Peter Verdée[§]
Inge De Bal[¶]

A new approach to classical relevance

Abstract. In this paper we present a logic that determines when implications in a classical logic context express a relevant connection between antecedent and consequent. In contrast with logics in the relevance logic literature, we leave classical negation intact - in the sense that the law of non-contradiction can be used to obtain relevantly implications, as long as there is a connection between antecedent and consequent. On the other hand, we give up the requirement that our theory of relevance can define a new standard of deduction. We present and argue for a list of requirements that such a logical theory of classical relevance needs to meet and go on to formulate a system that respects each of these requirements.

The presented system is a monotonic and transitive logic that extends the relevance logic **R** with a richer relevant implication that allows for Disjunctive Syllogism and similar rules. This is achieved by interpreting the logical symbols in the antecedents in a stronger way than the logical symbols in consequents. A proof theory and an algebraic semantics are formulated and interesting metatheorems (soundness, completeness and the fact that it satisfies the requirements for classical relevance) are proven.

Finally we give a philosophical motivation for our non-standard relevant implication and the asymmetric interpretation of antecedents and consequents.

Keywords: relevant implication, classical negation, relevance logic, algebraic semantics, non-transitive implication

1. Introduction

It is well known that the implication of classical propositional logic has very counterintuitive properties—we will call them henceforth the *paradoxes of material implication*[1]. Consider for example the fact that $(A \wedge \neg A) \supset B$, $A \supset (\neg A \supset B)$, $A \supset (B \vee \neg B)$ and $(A \supset B) \vee (B \supset C)$ are all theorems of Classical Logic. Nevertheless, no rational agent would ever assume that their communication partner utters something that is trivially true when he uses natural language sentences of that form. In natural language we have the tendency to assume that 'A implies B' means something like 'arguments to accept A give us arguments to accept B' or 'accepting A automatically leads to accepting B'. Hence, rational human agents assume that there is a (logical

Presented by **Name of Editor**; *Received* December 1, 2005

[1]We use this term only because the counterintuitive properties are usually called paradoxes in the relevance logic literature. Here we use the term without negative connotation.

or contingent) connection between the antecedent and the consequent of the implications they use in natural language discourse. If this is not the case, one usually does not use expressions like 'A implies B' but rather disjunctive expressions 'A is not the case or B is the case'.

In the literature, there are two ways of solving this problem. The first is the traditional relevance logic solution. After pioneering work by Church and Ackermann (cf. [1] and [7]), the systems are thoroughly elaborated by, among others, Anderson, Belnap, Meyer, Read, Dunn Routley, Urquhart and Brady, cf. [2], [3], [9], [8], [10], [11], [13], [12], and [17]. This group of solutions is basically revisionistic. They claim that the irrelevancies in Classical Logic are pure fallacies of reasoning. So they argue that Classical Logic (or at least its implication) is mistaken and needs to be fixed. They have developed different logical systems that do not have the counterintuitive properties Classical Logic has. These systems are weaker than Classical Logic and many aspects of Classical Logic are not valid in it. One of the more striking features is that they in fact give up the law of non-contradiction. This of course ensures that $(A \wedge \neg A) \rightarrow B$ is no longer valid, but it also invalidates $(A \wedge (\neg A \vee B)) \rightarrow B$ and $A \rightarrow (A \wedge (B \vee \neg B))$. In these implications, there at least seems to be a reasonable connection between antecedent and conclusion. And indeed there is from a classical point of view. But there is no implicational relation from a relevantist one, because they reject the general applicability of the law of non-contradiction (as well as of the law of excluded middle)[2]. So, from a relevantist point of view, $p \wedge (\neg p \vee q)$ does not imply q, but this is not the case because the content of the antecedent is irrelevant for the content of the conclusion, but because the implication is simply falsifiable in relevance logics.

The other way of solving it is due to Neil Tennant (cf. [14], [15] and [16]). He does not want to criticize classical reasoning, but wants to formalize actual classical reasoning as it is done in mathematics and in the hypothetico-deductive method of the sciences in such a way that irrelevancies are sieved out. This stance sees the irrelevant aspects of Classical Logic thus only as byproducts of how Boole and Frege formalized the meaning of logical connectives. For him there is no problem with the actual classical reasoning. He wants to show that everything that is usually formalized by means of Classical Logic entailment can also be formalized by means of his favourite logic which is free of the paradoxes of material implication. This

[2] To be more precise: these laws do not hold in all worlds in the Routley-Meyer semantics for the relevance logic **R**, although they do hold in the actual/normal worlds of that semantics.

logic, however, is not transitive, which means that he should find a way to eliminate all uses of transitivity in actual reasoning before he can start the formalization.

We, however, want to explore a third option. We do not want to reject Classical Logic. We do not even wish to reject the implication of Classical Logic. Especially in mathematics, computer science and other axiomatized sciences, Classical Logic has definitely proven its merits. We diagnose the problem of the paradoxes of material implication not as a problem of material implication as such, but as a problem of the relation between natural language and Classical Logic. The expressive power of Classical Logic is simply too limited to express that there is a special connection between antecedent and consequent of implicational statements in natural language. For this reason, we need a more expressive formal tool that is able to filter out the implications that do express a relevant connection from the big pool of implications that are merely materially/classically true. Compare it to the problem of finding causal relations. Sentences of the form 'A causes B' are not uttered for every case where there is a correlation between A and B. A good theory of causality is able to filter out those correlations where there is indeed a causal link in some sense of the word. In the same way we need a theory of relevance that can filter out those materially acceptable implications that also express a relevant connection between antecedent and consequent.

Our filter mechanism is presented as a transitive, monotonic and reflexive consequence relation, in the sense that, given a number of premises, the logic defines which of the implications in the classical consequences express a relevant connection between antecedent and consequent. Moreover, if one already has information that some of the implications in the premises are relevant, the logic will tell us which further statements about relevant implications follow from this. In the logic we will define, $A \to B$ is supposed to mean that A relevantly implies B. $\vdash A \to B$ will therefore mean that it is always the case that A relevantly implies B (remember that 'A relevantly implies B' is, in our approach, short for 'A implies B and there is a relevant connection between A and B'). $A \to B, C \to D \vdash E \to F$ will for example mean that given the information that A relevantly implies B and that C relevantly implies D, we know that E relevantly implies F.

That it is a logic does not mean that we want to sell it as a stand-alone deductive reasoning tool which can replace Classical Logic. Although it would be nice to obtain a stand-alone relevant deductive reasoning tool, in our opinion it is not a necessary requirement for a good theory of relevance. Even stronger, we shall argue that a good theory of relevance in a classical context cannot be a good theory of deduction at the same time. This is because of the fact that the classical relevant implication we defend will turn out to be necessarily non-transitive and Modus Ponens will turn out not to be a valid rule for this type of classical relevant implication.

We will now formulate a number of minimal requirements for a consequence relation $\vdash$ to function as a filter mechanism for classical relevant implications in the above described sense.

We accept the full classical meaning and functionality of the non-implicative symbols. This entails that B follows correctly from $A \wedge (\neg A \vee B)$. There is moreover a relevant connection between antecedent and consequent of this implication; the consequent would not be generally acceptable without the antecedent and from the antecedent does not follow no matter what. So we require that a filter for classical relevance picks out $(A \wedge (\neg A \vee B)) \rightarrow B$ as one of the generally valid relevant implications. The same holds for $A \rightarrow (A \wedge (\neg B \vee B))$. In general we can phrase this requirement by means of the following property of Classical Strength.

(CS) If $A \vdash_{\mathbf{CL}} B$ and $\nvdash_{\mathbf{CL}} \neg A$ and $\nvdash_{\mathbf{CL}} B$, then $\vdash A \rightarrow B$ (A relevantly implies B)

A second requirement is formality. We are looking for a formal theory of relevance. The relevant connection between antecedent and consequent we are interested in is of a formal nature. Without this presupposition one could for example argue that $(p \wedge q) \rightarrow p$ expresses a real relevant connection, but that $((r \vee \neg r) \wedge q) \rightarrow (r \vee \neg r)$ does not express such a connection, because the conclusion is already always true, independent of the truth of the antecedent. This is a valid line of reasoning, but it makes sense to look for a formal criterion anyway. Why? Exactly the same formal connection holds between $(p \wedge q) \rightarrow p$ and $((r \vee \neg r) \wedge q) \rightarrow (r \vee \neg r)$: if $A \wedge B$ holds then A holds. That the connection is not necessary for the consequent $r \vee \neg r$ in that particular instance of the connection could be seen as beside the point: one may claim that for implications to be relevant, it suffices that there is some kind of connection, necessary or not. Moreover, one may wonder what sentential letters like p correspond to in natural language. Maybe p itself corresponds to a necessary truth the truth of which is not formally expressible in propositional Classical Logic, e.g. bachelors are not married. There is no way to express that 'bachelors are not married' is a necessary truth in propositional logic and so one may simply formalize this as p. In that case, if one accepts that $(p \wedge q) \rightarrow p$ is a relevant implication, one should also accept that one of the sentences it could formalize, 'bachelors are not married and the moon is green implies that bachelors are not married' is a relevant implication. But of course the antecedent is in this case not needed for accepting the consequent. We conclude from this that requiring that the relevant connection is really necessary is too strict a requirement for an applicable theory of relevance. Without this requirement there seems to be no reason to give up formality in this case. This formal approach is also in line with all of the relevance logic tradition up to now.

The formality requirement is in this case specified as the validity of the law of Uniform Substitution.

(US) If $\vdash A$ then also $\vdash B$, where B is the result of substituting every occurrence of a sentential letter in A by a formula C.

A third requirement is monotonicity. This comes down to requiring that strengthening the antecedent can never take away the relevant connection between antecedent and consequent. We require only the existence of a relevant connection. Adding more information to the antecedent cannot take away a connection that was originally there. So the connection still exists if the antecedent is strengthened (even if it strengthened to inconsistency).

(MO1) If $\vdash A \to B$, then $\vdash (A \wedge C) \to B$

Given the idea that the existence of a relevant connection is enough to consider an implication as relevant and the classical meaning of conjunction, one should also accept the following related conjunction principle (but now with the conjunction in the antecedent).

(MO2) If $\vdash A \to B$ and $\vdash A \to C$, then $\vdash A \to (B \wedge C)$

We realize that it would make sense to remove these monotonicity requirements and still obtain a reasonable theory of (classical) relevance. One may for example defend a view on relevance where the information contained in the antecedent as a whole should be relevant for the consequent. In that case one would accept that $(p \wedge q) \to p$ is a relevant implication but $(p \wedge \neg p \wedge q) \to p$ is not. As the second antecedent is inconsistent, it does not contain any information whatsoever, and so there can be no relevant connection between this information and the information contained in the consequent. Indeed, that would make sense, but then one should also give up the formality requirement (substituting q by $r \wedge \neg r$, would transform $(p \wedge q) \to p$ once again into an irrelevant implication). In general the formality and the monotonicity requirement are closely related. A similar objection against treating $(p \wedge q) \rightarrow p$ as relevant but $(p \wedge \neg p \wedge q) \rightarrow p$ not may be employed as follows: let q stand for some sentence the impossibility of which is not expressible in propositional logic e.g. 'there is a round square'. Then the information expressed in the antecedent of $(p \wedge q) \rightarrow p$ as a whole is just as empty as the one contained in q itself and so there can be no relevant connection between $p \wedge q$ and p.

So, there may be good reasons to remove the monotonicity and formality requirement, but this requires a different way of looking at logical form and the relation with natural language (as is for example done in the Adaptive Logic programme, see [5] and [4]). Here we choose not to follow that road and stay closer to the formal relevance logic tradition.

A fourth requirement is that we agree with the positive characterization of relevance by the logic **R**. All theorems of **R** indeed express correct statements about relevant implications. Because the meaning of the non-implicational symbols of relevance logic is weaker than the meaning of their classical counterparts, every relevant implication between non-implicational formulas valid in the relevance logic tradition should a fortiori be valid in the classical relevance logic we want to design. Regarding implications between implicational formulas, we also see no reason to reject any of the implications that are validated by the relevance logic **R**. There is no deep reason for our requirement to start from the relevance logic **R**, but **R** seems to be the strongest available relevance logic to account for both contingent and logical relevance, without problematic theorems[3].

A final requirement is of course that the counterintuitive aspects of material implication disappear for the relevant implication. None of the so called paradoxes of material implication should be validated in a good theory of relevance. A list of these so called paradoxes can be found in section 6.

It is clear that none of the existing approaches to relevance can meet all of these requirements. This is not a criticism against their approach to relevance. It could only be a criticism in so far as they would want to give a filtering theory of classical relevance. Given that both traditions (the relevance logic tradition and the Tennant tradition) criticize and reject Classical Logic, probably none of the logical systems produced in these traditions aim to be such a theory of classical relevance.

From our point of view, however, it is very useful to give a theory of classical relevance that meets all the requirements. One may think that

[3]**RM** is an even stronger logic than **R**, which is often seen as a relevance logic. But this logic validates $A \rightarrow (A \rightarrow A)$ and we see no reason why A would be any more relevant for $A \rightarrow A$, than any other formula B.

all the requirements together are too strong to result in a coherent logical system. We however show in this paper that it is perfectly possible to give such a logical theory of classical relevance. The presented system will, admittedly, have properties that one would ideally not want, but we argue that this only shows that classical relevance does not behave as one would maybe expect. The system is coherent and well defined. The undesired properties are not artifacts of the technical machinery used in this paper but are inescapable properties of any theory that would respect all of the above requirements.

2. Language

We define three different languages. We first define $\mathcal{L}$ as the language of propositional Classical Logic with sentential letters $p, q, r, p_1, \ldots$, the binary connective $\vee$ and the unary connective $\neg$. $\mathcal{L}^{\twoheadrightarrow}$ will denote the extension of $\mathcal{L}$ with extra binary connectives $\rightarrow$ and $\twoheadrightarrow$ and the propositional constant t. $\mathcal{L}^{\mapsto}$ will denote the extension of $\mathcal{L}$ with binary connectives $\rightarrow$ and $\mapsto$, the unary connective $\oplus$ and the propositional constant t. $\mathcal{W}$, $\mathcal{W}^{\twoheadrightarrow}$ and $\mathcal{W}^{\mapsto}$ are the set of wffs of respectively $\mathcal{L}$, $\mathcal{L}^{\twoheadrightarrow}$ and $\mathcal{L}^{\mapsto}$ constructed in the usual way.

$\mathcal{L}^{\twoheadrightarrow}$ is the language we really want to study. The central logic **RR** defined and studied in this paper will have a set of theorems in $\mathcal{W}^{\twoheadrightarrow}$ and a consequence function in $\mathcal{P}(\mathcal{W}^{\twoheadrightarrow}) \rightarrow \mathcal{P}(\mathcal{W}^{\twoheadrightarrow})$. The $\twoheadrightarrow$-implication of this logic will be the actual classical relevant implication we are looking for and the $\rightarrow$-implication will be the implication that corresponds to the consequence/deduction relation in the logic **RR**. We were not yet able to axiomatize this logic **RR** in its own language. For this reason we will first define a logic **R2** in the language $\mathcal{W}^{\mapsto}$ from which the logic **RR** will be defined by focusing on specific defined symbols of **R2**. Hence, we do not aim to motivate or interpret the logic **R2** and its language. The reader is asked to see this logic and its language as a purely technical auxiliary means to define **RR**.

3. Syntactical characterization: axiomatization of R2

We start of with the Hilbert style proof theoretic characterization of the auxiliary logic **R2**. Given this auxiliary logic, we will be able define the logic **RR** which is the actual central system in this paper, i.e. the logical system that will fulfill all requirements (mentioned in the first section) to function as a theory for classical relevance.

The axioms of the system **R2** are the following ($A \wedge B$ abbreviates $\neg(\neg A \vee \neg B)$ and $A \leftrightarrow B$ abbreviates $(A \to B) \wedge (B \to A)$):

We start with the standard axioms of the logic **R** (see [2]) with a truth constant t.

(A1) $A \to A$
(A2) $(A \to B) \to ((B \to C)) \to (A \to C))$
(A3) $A \to ((A \to B) \to B)$
(A4) $(A \to (A \to B)) \to (A \to B)$
(A5) $(A \wedge B) \to A$
(A6) $(A \wedge B) \to B$
(A7) $((A \to B) \wedge (A \to C)) \to (A \to (B \wedge C))$
(A8) $A \to (A \vee B)$
(A9) $B \to (A \vee B)$
(A10) $((A \to C) \wedge (B \to C)) \to ((A \vee B) \to C)$
(A11) $(A \wedge (B \vee C)) \to ((A \wedge B) \vee C)$
(A12) $(A \to \neg B) \to (B \to \neg A)$
(A13) $\neg\neg A \to A$
(A14) $A \leftrightarrow (t \to A)$

We add axioms for another arrow $\mapsto$ that behaves exactly like the **R**-arrow $\to$ (in relation to the other connectives $\vee$, $\neg$, and t). The only difference between the two arrows $\to$ and $\mapsto$ is that the first is always a consequence of the second but not vice versa.

(A15) $(A \mapsto B) \mapsto ((B \mapsto C)) \mapsto (A \mapsto C))$
(A16) $A \mapsto ((A \mapsto B) \mapsto B)$
(A17) $(A \mapsto (A \mapsto B)) \mapsto (A \mapsto B)$
(A18) $((A \mapsto B) \wedge (A \mapsto C)) \mapsto (A \mapsto (B \wedge C))$
(A19) $((A \mapsto C) \wedge (B \mapsto C)) \mapsto ((A \vee B) \mapsto C)$
(A20) $(A \mapsto \neg B) \mapsto (B \mapsto \neg A)$
(A21) $A \leftrightarrow (t \mapsto A)$

The next axiom expresses the relation between the two arrows.

(A22) $(A \mapsto B) \mapsto (A \to B)$

Finally we axiomatize the symbol $\oplus$. It weakens the meaning of classical symbols in such a way that they cannot be eliminated/detached anymore (without extra information). $\oplus$ weakens the conjunction to the fusion operation $\circ$ so that $\oplus(A \wedge B)$ does not imply either $\oplus A$ or $\oplus B$ (nor A or B).

Disjunction is weakened in such a way it cannot lead to Disjunctive Syllogism. Because the standard disjunction of relevance logics like **R** is already not detachable (it cannot lead to Disjunctive Syllogism), $\diamondplus$ transforms $A \vee B$ into the **R**-disjunction of $\diamondplus A$ and $\diamondplus B$. $\diamondplus$ is moreover designed in such a way that it does not affect/transform the meaning of relevant implications or their negations.

(A23) $\diamondplus(\neg(A \vee B)) \leftrightarrow \neg(\diamondplus\neg A \rightarrow \neg\diamondplus\neg B)$
(A24) $\diamondplus(A \vee B) \leftrightarrow (\diamondplus A \vee \diamondplus B)$
(A25) $\diamondplus(A \rightarrow B) \leftrightarrow (\neg\diamondplus\neg A \rightarrow \diamondplus B)$
(A26) $\diamondplus(\neg(A \rightarrow B)) \leftrightarrow \neg(\diamondplus A \rightarrow \neg\diamondplus\neg B)$
(A27) $\diamondplus(A \mapsto B) \leftrightarrow (\neg\diamondplus\neg A \mapsto \diamondplus B)$
(A28) $\diamondplus(\neg(A \mapsto B)) \leftrightarrow \neg(\diamondplus A \mapsto \neg\diamondplus\neg B)$
(A29) $\diamondplus\neg\neg A \leftrightarrow \diamondplus A$,
(A30) $\diamondplus t \leftrightarrow t$,
(A31) $\diamondplus\neg t \leftrightarrow \neg t$,
(A32) $\diamondplus\diamondplus A \leftrightarrow \diamondplus A$,
(A33) $\diamondplus\neg\diamondplus A \leftrightarrow \diamondplus\neg A$.

We need to express that formulas are at least as strong as their $\diamondplus$-counterpart, and that $\neg\diamondplus\neg A$ is at least as strong as A. $\neg\diamondplus\neg A$ expresses that the classical symbols in A can be eliminated/detached: $\neg\diamondplus\neg$ transforms a conjunction into the detachable conjunction $\wedge$ and it transforms a disjunction $\vee$ into an (obviously detachable) intensional disjunction $+$ ($A + B =_{df} \neg A \rightarrow B$).

(A34) $\neg\diamondplus\neg A \rightarrow A$
(A35) $A \rightarrow \diamondplus A$

The following are the rules of **R2**.

(R1) $A, A \rightarrow B/B$ (Modus Ponens)
(R2) $A, B/A \wedge B$
(R3) $A \rightarrow B/A \mapsto B$

One might be inclined to think that (R3) and (A22) together entail that the two arrows $\rightarrow$ and $\mapsto$ have the same meaning. However, it is central for this paper that this is not the case: $(A \rightarrow B) \rightarrow (A \mapsto B)$ is not a theorem in **R2**. The two symbols only have the same meaning when they occur at the outmost level in formulas: $\vdash_{\mathbf{R2}} A \rightarrow B$ iff $\vdash_{\mathbf{R2}} A \mapsto B$. At the outmost level $A \rightarrow B$ and $A \mapsto B$ both express that A has a value in the algebra that is at least as low as B, given the algebraic semantics defined in the next section.

In the language of **R2** we can define the following other symbols.

(D1) $\boxplus A =_{df} \neg\oplus\neg A$
(D2) $A \circ B =_{df} \neg(A \to \neg B)$
(D3) $A \bullet B =_{df} \neg(A \mapsto \neg B)$
(D4) $A \twoheadrightarrow B =_{df} \oplus(A \mapsto B)$

To already see what the $\oplus$ symbol actually does, we list the following theorems of **R2**.

(T1) $\vdash_{\mathbf{R2}} \oplus A \to \oplus(A \vee B)$
(T2) $\vdash_{\mathbf{R2}} \oplus A \to (\oplus B \to \oplus(A \wedge B))$
(T3) $\vdash_{\mathbf{R2}} \boxplus(A \wedge B) \to \boxplus A$
(T4) $\vdash_{\mathbf{R2}} \boxplus(A \vee B) \to (\oplus\neg A \to \boxplus B)$

The idea is that one is able to decompose $\boxplus$-preceded formulas, but $\oplus$ formulas are too weak to be decomposed (but one is able to compose them into other $\oplus$-formulas). This enables us to define the implication $A \twoheadrightarrow B$ as a relevant implication with a strong antecedent and a weaker consequent. Exactly this enables us to obtain Classical Strength for the $\twoheadrightarrow$-implication. The idea of formally distinguishing decomposability from non-decomposability is based on the results presented in [18].

4. Algebraic semantics of R2

Definition 1. *De Morgan monoid.*

$\langle D, \vee, \circ, \neg, e\rangle$ is a de Morgan monoid iff D is a set, e is an element of D, $\wedge$ and $\circ$ are binary operations on D, $\neg$ is an unary operation on D, and where $a \leq b =_{df} a \vee b = b$, for all $a, b, c \in D$ it holds that

(DM1) $e \circ a = a$
(DM2) $(a \circ b) \circ c = a \circ (b \circ c)$
(DM3) $a \circ b = b \circ a$
(DM4) $a \leq a \circ a$
(DM5) $(a \vee b) \vee c = a \vee (b \vee c)$
(DM6) $a \vee b = b \vee a$
(DM7) $a \vee a = a$
(DM8) $\neg\neg A = A$
(DM9) $a \circ b \leq c$ *iff* $a \circ \neg c \leq \neg b$

(DM10) $\neg(\neg(a \vee b) \vee c) = \neg(\neg a \vee c) \vee \neg(\neg b \vee c)$
(DM11) $a \circ (b \vee c) = (a \circ b) \vee (a \circ c)$

Definition 2. **R2**-*structure.*
A sequence $\langle D, \vee, \circ, \bullet, \neg, \oplus, e\rangle$ *is a* **R2**-*structure iff, for all* $a, b \in D$,

(RS1) $\langle D, \vee, \circ, \neg, e\rangle$ *is a de Morgan monoid,*
(RS2) $\langle D, \vee, \bullet, \neg, e\rangle$ *is a de Morgan monoid,*
(RS3) $a \circ b \leq a \bullet b$
(RS4) $\neg\oplus\neg a \leq a \leq \oplus a$
(RS5) $\oplus e = e$
(RS6) $\oplus\neg e = \neg e$

Definition 3. **R2**-*valuation function.*
Where $\mathfrak{A} = \langle D, \vee, \circ, \bullet, \neg, \oplus, e\rangle$ *is a* **R2**-*structure, a function*

$$v_{\mathfrak{A}} : \mathcal{W}^{\rightarrow} \rightarrow D$$

is a **R2**-*valuation function iff*

(S1) $v_{\mathfrak{A}}(t) = e$,
(S2) $v_{\mathfrak{A}}(A \vee B) = v_{\mathfrak{A}}(A) \vee v_{\mathfrak{A}}(B)$,
(S3) $v_{\mathfrak{A}}(A \rightarrow B) = \neg(v_{\mathfrak{A}}(A) \circ \neg v_{\mathfrak{A}}(B))$,
(S4) $v_{\mathfrak{A}}(A \mapsto B) = \neg(v_{\mathfrak{A}}(A) \bullet \neg v_{\mathfrak{A}}(B))$,
(S5) $v_{\mathfrak{A}}(\neg A) = \neg v_{\mathfrak{A}}(A)$,
(S6) $v_{\mathfrak{A}}(\oplus(A \vee B)) = v_{\mathfrak{A}}(\oplus A) \vee v_{\mathfrak{A}}(\oplus B)$,
(S7) $v_{\mathfrak{A}}(\oplus\neg(A \vee B)) = v_{\mathfrak{A}}(\oplus\neg A) \circ v_{\mathfrak{A}}(\oplus\neg B)$,
(S8) $v_{\mathfrak{A}}(\oplus(A \rightarrow B)) = \neg(v_{\mathfrak{A}}(\neg\oplus\neg A) \circ \neg v_{\mathfrak{A}}(\oplus B))$,
(S9) $v_{\mathfrak{A}}(\oplus\neg(A \rightarrow B)) = v_{\mathfrak{A}}(\oplus A) \circ v_{\mathfrak{A}}(\oplus\neg B)$,
(S10) $v_{\mathfrak{A}}(\oplus(A \mapsto B)) = \neg(v_{\mathfrak{A}}(\neg\oplus\neg A) \bullet \neg v_{\mathfrak{A}}(\oplus B))$,
(S11) $v_{\mathfrak{A}}(\oplus\neg(A \mapsto B)) = v_{\mathfrak{A}}(\oplus A) \bullet v_{\mathfrak{A}}(\oplus\neg B)$,
(S12) $v_{\mathfrak{A}}(\oplus\oplus A) = v_{\mathfrak{A}}(\oplus A)$
(S13) $v_{\mathfrak{A}}(\oplus\neg\oplus A) = v_{\mathfrak{A}}(\oplus\neg A)$.

In what follows, we use the following abbreviations in **R2**-algebras: $a \leq b =_{df} (b = a \vee b)$, $a \rightarrow b =_{df} \neg(a \circ \neg b)$, $a \mapsto b =_{df} \neg(a \bullet \neg b)$, $a \wedge b =_{df} \neg(\neg a \vee \neg b)$, $\boxplus a = \neg\oplus\neg a$, $a \twoheadrightarrow b =_{df} \oplus(a \mapsto b)$ and $f =_{df} \neg e$.

Definition 4. **R2**-*validity.*
$\models_{\mathbf{R2}} A$ *iff* $v_{\mathfrak{A}}(A) = e \vee v_{\mathfrak{A}}(A)$ *(in other words:* $e \leq v_{\mathfrak{A}}(A)$*) for every* **R2**-*valuation function* $v_{\mathfrak{A}}$ *and every* **R2**-*structure* $\mathfrak{A} = \langle D, \wedge, \circ, \bullet, \neg, \oplus, e\rangle$

THEOREM 1. *Soundness.*
If $\vdash_{\mathbf{R2}} A$ *then* $\vDash_{\mathbf{R2}} A$.

THEOREM 2. *Completeness.*
If $\vDash_{\mathbf{R2}} A$ *then* $\vdash_{\mathbf{R2}} A$.

5. The logic RR

The language of the logic **RR** is $\mathcal{L}^{\twoheadrightarrow}$.

DEFINITION 5. $\vdash_{\mathbf{RR}} A$ *iff* $A \in \mathcal{W}^{\twoheadrightarrow}$ *and* $\vdash_{\mathbf{R2}} A$

DEFINITION 6. **RR**-*consequence. Where* $\Gamma \cup \{A\} \subset \mathcal{W}^{\twoheadrightarrow}$,

$$\Gamma \vdash_{\mathbf{RR}} A$$

iff, for some $\Delta \subseteq \Gamma$,

$$\vdash_{\mathbf{RR}} (A_1 \circ A_2 \circ \ldots \circ A_n) \rightarrow A,$$

where $\Delta = \{A_1, \ldots, A_n\}$.

In this notation, we interpret $(A_1 \circ A_2 \circ \ldots \circ A_n) \rightarrow A$ as A when $\Delta = \emptyset$.

Definitions:

(D5) $\blacklozenge A =_{df} t \twoheadrightarrow A$
(D6) $\blacksquare A =_{df} \neg(t \twoheadrightarrow \neg A)$

Formulas $\blacklozenge A$ in the language of **RR** have exactly the same meaning as $\oplus A$ in the language of **R2** and formulas $\blacksquare A$ in the language of **RR** have exactly the same meaning as $\boxplus A$ in the language of **R2**. They are equivalent but they are not literally identical. So we use other symbols to avoid confusion.

There are two implications in the language of **RR**. Both have a distinct motivation. The $\twoheadrightarrow$-implication is the implication that we wish to treat as the classical relevant implication; the one we are looking for in this paper. The $\rightarrow$-implication on the other hand is merely a device that corresponds to the **RR**-concept of consequence. This is an important distinction in the context of our project. We want to enable Disjunctive Syllogism and so $A, \neg A \vee B \vdash B$ but we do not want to derive $A \twoheadrightarrow B$ from $\neg A \vee B$, as this would clearly be paradoxical. We solve this problem by means of the $\rightarrow$-implication. We have as a theorem $\vdash_{\mathbf{RR}} \boxplus(\neg A \vee B) \rightarrow (A \rightarrow B)$; this merely express that Disjunctive Syllogism is valid. But we do not have $\vdash_{\mathbf{RR}} \boxplus(\neg A \vee B) \to (A \twoheadrightarrow B)$. We are not able to conclude that A relevantly implies B from the mere truth of $\neg A \vee B$, which is exactly as one would expect from a relevant implication.

6. Metatheoretic properties

In this section we list a number of important properties of the system **RR**. The proofs of non-evident results can be found in the appendix.

THEOREM 3. *Tarski relation.*
RR *is (1) reflexive, i.e.* $\Gamma \cup \{A\} \vdash_{\mathbf{RR}} A$*, (2) transitive, i.e. whenever* $\Gamma \cup \{B\} \vdash_{\mathbf{RR}} A$, $\Delta \vdash_{\mathbf{RR}} B$*, then also* $\Delta \cup \Gamma \vdash_{\mathbf{RR}} A$*, and (3) monotonic, i.e. whenever* $\Gamma \vdash A$*, also* $\Gamma \cup \Delta \vdash A$.

THEOREM 4. *Classical strength.*
Let $A, B \in \mathcal{W}$*. If* $A \vdash_{\mathbf{CL}} B$ *and* $\nvdash_{\mathbf{CL}} \neg A$ *and* $\nvdash_{\mathbf{CL}} B$*, then* $\vdash_{\mathbf{RR}} A \twoheadrightarrow B$.

This is provable by means of the next lemma, which presents a way to express the (meta-theoretic) lines of the goal directed proofs by Batens and Provijn (see [6]) in our object language of **RR**. Their procedural approach to Classical Logic (without their EFQ rule) gives all of Classical Logic in case the premises are consistent. Given that the next lemma allows us to embed their system into ours, we can also get all of Classical Logic. More specifically, if $A \vdash_{\mathbf{CL}} B$ and A is consistent, then there is a goal directed proof (without the use of the EFQ-rule) for B from the premise A. The next lemma says that every line of this proof can be translated into a particular **RR**-formula such that this formula **RR**-follows from the premises of the proof preceded by $\blacksquare$. Given that the goal directed proof successfully proves B, there is a line containing B without a condition. This line is translated into the formula $\blacklozenge B$. Hence, by the next lemma, we obtain $\blacksquare A \vdash_{\mathbf{RR}} \blacklozenge B$ which is equivalent to $\vdash_{\mathbf{RR}} \blacksquare A \to \blacklozenge B$ and therefore also to $\vdash_{\mathbf{R2}} \boxplus A \to \diamondplus B$. With (R3) we can conclude $\vdash_{\mathbf{R2}} \boxplus A \mapsto \diamondplus B$ from this. (A27) gives us $\vdash_{\mathbf{R2}} \diamondplus(A \mapsto B)$, which is the same as $\vdash_{\mathbf{RR}} A \twoheadrightarrow B$.

LEMMA 1. *Embedding goad directed proofs.*
Where $\Gamma \cup \{A\} \subset \mathcal{W}$*, if a line*

$$[A_1, A_2, \ldots A_n]A$$

can occur in a goal directed proof from Γ *then*

$$\Gamma^{\blacksquare} \vdash_{\mathbf{RR}} (\blacklozenge A_1 \circ \blacklozenge A_2 \circ \ldots \circ \blacklozenge A_n) \to B$$

where $B = \blacklozenge A$ *if* A *is the goal formula, and* $B = \blacksquare A$ *otherwise.*

THEOREM 5. *Extension of* **R** *(in two ways).*
(1) When one removes $\twoheadrightarrow$ *from the language of* **RR**, *one obtains the logic* **R**.
(2) For every theorem A *of the relevance logic* **R**, *when one replaces every occurrence of* $\rightarrow$ *in* A *by* $\twoheadrightarrow$, *one obtains a theorem of* **RR**.

THEOREM 6. *Paradox-free.*

$\nvdash_{\mathbf{RR}} B \twoheadrightarrow (A \twoheadrightarrow B)$
$\nvdash_{\mathbf{RR}} \neg A \twoheadrightarrow (A \twoheadrightarrow B)$
$\nvdash_{\mathbf{RR}} (\neg A \vee B) \twoheadrightarrow (A \twoheadrightarrow B)$
$\nvdash_{\mathbf{RR}} \neg(A \twoheadrightarrow B) \twoheadrightarrow (A \wedge \neg B)$
$\nvdash_{\mathbf{RR}} \neg(A \twoheadrightarrow B) \twoheadrightarrow (B \twoheadrightarrow A)$
$\nvdash_{\mathbf{RR}} (A \twoheadrightarrow B) \vee (B \twoheadrightarrow A)$
$\nvdash_{\mathbf{RR}} (A \twoheadrightarrow B) \vee (A \twoheadrightarrow \neg B)$
$\nvdash_{\mathbf{RR}} (A \twoheadrightarrow B) \vee (\neg A \twoheadrightarrow B)$
$\nvdash_{\mathbf{RR}} ((A \wedge B) \twoheadrightarrow C) \twoheadrightarrow ((A \twoheadrightarrow C) \wedge (B \twoheadrightarrow C))$
$\nvdash_{\mathbf{RR}} ((A \twoheadrightarrow B) \twoheadrightarrow A) \twoheadrightarrow A$
$\nvdash_{\mathbf{RR}} A \rightarrow (\neg A \twoheadrightarrow A)$
$\nvdash_{\mathbf{RR}} (A \wedge \neg A) \twoheadrightarrow B$
$\nvdash_{\mathbf{RR}} B \twoheadrightarrow (A \vee \neg A)$
$\nvdash_{\mathbf{RR}} \neg(A \twoheadrightarrow B) \twoheadrightarrow A$
$\nvdash_{\mathbf{RR}} \neg(A \twoheadrightarrow B) \twoheadrightarrow \neg B$

THEOREM 7. *No deduction theorem for the* $\twoheadrightarrow$ *implication.*
It is not the case that, $\Gamma \vdash_{\mathbf{RR}} A \twoheadrightarrow B$ *whenever* $\Gamma, A \vdash_{\mathbf{RR}} B$.

THEOREM 8. *Non-transitivity of the* $\twoheadrightarrow$*-implication.*
$A \twoheadrightarrow B, B \twoheadrightarrow C \nvdash_{\mathbf{RR}} A \twoheadrightarrow C$.

THEOREM 9. *No Modus Ponens for the* $\twoheadrightarrow$*-implication.*
$A, A \twoheadrightarrow B \nvdash_{\mathbf{RR}} B$.

7. Discussion

In section 6 we showed that the logic **RR** respects all of the requirements we listed in the introduction and that it is moreover a well defined logic with a proof theory, a semantics and a Tarski consequence relation. So the presented system does exactly what we expected from it. Nevertheless, the results of this section also showed us that the relevant implication we defined has some unexpected properties. It turns out that our classical relevant implication does not satisfy the deduction theorem, Modus Ponens nor transitivity. Are these not essential properties for a connective to be called an implication?

Well, in some sense one could say that it is indeed problematic that an implication connective would not have these properties. It means that we cannot use it as our main object language instrument for deduction, which is rightfully considered as the most important use of an implication connective.

Still this does not make our approach useless. Remember that our aim was never to define a new standard of deduction. Our aim was to define a system that could separate the irrelevant (but possibly correct) implications from the relevant ones in Classical Logic. There are no a priori reasons to expect that given the information of there being a relevant connection between A and B and between B and C, there would also be a relevant connection between A and C. Of course A surely implies C if we know that A implies B and B implies C, but why would one require that the relevance of this implication also stays intact? In fact, we suspect that the confusion here lies in the fact that we call a relevant implication an implication. Stating that $A \twoheadrightarrow B$ means not only that A implies B but also that there is a relevant connection between A and B. Given that our connective brings in this extra meaning, transitivity becomes a (maybe desired but) no longer expected property. Even in Classical Logic $(A \supset B) \wedge R_{A,B}$ and $(B \supset C) \wedge R_{B,C}$ does not entail $(A \supset C) \wedge R_{A,C}$. So, that the implication is not transitive is of course inconvenient, but cannot be seen as a problem for our logic as a filtering tool.

The lack of Modus Ponens is perhaps a bit harder to explain. This is related to the fact that we loosen the requirement that the connectives in the consequent of our relevant implication are at least as strong as the ones in the antecedent (in the sense that the same can relevantly be done with them). Of course a relevant implication should preserve the (classical) *truth* from antecedent to consequent. This criterion is obviously met in view of the fact that our logic **R2** is a sub-logic of Classical Logic if one reads the implication $\twoheadrightarrow$ as the classical $\supset$-implication. But there is no a priori reason why the relevant implication should also preserve the *strength of the relevant connection* between antecedents and consequents of implications that occur inside the antecedent and the consequent of the central implication. We live up to the requirements set in the beginning of this paper, precisely because we give up the presupposition that a relevant implication should preserve all logical strength. If one gives up the strength preservation of the relevant implication, it is obvious that one cannot have Modus Ponens.

From accepting $A \twoheadrightarrow B$, and accepting A in a strong way, it follows that we can accept B, but not necessarily in the same strong sense. From accepting $A \twoheadrightarrow B$ and accepting A in a weak sense, it may not even follow that we accept B at all. So, no matter in which way we accept the premises, A and $A \twoheadrightarrow B$, we cannot accept the conclusion B in the same sense.

In order to solve the somewhat strange properties of the $\twoheadrightarrow$-implication, one could let it function inside a non-transitive and possibly even non-monotonic consequence relation, which could be defined as follows:

$$\Gamma \vdash_1 A \text{ iff } \vdash_{\mathbf{RR}} B \to A,$$

where B is the formula that results from conjoining the members of Γ by means of the $\circ$-connective, or

$$\Gamma \vdash_2 A \text{ iff there is a } \Delta \subseteq \Gamma \text{ such that } \Delta \vdash_1 A.$$

$\vdash_2$ is not transitive but still monotonic and $\vdash_1$ is neither transitive nor monotonic. We have Modus Ponens now for both $\vdash_1$ and $\vdash_2$: $A, A \twoheadrightarrow B \vdash_1 B$ and $A, A \twoheadrightarrow B \vdash_2 B$. Also transitivity holds for both relations: $A \twoheadrightarrow B, B \twoheadrightarrow C \vdash_1 A \twoheadrightarrow C$ and $A \twoheadrightarrow B, B \twoheadrightarrow C \vdash_1 A \twoheadrightarrow C$. The deduction theorem holds for $\vdash_1$: if $\Gamma \cup \{A\} \vdash_1 B$, then $\Gamma \vdash_1 A \twoheadrightarrow B$, and a relevant version of it holds for $\vdash_2$: if A is relevant for the deduction $\Gamma \cup \{A\} \vdash_2 B$, then $\Gamma \vdash_1 A \twoheadrightarrow B$. By the way, this relevant deduction theorem already holds for **RR** itself.

However, we do not find sufficient reason to give up the transitivity of the consequence relation. Because it should function as a filtering mechanism, it is very practical that one can employ the full Tarski properties (monotonicity, transitivity and reflexivity) for our consequence relation. Moreover, even if the non-transitive consequence relation $\vdash_2$ gives us all the good implication properties for $\twoheadrightarrow$, it still remains the case that our relevant implication has odd properties. We think it is of no use to sweep the oddness of the classical relevant implication under the rug by defining an ill-motivated non-standard consequent relation.

It is important to remark that the failure of those central properties of implication is not a byproduct of the way in which we technically realize our solution. The requirements set out at the beginning of this paper immediately entail the failure of these properties. Let $\to$ be a classical relevant implication that satisfies all the requirements. By Classical Strength we obtain

$$\vdash p \to p$$

and

$$\vdash \neg p \to (\neg p \vee q).$$

Applying (MO1) to both gives us

$$\vdash (p \wedge \neg p) \rightarrow p$$

and

$$\vdash (p \wedge \neg p) \rightarrow (\neg p \vee q).$$

(MO2) allows us to bring those two results together into

$$\vdash (p \wedge \neg p) \rightarrow (p \wedge (\neg p \vee q)).$$

Moreover, by Classical Strength once again, we get

$$\vdash (p \wedge (\neg p \vee q)) \rightarrow q.$$

Transitivity would already allow us to derive from the latter two expressions that

$$\vdash (p \wedge \neg p) \rightarrow q,$$

which is obviously an irrelevant and therefore unwanted implication. So transitivity has to go.

But Modus Ponens gives us the same result. The fourth requirement says that all **R**-theorems should be valid, so also

$$\vdash (A \rightarrow B) \rightarrow (((A \wedge B) \rightarrow C) \rightarrow (A \rightarrow C))$$

and therefore

$$\vdash (A \rightarrow B) \rightarrow ((B \rightarrow C) \rightarrow (A \rightarrow C)),$$

which can be instantiated as follows:

$$\vdash ((p \wedge \neg p) \rightarrow ((p \wedge (\neg p \vee q)))) \rightarrow (((p \wedge (\neg p \vee q)) \rightarrow q) \rightarrow ((p \wedge \neg p) \rightarrow q))$$

If one would allow Modus Ponens and apply it twice, one would also be able conclude

$$\vdash (p \wedge \neg p) \rightarrow q.$$

So also Modus Ponens has to go.

That the deduction theorem is invalid is probably less of a surprise. Since the consequence relation is monotonic, one can add irrelevant premises as much as one likes without invalidating the consequence. It would of course be wrong to say that, in light of the other premises, there is a relevant connection between any of these irrelevant premises and the consequence.

8. Motivation of the proposed theory

So far we have argued that our proposed theory is not a priori inadequate to formalize human reasoning about relevance. We did this by refuting the potential objections against the strange properties (non-transitivity, no modus ponens) of our approach, unmasking them as prejudices. However, we did not motivate why our particular solution would be more plausible than any other solution that respects the requirements. This is exactly what we will do now.

Essentially, the core idea behind our proposal is that the classical symbols are interpreted differently in the antecedent and in the consequent. More particularly, disjunctions and conjunctions in antecedents are interpreted in a strong sense, while those symbols are interpreted in a weak sense when they occur in consequents.

One can pose the question whether this makes any sense. It seems that, if one wants to respects the classical meaning of conjunction and disjunction, there is only one semantics for them, not two. We however did not necessarily want to stick to the classical semantics of the classical symbols, we only wanted to save their classical behaviour. If one sticks to the classical semantics it is not possible to safeguard relevance and have the property of classical strength. Satisfying this classical strength property is for us sufficient to claim that the symbols have their classical meaning (we employ a 'meaning is use'-attitude towards logical symbols).

We realize relevance by making use of relevance logic. In relevance logic there are two disjunctions and two conjunctions. With these two we disambiguate the classical use of conjunction and disjunction. Our proposal assumes that in classical reasoning they are both referred to by means of the same token. The context makes it clear which of the two meanings is the correct one. In the context of (relevant) implications we would like to conjecture that it makes sense that the correct disambiguation is the one proposed in our logic: the conjunctions/disjunctions in antecedents are strong, i.e. extensional/intensional and the conjunctions/disjunctions in consequents are weak, i.e. intensional/extensional.

Why does this indeed make sense from an everyday reasoning point of view? Let us focus on the disjunction. An intensional disjunction is one which is intended to enable Disjunctive Syllogism. It seems reasonable that this is the kind of disjunction agents intentionally store in their memory as a disjunction in order to later apply Disjunctive Syllogism. An agent stores a disjunction because he is unaware which of the two disjuncts is the case (not because the disjunction is some correct weakening of one of the disjuncts;

it would be useless to store this kind of disjunction). He is unaware which one of the two disjuncts is the case, but he knows that at least one of the two is the case. If he later finds out that one of the disjuncts is not the case, he knows that the other one is the case. So it seems reasonable to say that rational agents want to apply Disjunctive Syllogism to the disjunctions stored in their memories. On the other hand, an agent will not simply have come to such a stored disjunction by means of the law of Addition (from A derive $A \vee B$), simply weakening one of the disjuncts. This would only weaken the stored knowledge and introduce a completely irrelevant and useless statement B. In this sense, the disjunctions in stored information are of an intensional kind. They express that if one of the disjuncts would not be the case, the other one is. Hence disjunctions that are stored in memories of agents are supposed to be ones with which one can apply Disjunctive Syllogism.

The extensional disjunctions correspond to mere weakenings of one of the disjuncts. They are true if one of disjuncts is true (for example, one accepts $A \vee B$ because A is known to be true) but will not explicitly be stored by rational agents, because they contain no new information and hence are redundant. One will moreover not apply Disjunctive Syllogism with them, because this cannot produce new information and can possibly cause loss of information in case former information turns out to be wrong (because of mistakes or mistaken sources or because nature has changed). Suppose one knows A and therefore accepts $A \vee B$. In order for this disjunction to be used in an application of Disjunctive Syllogism one needs extra knowledge $\neg A$ or $\neg B$. With the second one, one can only obtain A, which was already known. The first one indicates that there is a problem with regards to A. Concluding from this that the arbitrary formula B is the case, makes our knowledge trivial. Instead of applying rules like Disjunctive Syllogism in such circumstances, one probably wants to revise some of one's knowledge and for this, one wants to stick to safe rules, i.e. rules with which one does not risk triviality.

So in this picture of disjunction we can distinguish two kinds of information. First *core information*: information that is explicitly stored in order to be used later as valuable means to derive more information. This is the kind of information from which further information is implied. The second kind of information is only a consequence of what is already known. It is true because of the way our classical disjunction works, but it is not explicitly stored and only implicitly implied by what is explicitly stored. In the first kind of information the logical symbols are interpreted in a strong way (with a disjunction that allows for Disjunctive Syllogism). In the second kind the logical symbols do not need to be interpreted in the strong way. It is sufficient that all the weakenings of the core information allowed by classical logic can be part of the weak information. No further derivations are to be done with this kind of information (which will not be stored anyway).

Now we come back to the relevant implication. The antecedents of potentially relevant implications are statements from which other statements can be implied. Given our distinction between two types of information, this entails that the antecedents of implication are supposed to be of the core kind of information. The consequents, on the other hand, may be treated as containing information of the weak kind. All the weakenings of the core information are also implied by the antecedents. That this information is implied by the the antecedents does not entail that this information is again of the strong kind, i.e. the kind one wants to store for further deduction. This motivates why we defined our relevant implication $A \twoheadrightarrow B$ as $\oplus A \mapsto B$ which is equivalent to $\boxplus A \mapsto \oplus B$. The $\boxplus$ formalizes 'interpret what follows in the strong way', while $\oplus$ expresses 'interpret what follows (at least) in the weak way'. Given our intuitive framework, one may read $A \twoheadrightarrow B$ as 'B is at least a weakening of relevant consequences of the core information given by A'. This is a reasonable interpretation of the notion of a classically relevant implication. Antecedents are here read as core information (otherwise we would not want to imply stuff from it anyway). But given the classical logic context in which we situate the implication, the agent should still be able to imply every kind of classically true information from this core information (to be more correct: every kind of classically true information for which there is a relevant connection with the antecedent). So he should also be able to imply the kind of information that is merely a weakening of what is already in the antecedent. Consequents of relevant implications are therefore interpreted to contain the weak kind of information.

Consider a simple toy-example situation. Suppose there is a room with a lamp of which an agent knows that it can be lit by any of two switches. The agent observes that the lamp is on. He concludes that (E1) 'switch 1 is on or switch 2 is on'. This is information worth storing, because it can later result in the new information that either of the two switches is on. This disjunctive statement is core information. If the agent later observes that (E2) 'switch 1 is not on', he is able to conclude (E3) 'switch 2 is on' from (E1) and (E2) by means of Disjunctive Syllogism. The agent may also correctly conclude from what he knows now that (E4) 'switch 2 is on or the electricity is off' by Addition on (E3), but he will not store this in his memory and will not apply Disjunctive Syllogism to it. He will not apply this rule because it can at best result in again inferring (E3), which he already knows. In the worst case, there might be something wrong with (E3) (because of incorrect observation or a mechanical malfunctioning that makes (E1) false) and the agent now observes (E5) 'switch 2 is not on'. Applying Disjunctive Syllogism to (E4) may now result in the absurd conclusion (E6) 'the electricity is off'. So (E1) is here core information with a strong disjunction and (E4) is implicit information with a weak disjunction. It makes sense to say that in this situation (E2) and (E1) together relevantly imply (E3) and also (E4). On the other hand, (E4) and (E5) do not imply (E6) in this situation. Nevertheless, (E1) and (E4) seem to be both classical disjunctions. In this paper we formalize (E1) as a formula of the form $\boxplus(A \vee B)$ (it is core information) and (E5) as a formula $\oplus(A \vee B)$ (it is merely implicit information).

To wrap things up, we started this paper by sketching the outlines of a new approach to relevance: one in which Classical Logic is not rejected but treated as a useful and correct tool that is unfortunately insufficiently expressive to make a proper distinction between implications that do express a relevant connection between antecedent and consequent and the ones that do not. Because implications with a relevant connection have a special status in communication among human agents, we claim that it is very useful to try to develop a filtering mechanism that sorts out all the relevant implications in a Classical Logic context. Standard relevance logics do not accept the full meaning of the classical connectives, and so they for example reject the law of Disjunctive Syllogism. Our new approach on the other hand, explicitly starts from the requirement that all classical consequences with a relevant connection should be expressible by means of a valid implication. We achieved the realization of this project in the $\twoheadrightarrow$-implication of the logic **RR**. One can therefore conclude that the project of filtering out classical relevant implication is feasible. However, the necessary failure of essential principles like Modus Ponens and transitivity makes the filtering mechanism unsuitable as a stand-alone deductive logic (which was never the aim anyway).

Acknowledgements. We are indebted to Fabien Schang and Aleksandra Samonek for useful comments on a previous draft.

References

[1] Wilhelm Ackermann. Begründung einer strengen implikation. *Journal of Symbolic Logic*, 21:113–128, 1956.

[2] Alan R. Anderson and Nuel D. Belnap. *Entailment: The Logic of Rele-*

vance and Necessity, volume 1. Princeton University Press, Princeton, 1975.

[3] Alan Ross Anderson, Nuel D. Belnap, and J. Michael Dunn. *Entailment: The Logic of Relevance and Necessity*, volume 2. Princeton University Press, Princeton, 1992.

[4] Diderik Batens. Extending the realm of logic. The adaptive-logic programme. In Paul Weingartner, editor, *Alternative Logics. Do sciences need them?*, pages 149–164. Springer, Berlin, Heidelberg, 2004.

[5] Diderik Batens. A universal logic approach to adaptive logics. *Logica Universalis*, 1:221–242, 2007.

[6] Diderik Batens and Dagmar Provijn. Pushing the search paths in the proofs. A study in proof heuristics. *Logique et Analyse*, 173–175:113–134, 2001. Appeared 2003.

[7] Alonzo Church. The weak positive implication calculus. *Journal of Symbolic Logic*, 16:238, 1951.

[8] J. Michael Dunn. *The Algebra of Intensional Logics*. PhD thesis, University of Pittsburgh, 1966.

[9] J. Michael Dunn and Greg Restall. Relevance logic. In Dov M. Gabbay, editor, *Handbook of Philosophical Logic*, volume 6, pages 1–136. Kluwer Academic Publishers, second edition, 2002.

[10] Robert K. Meyer and Richard Routley. Classical relevant logics II. *Studia Logica*, 33:183–194, 1973.

[11] Stephen Read. *Relevant logic: a philosophical examination of inference*. Basil Blackwell, Oxford, 1988.

[12] Richard Routley and Robert K. Meyer. Semantics of entailment. In Hugues Leblanc, editor, *Truth Syntax and Modality*, pages 194–243. North Holland, 1973.

[13] Richard Routley, Val Plumwood, Robert K. Meyer, and Ross T. Brady. *Relevant Logics and their Rivals*. Ridgeview, 1982.

[14] Neil Tennant. Perfect validity, entailment and paraconsistency. *Studia Logica*, 43(1):181–200, 1984.

[15] Neil Tennant. The transmission of truth and the transitivity of deduction. In Dov Gabbay, editor, *What is a Logical System?*, volume 4 of *Studies in Logic and Computation*, pages 161–177. Oxford University Press, Oxford, 1994.

[16] Neil Tennant. Relevance in reasoning. In Stewart Shapiro, editor, *Handbook of Philosophical Logic and Mathematics*, pages 696–726. Oxford University Press, Oxford, 2004.

[17] Alasdair Urquhart. *The Semantics of Entailment.* PhD thesis, University of Pittsburgh, 1972.

[18] Peter Verdée. Strong paraconsistency by separating composition and decomposition in classical logic. In Lev D. Beklemishev and Ruy de Queiroz, editors, *Logic, Language, Information and Computation*, volume 6642 of *Lecture Notes in Computer Science*, pages 272–292. Springer Berlin Heidelberg, 2011.

Peter Verdée
Centre for Logic and Philosophy of Science
Ghent University
Blandijnberg 2
B-9000 Ghent, Belgium
`Peter.VerdeeUGent.be`

Inge De Bal
Centre for Logic and Philosophy of Science
Ghent University
Blandijnberg 2
B-9000 Ghent, Belgium
`Inge.DeBalUGent.be`

Appendix

8.1. Proof of Theorem 1

The reader can verify that all our axioms come out as verified by all **R2**-models.

8.2. Proof of Theorem 2

We prove that 'if $\vDash_{\mathbf{R2}} G$ then $\vdash_{\mathbf{R2}} G$', by showing that the contraposition 'if $\nvdash_{\mathbf{R2}} G$ then $\nvDash_{\mathbf{R2}} G$' holds.

Suppose $\nvdash_{\mathbf{R2}} G$.

We construct the Lindenbaum algebra of **R2**. Let $[A]$ denote the $\leftrightarrow$-equivalence class of A, i.e. $[A] =_{df} \{B \mid \vdash_{\mathbf{R2}} A \leftrightarrow B\}$. That this is indeed an equivalence class follows from $\vdash_{\mathbf{R2}} A \leftrightarrow A$ (reflexivity), if $\vdash_{\mathbf{R2}} A \leftrightarrow B$ then $\vdash_{\mathbf{R2}} B \leftrightarrow A$ (symmetry) and if $\vdash_{\mathbf{R2}} A \leftrightarrow B$ and $\vdash_{\mathbf{R2}} B \leftrightarrow C$ then $\vdash_{\mathbf{R2}} A \leftrightarrow C$ (transitivity). Let $S = \{[A] \mid A \in \mathcal{W}^{\rightarrow}\}$. The operations on S are $\circ$,$\bullet$, $\vee$ and $\neg$, such that $[A] \circ [B] = [\neg(A \rightarrow \neg B)]$, $[A] \bullet [B] = [\neg(A \mapsto \neg B)]$, $\neg[A] = [\neg A]$ and $\bigoplus[A] = [\bigoplus A]$. Let the identity $e = [t]$.

We need to prove that $\mathfrak{S} = \langle S, \vee, \circ, \bullet, \neg, \bigoplus, e\rangle$ is a **R2**-structure and that we can construct a valuation function $v_{\mathfrak{S}}$ for which it is not the case that $e \leq v_{\mathfrak{S}}(G)$.

As a helpful lemma we prove that $\vdash_{\mathbf{R2}} A \rightarrow B$ iff $[A] \leq [B]$ in our structure $\mathfrak{S}$, where $a \leq b$ abbreviates that $b = a \vee b$.

First we prove the left to right direction. Suppose $\vdash_{\mathbf{R2}} A \to B$. This is sufficient to prove that $\vdash_{\mathbf{R2}} B \to (A \vee B)$ and $\vdash_{\mathbf{R2}} (A \vee B) \to B$. The first part is evident because it is an axiom of **R2**. $\vdash_{\mathbf{R2}} (A \vee B) \to B$ holds because of $\vdash_{\mathbf{R2}} (A \to B) \to ((B \to B) \to ((A \vee B) \to B))$ hold and we obtain $\vdash_{\mathbf{R2}} (A \vee B) \to B$ form this by applying Modus Ponens twice.

For the right to left direction: suppose $\vdash_{\mathbf{R2}} B \leftrightarrow (A \vee B)$. Hence we obtain $\vdash_{\mathbf{R2}} (A \vee B) \to B$ and $\vdash_{\mathbf{R2}} A \to (A \vee B)$. Applying transitivity using both results we obtain $\vdash_{\mathbf{R2}} A \to B$.

By means of the following properties of **R2**, we can prove that $\langle S, \vee, \circ, \neg, e\rangle$ is a De Morgan monoid (for proofs of these properties, see [2]).

$\vdash_{\mathbf{R2}} (t \circ A) \leftrightarrow A$
$\vdash_{\mathbf{R2}} ((A \circ B) \circ C) \leftrightarrow (A \circ (B \circ C))$
$\vdash_{\mathbf{R2}} (A \circ B) = (B \circ A)$
$\vdash_{\mathbf{R2}} A \to (A \circ A)$
$\vdash_{\mathbf{R2}} ((A \vee B) \vee C) \leftrightarrow (A \vee (B \vee C))$
$\vdash_{\mathbf{R2}} (A \vee B) \leftrightarrow (B \vee A)$
$\vdash_{\mathbf{R2}} (A \vee A) \leftrightarrow A$
$\vdash_{\mathbf{R2}} \neg\neg A \leftrightarrow A$
$\vdash_{\mathbf{R2}} (A \circ B) \to C$ iff $\vdash_{\mathbf{R2}} (A \circ \neg C) \to \neg B$, in view of $\vdash_{\mathbf{R2}} (A \to (B \to C)) \leftrightarrow (A \to (\neg C \to \neg B))$
$\vdash_{\mathbf{R2}} ((A \vee B) \wedge C) \leftrightarrow ((A \wedge C) \vee (B \wedge C))$
$\vdash_{\mathbf{R2}} (A \circ (B \vee C)) \leftrightarrow (A \circ B) \vee (A \circ C)$

Next, we prove that also $\vdash_{\mathbf{R2}} A \mapsto B$ iff $[A] \leq [B]$. This follows immediately from the fact that $\vdash_{\mathbf{R2}} A \to B$ iff $\vdash_{\mathbf{R2}} A \mapsto B$. The latter fact holds in view of axiom $(A \mapsto B) \to (A \to B)$ and rule $A \to B / A \mapsto B$.

By means of the following properties of **R2**, we can then prove that also $\langle S, \vee, \bullet, \neg, e\rangle$ is a De Morgan monoid (the proofs that the theorems (substituting $\mapsto$ everywhere by $\to$) below are **R**-theorems can be found in [2], the proofs for the $\mapsto$ versions proceed the same way, because **R2** has all the **R**-axioms for $\mapsto$; it should however be noted that some of the standard **R**-axioms are absent in our axiomatization of the $\mapsto$-arrow, however the missing axioms follow immediately from their counterparts with the $\to$-arrow).

$\vdash_{\mathbf{R2}} (t \bullet A) \leftrightarrow A$
$\vdash_{\mathbf{R2}} ((A \bullet B) \bullet C) \leftrightarrow (A \bullet (B \bullet C))$
$\vdash_{\mathbf{R2}} (A \bullet B) = (B \bullet A)$
$\vdash_{\mathbf{R2}} A \to (A \bullet A)$
$\vdash_{\mathbf{R2}} ((A \vee B) \vee C) \leftrightarrow (A \vee (B \vee C))$
$\vdash_{\mathbf{R2}} (A \vee B) \leftrightarrow (B \vee A)$
$\vdash_{\mathbf{R2}} (A \vee A) \leftrightarrow A$
$\vdash_{\mathbf{R2}} \neg\neg A \leftrightarrow A$
$\vdash_{\mathbf{R2}} (A \bullet B) \to C$ iff $\vdash_{\mathbf{R2}} (A \bullet \neg C) \to \neg B$, in view of $\vdash_{\mathbf{R2}} (A \mapsto (B \mapsto C)) \leftrightarrow (A \mapsto (\neg C \mapsto \neg B))$
$\vdash_{\mathbf{R2}} ((A \vee B) \wedge C) \leftrightarrow ((A \wedge C) \vee (B \wedge C))$
$\vdash_{\mathbf{R2}} (A \bullet (B \vee C)) \leftrightarrow (A \bullet B) \vee (A \bullet C)$

In order to establish that $\mathfrak{S}$ is an **R2**-structure, we also need to prove that the five other properties for being a **R2**-structure are satisfied by $\mathfrak{S}$:

(1) $a \circ b \leq a \bullet b$. This holds because the contraposition of $\vdash_{\mathbf{R2}} (A \mapsto \neg B) \to (A \to \neg B)$ gives $\vdash_{\mathbf{R2}} \neg(A \to \neg B) \to \neg(A \mapsto \neg B)$ or $\vdash_{\mathbf{R2}} (A \circ B) \to (A \bullet B)$.

(2) $a \leq \oplus a$. Ths holds because of $\vdash_{\mathbf{R2}} A \to \oplus A$
(3) $\neg\oplus\neg a \leq a$. This holds because of $\vdash_{\mathbf{R2}} \neg\oplus\neg A \to A$
(4) $\oplus e = e$. This holds because of $\vdash_{\mathbf{R2}} \oplus t \leftrightarrow t$
(5) $\oplus\neg e = \neg e$. This holds because of $\vdash_{\mathbf{R2}} \oplus\neg t \leftrightarrow \neg t$
Now, define a valuation function in such a way that $v_{\mathfrak{S}}(A) = [A]$.
We need to prove that

(S14) $v_{\mathfrak{S}}(t) = e$,
(S15) $v_{\mathfrak{S}}(A \vee B) = v_{\mathfrak{S}}(A) \vee v_{\mathfrak{S}}(B)$,
(S16) $v_{\mathfrak{S}}(A \to B) = \neg(v_{\mathfrak{S}}(A) \circ \neg v_{\mathfrak{S}}(B))$,
(S17) $v_{\mathfrak{S}}(A \mapsto B) = \neg(v_{\mathfrak{S}}(A) \bullet \neg v_{\mathfrak{S}}(B))$,
(S18) $v_{\mathfrak{S}}(\neg A) = \neg v_{\mathfrak{S}}(A)$,
(S19) $v_{\mathfrak{S}}(\oplus(A \vee B)) = v_{\mathfrak{S}}(\oplus A) \vee v_{\mathfrak{S}}(\oplus B)$,
(S20) $v_{\mathfrak{S}}(\oplus\neg(A \vee B)) = v_{\mathfrak{S}}(\oplus\neg A) \circ v_{\mathfrak{S}}(\oplus\neg B)$,
(S21) $v_{\mathfrak{S}}(\oplus(A \to B)) = \neg(v_{\mathfrak{S}}(\neg\oplus\neg A) \circ \neg v_{\mathfrak{S}}(\oplus B))$,
(S22) $v_{\mathfrak{S}}(\oplus\neg(A \to B)) = v_{\mathfrak{S}}(\oplus A) \circ v_{\mathfrak{S}}(\oplus\neg B)$,
(S23) $v_{\mathfrak{S}}(\oplus(A \mapsto B)) = \neg(v_{\mathfrak{S}}(\neg\oplus\neg A) \bullet \neg v_{\mathfrak{S}}(\oplus B))$,
(S24) $v_{\mathfrak{S}}(\oplus\neg(A \mapsto B)) = v_{\mathfrak{S}}(\oplus A) \bullet v_{\mathfrak{S}}(\oplus\neg B)$,
(S25) $v_{\mathfrak{S}}(\oplus\oplus A) = v_{\mathfrak{S}}(\oplus A)$
(S26) $v_{\mathfrak{S}}(\oplus\neg\oplus A) = v_{\mathfrak{S}}(\oplus\neg A)$.

Clause (S14) to (S18) follow immediately from the construction of our algebra. The other clauses follow immediately from the axioms for $\oplus$ and the fact that $[A] = [B]$ iff $\vdash_{\mathbf{R2}} A \leftrightarrow B$.

It only remains to prove that it is not the case that $e \leq v_{\mathfrak{S}}(G)$ and hence that not $e \leq [G]$. Suppose that the latter would be the case. Then $\vdash_{\mathbf{R2}} t \to G$, but $\vdash_{\mathbf{R2}} t$ in view of $\vdash_{\mathbf{R2}} (t \to t) \to t$ and $\vdash_{\mathbf{R2}} t \to t$, whence, by Modus Ponens $\vdash_{\mathbf{R2}} G$ would be the case, which is in contradiction with our assumption made at the start of the proof.

We can conclude that $\nvdash_{\mathbf{R2}} G$, which completes our proof.

8.3. Proof of Theorem 6 (Paradox-freedom)

Paradox 1: $\nvdash_{\mathbf{RR}} B \twoheadrightarrow (A \twoheadrightarrow B)$**.** $\nvdash_{\mathbf{R2}} \boxplus p \mapsto (\boxplus q \mapsto \oplus p)$
Assuming that $(\boxplus p \mapsto (\boxplus q \mapsto \oplus p)$ would be a theorem, $\mapsto$ would be an irrelevant implication, because $\boxplus q$ is completely independent from p. This is impossible because $\mapsto$ is an **R**-implication.

Paradox 2: $\nvdash \neg A \twoheadrightarrow (A \twoheadrightarrow B)$**.** Similar reasoning as Paradox 1.

Paradox 3: $\nvdash_{\mathbf{RR}} (\neg A \vee B) \twoheadrightarrow (A \twoheadrightarrow B)$**.** Suppose $\vdash_{\mathbf{RR}} (\neg p \vee q) \twoheadrightarrow (p \twoheadrightarrow q)$ would be the case. Then also $\vdash_{\mathbf{R2}} \boxplus(\neg p \vee q) \mapsto (\boxplus p \mapsto \oplus q)$ and $\vdash_{\mathbf{R2}} (\oplus p \to \boxplus q) \mapsto (\boxplus p \mapsto \oplus q)$. We can construct a model that gives the same value a to each of $\neg \boxplus p$, $\boxplus q$, $\neg\oplus p$ and $\oplus q$ but different values $b = \neg(a \circ a)$ and $c = \neg(a \bullet a)$ to resp. $\oplus p \to \boxplus q$ and $\boxplus p \mapsto \oplus q$. There is no reason to require that $b \leq c$, the **R2**-algebra only requires that $c \leq b$, not vice versa. This model is therefore a counterexample to $\vdash_{\mathbf{R2}} (\oplus p \to \boxplus q) \mapsto (\boxplus p \mapsto \oplus q)$.

Paradox 4: $\nvdash_{\mathbf{RR}} \neg(A \twoheadrightarrow B) \twoheadrightarrow (A \wedge \neg B)$**.** If this paradox would be derivable, paradox 3 would also be derivable in view of the validity of contraposition for $\twoheadrightarrow$ and the inter-definability of disjunction and conjunction.

Paradox 5: $\nvdash_{\mathbf{RR}} \neg(A \twoheadrightarrow B) \twoheadrightarrow (B \twoheadrightarrow A)$**.** If this would be the case then $\neg(\boxplus p \mapsto \diamondplus q) \mapsto (\boxplus p \mapsto \diamondplus q)$ would be a **R2**-theorem. Given the **R2**-theorem $(\neg A \to A) \to A$, this would entail that $\boxplus p \mapsto \diamondplus q$ is a **R2**-theorem, but we can construct a model that assigns entirely independent values to $\boxplus p$ and $\diamondplus q$.

Paradox 6: $\nvdash_{\mathbf{RR}} (A \twoheadrightarrow B) \vee (B \twoheadrightarrow A)$**.** We can construct a **R2**-counter-model for $(\boxplus p \mapsto \diamondplus q) \vee (\boxplus q \mapsto \diamondplus p)$. Assign two completely different values a to both $\boxplus p$ and $\diamondplus p$ and b to both $\diamondplus q$ and $\boxplus q$. Assign another value $c = \neg(a \circ \neg b)$ to $\boxplus p \mapsto \diamondplus q$ and another value $d = \neg(\neg a \circ b)$ to $\boxplus q \mapsto \diamondplus p$. There is no reason why the join of c and d would be greater than or equal to the identity element t.

Paradox 7: $\nvdash_{\mathbf{RR}} (A \twoheadrightarrow B) \vee (A \twoheadrightarrow \neg B)$**.** We can construct a **R2**-counter-model for $(\boxplus p \mapsto \diamondplus q) \vee (\boxplus p \mapsto \diamondplus \neg q)$. Assign the same value a to both $\diamondplus p$ and $\boxplus p$ and a different value b to both $\diamondplus q$ and $\boxplus q$. Although the join of $\neg a$ and a is greater than e, there is no reason why the join of the independent $\neg(a \bullet \neg b)$ and $\neg(\neg a \bullet \neg b)$ would be greater than e.

Paradox 8: $\nvdash_{\mathbf{RR}} (A \twoheadrightarrow B) \vee (\neg A \twoheadrightarrow B)$**.** Similar to paradox 7.

Paradox 9: $\nvdash_{\mathbf{RR}} ((A \wedge B) \twoheadrightarrow C) \twoheadrightarrow ((A \twoheadrightarrow C) \wedge (B \twoheadrightarrow C))$**.** $\nvdash_{\mathbf{RR}} ((A \wedge B) \twoheadrightarrow C) \twoheadrightarrow ((A \twoheadrightarrow C) \wedge (B \twoheadrightarrow C))$ in view of

$\nvdash_{\mathbf{R2}} ((\diamondplus A \circ \diamondplus B) \mapsto \boxplus C) \mapsto ((\boxplus A \mapsto \diamondplus C) \circ (\boxplus B \mapsto \diamondplus C))$. This last expression is the case because

$\nvdash_{\mathbf{R2}} ((A \circ B) \mapsto C) \to ((A \mapsto C) \circ (B \mapsto C))$ (otherwise $\mapsto$ would not be an **R**-implication).

Paradox 10: $\nvdash_{\mathbf{RR}} ((A \twoheadrightarrow B) \twoheadrightarrow A) \twoheadrightarrow A$**.** Suppose $\vdash_{\mathbf{RR}} ((A \twoheadrightarrow B) \twoheadrightarrow A) \twoheadrightarrow A$. Then also $\vdash_{\mathbf{R2}} ((\boxplus p \mapsto \diamondplus q) \mapsto \boxplus p) \mapsto \diamondplus p$. One can construct a counter-model by assigning the same value a to $\diamondplus p$ and $\boxplus p$ and value b to $\diamondplus q$. In regular De Morgan monoids there is no reason why $((a \to b) \to a) \to a$, where $c \to b =_{df} \neg(a \circ \neg b)$, would be at least as great as the identity e. The same holds in **R2**-structures, where $\bullet$ behaves exactly like $\circ$ in a regular De Morgan monoid.

Paradox 11: $\nvdash_{\mathbf{RR}} A \twoheadrightarrow (\neg A \twoheadrightarrow A)$**.** Suppose $\vdash_{\mathbf{RR}} A \twoheadrightarrow (\neg A \twoheadrightarrow A)$. Hence $\vdash_{\mathbf{R2}} \boxplus p \mapsto (\boxplus \neg p \mapsto \diamondplus p)$. We construct a counter-model by assigning the value a to both $\diamondplus p$ and $\boxplus p$. This makes $\neg a$ the value of $\boxplus \neg p$. There is no reason why $a \bullet \neg a$ would be at least as small as a and so $(\boxplus p \bullet \boxplus \neg p) \mapsto \diamondplus p$ is not verified by this model.

Paradox 12: $\nvdash_{\mathbf{RR}} (A \wedge \neg A) \twoheadrightarrow B$**.** Suppose $\vdash_{\mathbf{RR}} (A \wedge \neg A) \twoheadrightarrow B$. Then also $\vdash_{\mathbf{R2}} (\boxplus p \wedge \boxplus \neg p) \mapsto \diamondplus q$. Construct a counter-model by assigning the value a to both $\boxplus p$ and $\diamondplus p$ and the value b to $\diamondplus q$. Although $a \wedge \neg a$ is at least as small as $\neg e$, this does not make $(a \wedge \neg a) \bullet \neg b$ at least as small as $\neg e$. So there is no reason why in this algebra $\neg((a \wedge \neg a) \bullet \neg b)$ would be at least as great as e, which means that this model does not verify $(\boxplus p \wedge \boxplus \neg p) \mapsto \diamondplus q$.

Paradox 13: $\nvdash_{\mathbf{RR}} B \twoheadrightarrow (A \vee \neg A)$**.** If this paradox would be derivable, paradox 12 would also be derivable in view of the validity of contraposition for $\twoheadrightarrow$ and the interdefinability of disjunction and conjunction.

Paradox 14: $\nvdash_{\mathbf{RR}} \neg(A \twoheadrightarrow B) \twoheadrightarrow A$. Suppose $\vdash_{\mathbf{RR}} \neg(A \twoheadrightarrow B) \twoheadrightarrow A$. Then also $\vdash_{\mathbf{R2}} (\boxplus p \bullet \boxplus\neg q) \mapsto \diamondplus p$. Construct a counter-model by assigning the value a to both $\boxplus p$ and $\diamondplus p$ and the value b to $\boxplus\neg q$. Although $a \wedge b$ is at least as small as a, this does not make $a \bullet b$ at least as small as a. So there is no reason why this model would verify $(\boxplus p \wedge \boxplus\neg q) \mapsto \diamondplus p$.

Paradox 15: $\nvdash_{\mathbf{RR}} \neg(A \twoheadrightarrow B) \twoheadrightarrow \neg B$. Similar to Paradox 14.

8.4. Proof of Theorem 5

First we recursively define what a positive and a negative part of a formula is. For the basic case: A is a positive part of A. For the resursion: $\neg B$ is a positive part of B iff B is a negative part of A. If $C \vee D$ is a positive part of A, then C and D are both positive parts of A. If $C \vee D$ is a negative part of A, then C, D are both negative parts of A. If $C \Rrightarrow D$ is a positive part of A, where $\Rrightarrow \in \{\rightarrow, \twoheadrightarrow, \mapsto\}$, then C is a negative part of A and D is a positive part of A. If $C \Rrightarrow D$ is a negative part of A, where $\Rrightarrow \in \{\rightarrow, \twoheadrightarrow, \mapsto\}$, then C is a positive part of A and D is a negative part of A. Finally, if $\diamond A$ is a positive/negative part of A then A is a positive/negative part of A. Let $pp(B, A)$ and $np(B, A)$ abbreviate that B is a positive resp. a negative part of A.

We prove that all **R**-theorems can be translated into theorems of **RR**. Take A to be a theorem of **R**. Change every occurrence of $\rightarrow$ in A to $\mapsto$ (for short $A_{\rightarrow/\mapsto}$) and call this B. This gives us

$$B = A_{\rightarrow/\mapsto}$$

Given that we defined $\mapsto$ as an arrow that behaves exactly like $\rightarrow$ (see $R3$), B will be an **R2**-theorem. And, because of $A35$, if B is an **R2**-theorem, so is $\diamondplus B$.

Now comes the tricky part. We prove that C, defined as the result of changing all occurrences of $\mapsto$ in B to $\twoheadrightarrow$, is equivalent to $\diamondplus B$: if

$$C = B_{\mapsto/\twoheadrightarrow}$$

then

$$C \leftrightarrow \diamondplus B$$

Note that, if A does not contain any $\rightarrow$-arrows, B will be equal to A and thus C to B and C will evidently be a theorem of **RR**. We thus presume that A contains at least one $\rightarrow$-arrow and thus B at least one $\mapsto$-arrow.

We can characterise B as a wff in which only $\mapsto$, $\vee$ and $\neg$ occur as connectives. The same thus holds for $\diamondplus B$. When we look more closely at the influence of the $\diamondplus$-operator, we can furthermore see that it spreads out over the subwffs of B in such a way that all positive parts of B are given a $\diamondplus$ and all negative parts of B are given a $\boxplus$. This is in agreement with $A23$-$A33$:

$$\diamondplus(p \mapsto q) \leftrightarrow \boxplus p \mapsto \diamondplus q$$

$$\diamondplus(p \vee q) \leftrightarrow \diamondplus p \vee \diamondplus q$$

$$\diamondplus\neg p \leftrightarrow \neg \boxplus p$$

We can thus state the following:

Lemma 2. *$\diamondplus B$ is **R2**-equivalent to a formula D defined as the wff resulting from substituting all positive parts P of B by $\diamondplus P$ and substituting all negative parts N of B by $\boxplus N$. Then, for any subwff S of B:*

if

$$pp(S, B)$$

then

$$\diamondplus S \textit{ is a subwff of } D$$

and if

$$np(S, B)$$

then

$$\boxplus S \textit{ is a subwff of } D$$

Given this, we now prove that if

$$C = B_{\mapsto/\twoheadrightarrow}$$

then

$$\vdash_{\mathbf{R_2}} C \leftrightarrow \diamondplus B$$

by mathematical induction. Since B only had $\mapsto$, $\vee$ and $\neg$ as connectives, C will only have $\twoheadrightarrow$, $\vee$ and $\neg$. Let n denote the depth of embedding of the connective in question.

Base case: $n = 1$: $B = D \mapsto E$, where D and E are implication-free. Then

$$C = D \twoheadrightarrow E$$

From $D4$ we can easily see that

$$(D \twoheadrightarrow E) \ \leftrightarrow \diamondplus(D \mapsto E)$$

Thus

$$C \leftrightarrow \diamondplus B$$

Note that since B must contain at least one $\twoheadrightarrow$, this is the only possibility.

Induction step: Assume that the equality holds to depth n. We now show that it holds for a connective on depth $n + 1$ as well. Take D (and E) to be the subwffs on the $n + 1$-depth. There are thus three possible subwffs of depth n: $D \twoheadrightarrow E$, $D \vee E$ and $\neg D$. Given that the equality holds to depth n, the subwffs will have gotten either a $\diamondplus$ or a $\boxplus$ operator, depending on whether it was a positive or negative part of B. We consider both possibilities for each of the possible subwffs:

If $D \twoheadrightarrow E$ was a positive part of B, then it received a $\diamondplus$. The analysis of the $\twoheadrightarrow$-arrow then delivers a $\boxplus$ to D and a $\diamondplus$ to E, because of the **R2**-axioms. This is in agreement with lemma 2. Since $D \twoheadrightarrow E$ was a positive part of B, D (a negative part of a positive part of B) will be a negative part of B and thus receive a $\boxplus$. And since E is a positive part of a positive part of B, it will also be a positive part of B and receive a $\diamondplus$.

If $D \twoheadrightarrow E$ was a negative part of B, it received a $\boxplus$. The analysis of the $\twoheadrightarrow$-arrow then delivers a $\diamondplus$ to D and a $\boxplus$ to E. This is in agreement with lemma 2. Since D is a negative part of a negative part of B, it will be a positive part of B and thus receive a $\diamondplus$. And E, a positive part of a negative part of B and thus a negative part of B, will get a $\boxplus$.

If $\neg D$ was a positive part of B, then it received a $\diamondplus$. The analysis of the $\neg$-connective then delivers a $\boxplus$ to D, since $\diamondplus\neg D \leftrightarrow \neg \boxplus D$. This is in agreement with lemma 2. Since $\neg D$ was a positive part of B, D (a negative part of a positive part of B) will be a negative part of B and thus receive a $\boxplus$.

If $\neg D$ was a negative part of B, then it received a $\boxplus$. The analysis of the $\neg$-connective then delivers a $\diamondplus$ to D, since $\boxplus\neg D \leftrightarrow \neg\diamondplus D$. This is in agreement with lemma 2. Since $\neg D$ was a negative part of B, D (a negative part of a negative part of B) will be a negative part of B and thus receive a $\boxplus$.

If $D \vee E$ was a positive part of B, then it received a $\diamondplus$. The analysis of the $\vee$-connective then delivers a $\diamondplus$ to D and a $\diamondplus$ to E, because of $A24$. This is in agreement with lemma 2. Since $D \vee E$ was a positive part of B, D (a positive part of a positive part of B) will also be a positive part of B and thus receive a $\diamondplus$. The same holds for E.

If $D \vee E$ was a negative part of B, then it received a $\boxplus$. The analysis of the $\vee$-connective following $A23$ (presuming that we account for the double negation) then delivers a $\diamondplus$ to $\neg D$ and thus a $\boxplus$ to D (see the previous case). It also delivers a $\boxplus$ to E. This is in agreement with lemma 2. Since $D \vee E$ was a negative part of B, D (a positive part of a negative part of B) will also be a negative part of B and thus receive a $\boxplus$. The same holds for E.

We conclude that for any occurring connective, no matter the depth, for every theorem of **R** A, if

$$B = A_{\rightarrow/\mapsto}$$

and

$$C = B_{\twoheadrightarrow/\mapsto}$$

then

$$C \leftrightarrow \diamondplus B.$$

Therefore, because $\vdash_{\mathbf{R}} A$, we obtain $\vdash_{\mathbf{R2}} B$ and by $\vdash_{\mathbf{R2}} \Diamond A \rightarrow A$, $\diamondplus B$ is a **R2**-theorem. Hence, C is a **R2**-theorem and moreover a **RR**-theorem. This concludes the proof that every **R**-theorem can be translated into a **RR**-theorem, by substituting $\rightarrow$ everywhere into $\twoheadrightarrow$.

8.5. Proof of Lemma 1

This proof of course relies on notions (the $*$-complement operator, α and β formulas, lines of the form $[A_1, A_2, \ldots, A_n]B$, etc.) related to goal directed proofs as defined and employed in [6]. The reader should consult this paper in order to be able to understand this proof.

Where $\Gamma \cup \{A\} \subset \mathcal{W}$, if a line

$$[A_1, A_2, \ldots A_n]A$$

can occur in a goal directed proof from Γ then

$$\Gamma^{\blacksquare} \vdash_{\mathbf{RR}} (\blacklozenge A_1 \circ \blacklozenge A_2 \circ \ldots \circ \blacklozenge A_n) \rightarrow B$$

where $B = \blacklozenge A$ if A is the goal formula, and $B = \blacksquare A$ otherwise.

We prove that all the rules of the goal directed proof-system hold in **RR**. Given that the proof-system works via a procedure, which restricts the applicability of the rules, this suffices to ensure that we have at least classical strength for **RR**. Take Δ to be $[\blacklozenge A_1 \circ \blacklozenge A_2 \circ \ldots \circ \blacklozenge A_n]$.

Goal-rule $\quad [G]\, G$

If we translate this into **RR**, we get $\blacklozenge A \rightarrow \blacklozenge A$. From A1, we immediately see that this rule holds in **RR**.

Formula-analysing rule: α $$\frac{[\Delta]\,\alpha}{[\Delta]\,\alpha_1 \quad [\Delta]\,\alpha_2}$$

If we translate this into **RR**, we get $(\blacklozenge A_1 \circ \blacklozenge A_2 \circ \ldots \circ \blacklozenge A_n) \rightarrow \blacksquare \alpha$. From this, $A5$ and $A6$ allow us to derive both $(\blacklozenge A_1 \circ \blacklozenge A_2 \circ \ldots \circ \blacklozenge A_n) \rightarrow \blacksquare \alpha_1$ and $(\blacklozenge A_1 \circ \blacklozenge A_2 \circ \ldots \circ \blacklozenge A_n) \rightarrow \blacksquare \alpha_2$. This rule thus holds in **RR**.

Formula-analysing rule: β $$\frac{[\Delta]\,\beta}{[\Delta \cup \{*\beta_2\}]\,\beta_1 \quad [\Delta \cup \{*\beta_1\}]\,\beta_2}$$

If we translate the first line into **RR**, we get $(\blacklozenge A_1 \circ \blacklozenge A_2 \circ \ldots \circ \blacklozenge A_n) \rightarrow \blacksquare \beta$. From this translation, $T2$ and $T4$ we can derive both $(\blacklozenge A_1 \circ \blacklozenge A_2 \circ \ldots \circ \blacklozenge A_n \circ \blacklozenge \neg \beta_1) \rightarrow \blacksquare \beta_2$ and $(\blacklozenge A_1 \circ \blacklozenge A_2 \circ \ldots \circ \blacklozenge A_n \circ \blacklozenge \neg \beta_2) \rightarrow \blacksquare \beta_1$. This rule thus holds in **RR**.

Notice that we have presupposed that neither the α- nor the β-formula were the goalformula, since they immediately got a $\blacksquare$. This has to do with the procedure of the goal directed proofs. This ensures that a formula-analysing rule is never applied to a goal-descendent, which is roughly put any formula that has the main goal as its formula element. This corresponds to the *where* $B = \blacklozenge A$ *if* A *is the goal formula, and* $B = \blacksquare A$ *otherwise*-part of the translation.

Condition-analysing rule: α $$\frac{[\Delta \cup \{\alpha\}]\,A}{[\Delta \cup \{\alpha_1, \alpha_2\}]\,A}$$

If we translate the first line into **RR**, we get $(\blacklozenge A_1 \circ \blacklozenge A_2 \circ \ldots \circ \blacklozenge A_n \circ \blacklozenge \alpha) \rightarrow \blacksquare A$. Given that $\blacklozenge(A \wedge B) \leftrightarrow \blacklozenge A \circ \blacklozenge B$, we can derive $(\blacklozenge A_1 \circ \blacklozenge A_2 \circ \ldots \circ \blacklozenge A_n \circ \blacklozenge \alpha_1 \circ \blacklozenge \alpha_2) \rightarrow \blacksquare A$. This rule thus holds in **RR**. Note that, since we do not know whether A is the goal formula or not, we proved it for the stronger $\blacksquare A$, but simply replacing $\blacksquare A$ by $\blacklozenge A$ results in the proof for the case where A is the goal formula. The same holds for the reasoning with respect to the following rule.

Condition-analysing rule: β $$\frac{[\Delta \cup \{\beta\}]\,A}{[\Delta \cup \{\beta_1\}]\,A \quad [\Delta \cup \{\beta_2\}]\,A}$$

If we translate the first line into **RR**, we get $(\blacklozenge A_1 \circ \blacklozenge A_2 \circ \ldots \circ \blacklozenge A_n \circ \blacklozenge \beta) \rightarrow \blacksquare A$. Given $A24$, $A8$ and $A9$, we can derive $(\blacklozenge A_1 \circ \blacklozenge A_2 \circ \ldots \circ \blacklozenge A_n \circ \blacklozenge \beta_1) \rightarrow \blacksquare A$ and $(\blacklozenge A_1 \circ \blacklozenge A_2 \circ \ldots \circ \blacklozenge A_n \circ \blacklozenge \beta_2) \rightarrow \blacksquare A$. This rule thus holds in **RR**.

Transitivity $$\frac{\begin{array}{c}[\Delta \cup \{B\}]\,A \\ [\Delta']\,B\end{array}}{[\Delta \cup \Delta']\,A}$$

If we translate the first line into **RR**, we get $(\blacklozenge A_1 \circ \blacklozenge A_2 \circ \ldots \circ \blacklozenge A_n \circ \blacklozenge B) \rightarrow \blacklozenge A$. The second line becomes $(\blacklozenge A'_1 \circ \blacklozenge A'_2 \circ \ldots \circ \blacklozenge A'_n) \rightarrow \blacksquare B$ (since B is definitely not the goalformula) . If we apply contraposition ($A12$) to both wffs, we can then apply $A2$ to eliminate B from the implications and derive $(\blacklozenge A_1 \circ \blacklozenge A_2 \circ \ldots \circ \blacklozenge A_n \circ \blacklozenge A'_1 \circ \blacklozenge A'_2 \circ \ldots \circ \blacklozenge A'_n) \rightarrow \blacklozenge A$ by contraposition. This rule thus holds in **RR**.

Excluded Middle $$\frac{\begin{array}{c}[\Delta \cup \{B\}]\,A \\ [\Delta' \cup \{*B\}]\,A\end{array}}{[\Delta \cup \Delta']\,A}$$

If we translate the first line into **RR**, we get

$$1.(\blacklozenge A_1 \circ \blacklozenge A_2 \circ \ldots \circ \blacklozenge A_n \circ \blacklozenge B) \rightarrow \blacksquare A$$

Because of $DM3$, this is equal to

$$2.(\blacklozenge B \circ \blacklozenge A_1 \circ \blacklozenge A_2 \circ \ldots \circ \blacklozenge A_n) \rightarrow \blacksquare A$$

This is furthermore equivalent to

$$3.\blacklozenge B \rightarrow ((\blacklozenge A_1 \circ \blacklozenge A_2 \circ \ldots \circ \blacklozenge A_n) \rightarrow \blacksquare A)$$

since

$$(A \rightarrow (B \rightarrow C)) \leftrightarrow (A \circ B) \rightarrow C.$$

The second line of the EM-rule translated into **RR** becomes

$$4.(\blacklozenge A'_1 \circ \blacklozenge A'_2 \circ \ldots \circ \blacklozenge A'_n \circ \blacklozenge \neg B) \rightarrow \blacksquare A$$

We can use the same principle we used on 3 to derive

$$5.(\blacklozenge A'_1 \circ \blacklozenge A'_2 \circ \ldots \circ \blacklozenge A'_n) \rightarrow (\blacklozenge \neg B \rightarrow \blacksquare A)$$

This then gives us, by means of $A12$

$$6.(\blacklozenge A'_1 \circ \blacklozenge A'_2 \circ \ldots \circ \blacklozenge A'_n) \rightarrow (\neg \blacksquare A \rightarrow \neg \blacklozenge \neg B)$$

or

$$7.(\blacklozenge A'_1 \circ \blacklozenge A'_2 \circ \ldots \circ \blacklozenge A'_n) \rightarrow (\blacklozenge \neg A \rightarrow \blacksquare B).$$

Because of $A34$ and $A35$, we can derive

$$8.(\blacklozenge A'_1 \circ \blacklozenge A'_2 \circ \ldots \circ \blacklozenge A'_n) \rightarrow (\blacklozenge \neg A \rightarrow \blacklozenge B)$$

and by

$$(A \rightarrow (B \rightarrow C)) \leftrightarrow (A \circ B) \rightarrow C$$

we can transform this to

$$9.(\blacklozenge A'_1 \circ \blacklozenge A'_2 \circ \ldots \circ \blacklozenge A'_n \circ \blacklozenge \neg A) \rightarrow \blacklozenge B.$$

We can now apply transitivity ($A2$) to 3 and 9 to derive

$$10.(\blacklozenge A'_1 \circ \blacklozenge A'_2 \circ \ldots \circ \blacklozenge A'_n \circ \blacklozenge \neg A) \rightarrow ((\blacklozenge A_1 \circ \blacklozenge A_2 \circ \ldots \circ \blacklozenge A_n) \rightarrow \blacksquare A).$$

This is equivalent to

$$11.(\blacklozenge A'_1 \circ \blacklozenge A'_2 \circ \ldots \circ \blacklozenge A'_n \circ \blacklozenge \neg A \circ \blacklozenge A_1 \circ \blacklozenge A_2 \circ \ldots \circ \blacklozenge A_n) \rightarrow \blacksquare A$$

and thus to

$$12.(\blacklozenge A'_1 \circ \blacklozenge A'_2 \circ \ldots \circ \blacklozenge A'_n \circ \blacklozenge A_1 \circ \blacklozenge A_2 \circ \ldots \circ \blacklozenge A_n) \rightarrow (\blacklozenge \neg A \rightarrow \blacksquare A)$$

and to

$$13.(\blacklozenge A'_1 \circ \blacklozenge A'_2 \circ \ldots \circ \blacklozenge A'_n \circ \blacklozenge A_1 \circ \blacklozenge A_2 \circ \ldots \circ \blacklozenge A_n) \rightarrow (\neg \blacksquare A \rightarrow \blacksquare A).$$

Because of $\vdash_{\mathbf{R2}} (\neg A \rightarrow A) \rightarrow A$, this gives us

$$14.(\blacklozenge A'_1 \circ \blacklozenge A'_2 \circ \ldots \circ \blacklozenge A'_n \circ \blacklozenge A_1 \circ \blacklozenge A_2 \circ \ldots \circ \blacklozenge A_n) \rightarrow \blacksquare A$$

which is what the EM-rule tells us. This rule thus holds in **RR**.

Since all the rules of the goal directed proof-system hold in **RR**, we can transform any proof of the goal directed system into a proof in **RR**. Any Classical Logic consequence for a consistent premise set thus holds in **RR**.