, \neg\varphi \nvdash \psi$.

We prove that if a formula ψ concludes a derivation in which a maximal formula φ occurs then there exists a normal derivation of ψ.

We first show how to reduce the length of maximum segments when necessary: Considering a maximum segment whose maximum formula is lowest in a branch (cf. the lowest occurrence of T below), apply the following permutation

$$\begin{array}{c}\dfrac{\begin{array}{c}\Delta\\X\vee Y\end{array}\quad \begin{array}{c}\Sigma_1\\T\end{array}\quad\begin{array}{c}\Sigma_2\\T\end{array}}{T}\\ \Pi\\ Z\\ \Theta\end{array} \quad\Longrightarrow\quad \begin{array}{c}\dfrac{\begin{array}{c}\Delta\\X\vee Y\end{array}\quad\begin{array}{c}\Sigma_1\\T\\\Pi\\Z\end{array}\quad\begin{array}{c}\Sigma_2\\T\\\Pi\\Z\end{array}}{Z}\\ \Theta\end{array}$$

In the resulting derivation, there can be no maximal segment through $X \vee Y$. Therefore, there can be no new maximum segment in the leftmost branch. Similarly with the other two branches. However, the length of the maximum segment we considered is now decreased by one. The process can be iterated until the desired length is obtained for any maximal segment(s) we consider.

When considering a maximal segment of length 1 whose maximal formula is $X \to Y$, we apply the following transformation:

$$\begin{array}{c}\dfrac{\dfrac{\begin{array}{c}^{(1)}X\\\Sigma_2\\Y\end{array}}{X\to Y}\,{\scriptstyle (1)\ (I\to)}\quad \begin{array}{c}\Sigma_1\\X\end{array}}{Y}\\\Pi\\Z\end{array}\quad\Longrightarrow\quad\begin{array}{c}\Sigma_1\\X\\\Sigma_2\\Y\\\Pi\\Z\end{array}$$

Note that all formulas discharged in Σ_1 or Σ_2 are still discharged in the resulting derivation (if n occurrences of X are discharged within Σ_2 then the resulting derivation is to display n additional copies of Σ_1). Of course, no new discharge is introduced. Clearly, every other proviso of the rules is also satisfied. Therefore, we have obtained a derivation with the same hypotheses and the same conclusion.

Observe that there can be a new maximal segment (in the resulting derivation) only if X or Y is involved in it. Further observe that there can be a new maximal segment involving X only if either X or a formula in Σ_2 is a maximal formula. Whatever is the case, the definition of a maximal segment forces the new maximal formula to be an occurrence of X. The same applies to Y wrt Π. All in all, the (at most two) new maximal formulas X and Y are sub-formulas of the initial maximal formula $X \to Y$.

When considering a maximal segment of length 2 (cf. (†)) whose maximal formula is $X \vee Y$, we apply the following transformation:

$$\begin{matrix}\dfrac{\dfrac{\begin{matrix}\Delta\\ W\end{matrix}\quad\begin{matrix}\Sigma_1\\ X\\ X\vee Y\end{matrix}\quad\begin{matrix}\Sigma_2\\ Y\\ X\vee Y\end{matrix}}{X\vee Y}\quad\begin{matrix}{}^{(1)}X\\ \Theta_1\\ T\end{matrix}\quad\begin{matrix}{}^{(1)}Y\\ \Theta_2\\ T\end{matrix}}{T}\,{\scriptstyle (1)\ (E\vee)}\\ \Pi\\ Z\end{matrix}$$

$$\Longrightarrow \qquad \begin{matrix}\dfrac{\begin{matrix}\Delta\\ W\end{matrix}\quad\begin{matrix}\Sigma_1\\ X\\ \Theta_1\\ T\end{matrix}\quad\begin{matrix}\Sigma_2\\ Y\\ \Theta_2\\ T\end{matrix}}{T}\\ \Pi\\ Z\end{matrix}$$

Note that all formulas discharged in Σ_1 or Σ_2 are still discharged in the resulting derivation (observe that, in the original derivation, discharging within any Θ_i an hypothesis introduced in some Σ_j is incorrect). Also, all formulas introduced in Θ_1 and Θ_2, if discharged, are still discharged in the resulting derivation. Similarly, every other proviso of the rules is also satisfied. Again, we obtain a derivation with the same hypotheses and the same conclusion.

Observe that there can be a new maximal segment (in the resulting derivation) only if X or Y is involved in it. Further observe that there can be a new maximal segment involving X only if either X or a formula in Θ_1 is a maximal formula. Whatever is the case, the definition of a maximal segment forces the new maximal formula to be an occurrence of X. The same applies to Y wrt Θ_2. All in all, the (at most two) new maximal formulas X and Y are sub-formulas of the initial maximal formula $X \vee Y$.

In the resulting derivation, only X and Y can be new maximal formulas (whether the initial maximal formula is $X \to Y$ or $X \vee Y$). So, the transformation either decreases the number of maximal formulas or replaces a maximal formula φ by simpler maximal formulas (actually, one or two sub-formulas of φ). Of course, formulas have a finite number of occurrences of the connectives and atomic formulas never are maximal formulas. Repeatedly applying the transformation is then a finite process, ending with a derivation in which no maximal formula occurs.

For the purpose of applying this theorem together with a result due to Tennant, we consider the case where the $(E\vee)$ rule followed by the simplification rule can be normalized in a special $(E\vee)$ rule as follows: When some minor assumption ψ_1 or ψ_2 of $(E\vee)$ is $\perp$, then the conclusion ψ is a copy of the other minor assumption. Clearly, the above transformation still gives us the desired outcome and the proof is over.

In view of the system defined by Tennant in [1987, p. 672], this theorem yields the desired result that our system is paraconsistent.

Acknowledgements

This work was supported by the French foreign ministry within programme PROCOPE under grant number 99027 and the German ministry for education, science, research and technology within programme PROCOPE under grant number 9822866.

References

[Anderson and Belnap, 1975] A. Anderson and N. Belnap. *Entailment: The Logic of Relevance and Necessity, Vol. I.* Princeton University Press, 1975.

[Dummett, 1977] M. Dummett. *Elements of intuitionism.* Oxford University Press, 1977.

[Epstein, 1979] R. Epstein. Relatedness and implication. *Philosophical Studies*, 36:137–173, 1979.

[Gentzen, 1935] G. Gentzen. Untersuchungen über das logische Schließen. *Mathematische Zeitschrift*, 39:176–210 and 405–431, 1935. English translation in [Szabo,1969].

[Girard, 1987] J.-Y. Girard. Linear logic. *Theoretical Computer Science*, 50:1–102, 1987.

[Krajewski, 1986] S. Krajewski. Relatedness logic. *Report on Mathematical Logic*, 20:7–14, 1986.

[Parry, 1989] W. Parry. Analytic implication: Its history, justification, and varieties. In R. Routley and J. Norman, editors, *Directions in Relevant Logic.* Nijhoff, Amsterdam, 1989.

[Prawitz, 1965] D. Prawitz. *Natural Deduction.* Almqvist & Wiksell, Stockholm, 1965.

[Read, 1989] S. Read. *Relevant Logic. A Philosophical Examination of Inference.* Basil Blackwell, 1989.

[Robinson, 1965] J. Robinson. A machine-oriented logic based on the resolution principle. *Journal of the ACM*, 12(1):23–41, 1965.

[Szabo, 1969] M. Szabo, editor. *The collected papers of Gerhard Gentzen*. North-Holland, 1969.

[Tennant, 1987] N. Tennant. Natural deduction and sequent calculus for intuitionistic relevant logic. *Journal of Symbolic Logic*, 52(3):665–680, 1987.