May 28, 2006

Between Saying and Doing:
Towards an Analytic Pragmatism

Lecture 5 (May 31, 2006):

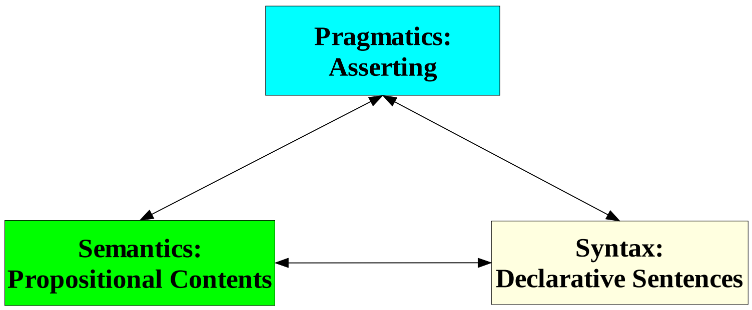
Incompatibility, Modal Semantics and Intrinsic Logic[1]

Section 1: Introduction

I closed my lecture last week with an argument building on the idea that every autonomous discursive practice, in order to count as a *discursive* or *linguistic* practice, in order to count as deploying any *vocabulary*, must include performances that have the *pragmatic* significance of *assertions*, which on the *syntactic* side are utterances of *declarative sentences*, and whose *semantic* content consists of *propositions*. These pragmatic, syntactic, and semantic conditions form an indissoluble package in the sense that one cannot properly understand any of the concepts assertion, sentence, and proposition apart from their relation to each other. This is the *iron triangle of discursiveness*:

[1] The work reported here was conducted with the support of the A.W. Mellon Foundation, through their Distinguished Achievement in the Humanities Award, and the Center for Advanced Study in the Behavioral Sciences at Stanford University (where I was also supported by the Mellon Foundation).

The Iron Triangle of Discursiveness



This structure can be further articulated as a triad of triads: On the pragmatic side of kinds of *doing* are asserting, inferring, and referring. On the semantic side of corresponding worldly items are facts, laws (more generally: counterfactually robust consequence relations[2]), and objects (and their properties). And on the syntactic side of vocabulary are sentences, (modally qualified quantified) conditionals, and singular terms (and predicates). These are the items *used* in the various kinds of doings so as to express semantic contents in the sense of representing the various kinds of categorial features of the objective world. These relations will be further explored in the sixth lecture.

I then proceeded to look at the pragmatic presuppositions of the assertional practices that are, on this account, PV-necessary to deploy any autonomous vocabulary. Here my claim was that no set of practices could count as according some performances the pragmatic significance of assertions unless it includes practices of giving and asking for *reasons*. That is the claim that within the pragmatic dimension of the triad, *asserting* and *inferring* also form an indissoluble package, each element of which is in principle intelligible only in a context that includes the other. Assertional and inferential practices are reciprocally PP-necessary. Notice that these claims do not depend on the sort of methodological pragmatism that seeks to understand the semantic and syntactic

[2] Since, as I argued in the previous lecture, counterfactually robust material inferences are not always expressions of laws.

dimensions of the triadic constellation of assertion, sentence, and proposition in terms of the pragmatic one. Nor am I claiming that practices of assertion-and-inference are PV-*sufficient* to deploy an autonomous vocabulary—that is, sufficient to *confer* propositional semantic contents. (Although I did pursue that strategy and make that claim in *MIE*.) What is needed here is only something much weaker: that practices of assertion-and-inference are PV-*necessary* to deploy any autonomous vocabulary.

I then argued that any constellation of social practices is intelligible in principle as including the giving and asking for *reasons*—making *claims* whose status depends on their *inferential* relations to other claims that are their consequences, or have them as their consequences, or rule them out—only if it includes the capacity to distinguish two sorts of normative status as part of the pragmatic significance practically attributed to a speech act. To be giving and asking for reasons, practitioners must distinguish (be able to respond differentially to) the sentences to which their interlocutors and they themselves are *committed* (based on those they are disposed to assert). And they must distinguish the sentences to which their interlocutors and they themselves are *entitled* (based on those they are committed to). These practical discriminative capacities need not be infallible (by any standard of ultimate correctness), and they need not be complete. And I am not claiming (here) that including these two sorts of attribution and acknowledgment of pragmatic significance is *sufficient* for the practices in question to count as vocabulary-deploying. But unless interlocutors make at least these two sorts of discrimination, what they are doing does not deserve to count as producing and consuming *reasons*, hence not as practically according some performances the pragmatic significance of *assertions*, hence not as deploying any autonomous *vocabulary*.[3]

[3] These are strong claims, no doubt contentious because tendentious: framed from the point of view of a normative pragmatist rationalism about the discursive. Those who are not convinced, those not tempted, and those not even willing to suspend disbelief on these points should just consider the remarks that follow as restricted to that sub-class of discursive practices that *does* exhibit the structure being considered—a sub-class that should at least be admitted to be large and significant, even by those who doubt that it plays the foundational and demarcational role

My interest last time was in arguing that these practices-or-abilities to discriminate commitments and entitlements are, in the terms of the sort of meaning-use analysis I have been developing here:

- PV-necessary for deploying any autonomous vocabulary,
- PP-sufficient for engaging in practices that are
- PV-sufficient to deploy normative vocabulary, which is
- VP-sufficient to specify those original universally PV-necessary practices-or-abilities.

In sum, it was to argue that normative vocabulary—paradigmatically 'commitment' and 'entitlement'—stands in the complex resultant meaning-use relation of being elaborated-explicating (LX) with respect to every autonomous vocabulary. Whatever the status of that argument may be, my purpose here is to consider a different complex resultant meaning-use relation that the explicitly normative vocabulary of commitment and entitlement stands in to other vocabularies of philosophical interest, principal among them being *modal* vocabulary. The relation I will focus on is that of one vocabulary's being a *pragmatic metavocabulary* for another. I want to explore a particular construction according to which normative vocabulary can serve as a pragmatic metavocabulary for *logical* vocabulary, including *modal* vocabulary, and how in those terms it can be seen to serve as such a metavocabulary for *semantic* vocabulary more generally. Along the way we will learn some lessons about logic and modality, and especially about the relation of *truth* and *compositionality* to semantics, that I think are of general interest, quite apart from the way in which they emerge from the particular analytic project I am pursuing here.

---

here attributed to it. In any case, the principal arguments and constructions to be presented here as articulating incompatibility semantics do not depend on the particulars of the normative pragmatic metavocabulary in terms of which I want to understand incompatibility.

Section 2: Incompatibility

The story I told about how engaging in practices of giving and asking for reasons requires the practical differential responsive ability to take or treat someone as committed and as entitled to the claims expressed by various sentences lets us make sense straight away of two sorts of inferential relations between propositional contents on the semantic side, on the one hand, and corresponding practical dispositions on the pragmatic side, on the other. One takes or treats *q* as an inferential consequence of *p* in one sense by being disposed to attribute *commitment* to (what is expressed by) *q* to whomever one credits with commitment to (what is expressed by) *p*. And one takes or treats *q* as an inferential consequence of *p* in another sense by being disposed to attribute *entitlement* to the claim that *q* to whomever one credits with entitlement to the claim that *p*.[4] The first sort, commitment-preserving inferential relations, is a generalization, to include the case of non-logical, material inferences, of obligatory, *deductive* inferential relations. The second sort, entitlement-preserving inferential relations, is a generalization, to include the case of non-logical, material inferences, of permissive, *inductive* inferential relations. For example, anyone who is committed to a plane figure being rectangular is committed to its being polygonal. And the old nautical meteorological homily "Red sky at night, sailor's delight; red sky in morning, sailor take warning" tells us that anyone who sees a colorful sunrise is entitled to the claim that a storm that day is probable. But here the reasoning is only probative, not dispositive. The colorful sunrise provides *some* reason to predict a storm, but does not yet settle the matter. Other considerations, such as a rising barometer, may license one not to draw the conclusion one would otherwise be entitled to by the original evidence.

[4] As will appear, entitlement-preserving inferences are always defeasible; the entitlement one acquires thereby is only *prima facie*. One is not entitled to the conclusion of a good entitlement-preserving inference if one is committed to something incompatible with it.

The abilities to take or treat interlocutors (including oneself) as committed or entitled to propositional contents expressed by various declarative sentences are PP-sufficient for the practical responsive recognition of another sort of semantic relation among propositional contents. For being disposed to respond to anyone who is *committed* to *p* as thereby precluded from counting as *entitled* to *q* (and vice versa) is treating *p* and *q* as *incompatible*. On the pragmatic side, this is a *normative* relation. It is not that one *can*not undertake incompatible commitments, make incompatible assertions. Finding that one has done so is an all-too-common occurrence. But the effect of doing so is to alter one's normative status: to undercut any entitlement one might otherwise have had to either of the incompatible commitments. For each commitment counts as a decisive reason against entitlement to the other.

On the pragmatic side, incompatibility can accordingly be thought of as a consequential relation like the other two:

- <u>Incompatibility of *p* and *q*</u>: If *S* is committed to *p*, then *S* is not entitled to q.
- <u>Committive consequence</u>: If *S* is committed to *p*, then *S* is committed to q.
- <u>Permissive consequence</u>: If *S* is committed and entitled to *p*, then *S* is (*prima facie*) entitled to q.

But it is not immediately an *inferential* relation, since the conclusion is the *withholding* of a normative status, rather than the inheritance of one. Incompatibility relations do, however, underwrite a kind of inferential relation. The idea is an old one. Sextus Empiricus says, perhaps referring to Chrysippus:

> And those who introduce the notion of connexion say that a conditional is sound when the contradictory of its consequent is incompatible with its antecedent.[5]

My concern is not with when a conditional is sound, but with when the underlying inference that it is VP-sufficient to specify is a good one, in the material (that is, non- or better pre-logical) sense of "good inference" we are trying to articulate. And I do not want to assume at this stage that we are in a position to identify the *contradictory* of any claim. But the notion of material incompatibility can be put in its place. Making those adjustments yields the following definition:

> *p incompatibility-entails q* just in case everything incompatible with *q* is incompatible with *p*.

Thus 'Pedro is a donkey,' incompatibility-entails 'Pedro is a mammal,' for everything incompatible with Pedro's being a mammal (for instance, Pedro's being an invertebrate, an electronic apparatus, a prime number…) is incompatible with Pedro's being a donkey. (Notice that this is an asymmetric relation. Pedro's being a cat is incompatible with his being a donkey, but not with his being a mammal.)

I said before that the inferential relations among the propositional contents expressed by declarative sentences that correspond on the semantic side to inheritance of *commitment* can be thought of as a generalization (to the material case) of *deductive* inferential relations, and that those corresponding to inheritance of *entitlements* can be thought of as a generalization to the material case of *inductive* inferential relations. So we may ask: do incompatibility-entailments similarly generalize some kind of inferential relation that we already recognize in other terms? I think that they do, and that the inferences in question are counterfactual-supporting, *modally*

[5] Sextus, *Pyrrhoneiae Hypotyposes,* ii, 110-112, translated by Kneale in William and Martha Kneale, *The development of logic* [Oxford U. Press, 1962], p. 129.

robust inferential relations: a sort of *strict implication*, the kind of inferences made explicit by *modally qualified conditionals*. The fact that the *properties* of being a donkey and being a mammal stand in the relation of incompatibility-entailment means that every *property* incompatible with being a mammal is incompatible with being a donkey. If two properties (such as being a mammal and being an invertebrate) are incompatible then it is *impossible* for any object simultaneously to exhibit both. And that means that it is *impossible* for anything to be a donkey and not be a mammal. That is why the incompatibility-entailment in question supports counterfactuals such as "If my first pet (in fact, let us suppose, a fish) *had been* a donkey, it *would have been* a mammal." We could say: "Necessarily, anything that is a donkey is a mammal."

On the semantic side, incompatibility is an implicitly *modal* notion. On the pragmatic side, the *normative* concepts of commitment and entitlement provide a pragmatic metavocabulary VP-sufficient to *specify* practices PV-sufficient to *deploy* that modal notion. That is, they let us *say* what it is one must *do* in order thereby to be taking or treating two claims *as* incompatible.[6] To begin to explore the consequences of this pragmatically mediated semantic relation between normative and modal vocabularies, we may consider the sort of grip on the semantics of expressions—the meanings expressed by deploying vocabularies—that one gets by thinking of their contents in terms of incompatibilities. We know in some sense that this grip will be incomplete, since the inferential relations corresponding to commitment-preserving and entitlement-preserving pragmatic consequential relations are left out of the picture. But I argued last time that there is an intimate connection between the conceptual contents expressed by vocabularies and the counterfactually

6 In my final lecture I shall be concerned to explore in much further detail this relation between what is expressed by modal and by normative vocabulary, as a way of thinking about the intentional nexus beween objects and the subjects who make claims about and act upon them.

robust inferences they are involved in. We might hope that a semantic metavocabulary centered on *incompatibility* would have the right expressive resources to make explicit important features of such contents. One case where we have particularly clear criteria of adequacy for our semantics is *logical* vocabulary. So I will be specifically concerned to offer an incompatibility semantics for logical vocabulary. Again, since incompatibility is at least implicitly itself a *modal* notion, we will want to see what an incompatibility semantics for modal vocabulary might look like.

Section 3: Incompatibility Semantics

Incompatibility is a relation of *exclusion*. That a plane figure is a circle *rules out* its being a triangle. There is a long philosophical tradition that understands *determinateness* in terms of exclusion. A central principle of Spinoza's metaphysics was "Omnis determinatio est negatio." Inspired by that thought, Hegel takes what he calls "determinate negation"—for which we can read "material incompatibility"—to be the central relation that articulates the conceptual structure both of claims and other commitments on the side of subjective discursive activity, and of states of affairs and the objects and properties they comprise on the side of the objective world. It is important for Hegel that exclusive [ausschließend] difference, or contrariety, is stronger than what he calls 'mere' or 'indifferent' difference: the relation that holds between triangular and red. In a sense that we will see can be made precise, it is to be understood as more basic than 'formal' or 'abstract' negation: the relation between the contradictories triangular and not-triangular.

One strategic motivation for this starting-point is that, as we have seen, the vocabulary of incompatibility is VV-sufficient for the expression of *inferential* relations. 'Mediation' [Vermittlung] is his generic term for *material inference*. A paradigm would be the non-logical (hence 'material') inference from something's being a horse to its being a vertebrate. Together with all his contemporaries and predecessors, he thought of inference in terms of Aristotelian syllogisms. The middle term 'mammal' in the syllogism (in Barbara):

All horses are mammals.

All mammals are vertebrates.

∴ All horses are vertebrates.

is said to 'mediate' the inferential connection between horses and vertebrates.[7]

To possess determinate conceptual content is then to stand in relations of both determinate negation and mediation. This is a resolutely *non-psychological* conception of the conceptual. For although the commitments that characterize subjects, in the form of beliefs and intentions, have contents that stand in relations of material incompatibility and consequence to one another, so too do the properties that characterize objects. The property of being a horse is incompatible with the property of being a cat, and both entail the property of being a vertebrate. In virtue of being articulated into *facts* about *objects* possessing *properties* and standing in *relations*, the objective world is intelligible as having what is in this sense a *conceptual* shape. Since properties exclude and entail one another, so do facts, and that is enough for them to count as conceptually articulated in Hegel's sense. Consequential relations among facts underwrite inferences involving the statements that express them. Inferring is reasoning, and in that sense inferential-consequential relations are *rational* relations.

Another reason to begin with exclusive difference or incompatibility is that Hegel understands these conceptual relations as *modally robust*. To say that one way things could be *entails* another is to say that it is not *possible* that the first obtain and the second not—that if the first obtains, then the second *necessarily* does. And to say that one way things could be is *incompatible* with another is to say that it is not *possible* that the second obtain if

[7] Hegel's view is that *all* inferences are mediated—in the sense that for *any* consequential connection between two concepts, a concept can be found that mediates it in the sense of providing a middle term for a syllogism connecting them. Producing such a middle term is making *explicit* some of what is *implicit* in the inference, and hence in the concepts it connects. Expressing what is in this sense implicit in ordinary determinate empirical/practical concepts is an infinite, in-principle uncompletable task. This is part of what he means by saying that the Concept—the holistic system comprising all the determinate concepts and their relations—is 'infinite'.

the first does—that if the first does, it is *necessary* that the second does not. Understanding the conceptual in terms of material incompatibility and material inference is his way of working out Kant's thought that *conceptual* relations are essentially *modal* in nature. One of the characteristic commitments they share is that to be a modal realist about the objective world—to take it that what is true is intelligible only in terms of *laws* expressing what is genuinely *possible* and *necessary*—*is* to take it that the objective world is *conceptually* articulated. Among the core commitments of their brand of idealism are two that are of considerable interest to us today: to understanding metaphysics in terms of semantics, and to understanding semantics in terms of modality.

So here is a semantic suggestion: represent the propositional content expressed by a sentence by the set of sentences that express propositions incompatible with it. More generally, we can associate with each *set* of sentences, as its semantic interpretant, the set of *sets* of sentences that are incompatible with it.[8] The generalization from seeing incompatibility as a relation among sentences to seeing it as a relation among *sets* of sentences acknowledges an important structural fact about incompatibility: one claim can be incompatible with a set of other claims without being incompatible with any of its members. So on the formal, logical side, where incompatibility is just inconsistency, p is incompatible with the set consisting of $p \rightarrow q$ and $\sim q$, but not with either individually. And on the side of non-logical content, the claim that the piece of fruit in my hand is a blackberry is incompatible with the *two* claims that it is red and that it is ripe, though not with either individually—in keeping with the childhood slogan that blackberries are red when they're green.

[8] This generalization opens up a number of possibilities for correspondingly generalizing the incompatibility entailment relation. One very natural way to do that is to take it that a set of sentences X incompatibility entails a set Y in just case every set Z that is incompatible with Y is incompatible with X. In this case, $X|=\{y_1, ..., y_n\}$ has the meaning, X entails (y_1 and .... and y_n). It turns out to be more formally convenient if instead one requires that X incompatibility entails Y in case every set Z incompatible with *every sentence in Y* is incompatible with X. In this case, $X|=\{y_1, ..., y_n\}$ has the meaning, X entails (y_1 or .... or y_n).

Aiming at maximal generality, I will impose only two conditions on the incompatibility relations whose suitability as semantic primitives I will be exploring here. First, I will only consider *symmetric* incompatibility relations. This is an intuitive condition because it is satisfied by familiar families of incompatible properties: colors, shapes, quantities, biological classifications, and so on. (Though it is perhaps worth noting in passing that the pragmatic metavocabulary in terms of which we understand the use of this semantic primitive does not entail this restriction. It is in principle intelligible that commitment to *p* should preclude entitlement to *q*, while commitment to *q* does not rule out entitlement to *p*.) Second, if one set of claims is incompatible with another, so is any larger set containing it. That is, one cannot remove or repair an incompatibility by throwing in some further claims. I call this the 'persistence' of incompatibility. If the fact that the monochromatic patch is blue is incompatible with its being red, then it is incompatible with its being red and triangular, or its being red *and* grass being green. (Once again, this plausible condition is in principle dispensable. The result of relaxing it will be a non-monotonic consequence relation.) Any incompatibility relation satisfying symmetry and persistence can be represented by partitioning all the sets of sentences over some vocabulary into those that are and those that are not *incoherent*. Examples of incoherent sets of sentences are {p,p→q,~q} and {'The fruit in my hand is a blackberry,' 'The fruit in my hand is red,' 'The fruit in my hand is ripe.'}. The connection between incoherence and incompatibility is established by the principle of *partition*: two sets of sentences are incompatible just in case their union is incoherent. Symmetry of incompatibility then is a consequence of the symmetry of the union operation, and persistence of incompatibility will hold so long as it holds for incoherence—that is, so long as we insist that all the supersets of any incoherent set are also incoherent. Incoherence properties that satisfy this persistence condition—and their associated incompatibility relations—I will call '*standard*'.

Given any set of sentences, we can then define a standard *incompatibility interpretation* over that vocabulary as an incoherence partition of its power set that satisfies persistence. Each such incompatibility interpretation induces an *incompatibility consequence* (or entailment) *relation* |= in the way already indicated: Being a cat entails being a mammal in this sense because every set of properties incompatible with being a mammal is also incompatible with being a cat.[9] A *defeasor* of a putative entailment is any set of sentences that is incompatible with its conclusion but not with its premises. Thus that Moby is a finned creature that lives in the sea does not entail that Moby is a fish, since 'Moby is a whale' is incompatible with the latter but not the former. Persistence ensures that incoherent sets are incompatible with everything, and it follows both that they cannot defeat any putative entailment and that they entail everything. We can note for future reference that a notion of *validity* for consequences applies to entailments that hold in all standard frames. And it can be extended to sets of sentences that are entailed by everything in all frames. (We don't yet know of any such sentences, since we haven't yet talked about how to introduce *logical* vocabulary. But that is coming soon.)

Before turning to the consideration of logical vocabulary in the framework of incompatibility semantics, it is worth pausing for a few general observations about this semantic framework. To begin to place it in philosophical space, we might compare the idea of doing semantics in terms of incompatibility relations to some of its neighbors. One comparison that might come to mind is with Sheffer's demonstration, early in the twentieth century, that if one defined a logical connective of incompatibility from the usual connectives of classical propositional logic, as not both *p* and *q*, one could then define both negation and conjunction (and hence the rest of the classical connectives) in terms of that Sheffer stroke alone.[10] Another is the fact that at around the same time, at the dawn of modern modal logic, Sheffer's Harvard colleague C.I. Lewis was already exploring modal systems that used a dyadic *impossibility*—that is, non-compossibility, *necessarily* not both *p* and *q*—operator, in place of the now-standard monadic *necessity* and *possibility* operators.[11] (In fact, though it is less widely known, these two lines of thought can in a sense be combined: Sheffer operators for *modal* systems have been discovered. So, using the standard connectives, one can define single Sheffer operators sufficient in turn to define all the connectives of S5, and of S4,

[9] See the previous note. For present purposes, we can think of this talk of the incompatibility of sets of *properties* as shorthand for a generalization about incompatibilities of sets of sentences.
[10] For the definitions, see note 10 below.
[11] C. I. Lewis *A Survey of Symbolic Logic* [University of California Press, Berkeley, 1918].

although these lack the intuitive appeal of Sheffer's original.[12]) There are indeed some conceptual filiations between these enterprises and the present one. But the salient difference is that both Sheffer and Lewis were deploying operators within the *object* vocabulary (of logic). The proposal here is to use incompatibility (itself introduced by a normative pragmatic metavocabulary) as the basic element of the *semantic metavocabulary*—and not just for *logical* expressions, but for ordinary *non*-logical vocabulary as well. The semantic interpretant of an object-vocabulary sentence is taken to be the set of sets of sentences materially incompatible with it.

The result is a *modal* semantics. For *incompatibility* is a *modal* notion (albeit one that has been introduced via a normative pragmatic metavocabulary)—essentially C.I. Lewis's *non-compossibility* raised from the level of syntax to that of semantics. Now the development of modal semantic metavocabularies—in particular, the extension of possible world semantics from its initial home as a semantics for modal *logical* vocabulary to a modal semantics for ordinary, non-logical expressions in general—is perhaps the principal technical philosophical advance of the past forty years. I want to take that hint, but to apply modal vocabulary to semantic projects in a somewhat different way, using the notion of incompatibility to provide a *directly* modal semantics. By that I mean one that does not approach modality by beginning with a more basic semantic notion of *truth*. Classical possible worlds semantics proceeds in two stages. Like more traditional semantics, its basic semantic notion is that of truth. It begins by relativizing evaluations of truth to points of evaluation—paradigmatically, possible worlds (self-consciously modeling this procedure on Tarski's relativization of satisfaction of a quantified statement to an assignment of values to its variables). Then, at the second stage, necessity and possibility can be introduced by quantification over such points

[12] The S5 connective, due to my colleague Gerry Massey, is $p*q \approx_{df.} \sim\Diamond(p\&q) \vee (\Diamond(p\&q)\&\Diamond(p\&\sim q)\&\sim(p\&\sim q)) \vee (\Diamond(p\&q)\&\sim\Diamond(p\&\sim q)\&\sim(p\&q))$. Then $\sim p \approx p*p$, $p\supset q \approx p*(p*q)$, and $\Diamond p \approx ((\sim p*p)*p)\supset p)$. The S4 connective is simpler: $p\#q \approx \sim p\supset(\Diamond p\supset\Diamond q)\ \&\ (p\supset\sim q)$. For references and a further development, see my "A Binary Sheffer Operator which Does the Work of Quantifiers and Sentential Connectives", pp. 262-264 in *The Notre Dame Journal of Formal Logic* Volume XX, Number 2, April 1979.

of truth-evaluation—possibly exploiting structural relations among them, such as accessibility relations among possible worlds, or the ordering of time and place co-ordinates. The semantic interpretants of expressions are in the first instance functions from points of evaluation to extensions or truth-values. This is one natural way to capture the element of *generality* that Ryle insisted was present in all endorsements of inferences:

> …some kind of openness, variableness, or satisfiability characterizes all hypothetical statements alike, whether they are recognized "variable hypotheticals" like "For all *x*, if *x* is a man, *x* is mortal" or are highly determinate hypotheticals like "If today is Monday, tomorrow is Tuesday.[13]

By contrast to such two-stage approaches, semantics done in terms of incompatibility is *directly* modal. One may, if one likes, think of the incompatibility of *p* and *q* as the impossibility of both being *true*. But that characterization in terms of truth is entirely optional. Incompatibility is itself already a modal notion, and for semantic purposes we can treat it as primitive. The explication I have offered is in pragmatic terms: saying (in terms of the normative notions of commitment and entitlement) what one must *do* in order to be taking or treating two claims *as* incompatible. The element of generality comes in because in assessing entailments we look at *all* the claims that are incompatible with the conclusions and the premises. One claim is an incompatibility consequence of another only if there is *no* set of sentences incompatible with the conclusion and not with the premises. And here it is important that the potential defeasors are not limited to sentences that are *true*. Even if as a matter of fact all the coins in my pocket are copper, that a coin is in my pocket does not *entail* that it is copper,

[13] Gilbert Ryle " 'If', 'So', and 'Because' ", pp. 302-318 in Black, Max (ed.) *Philosophical Analysis* [Prentice Hall, 1950], p. 311.

since ‘This coin is silver’ is incompatible with its being copper, but not with its being in my pocket, even though it is not *true* that it is in my pocket. For, as we want to say, it *could* be in my pocket: that non-actual state of affairs is *possible*. That modal fact is reflected in the fact that a coin’s being silver is not *incompatible* with its being in my pocket. The idea that I want to explore is that once we have properly learned the lesson that modality matters in semantics because counterfactually robust inferences are an essential aspect of the articulation of the conceptual contents of sentences, the way is opened up to a *directly* modal semantics, which does *not* make what now appears as an unnecessary preliminary *detour* through assessments of *truth*.

This is all very abstract. In order to see incompatibility semantics in action, we should look to the case where the criteria of adequacy of a semantics are clearest: namely, to semantics for *logical* vocabulary. That, after all, is where possible worlds semantics cut its teeth.

Section 4: Introducing Logical Operators

The notion of *incompatibility* can be thought of as a sort of conceptual vector-product of a *negative* component and a *modal* component. It is *non*-com*possibility*. To use this semantic notion to introduce a negation operator into the object vocabulary, we must somehow isolate and express explicitly that negative component. The general semantic model we are working with represents the content expressed by a sentence by the set of sets of sentences incompatible with it. So what we are looking for is a way of computing what is incompatible with negated sentences (and, more generally, with sets of sentences containing them). Since we do not have any sort of yes/no evaluation of sentences in the picture (not even a relativized one), we cannot approach negation as a kind of reversal of polarity. How else might we think about it? Incompatible sentences are Aristotelian contraries. A sentence and its negation are contradictories. What is the relation between these? Well, the contradictory is *a* contrary: any sentence is incompatible with its negation. What distinguishes the contradictory of a sentence from all the rest of its contraries? The contradictory is the *minimal* contrary: the one that is entailed by all the rest. Thus every contrary of 'Plane figure *f* is a circle,'—for instance '*f* is a triangle,' '*f* is an octagon,' and so on—entails '*f* is *not* a circle.' *Blue, green, yellow* all entail *not-red*. For any sentence *p* we can already pick out its contraries, that is, the (sets of) sentences that are incompatible with it. And we already have an entailment relation, defined wholly in terms of incompatibility. So we have all the resources needed to say that some other sentence *q* is the negation of *p* just in case *q* is the *minimal* incompatible of *p*: the one entailed by everything else incompatible with it. I take it that Hegel means to be endorsing this order of explanation when he insists that *determinate* negation is to be understood as more fundamental in the order of explanation than *formal* negation, which is to be understood in terms of it.

It may happen that in some standard interpretation of the vocabulary to which *p* belongs, there already is such a *q*. Then we are in a position to identify it as the negation of *p* just in case $\forall X \subseteq L[X \in I(\{p\}) \rightarrow X|=q]$. But we cannot count on every sentence already having such a negation in every interpretation. So we need to introduce new sentences, of the form *Np*, on the basis of this relation. Inspection of the definition of incompatibility entailment yields the result that *Np* will be an inferentially minimal incompatible of *p* if and only if a set of sentences is incompatible with it just in case that set entails *p*. This is equivalent to saying that what is incompatible with the negation of *p* is what is incompatible with *every* set of sentences incompatible with *p*—that is, that the incompatibility set of *Np* is just the *intersection* of the incompatibility sets of everything incompatible with *p*.

This definition lets us recursively add, for every sentence *p* of the language, its negation *Np*, and to compute the incompatibility sets of those negations so as to satisfy the principle that everything incompatible with *p* entails *Np*. That is, if L is our initial vocabulary, thought of as consisting of primitive proposition letters, and I is a standard incompatibility interpretation on it, this definition lets us extend that incompatibility interpretation I to I_N, defined on the extended language L_N which contains L, and for every sentence *x* in L_N contains a sentence of the form *Nx*. Extending the incompatibility relation to apply to sets of sentences that include arbitrarily iterated negations automatically extends the incompatibility consequence relation, which is defined in terms of it. And it is easy to show that that extension is inferentially conservative—that is, that the extended consequence relation does not add or subtract any consequences that involve only the old vocabulary.

What are the properties of negation, given this incompatibility semantics? It turns out to have all the familiar and desirable properties we expect in a negation:

- Because *p* and *Np* are guaranteed to be incompatible, every set of sentences that contains or entails both—what we are now in a position to characterize as the *inconsistent* sets of sentences—is guaranteed to be incoherent.
- Negation contraposes appropriately with incompatibility entailment. That is, p|=q if and only if Nq|=Np.
- And every sentence is incompatibility-equivalent to its double negation: p|=NNp and NNp|=p.

Further logical properties of negation depend on its interaction with other connectives, and so must be considered after we have introduced them.

~~1.~~ So the procedure is to start with a *material* incompatibility-and-consequence structure that articulates the contents of *non*-logical vocabulary, and on that basis introduce *logical* vocabulary—in this case negation—whose content is derived from that of the non-logical vocabulary on which it is based. A corresponding procedure permits the introduction of *conjunction*. Here the most important fact to acknowledge is that something can be incompatible with a conjunction even though it is not incompatible with either conjunct. That the fruit in my hand is a blackberry is incompatible with its being red *and* ripe, even though it is not incompatible with either one individually. This is the phenomenon that led us to think about incompatibility relations among *sets* in the first place. And that is the clue as to how to compute the incompatibilities of conjunctions. What is incompatible with the conjunction *Kpq* should just be whatever is incompatible with the *set* {p,q}—more generally, what is incompatible with any set containing conjunctions should just be whatever is incompatible with the set that results from replacing those conjunctions by their conjuncts. Once again we can introduce conjunctions recursively and conservatively in this way, along with negations, so as to extend any standard incompatibility relation by computing incompatibilities for all sentences formed from basic vocabulary of primitive proposition letters by arbitrary iterations of K and N.[14]

Formatted: Indent: First line: 36 pt, Space Before: 0 pt, No bullets or numbering, Don't adjust space between Latin and Asian text

It is easy to show that under this definition, conjunction acts like conjunction. It is obvious from the semantic definition that {p,q}|={Kpq}, and it follows immediately from the

[14] Here we must make sure that the recursion is *path-independent*: that it does not make any difference in what order we apply the semantic rules in computing the incompatibility sets interpreting compound sentences. The end of Appendix 1 includes such a proof.

persistence of incompatibility that {Kpq}|={p} and {Kpq}|={q}. It is less obvious, but true, that this definition also validates the principle that if p|=q and p|=r, then p|=Kqr: that if p entails q and p entails r, then p entails their conjunction.[15]

In fact, conjunction behaves classically, and furthermore, interacts with negation in the familiar ways: the full logic is distributive. To make a long story short (the details can be found in the Appendix), the logic generated by these semantic definitions of negation and conjunction in terms of incompatibility is just classical logic. I will say something a bit further along about why, on reflection, this ought to be a surprising result. For now it is perhaps enough to notice that negation and conjunction are not interpreted semantically by anything at all like *truth*-functions. As we'll see, their semantics is in both cases intensional in definition, but nonetheless extensional in result—at least in the sense that the it yields just the theorems of classical two-valued logic.

The conditional that corresponds to the classical horseshoe, DeMorgan-defined from negation and conjunction as not both *p* and *not-q* (NKpNq), codifies the incompatibility-consequence relation, in that in every standard incompatibility-interpretation Cpq is valid (|=Cpq) just in case *p* incompatibility-entails *q* (p|=q).

What sets incompatibility semantics apart, however, is that we can exploit the fact that *incompatibility* is a *modal* semantic primitive to introduce *modal* logical vocabulary in the very

---

[15] For as we have seen, the definition of conjunction in terms of incompatibilities meeting the conditions of Persistence and Partition allows that what is incompatible with the conjunction of *q* and *r* may be much larger than the union of the incompatibility-set of *q* and that of *r*. And the intended interpretation requires it. Thus s=<u>a is a McIntosh</u> is compatible with q=<u>a is a ripe apple</u> and it is compatible with r=<u>a is a green apple</u>. But it is *in*compatible with their conjunction. So in this example, s∈ I({q,r})-I({q})∪I({r}). Knowing that p|=q and p|=r settles it that I({q})⊆ I({p}) and I({r})⊆ I({p}), and so that I({q})∪I({r})⊆ I({p}). But that is not enough to show that I({q,r})⊆ I({p}). What guarantees that I({q,r})-({q})∪I({r})⊆I({p})? The answer is that for the special case in which p|=q, so that I({q})⊆ I({p}), I({p,q})=I({p})∪I({q})=I({p}). And that turns out to be enough to ensure the result. What makes entailment of the conjuncts nonetheless sufficient for entailment of the conjunction is that the incompatibility framework validates the principle of

***Redundancy*:** **∀X,Y[(I(Y)⊆I(X))→(I(X∪Y)=I(X))]**

(Proven as a lemma in Appendix 1). If *p* entails *q*, then what is incompatible with both of them is just what is incompatible with *p*.

same setting, and the very same terms, in which we introduce the classical non-modal logical vocabulary.

On the semantic approach I am pursuing, to introduce a connective, one specifies how to compute its incompatibilities. So the question is: what intuitively should be taken to be *incompatible* with *necessarily p* (□p)—that is, with the *necessity of p*? Put otherwise, what claims rule out the *necessity* of *p*? Clearly anything incompatible with *p* is incompatible with *necessarily-p*. (Given the definition of entailment, this just says that the rules for computing the incompatibilities of *necessarily p* (I({Lp}) should ensure that *necessarily-p* entails *p* (Lp|=p.) But what else is incompatible with the *necessity* of *p*, besides the things that are incompatible with *p*? Here is the basic thought: **To be incompatible with *necessarily-p* is to be** (self-incompatible or) **compatible with something that *does not entail p***. For anything compatible with something that does not entail *p* is compatible with something that does not necessitate *p*, and so leaves open the possibility that *p* is not necessary.[16] This idea motivates the following definition:

***L Introduction*:**

$X \in I(\{Lp\})$ iff $X \in I(X) \vee \exists Y[X \notin I(Y) \& (Y |\neq \{p\})]$. [17]

A similar line of thought applies to *possibility* in relation to incompatibility, permitting us to introduce *possibly-p* (Mp) as well as *necessarily-p* (Lp). Whatever is incompatible with *possibly-p* should be incompatible with *p* (ensuring that *p* entails *possibly-p*, p|=Mp). But only

[16] If instead (as I first tried) one takes it that what is incompatible with Lp is what is compatible with something *incompatible* with p (rather than the wider class of things that do not entail p), the resulting operator combines with N in such a way that p≈LMp≈MLp—that is, it is one in which not only the Brouwer formulae MLp|=p and p|=LMp hold, but also their converses. As Tim Williamson (and, independently, Cian Dorr) pointed out—for which many thanks—this yields a system that is modally degenerate in the sense that KpNLp, which says that p is contingently true, turns out to be self-incompatible (the so-called "Fitch Paradox"). For KpNLp|=MLKpNLp, and LKpNLp is self-incompatible because LKpNLp|=KLpLNLp|=KLpNLp, which is a contradiction.

[17] The first clause ensures both that Lp|=p and that self-incompatibles are incompatible with {Lp}, if they are with {p}, as is needed to preserve the fundamental structural principles of Persistence and Partition.

some things that rule out *p* also rule out the *possibility* of *p*. Which are those? Here is an idea: **To be incompatible with *possibly-p* is to be incompatible with *everything* that is *compatible* with something *compatible* with *p*.** For anything compatible with something compatible with *p* is compatible with something that leaves the *possibility* of *p* open.[18] In symbols: ***M Introduction*:** $\mathbf{X{\in}I(Mp)\ iff\ \forall Y[Y{\in}I(Y){\vee}\exists Z[Y{\notin}I(Z)\&Z{\notin}I(p)]\rightarrow Y{\in}I(X)]}$. It turns out to be straightforward to show that according to these definitions, *possibly-p* is incompatibility-equivalent to *not-necessarily-not-p*, and *necessarily-p* is incompatibility-equivalent to *not-possibly-not-p* (Mp≈NLNp, and Lp≈NMNp), given the way we have defined negation above. So these definitions fit together in the way we would expect.

Given this vocabulary, we can introduce an incompatibility connective *I* into the language, defined by *not-possibly-both-p-and-q*, or, equivalently, *necessarily-not-both-p-and-q*. Intuitively, *Ipq says that p* and *q* are incompatible. Then in any given incompatibility interpretation, *Ipq* is valid in that interpretation just in case *p* is semantically incompatible with *q* according to it. ({p}∈I({q} iff |=Ipq).)

~~2.~~ To make another long story short (as before, the details can be found in the Appendices), the modal-logical theorems that are valid on all standard incompatibility frames given these definitions are just those of the familiar Lewis system S5. This is the system in which it is true both that whatever is necessary is necessarily necessary and that whatever is possible is necessarily possible. In the usual Kripke semantics, this is the modal logic generated by accessibility relations among possible worlds that are reflexive, symmetric, and transitive. In the

Formatted: Indent: First line: 36 pt, Space Before: 0 pt, No bullets or numbering

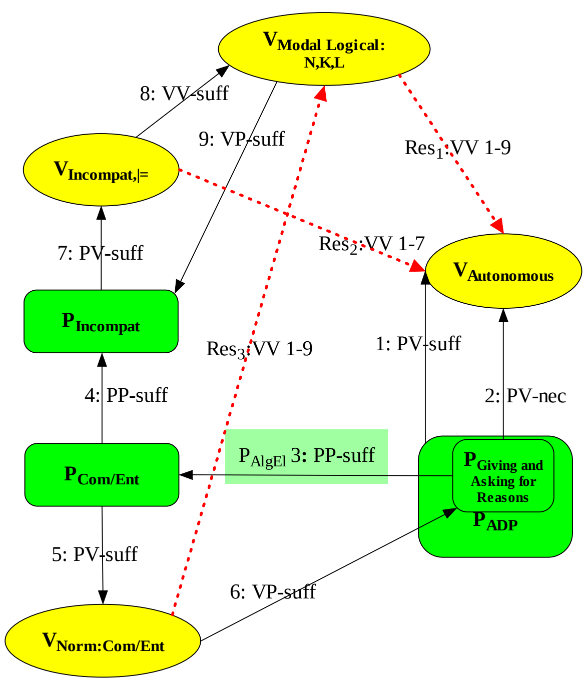
[18] If one says instead that what is incompatible with Mp is what is incompatible with everything compatible with p (instead of its having to rule out everything *compatible* with any such p-compatible), this turns out to be the N-dual of the modally degenerate semantic definition of L discussed in note 11, which results in p≈LMp≈MLp and so makes KpNLp contradictory.

tangled jungle of modal-logical systems, this is the unexciting, well-studied, well-behaved, plain-vanilla modal analogue of the classical non-modal propositional calculus.

Section 5: Meaning-Use Analysis

Here is a meaning-use diagram corresponding to this incompatibility semantics for modal logical vocabulary:

Incompatibility Semantics for Modal Logic

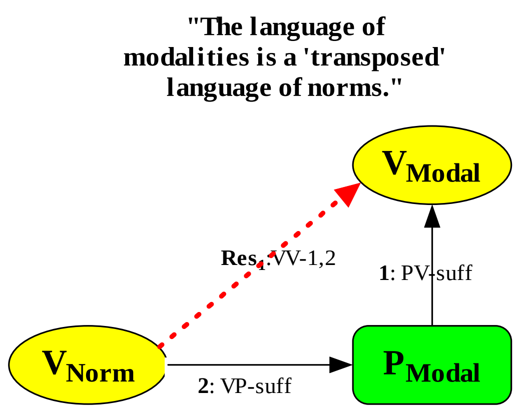


Here is some help in reading it:

- Basic MURs (1)-(3) are by now familiar. I have argued that every autonomous discursive practice must include practices of giving and asking for reasons—as part of the iron triangle of discursiveness, and that that involves distinguishing in practice between the deontic statuses of commitment and entitlement.

- We saw last time that that is sufficient to introduce normative vocabulary, specifically the deontic modal vocabulary of ‘commitment’ and ‘entitlement’, which is VP-sufficient to specify the triadic inferential substructure of practices of giving and asking for reasons. Those facts are represented by MURs (5) and (6).
- We also saw how practically distinguishing commitments and entitlements underwrites a notion of practical incompatibility of commitments, where commitment to one claim is taken or treated as sufficient to rule out entitlement to another. That is MUR (4), which permits the introduction of semantic metavocabulary letting one say that two claims are incompatible, and that claims stand in the relation of incompatibility-entailment, which is MUR (7).
- We have now seen how that semantic metavocabulary allows one to extend the original vocabulary by introducing modal-logical vocabulary (MUR (8)), which has the expressive power to define a connective that says in that object-vocabulary *that* two claims are incompatible: LNKpq. Basic MUR (9) accordingly exhibits modal-logical vocabulary as a kind of *semantic* metavocabulary for incompatibility.
- Complex resultant MUR Res_1 analyzes the sense in which the vocabulary of modal logic S5 is *implicit* in the use of any autonomous vocabulary. This analysis is a further cashing-out of what last time I called “The modal Kant-Sellars thesis.”
- Complex resultant MUR Res_2 codifies an analysis of the possibility of using <u>incompatibility</u> and <u>incompatibility-entailment</u> as a semantic metavocabulary for any autonomous vocabulary.

- Finally, complex resultant MUR Res_3 represents a new relation between the *normative* vocabulary of commitment and entitlement and the *modal* vocabulary of necessity and possibility. It represents a detailed analysis of a sense in which we could understand Sellars's dictum that "the language of modality is a 'transposed' language of norms." When I introduced that slogan in my third lecture, I suggested that the way to fill in Sellars's black-box notion of 'transposition' was in terms of the pragmatically mediated semantic relation of providing a *pragmatic* metavocabulary. I offered the following simple MUD as a representation of this relation:



We are now in a position to unpack what were there represented as two basic MURs. The PV-sufficiency of a set of modal practices for the deployment of a modal vocabulary in this simple diagram corresponds to the complex MUR that is the resultant of basic MURs (7) and (8) in the MUD for modal logic above. The VP-sufficiency of a normative vocabulary to specify those implicitly modal vocabulary-deploying practices now shows up as the complex MUR that is the resultant of relations (3), (4), (5), and (6). In fact we ought to include the other basic MURs that occur in our diagram, and identify the resultant representing the fact that normative vocabulary can serve as a pragmatic metavocabulary for modal vocabulary in the simple MUD with the complex resultant relation (3) in the more complex MUD.

That meaning-use diagram accordingly offers a detailed analytic reading of the Sellarsian claim that “the language of modality is a ‘transposed’ language of norms,” understood as asserting a complex pragmatically mediated semantic relation between deontic and alethic modal vocabularies. Now, Sellars’s claim might or might not be correct. And this interpretation-as-analysis of it might or might not be correct. But I take it to be a signal measure of the power of the metaconceptual apparatus of meaning-use analysis that it so much as permits the expression of this detailed a reading. And I have worked hard here today to justify MURs (4), (7), (8), and (9). Regardless of how successful those efforts have been, the fact that the meaning-use analysis tells us *exactly* what constellation of sub-claims we must argue for in order to justify the overall account seems to me to constitute concrete progress in our grasp of and control over the philosophic claims we make.

The complex pragmatically mediated semantic relation between deontic and alethic modal vocabularies that shows up here indicates that there is a deep relation between what in the previous lecture I called the “*modal* and *normative* Kant-Sellars theses.” In the final lecture I will have more to say about this relation, and about what it has to do with what is expressed by *intentional* vocabulary.

Section 6: Semantic Holism: Recursive Projectibility without Compositionality

Returning to ground-level, however, there are two more lessons I take to be of some potential philosophical significance that can be drawn from the construction of an incompatibility semantics for modal-logical vocabulary. The first concerns debates about semantic holism and compositionality.

As with the familiar Kripke semantics for modal vocabulary, the metavocabulary in which incompatibility semantics is conducted is entirely *extensional*. The semantic interpretants of sentences (and theories) are just sets (of sets of sentences), and the semantic interpretants of logically compound sentences are computed by purely set-theoretic operations on those sets. Also as with the Kripke semantics, this is possible because an overtly modal semantic primitive is appealed to: in the one case *accessible possible world*, in the other case *incompatibility*. (A principal difference is that I have offered a normative, deontic, pragmatic metavocabulary in which to *say* what you have to *do* to deploy that modal semantic primitive, and hence, eventually, the modal operators semantically defined in terms of it.)

The operators defined by the extensional incompatibility semantics are strongly *intensional*, however. We have noticed that one cannot in general compute the incompatibilities of a conjunction from the incompatibilities of its conjuncts. For something can be incompatible with a conjunction without being incompatible with either of its conjuncts. One might think that the fact that one *can* compute the incompatibility-interpretant of the conjunction *p&q* from the incompatibilities of the *set* {p,q}, though not from those of its individual members, to some extent mitigates this semantic non-extensionality. And things are in any case much worse with negation. The two commitments:

- to defining *p* as incompatibility-entailing *q* just in case everything incompatible with *q* is incompatible with *p*, and

- to understanding the negation of *p* as its inferentially weakest incompatible, that is, as what is incompatibility-entailed by everything incompatible with *p*

together have as a consequence that to be incompatible with *not-p* is just to be in the intersection of the incompatibility-sets of everything incompatible with *p*. But that means that we can hold fixed what is incompatible with *p*, and by varying the incompatibility-sets of some of *those* elements alter the incompatibility-set of *not-p*. It follows that in each incompatibility-interpretation, the semantic value of *not-p* is not determined by the semantic value of *p* alone, but only by it together with the semantic values of a *lot* of other sentences not mentioned in the formula—namely those incompatible with those incompatible with *p*.

This is what I meant earlier when I mentioned how surprised we ought to be that in the context of the definition of incompatibility-consequence, the highly intensional semantic definitions of negation and conjunction combine to yield classical propositional logic—a logic also generated by the paradigmatically extensional semantics of Boolean truth-functions. Now one might conclude from the fact that there is such a simple way to understand the semantics of that logic that the more complex incompatibility semantics for it amounts to *falsche Spitzfindigkeit*. Why bother with the fancier incompatibility semantics? One reason to think that the incompatibility semantics reveals something important even about the vocabulary of classical logic is that it exhibits that vocabulary as being in an important sense an expressively impoverished fragment extracted from a richer modal vocabulary, whose semantics is intelligible in exactly the same terms. In the next section I will present another argument to the effect that the incompatibility approach to the semantics of logic reveals deeper features of its relation to the use of ordinary non-logical vocabulary than do more familiar ones.

It is perhaps less surprising that the incompatibility definitions of what is expressed by necessity and possibility are also intensional, in much the same sense that negation is. So for instance, what is incompatible with *possibly-p* is what is *in*compatible with everything compatible with something compatible with *p*. Once again, we can fix the semantic interpretant of *p*, its incompatibility set, and still vary the semantic interpretant of *possibly-p*, by varying the semantic interpretants of things compatible with what is compatible with *p*. And the same phenomenon is exhibited by the incompatibility definition of *necessarily-p*.

This is to say that the classical and modal-logical connectives, as semantically defined by incompatibilities, do not have the *semantic sub-formula property*. That is, it is *not* the case that the semantic interpretants of logical compounds formed by applying those connectives is a function of the semantic interpretants of their components. It is *not* possible to compute the semantic values of arbitrary logical compounds of primitive sentences from the semantic values of the sentences and connectives from which they are formed. Another way to put this point is that the incompatibility semantics for these connectives is *not compositional*. It is in this precise sense a *holistic* semantics, in that what is incompatible with (and hence an element of the semantic value of) *not-p* or *necessarily-p* or *possibly-p* depends on what is incompatible with (and hence on the semantic value of) other sentences *q* linked with *p* in that they are compatible or incompatible with it, or incompatible with something that entails it, or compatible with something compatible with it. The holistic character of incompatibility semantics—whether for logical expressions such as ‘not’ or material, non-logical ones such as ‘triangular’—is a result of its codifying the so-to-speak *horizontal* dimension of semantic content, the one that is articulated by the relations of sentences to each other, rather than the *vertical* dimension, which consists in their relations to things that are not themselves sentences.

It is widely believed, and has been particularly forcefully argued by Jerry Fodor, that no holistic semantics can account either for the *projectibility* of language or for its *systematicity*, and hence not for its *learnability*. That is, it is argued that *only* on the assumption that semantics is *compositional* can we account for the determinateness of the semantic values of an indefinite number of novel compounds of simple expressions, for the fact that wherever some syntactic combinations of those simple expressions have semantic values so do others systematically related to them, and for the fact that speakers of a language can produce and understand an indefinite number of novel compounds, systematically related to one another by their modes of formation, upon mastering the use of the simple expressions and modes of formation.

But I think we are now in a position to see that those arguments cannot be right. They depend upon systematically overlooking the possibility of semantic theories that have the shape of the incompatibility semantics for classical and modal logical vocabulary we have been considering. For—and this is the key point—although that semantics is *not compositional* it *is* fully *recursive*. The semantic values of logically compound expressions are wholly determined by the semantic values of logically simpler ones. It is holistic, that is, noncompositional, in that the semantic value of a compound is not computable from the semantic values of its components. But this holism *within* each level of constructional complexity is entirely compatible with recursiveness *between* levels.

The semantic values of all the logically compound sentences are computable entirely from the semantic values of *less complex* sentences. It is just that one may need to look at the values of *many*—in the limit *all*—the less complex sentences, not just the ones that appear as sub-formulae of the compound whose semantic value is being computed. The semantics is projectible and systematic, in that semantic values are determined for all syntactically admissible compounds, of arbitrary degrees of complexity. It is learnable—at least, putting issues of contingent psychology aside, in the ideal sense we have been working with. For the capacity to distinguish the incompatibility-sets of primitive propositions is, in the context of the semantic definitions of the connectives in terms of incompatibilities I have offered, PP-sufficient by algorithmic elaboration for the capacity to distinguish the incompatibilities of all their logical (including modal-logical) compounds—and hence for the practical capacity to distinguish what is a consequence of what.

What semantic projectibility, systematicity, and learnability-in-principle require, then, is not semantic atomism and compositionality, but semantic recursiveness with respect to complexity. That is entirely compatible with the semantics being holistic, in the sense of lacking the semantic sub-formula property. And the argument for this claim is not merely the description of an abstract possibility. The incompatibility semantics for logical vocabulary provides an up-and-running counterexample to the implicit assumption that semantic recursiveness is achievable only by compositionality. Having compound expressions exhibit the semantic sub-formula property is only *one* way of securing recursiveness. The standard arguments for semantic compositionality are fallacious.[19]

So here is another side-benefit or philosophical spinoff from looking analytically at pragmatically mediated semantic relations between antecedently philosophically interesting vocabularies—a potentially valuable lesson we can learn from incompatibility semantics.

---

[19] A more charitable way to put things would be that compositionality—which really amounts to semantic recursiveness—has been confused with the semantic sub-formula property. Thus Fodor and Lepore's *The Compositionality Papers* [Oxford U. Press, 2002] opens with this definition:

> Compositionality is the property that a system of representations has when (i) it contains both primitive symbols and symbols that are syntactically and semantically complex; and (ii) the latter inherit their syntactic/semantic properties from the former. (p.1)

On this definition, the incompatibility semantics is fully compositional. But Fodor and Lepore go on to draw consequences from compositionality that in fact only follow from the semantic sub-formula property. So the confusion I am concerned to point out is in play, however we decide to specify it.

<u>Section 7: Consequence-Intrinsic Logic</u>

The order of explanation I have been pursuing up to this point

- starts with practices of giving and asking for reasons,
- argues that they are PP-sufficient for practices of deploying basic normative vocabulary—in particular the deontic modal vocabulary of 'commitment' and 'entitlement',
- uses that as a pragmatic metavocabulary that specifies how to deploy a modal concept of <u>incompatibility</u>,
- uses that as the basic semantic metavocabulary in which to define a *consequence* relation of incompatibility-entailment,
- and on that basis offers semantic definitions of logical vocabulary, including modal operators.

It is possible to exploit the pragmatic and semantic relations appealed to in this approach in service of a different order of explanation, however. In particular, instead of defining a semantic consequence relation in terms of a prior notion of *incompatibility*, we can start with a consequence relation—either a logical consequence relation or a material one that depends on the contents of the non-logical vocabulary articulating its premises and conclusions—and *impute* an incompatibility relation on that basis that will semantically generate just that consequence relation.

The idea is to hold fixed the principle that Y is a semantic consequence of X just in case everything incompatible with Y is incompatible with X, but to use that principle relating them to define an incompatibility relation among sets of sentences of the language that would generate whatever consequence relation we are given to begin with. To make this work, we have to ask what conditions a consequence relation defined on an arbitrary set of sentences must meet in order to make it possible to define from it an incompatibility relation such that sets of sentences X and Y stand in the consequence relation (which I'll write '$X \vdash Y$') just in case everything incompatible with Y is incompatible with X (which I will continue to write '$X \models Y$').

It turns out that two conditions suffice:

i. <u>General Transitivity</u>: $\forall X,Y,Z,W \subseteq L[\ (X \vdash Y \ \&\ Y \cup W \vdash Z) \rightarrow X \cup W \vdash Z)]$.

ii. <u>Defeasibility</u>: $\forall X,Y \subseteq L[\sim(X \vdash Y) \rightarrow$
$\exists Z \subseteq L[[\forall W \subseteq L[Y \cup Z \vdash W] \& \exists W \subseteq L[\sim(X \cup Z \vdash W)]\ ]$.

I will call any consequence relation meeting these conditions 'standard'. The first is a very minimal condition on consequence relations, which corresponds to the usual 'Cut' rule of sequent calculi:

$$\frac{\Gamma: A \text{ and } \Delta, A: B}{\Gamma, \Delta: B.}$$

The second says that if Y is *not* a consequence of X, then there is something that yields an absurdity—something that has *everything* as a consequence—when added to Y but not when added to X. In an Appendix, I show that if a consequence relation meets these two conditions, then it is possible to define an incompatibility relation that will generate exactly that consequence relation as incompatibility-entailment, by identifying incoherent sets as those that have all sets as their consequences, and then taking two sets to be incompatible if and only if their union is incoherent. That is, for every standard consequence relation, we can find a standard incompatibility relation that semantically validates it.

Inspection of that argument shows that the first, more natural, condition suffices to ensure that any consequence relation meeting them will induce, by the procedure indicated, an incompatibility relation guaranteed to validate all its consequences—that is, to show the *soundness* of the semantic codification by incompatibility of the consequence relation. The less common second condition is needed only for the converse result: to ensure that every pair of sets of sentences that stands in the incompatibility-entailment relation also stand in the original consequence relation—that is to show the *completeness* of the semantic codification by incompatibility of the consequence relation. So how interesting the representation theorem that comprises these soundness-and-completeness results is depends on how severe a constraint on consequence relations that second condition is.

The easiest way to assess that is to see what familiar, or otherwise interesting, consequence relations do and do not satisfy it. I have already argued, in effect, that sentences attributing ordinary compatible families of incompatible properties—paradigmatically, shapes and colors, and membership in various biological or physical kinds, but encompassing a great many others as well—exhibit material consequence relations that are *standard* in this sense. For those sentences stand in the material consequence relations that are defined by their incompatibilities, and those, the representation theorem shows, meet the two conditions of standardness. But what of others, which are not defined to begin with in terms of incompatibility? The consequence relations we understand best are *logical* consequence relations, defined on logical vocabulary by various sets of axioms concerning derivability. Perhaps at this point it comes as no surprise that the consequence relation characteristic of classical logic is a standard one. In that setting, a set of sentences can be taken to be *incoherent* just in case it is *inconsistent*, in that for some *p*, both *p* and ~*p* can be derived from it. Any superset of an inconsistent set is inconsistent, and from inconsistent sets one can derive everything. Treating two sets as *incompatible* just in case their union is inconsistent then yields the classical consequence relation under our usual definition of semantic incompatibility-entailment. And this result holds for the first-order quantificational calculus just as it does for the classical propositional calculus. The consequence relations of both logics are standard, and so can be completely codified semantically by incompatibility relations.

Logically inconsistent sets play just the same role in the consequence relations of standard Lewis modal systems, such as S4 and S5, as they do in the classical logic on which they are based. So these too are standard, and so semantically codifiable as incompatibility-entailment relations. Indeed, the features of inconsistency on which that codification depends—those in virtue of which it can serve as an incoherence relation on the basis of which we can define incompatibility—for instance, persistence and absurdity entailing everything, are shared also by intuitionistic negation.[20] Cut (and so the general transitivity condition), of course, holds in the intuitionistic setting. But Defeasibility does *not* hold, so its consequence relation is not standard.[21] Another important class of logics whose consequence relation is *not*

[20] In the logical case, that an inconsistent set of premises entails every proposition (*ex falso quodlibet*) follows from the elimination rule for intuitionistic negation. In Dummett's version (*Elements of Intuitionism*, 2nd Edition [Oxford University Press, 2000], p. 89) of a sequent-calculus formulation, this rule is:

$$\frac{\Gamma : A \quad \Delta : \neg A}{\Gamma, \Delta : B}$$

which evidently has as a special case:

$$\frac{\Gamma : A \quad \Gamma : \neg A}{\Gamma : B.}$$

[21] In looking into the intuitionist consequence relation and comparing it to that of classical logic, I noticed a fact of which I had previously been unaware. Unlike that of classical logic, the consequence relation of propositional intuitionistic logic is *categorical* in the following sense: given only the role of a sentence in the consequence relation, one can infer its logical form, up to logical equivalence. In both cases conjunctions can be distinguished by the facts that {p,q}|—p&q and p&q|—p and p&q|—q. (Though one cannot in those terms tell p&q from the logically equivalent p&q&q.) The difference (not surprisingly) lies in negation. In classical logic one cannot, from its role in the consequence relation alone, distinguish something of the form ~*p* from something of the form *p*, even those these are, to say the least, not logically equivalent. In the intuitionist case, by contrast, negations behave differently from non-negated sentences, in that double-negation elimination holds for them. That is, for any sentence p|—~~p|—~~~~p and so on. The converse, of course, does not hold generally. But if *p* is atomic, we also have the converse sequence: ~~~~~p|—~~~p|—~p. That is, compounds of atomic *p* with odd numbers of negations are all equivalent, but not those with even numbers of negations. This property suffices to distinguish negated propositions from non-negated ones, just in terms of their role in the consequence relation. (Though we still cannot tell ~~~p from the logically equivalent ~~~p&~p.) I take it that the fact that its consequence relation is in this sense categorical for logical form (up to logical equivalence)

standard is relevance logics. For they have as a defining feature that inconsistent sets do *not* act like incompatibility-incoherent ones in having everything as a consequence. In fact—and I find this an at least mildly curious fact—there is *nothing at all* about their behavior with respect to the logical consequence relation that distinguishes inconsistent from consistent sets in the relevance context. Whether that fact seems a virtue or a blemish turns on issues that would take us very far afield indeed.

Now the observation with which I shall close is this: our representation theorem shows that any consequence relation that meets the conditions of standardness—whether it be a material or a logical consequence relation—can be codified by a standard incompatibility relation definable in a natural way from that consequence relation. And we have seen that any standard incompatibility relation has a logic whose non-modal vocabulary behaves classically and whose modal vocabulary is S5, in the sense that the natural semantic definitions of such vocabulary in terms of incompatibility yields that logic. Putting these results together, we can say that in this precise sense, S5 (whose non-modal fragment is just classical logic) is the logic *intrinsic to* standard incompatibility relations, and hence standard consequence relations. But since not only classical logic, but *all* the usual modal logics—not only S5, but K, T, S3, S4, and B, have standard consequence relations, classical logic and S5 are the intrinsic logic of, for instance, S4, as well as the others. And although the consequence relation of intuitionistic logic is not standard, so not codifiable by a standard incompatibility relation, in a natural sense (explained in the Appendix) it does implicitly contain a standard consequence relation, and so in this somewhat extended sense it, too, has PC+S5 as its intrinsic logic. And in the same sense, so does intuitionistic S4. Relevance logic aside, the logic that is in this sense *intrinsic* to the

---

constitutes a comparative virtue of intuitionist over classical propositional logic, since it shows that the logical consequence relation and logical form fit together better in the one case than in the other.

consequence relations of *most* other familiar logics is classical S5. S5 accordingly has some claim to being *the* modal logic of consequence relations, whether material or logical.

The claim that classical S5 is the logic that is *intrinsic* to consequence relations very generally—to those of other logics as well as of the material consequence relation generated by the contents of non-logical vocabulary—in the sense of 'intrinsic' defined by the two processes of:

- imputing an incompatibility relation on the basis of a consequence relation, and
- introducing logical connectives semantically in terms of that incompatibility relation,

and is in that sense privileged as *the* logic of inference, runs in many ways counter to the expressive approach to distinctively logical vocabulary I have been pursuing in these pages. I take it that the three big traditional questions in the philosophy of logic are:

i. how to demarcate *logical* vocabulary from non-logical vocabulary,

ii. how to understand the *relation* between logical and non-logical vocabulary,

iii. what is—and what it means to say that it is—the *correct* logic.

Taking our lead from Ryle's conception of conditionals as inference-tickets, I have suggested answering the first two questions in terms of logical vocabulary being elaborated from and explicative of the inferential articulation of non-logical vocabulary. But if that is right, then the third question is ill-put.

For if we think of conditionals as codifying proprieties of inference, expressing the endorsement of inferences (that is, this is their pragmatic function, this is what one is *doing* in endorsing a conditional or making a conditional claim), then we need to think of the different *kinds* of pragmatic propriety (normative), the different kinds of Rylean permission that could be codified:

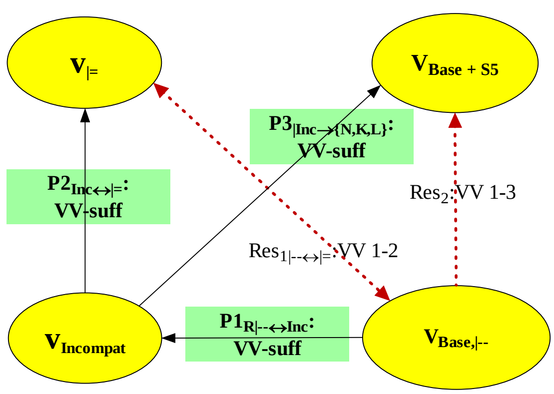
- The classical conditional: the sense of 'good' in which an inference is good if it does not have true premises and a false conclusion. This may be a poor sort of virtue, but at least we can agree that it is a bad thing about an inference if it does not have this sort of good-making property.
- The intuitionist conditional: the sense of 'good' in which an inference is good if there is a procedure for turning a proof of or dispositive argument for its premises into a proof of or dispositive (commitment-preserving) argument for its conclusion.

- The strict implication conditional: the sense of 'good' in which an inference is good if it is impossible for its premises to be true and its conclusion not be true.

On this line, the important philosophical question is not which one of these conditionals is *correct*, but how to understand, for *any* conditional, what sort of propriety, virtue, or evaluation of inference it makes explicit. But the discovery that one logic is privileged above the others in that it is the one that is *intrinsic* to the consequence relations they induce suggests the possibility of a different sort of answer to the third question above, even within the context of the expressive view of logical vocabulary.

The concept of the logic that is *intrinsic* to the consequence relation characteristic of some vocabulary (whether logical or not) is the concept of a new kind of semantic relation between vocabularies. It is mediated by the vocabulary of incompatibility, in terms of which, on the one hand, the consequence relation is codified (according to the representation theorem for extracting standard incompatibility relations from standard consequence relations), and in terms of which, on the other hand, logical vocabulary is semantically introduced. Here is a Meaning-Use Diagram for this new complex resultant semantic relation:

Consequence-Intrinsic Logic



I have been concerned to fill in the three sets of practices that implement the basic VV-sufficiency relations of which the relation of *intrinsicness* of a logic to a vocabulary is the resultant:

- imputing a standard incompatibility relation from a standard consequence relation,
- defining incompatibility-entailment in terms of that incompatibility relation, and
- semantically introducing logical vocabulary, including modal vocabulary, in terms of incompatibility.

The complex resultant MUR this constellation of basic MURs defines is a semantic relation that, apart from this methodology, we would never have been in a position to notice: the relation between logical vocabularies and other vocabularies, when the logical vocabulary is *intrinsic* to the consequence relation *characteristic* of the other vocabulary.

Having put the technical material behind us, next time I will shift focus by turning attention to what is expressed by *intentional* vocabulary and take up once again the issue of the relations between normative and modal vocabularies—or if you prefer, between deontic and alethic modalities—as it bears on the nature of intentionality.

END

Appendix 1: Incompatibility Semantics

Note: The formal work presented in this appendix is the result of a collaboration with my colleague Alp Aker. I came up originally with versions of the semantic definitions, the introduction rules for the connectives, proofs of the validity of the various logical principles involving those connectives, and most of the other results reported in these Appendices. Alp vastly improved our understanding of the incompatibility semantics by shifting to a definition of incompatibility entailment that is disjunctive on the right (I had used one that was conjunctive on the right). This made it possible for him to formulate the reduction formulae for the connectives, which made all the proofs cleaner and easier. (I had been working directly from the basic definitions, which required extremely laborious derivations from very quantificationally complex formulae.) It also made it possible for him to prove the crucial metatheorems showing that the semantic connective definitions determine extensions of incompatibility frames over a set of non-logical sentences to arbitrarily complex logical compounds of them in a way that is inferentially conservative and unique. (The explicit recursions I had attempted to use to the same end proved unworkable.)

1 Definitions and Axioms

1.1 Incoherence, Incompatibility, and Entailment

We are given a language *L*, which is a set of sentences. An *incoherence property* Inc is a subset of $P_f(L)$ (the collection of finite subsets of *L*) that is upward closed. That is, incoherence properties satisfy:

Axiom (Persistence). For finite $X,Y \subseteq L$, and $X \subseteq Y$, if $X \in Inc$ then $Y \in Inc$.

Incoherence is a generalization of *inconsistency* to the case of non-logical properties. Persistence says that if a set is incoherent, adding more sentences to it cannot cure that condition. An ordered pair <L,Inc> whose second element is an incoherence property defined over the first element is a *standard incoherence frame* on L.
(Henceforth all subsets are understood to be finite, and all frames to be standard—that is, their incoherence property to satisfy Persistence.)

Incompatibility An *incompatibility function* I is a function from $P_f(L)$ to $P(P_f(L))$. Incoherence relations are related one-to-one to incompatibility functions by:

Axiom (Partition). $X \cup Y \in Inc$ iff $X \in I(Y)$.

That is, two sets of sentences are incompatible just in case their union is incoherent. An ordered pair <L,I> is a *standard incompatibility frame* on L.
Persistence of incompatibility: It follows from the Persistence of incoherence that if $X \in I(Y)$ and $X \subseteq X'$, then $X' \subseteq I(Y)$.
Symmetry of incompatibility: It follows that if $X \in I(Y)$, then $Y \in I(X)$.

Given an incoherence relation or an incompatibility function we have the following relation of *incompatibility-entailment*:

Entailment $X \models_I Y$ **iff** $\bigcap_{p\in Y} I(\{p\}) \subseteq I(X)$**.**

When Y is empty we read $\bigcap_{p\in Y} I(\{p\})$ as equivalent to $P_f(L)$. (Thus $X \models_I \varnothing$ is equivalent to $X \in$ Inc.) We index entailment relations by incompatibility functions (or, equivalently, by incoherence properties). The underlying idea is that one sentence incompatibility-entails another if and only if everything incompatible with the conclusion is incompatible with the premise. That idea is generalized to a relation between sets in a convenient and natural way. The heuristic meaning of $X \models \{y_1, ... ,y_n\}$ is that X entails y_1 *or* .... *or* y_n.

Validity X is *valid* **iff** $Y\in \bigcap_{p\in X} I(\{p\}) \rightarrow Y\in$ **Inc.**

Thus $\models X$ is equivalent to X's being valid.

This definition of entailment has the following consequences:

1.1.1 (Weak Transitivity): If (p|=q and q|=r) then p|=r.

Proof. Since $\subseteq$ is transitive, so is |=. By the definition of entailment, the antecedent holds just in case everything incompatible with r is incompatible with q and everything incompatible with q is incompatible with p, that is, $I(\{r\})\subseteq I(\{q\})\subseteq I(\{p\})$. So $I(\{r\})\subseteq I(\{p\})$ and p|=r.

Notice that the corresponding transitive principle for *sets* of premises does *not* hold, since adding premises (on the left of '|=') is conjunctive, while adding conclusions (on the right of '|=') is disjunctive.

1.1.2 If p|=q and |=p, then |=q.

Proof: |=p just in case everything incompatible with p is incoherent. If p|=q, then everything incompatible with q is incompatible with p, hence incoherent. So everything incompatible with q is incoherent. So |=q.

1.2 Connective Definitions

We have three axioms that govern the behavior of the connectives N, K, and L:

Axiom (Negation Introduction; NI). $X \cup \{Np\} \in$ **Inc iff** $X \models \{p\}$**.**

In order to introduce conjunction, it is helpful to define functions f and F as follows. For $p \in L$, $f(p) = \{q, r\}$ if p = Kqr and {p} otherwise. For $X \subseteq L$, $F(X) = \bigcup_{p \in X} f(p)$.

Axiom (Conjunction Introduction; CI). $X \in$ **Inc iff** $F(X) \in$ **Inc.**

Axiom (L Introduction; LI). $X \cup \{Lp\} \in Inc$ iff $X \in Inc$ or $\exists Y[X \cup Y \notin Inc$ and $Y|\neq\{p\}]$.

For ease of reading we sometimes drop brackets around sets and sometimes use the comma to denote set union. Thus we can, e.g., write I(p,X) instead of $I(\{p\} \cup X)$ and X,Np|= q instead of $X \cup \{Np\}|=\{q\}$. We also write Apq as an abbreviation of NKNpNq and Mp as an abbreviation of NLNp.

2 Basic Lemmas

<u>2.1 (Weakening)</u>: If X |= Y, then X, W |= Y, V.
Proof: Suppose X |= Y. Then $\cap_{p\in Y} I(p) \subseteq I(X)$. Since $I(X) \subseteq I(X,W)$ and $\cap_{p\in Y \cup V} I(p) \subseteq \cap_{p\in Y} I(p)$ we know $\cap_{p\in Y \cup V} I(p) \subseteq I(X,W)$. Thus X,W |= Y,V.

<u>2.2 (Cut)</u>: If X |= q,Y and q,W |= V then X,W |= Y,V.
Proof: We want $\cap_{p\in Y \cup V} I(p) \subseteq I(X,W)$. So suppose $S \in \cap_{p\in Y \cup V} I(p)$. We then want $S\in I(X,W)$. This is equivalent to $S\cup W\in I(X)$, which—because X |= q,Y —holds just in case both $S\cup W\in I(q)$ and $S\cup W\in I(p)$ for all $p\in Y$.
Now, because q,W |= V and by supposition $S\in I(p)$ for all $p\in V$, we know $S\in I(W, q)$ and thus $S\cup W\in I(q)$. We also know by supposition that $S\in I(p)$ for all $p\in Y$ and so $S\cup W\in I(p)$ for all $p\in Y$.

3 Some Modal Properties

We begin with two small points:

<u>3.1:</u> $\exists Y[Y|\neq \varnothing$ & $Y|\neq p]$ iff $|\neq p$.

Proof: ($\Rightarrow$) Instantiate to get $X|\neq p$. This implies $|\neq p$.
($\Leftarrow$) We show the contrapositive. Suppose $\forall Y(Y |= \varnothing \vee Y|= p)$. Since $Y|= \varnothing$ implies Y|= p we have $\forall Y(Y|= p)$. Then if X is incompatible with p it is incompatible with everything and so |= p.

<u>3.2</u>: Lp $|= \varnothing$ iff $|= \varnothing$ or $|\neq p$.

Proof: Instantiating the L Introduction rule with $\varnothing$ as X we get Lp $|= \varnothing$ iff $|= \varnothing$ or $\exists Y[Y|\neq \varnothing$ & $Y|\neq p]$. By 3.1 the latter disjunct is equivalent to $|\neq p$, and we thus have Lp $|= \varnothing$ iff $|= \varnothing$ or $|\neq p$.

Note that the disjunct $|= \varnothing$ ("the True implies the False") is not necessarily false. It is equivalent to $P_f(L) \subseteq Inc$, which condition is fulfilled in the degenerate case in which every (finite) set of sentences is incoherent.

We now get the **basic observation** about modal formulae:

<u>3.3</u>: $X,Lp|=\varnothing \Leftrightarrow X|=\varnothing \vee Lp|=\varnothing$.

Proof: ($\Rightarrow$) X,Lp |= $\varnothing$ is by definition equivalent to X|=$\varnothing$ $\vee$ $\exists$Y[X,Y|$\neq\varnothing$ & Y|$\neq$p]. Since X,Y|$\neq\varnothing$ implies Y|$\neq\varnothing$ we know $\exists$Y[X,Y|$\neq\varnothing$ & Y|$\neq$p] implies $\exists$Y[Y|$\neq\varnothing$ & Y|$\neq$p], which, instantiating L-Introduction with X= $\varnothing$, implies Lp|= $\varnothing$.
($\Leftarrow$) This follows from Persistence.

It is now easy to prove:

<u>3.4 (Necessitation)</u>: |=p $\Rightarrow$ |=Lp.

Proof: Suppose |= p and X,Lp|= $\varnothing$. We want X incoherent. Then by the basic observation either X|=$\varnothing$, as desired, or Lp|=$\varnothing$, in which case either |=$\varnothing$ or |$\neq$p. In the former case every set is incoherent, including X. The latter case contradicts our supposition and so can't occur.

The entailment that corresponds to the T axiom CLpp is also easy:

<u>3.5</u> : Lp|=p.

Proof: Suppose not. Then there is some Z such that Z,p|= $\varnothing$ and Z,Lp|$\neq$ $\varnothing$. From the latter it follows that Z|$\neq\varnothing$ and Lp|$\neq\varnothing$. From Lp|$\neq\varnothing$ it follows that |= p by 3.1. But then since Z,p|=$\varnothing$ we know that Z|=$\varnothing$, which implies Z,Lp|=$\varnothing$, which is a contradiction.

Using 3.5 and Cut we easily get the following useful rule:

<u>3.6</u>: |=Lp $\Rightarrow$ |=p.

We can also extend the basic observation:

<u>3.7</u>: X,Lp |= Y iff X|= Y $\vee$ Lp|= $\varnothing$.

Proof: X,Lp|= Y is $\forall$Z(Z $\in \cap_{p\in Y}$ I(p) $\Rightarrow$ X,Lp,Z|=$\varnothing$), which by 3.3 is $\forall$Z(Z $\in \cap_{p\in Y}$ I(p) $\Rightarrow$ X,Z|= $\varnothing$ $\vee$ Lp|= $\varnothing$). This is in turn is equivalent to $\forall$Z(Z $\in \cap_{p\in Y}$ I(p) $\Rightarrow$ X,Z|=$\varnothing$) $\vee$ Lp|= $\varnothing$, which is X|=Y $\vee$ Lp|=$\varnothing$.

We can also show that with respect to any particular incompatibility frame, every necessary proposition is either contradictory or valid:

<u>3.8</u>: Lp|=$\varnothing$ or |=Lp.

Proof: Suppose Lp|$\neq\varnothing$. Then |= p by 3.2 and so |=Lp by 3.4.

Of course, this does not mean that every necessary proposition is either incoherent in all frames or valid in all frames, but only that it is incoherent-or-valid in all frames.

The next result is a dual of 3.7:

<u>3.9</u>: X|= Y,Lp iff X|=Y ∨ |=Lp.

Proof. (⇒) Suppose not. Then not X|= Y and |≠ Lp. By 3.8 we then know Lp|=∅. Since X |≠ Y there is some Z ∈ $\cap_{p\in Y}$ I(p) with Z ∉ I(X). Since Lp|= ∅ we know Z is incompatible with Lp. Since X|=Y,Lp we then have Z incompatible with X, which is a contradiction.
(⇐) This follows from Weakening.

4 *Semantic Reduction*

We can, for any given entailment X |= Y, show that it is equivalent either to another entailment that mentions only atomic formulae or to a Boolean combination of entailments that mention only atomic formulae. This shows that the introduction rules we have given for the connectives unambiguously determine standard incoherence frames for languages including arbitrarily logically complex formulae defined by sequential applications of those connectives to atomic formulae, relative to every standard incoherence frame defined on those atomic formulae.

4.1 Reduction Schemata for Non-modal Connectives

<u>4.1.1 (Left Negation)</u>: X |= Y, p iff X, Np |= Y.

Proof: (⇒) Suppose X |= Y, p. We want $\cap_{r\in Y}$ I(r) ⊆ I(X, Np). Now suppose Z ∈ $\cap_{r\in Y}$ I(r). Then Z ∪ {Np} ∈ $\cap_{r\in Y}$ I(r). We also know p ∈ I(Np) and so Z ∪ {Np} ∈ I(p). Then since X |= p, Y we have Z ∪ {Np} ∈ I(X) and so Z ∈ I(X, Np).
(⇐) Suppose X, Np |= Y. We want $\cap_{r\in Y \cup \{p\}}$ I(r) ⊆ I(X). So suppose Z ∈ $\cap_{r\in Y \cup \{p\}}$ I(r). Since X, Np |= Y we have Z ∈ I(X, Np). Then Z, X |= p. Since Z ∈ I(p) Cut gives Z ∈ I(Z, X) and thus Z ∈ I(X).

<u>4.1.2 (Right Negation)</u>: X |= Y, Np iff X, p |= Y.

Proof: (⇒) Suppose X |= Y, Np. We want $\cap_{r\in Y}$ I(r) ⊆ I(X, p). So suppose Z ∈ $\cap_{r\in Y}$ I(r). Then Z ∪ {p} ∈ $\cap_{r\in Y}$ I(r). We also know Z ∪ {p} ∈ I(Np). Since X |= Y, Np it follows that Z ∪ {p} ∈ I(X), or Z ∈ I(X,p).
(⇐) Suppose X, p |= Y. We want $\cap_{r\in Y \cup \{Np\}}$ I(r) ⊆ I(X). So suppose Z ∈ $\cap_{r\in Y \cup \{Np\}}$ I(r). Then since Z ∈ $\cap_{r\in Y}$ I(r) and X, p |= Y we have Z ∈ I(X, p) or Z, X, p |= ∅. Since Z ∈ I(Np) we have Z |= p. From Z, X, p |= ∅ and Z |= p we can apply Cut to get Z, X |= ∅, or Z ∈ I(X), as desired.

<u>4.1.3 (Left Conjunction)</u>: X, Kpq |= Y iff X,p,q |= Y.

Proof: By definition I(X, Kpq)=I(X, p, q). The result follows immediately.

<u>4.1.4 (Right Conjunction)</u>: X |= Y, Kpq iff X |= Y, p and X |= Y, q.

<u>Proof</u>: (⇒) Suppose X |= Y, Kpq. Then $\cap_{r\in Y \cup \{Kpq\}}$ I(r) ⊆ I(X). Equivalently, if Z ∈ I(p,q) and Z ∈ $\cap_{r\in Y}$ I(r) then Z ∈ I(X). But I(p) ⊆ I(p,q). Then if Z ∈ I(p) and Z ∈ $\cap_{r\in Y}$ I(r) it

follows that $Z \in I(p,q)$ and $Z \in \cap_{r\in Y} I(r)$ and so $Z \in I(X)$. Thus X |= Y, p. We can argue similarly to get X |= Y,q.
($\Leftarrow$) Suppose X |= Y, p and X |= Y, q. We want $\cap_{r\in Y \cup \{Kpq\}} I(r) \subseteq I(X)$. So suppose $Z \in I(Kpq)$ and $Z \in \cap_{r\in Y} I(r)$. If $Z \in I(Kpq)$ then Z, p, q |= $\varnothing$. By Cut and the fact that X |= Y, p we then have Z, X, q |= Y. Applying Cut again, this time with X |= Y, q, we get Z,X |= Y. Since by supposition $Z \in \cap_{r\in Y} I(r)$ we then know $Z \in I(Z,X)$, which is $Z \in I(X)$, as desired.

4.2 Reduction Schemata for Modal Connectives

<u>4.2.1 (Left Necessity)</u>: X, Lp |= Y iff X |= Y or |≠ p.

Proof. If we are in the degenerate frame everything implies everything else, so the result holds. If not, it follows from 3.7 and 3.2.

<u>4.2.2 (Right Necessity)</u>: X|= Y, Lp iff X |= Y or |=p.

Proof. Apply 3.9 and 3.6.

4.3 Semantic Reduction for Arbitrary Entailments

We now show the result alluded to in the beginning of this section.

<u>4.3.1 (Semantic Reduction Lemma; SR)</u>: Given a language L and frame Inc, let $L_{\varnothing}$ be the atomic fragment of L. Then X $|=_{Inc}$ Y iff $F(X_1|=_{Inc}Y_1, \ldots, X_n|=_{Inc}Y_n)$, where F is a Boolean function on n propositions and each X_i,Y_i is in $L_{\varnothing}$. Further, $F(X_1|=_{Inc}Y_1, \ldots, X_n|=_{Inc}Y_n)$ is determined only by the syntactical structure of the members of X and Y.

Proof: We actually show a slightly stronger result, viz. $F(X_1|=_{Inc}Y_1, \ldots, X_n|=_{Inc}Y_n)$ iff $G(X'_1|=_{Inc}Y'_1, \ldots, X'_m|=_{Inc}Y'_m)$, where each X_i,Y_i is in L and each X'_i,Y'_i is in $L_{\varnothing}$.

Fix an ordering of the formulae of L. We prove the result by induction on the number of total connectives contained in all the X_i,Y_i.

Given $F(X_1|=_{Inc}Y_1, \ldots, X_n|=_{Inc}Y_n)$ choose the first $X_i|=_{Inc}Y_i$ that mentions a non-atomic formula and choose the first such formula in $X_i|=_{Inc}Y_i$ according to our chosen ordering. We have six possibilities, according as the major operator is N, K, or L, and according as the formula is in the consequent or the antecedent of the entailment. In the first case (an N-formula in antecedent position), we have

$F(X_1|=_{Inc}Y_1, \ldots, X_n|=_{Inc}Y_n)$ iff $F(X_1|=_{Inc}Y_1,\ldots, Np, Z_i|=_{Inc}Y_i ,\ldots, X_n|=_{Inc}Y_n)$,

where $X_i = \{Np\} \cup Z_i$. Applying 4.1.2 we get:

$F(X_1|=_{Inc}Y_1, \ldots, X_n|=_{Inc}Y_n)$ iff $F(X_1|=_{Inc}Y_1,\ldots, Z_i|=_{Inc}Y_i, p,\ldots, X_n|=_{Inc}Y_n)$.

The right-hand side is a Boolean combination of entailments that mention one fewer connective than the left-hand side, so we may apply the induction hypothesis to get:

$F(X_1|=_{Inc}Y_1, \ldots, X_n|=_{Inc}Y_n)$ iff $G(X'_1|=_{Inc}Y'_1, \ldots, X'_m|=_{Inc}Y'_m)$

as desired. The remaining five cases are treated similarly, applying 4.1.1, 4.1.3, 4.1.4., 4.2.1, or 4.2.2 as the case may be.

5 Incompatibilities for Extensions of the Atomic Language

Given an atomic language L and a frame Inc for L, we are interested in characterizing the frames for arbitrary extensions of L. At a minimum we want the frame Inc' for an extension L' to be a conservative extension of Inc:

5.1 (Inferential Conservativeness; IC): Suppose L⊆ L' and that Inc and Inc' are frames for L and L' respectively. Then for all X,Y⊆L, X $|=_{Inc}$ Y iff X $|=_{Inc'}$ Y.

We could independently motivate a construction of Inc' from Inc and show that it satisfies IC. However, it turns out that satisfaction of IC is sufficient to determine Inc', and so we need not suppose anything further about the relation between Inc and Inc'. Further, satisfaction of IC is sufficient to ensure that frames for different language extensions agree on their shared vocabulary. Putting this latter result in more precise form, we have:

5.2 (Frame Harmony): Suppose L⊆ L',L'' and that Inc, Inc', and Inc'' are frames for L, L', L'' respectively. Suppose further that Inc' and Inc'' are conservative extensions of Inc. Then for any X,Y⊆ L',L'' we have X $|=_{Inc'}$ Y iff X $|=_{Inc''}$ Y.

Proof. We can write:

X $|=_{Inc'}$ Y iff $F(X_1|=_{Inc'}Y_1, \ldots, X_n|=_{Inc'}Y_n)$ (Semantic Reduction)
iff $F(X_1|=_{Inc}Y_1, \ldots, X_n|=_{Inc}Y_n)$ (IC)
iff $F(X_1|=_{Inc''}Y_1, \ldots, X_n|=_{Inc''}Y_n)$ (IC)
iff X $|=_{Inc''}$ Y (Semantic Reduction).

Note that in the last step we make use of the fact that $F(X_1|=Y_1, \ldots, X_n|=Y_n)$ is determined only by the syntax of the formulae in X and Y.

It follows immediately that satisfaction of IC is sufficient to uniquely determine the conservative extensions of any Inc for an atomic L.

5.3 (Frame Determination): Suppose that L⊆ L',that Inc is a frame for L, and that Inc' and Inc'' are both frames for L'. Suppose further that Inc' and Inc'' are both conservative extensions of Inc. Then Inc'=Inc''.

Proof. Applying 5.2 with L'=L" we get X $|=_{Inc'} \varnothing$ iff X $|=_{Inc''} \varnothing$. But this is equivalent to X ∈ Inc' iff X ∈ Inc".

Appendix 2: Logic Using the Reduction Formulae

Appendix 1 showed that the semantic definitions of the principal connectives in terms of incompatibility/incoherence underwrites these six reduction schemata:

LN: $X, Np \models Y \Leftrightarrow X \models Y, p.$
RN: $X \models Y, Np \Leftrightarrow X, p \models Y.$

LK: $X, Kpq \models Y \Leftrightarrow X,p,q \models Y.$
RK: $X \models Y, Kpq \Leftrightarrow X \models Y, p$ **and** $X \models Y, q.$

LL: $X, Lp \models Y \Leftrightarrow X \models Y$ **or** $\not\models p.$
RL: $X \models Y, Lp \Leftrightarrow X \models Y$ **or** $\models p.$

Using them, it is easy to show that the incompatibility semantics validates classical logic for N and K (hence for A), and S5 when we add L (and hence M):

1 Negation

1.1: If $\{p\} \in I(X)$ and $\{Np\} \in I(X)$, then $X \in Inc$.

Proof: If $\{Np\} \in I(X)$ then $X \models p$, that is, $I(p) \subseteq I(X)$. Since $\{p\} \in I(X)$, we have $X \in I(p)$. Thus $X \in I(X)$, or $X \in Inc$.

It follows that $Np \in I(p)$:

1.2 (Double Negation Equivalence): $NNp \approx p$.

Proof:
($\Rightarrow$): $NNp \models p \Leftrightarrow \emptyset \models Np,p$ (LN) $\Leftrightarrow p \models p$ (RN).
($\Leftarrow$): $p \models NNp \Leftrightarrow p,Np \models \emptyset$ (RN) $\Leftrightarrow p \models p$ (LN).

1.3 (Contraposition 1): $p \models q \Leftrightarrow Nq \models Np$.

Proof: $Nq \models Np \Leftrightarrow \emptyset \models Np,q$ (LN) $\Leftrightarrow p \models q$ (RN).

1.4 (Material Consistency): $\models X \Rightarrow \sim\exists Y[Y \in \cap_{p \in X} I(p)\ \&\models Y]$.

That is, if X and Y are materially incompatible, they cannot both be valid, or, equivalently, nothing incompatible with anything valid is valid. This result, we should note, depends on our not being in the degenerate frame in which all sets are incoherent. In that case, all sets are also valid.

Proof: Suppose we are not in the degenerate frame. By definition, $|\!=X \Leftrightarrow \forall Z[Z\in\cap_{p\in X} I(p)\rightarrow Z\in(Z)]$. So only self-incompatible Y could be incompatible with valid X. But since *everything* is incompatible with any such self-incompatible Y, it cannot be that |=Y. For instance, X would be a non-self-incompatible counterexample to |=Y. (If a valid X is self-incompatible, it follows that all sets are incoherent, and hence that we are in the degenerate frame.)

<u>1.5 (Formal Consistency)</u>: (X∉I(X) & X|=p) ⇒ X|≠Np.

Proof: By NI and Partition, X∈I(Np) ⇔ X|=p. But then X itself is a counterexample to X|=Np, since X∈I(p) and X∉I(X).

It follows immediately that if |=p, then |≠Np (assuming, again, that we are not in the degenerate frame). So the incompatibility logic of negation is consistent.

Since incoherent sets entail everything, it also follows that (X|=p & X|=Np) ⇒ ∀Y[X|=Y].

2 Conjunction

<u>2.1</u>: Kpq|=p and Kpq|=q.

Proof: By LK, Kpq |= Y ⇔ p,q |= Y. So Kpq|=p ⇔ p,q|=p, and Kpq|=q ⇔ p,q|=q. But p,q|=p and p,q|=q hold by Weakening.

<u>2.2</u>: (X|=p & X|=q) ⇔ X|=Kpq.

Proof: By RK, X |= Y, Kpq ⇔ (X |= Y, p & X |= Y, q). Letting Y=∅, then, X |= Kpq ⇔ (X |=p & X |=q).

3 Negation and Conjunction Together

3.1 KpNp|=Y.

Proof: Immediate from LK and the final observation under 1.5 above.

Anything that satisfies 1.2, 1.3, 2.1, 2.2, and 3.1 and *distributivity* is classical (Boolean) logic.

3.2 (Distributivity): KpAqr ≈ AKpqKpr.

Proof:

a) Axy $\approx_{df.}$ NKNxNy. So KpAqr ≈ AKpqKpr iff KpNKNqNr ≈ NKNKpqNKpr.
b) First direction: KpNKNqNr |= NKNKpqNKpr.
c) KpNKNqNr |= NKNKpqNKpr iff KpNKNqNr, KNKpqNKpr|=∅ (RN).
d) KpNKNqNr, KNKpqNKpr|=∅ iff KpNKNqNr,NKpq,NKpr|=∅ (LK) iff p, NKNqNr,NKpq,NKpr|=∅ (LK).

e) p, NKNqNr,NKpq,NKpr|=∅ iff p,NKpq,NKpr|=KNqNr (LN) iff p,NKpr|=Kpq,KNqNr (LN).
f) p,NKpr|=Kpq,KNqNr iff p,NKpr|=Kpq,Nq and p,NKpr|=Kpq,Nr (RK).
g) p,NKpr|=Kpq,Nq iff p,NKpr|=Nq,p and NKpr|=Nq,q (RK).
h) p, NKpr|=Kpq,Nr iff p,NKpr|=Nr,p and p,NKpr|=Nr,q (RK).
i) So, plugging (h) and (g) into (f): p,NKpr|=Kpq,KNqNr iff
    i. p,NKpr|=Nq,p and N
    ii. Kpr|=Nq,q and
    iii. p,NKpr|=Nr,p and
    iv. p,NKpr|=Nr,q.
j) Now (i-i), (i-iii), and (1-iv) all hold because p|=p; and (i-ii) holds because Kpr|=Nq,q iff Kpr,q|=q (RN), and Kpr,q|=q because q|=q.
k) So p,NKpr|=Kpq,KNqNr (f), and by (c), (d), and (e), KpNKNqNr |= NKNKpqNKpr. QED.
l) Other direction: NKNKpqNKpr|=KpNKNqNr.
m) NKNKpqNKpr|=KpNKNqNr iff
    i. NKNKpqNKpr|=p and
    ii. NKNKpqNKpr|=NKNqNr (RK).
n) Unpack (m-i): NKNKpqNKpr|=p iff |=p, KNKpqNKpr (LN) iff
    i. |=p,NKpq and
    ii. |=p,NKpr.
o) Unpack (n-i): |=p,NKpq iff Kpq|=p (RN). So (n-i) holds.
p) Unpack (n-ii): |=p,NKpr iff Kpr|=p (RN. So (n-ii) holds.
q) So (m-i) holds.
r) Unpack (m-ii): NKNKpqNKpr|=NKNqNr iff |=KNKpqNKpr,NKNqNr (LN).
s) |=KNKpqNKpr,NKNqNr iff KNqNr|=KNKpqNKpr (RN).
t) KNqNr|=KNKpqNKpr iff
    i. KNqNr|=NKpq and
    ii. KNqNr|=NKpr.
u) Unpack (t-i): KNqNr|=NKpq iff KNqNr,Kpq|=∅ (RN).
v) KNqNr,Kpq|=∅ iff Nq,Nr,Kpq|=∅ (LK) iff Nq,Kpq|=r (LN) iff Kpq|=r,q (LN). But Kpq|=r,q because Kpq|=q, since q|=q, by (LK). So (t-i) holds.
w) Unpack (t-ii): KNqNr|=NKpr iff KNqNr,Kpr|=∅ (RN) iff Nq,Nr,Kpr|=∅ (LK) iff Nr,Kpr|=q (LN) iff Kpr|=q,r (LN). But Kpr|=q,r iff p,r|=q,r (LK), and p,r|=q,r because r|=r. So (t-ii) holds.
x) So, by (t): KNqNr|=KNKpqNKpr, so (m-ii) holds.
y) By (x) and (q), (m) holds: NKNKpqNKpr|=KpNKNqNr. QED
z) So by (y) and (k): KpNKNqNr ≈ NKNKpqNKpr. So KpAqr ≈ AKpqKpr, and distributivity holds. QED.

It follows that N, K behave entirely classically.

Besides A, it is useful to define the conditional Cpq$\approx_{df.}$NKpNq. A conditional is valid just in case the corresponding entailment holds:

3.3 p|=q ⇔ |=Cpq.

Proof: By the definition of C, |=Cpq ⇔ NKpNq. By RN, |=NKpNq ⇔ KpNq|=∅. By LK, KpNq|=∅ ⇔ p,Nq|=∅. By LN, p,Nq|=∅ ⇔ p|=q.

Given 3.3, contraposition across '|=' (p|=q ⇔ Nq|=Np), proven in 1.3 has as an immediate consequence contraposition for C:

3.4 (Contraposition 2): |=Cpq ⇔ |=CNqNp.

We get detachment (*modus ponens*) as a derived rule:

3.5 (Detachment): KCpqp|=q.

Proof: By LK, KCpqp|=q ⇔ Cpq,p|=q. By the definition of C, Cpq,p|=q ⇔ NKpNq,q|=q. By LN, NKpNq,q|=q ⇔ q|=KpNq,q. By RK, q|=KpNq,q ⇔ (q|=p,q and q|=Nq,q). Both of these hold, by Weakening.

The following two results may help to impart a better intuitive grasp of the behavior of the connectives. The first vindicates the heuristic reading of "X|=p,q" as "X entails p or q".

3.6. X|=Apq ⇔X|=p,q.

Proof. X|=Apq ⇔X|=NKNpNq (definition) ⇔ X, KNpNq|=∅ (RN) ⇔ X, Np, Nq|=∅ (LK) ⇔ X|=p,q (LN).

The next result generalizes Negation Introduction. Where the latter claims that X is incompatible with {Np} just in case X entails p, we now show how X relates to multiple negated formulas. In essence, X is incompatible with $\{Np_1, \ldots, Np_n\}$ just in case X entails (p_1 or ... or p_n).

3.7. $X \cup \{Np_1, \ldots, Np_n\} \in$ Inc iff $X|= p_1, \ldots, p_n$.

Proof: The claim is equivalent to $X, Np_1, \ldots, Np_n|= \emptyset$ iff $X|= p_1, \ldots, p_n$. This latter claim follows by n applications of LN.

4 Modality

The K axiom is validated by the incompatibility semantics:

4.1 (K): |= CLCpqCLpLq.

Proof:

a) Since we have already shown in 3.3 that p|=q ⇔ |=Cpq, it suffices to show LCpq|=CLpLq.
b) By RL, LCpq|=CLpLq ⇔ (|=CLpLq ∨ |≠Cpq).
c) By 3.3, p|=q ⇔ |=Cpq, |=CLpLq ⇔ Lp|=Lq.

d) By RL, Lp|=Lq ⇔ Lp|=∅ ∨ |=q.
e) By LL, (Lp|=∅ ∨ |=q) ⇔ (∅|=∅ ∨ |≠p ∨ |=q) ⇔ (|≠p ∨ |=q).
f) Since p|=q ⇔ |=Cpq, |≠Cpq ⇔ p|≠q.
g) So LCpq|=CLpLq ⇔ (|≠p ∨ |=q ∨ p|≠q).
h) Suppose not. Then (|=p & |≠q & p|=q). But we showed in 1.1.2 of Appendix 1 that if p|=q and |=p, then |=q. So this is a contradiction. So LCpq|=CLpLq.

Since PC with *modus ponens* (and substitution) is validated, so is the minimal modal system K.

We already showed in 3.4 of Appendix 1 that the rule of necessitation, **|=p ⇒ |=Lp** holds.

From this it is easy to show that the T axiom—and hence the modal system T—is validated:

4.2 (T): |=CLpp.

Proof: We also showed in 3.5 of Appendix 1 that Lp|=p, and in 3.3 that p|=q ⇔ |=Cpq.

4.3 (S4): |=CLpLLp.

Proof:
a) By 3.3, |=CLpLLp ⇔ Lp|=LLp.
b) By RL, Lp|=LLp ⇔ Lp|=∅ V |=Lp.
c) By LL, Lp|=∅ ⇔ ∅|=∅ ∨ |≠p.
d) By RL, ∅|=Lp ⇔ ∅|=∅ ∨ |=p.
e) So, plugging (b) and (c) into (a): Lp|=LLp ⇔ (∅|=∅ ∨ |≠p ∨ ∅|=∅ ∨ |=p).
f) So Lp|=LLp ⇔ ∅|=∅ ∨ (|≠p ∨ |=p). But this second disjunct always holds.

Since the system S4 is just T plus the S4 axiom, the incompatibility semantics validates S4.

4.3 (S5): |=CMpLMp.

Proof:
a) By 3.3, |=CMpLMp ⇔ Mp|=LMp.
b) Mp|=LMp ⇔ NLNp|=LNLNp, since Mp≈NLNp.
c) By LN, NLNp|=LNLNp iff ∅|=LNLNp,LNp.
d) By RL, ∅|=LNLNp,LNp ⇔ ∅|=LNLNp ∨ |=Np.
e) By RN, |=Np ⇔ p|=∅.
f) So, plugging (d) into (c) and (c) into (b): NLNp|=LNLNp ⇔ ∅|=LNLNp ∨ p|=∅.
g) By RL, ∅|=LNLNp ⇔ ∅|=∅ ∨ |=NLNp.
h) By RN, |=NLNp ⇔ LNp|=∅.
i) By LL, LNp|=∅ ⇔ ∅|=∅ ∨ |≠Np.
j) By RN, |=Np ⇔ p|=∅, so |≠Np ⇔ p|≠∅.
k) Plugging (i) into (h) and (h) into (g): |=NLNp ⇔ ∅|=∅ ∨ p|≠∅.
l) Plugging (g) into (f) and (f) into (e): NLNp|=LNLNp ⇔ ∅|=∅ ∨ ∅|=∅ ∨ p|≠∅ ∨ p|=∅.

m) But $p|\neq\varnothing \vee p|=\varnothing$ always holds, so $NLNp|=LNLNp$ always holds.

So the incompatibility semantics validates S5.

Appendix 3: Representation of Consequence Relations by Incompatibility Relations

I. Imputing Incompatibility Relations from Consequence Relations[22]

1. Preliminary Remarks

We assume that we have a consequence relation |— whose consequent positions is either empty or filled by a single sentence. That is, |— is a relation between sets of sentences and single sentences (or the empty set) for some language L. (We consider other types of consequents relations below; see section 6.) This may be a *material* consequence relation, if the sentences do not have any internal logical complexity (or if we are ignoring what they do have), or it may be a *logical* consequence relation, perhaps defined axiomatically, or by a natural deduction system, or by a sequent calculus.

The Representation Theorem for turnstile |— has two conditions.

General Transitivity: $\forall X,Y\subseteq L\ \forall p,q\in L[(X|\!\!-\!\!-p\ \&\ \{p\}\cup Y|\!\!-\!\!-q)\rightarrow X\cup Y|\!\!-\!\!-q]$.
Defeasibility: $\forall X\subseteq L\forall p\in L\ [\sim(X|\!\!-\!\!-p)\rightarrow \exists Y\subseteq L[\forall q\in L\ [\{p\}\cup Y|\!\!-\!\!-q]\&\exists q\in L\ [\sim(X\cup Y|\!\!-\!\!-q)]\ ]]$.

Note that General Transitivity has Pure Transitivity as a special case, where Pure Transitivity is:

Pure Transitivity: $\forall X\subseteq L\forall p,q\in L\ [(X|\!\!-\!\!-p\ \&\ \{p\}\ |\!\!-\!\!-q)\rightarrow X|\!\!-\!\!-q]$.

We simply take $Y=\varnothing$.

2. Representation Definitions

[22] The original representation theorem was proved by the author. But it has been substantially sharpened, and the proof improved, by Alp Aker. Besides the Defeasibility condition required for completeness, the first proof appealed to four conditions that were sufficient for soundness:

i. Reflexivity: $\forall X\subseteq L[\ X|\!\!-\!\!-X]$.
ii. Transitivity: $\forall X,Y,Z\subseteq L[(X|\!\!-\!\!-Y\ \&\ Y|\!\!-\!\!-Z)\rightarrow X|\!\!-\!\!-Z]$.
iii. Monotonicity: $\forall X,Y,Z\subseteq L[(X|\!\!-\!\!-Y\ \&\ X\subseteq Z)\rightarrow Z|\!\!-\!\!-Y]$.
iv. Amalgamation: $\forall X,Y,Z\subseteq L[(X|\!\!-\!\!-Y\ \&\ X|\!\!-\!\!-Z)\rightarrow X|\!\!-\!\!-Y\cup Z]$.

Aker showed that, although these conditions are indeed sufficient for the imputed incompatibility relation to generate a semantic consequence relation $|=_I$ that would hold whenever the original consequence relation |— did, they were not in fact necessary for that result. He showed further that General Transitivity is both necessary and sufficient. The "Converse Results" presented below are also due to Aker.

i. Inch(X) iff $\forall p \in L\ [X|\!\!\sim p]$.

The basic idea is to read off an incoherence relation from the consequence relation by taking the incoherent sets to be the ones that have *everything* as their consequence. If we start with a *logical* consequence relation, generated by a logic that has *ex falso quodlibet* as a basic or derived rule, this will just be the inconsistent sets: the ones that have as a consequence some sentence and its negation.

We then define incompatibility from incoherence in the usual way:

ii. $X \in I(Y)$ iff Inch($X \cup Y$).

And also define the incompatibility-consequence relation as usual:

iii. $X|=_I p$ iff $\forall Z \subseteq L[(Z \in I(p)) \rightarrow (Z \in I(X))]$.

3. Soundness and Completeness

Representation Theorem:

|— is sound and complete with respect to $|=_I$ if, and only if, |— satisfies General Transitivity and Defeasibility.

We first give the proof from right to left. That is, we show soundness and completeness assuming General Transitivity and Defeasibility.

3.1 (Soundness). If X|—p then X|=p.

Proof. Suppose not. Then X|—p but not X|=p for some X and p. Then by definition of |= there is some $Z \in I(p)$ while $Z \notin I(X)$. Unpacking definitions we have $\forall q \in L\ [\{p\} \cup Z|\!\!-q]$ and $\forall r \in L[\sim(X \cup Z|\!\!-r)]$]. Choose some such witnessing r so that $\sim(X \cup Z|\!\!-r)$. Instantiating $\forall q \in L[\{p\} \cup Z|\!\!-q]$ we also know $\{p\} \cup Z|\!\!-r$. Since X|—p, it follows from General Transitivity that $X \cup Z|\!\!-r$, which is a contradiction.

3.2 (Completeness). If X|=p then X|—p.

Proof. Suppose not. Since ~(X|—p) we know by Defeasibility that there are V and r such that $\forall q \in L[\{p\} \cup V|\!\!-q]$ and $\sim(X \cup V|\!\!-r)$. Since X|=p we have $\forall Z \subseteq L[(Z \in I(p)) \rightarrow (Z \in I(X))]$. Unpacking the definition further we have $\forall Z \subseteq L[\forall q \in L\ (\{p\} \cup Z|\!\!-q) \rightarrow \forall q \in L(X \cup Z|\!\!-q)]$. Instantiating with Z=V we have $\forall q \in L\ (\{p\} \cup V|\!\!-q) \rightarrow \forall q \in L\ (X \cup V|\!\!-q)$. By modus ponens we have $\forall q \in L\ (X \cup V|\!\!-q)$. Instantiating with q=r we have $X \cup V|\!\!-r$, which contradicts $\sim(X \cup V|\!\!-r)$.

4. Converse Results

We now show that |— satisfies General Transitivity and Defeasibility, assuming |— is sound and complete with respect to |=.

4.1 (General Transitivity). $\forall X,Y \subseteq L\ \forall p,q \in L[(X|\!\!-\!\!-p\ \&\ \{p\} \cup Y|\!\!-\!\!-q) \rightarrow X \cup Y|\!\!-\!\!-q]$.

Proof. Suppose X|—p and {p}∪Y|—q. By soundness X|=p and {p}∪Y|=q. We show X∪Y|=q. To show this, we need to show that V ∈I(q) implies V ∈I(X∪Y). So suppose V ∈I(q). Since {p}∪Y|=q, this implies that V ∈I({p}∪Y), which implies Inch(V∪{p}∪Y), which in turn implies V∪Y ∈I(p). Since X|=p, we then know that V∪Y ∈I(X). This is equivalent to Inch(V∪X∪Y), which is in turn equivalent to V ∈I(X∪Y). Hence X∪Y|=q. By completeness, X∪Y|—q.

4.2 (Defeasibility). $\forall X \subseteq L \forall p \in L\ [\sim(X|\!\!-\!\!-p) \rightarrow \exists Y \subseteq L[\forall q \in L\ [\{p\} \cup Y|\!\!-\!\!-q] \& \exists q \in L\ [\sim(X \cup Y|\!\!-\!\!-q)]\ ]]$.

Proof. Suppose ~(X|—p). By completeness, ~(X|=p). Unpacking the definition of |=, we have ∃Y[Y∈I(p) & ~(Y∈I(X))]. Unpacking the definitions of Y∈I(p) and ~(Y∈I(X)), we have ∀q∈L [{p}∪Y|—q and ∃q∈L [~(X∪Y|—q)].

5. Discussion

We have shown that General Transitivity and Defeasibility are jointly equivalent to soundness and completeness. As noted, this initially looks like an ideal result. But the reader might have noticed that our proofs allow for a more precise characterization of the logical relation between these four properties. Put briefly, the situation is this:

> Completeness if, and only if, Defeasibility.

That is, we appealed only to Defeasibility in the proof of completeness, and *vice versa.* One might expect, then, that we would have:

> Soundness if, and only if, General Transitivity.

But a look at the proofs reveals instead that we have:

> General Transitivity implies soundness.
> Soundness and completeness imply General Transitivity.

We haven't been able to eliminate an appeal to completeness in the proof of General Transitivity.

6. Generalizations

Having identified the incoherent sets and the semantic entailments, we could proceed to reason logically in the language L. The rules of incompatibility logic are not directly applicable because those rules in general require that the consequents of |= can be sets of formulae, not just single sentences. But having identified the incoherent sets, the entailment relation we have been using in previous sections is perfectly well-defined. Of course, when we allow claims such as X|=Y we will not have a corresponding consequence X|—Y because of expressive limits on |—.

In cases where |— is more expressive, in the sense of allowing multiple formulae in consequent position, there is still a representation result, but the conditions of course must be slightly different. If X|—{y_1, ... , y_n} has the meaning "X implies y_1 and ... and y_1" then the conditions for representation are:

<u>General Transitivity</u>: $\forall X,Y,W,Z\subseteq L$ [(X|—Y & Y∪W |—Z)→X∪W|—Z].
<u>Defeasibility</u>: $\forall X,Y \subseteq L$[~(X|—Y) →
∃Z⊆L[[∀W⊆L[Y∪Z|—W]&∃W⊆L[~(X∪Z|—W)]].

And we must, naturally, also adjust our representation definitions:

i. Inch(X) iff ∀U⊆L[X|—U].
ii. X∈I(Y) iff Inch(X∪Y).
iii. X$|=_I$Y iff ∀Z⊆L[(Z∈I(Y))→(Z∈I(X))].

With these changes the representation theorem again holds. Indeed, the proofs require only trivial changes.

If |— is instead a disjunctive-consequence turnstile (that is, X|—{y_1, ... , y_n} has the meaning "X implies y_1 or ... or y_1"), then the conditions are naturally different. We can retain Defeasibility as in the previous case, but our transitivity condition becomes:

<u>General Transitivity</u>: ∀X,W,Z⊆L∀y_1, ... , y_n ∈L [(X|— y_1, ... , y_n & {y_1}∪W |—Z & ... & {y_n}∪W |—Z)→X∪W|—Z].

The representation definitions are also as in the previous case, except of course for the definition of entailment:

iii. X$|=_I$Y iff ∀Z⊆L[(Z∈$\cap_{p\in Y}$I(p))→(Z∈I(X))].

The proofs again require only obvious changes.

II. Discussion of some Logical Consequence Relations

All the logics we consider satisfy General Transitivity. For, as pointed out in the text, that condition is equivalent to the Cut rule:

$$\frac{\Gamma\text{: A and }\Delta\text{,A: B}}{\Gamma,\Delta\text{:B.}}$$

This will hold as a derived rule in any system that permits the argument:

$$\frac{\Delta\text{,A: B}}{\Delta\text{: A}\rightarrow\text{B}}$$

by '→' Intro, and then:

$$\frac{\Gamma: A \text{ and } \Delta: A \rightarrow B}{\Gamma,\Delta: B}$$

by '→' Elimination.

And all except relevance logics include *ex falso quodlibet*, equivalent to the rule:

$$\frac{\Gamma: A \text{ and } \Gamma: \sim A}{\Gamma: B}$$

which is required to identify incoherent sets on the basis of their role in the consequence relation, so as to impute the incompatibility relation which in turn determines the incompatibility-entailments.
So the significant condition to consider with respect to various logical consequence relations is Defeasibility.

1. Classical Logic

It is easy to show that classical logic *does* satisfy Defeasibility. For in this context, defeasibility just comes to the condition that if it is not the case that $p \vdash q$, then there is something that is *inconsistent* with with q and not with p. If q is not a consequence of p, there must be some interpretation on which p is true and q is not true. Pick one such. Now it might, or it might not, be the case that p and q are incompatible or inconsistent (that is, that $\forall U[\{p\} \cup \{q\} \vdash U]$). If they are *not* incompatible, then ~q is incompatible with q (that is $\forall U[\{q\} \cup \{\sim q\} \vdash U]$) and not with p. If p and q *are* incompatible, then p itself is something that is incompatible with q and not with p—unless p were itself incoherent (=inconsistent: $\forall U[p \vdash U]$), in which case $p \vdash q$, contrary to our hypothesis. So if q is not a logical consequence of p, then there is something that is incompatible with q and not with p, which is the defeasibility condition.

2. Modal Logics

Most familiar modal logics, including T (sometimes called 'M'), K, B, S4 and S5 (indeed, all the Lewis systems), and many less familiar ones (such as Boolos's GL modal logic of provability) are *normal*, in that they contain all the theorems of the classical propositional calculus PC. Defeasibility and the arguments and constructions concerning it depend only on the effects of classical negation on the logical consequence relation, so they go through straightforwardly for all normal modal logics.

3. Intuitionism

The consequence relation of intuitionistic logic does not satisfy defeasibility. It is the case that whenever an intuitionistic consequence is a good one, everything incompatible (here, inconsistent) with the consequent is incompatible with the antecedent. (That much follows from the soundness result, which depends only on Cut.) But it is not the case that wherever an intuitionistic consequence *fails* there is something that is inconsistent with the consequent but not the antecedent. For instance, it is characteristic of intuitionism that although $\neg\neg p$ does follow from p, p is not a consequence of $\neg\neg p$. The Defeasibility condition requires that there be a 'witness' of the badness of this inference, in the form of something incompatible with p, but not with $\neg\neg p$. In this setting, what is incompatible with p is what is inconsistent with it, and that is whatever entails ¬p. But everything that entails ¬p is inconsistent *both* with p *and* with ¬¬p. So there can be no such witness. So Defeasibility fails for the consequence relation of intuitionistic logic. Indeed, the cases where it fails, the non-consequences that fail to have the witnesses incompatibility-defeasibility demands, are just those classical inferences that do not hold good in the intuitionist setting. So intuitionism can be characterized precisely by the cases in which incompatibility-defeasibility fails.

Since the second condition of the representation theorem proved above does not hold for intuitionism, the intuitionistic logical consequence relation is not fully captured by the incompatibility-consequence relation implicit in it. Does that mean that the intuitionistic propositional calculus (and its modal extensions such as intuitionistic S4) does *not* have PC+S5 as its consequence-intrinsic logic? That conclusion would be hasty. For the intuitionistic notion of negation defines a notion of inconsistency that when made to play the role of incompatibility generates a standard incompatibility-consequence relation: that is, one whose proper elaborated-explicating (LX) implicit logical vocabulary is PC+S5. It follows that the techniques introduced here show that alongside the logical consequence relation explicitly defined by the axioms, natural deduction rules, or sequents of intuitionistic propositional calculus, there is *another* logical consequence relation implicitly put in play by the relation of intuitionistic inconsistency defined by intuitionistic negation. Defeasibility *does* hold for that one, and it permits the introduction of the classical connectives plus the S5 modal connectives, by the means outlined in these Appendices. In this somewhat extended sense, PC+S5 *is* the intrinsic logic of intuitionism (and its modal extensions such as intuitionistic S4) too.

To these considerations we may add another, which may be instructive in comparative assessments of the expressive power of intuitionistic versus classical logical connectives (the issue that supersedes concern over which is the *true* or *correct* logic, on the expressive view of the demarcation of logical vocabulary pursued here). If we look at small finite numbers of propositions—say *n* atomic propositions, along with some, but not all of their negations, and some, but not all of the conditionals relating them—it will often happen that for some incompatibility interpretations (even those that respect the meanings of the connectives to the extent possible, for instance by ensuring that any set containing ~p is incompatible with any containing p), some inferences we take to be bad ones are endorsed, because everything incompatible with the consequent is incompatible with the antecedent. Intuitively, this is because there just are not *enough* propositions—not enough, that is, to provide witnesses, incompatibility-defeasors, for all the bad inferences. Throwing in some more propositions, for instance, adding more negations, more conditionals, negations of conditionals, and so on, provides the desired defeasors. As *n* gets larger, and as we more completely form the logical compounds of those atomic propositions, the incompatibility-consequence relation converges on the intuitively—and logically—correct one. One might think of the situation with the two consequence relations generated by intuitionistic logic—the one it defines directly and the one generated by its notion of inconsistency—along these lines. The intuitionistic consequence relation tells us that some consequences are bad, that they do not hold: paradigmatically, the inference from ¬¬p to p. But while for *most* of the consequences that fail in the intuitionistic setting (for instance, that from p∨q to p&q) it is possible to give *reasons* justifying the claim that the inference is a bad one, in the form of inconsistency-defeasors, sets of claims that are inconsistent with the conclusion but not with the premises, for *some* (indeed, for just those whose failure distinguishes intuitionistic from classical logic), it is *not* possible to formulate such defeasors, to give reasons of that kind. From the incompatibility point of view—and keeping in mind the way failures to yield incompatibility-defeasors for intuitively bad inferences can be seen to be due to the expressive impoverishment of systems with "too few" propositions—the failure of Defeasibility for what we may call the 'official' consequence relation of intuitionistic logic amounts to an admission of expressive impoverishment. The intuitionistic logical vocabulary does not have the expressive power to formulate defeasors that could serve as witnesses, as reasons for denying the goodness of inferences the logic nonetheless insists are bad. Such reasons can be given for some of the inferences it rejects (those that are rejected also by classical logic), but not for the rest.

Again from this point of view, intuitionistic logic shows itself to be incomplete. To defeasor-complete a system containing intuitionistic negation, one would want to add another kind of negation, so contrived that it would supply defeating witnesses inconsistent with the conclusions but not the premises of the inference-forms intuitionism characteristically rejects: paradigmatically, a kind of negation of p (which could be neither intuitionistic nor classical negation) inconsistent with p but not with ¬¬p). Intuitionistic negation provides defeasors only for inferences rejected by *classical* logic. What stands to intuitionistic negation in this respect as it stands to classical negation (which of course is already in equilibrium in the sense of being defeasor-complete)?

Notice that nothing in this discussion of relations between the consequence relations of intuitionistic and classical logic requires the appeal to notions of <u>truth,</u> or <u>truth-value</u>, or <u>bivalence</u>. The difference in the contribution of the two different sorts of negation to the consequence relation is adequately characterized entirely in terms of the notion of logical incompatibility, in the form of inconsistency, that they codify. From the point of view of the pragmatic expressive approach to the demarcation of logical vocabulary pursued here, understanding those negations is a matter of understanding which aspects of *material* incompatibility they make explicit.