Relevance Logic

UC Berkeley, Philosophy 142, Spring 2016 — John MacFarlane

1 Motivations for relevance logic

Robert Meyer [9, pp. 812–816] gives five arguments for relevance logic.

Paradoxes of implication are counterintuitive

> ...the claim that '$q \& \sim q$' entails 'p', in general, signals a breakdown in our intuitions not different in kind, though different perhaps in severity, from the kind of breakdown whose result is outright inconsistency... [9, p. 812]

Q: Do we *have* intuitions about validity? About what follows from what? (We'll revisit this later.)

Symmetry between overdetermination and underdetermination A theory may go wrong in either of two ways with respect to a pair of sentences A, $\neg A$. It may fail to tell us which is true (underdetermination), or it may tell us that both are true (overdetermination). Both of these are failures of discrimination; in both cases, we get no useful information about A:

> Thus, when a definite interpretation is in mind, there is an intuitive parity between underdetermination and overdetermination with respect to a given A; in either case, we get no usable information about A. [9, p. 812]

But given classical logic there is a huge difference: overdetermination trivializes our theory, while underdetermination is no big problem.

No reason to suppose mathematics is consistent There's no good reason to assume that mathematics must be consistent. If math is about a supersensible realm of objects, why should we assume they're like ordinary empirical objects with respect to consistency? But if math is a free human creation, why can't it be inconsistent?

> ...for certain purposes an inconsistent system might be more useful, more beautiful, and even—at the furthest metaphysical limits—as the case may be, more accurate. [9, p. 814]

Classical logic *forces* math to be consistent or trivial.

Contradictory beliefs and conflicting obligations People sometimes have contradictory beliefs. If they are committed to the logical consequences of their beliefs (as seems plausible), then according to classical logic, they are committed to everything! This is implausible.

People are sometimes subject to conflicting obligations (e.g., to tell Jim about his wife, because he is your friend, and not to tell him, because the information was given to you in confidence). Plausibly, "one is obligated to bring about what follows logically from what one is obligated to bring about." In cases of conflicting obligation, then, according to classical logic one is obligated to bring about every state of affairs! This is implausible.

Classical logic goes with a bad theory of the conditional (The material conditional, of course.) I'll skip this one, because it gets into issues we won't be able to deal with well until later. Many people nowadays think that natural language indicative and counterfactual conditionals shouldn't be expected to indicate *or* to express entailment. So the debate about relevance has moved away from conditionals and is better thought of as purely a debate about entailment.

However, the distinction he makes between a conditional that *expresses* entailment and one that *indicating* entailment is a useful one.

$\rightarrow$ *expresses* implication: $A \rightarrow B$ means that A logically implies B. *expresses*

$\rightarrow$ *indicates* implication: $A \rightarrow B$ is logically true iff A logically implies B. *indicates*

The material conditional *indicates* classical implication but does not express it.

2 The Lewis Argument

This argument appears in C. I. Lewis and C. H. Langford's *Symbolic Logic* [6, pp. 248–51]. Anderson and Belnap note that the argument was known to Albert of Saxony and other medieval logicians:

1	$P \wedge \neg P$	
2	P	(1, $\wedge$ Elim)
3	$\neg P$	(1, $\wedge$ Elim)
4	$P \vee Q$	(2, $\vee$ Intro)
5	Q	(3, 4, Disjunctive Syllogism)

The argument establishes the validity of the inference from $P \wedge \neg P$ to Q: an inference called *explosion* or *ex falso quodlibet*. So, if we want to reject explosion, we need to reject one of the assumptions of this argument. *explosion* *ex falso quodlibet*

No one wants to reject $\wedge$ elimination. These options have all been tried:

(a) Reject $\lor$ Intro (a.k.a. "disjunctive weakening")
(b) Reject the transitivity of entailment
(c) Reject Disjunctive Syllogism

3 Rejecting disjunctive weakening

How can relevance be represented formally? *Variable sharing.*

Parry, "analytic implication": the conclusion may not contain any variables not contained in the premises. Compare Kant's definition of an analytic judgment as one in which the predicate is *contained in* the subject.

Note that P does not analytically imply $P \lor Q$.

Disadvantages:

- Disjunctive weakening is an absolutely crucial rule!
- The "containment" explication of analyticity forces us to reject disjunctive definitions. For example, if "all husbands are spouses" is to be analytic, we must define "husband" as "male spouse" rather than defining "spouse" as "husband or wife."
- What could be more relevant to $P \lor Q$ than P?

Parry's view is discussed in [2, §29.6.1].

4 Rejecting transitivity

Motivation: In every single step of the Lewis argument, the premises seem relevant to the conclusion. It is only when we chain these together and conclude that (1) entails (5) that we get problems. So, maybe,

- (1) entails (2)
- (1) entails (3)
- (2) entails (4)
- (3) and (4) together entail (5)
- But (1) does not entail (5)

Can we describe a system that works this way? We want to exclude arguments with contradictory premises or tautologous conclusions. But not *all* such arguments! A relevantist still wants $P \land \neg P$ to entail $\neg P$, for example. But perhaps this is okay only because it is a substitution instance of a valid argument without a contradictory premise, namely: $P \land Q$, therefore Q.

Proposal: An argument is *valid* iff it is a substitution instance of an argument that

- is classically valid,
- does not have a contradictory premise, and
- does not have a tautologous conclusion.

Advantages:

- allows that each step of the Lewis argument is valid
- rejects very little classical reasoning (indeed, Neil Tennant has proved that in almost every case where there is a non-relevant proof, there will be a relevant one)

Disadvantages:

- What about the argument from $(P \wedge \neg P) \vee R$ to $Q \vee R$? This is classically valid, the premise is not contradictory, and the conclusion is not tautologous. Yet it seems objectionable from a relevance point of view. (Consider how you'd prove it.)

- Transitivity of entailment seems fundamental!

> Any criterion according to which entailment is not transitive, is *ipso facto* wrong. It seems in fact incredible that anyone should admit that B follows from A, and that C follows from B, but feel that some further argument was required to establish that A entails C. [2, p. 154]

The idea of rejecting transitivity was first proposed by Timothy Smiley [12].

5 Rejecting Disjunctive Syllogism

We'll focus on Belnap and Anderson's logic of first-degree entailment (called $\mathbf{E_{fde}}$), which is equivalent to a logic devised by Wilhelm Ackermann [1, pp. 113–128].

Exercise:

1. How would you show that Disjunctive Syllogism is valid using our introduction and elimination rules for $\vee$ and $\neg$?

Because we can always combine (a finite number of) premises into a single conjunction, we'll consider here only one-premise arguments, which we'll call "entailments." The goal will be to define a notion of "tautological entailment" that captures just the ones that are relevantly valid in virtue of their propositional forms.

An *atom* is a propositional constant or its negation:

- atoms: P, $\neg P$, Q
- not: $P \vee Q$, $P \wedge \neg R$

A *primitive conjunction* is a conjunction of atoms. A *primitive disjunction* is a disjunction of atoms.

- primitive conjunctions: $P \wedge \neg P \wedge Q$, $P \wedge Q \wedge R \wedge \neg S$

- primitive disjunctions: $P \vee Q \vee \neg P$, $P \vee R$
- neither: $(P \wedge Q) \vee R$, $Q \wedge (P \vee \neg R)$

$\phi \Rightarrow \psi$ is a *primitive entailment* if ϕ is a primitive conjunction and ψ a primitive disjunction.

- example: $P \wedge \neg P \wedge Q \Rightarrow P \vee R$
- not: $P \wedge \neg P \wedge Q \Rightarrow P \vee (R \wedge S)$

A primitive entailment $\phi \Rightarrow \psi$ is *explicitly tautological* if some (conjoined) atom of ϕ is identical with some (disjoined) atom of ψ.

- examples: $P \wedge \neg P \wedge Q \Rightarrow P \vee R$, $\neg P \wedge \neg Q \wedge \neg R \Rightarrow S \vee \neg Q$
- not: $P \wedge \neg P \Rightarrow Q$

This captures a certain kind of "containment" of conclusion in premises.

What about entailments that are *not* primitive entailments, like $P \vee Q \Rightarrow P \vee \neg(R \wedge \neg R)$? How do we test them?

1. Put the premise into *disjunctive normal form* (i.e., convert it into a disjunction of primitive conjunctions) and the conclusion into *conjunctive normal form* (i.e., a conjunction of primitive disjunctions). You should now have something of the form $\phi_1 \vee \phi_2 \vee \ldots \vee \phi_n \Rightarrow \psi_1 \wedge \psi_2 \wedge \ldots \wedge \psi_n$, where each ϕ_i is a primitive conjunction and each ψ_i is a primitive disjunction.

2. The entailment is a *tautological entailment* iff for every ϕ_i and ψ_j, $\phi_i \Rightarrow \psi_j$ is explicitly tautological.

This procedure depends on the fact that every entailment can be put into normal form using relevantly acceptable rules. Here is the algorithm:

1. Apply DeMorgan's laws and Double Negation Elimination to drive negations inward as far as possible.

2. Use Distribution and Commutation to move all disjunction signs outside conjunctions (for disjunctive normal form) or inside conjunctions (for conjunctive normal form).

3. Use Association to group things to the left. (Remember, $\psi_1 \wedge \psi_2 \wedge \psi_3$ is short for $(\psi_1 \wedge \psi_2) \wedge \psi_3$.)

This can always be done, and although there is not a unique normal form for each entailment, it can be proved that if one normal form passes the test, they all do.

Let's try a simple example:

$$P \vee Q \Rightarrow P \vee \neg(R \wedge \neg R)$$

The premise is already in DNF so we can leave it.

We need to put the conclusion in CNF. First, drive in the negation:

$$P \vee (\neg R \vee \neg\neg R)$$

$$P \vee (\neg R \vee R)$$

Now associate:

$$(P \vee \neg R) \vee R$$

which is

$$P \vee \neg R \vee R$$

We're there! (Note: it's a conjunction of a single disjunction.)

Now we ask:

- Is $P \Rightarrow P \vee \neg R \vee R$ explicitly tautological? YES
- Is $Q \Rightarrow P \vee \neg R \vee R$ explicitly tautological? NO

So it's not a tautological entailment. Try disjunctive syllogism!

Exercises:

2. Using De Morgan's laws, Double Negation Elimination, Commutation, Association, and Distribution, put the following sentences into *both* disjunctive normal form and conjunctive normal form. Show your work.

   (a) $(P \wedge Q) \vee (P \wedge \neg Q)$

   (b) $P \wedge Q \wedge \neg Q$

   (c) $\neg(P \vee Q) \wedge \neg(P \vee \neg Q)$

3. Use the procedure described above to determine whether the following classical entailments are tautological entailments:

   (a) $P \wedge \neg Q \Rightarrow P$

   (b) $P \Rightarrow (P \wedge Q) \vee (P \wedge \neg Q)$

   (c) $\neg(P \vee \neg Q) \Rightarrow \neg(P \vee (P \wedge R))$

Note that Disjunctive Syllogism is essentially *Modus Ponens* for the material conditional, since $\phi \supset \psi$ is equivalent to $\neg\phi \vee \psi$. Relevantists do not think the material conditional is a real conditional at all, so giving up this principle is not a problem for them.

Advantages:

- This is the only solution if you want to keep disjunctive weakening and transitivity.

Disadvantages:

- There are fewer valid arguments than with the "rejecting transitivity" approach, so this approach is more revisionary.
- It is hard to see intuitively how there is not a relevant connection between premises and conclusion in disjunctive syllogism.

This approach is developed in [2, §15.1], which can be found in your reader.

6 Four-valued tables

It turns out that the logic of tautological entailment can be captured using four-valued truth tables [3, §81] The four truth values are sets of regular truth values: $\{\mathbf{T}\}$, $\{\mathbf{F}\}$, $\{\}$, $\{\mathbf{T},\mathbf{F}\}$. Here are the truth tables for $\neg$ and $\wedge$:

	$\{\}$	$\{\mathbf{F}\}$	$\{\mathbf{T}\}$	$\{\mathbf{T},\mathbf{F}\}$
$\neg$	$\{\}$	$\{\mathbf{T}\}$	$\{\mathbf{F}\}$	$\{\mathbf{T},\mathbf{F}\}$

$\wedge$	$\{\}$	$\{\mathbf{F}\}$	$\{\mathbf{T}\}$	$\{\mathbf{T},\mathbf{F}\}$
$\{\}$	$\{\}$	$\{\mathbf{F}\}$	$\{\}$	$\{\mathbf{F}\}$
$\{\mathbf{F}\}$	$\{\mathbf{F}\}$	$\{\mathbf{F}\}$	$\{\mathbf{F}\}$	$\{\mathbf{F}\}$
$\{\mathbf{T}\}$	$\{\}$	$\{\mathbf{F}\}$	$\{\mathbf{T}\}$	$\{\mathbf{T},\mathbf{F}\}$
$\{\mathbf{T},\mathbf{F}\}$	$\{\mathbf{F}\}$	$\{\mathbf{F}\}$	$\{\mathbf{T},\mathbf{F}\}$	$\{\mathbf{T},\mathbf{F}\}$

$\phi \Rightarrow \psi$ is a tautological entailment iff any assignment of values to propositional constants that makes ϕ *at least* $\mathbf{T}$ makes ψ *at least* $\mathbf{T}$, and any assignment of values to propositional constants that makes ψ *at least* $\mathbf{F}$ makes ϕ *at least* $\mathbf{F}$. (The values $\{\mathbf{T}\}$ and $\{\mathbf{T},\mathbf{F}\}$ are at least $\mathbf{T}$, and $\{\mathbf{F}\}$ and $\{\mathbf{T},\mathbf{F}\}$ are at least $\mathbf{F}$.) In other words: validity is preservation of truth and non-falsity.

Exercises:

4. What should the table for $\vee$ look like? Study the table for $\wedge$ and figure out the principles behind its construction, and apply these to $\vee$. Explain your reasoning.

5. Use your tables, and the definition of tautological entailment above, to test (3a) and (3b) for tautological entailment. You don't need to give the whole truth table (which can be pretty large with a four-valued logic), but be sure to show your work.

7 Natural Deduction for Relevance Logics

So far we have been looking at a system for assessing "first-degree entailments," which are so-called because they contain only one occurrence of the entailment arrow. But, supposing we think of this as a kind of conditional, it should be able to nest. Can we do logic for a conditional that expresses entailment?

One way to get a feel for this is to look at natural deduction systems (but for simplicity, we'll consider systems with just the conditional $\rightarrow$, no other operators).[1] Let's first consider how we can get "paradoxes of material implication" using standard rules:

$$
\begin{array}{lllll}
1 & \quad P & & & \\
2 & \quad\quad Q & & & \\
3 & \quad\quad P & & \text{Reit} & \\
4 & \quad Q \rightarrow P & & \rightarrow \text{Intro } 2\text{–}3 & \\
5 & P \rightarrow (Q \rightarrow P) & & \rightarrow \text{Intro } 1\text{–}4 &
\end{array}
\tag{1}
$$

This is "irrelevant," and relevance logicians want to reject it.

One way to reject it is to restrict the Reit move in line 3. We've seen a similar strategy in the natural deduction rules for modal operators, which use subproofs restricting reiteration. We can get the same effect here by allowing only conditionals to be reiterated into $\rightarrow$ Intro subproofs. The resulting system is $\mathbf{S4}_{\rightarrow}$ (the conditional fragment of **S4**, in which $\phi \rightarrow \psi$ can be defined as $\Box(\phi \supset \phi)$). But this still allows us to get some "irrelevant" theorems, for example:

$$
\begin{array}{lllll}
1 & \quad P \rightarrow R & & & \\
2 & \quad\quad Q & & & \\
3 & \quad\quad P \rightarrow R & & \text{Reit} & \\
4 & \quad Q \rightarrow (P \rightarrow R) & & \rightarrow \text{Intro } 2\text{–}3 & \\
5 & (P \rightarrow R) \rightarrow (Q \rightarrow (P \rightarrow R)) & & \rightarrow \text{Intro } 1\text{–}4 &
\end{array}
\tag{2}
$$

The relevance logic **R** blocks the original proof differently. According to **R**, there's nothing wrong with reiterating P into the subproof. The problem, rather, is that the hypothesis, Q, is not *used*. We fix this by requiring that the hypothesis be used.

Formally, this is done by subscripting. Each formula has, as subscript, a set of numerical indices. When a formula is reiterated, its indices are carried along with it. When a rule uses two or more premises, all of the indices are combined (we take the union of the sets of indices for the premises). When a hypothesis is introduced, it comes with a new index of its own (a singleton set). When a hypothesis is discharged ($\rightarrow$ Intro), its index is subtracted from the index of the consequent, *in which it must be included*. (This is how we

[1] I am indebted in my presentation to seminars given by Nuel Belnap. For a presentation that closely parallels this one, see the first five sections of [2].

require that the hypothesis actually be used.)

So let's see how our proof is blocked:

1	$\mid P_{\{1\}}$ (hypothesis)		(3)
2	$\mid \mid Q_{\{2\}}$ (hypothesis)		
3	$\mid \mid P_{\{1\}}$	Reit	
4	$\mid Q \to P$	Illegal! $2 \notin \{1\}$	
5	$P \to (Q \to P)$		

Proof (2) is blocked for similar reasons. But there are still some things we get that we don't want. For example, the "law of assertion:" $A \to ((A \to B) \to B)$:

1	$\mid A_{\{1\}}$ (hypothesis)		(4)
2	$\mid \mid A \to B_{\{2\}}$ (hypothesis)		
3	$\mid \mid A_{\{1\}}$	Reit	
4	$\mid \mid B_{\{1,2\}}$	$\to$ Elim, 2, 3	
5	$\mid (A \to B) \to B_{\{1\}}$	$\to$ Intro, 2–4	
6	$A \to ((A \to B) \to B)_{\{\}}$	$\to$ Intro, 1–5	

What's wrong with this law, from a relevance point of view? Intuitively: the truth of A does not seem relevant to whether $(A \to B)$ implies/entails B.

This is not a problem unless we want our conditional to *express* (rather than merely indicating) entailment. (Here I'm using terminology from Meyer's article "Entailment.") For it is only in that case that we can read $A \to B$ as "A entails B."

To get a conditional that expresses entailment (and this is system **E**), we need to impose *both* the **S4** restriction on Reit *and* the indexing system of **R**. Then the proof is blocked. So, **E** = **S4** + **R**.

You might conjecture that a sentence is a theorem of $\mathbf{E}_{\to}$ just in case it is a theorem of $\mathbf{S4}_{\to}$ *and* of $\mathbf{R}_{\to}$. But this isn't the case! Kripke came up with a countexample:

$$A \to ((A \to (A \to B)) \to (A \to B))$$

is not a theorem of $\mathbf{E}_{\to}$, but it is a theorem of both $\mathbf{R}_{\to}$ and $\mathbf{S4}_{\to}$.

Try proving it in all three systems…

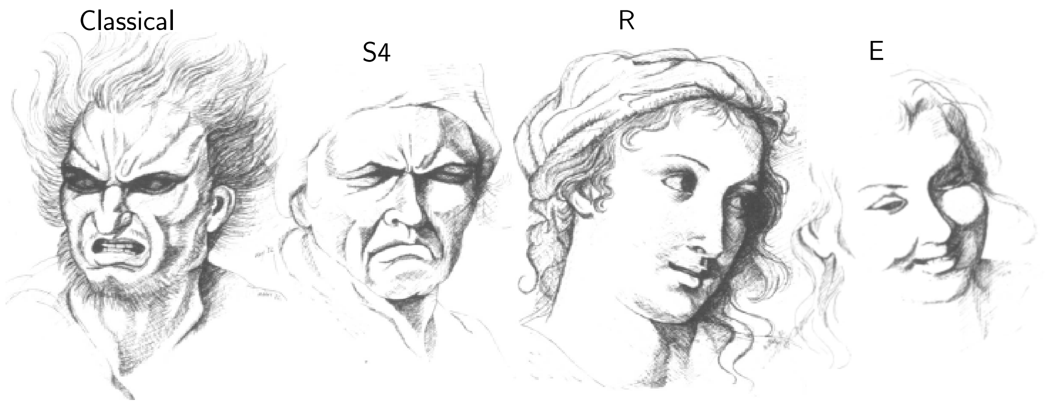

Figure 1: Images from frontispiece of [3].

1	$A_{\{1\}}$	
2	$A \to (A \to B)_{\{2\}}$	
3	?	$\longleftarrow$ can't reit A, so can't finish
4	?	In **R**, we can reit A; in **S4**, we don't need to use it
5	$(A \to B)$	in this subproof.
6	$(A \to (A \to B)) \to (A \to B)$	?
7	$A \to ((A \to (A \to B)) \to (A \to B))$	?

(5)

Exercise:

6. Give a proof of '$A \to ((A \to A) \to A)$' in **R**. Is your proof a valid proof in **E** too? Why or why not?
7. Give a proof of '$(A \to B) \to (((A \to B) \to C) \to C)$' in **E**.

8 Assessing relevantism

The initial arguments for relevance logic were willing to grant that classical forms of inference preserved truth. The idea was that something more was required, in addition to truth preservation: the premises had to be relevant to the conclusion.

Later relevantists tend to move towards the view that propositions can be both true and false, so classical inference principles *don't* preserve truth.

We'll first look at reasons for **wanting relevance in addition to truth**, and then at reasons for **thinking classical inference principles don't preserve truth**.

The first motivation is clearly articulated by Graham Priest:

> ...the notion of validity that comes out of the orthodox account is a strangely perverse one according to which any rule whose conclusion is a logical truth is valid and, conversely, any rule whose premises contain a contradiction is valid. By a process that does not fall far short of indoctrination most logicians have now had their sensibilities dulled to these glaring anomalies. However, this is possible only because logicians have also forgotten that logic is a normative subject: it is supposed to provide an account of correct reasoning. When seen in this light the full force of these absurdities can be appreciated. Anyone who actually reasoned from an arbitrary premise to, e.g., the infinity of prime numbers, would not last long in an undergraduate mathematics course. [10, p. 297]

True, no mathematician would conclude Fermat's Last Theorem on finding that she had inconsistent premises. Nobody, on finding that she had inconsistent beliefs, would infer that she was a pumpkin. But what does that show?

9 Logic and reasoning

Gilbert Harman [5, ch. 1–2] thinks that the whole line of thought we can find in Priest is based on an overly simplistic assumption about the relation between implication or entailment and inference.

Inference—two senses.

Reasoning Inference in the broad sense is reasoned change in view, revision of beliefs in light of new information or reflection.

Argument Inference in the narrow sense is a process of drawing out the consequences of a given set of premises, in isolation from one's other beliefs.

Reasoning is belief revision. It can involve both additions and subtractions to one's body of beliefs. Argument is what is modeled by practices of formal proof. It is a fundamental mistake, Harman argues, to confuse this artificial practice with belief revision. If you confuse these, you'll think that it's obvious that logic has a special role to play in reasoning. But we shouldn't confuse them. Reasoning is not the same as argument or proof.

For example, it's a confusion to think that modus ponens is a rule of Reasoning. It's not. It's a rule of Argument. Doing modus ponens in Reasoning is not always a good idea!

Suppose you find yourself believing

1. Trees must shed their leaves every year in order to live.
2. Pine trees do not shed their leaves.

It would be bad to reason as follows: it follows from (1) that

> If pine trees do not shed their leaves, they cannot live more than one year.

Therefore,

3. Pine trees cannot live more than one year.

Rather than reasoning this way, you should reject one of the premises.

But in Argument, it's fine to write down (3) when you've written down (1) and (2).

When Graham Priest notes that nobody would (or should) reason from a contradiction to, say, the infinity of prime numbers, this is plausible if we are talking about "reasoned change in view." This would be a bad way to revise your beliefs; it would be better, in most cases, just to give up believing the contradiction.

Perhaps GP's claim is also right about the second sense of inferring, the process of drawing out consequences. But it's not at all *obvious* that it is right about this.

Do ordinary folks have intuitions about inferring in the narrow sense (Harman's "argument") at all? When you teach an introductory logic course, it takes quite a bit of time at the beginning just to convey the idea that we're trying to capture a kind of "validity" that is independent of the truth of the premises or the plausibility of the conclusion. Think about what needs to be learned when you learn how to "draw out consequences" in this refined sense: you need to learn to ignore the obvious implausibility of the steps you're generating, to ignore things you know to be true that aren't among the premises, to ignore the falsity of the premises. Try asking untrained people,

> Can "the moon is made of green cheese" be inferred from "Everything in the sky is made of green cheese" and "the moon is in the sky"?

and see what range of responses you get. But if you ask people who *have* been trained in a formal, artificial practice of inferring, they will also reply that Q can be inferred from $P \wedge \neg P$—assuming they were trained in classical logic.

By contrast, people *do* have intuitions (not entirely due to training) about how it is correct to infer in the broad sense. But it is not clear (as Harman shows) how to move from normative claims about belief revision to claims about logical entailment. We need some kind of bridge principle connecting the two.[2] Two obvious bridge principles are

Ought-believe If you believe A and believe B, and A and B together entail C, you ought to believe C.

[2] See [8] for a systematic classification of possible bridge principles.

Entitled-believe If you believe A and believe B, and A and B together entail C, you are entitled to believe C.

Clearly these would help to rule out *ex falso*. Indeed, Priest seems to be appealing implicitly to something like these principles. But neither is plausible.

Ought-believe requires you to believe many trivialities, and sets an impossible standard. (To conform to this norm, you're required to believe every true mathematical theorem if you believe the axioms. Could this be fixed by inserting "you believe that" before "A and B together entail C"?)

Entitled-believe doesn't *require* you to believe all these things; it just permits it. But it doesn't take into account the possibility that, on recognizing that you believe A and B which together entail C, you might sometimes be required to reject C and stop believing either A or B or both. (See the argument above involving pine trees.)

Perhaps logic gives us no positive guidance, but only *negative* guides to belief revision:

Ought-not-believe-strong If A and B together entail C, you ought not believe all of A, B, and not-C.

Ought-not-believe-weak If you believe that A and B together entail C, you ought not believe all of A, B, and not-C.

These principles allow that we may choose between rejecting a premise and accepting the conclusion of a valid argument, but they require us to do one of these things.

However, Harman criticizes even these "negative" norms:

> sometimes one discover's one's views are inconsistent and does not know how to revise them in order to avoid inconsistency without great cost. In that case the best response may be to keep the inconsistency and try to avoid inferences that exploit it. This happens in everyday life whenever one simply does not have time to figure out what to do about a discovered inconsistency. It can also happen on more reflective occasions. For example, there is the sort of inconsistency that arises when one believes that not all one's beliefs could be true. One might well be justified in continuing to believe that and each of one's other beliefs as well. [5, pp. 15–16]

How is it rational to revise your beliefs when you find that they are inconsistent, but can't identify the source of the inconsistency? It's not at all clear that rationality demands that you just give up all the beliefs, or all but one—but these are the only revisions you can know will restore consistency. (Harman: there is a tension between demands of conservativeness and coherence.)

Consider:

- Paradoxes in theories (naive theory of truth, inconsistency of quantum mechanics with gravitation)

- Paradox of the Preface ("Despite my best efforts, I know that some of the claims I make below will turn out to be false.")
- Generalized Paradox of the Preface ("I have at least one false belief.")

Perhaps there is some connection between logical implication/entailment and belief revision, but it is not at all obvious.

But if we can't argue from robust intuitions about belief revision to the rejection of *ex falso quodlibet*, how can we argue for relevance?

10 Relevance logic as the correct logic of truth-preservation

It's actually quite hard to maintain the position that an inference principle like disjunctive syllogism is truth-preserving, but should be rejected because it fails to satisfy an additional requirement of relevance. Consider this argument:

1. $P \vee Q$ (premise)
2. $\neg P$ (premise)
3. If '$P \vee Q$' is true and '$\neg P$' is true, then 'Q' is true (premise)
4. If $P \vee Q$ then '$P \vee Q$' is true (premise)
5. If $\neg P$ then '$\neg P$' is true (premise)
6. '$P \vee Q$' is true (from 1, 4)
7. '$\neg P$' is true (from 2, 5)
8. 'Q' is true (from 3, 6, 7)
9. If 'Q' then Q (premise)
10. Q (from 8, 9)

Note that this proof only requires modus ponens and some plausible premises involving the truth predicate. It does not use disjunctive syllogism or other relevantly questionable inferences. So this looks like a proof, using only relevantly acceptable steps, of disjunctive syllogism. (What steps do you think a relevantist should reject?)

More recently, some have taken to arguing that relevance logics are to be preferred because classical inferences forms don't preserve truth! The idea is to keep the classical idea that validity is truth preservation, but give up the classical assumption that the same sentence cannot be both true and false.

If we look at the 4-valued truth tables, they suggest this. In *ex falso*, you can go from a premise that is {T,F} to a conclusion that is F. So Truth is not preserved.

We need to think more about what these values mean, though.

10.1 Dialetheism

One possibility is to say that {T,F} means: the thing really is both true and false!

Graham Priest takes this line in [11]. *Dialetheism* is the view that some propositions are both true and false (that is, both they and their negations are true).

If you take this view, you have a reason to reject *ex falso* as non-truth-preserving. Why? Because '$P \wedge \neg P$' can be both true and false, when 'Q' is just plain false. Moving from something both true and false to something just false does not preserve truth.

Verify that *disjunctive syllogism* is not truth-preserving, in this sense.

Priest's candidates for sentences that are both true and false:

- The Liar
- "This dissertation is accepted" – supposing the department's word makes it so, and supposing the department said both that it is and that it isn't.
- "I am in the room" (said when my center of gravity is on the vertical plane containing the center of gravity of the door.) (By symmetry, I'm either both in and out, or neither. But if I'm neither in nor out, then I'm both (applying DeMorgan's and DN, and assuming that I'm in iff I'm not out and vice versa).

10.2 Lewis's rejection of dialetheism

> The reason we should reject this proposal is simple. No truth does have, and no truth could have, a true negation. Nothing is, and nothing could be, literally both true and false. This we know for certain, and *a priori*, and without any exception for especially perplexing subject matters. The radical case for relevance should be dismissed just because the hypothesis it requires us to entertain is inconsistent.
>
> That may seem dogmatic. And it is: I am affirming the very thesis that Routley and Priest have called into question and—contrary to the rules of debate—I decline to defend it. Further, I concede that it is indefensible against their challenge. They have called so much into question that I have no foothold on undisputed ground. So much the worse for the demand that philosophers always must be ready to defend their theses under the rules of debate. [7, p. 101]

10.3 The moderate approach

In "A useful four-valued logic" [3], Belnap gives a somewhat different kind of argument. He is not here arguing for full-blown relevantism, the view that we should use relevance logic for all purposes. Rather, he focuses on a particular practical problem: extracting information from a database. The computer gets fed information from several fairly reliable (but not infallible) sources. We want to be able to ask it questions and get answers.

Let's say the sources *only* report on atomic sentences. Each source says either T or F or nothing about a given sentence. So with respect to each sentence the computer can be in four possible states:

T	told true only	**F**	told false only
Both	told both true and false	**None**	told neither

We want to be able to ask the computer not just about atomic sentences (and here it can just spit out the value) but about compounds. To answer our questions, it needs to do some *deduction*. Belnap proposes the four-valued logic (which turns out to capture the first-degree entailments) as a procedure the computer can be programmed to follow in doing this "reasoning."

Why not just use classical logic? The database may contain inconsistent information. For a given A, it might have been told both True and False, so it might reckon both A and its negation True. If it uses classical logic as a canon for reasoning, then as soon as it gets into an inconsistent state, it will start saying "Yes" to every question…which is not useful. We need to be able to contain contradictory information so it doesn't trivialize the whole thing:

> our computer is *not* a complete reasoner, who should be able to do something better in the face of contradiction than just report. The complete reasoner should, presumably, have some strategy for *giving up* part of what it believes when it finds its beliefs inconsistent. Since we have never heard of a practical, reasonable, mechanizable strategy for revision of belief in the presence of contradiction, we can hardly be faulted for not providing our computer with such. In the meantime, while others work on this extremely important problem, our computer can only accept and report contradictions without divesting itself of them. [3, p. 508]

Belnap concedes that it would be preferable to program the computer so that it could subtract beliefs as well as adding them, as in human belief revision. But this is a very difficult problem. No one has given a good algorithm for this. So the relevance logic comes out as a kind of "second best" approach.

This gives a nice explanation of why it makes sense to reject disjunctive syllogism. Suppose the database has been told both that P is true and that P is false, but has been told only that Q is false. Then $P \vee Q$ will be $\{T, F\}$, P will be $\{T, F\}$, and Q will be $\{F\}$. Using DS would allow one to go from premises that are told-true to a conclusion that is not told-true.

Note: the inference from '$(A \vee B) \wedge \neg A$' to '$(A \wedge \neg A) \vee B$' is relevantly okay.

> That is, having determined that the antecedent is at least told True, we allow the computer to conclude: either B is at least told True, or something funny is going on; i.e., it's been told that A is both True and False. And this, you will see, is right on target. If the *reason* that $(A \vee B) \& \sim A$ is getting thought of as a Truth is because A has been labeled as both told True and told False, then we certainly do *not* want to go around inferring B. The inference is wholly inappropriate in a context where inconsistency is a live possibility. [3, p. 520]

10.4 Compartmentalization

Lewis [7] likes the idea of trying to make sense of "truth in a (possibly inconsistent) corpus," but thinks relevance logic isn't the right way to do it. He prefers something he calls "compartmentalization."

Desiderata on "truth in a corpus":

1. Anything explicitly affirmed is true in the corpus.
2. There is more in the corpus besides what is explicitly affirmed—truth in a corpus is to some extent closed under implication.
3. But a corpus can contain inconsistency without containing everything.
4. S is false in the corpus iff $\neg S$ is true in the corpus.
5. The orthodox rules for '$\wedge$' and '$\vee$' apply without exception.

Lewis agrees that if you want 1–5, you get something like the logic E. But he thinks a natural, useful conception of truth in a corpus will reject (5).

Why?

> I am inclined to think that when we are forced to tolerate inconsistencies in our beliefs, theories, stories, etc., we quarantine the inconsistencies entirely by fragmentation and not at all by restrictions of relevance. In other words, truth according to any single fragment is closed under unrestricted classical implication.
>
> [E] cannot be trusted to preserve truth according to a fragmented corpus, nor can any logic that ever lets us mix fragments in many-premise implications. [7, p. 105]

He illustrates this with the following example (see Fig. 2 for a map of Princeton). At one time he believed:

1. Nassau Street ran roughly EW.
2. The railroad ran roughly NS.
3. Nassau Street was roughly parallel to the railroad.

Say these are all true according to his inconsistent belief-corpus.

But, he says, he didn't believe their conjunction—that Nassua Street ran roughly EW *and* the railroad ran roughly NS *and* they were roughly parallel. This wasn't true according to the "corpus" of his beliefs.

The corpus was broken up into overlapping fragments, and different pieces came in to reasoning in different situations, but never all of these at once. Triviality is avoided by allowing inferences only when all the premises are in the same fragment.

Advantages: You can do DS when you're within a fragment.

Disadvantages: Have to figure out how to break the corpus into fragments. Perhaps Lewis's psychology does that for us, but what about in our database? One way would be

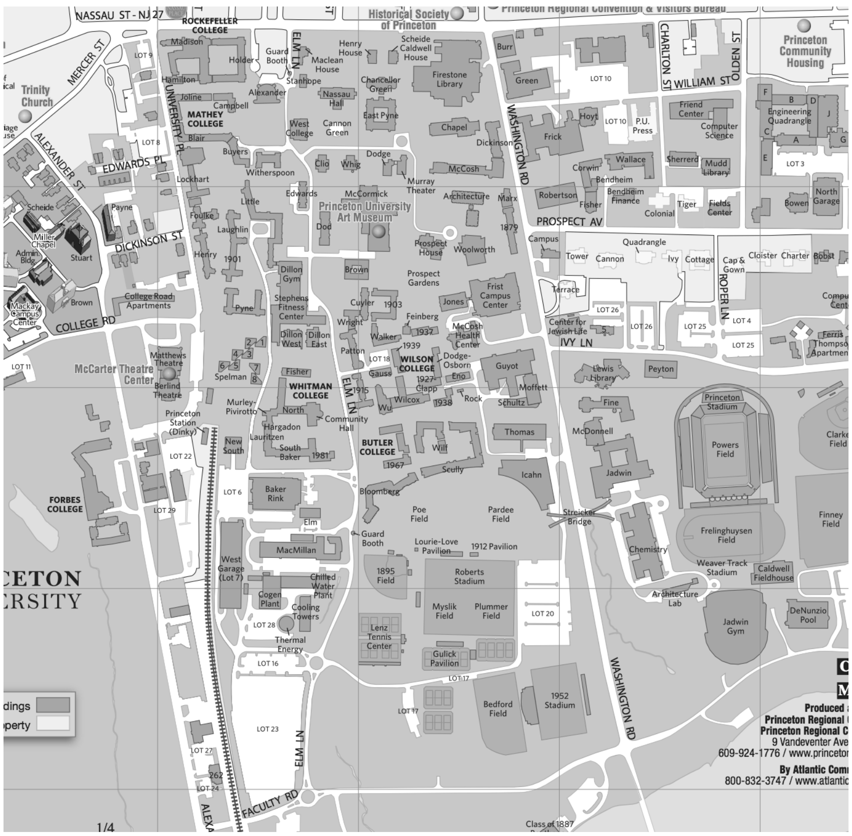

Figure 2: Map of Princeton

to keep track of source of information (assuming no one source will enter contradictory information!), but then the fragments might be too small to get useful inferences.

Having argument that relevance logic is not needed to deal with inconsistent corpora, Lewis suggests that if it is good for anything, it's for preventing damage from potential equivocation:

> We teach logic students to beware of fallacies of equivocation. It would not do, for instance, to accept the premise $A \vee B$ because it is true on another disambiguation of A, and then draw the conclusion B. After all, B might be unambiguously false. The recommended remedy is to make sure that everything is fully disambiguated before one applies the methods of logic.
>
> The pessimist might well complain that this remedy is a counsel of perfection, unattanable in practice. He might say: ambiguity does not stop with a few scattered pairs of unrelated homonyms. It includes all sorts of semantic indeterminacy, open texture, vagueness, and whatnot, and these pervade all of our language. Ambiguity is everywhere. There is no unambiguous language for us to use in disambiguating the ambiguous language. So never, or hardly ever, do we disambiguate anything fully. So we cannot escape fallacies of equivocation by disambiguating everything. Let us rather esacpe them by weakening our logic so that it tolerates ambiguity; and this we can do, it turns out, by adopting some of the strictures of the relevantists. [7, pp. 107–8]

Read "true osd" as "true on some disambiguation." Then we get three values: "true-osd only," "false-osd only," "both true-osd and false-osd." If we define validity as preservation of true-osd only, we get Priest's LP; if we define it as preservation of true-osd and non falsity-osd, we get the relevance logic R-mingle. Hence, relevance logic is "logic for equivocators."

References

[1] Willhelm Ackermann. "Begründung einer strengen Implikation". In: *Journal of Symbolic Logic* 21 (1956), pp. 113–128.

[2] Alan Ross Anderson and Jr. Nuel D. Belnap. *Entailment*. Vol. 1. Princeton: Princeton University Press, 1975.

[3] Alan Ross Anderson, Jr. Nuel D. Belnap, and J. Michael Dunn. *Entailment*. Vol. 2. Princeton: Princeton University Press, 1992.

[4] Hartry Field. "The Normative Role of Logic". In: *Proceedings of the Aristotelian Society* s.v. 83 (), pp. 251–268.

[5] Gilbert Harman. *Change in View: Principles of Reasoning*. Cambridge, MA: MIT Press, 1986, pp. 1–20.

[6] C. I. Lewis and C. H. Langford. *Symbolic Logic*. second. New York: Dover, 1959.

[7] David Lewis. "Logic for Equivocators". In: Cambridge: Cambridge University Press, 1998, pp. 97–110.

[8] John MacFarlane. "In What Sense, If Any, Is Logic Normative For Thought or Reasoning?" 2004. URL: `http://johnmacfarlane.net/normativity_of_logic.pdf`.

[9] Robert K. Meyer. "Entailment". In: *Journal of Philosophy* 68 (1971), pp. 808–818.

[10] Graham Priest. "Two Dogmas of Quineanism". In: *Philosophical Quarterly* 29 (1979).

[11] Graham Priest. "What Is So Bad About Contradictions?" In: *Journal of Philosophy* 95 (1998), pp. 410–426.

[12] Timothy Smiley. "Entailment and Deducibility". In: n.s. 59 (1959), pp. 233–254.