September 21, 2022

Notes for Week 4: Reason Relations I

1) Recap on demarcating discursive practices, commitment/entitlement, defending/challenging. From pragmatic relations among claim*ings* to reason relations among claim*ed*s.

I have picked out discursive practices as practices in which participants undertake or acknowledge commitments by producing assertional performances.
Entitlement to those commitments can be challenged, and must be defended.
In the most basic case, those challenges are reasons against the contents of those commitments, and the defenses are reasons for them.
Claimable (assertible/deniable) contents are what can both serve as reasons in challenges and defenses, and stand in need of reasons.
That is, they can play the role both of premises and of conclusions in broadly inferential relations of implication (reasons for) and incompatibility (reasons against).
Discursive practices accordingly are practices of giving and asking for reasons (defending and challenging claims).
As such, they are normatively governed by reason relations of implication and incompatibility.
We saw both how to define such reason relations in terms of commitments and entitlements to acts (claimings) and practical attitudes (of accepting and rejecting), and how to understand the normative governance of those acts and attitudes by those reason relations.

I have said that reasons for are reasons to accept and reasons against are reasons to reject. But that is talking about reasons for and against *claimables*, not reasons for and against *commitments* that can take the form either of acceptance or rejection. In some of my remarks, I have probably run together these two notions of "reasons for/against"—one concerning claimables, and so directly connected to reason relations, and the other concerning commitments, and only indirectly connected to reason relations.

So let me try to sort things out.
The objects of assertions or denials that have the significance of defenses or challenges (their objects *as* defenses or challenges) is in the first instance practical *attitudes* or *commitments* either acceptances or rejections), which are avowed overtly by *acts* and only secondarily their claimable (acceptable/rejectable) contents.
To challenge an assertion-acceptance is to offer a reason *against* the asserted-accepted claimable.
To defend an assertion-acceptance is to offer a reason *for* the asserted-accepted claimable.

To defend a denial-rejection is to offer a reason *against* the denied-rejected claimable.
To challenge a denial-rejection is to offer a reason *for* the denied-rejected claimable.
To defend an assertion-acceptance is to offer a reason *for* the asserted-accepted claimable.

This way of distinguishing between the relations between *commitments* (to accept and to reject) and *claimables* = commitables, what can be accepted/reject*ed*, asserted/deni*ed*, as opposed to the *attitudes* of accept*ing*/reject*ing* and the (speech) *acts* of assert*ing*/deny*ing*, that is, between the two sides of what Sellars called "the (in)famous 'ing'/'ed' distinction"—which, I think it is important to point out, is a natural (ordinary) language distinction, not a theory-laden philosophers' invention laden—is articulating the fine structure of a pragmatics-first order of explanation.
The transition from the 'ing' side to the 'ed' side, from practical acts and attitudes to reason relations among claimables, takes us half way to a broadly inferential semantics.
That is, it sets up a further move, which understands those claimables in terms of the roles they play in the network of reason relations, in which sets of claimables are nodes or positions.

[Rehearse definitions of implication and incompatibility in terms of commitment to accept premises precluding entitlement to reject/accept conclusion.]

Reminder (of something I mentioned last week):
In these relations, the *reasons* for/against are always considered to be *acceptance* commitments. I don't consider *rejections* on this side of the turnstiles. Doesn't this constitute a *restriction* on the conception?
Smiley and Rumfitt (and following them, Simonelli) would say so. But if and insofar as that is true, the expressive deficit is recovered once we have introduced *logical* vocabulary of conditionals and negation (along with the comma-codifying Boolean helper monkeys of conjunction and disjunction), as I argue further along in these notes. For now we can look at commitments to accept negations, to do the work of commitments to reject, as premises.

2) The topic this week is the **structure of non- or prelogical (material) reason relations**.

Once we have distinguished these reason *relations* from the practice or activity of reasoning that they normatively govern, we can ask after the *algebraic* structure of such relations.

3) First bit of structure is that **there *are* two kinds of reason relations**, and neither is in general definable in terms of the other.

First element of structure is that there are *two* (kinds of) reason relations.
In addition to *implication* or *consequence*, there is *incompatibility* or *incoherence*.
This duality has often been ignored.
(Dummett quote on topic of logic being logical consequence, not logical truth.
But it is not *only* consequence.)
Cannot reduce one of these kinds of reason to the other.

a) In one direction, *can* define *a* consequence relation in terms of incompatibility, and it is a modally robust one. I do this in *BSD*. But it would be a mistake to think this is the *only* kind of material consequence. Not all reasons for can be reconstructed in terms of reasons against.

Ordinary, (implicitly) committive or commitment-preserving implications can't be reconstructed this way.

b) In the other direction, the *only* implicational feature that distinguishes incoherent sets from coherent ones is in systems (like classical and intuitionist logic) that endorse *explosion*. I will say more about explosion later on. But systems that do not admit **explosion**, such as relevance logics, have *no* features that systematically distinguish coherent from incoherent premise-sets, in terms of their consequences.

Q: **Why is it that there are *two* kinds of reason relation: implication and incompatibility**?
Why two and not one or three?
And if two, why just *these* two?

A: I have offered a *pragmatic* reason (a reason expressible in a *pragmatic* metavocabulary).
This dyadic structure of reason relations reflects deep features of discursive practices.
The two practical attitudes of accepting/rejecting, corresponding to two flavors of *commitment* (a normative, deontic status), are expressed in two kinds of speech act: assertion/denial.
And those speech acts, and attitudes, those kinds of commitment, essentially (and not just accidentally) involve the possibility of making more assertions that have the pragmatic significance of defending and challenging *entitlement* to those commitments. They do that by being commitments to accept claimables that stand to the claimables being defended or challenged as reasons *for* or reasons *against* them. And those are the reasons relations among claimables (acceptables/rejectables) of *implication* and *incompatibility*.
That is why there are two kinds of reason relations, and only two.
I think this response is colorable and defensible, and offers a satisfying explanation of the phenomenon in question.

4) Given the parallels between reasons *for* and reasons *against* claimables, between practices of rationally *defending* doxastic commitments by giving reasons *for* the claimables involved in those commitments and *challenging* doxastic commitments by giving reasons *against* the claimables involved in those commitments, and ultimately, the practical attitudes (kinds of commitments) of *accepting* or *rejecting* claimables (acceptables/rejectables) and the speech acts (expressing those commitments) of *asserting* and *denying* them, **one would expect that the internal structure** of reason relations of *implication* or *consequence*, determining reasons *for* claimable contents and relations of *incompatibility* or *incoherence*, determining reasons *against* claimable contents **would be the same**.

The astonishing fact is that they are fundamentally different.
There is **a fundamental *structural difference* between the two sorts of reason relations**:

a) Implication is and must be largely (though not without exception) *non*symmetric. "Pedro is a donkey," implies "Pedro is a mammal," but not the other way around. (Maybe Pedro is a cat.)

b) Incompatibility is and must be exclusively *symmetric*.
"Pedro is a donkey," is incompatible with "Pedro is a prime number," and "Pedro is a wallaby." And his being a prime number or a wallaby is also incompatible with his being a donkey.

Relations of *implication* must be substantially *non*symmetric—though they can include symmetric implication equivalences as special cases.

Relations of *incompatibility* must be, in all cases, *symmetric*.

The first big question is: why is this?
Why this fundamental structural difference between reasons *for* and reasons *against*?
Although there is a certain residual asymmetry between truth and falsity, assertion and denial, acceptance and rejection, with the first element of each dyad having a certain conceptual priority, that priority is subtle and not easy to characterize. (I think, following Dummett, that one must look to pragmatics to do so.) But one would not expect *that* priority to manifest itself in the sort of difference of fundamental algebraic structure between *symmetric* incompatibilities and *non*symmetric implications.

What's going on?

There is a subsequent, less important, but nonetheless nontrivial and even striking observation in the vicinity here, concerning the sociology of philosophy, specifically philosophy of logic, where one would think this issue would be addressed.

It is a measure of the extent to which incompatibility has been systematically snubbed and swept under the rug by the logical tradition that this basic question about why it—by substantial contrast to implication—must be, is *de jure*, and not just *de facto*, *symmetric* has not been addressed by philosophers *at all*.
Simonelli is good on this, looking at recent authors who have been most concerned with incompatibility as a basic reason relation.

5) <u>Incompatibility</u>:

The first question about incompatibility, then, is:

Q: Why is it *de jure* structurally symmetric?

This is a deep question that seems to have attracted no philosophical attention at all.
(The first half of Simonelli's paper offers a good general characterization of the situation.)

It is not obvious why incompatibility should be symmetric, if we define it in terms of commitment and preclusion of entitlement, as I do in *MIE*.

Why should it not be that commitment to A precludes entitlement to B, but commitment to B does not preclude entitlement to A?

Introduce incoherence in terms of irreducible triadic incompatibilities: fruit examples (berries and apples).
Sellars challenge for perceptual-gustatory case.
Understanding incompatibility in terms of incoherence—two sets of claimables are incompatible iff their union is incoherent—(as I do in *BSD*) builds in symmetry.
Defining *incompatibility* (of two sets of sentences-claimables) from *incoherence* (of single sets of claimables) builds in symmetry.
So, in sequent terms, if we say that for A to be incompatible with Γ, Γ#A, is for Γ,A|~⊥ (or just Γ,A|~), which is what we do if we say that Γ|~¬A iff Γ#A, then we build in symmetry because whenever Γ,A,B|~⊥, so that Γ,A#B, it will also be the case that Γ,B#A.
This is how I do things in *BSD*: two sets of claimables are *incompatible* just in case their union is *incoherent*. "Incoherent" here means: a set of commitments-to-accept one cannot be entitled to all of. That is consonant with RR-bilateralism's "out of bounds."

6) A couple of arguments:

a) Direct [This is basically my rendering of what Ryan presents as the first move in his dialogue.]

The idea being assessed is that *q* could be incompatible with *p*, in the sense that commitment to *q* precludes entitlement to *p*, without *p* being incompatible with *q*, in the sense that commitment to *p* precludes entitlement to *q*.
If original assertor, A, claims that *p*, and B objects that *q* (with better reasons—a notion I'll talk about next time), it seems open to A to accept *q* (with the entitlements it gets from *its* reasons for) and hold onto *p*. Then A will be committed *and* entitled to *q*, but only committed, but *not* entitled, to *q*.
Since the incompatibility by hypothesis does not go the other way around, A could not end up committed and entitled to *p*, but only committed to *q*.
But is this scenario in fact intelligible?
Could A *legitimately* accept *q*? That is, could A be *entitled* to accept *q*?
After all, undertaking that commitment by hypothesis precludes entitlement to *p*, something A is already committed to.
In the light of that commitment to accept *p*, isn't A thereby precluded from being *entitled* to a commitment to accept *q*?
If so, then incompatibility *must* be symmetric: commitment to accept either precludes entitlement to accept the other.
Now A might just accept the challenge mounted by B's assertion of *q* and relinquish *p*.

A is then entitled (by B's reasons, assuming they are not incompatible with anything *else* A is committed to accept) to accept *q*.
But if A does *not* respond to the challenge by giving up *p*, and persists in his commitment accepting *p*, then, though that does not make it impossible for him to undertake commitment to *q*, surely he can't be *entitled* to do so. For accepting *q* will preclude entitlement to a commitment A continues to have. How could A be entitled to do that?
It seems that the commitment to accept *p* precludes entitlement to accept anything whose acceptance would preclude entitlement to *p*.
A's commitment to accept *p implicitly* commits A to reject *q*, which is just what *p*'s being incompatible with *q* consists in.

I *think* this argument reflects the following observation:
We read Γ#A as commitment to accept Γ precluding entitlement to commitment to accept A.
But that just means that one cannot be entitled to commitment to accept *all* of Γ,A.
That means that the set Γ∪A is incoherent. Any element of it is incompatible with the rest of it.
And that is symmetry of incompatibility.
Implication does not have this property, because it has commitment to *accept* Γ precluding entitlement to *reject* A. It does not (explicitly) govern a whole *set* of acceptances only.

b) Indirect: Ryan's argument, and its generalization in terms of holistic character of entitlement, which must be assessed relative to *sets* of claimables, regardless of the actual *history* of any *commitment* to any of them.

Simonelli [26-27]
The claim that B makes, concluding the above dialogue, is that asymmetric incompatibility is no real incompatibility at all. Incompatibility, understood pragmatically as I am understanding it here, must, at least potentially, be able to bear in challenging a claim. If p is incompatible with q, in the sense that commitment to p precludes entitlement to q, then an act of claiming q must be able to function to challenge to someone's commitment to p, even if we suppose that the converse does not hold. However, in this case, A claims to take p to be incompatible with q, and yet, A's claim of q is not able to function to challenge to B's commitment to p, since B can simply rejuggle their commitments and, maintaining a commitment to p, commit themself to q in a way that is perfectly licit by A's lights. In this way, A's attitude of taking p to be incompatible with q is utterly pragmatically inefficacious, and, for this reason, B claims that A does not really have this attitude at all.

Can offer Ryan's argument.
What underlies it is the idea that entitlements must be calculated *holistically*, for the whole *set* of commitments. This can depend only on the current constellation of commitments, not on the history of their acquisition.
That would be *deontic hysteresis*.
[But a kind of hysteresis of the sort we argued against in considering the symmetry of incompatibility is what we get if we deny closure in the strong, hypernomonotonic sense of

denying CM: it now matters what path we take in extracting consequences. Should I talk about this at the very end?]

7) Implication:

a) By far the dominant structural model of **implication** understands it as **a *closure* operation, in a strict topological sense**.

It is important to understand this traditional closure model, because the line of thought I will be pursuing in the rest of the course turns on *rejecting* this model.
We will be looking at the substantial consequences of considering *open* (rather than *closed*) reason relations—in particular the consequences for how we should do *logic*, and for how we should do *semantics*.
Commitment to the closure structure of implication is built very deeply into inherited approaches to both logic and semantics.

b) For there is wide agreement about the structure of *logical* reason relations. In classic papers published in 1935-36 Alfred Tarski and Gerhard Gentzen founded rival traditions by introducing, respectively, model-theoretic and proof-theoretic metavocabularies for codifying relations of logical consequence.[1]
In spite of their fundamentally different approaches, they impose essentially the same structural constraints (if we divide through by the fact that Tarski worked with sets and Gentzen with lists).
Tarski expresses them using Kuratowski's axioms defining topological closure operators. Omitting irrelevances, they are[2]:

CO: $\Gamma \subseteq \mathrm{Con}(\Gamma)$.
MO: $\mathrm{Con}(\Gamma) \subseteq \mathrm{Con}(\Gamma \cup \Delta)$.
CT: $\mathrm{Con}(\mathrm{Con}(\Gamma)) = \mathrm{Con}(\Gamma)$.

These are basically Kuratowski's axioms defining topological closure operators.

The first says that the premise-set is contained in the consequences of those premises.
I will call it "Containment" (CO).
In the variant of Gentzen's sequent calculus formulation that I am using, this is:
CO: $A \in \Gamma \Rightarrow \Gamma \mid\sim A$.
If one of the premises is that Pedro is a donkey, then among the conclusions one can draw from that premise-set—indeed, one of the very safest—is that Pedro is a donkey.

[1] Alfred Tarski's 1936 classic "On the Concept of Logical Consequence"
pp. 409-420 in *Logic, Semantics, Metamathematics* [Oxford University Press, 1956].
Gerhard Gentzen's 1935 "Investigations into Logical Deduction" [English translation: *American Philosophical Quarterly* Volume I, Number 4, October 1964, pp. 288-306].

[2] The fourth axiom, omitted here, is $\mathrm{Con}(\varnothing)=\varnothing$.

The second of these Monotonicity (MO).
It says that the consequences of any superset of Γ is a superset of the consequences of Γ.
In Gentzen-style notation, this principle, which he called "Thinning," is
MO: $\Gamma|\sim A \Rightarrow \Gamma,\Delta|\sim A$.
It says that if A is a consequence of the set of premises Γ, then adding further, collateral premises doesn't interfere with that consequence.
You can get *more* conclusions (indeed, CO says that you'll get at least all of Δ that isn't already in Γ), but you won't get *fewer*. The effect of adding premises can only be to add conclusions.
If the premise-set contains "Pedro is a donkey," and one of the consequences of that premise-set is that "Pedro is a mammal," then adding the further premise that "Albertina is an echidna," will yield a premise-set that still implies "Pedro is a mammal."

The third Tarskian condition is transitivity: the consequences of the consequences of a premise-set count as consequences of that set. (Mathematically, this property of the consequence relation is called "*idempotence*," but never mind about that.)
This principle generalizes to sets the idea that if A implies B and B implies C, then A implies C.
It is what allows mathematicians to string together inferences into long chains that lead from axioms, definitions, and rules to ever more distant theorems.
(Transitive chains of inferences are what Kant, working in the old logic, called "*ratiocinatio polysyllogistica*.")

There are good reasons to impose these structural conditions on logical relations of consequence or implication.
As to Containment, it does seem that among the things that follow from any set of premises we should count those premises themselves.
Monotonicity says that if an implication is good, if some conclusion genuinely follows from a set of premises, that it remains good, the same conclusion still follows, if we add further premises.
Transitivity says that we can use the conclusions we draw from some set of premises as further premises from which to draw further conclusions.
These all seem like generally plausible principles of logical reasoning, and they are all satisfied in mathematical reasoning, which is the solidest and in many ways the best understood sort of reasoning we engage in.
Further, sentential logical connectives can be introduced in particularly simple ways, and are particularly well-behaved in settings with this structure.
There has been growing interest in substructural logics, especially nonmonotonic logics, and logicians addressing semantic paradoxes have found it useful to consider relaxing other conditions besides MO.

Such enterprises, though, still have a somewhat suburban status, and are located by their relations to the downtown of the fully structural settings in which core logics, paradigmatically classical bivalent and intuitionist logics, are at home.

Tarski and Gentzen agreed that logical consequence satisfies Containment, Monotonicity, and Cumulative Transitivity.
(In Tarski's terms: $X \subseteq Cn(X)$, $X \subseteq Y \Rightarrow Cn(X) \subseteq Cn(Y)$, and $Cn(Cn(X))=Cn(X)$.)
Gentzen imposed three other structural conditions, Permutation, Contraction, and Expansion, but they are just consequences of his working with *lists* of premises rather than, as Tarski did, *sets* of them. They serve to turn lists into sets. I'm going to ignore them—though Contraction, the principle that if $\Gamma,A,A \mid\sim B$ then $\Gamma,A \mid\sim B$ will come up again later in the course.

8) Tarski and Gentzen might be right about the algebraic structure of *logical* reason relations.

We will not contest that—and indeed, the notion of *logical* consequence generated by the logics we endorse satisfies these classical closure constraints.
But what about relations of rational consequence (and incompatibility) more generally?
A language that does not contain any logical vocabulary at all still must distinguish what follows from what and what doesn't, what claimables provide reasons for or against what other claimables.
Implications such as

- Pedro is a donkey, so Pedro is a mammal,
- Princeton is to the East of Pittsburgh, so Pittsburgh is to the West of Princeton, and
- It is raining, so the streets will be wet,

are all good implications that don't involve any specifically *logical* concepts.
(The "so" is marking an implication, not functioning as a conditional.)
They hold in virtue of the contents of the *non*logical concepts donkey and mammal, East and West, rain and wet.
They are what Sellars calls "*materially* good" implications—by contrast to those that do essentially involve logical vocabulary and are good in virtue of their *logical form*, such as

- *p* and $p \rightarrow q$, so *q*.

What are the structural principles that govern *material* consequence (and incompatibility), *non-* or *pre*logical (material) reason relations?
I've already argued that material incompatibility relations are *de jure* symmetric.

What can we say about the algebraic structure of implication in general?
Let's look more closely at the structural closure principles retailed above.

9) **COntainment**.

Containment (CO) says that among the consequences of any set of premises is those premises themselves. Alone among the metainferential principles that make up a closure structure, we do not consider this one controversial.
It is a generalization of:
Reflexivity (RE): A|~A.
Gentzen builds this in by treating all instances of RE in effect as axioms. Sequents of this "stuttering" form are the leaves of all proof trees in his sequent calculi.
CO follows from RE in the context of monotonicity (MO).
It is in fact a very restricted form of monotonicity.
Instead of saying that *all* implications are robust under arbitrary additions of auxiliary premises, it says that a very special class of implications is robust under such arbitrary additions of premises, namely, sequents that are instances of Reflexivity.
That seems harmless to us, and it is almost universally endorsed.

There is one class of dissenters: relevance logicians.
For them a good implication is turned into a bad one, one where the conclusion does not *follow from* the premises in their sense of "follows from" by adding any further premises that are *irrelevant* to the implication of the conclusion—that is, roughly, any premises that could be omitted without infirming the implication.
That is just what Containment does, so they reject even this extremely limited form of monotonicity.

10) **MO for implication**.

The structural principle of monotonicity says that adding new premises can only *add* consequences. It cannot *subtract* them. If a given premise-set has A as a consequence, then A will still be a implied by that premise-set even if we add some further commitments to it. Those further premises might give us some new consequences, but what followed from the original premise-set should still follow when we add information.

Note that there is a version of monotonicity for incompatibility, too.
It says that if two Γ#A, then for *any* set X of sentences, Γ,X#A.
The "empty RHS" notation for incoherence builds Monotonicity for incompatibility into (makes it a consequence of) Monotonicity for implication. For if Γ|~ (or Γ|~⊥), then MO for implication entails that Γ,X|~ or Γ,X|~⊥.

Both of the kinds of logic Tarski and Gentzen were looking at, classical two-valued logic and intuitionist logic, are strictly monotonic, as are traditional multivalued logics.
Relevance logic is not, because it requires that for a *real* implication *all* the premises have to be essential for drawing the conclusion. It forbids the addition of irrelevancies. MO licenses exactly that.

11) Q: How plausible a structural constrain on implication *in general*, that is, relations material consequence is MO?

A: Not very.

a) Monotonicity, by contrast to Containment, is *not* a plausible constraint on *material* consequence relations.

It requires that if an implication (or incompatibility) holds, then it holds no matter what additional auxiliary hypotheses are added to the premise-set.
But outside of mathematics, almost all our actual reasoning is *defeasible*.

Most of the reasons we give to justify conclusions in ordinary life admit unspoken exceptions or qualifications.

- Why is it that critter can fly? Because it's a bird.

But not if it's a penguin, cassowary, dodo…and so on.

- There are cookies in the cabinet. So, you can eat them in case you're hungry.

But not if the cabinet is stuck, nailed, or locked shut.
And not if it has been infested by thousands of dangerous insects, or the cookies have been poisoned, or mined, or have shrunk to microscopic size, or are encased in glass, or defended by ferocious ferrets….
This sort of example shows why we usually *do* not bother explicitly acknowledging the possible exceptions, so as to make our argument bullet-proof in the sense of monotonic. It would be a real bother.
But it also suggests that perhaps we *cannot* explicitly exclude all the possible further conditions that would make the conclusion not follow.
For the list of possible ways the implication could go wrong—the additional considerations that are compatible with the truth of the original premise but rule out the conclusion—might not only be *long*, it might actually be *infinite*.
And even if it is not, it might be *indefinite*, or simply *unsurveyable*.
How would one go about making sure that one had excluded *every* possibility, however outré, that would infirm the implication—the intervention of freak gravity waves, black holes, or even totally unknown effects of dark energy or dark matter?

b) At this point one is liable to think of ***ceteris paribus* clauses**. They involve reformulations such as:

- There are cookies in the cabinet. So, other things being equal (which is roughly what "*ceteris paribus*" means) you can eat them in case you're hungry.

The possibility of this sort of rephrasing of the implication one is endorsing is significant. But it is important to realize what expressive work is, and what expressive work can*not* be, done by the addition of such a qualification.

The effect of adding a "*ceteris paribus*" clause is just to *acknowledge* the nonmonotonicity of the implication. It is to register explicitly that there *are* exceptions and qualifications, that there *are* further premises that, if they *were* added, *would* defeat the implication of the conclusion. And it registers the existence of such defeaters *without* having to try, tediously, to enumerate them. That is an important expressive function, and it is good to have tools such as *ceteris paribus* clauses to do that work.

But what such clauses do *not* do is to *remove* the nonmonotonicity—to somehow make it that the implication *is* robust under the addition of arbitrary additional premises. For if *that* were what the *ceteris paribus* clause did, it would trivialize the claim. It would be equivalent to saying "Γ implies A—except in those cases in which it doesn't." True, but not substantive, interesting, or informative.

Nor does it have the effect of quantifying over all possible defeaters, and stipulating that *none* of them actually holds. (Think about adding various defeaters to the claim that already has the *ceteris paribus* claim in it. Doing that does not make the premise-set incoherent or inconsistent, and it wouldn't make the implication hold.)

The proper term for a Latin expression whose utterance can turn a nonmonotonic implication into a monotonic one is "magic spell." That's not what *ceteris paribus* clauses are. The expressive function characteristic of *ceteris paribus* clauses is rather explicitly to *mark* and *acknowledge* the defeasibility, hence nonmonotonicity, of an implication codified in a conditional, not to cure it (magically) by *fiat*.

c) [Sobel sequences] One phenomenon that

And things get worse. Not only do implications typically have defeaters testifying to their nonmonotonicity—their fragility under arbitrary augmentations by further premises. Those defeaters also have defeaters—sometimes called "defeater eaters." And it doesn't stop there. One can often find sequences of additional premises that, added one by one, sequentially flip the valence of implications from good to bad. When the implications are expressed explicitly in the form of subjunctive conditionals, this phenomenon has come to be called "Sobel sequences."

- If I were to strike this dry, well-made match, it *would* light.

(Note the explicit qualifications that have already been put in place.)

- If I were to strike a dry, well-made match, and it was in a strong magnetic field, it would *not* light.
- If I were to strike a dry, well-made match and it was in a strong magnetic field, but inside a Faraday cage, it *would* light.

- If I were to strike a dry, well-made match and it was in a strong magnetic field, but inside a Faraday cage, and the Faraday cage was in a vacuum, it would *not* light.
- If I were to strike a dry, well-made match and it was in a strong magnetic field, but inside a Faraday cage, and the Faraday cage was in a vacuum, and it was an underwater match that supplies its own oxygen, it *would* light.

And one could go on.

d) Outside of **mathematical** reasoning, we see nonmonotonicity all the time. Even there, the sort of phenomenon Imre Lakatos recounts in *Proofs and Refutations* concerning the vicissitudes of the concept polyhedron during the late nineteenth century—when the advent for instance of continuously bounded but nowhere-differentiable figures (which by previous criteria consist entirely of vertices, without sides)—suggest that mathematical reasoning *in general* might well not be monotonic, even though stably achieving that status is clearly a regulative ideal.

Physics is no exception. Monotonic models and idealizations must confront nonmonotonic reasoning as soon as they are applied. The very idea of "laws of nature" reflects an obligatorily monotonic approach to subjunctive reasoning that is deformed by a historically conditioned, Procrustean ideology whose shortcomings show up in the need for idealizations (criticized by Cartwright in her provocatively titled book *How the Laws of Physics Lie*) and for "physics avoidance" (diagnosed by Wilson in *Wandering Significance* on the basis of the need to invoke supposedly "higher-level" physical theories in *applying* more "fundamental" ones). Defeasibility of inference, hence nonmonotonicity of implication relations, is a structural feature not just of probative or permissive reasoning, but also of dispositive, commitive reasoning.

Outside of the realm of high theory, the nonmonotonicity of reasoning is an uncontroversial part of the everyday practical experience of auto mechanics and computer help desks. And it is taken for granted in more institutionalized reasoning contexts—for instance in the rules of evidence in courts of law and in medical diagnosis and treatment decisions. (Indeed, the defeasibility of medical diagnoses forms the basis of the plots of every episode of the television show "House" you have ever seen—besides all those you haven't.)

e) Assigning probabilities to kinds of events based on frequencies in reference classes of cases (one kind of statistical reasoning) is notoriously nonmonotonic. Adding an additional premise that further restricts the reference class can result in Sobel-sequence type valence flipping of implications, as in barn-façade cases (of the sort I discuss in "Insights and Blindspots of Reliabilism," Chapter Three of *Articulating Reasons*). Crucial forms of probabilistic reasoning are for this reason not monotonic—even though at the limit, when all probabilities are either 1 or 0, the monotonic probability calculus becomes monotonic classical logic.

f) I think that the nonmonotonicity of reason relations is a manifestation of the practical need for reasons to be finitely statable, in a world where, for any possible state of affairs and any implication there are auxiliary hypotheses that would make that state of affairs relevant to the goodness of the implication. (Jerry Fodor thinks a phenomenon like this is the source of the "epistemological frame problem" in cognitive science.) Could the world be such that this issue would not arise for—and force the nonmonotonicity of the reason relations of—any sufficiently expressively powerful vocabulary? I doubt it, but I don't know how to fill in an argument to that effect.

g) Presuppositions of the monotonicity (and, more generally, full closure-structurality) of reason relations are very common in philosophical reasoning. This is so even though some classical arguments are actually arguments against monotonicity. For instance, skeptical arguments traidtionally proceed by offering a single hypothesis that is claimed to defeat all implications of a certain kind (from the contents of our thoughts to the existence of the external world, or of other minds)—evil demons, brains in vats, and so on.

h) Realizing that reasons are not in general monotonic makes people want nonmonotonic logics.

But, as we'll see maybe that isn't what they should be looking for.

i) Note that there is a version of monotonicity that applies to incompatibilities, too.

It stipulates that if Γis incompatible with A, then so are all of Γs supersets.
Adding further premises cannot cure an incompatibility.
Most of the considerations I have advanced against thinking of implication as globally monotonic apply to incompatibility too.
I won't discuss this issue here.

12) **Transitivity**:

I have accepted Containment (CO) as a plausible condition on reason relations consequence or implication in general, outside of logic. And I have offered reasons to reject monotonicity (MO) as a *global* condition on such reason relations.
Of course, that is compatible with admitting that *some* nonlogical implications *can* be monotonic. Scarlet implies red and is incompatible with green, being an electron implies being smaller than a cat, committing a murder implies doing something illegal, and so on.
But what about the other structural closure condition: **Cumulative Transitivity** (**CT**)?
I'll return to this issue in a bit, when we consider the plausibility of weaker monotonicity principles.

But it is worth pointing out that there is a recipe for turning at least some kinds of failures of monotonicity into failures of CT. This was pointed out by Ryan Simonelli.
(If you could *always* do this, we'd have $\neg MO \Rightarrow \neg CT$, and so, presumably, $CT \Rightarrow MO$, which is not so.)

Failures of MO generating failures of CT:
Here the presence of '(**not** |~)' where MO/CT requires '|~' shows failure of the principle.

Γ = Tweety is a bird.
A = Tweety flies.

Failure of MO:

B = Tweety is a penguin.

Tweety is a bird. |~ Tweety flies.
Tweety is a bird, Tweety is a penguin. (**not** |~) Tweety flies.

B'= Tweety is a nonpenguin.

Failure of CT:

Tweety is a bird |~ Tweety flies, Tweety is a bird, Tweety flies |~ Tweety is a nonpenguin.
Tweety is a bird (**not** |~) Tweety is a nonpenguin.

Here I follow Gentzen in using the solid horizontal line to mark a permissible metainference. It says that if all the sequents above the line are good, then so is the sequent below the line.

13) Cut (CT) is the dual of what is usually thought of as the weakest acceptable structural principle that must be required if full monotonicity is not.[3]

"**Cautious monotonicity" (CM)** is the structural requirement that adding to the explicit content of a premise-set sentences that are already part of its implicit content not defeat any implications of that premise-set. Even though there might be *some* additional premises that *would* infirm the implication, sentences that are *already implied* by the premise-set are not among them.

---

[3] On holding onto both Cut and Cautious Monotonicity, see Gabbay, D. M., 1985, "Theoretical foundations for nonmonotonic reasoning in expert systems", in K. Apt (ed.), *Logics and Models of Concurrent Systems*, Berlin and New York: Springer Verlag, pp. 439–459. Gabbay agrees with the criteria of adequacy laid down by the influential KLM approach of Kraus, Lehman, and Magidor: Kraus, Sarit, Lehmann, Daniel, & Magidor, Menachem, 1990. Nonmonotonic Reasoning, Preferential Models and Cumulative Logics. *Artifical Intelligence*, 44: 167–207.

CM says that even if monotonicity fails in general—because you can't *always* add arbitrary further premises and be sure that that won't cost you some consequences—still, there are *some* claimables that you *can* always add as auxiliary premises without danger of losing any consequences. These are ones that are already consequences of the premise-set.

Almost all contemporary approaches to nonmonotonic logic endorse CM. Dov Gabbay speaks for a contemporary consensus when he takes it that CO, CT, and CM (his Reflexivity, Cut and Cautious Monotony) are critical properties for any well-behaved non-monotonic consequence relation. It is very difficult to add any condition to these without forcing monotonicity, once even minimal logical machinery is added. And he takes it that these are the minimal structural conditions that still permits reasoning in nonmonotonic contexts. (We'll claim he's wrong about that.) I think he is right that the next fallback position is to impose *only* CO. Although it is surprising, we think we can show how that is an intelligible position and leads to tractable logics.

Here the first thing to note is that
CM and CT are duals:

CO: $\Gamma,A|\sim A$

MO: $\dfrac{\Gamma|\sim B}{\Gamma,A|\sim B}$

CM: $\dfrac{\Gamma|\sim A \quad \Gamma|\sim B}{\Gamma,A|\sim B}$

CT: $\dfrac{\Gamma|\sim A \quad \Gamma,A|\sim B}{\Gamma|\sim B}$

What I mean by saying that CM and CT are "duals" is that they consist of the same three sequents, permuted. What in the one is a second premise is in the other the conclusion of the metainference indicated by the horizontal line.
More specifically, CM says that adding to a premise-set Γ any consequence that it implies does not *lose* any consequences Γ already had.
Conversely, CT says that adding to a premise-set Γ any consequence that it implies does not *gain* any consequences beyond those Γ already had.

14) **Explicitation**:

Here is what seems to me to be an illuminating and suggestive way of thinking about implications that also helps in understanding the relations between this restricted form of monotonicity and transitivity in the form of CM and CT.

When we express an implication Gentzen-wise, by writing "Γ|~A," we can think of it as indicating two aspects of the content of the premise-set Γ.
On the one hand, Γ is some set (usually finite) $\{G_1 \ldots G_n\}$ of sentences of the nonlogical language we are working in (so far).
Those sentences G_i, which are elements of the set Γ in the set-theoretic sense, can be thought of as expressing the *explicit* content of Γ. They are what the set Γ literally *contains*: its members.
Now the implication Γ|~A tells us that Γ implies A, so that in *another* sense A is part of the content of Γ. Γ *implies* A, and so "contains" it *implicitly*.
A is part of the *implicit* content of Γ in the *literal* sense of being *implied by* it.

Connect this notion of implicitness with that from my pragmatics.
There we read "Γ|~A" as saying that commitment to accept all of Γ precludes entitlement to reject A, and in that sense commitment to accept all of Γ *implicitly commits* one to *accept* A.
In the pragmatic metavocabulary for reason relations offered last time, we read "Γ|~A" as saying that commitment to accept all of Γ precludes entitlement to reject A, and in that sense commitment to accept all of Γ *implicitly commits* one to *accept* A.
That is, commitment to accept Γ includes **implicit commitment to accept** (what we can now describe as) Γ's *rationally* **implicit content**.

So far, this is just a way of talking. What do we get by talking this way?

This way of distinguishing explicit and implicit content gives us a new way of saying why CO is an unobjectionable structural principle:
CO just says that whatever a premise-set Γ contains *explicitly* it also counts as containing *implicitly*. Its explicit content is a subset of its implicit content.

But we also get something much more.
This evocative idiom of explicit and implicit content motivates a way of talking about a distinctive kind of process. Moving a sentence from the right side of the implication turnstile to the left side is changing the role of the sentence moved from that of conclusion to that of premise of implications. It is changing the status of the conclusion from that of being part of the *implicit* (that is implied) content of the premise-set to being part of the *explicit* content of the new premise-set, which consists of the old one supplemented by that implied consequence.
We can call that process "*explicitation*." Because the sequent expresses implication, a relation of *rational* consequence (a *reason* relation, namely being a reason *for*), we can think of it as *rational* explicitation.
It is making *explicit* (in a new sequent) something that was *implicit* (in the original sequent): upgrading it from being merely implicitly contained to being explicitly contained.

Explicitation in this sense is not at all a *psychological* matter. And it is not even yet a strictly *logical* notion. For even *before* logical vocabulary has been introduced, we can make sense of explicitation in terms of the structure of *material* consequence relations. Explicitation in this sense is a *rational* concept, in the sense that it depends only on the reason relation of implication. Noting the effects on implicit content of adding as an explicit premise sentences that were already implied is already a process available for investigation at the level of the *prelogic*.

And here is the payoff:
We can think in these terms about the structural metainferential principles CM and CT as telling us something about the process of explicitation.
CM tells us that explicitation never *loses* consequences—that is, implicit content.
The premise-set that results from explicitation still has all the consequences, all the implicit content, that the original premise-set had.
CT tells us that explicitation never *adds* consequences—that is, implicit content.
The premise-set that results from explicitation *only* has the consequences, the implicit content, that the original premise-set had.

Together they require that ***explicitation* is *inconsequential***. If CM and CT *both* hold, then moving a sentence from the right-hand side of the implication-turnstile to the left-hand side does not change the consequences of the premise-set. Explicitation has no effect whatever on the implicit content, on what is implied.

It might well be sensible to require the inconsequentiality of explicitation as a structural constraint on *logical* consequence relations. But just as for the logical expressivist there is no good reason to restrict the rational relations of implication and incompatibility we seek to express with logical vocabulary to monotonic ones, there is no good reason to restrict our expressive ambitions to consequence relations for which explicitation is inconsequential. On the contrary, there is every reason to want to use the expressive tools of logical vocabulary to investigate cases where explicitation *does* make a difference to what is implied.

One such case of general interest is where the explicit contents of a premise-set are the records in a *database*, whose implicit contents consist of whatever consequences can be extracted from those records by applying an *inference engine* to them. (The fact that the "sentences" in the database whose material consequences are extracted by the inference engine are construed to begin with as *logically* atomic does not preclude the records having the "internal" structure of the arbitrarily complex datatypes manipulated by any object-oriented programming language.) It is by no means obvious that one is obliged to treat the results of applying the inference-engine as having exactly the same epistemic status as actual entries in the database. A related case is where the elements of the premise-sets consist of experimental *data*, perhaps measurements, or observations, whose implicit content consists of the consequences that can be extracted from

them by applying a *theory*. In such a case, explicitation is far from inconsequential. On the contrary, when the CERN supercollider produces observational measurements that confirm what hitherto had been purely theoretical predictions extracted from previous data, the transformation of rational status from *mere* prediction *implicit* in prior data to actual empirical observation is an event of the first significance—no less important than the observation of something incompatible with the predictions extracted by theory from prior data. This is the very nature of empirical *confirmation* of theories. And it often happens that confirming *some* conclusions extracted by theory from the data infirms *other* conclusions that one otherwise would have drawn.

Imposing Cut and Cautious Monotonicity as global structural constraints on material consequence relations amounts to equating the epistemic status of premises and conclusions of good implications.
But in many cases, we want to acknowledge a distinction, assigning a lesser status to the products of risky, defeasible inference.
In an *ideal* case, perhaps this distinction shrinks to nothing.
But we also want to be able to reason in epistemically messier situations where it is important to keep track of the difference in status between what we take ourselves to know and the shakier products of our theoretical reasoning from those premises.
We shouldn't build into our global structural conditions on admissible material relations of implication and incompatibility assumptions that preclude us from introducing logical vocabulary to let us talk about those rational relations, so important for confirmation in empirical science.

Thinking about CM and CT in terms of explicitation shows us that they jointly entail the *inconsequentiality of explicitation*.
But for at least some real material consequence relations, explicitation is *not* inconsequential.
It matters whether something is part of the *implicit* content of a set of premises, or part of its *explicit* content.
And real work might need to be done, a real difference made, to promote it from the one status to the other, that is to make that consequence explicit, or to acknowledge it explicitly.
That amounts to an argument that we should *not* think of the structures expressed by CM and CT as characterizing relations of material consequence or implication *generally*.

Reason relations that are not universally monotonic—that is, for which the structural principle MO does not hold for all implications—are *nonmonotonic*.
Cautious monotonicity, CM, is a (much) weaker form of monotonicity.
Reason relations that are not universally even *cautiously* monotonic—that is, for which neither the structural principle CM and *therefore* the structural principle MO hold for all implications—are *hyper*nonmonotonic.

From here on out, we are going to be concerned with **reason relations that are radically structurally *open*, in the sense of being *nontransitive* and *hypernonmonotonic*.**

15) Looking forward to logic:

We would like our logic to be able to express open-structured, that is, substructural, reason relations, and not just those with closure structures.

What we want is a logic to codify reason relations. And the only structural principle we can endorse generally is CO. Beyond that, we want the expressive power to codify *any* substructural, structurally *open* (as opposed to closed) constellation of reason relations.

That is different in principle from a nonmonotonic logic.

It is a logic codifying nonmonotonic reason relations (and nontransitive, and *hyper*nonmonotonic ones). The logic itself—the purely *logical* reason relations—might be fully structural.

And in fact that is true of the logics we will propose.