Scuola Normale Superiore di Pisa

Tesi di perfezionamento in Filosofia

Explicit Normativity

The Logic of Brandom's Scorekeeping

Relatore:
Prof. Massimo Mugnai

Candidato:
Giacomo Turbanti

Anno Accademico 2010–2011

Explicit Normativity
The Logic of Brandom's Scorekeeping

Giacomo Turbanti

Contents

Chapter 1

Normative semantics for normative meaning

1.1 Concepts without norms?

1.1.1 Is meaning normative?

Is meaning normative? This question has been haunting philosophy of language since the wittgensteinian reflection on rule-following, but it is nourished by roots which dig deeper into epistemology and modality, and whose semantic consequences are somehow only their superficial products. Nonetheless the formal apparatus standardly deployed to account for these consequences often falls short from providing a perspicuous representation of what is at play. In this sense the suggestion is that normative facts is what it should be looked for in order to show meaning to be normative. As a result, the whole question has been hinged on the debate about how to extract an account of meaning from a naturalistic account of speakers' behavior. The frustration of this enterprise usually produces two main outcomes as a stark choice. The first option is to accept normativity of meaning but renounce to use meanings to explain many of the things about human rationality we intended it to explain: meanings are normative but we don't grasp them good enough. The second option is to keep meanings with their explanatory power but try to defuse all the issues about underdetermination gushing out from the concept of normativity. I reject both these pidgeon holes, and in what follows I'll try to slip under them a sellarsian lever to unhinge the whole framework.

1.1.1.1 Rule following Wittgenstein

Kripke [80], notoriously, interpreted Wittgenstein's *Philosophical Investigation* as an argument against naturalistic theories of meaning that reaches a skeptical, though acceptable, conclusion. The story is well known. The initial question is: how is the meaning of a symbol determined? The answer is that there is no collection of facts that can fix such a relation, as it is proved by the arguments on rule following and on private language: meanings are underdetermined with respect to facts about speaker's linguistic practice because every piece of linguistic behaviour can be construed as conforming with the application of infinitely different contents; meanings are underdetermined with respect to facts about speaker's inner ('mental') states because these don't make any friction on the public dimension required for linguistic communication. Kripke is mainly concerned with dispositional theories which intend to cope with these difficulties by providing a way to tune future linguistic behaviour on present, finite coordinations between stimuli and speakers' inner states. Dispositions to apply words are shown to be insufficient to determine contents for two main reasons: (i) our dispositions are *finite* as contrasted with the infinite number of applications required to fix contents, and (ii) they can't determine the *correct* use of a word because we could have dispositions to make errors. Obviously these two objections are strongly related. In fact, Kripke notices, one could think that objection (i) can be met by an account of the process of learning dispositions that develops them as including *ceteris paribus* clauses: roughly, perceptive mechanisms that allow people to distinguish regular patterns in the experience of the world and to tune behavior upon them have already been forged (by nature) to produce responses properly differentiated according to *relevant* patterns – where "relevant" is to be related to natural purposes people have as human beings, like surviving and reproducing, for instance. But this answer gets patently exposed to objection (ii) as soon as the cartesian question about justification is raised. No matter how much nature can be trusted, what does prevent people from making mistakes in their processes of forming beliefs? What does justify knowledge?

Now, it's crucial to notice that objection (ii) can only work under the assumption of the normative character of meaning. Kripke clearly states that the main flaw of dispositionalist theories is to miss this point:

> "the relation of meaning and intention to future action is *normative*, not *descriptive*"[1]

[1] *Ibid.*, p. 37.

Kripke's well known answer is that the responsibility to evaluate such a normative character can be delegated to the community, so that any request for justification of knowledge can be watered down into the description of the community behavior. The flaws of this proposal are well known too.

What I want to stress about this old story is not to be found in the details, but in the whole the picture it presents. I'm afraid that Kripke's way to frame things (later established as a standard benchmark) echoed some misunderstandings about the nature of normativity of meaning and engendered the idea that the choice between a naturalistic and a normative account depends after all on a prior intuition to be justified against criticisms from the opposite side. Thus, while intuitions about the normative character of meaning go either ways, the debate follows by now well trodden paths.

Thus, I would rather pull for a slight but insightful and consequence-laden revision of this frame, which I borrow from John McDowell's analysis of Kripke's story. McDowell[2] pictured the bundle of problems Wittgenstein was dealing with as a dilemma: on the one horn the familiar correspondentistic representation of truth as congruence between meanings and facts that Wittgenstein refutes, and, on the other horn the whole famous paradox of *Philosophical Investigation*:

> "no course of action can be determined by a rule, because every course of action can be made out to accord with rule"[3].

I think this picture is still very useful to take a stock of the different structures of the arguments directed against the normativity of meaning. On the one side there are those who, in order to preserve a 'realist'[4] stance towards meaning, have to move conceptual contents in platonistic reserves where, later, it's hard to explain how speakers can get in touch with them. On the other side there are those who assume an 'anti-realist' stance towards meanings themselves and thus are forced to choose between reconstruing a weakened sense of objective content, or abandoning the idea that meanings have any substantial role in the explanation of linguistic communication.

The whole wittgensteinian reflection on rule-following and private language is purported to attack the former realist thesis: the idea that conceptual content can stand by itself

[2] *Ibid.*, p. 331.

[3] Wittgenstein [174], §201.

[4] Here I stick to McDowell's usage of "realism" and "anti-realism": according to realism meanings are 'out there' autonomous from our grasping them, while according to 'anti-realism' they are only the superficial effetcs of deeper (social, psychological, etc.) structures.

oustide the communicative practice. The latter line of argument, instead, in this paradoxical form, apparently can rest only in a rortian position where the *very* skeptical doubts that generate the paradox are discharged and successful communication is accounted for in communitarian terms, or in a quinean position where eventually intensional contents themselves are discharged, in favour of a fully naturalistic analysis. For those who acknowledged the results of Wittgenstein's criticism the classical semantic picture is irreparably compromised. Among these, there are some who are content with a communitarian solution, within an overall skeptical framework about an objective definition of meaning. For those who, instead, have too many philosophical scruples and complain that the community's perspective is subject to corresponding criticism, this latter naturalistic way out seem to be the best option left. But everyone, so to say, is free to choose the position she prefers before the game begins, and keep it quite harmlessly till it ends. The only unteanable position would be to try to eat the cake and have it, as sometimes philosophers and epistemologists try to do by theorizing about how people would build normative meanings up from experience.

But rarely things are so easy with Wittgenstein, and the very idea that *Philosophical Investigations* would entitle one to rest peacefully in any position shoule make him suspicious that it is but a bad misunderstanding. This is where McDowell's reframing of Wittgenstein's argument can be appreciated in all its importance. In fact, by presenting the rule-following paradox just as the second horn of a more complex dilemma, it make sense of Wittgenstein's argument not as resting in the formulation of a paradoxical conclusion but as a path that moves from the rejection of the augustinian correspondentisic interpretation of meanings (the first horn of the dilemma) through the analysis of the difficulties of this second horn and eventually to their solution. This sheds light on the fact that Wittgenstein himself suggests the way out of the paradox and of the whole dilemma: just in §201 of *Philosophical Investigations* he maintains that the paradox is generated by the misunderstanding that there could be "a way to grasp a rule which is *not* an *interpretation*". And again the same topic is even more clearly stated in the *Remarks on the foundations of Mathematics*[5] (VI-28), where he writes

> "Following according to the rule is FUNDAMENTAL to our language game. It characterizes what we call description".

Thus, Wittgenstein is looking for an actual way out of his dilemma, not just a skeptical solution of it. Such a way out certainly goes right through the acknowledgement of the

[5] Wittgenstein [176].

essentially social dimension of meaning, but it construes such a social dimension not as a renounce to objectivity, but as the horizon in which meaning is possible. As it's well known, Wittgenstein has an evocative but not quite helpful name for such an horizon:

> "What has to be accepted, the given, is – so one could say – *forms of life*."[6]

Now, although this matches quite neatly with McDowell's general idea of conceptual content and its 'space', I think it still begs the question about how to account for a 'logic of meanings', which is what is needed for semantical tasks – from the broader to the more formal ones.

1.1.1.2 Meaning as describing extensions

Here I want to point out one more thing McDowell remarks in his analysis that I find quite crucial.

The same formal apparatus developed to represent meaning within the 'realist' framework doesn't have to be rejected together with such a framework[7]. Consider T-sentences, for example

> "emeralds are green" is true iff emeralds are green.

What is rejected together with the realist picture is the idea that the right hand side of the biconditional above depicts a fact out there, given to us in some nonconceptual way. This is not an original remark: Davidson's reflections on this point are well known. He pointed out that T-sentences can be used to deploy an extensional definition of truth in a language only if the notion of meaning is taken for granted – more precisely, but equivalently, only if rules are available to translate expressions of objective language into expressions of the metalanguage. He thus proposed to reverse the explanatory direction of the argument and to use the tarskian biconditionals to define meaning from a given notion of truth. Notice however that, while he construed T-sentences as a theory of meaning, he also pointed out that they just constitute the form in which such a theory can be expressed, but they don't say anything new about meanings themselves, because they don't explain how the extension of the T-predicate has to be defined[8].

But, I think, if Wittgenstein is right, to ask "what is truth *then*?" is the to raise the wrong question.

---

[6] *Ibid.*, p. 226.
[7] See McDowell [96], p. 352
[8] See Davidson [38].

In this sense I also think that the lack of any formal tool alternative to truth-functional semantics encouraged to move around this point and negatively influenced the possibility to recognize and develop Wittgenstein's suggestion. I want to argue that truthfunctional extensional semantics contributes to blur the dependence of descriptive vocabulary from modal vocabulary in the analysis of meaning. The result is that normativity of meaning remains hidden.

This point stands out very clearly in Boghossian:

> "What, however, is the intuitive normative truth that falls directly out of the attribution of meaning [...] ?
>
> One thought that might seem to be in the right neighborhood is this:
>
> > If I mean addition by '+' then, although I may not be disposed to say '125', in response to the question 68+57=?, it is *correct* for me to say '125'.
>
> The trouble is that it is not clear that, at least as it is being used here, "correct" expresses a normative notion, for it may just mean "true". [...] But there is no obvious sense in which truth is a normative notion."[9]

If my semantic *explanans* is the intuitive correspondentist notion of truth as *adequatio rei et intellectus*, why should I also assume that the *intellectus* "ought" to adjust to *res*? Suppose I believe with Davidson that meanings are not given to us but they have to be evinced from our semantic practices of treating some statements as true. Why should I put here a 'no access beyond this point' limit? Why couldn't I go ahead and explain these practices in terms of my naturalistic theories about human psychology or sociology? What normativists do, in this sense, would be to take the uncertainty of scientific explanation for a conceptual gap between actual meaningful uses and correct meaningful uses of language expressions.

Let me pick Anandi Hattiangadi's arguments against normativity of meaning[10] as an example of this sort of reasoning. She puts forward as obvious the following meaning platitude (MP) as a relation between the meaning of a term and the conditions of its application[11]:

[9] Boghossian [11], p. 207.
[10] See Hattiangadi [69, 70].
[11] Hattiangadi [70], p. 222.

Meaning Platitude: t means $F \rightarrow (x)(t$ applies correctly to $x \leftrightarrow x$ is $f)$

where t is a term, F is its meaning, "t applies correctly to x" can be substituted for any representational semantic relation we want to consider (for example, "t refers to x", "t is true of x", etc.) and f is the collection of features in virtue of which t applies[12]. From (MP) Hattiangadi trivially derives

Correctness: S means F by $t \rightarrow (x)(S$ applies correctly t to $x \leftrightarrow x$ is $f)$

Now, Hattiangadi finds it puzzling that *Correctness* can be used to support normativity of meaning even if it doesn't *prescribe* anything to speakers. This is less trivial: as we've just seen, the assumption implicit here is that correctness of could be vindicated in naturalistic terms. Thus normativity of meaning can't be presented as the unavoidable consequence of this premise. What really lies behind *Correctness*, according to Hattiangadi, is something different, that she proposes to represent as:

Prescriptivity: S means F by $t \rightarrow (x)(S$ ought (to apply t to $x \leftrightarrow x$ is $f))$

This would really represent an explanatory gap in the resources of naturalism.

But then she proves (after Boghossian[13]) that *Prescriptivity* is false, because the right-left direction of the biconditional is false: even if one could accept to infer that something is f from the fact that speakers ought to apply t to it – given (MP), in fact, something is f if and only if speakers apply t correctly to it –, it is simplt not true that one ought to apply t to everything that is f – what if one wants to lie, for instance?

Hattiangadi then shows that there seems to be no obvious way to fix *Prescriptivity* and also provides several examples in which it wouldn't fit. Eventually she concludes meaning is not normative.

I think this conclusion follows from a misunderstanding of what normativity of meaning is: there's no reason to be puzzled by the absence of explicit "oughts" in *Correctness*. Let me focus on the formal conclusion of this argument. She's right: the right-left direction

[12]Here Hattiangadi has in mind the meaning of subpropositional expressions, i.e. something like

"Socrates" means Socrates $\rightarrow (x)$("Socrates" applies correctly to $x \leftrightarrow x$ is Socrates),

but I think (MP) can be applied also to propositional expressions, i.e.

"Snow is white" means Snow is white $\rightarrow$ (x)("Snow is white" is true in x $\leftrightarrow$ Snow is white in x)

where Tarski's T-sentences can be recognized.

[13]See Boghossian [11], p. 210.

of the biconditional fails in *Prescriptivity*. But she's wrong about the reason: she thinks that in *Prescriptivity*, as opposed to *Correctness*, the biconditional is threatened by what speakers want or are able to do[14]. There's another deeper reason for which it fails: it's not modally robust. Indeed, it's easy to see that it may fail in (MP) too.

Kripke would have criticised it not because of the normative character of meaning, but because in this form it can be construed as supporting the Descriptivism he rejected in *Naming and Necessity*[15]: there's no set of properties f that can identify the object x in a necessary and sufficient way according to our communicative practices and thus, crucially, (MP) can't provide any criteria to define meaning after correctness of application of a term. Consider

"Socrates" means Socrates$\rightarrow (x)$("Socrates" applies correctly to $x \leftrightarrow x$ is Plato's teacher).

But what if Plato was instructed by Isocrates?

Thus (MP) should be modified like this:

"Socrates" means Socrates$\rightarrow \Box(x)$("Socrates" applies correctly to $x \leftrightarrow x$ is Socrates).

The solution, in this sense, would be to keep truthfunctional extensional semantics to represent (MP), while exploiting the modal rigidity of some non-descriptive vocabulary to fix x is$f$ in all possible worlds.

Notice that this is a quite different and, I think, deeper issue about (MP), that shakes its status of 'platitude': in the absence of analytic criteria to govern the righthand-side of the biconditional there is no clear sense in which correcteness of application of a term could define its meaning.

Quine's skeptical results about meaning's interpretation hinge on the same point. In his legendary reconstruction of the linguistic work in the field he adopts a Naturalist account – which is supposed to be welcome by Hattiangadi – of meanings in terms of classes of responses to stimulations: indeed his conclusion is that simulus-meanings thus defined

[14] I'm not equally sure she's right in the examples she provides to support the thesis that meaning isn't normative. Among those there's the example of *lying*, which is quite typical in this kind of argument. I don't find it convincing though. Let me quote Boghossian's use of it:

> "Is it really true that, if I mean addition by '+', then, if am asked what the sum of 58 and 67 is, I *should* answer '125'? What if I feel like lying or misleading my audience? Is it still true then that I should answer '125'? If I want to mislead, it looks as though I should not say '125' but rather some other number." (Boghossian [11], p. 207.)

Now, we are told, if I want to say the truth I *should* answer 125. But we are told as well that if I want to lie I *should not* answer 125. I venture to conclude that meaning is normative, no matter what I want to.

[15] Kripke [79].

(the only one we can legitimately grasp) are undertermined with respect to all possible analytical hypotheses. Thus, for instance, in

> "Gavagai" means Rabbit →
> (x)("Gavagai" applies correctly to $x \leftrightarrow x$ belongs to the S-meaning of "Rabbit"),

even if the right-left direction of the biconditional holds, i.e. even if the application of a term is guaranteed to be correct with respect to stimulus meaning, there still no obvious way to define meaning upon this correctness.

This point is so important that in order to avoid it Quine proposed to drop the very notion of analyticity, by 'naturalizing' the double implication on the right, i.e. by considering such a coimplication, in humean terms, as a *contingent* conjunction to be vindicated through an indictively established universal generalization by natural sciences.

It should be clear by now where the problems of (MP) come from.

It's not what it explicitly states, but what it implicitly says. It crucially *conveys* a relation between F and f, namely that

> Fs do f.

It is this relation that deals with meaning, and it has to be modally robust. *That* is what makes (MP) a platitude.

1.2 Norms without modality?

1.2.1 Digging misunderstandings out

Traditional representationalist semantics leans on a certain designational model. Section 1.1.1 lifted it enough to spot the *fundamental* place normativity occupies in the characterization of the very notion of meaning. This is something standard semantics oversighted. Notice however that this is not yet a diagnosis of what goes wrong with standard semantics, it is just the acknowledgement of a symptom. Let's try to put it otherwise. So Wittgenstein invites us to reject the idea that a basic non interpretative grasp on meaning is required:

> "If I have exhausted the justifications I have reached bedrock, and my spade is turned."[16]

[16] Wittgenstein [174], §217.

But, what lies below the bedrock? McDowell suggests it's not so difficult to figure it out, in first approximation, what lies down there:

> "a web of facts about behaviour and 'inner' episodes, describable without using the notion of meaning."[17]

As a matter of fact, to an empiricist-minded philosopher the analysis of this web appears much more promising than the wittgensteinian notion of a 'form of life': what does really prevent us to provide a scientific description of these facts, such that the linguistic behaviour, commonsensically described in terms of meaning-talk, could be reduced to them?

In what follows I present a not surprising common pattern in some paradigmatic ways to raise this issue among empiricistically minded philosophers and a quite more surprisingly common pattern in the ways to answer to it.

1.2.1.1 Descriptive vocabulary and modal vocabulary

There is a more fundamental misunderstanding on which this illusion is rooted (or maybe just another side of the same illusion): the idea that modal vocabulary has to be considered a dispensable metavocabulary which should be better explained away in descriptive terms. This is the core thesis that supported both Hume's rejection of causality and Quine's rejection of modality. Indeed, it is one of the standing points of traditional Empiricism, against which Sellars directed his criticism of 'the Given'[18]. His target is the Humean idea that laws of nature must be reduced to inductively established generalizations: there's nothing in the factual texture of the world that can *justify* the necessity of lawlike generalizations of relations among facts, thus such a necessary character has to be explained as a psichological projection over the set of empirical *data*. Let me recall here the contention about inductive arguments between the humean supporter of a descriptive account and the sellarsian opponent claiming the irreducibility of the modal analysis. In his attempt to combine the lawlikeness and the non analiticty of laws of nature the humean finds a hurdle he can't clear. This result triggers Quine's *reductio ad absurdum* of modality, and engenders the skeptical *impasse* that Goodman [60] tries to overcome by appealing to the notion of "projection" and which Kripke surrenders to.

Wittgenstein himself explicilty grouped together the problem of induction and the problem of the normativity of meaning:

[17] McDowell [96], p. 348.
[18] See Brandom [27].

> "How do I know that in working out the series +2 I must write '20.004, 20.006' and not '20.004, 20.008'? – (The question 'How do I know that this color is 'red'? is similar.)"[19]

The two paradoxes proposed by Goodman and Kripke (as well as Quine's argument for the indeterminacy of translation) obviously hit on the same point. But since we are also dealing here with formal semantics, it's important to spell this point out, in order to rehearse its relation with induction and mathematical reasoning, the main test-bed of formal logic and its standard tarskian semantics.

The barebone structure of the paradoxical argument goes like this: consider some conceptual content and put forward a definition of it which is adequate to its public usage, then, either you have any independent peg to hang an analysis of such an adequacy or you inevitably can produce another conceptual content which satisfies your definition but which is not equivalent (not synonym) to the first one.

Thus, Goodman invites us to consider the usage of color words. It doesn't really matter how we define the content of those words (for example we can associate perceptive experiences of green objects to particular ranges of wavelengths that hit our senses), what is important is that our definition has to be adequate to their usage, in the sense that it allows us to say that, for example, it is correct to examine an emerald and to utter "this is green" (or, we may prefer to say that it is *true*). But now Goodman presents the color word "grue", whose usage is s.t. it is correct to examine something before t and utter "this is green", and to examine it and utter "this is blue" otherwise. It's easy to realize that however we define the content of "green" the same definition will apply to "grue" before t, because their usages correspond before t. And yet "green" and "grue" can't have the same content, if the definition of content has to be adequate to their usage, because if we examine an emerald after t it will be still correct to utter "this is green", but it will be wrong to utter "this is grue".[20]

[19] Wittgenstein [176], I, §3.

[20] I'm aware I've just presented the 'grue paradox' without any reference to the problem of confirmation which is its setting in Goodman's analysis. I actually think that is not its point, and that Goodman is clear about that when he writes that the problem of induction is not the problem of justify induction:

> "An inductive inference, too, is justified by conformity to general rules, and a general rule by conformity to accepted inductive inferences [...] The problem of induction is not a problem of demonstration but a problem of defining the difference between valid and invalid predictions". (Goodman [60], p. 65.)

There are several ways to misunderstand Goodman's paradox, and several *nota bene* are usually added to its presentation, for example that nothing changes color after t.

In the same fashion Kripke invites us to consider the arithmetical sum. We all have performed many computations by applying the rule of addition we learned at some time as young students, and we can produce a definition of the arithmetical sum that matches every application of the rule of addition in every computation we've ever done. We are confident we are able to apply correctly such a rule to any computation we will have to perform in the future, even if we've never encountered it before, so when we face for the first time, for example, '68+57' without further ado we proceed to compute the result: '125'. But here comes the skeptic to ask how can we be so sure that the function we've been called "sum" and applied in the past isn't actually the "quum" function, "$\oplus$", defined as follows:

$$x \oplus y \begin{cases} = x + y, & if\ x, y < 57 \\ = 5, & otherwise \end{cases}$$

This definition of the "quum" function also matches every application of the rule of addition in every computation we've ever done, but the correct computation of our new example would yeld '5' as a result, not '125' as expected according to the "sum" function.

Notice how the outline of the presentation can be switched between the two paradoxes and we can talk about a skeptic who invites us to wonder whether we were not applying the concept "grue" while we thought we were using the concept "green"[21], and talk about the content of "quus" which corresponds to that of "plus" before t and differs from it after t, even if any definition we can produce is undetermined with respect to this distinction. This leads to Quine's thesis of the 'indeterminacy of translation'.

Is traditional empiricism completely harmless against these struggles? Let's see.

If the thesis of the rejection of modality has to be accepted, it's easy to understand why dispositional account is so entrusted. Goodman can be construed as trying to develop the outline of a semantics for dispositional predicates. Kripke himself, while presenting possible responses to his wittgensteinian paradox, chose the dispositional analysis of meaning as the main opponent theory and showed that it had to face inextricable perplexities.

But such an analysis of dispositions and its application to the induction problem is not just an independent optional premiss that can be proved to be wrong and substituted with

Now, Goodman maintains, the point of the question about confirmation is not to justify our concepts, i.e. to tell *if* emeralds are green, but to part valid from invalid usages, i.e. to tell *why* it is correct to say that emeralds are green. The 'grue paradox' shows that this second problem (the real one) is not easy to be solved. This is why, according to Goodman, Hume's answer to the problem of induction was right but incomplete.

[21] See Kripke [80], p. 20.

something better. It's rather a proper consequence of the traditional empiricist attempt to reduce normativity to regularity. Goodman was clearly aware of this close relation:

> "The problem of dispositions looks suspiciously like one of the philosopher's oldest friends and enemies: the problem of induction. Indeed, the two are but different aspects of the general problem of proceeding from a given set of cases to a wider set. The critical questions throughout are the same: when, how, why is such a transition or expansion legitimate?"[22]

But even in the empiricist tradition, the definition of conceptual content must cope with adequacy of such a regularity to the public usage we make of contents. The only way to account for such regularities seems to be that of interpreting extensions over possible particulars, and this is the burden of modality. Dipositional predicates fit this analysis, because their semantics seem to require possible particulars, in the sense, for example, that the reference of "flexible" is the class of everything that *can* bend – now if we add "under appropriate circumstances", we may hopefully be tempted to catch the glimpse of an analysis of counterfactuals. What does "green" mean then? How can we define its content in order to be sure we are applying the concept "green" rather than the concept "grue"? The answer seems to be: "green" *is somehow related* to the class of all possible occurrences of "looking green", which *obviously* differs from the class of all possible occurrences of "looking grue".

1.2.1.2 Epistemology naturalized

But what if one feels too much philosophical discomfort towards this metaphysically loose turn that calls for possible particulars and wants to keep extensions within the limits of the actual ones? Obviously there's no way to verify generalizations against all possible particulars in our everyday linguistic practice. Here Quine registers the defeat of analyticity and, in his idea of naturalized epistemology, invokes psychology to sort out sets of actual particulars relevant for meanings:

> "Insofar as theoretical epistemology gets naturalized into a chapter of theoretical science, so normative epistemology gets naturalized into a chapter of engineering: the technology of anticipating sensory stimulations"[23].

[22] Goodman [60], p. 58.
[23] Quine [128], p. 19.

This is the view whose details Quine had taken the responsibility to make fully explicit in his famous paper on naturalized epistemologies, Quine [127]:

> "The stimulation of his sensory receptors is all the evidence anybody has had to go on, ultimately, in arriving at his picture of the world. Why not just see how this construction really proceeds? Why not settle for psychology?"[24]

I'm not really concerned here with the exegesis of quinean though, so I'll avoid to deal with the question if Quine was ever committed to this view. I'll pick up an hypothetical supporter of every quinean thesis included naturalized epistemologies, let's call him the 'Naturalist', no matter if it reveals to be a strawman: his role here is to make room for this view which, though somehow extreme, represents a real answer to Kripke's skeptical account against dispositions[25]. However, since anticipations along this direction were already contained in *Word and Object*[26], the Naturalist won't have scruples about using them as well. What the Naturalist wants to show is that, to the contrary of the epistemological tradition sprung from Descartes, there is no question about justification for empirical knowledge. His argument notoriously goes like follows[27].

Two main tenets remain stable in the empiricist tradition, namely:

a) evidence for science is sensory evidence;

b) word meanings rest on this evidence.

There's an obvious implicit third tenet to be added here in order to complete the picture, i.e. that while sensory evidence is immediate, word meanings that provide empirical knowledge have to be acquired. Let me paraphrase this with Willfrid Sellars[28]:

A. *X senses red sense content s* entails *x non-inferentially knows that s is red.*
B. The ability to sense sense content is unaquired.
C. The ability to know facts of the form *x is* ϕ is aquired.[29]

Now, by using this inconsistency the skeptic proves that given (a) and (b) empirical knowledge can't be justified. This corresponds for the empiricist to giving up A above, as Sellars notices:

[24] *Ibid.*, p. 75.
[25] See Boghossian [9], p. 532ff.
[26] Quine [126].
[27] This is adapted from Quine [127].
[28] Sellars [147], p. 132.
[29] Howether things are with Quine's endorsement of this view, it's striking how this triplet were explicitly reformulated a decade later Sellars showed how its inconsistency is due to the Myth of the Given.

> "He can abandon A, in which case the sensing of sense contents becomes a noncognitive fact – a noncognitive fact, to be sure which may be a necessary condition, even a *logically* necessary condition, of non-inferential knowledge, but a fact, nevertheless, which cannot *constitute* this knowledge."

But one could still try to eat the cake and have it by suggesting that actually A. is a product of the linguistic misconception that empirical vocabulary of red things is to be considered as authonomous: there's no need to justify empirical knowledge upon sensory datas because every sentence containing empirical predicates can be contextually defined in observational and logical terms.

But here is where the theme of analyticy comes into play as we already saw. Let me paraphrase again in Sellars's words:

> "But what then are we to make of the necessary truth – and it is, of course a necessary truth – that
>
> x is red $=_{Def}$ x would look red to standard observers in standard conditions?"[30]

And here is also where Sellars's argument diverges from the Naturalist's one: let's follow the latter to his conclusion while reserving the right to come back to the former at proper time. The Naturalist puts forward as premises two theses so eminent they barely require justification, since they were maintained respectively by Charles Peirce and Pierre Duhem and had so deep and widespread influence:

> i) truthfunctional semantics: meaning of a statement consists in the difference its truth would make to possible experience;
>
> ii) holism: theoretical sentences have their evidence not as single sentences but only as larger block of theory.

Together (i) and (ii) imply the so called indeterminacy of translation thesis, which prevents the strategy described above to rescue the empiricist standard epistemology:

> "The crucial consideration behind my argument for the ideterminacy of translation was that a statement about the world does not always or usually have a separable fund of empirical consequences that it can call its own.

[30] *Ibid.*, p. 142.

> That consideration served also to account for the impossibility of an epistemological reduction of the sort where every sentence is equated to a sentence in observational and logico-mathematical terms." [31]

All things considered – the Naturalist presses – empirical knowledge, as standard epistemology saw it, doesn't exist, and yet things are not so bad for a theory of knowledge: if one takes stock of what is left in the field he finds observation sentences with which speakers regularly react to different stimulations, for instance

> Every time x presents the stimulus σ to S, S applies "red" to x.

Now one can ask what these regularities depend on. But since, and this is the crucial point, *there's nothing like a red content in the stimulus* σ, the Naturalist concludes that this is now become a question for psychology.

The enterprise of psychology[32] (in this sense, as that of any other natural science) is to produce lawlike generalizations of those regular phenomena, something like,

> $\forall x$(x presents the stimulus σ to S $\rightarrow$ S applies "red" to x).

Let me put forward two passing remarks. First, these generalizations then raise the issue of their verification, but this again is an issue for psychological research rather than for epistemology. Second, the expression "red" stands for a candid utterance like "it's red" and there is no ontological commitment on the variables independent from the tenets of psychology.

Once these points are neatly stated, the Naturalist could even paraphrase those lawlike generalizations of psychology in dispositional talk[33]:

> S is disposed to apply "red" to things that present stimulus σ to him.

Notice again that this would be mere paraphrase, to be vindicated in terms of the bare language of the above psychological generalization as soon as any epistemological issue is raised about the status of dispositions; it represents just another, possibly more suitable, way to describe as a natural phenomenon the psychological human subject that

[31] Quine Quine [127], p. 80.

[32] The nominalism-flavoured story I'm about to tell about how to contextually define linguistic resources to talk about meanings upon first order language is modelled on the typical structure of several sellarsian arguments, especially the one in Sellars [151].

[33] Stepping through dispositional talk is dispensable for the argument to go through, but it presents two useful advantages: it allows smoother paraphrases and, by disclosing the obvious fact that the problem of disposition is but a reformulation of the problem of normativity of meaning, increases the appeal and the range of the argument itself.

> "is accorded a certain experimentally controlled input – certain patterns of irradiation in assorted frequencies, for instance – and in the fullness of time the subject delivers as output a description of the three-dimensional external world and its history."[34]

Here a further step can be introduced. At this point, if one really feels the need, one could even venture to contextually define a meaning-talk upon these psychological generalizations, without disappointing the Naturalist too much. It's easy to show how. In fact by the time psycholgy has provided his lawlike generalizations by using first order variables, cathegorizations of those things that present certain stimula to speakers can be produced as well, according to the different stimula they present. In fact, if

> S is disposed to apply "red" to things that present stimulus σ_i to him.

while

> S is disposed to apply "blue" to things that present stimulus σ_j to him.

Then one can cathegorize unpon different labels things that present stimulus σ_i and things that present stimulus σ_j, say label $\ulcorner Red \urcorner$ and label $\ulcorner Blue \urcorner$; thus for instance

> $\ulcorner Red \urcorner$ sticks to $x =_{Def} x$ presents the stimulus σ to S.

Here one can legitimately use the plural $\ulcorner Red \urcorner$s to refer to things to which label $\ulcorner Red \urcorner$ sticks.

Now, once these labels are available to be attached to what the Naturalist calls different *stimulus-meanings*, one can use them to paraphrase usual meaning-talk as a metalinguistic language like follows:

> S means *Red* by "red" $=_{Def}$ S is disposed to apply "red" to $\ulcorner Red \urcorner$s.

After all, the Naturalist himself seem to acknowlede something like that:

> "Occasion sentences whose stimulus meanings vary none under the influence of collateral information may naturally be called *observation sentences*, and their stimulus meanings may without fear of contradiction be said to do full justice to their meanings. These are the occasion sentences that wear their meanings on their sleeves."[35]

[34] Quine [127], pp. 82-83.
[35] Quine [126], p. 42.

Here the truth of psychological generalizations guarantees the impermeability to collateral informations, thus observational sentences can legitimately receive their meaning-label.

But as a philosopher, it's hard to resist to the temptation to push the Naturalist's patience one last time in order to take one final step. At this point we legitimately dispose of contextually defined meaning-labels: we are entitled to use statements like "x is *Red*", with the proviso that we are making explicit in a code-language the application of the espression "red" to something our psychological laws classify as $\ulcorner Red \urcorner$[36]. The temptation is poked by the question about analyticity we stated above in Sellars's terms of a subjunctive conditional: have we an answer to that question now? Indeed, we may have it, so let's push the Naturalist's patience and ask: what would invalidate the definition

x is *Red* $=_{Def}$ S *would be* disposed to apply "red" to x?

Our answer must be that it would be a case that invalidates the psychological generalization

$\forall x(x$ presents the stimulus σ to S $\rightarrow$ S applies "red" to $x)$

upon which we built such a definition, i.e. a case in which speakers do not react to stimulus σ by applying "red".

Once we have defined label-meanings, we can cathegorize them into labels supported by true psychological generalizations in the sense just stated, say T-labels, and labels which are supported by generalizations which are discovered to be false, say F-labels. This would provide the means for the contextual definition of a full-blown meaning-talk with variables ranging over the meaning labels introduced above. In this way, we would have something like

$\forall \ulcorner \Phi \urcorner, x((T(\ulcorner \Phi \urcorner) \wedge \ulcorner \Phi \urcorner(x)) \rightarrow$ x is $\Phi)$[37],

were $\ulcorner \Phi \urcorner$ is a variable ranging over meaning-labels, T is a predicate that characterize meaning-labels supported by true generalizations, and "x is Φ" is a statement expressing the usage of the meaning-label $\ulcorner \Phi \urcorner$ inside our contextually defined meaning-talk.

[36] I'm oversimplifying here for the argument's sake: things with predication are a little more complex. For a more detalied presentation of the topic along these lines see Danielle Macbeth's analysis of Frege's account of predication in Macbeth [89], especially section 3.

[37] Notice that this is not equivalent to

$\forall \Phi, x((T(x) \wedge \Phi(x) \rightarrow p)$

This latter would amount to provide a reliable way to specify *standard conditions* in counterfactual conditionals.

Notice that this, in a sense, would solve the problem of analyticity that triggered the argument for naturalized epistemology and it would provide an extentional generalization suitable to play the role of the lawlike major premise in a deductive argument of standard epistemological talk about empirical knowledge. But, it may also look like a cheap taunt against the Naturalist's nominalism. So what's the trick? I think there is no trick, so let me briefly scatter these clouds in order to see clearly where the real problem is.

First, here the second order property T is obviously far from implying the confutation of the Naturalist's theses about second order logic and about meanings. Instead, in a sense it represents an expected consequence: meanings are not things in the ontological texture of a scientific description of the world, and if we want to talk about and identify them, we have to move away from the proper scientific vocabulary. The vantage point we reached hinges on a paraphrase: we can discharge the ontological commitment due to the second order predicate T on the contextual definition given above, since we know how to track it back to the lawlike generalization of psychology. Does this represent a confutation of the Naturalist's ontological criteria: "to be is to be the value of a variable"? I don't think so. At least no more than the contextual definitions proposed in *Word and Object* he himself would endorse. I don't want to hide behind my finger though. If these remarks do not require a confutation of the Naturalist ontological criteria, they do not leave them as they stand: rather than ontological criteria what is at play is the interpretation of variables in formal language[38].

Second, to the extent in which the Naturalist's premises about the vocabulary of naturalized epistemologies are correct, one could not raise skeptical doubts about the definition of $\ulcorner\Phi\urcorner s$ as well: they deal with the good – maybe naturally selected – set up of our neurophisiological configurations, and not with meanings, so they can't be jerrymandered[39]. Since $\ulcorner\Phi\urcorner$s *are not* meanings, the argument for indeterminacy of translation doesn't apply to them in the sense that they can't be underdetermined with respect to things that produce stimulus σ since they are simply a code to paraphrase regularities expressed by psychological generalizations. For instance, the meaning-label $\ulcorner Gavagai\urcorner$ can't be undetermined with respect to certain things that produce stimulus σ given that speakers would apply "gavagai" as a reaction to possibly different kinds of stimulations, because $\ulcorner Gavagai\urcorner$ simply sticks to things that produce stimulus σ in that speakers react to them by applying "gavagai", as stated by a psychological generalization such as

[38]Again, for further details, I refer here to Macbeth [89].

[39]For a sellarsian discussion of this point see Millikan [105].

$\forall x$(x presents the stimulus σ to S $\rightarrow$ S applies "gavagai" to x)

If there is a question about undeterminacy of description with respect to phenomena it will be a trouble for psychology.

Third, the defeasibility of such a meaning definition would not depend on any normative character of meaning, but on the epistemic uncertainty embedded in the scientific enterprise of determining Φs which are T, since like any other enterprise in natural sciences, inductively established generalizations of psychology might reveal to be false.

Fourth, the predicate T doesn't provide any analytical guarantee on the usage of meaning-labels inside our meaning-talk. Consider for instance the following meaning statement:

> S means *Gavagai* by "gavagai" iff S has dispositions to apply "gavagai" to ⌜*Gavagai*⌝s.

Here, in the Naturalist's terms, the *stimulus-meaning* of "gavagai", *Gavagai*, is simply defined over a particular set of stimulations embedded in a presumably true psychological generalization. That means that the meaning-label ⌜*Gavagai*⌝ is presumably T and that we are *justified* in using it in our meaning-talk. But, again, here the justification doesn't come from an implication but from the availability of a contextual definition to discharge ontological commitments. As a consequence, this guarantees that any expression which is *stimulus-synonym* with "gavagai" can't fail to be such, but, notice, not in the sense that any established synonymity relation won't be found failing, rather in the sense that no weird analytic hypothesis can be invoked to undermine any synonymity relation established among T labels. Again, the application of the label ⌜*Gavagai*⌝ may be discovered be supported by a false generalization between stimulations and speaker reactions, but it can't be found that, if a meaning-label is T – i.e. it is contextually definable in terms of true gerenalizations – it determines wrong meaning relations. In other words, if ⌜*Gavagai*⌝ and ⌜*Rabbit*⌝ are both T, then if *Gavagai* = *Rabbit* then it is *necessarily* so: this would make them "synonyms", in the sense of "guaranteed to be coextensional".

Whilst the hardcore nominalist could still find something to complain about in this reconstruction of a paraphrase for meaning-talk based on the ground level of psychological laws, I'm not much interested here in accepting that kind of challenge. Moreover, I don't really think this is a really sound argument. However I think the whole structure of the argument it supports is correct enough for us to appreciate the sort of answer the Nominalist could respond to the challenge of modality with his thesis of naturalized epistemology.

1.2.1.3 Predicates projection

I've tried to develop the idea of naturalized epistemology to its deails in order to show how it could really represent an answer to the kripkean skepticism, but I'm not sure I didn't lose the Naturalist at some point along the way: would he really be willing to endorse this solution to the analyticity problem in epistemology? What's for sure is that I had an easier time drag behind me a supposed strawman than Quine himself. In my defense I have to note that the conclusion to which I lead the argument above is not fictious at all. Indeed it was put forward and maintained by someone how shared Quine's urges to deal with what Sellars called "nominalistic proclivities"[40] of traditional empiricism. I'm referring to Nelson Goodman.

It is because of these proclivities, for instance, that they were forced to face the above inconsistency of theses A, B and C, because they couldn't concede that empirical knowledge, as manifested in propositional usage of concepts, is immediate and unacquired in the same way as sense datas are supposed to be: they couldn't simply reject thesis C.. I'm not questioning this choice, nor does Sellars himself propose an argument against thesis C: he seems to take it for granted that there are some good point lying behind certain tenets of nominalism[41]. This is why, even if Goodman's collaboration with Quine on this topic culminated, as it's well knwon, in the paper they wrote together in 1947[42], just one year after Quine's conference on Nominalism[43], I'm not interested here in Goodman's "calculus of individuals" and its possible applications. Rather than in Goodman's nominalism *per se*, I'm interested in evaluating how he applied his tenets – which produced his nominalism and are thus, in this sense, very similar to Quine's – to solve the above inconsistency, and expecially how he tried to solve the puzzle about analiticity and modality.

What I'm interested in are the results of Goodman's analysis of inductive inference which is *consequent* to the analysis of those counterfactuals such as

> If x looked red then it would be red.

The problem of counterfactuals is well known. Goodman points out two main issues about it.

[40] Sellars [147], p. 132.

[41] Is Sellars a nominalist? There's a vast bibliography on Sellars's nominalism. Yet, I think the right question to ask would be: what is it for Sellars to be a nominalist? While a proper answer to such a question would require another book, maybe more, what Sellars thought to be the insightful core of Nominalism, if I'm following him right as I hope, is also at play in the present discussion and can be evinced from it.

[42] Especially Goodman and Quine [62].

[43] Quine [120].

The first one is that the specification of standard conditions has to be added to the antecedent for the counterfactual conditional to be valid, but this procedure is not unproblematic because it may reveal problems of consistency. Goodman temptatively proposes the requirement of "cotenability" to deal with them. I'll come back to this later and I hope I'll be able to show how this issue is but another side of the same puzzles that originates the second issue.

The second issue is about laws. Goodman's idea is that countefactual conditionals have to be supported by a general principle for them to be used in an inference:

> "In order to infer the consequent of a counterfactual from the antecedent A and a suitable statement of relevant conditions S, we make use of a general statement; namely the generalization of the conditional having $A \cdot S$ for antecedent and C for consequent."[44]

Thus the generalization supporting our above example of counterfactual conditional would be

$$\forall x(x \text{ looks red } \cdot S \rightarrow x \text{ is red}).$$

Now, the point is that this generalization must not be merely accidental, but has to be lawlike. In fact, if it simply said that all things that happend to look red were red – if it merely was a universal generalization over actual particulars –, it would be falsified by this orange which, though it never happened to look red, if it will ever happen to look red, wouldn't be red nonetheless[45]. Thus, for a generalization to be lawlike it has to deal with possible particulars. This is what convinces Goodman to focus on the analysis of dispositional translations of counterfactual conditionals, since dispositional predicates apply to objects not only in virtue of their actual instantiations, but also in virtue of their

[44] Goodman [60], pp. 17-18.

[45] Here one could correctly object that if that orange would ever look red, it would be in non standard conditions. In this sense, the problem with

$$\forall x(x \text{ looks red } \cdot S \rightarrow x \text{ is red}),$$

considered as a generalization over actual particulars, is that S may not specify the non-standardness of the conditions in which that orange would look red, since it never looked red.

This points out that the cotenability issue and the laws issue are strictly related, as Goodman himself recognizes in Goodman [60], p. 19:

> "The problem illustrated by the example [...] is closely related to that which led us earlier to require the cotenability of the antecedent and the relevant conditions, in order to avoid resting a counterfactual on any statement that would not be true if the antecedent were true."

What I want to show is *why* these issues are related and how they can be both solved.

possible instantiations. Thus, for instance, the dispositional predicate "is flexible" extends or 'projects' the set of actual instantiations of the manifest predicate "flexes" to the class of possible instantiations, because while x may not actually flex because it is not actually under suitable pressure, it is flexible because it would flex under suitable pressure. In the same way the dispositional predicate "is red" projects the set of actual instantiations of the manifest predicate "looks red" to the class of possible instantiations, because while x may not actually look red because it is not actually under suitable light conditions, it is red because it would look red under suitable light conditions.

The next step is to get clear of this relation of projection between manifest and dispositional predicates.

Let's pause here for a moment to make our point on the map: it's easy to realize that, even if we are not walking on his same path, we are traveling towards Quine's same destination. But it's not nominalism to make the difference: Goodman and Quine both moved from the same starting point and, as empiricists they both had to stop in front of the inconsistency pointed out above in Sellar's formulation. Here the Naturalist opted for a detour in order to get around the obstacle by declining from standard positions on epistemology; later he could try to reach his original destination through a nominalistic reconstruction like the one proposed above. But if one sticks to the standard epistemology, as Goodman does, he's forced to find a way to throw – or to 'project' – a bridge over the gap between actual experience and possible experience.

Induction is the more handy tool the empiricist Goodman has at his disposal to build such a bridge. Yet, as Goodman realizes, induction can't provide a modal bootstrap, and this latter insight is what makes his path parallel to Quine's.

The problem of induction, Goodman says, has been misconceived, since it has been construed as the problem of justifying predictions: how can we be sure that the sun will rise tomorrow morning, given that it's always risen till today? Hume already gave the correct answer: we can't. The point is rather to justify inductive inferences themselves, i.e. to establish which inductive inferences are good inferences, in the same way as it has been done for deductive inferences. Here is where the so called "new riddle of induction" comes to highlight the same problem with analyticity we found above when we faced Sellars's question about the necessity of a definition such as

> x is green $=_{Def}$ x would look green to standard observers in standard conditions.

What if the correct predicate to project were "grue"? How can we tell if the inductive

inference from "x looks green" to "x is green" is valid?

Goodman thinks that the analysis of the relation of prejectibility would lead to a distinction between valid and invalid inductive inferences[46]. Roughly, the idea lying under the notion of projection is to consider not only the inductive relation between a universally quantified statement and the particular statements upon which it is established, but also the validity of past inductively established generalizations. This, Goodman hurries to notice, has not to be construed as a meta-induction on the validity of inductively established generalizations: an induction from the actual episodes of verification of the inductively established generalization to the validity of the generalization itself. Such a meta-induction would manifest at the meta-level the same problems of the object-level induction. Projection is rather to be defined as an induction on second order concepts: the definition of projection is the projection of the manifest predicate "projected" to the dispositional predicate "projectible", which means a generalization of the class of actually verifi*ed* inductively established generalizations to the class of possibly verifi*able* inductively established generalizations:

> "We ask not how predictions come to be made, but how – granting that they are made – they come to be sorted out into valid and invalid."[47]

From his empiricist point of view, Goodman thinks this can be done within the scientific practices themselves: let alone the risk of mistakes, it's possible to establish by induction on actual projections lawlike generalizations concerning projectible predicates. In his informal account he sketches some rules for evaluating the projectibility of predicates based on the criteria of *entrenchment*: roughly, different predicates can be more or less entrenched within a theory depending on how many times they have been already *actually* projected. When conflicts arise about predicates projectibility that would produce inconsistencies within the theory, as in the case of "green" and "grue", the more entrenched predicate is to be considered the projectible one. To this basic idea some details should be added to account for the consequences the logical relations among predicates have in their projectibility.

However vague this may be, it represents an answer to the 'grue paradox': obviously "green" is a much better entrenched predicate than "grue". Thus, in this sense, it also represents an answer to the challenge about analyticity we posed to Goodman's empiricism, because the notion of projectibility is a mean to justify an account of empirical knowledge on inductive grounds.

[46] *Ibid.*, pp. 86-87.
[47] *Ibid.*, p. 87.

This explains what Goodman has in mind when he writes:

> "the dispositional predicate 'is orange' is a projection of the manifest predicate 'looks orange"'.[48]

Here again the idea can be formalized like follows

$\forall x$ [("looks orange" is projectible $\wedge$ x looks orange)$\rightarrow x$ is orange],

which is an instantiation of

$\forall \Phi, x$ [(Φ is projectible $\wedge$ Φx)$\rightarrow$x is $pr(\Phi)$],

where $pr(\Phi)$ is the predicate projected from Φ.

In this sense, if a predicate is projectible, the problem of evaluating all possible particulars for its extension can be discarded. In other words, since "looks orange" is a projectible predicate, one can legitimately adopt the hypothesis, for example, that "pumpkins *are* orange", while not every possible pumpkin is orange because in some possible world the orange-looking might be due to the prismatic effect of some particular protein pumpkins have on their peel that makes them look orange while they are actually pink. Whenever such a pumpkin would be found and the generalization that supports projectibility of the manifest predicate "looks orange" would be tested against it, it would be falsified.

In fact, the crucial point to notice here, as before, is that projections may represent predictions that might be falsified in the future: the theory of projection doesn't tell if the hypothesis "emeralds are green" is true, it rather tells that 'grue paradox' doesn't hold. A predicate is to be considered as projectible in virtue of its projections in possible situations, but in virtue its actual projections and thus there are no evaluation of correctness involved in the procedures of distinguishing projetible from unprojectible predicates.

> "The reason why only the right predicates happen so luckily to have become well entrenched is just that the well entrenched predicates have thereby become the right ones. If our critic is asking, rather, why projections of predicates that have become entrenched happen to be those projections that will turn out to be true, the answer is that we do not by any means know that they will turn out to be true."[49]

[48] *Ibid.*, p. 56n.
[49] *Ibid.*, pp. 98-99.

This formally corresponds to the results of the nominalistic reconstruction of meaning-talk we proposed for the Naturalist. There is no guardantee that the scientific practice will ever falsify inductively established generalizations of presently projectible predicates, as there is no guarantee that generalizations of psychology won't ever be falsified. What entitles both Goodman and the Naturalist to talk about empirical knowledge is that contents are guaranteed to be counterfactually robust as far as the generalizations that supports them have not yet been invalidated. As before, suppose now that both "looks like a rabbit" and "looks like a gavagai" were predicates respectively projectible into "is a rabbit" and "is a gavagai", and suppose that they both applied to the same class of actual particulars; then "is a rabbit" and "is a gavagai" would be "synonyms" in the sense stated above.

1.2.2 Taking directions

The previous section highlighted the general structure of an argument which is broadly common to the empiricism-laden reasoning when confronted with the question about meaning. Its declared target is to establish a bridge from statements about particulars given in empirical experience to conterfactually robust generalizations. The purpose is double and the stakes are high: on the one side such a bridge would settle the question about the status of empirical laws, on the other side it would describe an account of meaning in descriptive terms avaliable to empiricist analysis. The main obstacle is the notion of analyticity. The strategy to overcome it is to reduce the question about meaning identity to the question about validity of regularities. With respect the the inconsistency stated above in standard Empiricism's tenets, this amounts to reject thesis A in Section 1.2.1.2 above by showing how to construct an inferential relation between experience and propositional empirical knowledge. We examined two samples of this strategy that however quite different theoretical contents to put some flash on these same formal bones.

Now it's time to venture an evaluation of this argument.

1.2.2.1 Myth busting

I already introduced Sellars's criticism of 'the Given', so it will not be a surprise to find it here at play against such a myth as it infests the structure of the argument just presented. I'm not about to rehearse now the whole path through *Empiricism and the Philosophy of Mind*[50] which, presumably, has already been well trodden by any responsible philosopher.

[50]Sellars [147].

My aim is rather to show how it applies here as a part of the wider struggle Sellars engaged with this thick mythology so unconsciously rooted in the arguments of traditional empiricism. The other main part, as Brandom [27] began to show, is the analysis of modality. Let me venture to approach this huge topic from the position we got here.

First of all I need to evoke another strawman, a goodmanian one this time, to support a flawed interpretation of the theory of projection doomed to fall under sellarsian criticism.

Now, what if one, suspecting a regress, asks how to single out the class of possible occurrences of the manifest predicate "looks orange"? The goodmanian strawman could answer that such a justificational scruple has nothing to do with the the problem of induction and of dispositions – and thus, in our perspective, of meanings: the fact that "looks orange" is a manifest predicate means that orange-lookings are a very special sort of particulars whose content is given. The distinction between actual and possible orange-lookings doesn't make any sense because, while something that looks orange here and now may look pink under some other circumstances (in another possible world), an orange-looking can't be a pink-looking, ever. Notice that the fact that one *can't* distinguish before certain time t a green-looking from a grue-looking (and correspondingly that one can't distinguish after that certain time t a blue-looking from a grue-looking) doesn't mean, according to this reasoning, that green-lookings *are* grue-lookings before t (and correspondingly that blue-lookings *are* grue-lookings after t): that's why "looks grue" is to be considered a manifest predicate.

In *Empricism and the Philosophy of Mind* Sellars frustrated this cartesian line of argument once and for all. Here in fact is where the mythology of the given gets exposed. As we learned from John's troubles with reporting colours of his ties, *look*-talk is not authonomous from *is*-talk, in the sense that stating that something looks green (or grue) is not to report an empirical fact more basic than the fact that it is green (or grue): it is rather the way we make explicit the disavowal of the full commitment to report something as green. The notion of projection was supposed to build a bridge from experience to knowledge. But if this whole building has to lean against the idea that there must be some basic empirical contents immediately given in experience, the design is flawed and the construction must be stopped before it starts.

Notice that it is not trivial to blame our Naturalist for this sort of commitment to the Myth of the Given: his original holism about meaning forced him to retreat from analyticity and to take a long detour through naturalized epistemology and nominalism to get to an account of empirical knowledge. Such an effort now seems to pay him back

because he's free from any commitment to immediately given contents: his ⌜Φ⌝, far from being a manifest predicate like "looks Φ", is merely a label that sticks to things producing certain pattern of stimulations that do not *carry* any content.

In this sense the main mistake in the reasoning of the goodmanian strawman would be to slip over the distinction between green-lookings and grue-lookings. Without any previous grasp on meanings conveyed by dispositional predicates "is green" and "is grue", there's no way to make sense of such a distinction because grue-lookings are perceptively indistinguishable from green-lookings before t. The lesson to be learned, then, would be that meanings are underspecified with respect to observational experience: there is no univocal way to track down conceptual contents to particular patterns of experience in order to produce a *logische Aufbau der Welt*.

Thus, while the structure of the argument is equivalent, the aims of the Naturalist are extremely less ambitious: there is no *epistheme* of the world to be established but a psychological theory about how we make experience and a nominalistic reduction of epistemological talk to psychological talk – in Sellars's words, a reduction to a *psychological nominalism*[51]. His account seems to be successful in avoiding the Myth of the Given.

Unfortunately for the Naturalist, this was just the first test. First of all, it must be noticed, this idea of a reduction to psychological nominalism is too rough to provide an account of conceptual content: we were able to represent synonimity in terms of guaranteed equiextensionality, but how could we introduce logical relations like negation or implication[52]? In this sense the Naturalist's meaning-talk is but aping proper meaning-talk. Too bad for this quinean strawman: Quine, after all, explicitly rejected meaning-talk. But secondly, and crucially, there's a deeper lurk where the Myth hides. Let me clarify this by introducing one of Sellars's more clear-cut representation of the Myth:

> "The idea that observation 'strictly and properly so called' is constituted by certain self-authenticating nonverbal episodes, the authority of which is transmitted to verbal and quasi-verbal performances when these performances are made 'in conformity with the semantical rules of the language', is, of course,

[51] Sellars [147], p. 160.

[52] It's easy to see that we can't exploit extensional framework because there are no incompatibilities in meaning-labels, as for instance

$$\forall x(\ulcorner\Phi\urcorner x \rightarrow \neg\ulcorner\Psi\urcorner x).$$

And there are no such incompatibilities because they are not traceable back to incompatibilities among patterns of stimulations, in the absence of conceptual cathegorization.

the heart of the Myth of the Given."[53]

Quine's point is that since there's no way to establish determinate relations from nonverbal episodes to verbal performance that may transmit authority, verbal performances can't be justified in terms of semantic rules. This obviously defuses the Myth but it doesn't get rid of it. In fact, it shows up again as soon as another way is found to deploy the authority of self-authenticating non verbal episodes, like, for instance, the naturalistic reduction of epistemology to psychology. In this sense the Naturalist's main mistake is to take for granted the possibility to cathegorize regular pattern of stimulations as independent from the deployment of a whole space of conceptual contents. This mistake he shares with his real prototype, Quine himself. Notice that no cathegorization is realized by the application of meaning-labels to things that produces certain patterns of stimulations, since that is a nominalistically acceptable procedure that produces what Sellars would call distributive singular terms, i.e. collective names for groups of objects or names for distributive objects[54]. Cathegorization is realized when these labels are embedded into lawlike generalizations with the purpose to account for their features within a theory of psychology.

In other words, when psychology establishes laws that make explicit regularities in the relations between patterns of stimulations and linguistic behaviour, it produces cathegorizations of patterns of stimulations that can be formulated as:

$\ulcorner\Phi\urcorner$s do f.

Notice how this matches the blindspot of extensional semantics diagnosed in Section 1.1.1.2.

1.2.2.2 Strawmen and myths

The reason why I decided to argue with so many strawmen instead of real authors was neither because I needed someone at my level to confront with nor because I felt too much respect for the real authors to arraign them patently and hand down a sentence. I rather needed, for the clarity and completeness of my argument, to cover some positions that the real authors, legitimately, would have refuted to occupy. In this sense, it seems that I've been following Sellars more closely than I admitted at first. On the one side the unteanable position of the goodmanian strawman can be recognized in the theses Sellars criticise in Sections 13-19 of *Empiricism and the Philosophy of Mind*. Yet, I think that the real Goodman had clear enough the point about induction to avoid to fall prey of the

[53]Sellars [147], p. 169.
[54]See Sellars [154].

Myth of the Given so naïvely: he explicitly maintained that the choice between concurrent projectible predicates was not to be decided by observation[55]. On the other side the thesis of the quinean – or, should I say, humean, – strawman corresponds to what Sellars represents in Sections 29-31 as a more subtle but equivalently risky trigger for the mythology of the given:

> "if we *should* happen, at a certain stage of our intellectual development, to be able to classify an experience *only* as *of the kind* which could be common to a *seeing* and corresponding qualitative and existential *lookings*, all we would have to do to acquire a 'direct designation' for the kind of experience would be to pitch in, 'examine' it, locate the kind which it exemplifies and which satisfies the above description, name it – say 'ϕ' – and, in full possession of the concept of ϕ, classify such experiences, from now on, as ϕexperiences."[56]

Notice, as Sellars makes abundantly clear till the very last section of *Empiricism and the Philosophy of Mind*, that here the problem is not with this conditional itself but with a certain way to verify its antecedent: the idea that the ability to sort out determinate repeatables could be taken for granted.

Now, the real Quine would have been happy to endorse this conclusion and would have dismissed any further criticism of mine, simply by pointing me at *Word and Object*. And he would have not accepted to let meanings thrown out of the door come back from the window in Section 1.2.1.2, even after a strict nominalistic treatment: he simply wouldn't see the point in doing that, since he thought to have proven them to be dispensable. This is why I needed another strawman, the Naturalist, to endorse the task to confront patently with this blindspot. My point in doing that was to show that in his appeal to naturalism Quine was however blind to the problem of determinate repeatables, and that this blindspot crucially affects the premises of his argument for the indeterminacy of translation[57]. What he couldn't see is that the very logical space of repeatables comes together with their determinations. Thus the the failure of the linguist's work in the field follows from the description of a procedure for generating alternative ways to reidentify competing determinations of repeatables: first a certain determination of repeatables is identified in the native's language, then an alternative determination of those repeatables is provided in the linguist's language, and at last the identity between the two determinations

[55] Goodman [60], p.98-99.
[56] Sellars [147], pp. 156-157.
[57] For the diagnosis of this blindspot I refer to §30 of Sellars [147] and to Brandom [14], pp. 409-412.

is projected again into the native's language. But this procedure crucially hinges on the assumption that the whole logical space of particulars as acknowledged by the interpreter, with its sortals and the relations among them, can be projected to the speaker.

The proclivity not to question this assumption is but the form the Myth of the Given assumes in truthfunctional extensional semantics.

1.3 Modality without analyticity

1.3.1 Recourse to intensions

The reader who ran through all the previous sections may be convinced by now that modality is not only crucially involved in meaning-talk, but also that it has to be embedded in any theory of knowledge. And yet whoever is enough acquainted with the current analysis of modality will presumably refuse to consider this a really problematic remark. Has not Kripke already shown how to handle modality? Don't we have at our disposal technical tools, like possible worlds semantics and its developments, suitable enough to handle tightly that intensional part of meaning which is responsible for the bundle of problems we have been dealing with in the previous sections? We have seen that Hume's and Quine's arguments against modality were wrong, but is there something else? I think there is.

1.3.1.1 Kripke on 'extensionalizing' intensions

To begin with let me borrow from Brandom the following analysis of Kripke's contribution:

> "Kripke's provision of a complete extensional semantic metavocabulary for intensional modal logical vocabulary [...] is an adequate response to worries stemming from the extensional character of the logical vocabulary in which semantics had been conducted."[58]

I suggest, then, that maybe the grasp we have on modality seems so tight because it follows from a representation of intensionality in extensional terms and the adequacy of this representation is evaluated within the perspective of extensional semantics itself. This may be clear enough not to surprise anyone, untill it is shown that the perspective of extensional semantics itself may not be fully adequate. Now, since we've just cast some

[58] Brandom [23], p. 94.

doubts on this very extensional account of semantics, it's worth questioning our extensional representation of intensionality as well.

Kripke's formal point is well known and easy to state. Given a system of modal logic, an identity predicate can be introduced by adding to its axioms

I1 $x = x$

I2 $x = y \rightarrow (\phi \rightarrow \phi[x/y])$

where in I2 Leibniz's principle of indiscernibility of identicals can be recognized.

In this system the following is a theorem:

LI $x = y \rightarrow \Box x = y$[59]

The theory of rigid designation is the consequence Kripke derived from (LI) in semantics: this is the door through which intensions can be welcome in the extensional framework. And yet, the plausibility of rigid designation is not obvious, and there are those like Quine, paradigmatically, who are not willing to welcome intensions at all and who think that the door should be shut even at the price of renouncing to modal logic. After all, there are several reasons not to be content even with the burden of assumption the formal apparatus of modal logic forces to accept.

Rather than bog down into the debate about reference in semantics, let me try to propose a bird's-eye view on the whole picture possible worlds semantics. It will provide a rough definition of the details, but, I hope, it will make easier to trace the outline of how conceptual content is represented in this account.

From an empiricist point of view, there are two main sorts of complaints against quantified modal logic: (i) the ontological issue about the status of the possible worlds and (ii) the epistemological issue about how we can know things about them. As David Lewis made clear in his characteristic provocative way, issue (ii) is the main and more problematic one. In assuming Leibniz's Law as an axiom in I2, Kripke forces his system to grant susbtitutivity *salva veritate* of coreferential expressions, and this implies that coreference has to be fixed independently of intensional variability: given that intensional variability is represented as possible worlds variability, it implies that reference is to be considered as modally robust.

---

[59] Proof.

(1)	$x = y \rightarrow (\Box x = x \rightarrow \Box x = y)$	I2
(2)	$\Box x = x \rightarrow (x = y \rightarrow \Box x = y)$	Predicate Calculus
(3)	$\Box x = x$	I1 Nec
(4)	$x = y \rightarrow \Box x = y$	(2)(3) MP

But what does it mean to consider reference as modally robust? It means to introduce a sharp distinction between two sorts of vocabularies, two sorts of modalities and two sorts of knowledge: in epistemic terms it is a distinction between the knowledge of substances and the knowledge of accidents, in modal terms it is a distinction between metaphysical modality and epistemic modality, in semantic terms it is a distinction between the modally robust vocabulary of rigid designators and the modally variable vocabulary of descriptive designators. Intensions are regained in the extensional semantics by separating their role as means to fix the reference from their role as means to represent cognitive value: reference is fixed independently of cognitive value.

Then, if the modal character of conceptual content has to be vindicated according to this account of modality, it seems that we are forced to a stark choice: either we stick to metaphysical modality and we give up on epistemic modality or we stick to conceptual content we give up on modality itself (compare with the stark choice presented in Section 1.1.1).

On the one side, Kripke contrasted the modally constant interpretation function of some non descriptive vocabulary with the modally variable intepretation function of some descriptive vocabulary. Analyticity and necessity were bound to the the former part of vocabulary, the vocabulary of rigid designators. Thus notoriously

$$\text{Hesperus} = \text{Phosphorus} \Rightarrow \text{Necessarily (Hesperus} = \text{Phosphorus)}$$

i.e. if Hesperus is Phosphorus then Hesperus is Phosphorus from the beginning till the end of time, in any possible world.

The epistemic dimension of aprioricity was abandoned to the intensional variability of descriptive vocabulary, thus

Saul believes that the morning star is the evening star

can be false, because Saul may not know how things are. Too bad for Saul.

In this way Kripke made room for a new sort of *a posteriori necessity* for those cases in which constant interpretation functions would be discovered to be the same function, i.e. those cases in which two referential causal chains would be discovered to be originated by the same initial baptism. No hope for descriptive vocabulary to be modally robust since its interpretation may vary along possible worlds and these can't be all checked even *a posteriori*. Then if conceptual content is to be found in descriptive vocabulary, no conceptual (*non-extensional*) necessity is available.

On the other side, if one refutes to distinguish different referential behaviour of parts of vocabulary then he can't appeal to any modal distinction to handle intensional variability of conceptual content. This is Quine's position: for logical reasons he only recognizes the descriptive referential behaviour and maintains that singular terms are better to be eliminated and substituted with quantified statements:

> "Ultimately the objects referred to in a theory are to be accounted not as the things named by the singular terms, but as the variables of quantification."[60]

He also rejects the essentialism he sees as springing from the distinction among different ways to specify referents:

> "This means adopting an invidious attitude toward certain ways of uniquely specifying x [...] and favouring other ways [...] as somehow better revealing the 'essence' of the object."[61]

Now, the only way to regain the extensional grasp on the intensional variability of conceptual content would be to establish a corresponding distinction in epistemic terms. But if one reject the *synthetic / analytic* as well, as Quine does, no hope is left for a modal account of conceptual content.

1.3.2 Revisable analyticity

So the rejection of the *synthetic / analytic* distinction plays a pivotal role in the argument for the rejection of modality. But what if one is not fully convinced by Quine's arguments against such a distinction? This suggests that maybe a third way can be found out of the above dilemma between the renounce to an extensional analysis of conceptual content and the renounce to an extensional analysis of modality.

1.3.2.1 Going normative

Lance and Hawthorne, with their precious work on the normativity of meaning, *The Grammar of Meaning*[62], describe a possible path. They are not really concerned with modality *per se*, but with normativity, their main thesis being:

[60] Quine [124], pp. 144-145. See also Quine [121].
[61] Quine [124], p. 155.
[62] Lance and O'Leary-Hawthorne [82].

> "[A] claim of the form 'S_1 means S_2' *licenses* a pair of inferences: from S_1 to S_2 and from S_2 to S_1. [...] That meaning claims offer inference licenses suggests that while they do not report a descriptive regularity as the description picture would have us believe, they do state a normative property."[63]

Thus, their first crucial move is to adopt from the beginning a normative representation of semantics, and this makes their work particularly important for our discussion here, since it can be used as an advanced platform for a first recognition of the relatively unexplored land we want to colonize. It must be acknowledged, however, that this platform leans on the imposing but potentially controversial pillar built by Robert Brandom's deep analysis of language and normativity in his *Making it Explicit*[64]: one main component of this analysis is the adoption of an inferentialist semantics, which is what makes sense of the above construal of meaning statements as making explicit inferential rules. To accept all this obviously implies a huge step beyond the standard representational account of meaning, but I choose here to postpone any proper evaluation of this framework till Chapter 3. I have three reasons to do that: first, here I want to stick to the question about modal character of conceptual content and focus on some consequences this move has on this issue; second, Lance and Hawthorne themselves deploy the normative account of meaning before they evaluate its foundations; third, I think that while Lance and Hawthorne decisively opt for a normative account of meaning, they keep a certain standard understanding of their choice and this eventually limits the range of their insights so that they fall prey of some of the old misunderstandings. Let's keep following Lance and Hawthorne without further ado then.

To begin with, it must be noticed that even in this normative framework the foundational problem of the 'bedrock' can't be discharged for free. At first glance it may seem that since normative properties articulate justifications, then the 'bedrock' of justificational space has to be the limit of meaning analysis, but this reasoning actually does presuppose rather than explain a certain account of the norms that regulate successful justifications. Meanings are now represented directly in terms of normative rules of inference, but what does really prevent the contents of these rules to be underdetermined with respect to facts about meaning, even if these are normative facts? In this sense Quine's indeterminacy of translation thesis is still an unavoidable test bed for Lance and Hawthorne's approach. It is an easy test, though, for anyone standing on Brandom's shoulders.

[63] *Ibid.*, pp. 58-59.
[64] Brandom [14].

Doubts against Quine's main argument for this thesis were raised already at the end of Section 1.2.2.2, but here it's worth quickly reconsidering it in its bare structure. The indeterminacy of translation follows from two premises: the first one establishes epistemic holism, while the second provides a way to describe facts about meanings. Quine's favourite theses to support this latter premise are verificationism and behaviorism: verificationism, in a broad sense, describes facts about meaning as testable dispositions to react possessed by things in the world, behaviorism describes facts about meaning as verbal dispositions to behave possessed by members of a speech community. Lance and Hawthorne deal with Quine's arguments against analyticity and his thesis about indeterminacy at length[65]: by exploiting Brandom [14]'s normative account of meaning, they argue that, since meaning claims institute norms rather than describing practices, no question of inderdeterminacy can be raised. This claim is coarse indeed, but I'll postpone the careful analysis it deserves to Chapter 3. Here let's just be content with acknowledging it and move on. Once the result is established, Lance and Hawthorne hurry to reverberate it on Quine's preferred ground of analysis, the relation between semantics and epistemology. In fact their second move is to make room for a weakened *synthetic / analytic* distinction with a pragmatic flavour. They maintain that since Quine raises complaints against analyticity only to the extent that it teams with *aprioricity* and *epistemic privilege* in an holistic picture, it can be rescued from his criticism by showing how it can avoid bad companies. Then they ask a non quinean question: what does it mean to *treat* a claim as analytic?[66] The answer is purported to treasure some insights of Wittgenstein's analysis of linguistic practices in *On Certainty*[67]:

> "[Analytic sentences] are sentences against which bare challenges are, *de jure*, not in order. Their special and central status in our language game consists primarily in this: that challengers are forbidden to demand evidence without first supporting some claim incompatible with them."[68]

Here *de jure* unchallengeability is contraposed to *de facto* unchallengeability: the former is a pragmatic feature of some statements within a linguistic practice, the latter is a contingent feature of matters of facts and semantic rules. Meaning statements are thus provided with a *prima facie* entitlement that accounts for their being basic elements in the sellarsian

[65] *Ibid.*, pp. 19-54.
[66] *Ibid.*, p. 95.
[67] Wittgenstein [175].
[68] Lance and O'Leary-Hawthorne [82], p. 110.

game of giving and asking for reasons. And yet they are proper elements of the game because they can be questioned and eventually revised.

On the one side this makes the *synthetic / analytic* distinction a *local* distinction: our linguistic practices may be locally revised when proper challenges are supported within the practices themselves. On the other side it makes real sense of the choice to move the problem of meaning from the context of truth to the context of justification: the real point of determining the content of linguistic rules is not to describe a practice in a more accurate way, but to improve and amend the rules themselves.

1.3.2.2 A look back to Kripke

Now, this reconciliation with analyticity obviously seems to reassure our hopes to find a way out of the blind alley we have been driven into. But we are still far away from an account of conceptual modality. So let's try to take stock before we move on. If Lance and Hawthorne's pragmatic notion of analyticity is accepted, we are confronted with a picture in which analytic sentences (i) may not be known *a priori* and (ii) may be revisable.

Now, (i) is not something that may surprise anyone since *Naming and Necessity*: there Kripke took great care to stress the distinction between the modal plane and the epistemic one and to establish the metaphysical character of necessary truths which hold independently of our knowledge of them. But then, (ii) clashes directly with this very metaphysical character of Kripke's picture of modality: once epistemic uncertainty has been distinguished from intensional variability we have to accept that if something is true in every possible world, even if we may well not be aware of it, it won't ever cease to be true in any possible world[69].

[69]This can be misconstrued because sometimes Kripke's picture may be tricky. Let me recall it just to be sure to make my point clear.

Consider the sentence p, "water is H_2O". Since "water" and H_2O are names for substances and thus rigid designators, p is true in all possible worlds. Then $\Box p$.

Some well known observations follow. We, inhabitants of W_{Earth}, may not know at t_1 (let's say 1650) that water water is H_2O. We discover it only later, at t_2 (late XVIII century). Thus p is necessary *a posteriori*.

Obviously the necessary truth is not the *de dicto* one that *the substance that we, inhabitants of* W_{Earth}*, call water* is H_2O, but the *de re* one that *the substance* that we, inhabitants of W_{Earth}, call water is H_2O (where the scope distinction is crucial). The inhabitants of $W_{Twin\,Earth}$ may call "water" a substance whose chemical composition is XYZ, but it would be another substance, accidentally called "water" like the first one.

What could not happen is that p ceases to be necessary, i.e. that what is water at t_1 ceases to be water at t_2.

Here again, there could be a possible world in which *the substance that we, inhabitants of* W_{Earth}*, call water* is H_2O at t_1 and changes into XYZ at t_2, but that would not be the same substance that doesn't

Here I want to recall McDowell's warning not to be so hurry in rejecting the realist account of meaning together with its formal semantic apparatus. As I already said Lance and Hawthorne are not really concerned with modality and they are quite content with their epistemic results about *non a priori* analyticity, because, as they explain in the latter part of their work, it entitles them to beat Quine on his own naturalistic field. Yet, while focusing on the relation between modality and epistemology, they seem to miss the role of the other dimension where the acknowledgement of the normativity of meaning reverberates, i.e. the relation between semantics and modality. And what is at play on this second dimension is the determination of conceptual contents. In other words, if what really matters is to distinguish epistemic uncertainty from semantic indeterminacy, why bother with normativity of meaning? Why should one not be content with (i) and with Kripke's *a posteriory necessity*?

Lance and Hawthorne do not ignore this problem, but when they briefly deal with modality[70] they seem to be content with the proposal of a reformulation of the distinction *necessary / contingent* that may defuse the *aprioricity* of necessary truths: they suggest that *if the point is justification*, then it's enough just to consider *a posteriori* justification of necessary statements. In a sense they are quite right, because *if a normative semantics is taken for granted* there is no need moreover to require norms to be defined *a priori*. Unfortunately the point for Kripke's metaphysical modality is not justification but the determination of extensions in possible worlds, i.e. an extensional semantics is assumed.

Thus they complain against rigid designation, and their unsatisfaction is authentically quinean there:

> "First, consider the claim: 'Necessarily, the stuff that is H_2O contains hydrogen atoms.' Is this true? It may well be true. But if it is read *de re*, it is far from a trivial truth. [...] This does not, at all, impugn the claim that 'Necessarily, water is H_2O.' The trouble is that this is taken to be at once a *de re* claim and also to imply trivially that water essentially has 2 hydrogen atoms etc." [71]

What does "trivial" mean here? It, obviously to me, means "analytic". Thus the above means that if we consider sentences like "necessarily, water contains hydrogen atoms" in

change. How do we know that our water is H_2O and not $H_2O - till - t_2 - then - XYZ$? We don't know, but, again, according to Kripke, that is an epistemic point and not a modal one.

[70] *Ibid.*, pp. 165-171.

[71] *Ibid.*, pp. 167-168.

their *de re* interpretation, these, far from being immediate truths that simply depend on semantics and reference mechanisms, are scientific results with relevant cognitive value. Then, in order to justify "necessarily, water contains hydrogen atoms", one can't just rely on semantics, but one should argue on the scientific level.

And they keep beating the point:

> "The expression, 'the kind of stuff in the rivers,' by itself does not secure unique reference. For 'the stuff' in the rivers can be described in more or less specific ways".[72]

Here, they recall that the usage of definite description in initial baptisms doesn't guarantee the uniqueness of the substance to be rigidly related to an expression: if I state "the substance in this glass is water", how can I be sure that I'm baptizing one unique substance and not some spurious compound?

Now, it's worth asking what these remarks amount to, since they can be easily misunderstood. If these are construed as direct objections to Kripke, either they rehearse the doubts about the epistemology of direct reference in order to clarify his position, or they aim to a complete rejection of rigid designation.

In the first case they amount to the well known and quite innocuous objections Kripke himself deals with in *Naming and Necessity*[73]. For what concerns the first of the above supposed objections, indeed, it touches the very reason why Kripke's modality has been qualified *metaphysical*. But there's nothing trivial in this, at least if what is implied by "trivial" is that it would reduce synthetic truths to the analytic truths of semantics. Rather, Kripke maintains that there's a part of vocabulary, proper names and substance names, whose semantics is modally rigid, and another part of vocabulary whose semantics is modally variable. Maybe, one could provisionally say that the relation between a rigid designator and its extension is epistemologically trivial, but the crucial point to clarify is that this happens because, in Kripke's picture, the epistemic plan has nothing to do with the modal plan. As Putnam put it clearly, thanks to Kripke meanings can be treated as extensions and extensions are out there: we may be in trouble in identifying them, but they are given, and so is semantics. Once this is acknowledged, the answer to the second supposed objection is even straighter. Since epistemic mistakes can't affect the structure of possible worlds, it could well happen that the descriptive definitions we use in initial baptisms do not match our referring intentions: we may baptize with one only

[72] *Ibid.*, p. 168.

[73] Respectively Kripke [79], pp. 123-127, and pp. 135-140.

name different substances, as in case of impure samples, or even nothing at all, as for example in "Gabriel is the angel of the Annunciation", or "Mark Lance is the author of *The Grammar of Meaning*". Nonetheless those expressions would rigidly keep their meaning (or their absence of meaning), independently of our epistemic discoveries about water, angels or philosophical bibliographies. Notice, in fact, that by introducing the correct meaning statement for those expressions we would simply perform another baptism of a different substance with an omophone expression.

In the second case the picture gets dark for us. If direct reference is really to be considered a way to trivialize epistemology and thus it is to be rejected, then it is hard to rescue quantified modal logic from the criticism against *de re* modality. In this picture analyticity may be considered as a feature of the practices to justify statements, but, given the equivalence between necessity and analiticity, necessary statements are reduced to revisable analytical statements: Necessarily(water is H_2O) is true because, as far as we know, "water is H_2O" is analytical, that is it is preserved from critcism. A place is left in our tool-box for modality, but quinean doubts might be raised on its utility.

I don't think Lance and Hawthorne's remarks should be construed as direct objections to Kripke, since they explicitly state the following:

> "Suppose now one does away with *a priori* knowledge (in the traditional sense) and entertains only an innocuous version of the analytic / synthetic distinction. Need one thereby do away with the necessary / contingent distinction or else align it with the innocuous analytic / synthetic distinction? We deny that either move is required. It appears, rather, that some important, but distinct, theoretical role for necessary / contingent distinction can be made out."[74]

In fact this means that the traditional equivalence between the notions of analyticity and necessity has not to be rejected and, at the same time, that the analysis of modality is to be considered a worthy task. Thus, both the simple solution of the first case above and, more importantly, the dark picture of the second case shoud be avoided. But, the question now rises, how can we avoid them? How can that theoretical role for necessary / contingent distinction can be made out?

What I think Lance and Hawthorne's objections really blame Kripke for is the unsatisfactory account of the epistemological grasp we have on direct reference. Here we can fully

[74] Lance and O'Leary-Hawthorne [82], p. 166.

appreciate their concern about the point of justification: they ask how modal truths are to be justified and they construe this as a question about the "epistemology of modality"[75]. In this sense they complain against Essentialism: they reject the triviality with which, once reference is directly stuck to designators, sentences like "Necessarily water contains two hydrogen atoms" is to be considered as true even if we may not know anything about the stuff we call "water" except that it is identical to the stuff that is H_2O. And again they complain against initial baptizms: how can we stick a rigid designator to stuff we may not know anything about? What they ask for then is an account of how modal statements are justified in linguistic practices.

Here however, their reasoning dangerously bends out:

> "What picture, then, should we embrace concerning how necessary *a posteriori* truths get justified? Perhaps one way to an adequate answer here is to realize that *there does not have to be a special answer*. The one obvious epistemic barrier to such truths – that we do not have direct perceptual access to possibility space – is shared by plenty of other sorts of truths. For example: universal generalizations about all places and times."[76]

Now, the risk to enter this slope is to let the distinction between modally robust conditionals and universally quantified statements completely collapse. Rather than arguing directly in favour of such a distinction for the moment[77], what I want to notice here is that this is not just a minor concession to quinean epistemology, but a quite serious blind spot in Lance and Hawthorne's argument. In fact, as it's easy to realize, the distinction between lawlike generalizations and contingent generalizations corresponds to the distinction between *de jure* and *de facto* unchallengeability that supports the whole idea of *prima facie* analyticity. It's worth going back then to this distinction:

> "a view is treated as unchallengeable in the first sense [*de jure*] if it is a convention of a practice or consitutive of a practice that it is improper to look for reasons against it. [...] A view is unchallengeable *de facto* if no-one is able, or ever will be able, to mount a halfdecent challenge to it. We must, that is, distinguish between a constitutive rule against doing something and something that one is unable to do given the rules and background facts."[78]

[75] *Ibid.*, p. 168.
[76] *Ibid.*, p. 168.
[77] See Chapter 5 for a wider discussion of this point.
[78] *Ibid.*, p. 119.

Now, a view can be *de jure* unchallengeable without being *de facto* unchallengeable: for instance "'Gavagai' means rabbit" *could*, *de facto*, be challenged by providing alternative meaning statements that happen to equivalently match meaning facts, as "'Gavagai' means undetached rabbit part", and yet, since it is treated as analytic within a certain linguistic practice it *can't*, *de jure*, be challenged. But the crucial point to realize is that a view can be *de facto* unchallengeable without being *de jure* unchallengeable: suppose for instance that all the coins in my pocket are copper, then "if a coin is in my pocket it is copper" *can't*, *de facto*, be challenged, but it *could*, *de jure*, be challenged by providing a different description of how things might have been with my coins, since there is no rule dealing with the material of the coins in my pocket. If one misses this point, one could still think that *de facto* unchallengeability implies *de jure* unchallengeability and come to believe that the reason why the implication fails in the other direction is just epistemic uncertainty: *de jure* unchallengeability fails to imply *de facto* unchallengeability because we don't know which rules are actually unchallengeable. And we don't know which rules are *de facto* unchallengeable because we don't know which descriptions of matters of facts are *true*. In this sense, *prima facie* analyticity really turns out to be *uncertain* analyticity, i.e. analyticity not yet established, and where Lance and Hawthorne want to recognize *prima facie* valid rules, there is just epistemic ignorance about the relevant normative facts that would tell which rules are valid. Notice that this would jeopardize the argument against Quine's indeterminacy of translation: if the practices of instituting rules are really practices of improving the description of normative facts, the question still rises about the underdetermination of our best description of these facts with respect of how these facts might turn out to be. If one blurs the distinction between modally robust conditionals and contingent conditionals, one can't really escape the descriptivist framework through a normative account, and gets stuck in the familiar dilemma: either stick to modally robust representation of analytic relations and give up on the variability of the descriptions we give of such relations that depends on our epistemic uncertainty, or stick to the epistemic description of analytic relation and give up on the attempt to provide an *apriori* modally robust representation of them. Here we have to appreciate again the consequences of the proclivity to reduce semantics to epistemology, which seems to become irrepressible as soon as the issue of the definition of contents is raised.

I have to admit that maybe I've been a little too hasty with my call here: after all what Lance and Hawthorne really have in mind, while comparing the practices of justification of modal statements to those of universally generalized statements, is the process of amending

scientific theories rather than the process of establishing semantic relations, and they would easily discharge the above accusation by rehearsing Brandom's account of normativity. Yet, I think it's worth enough being resolute in warning against this oversight, which is so deeply buried, if not in Lance and Hawthorne's, certainly in standard approaches to normativity.

Thus, a redefinition of modality may not be one of Lance and Hawthorne's tasks, but if their normative analysis of meaning has to be taken seriously, it is an urgent task indeed: the question to be asked is how the revisability of analyticity could be received in an account of modality. As we will see in Chapter 5, an account of modality in terms of justifications is much more than a slight revision of *a posteriori* necessity, but this is the real way out of the dilemma between Quine's rejection of an extensional anslysis of modality and Kripke's rejection of an extensional analysis of descriptive vocabulary.

1.4 Conclusions

The question about meaning can be legitimately construed as a question about meanings' nature. What sort of 'things' meanings are? What does it mean to get in touch with them? What do they have to do with our linguistic practices? But the question about meaning can also be construed as a question about meanings' content. How are meanings determined? What are they about? What is their logic? A certain tradition pictures these as separate sorts of questions and standard extensional semantics encourages a separate account of them: on the one side an epistemological demand, on the other side a logical demand. As a consequence any attempt to meet both of these demands has to face the difficult task to put together epistemic uncertainty and logical objectivity. It is a spread and lasting tradition indeed. On the one side for instance, Lance and Hawthorne, who do reject standard semantics, yet suggest that the right question to ask about meaning is optional:

> "The questions we have in mind address not what *constitutes* meaning but instead the *point* of meaning talk and, relatedly, how such talk functions in a broader socio-linguistic practice. [...] One could approach an investigation of meaning with either sort of questions. [...] [T]he difference is (roughly and uncritically) that one approach asks what meaning talk does for us, another asks what corresponds to it."[79]

[79] *Ibid.*, p. 11.

On the other side for instance, still in 2006, Hattiangadi recognizes that the choice between normative and truthfunctional semantics is seldom supported by a proper argument[80], but then she proceeds to reclaim the demand for content representation against any account that sets for normative semantics simply by reasoning about meanings' nature[81].

Here I've been questioning this very picture and I traced its origins, at least in the explicit form it acquired in the recent philosophical debate, back to a certain understanding, or mis-understanding, of Wittgenstein's remarks, due to Kripke's famous account. Then I tried to exploit some of the accounts of meaning developed within the borders of this so defined *standard* framework in order to single out some major blindspots. In correspondence to these blindspots I tried to put forward suggestions that could aid to break the borders of this frame and achieve a more perspicuous representation of what is at play when the question about meaning is raised.

1.4.1 The first blindspot

In Section 1.1.1.2, after a quick analysis of the extensional approaches to the problem of normativity, we got a first glance of what was going wrong. Yet it took us other two sections of careful excavation into epistemology and logic to dig up what we found. We had to penetrate a kernel of assumptions deeply rooted in traditional Empiricism and yet unwarranted, that Sellars diagnosed under the rubric 'Myth of the Given'. Among the many forms of the phenomenology this myth manifests itself there's the idea that the structure of repeatables may be accounted independently of the structure of their determinations. This is the furrow where truthfunctional extensional semantics can take roots and grow into a tangled shell that soon may easily become impermeable to the acknowledgement of the normative features of meaning: Russell's reinterpretation of fregean quantifiers as deprived of modal character[82], which lies at the starting point of modern first order logic, can be applied only to a previously given domain of bare particulars. The therapy against this flaw in semantics steps through the acknowledgement of the interdependence between descriptive and modal vocabulary. We'll come back to this in Chapter 3 where we'll put it explicitly into play, but in the meanwhile it's worth store this suggestion explicitly and keep it for future use: conceptual content is intrinsically characterized by modally robust relations that are not reducible to the description of states of affairs.

[80] Hattiangadi [70], pp. 220-221.
[81] *Ibid.*, pp. 232-238
[82] See Macbeth [89].

1.4.2 The second blindspot

But even when this modal character is acknolwedged, things do not move on smoothly. In fact, while the analysis of the relation between analyticity and modality is not really a brand new enterprise, the results that have been obtained still manifest the same difficulty to keep together a certain view of logic with a certain view of epistemology. Thus, we opened Section 1.3.2 with the diagnosis a clash between the need for an extensional account of intensions and the need for an account of conceptual content: on the one side Kripke's picture of modality that provides an extensional account for intensions but splits conceptual content into metaphysical content and epistemic content, on the other side Quine's holistic picture that keeps together content but gives up on its semantical analysis. This means that Kripke's extensionalization of intensions still falls short of a representation of meanings that may be perspicuous enough to account peacefully for both epistemological and logical demands.

In order to take one step forward there we had to take for granted a normative account of meaning that could properly develop our first suggestion and so we followed Lance and Hawthorne along the path of Brandom's account of inferential semantics. In this sense we had to sign a committive promissory note, but it let us to establish two important points: we acknowledged that meaning statements do not describe norms, but institute them, and we also acknowledged that analyticity of sentences depends on their role in the practices of justify statements. The practices of instituting norms determine meanings by establishing normative relations among sentences, for instance by establishing that certain sentences are *de jure* unchallengeable. Because of their normative character, these relations are to be properly represented as modally robust: since they are not descriptions of regularities that may turn out to be more or less accurate or *correct*, it is not the case that they may be *false* in counterfactual situations. Yet, crucially, because of the dynamics of the practices, the norms may be modified and amended, and this means that the unchallengeability of analytic statements is modifiable and amendable too. Thus we realized that conceptual content is revisable.

To be fully explicated, this second suggestion requires a deeper analysis of normative facts and of how they supports meanings, i.e. of the dynamics of the linguistic practices. But it also requires an account of modality that could cope with a notion of revisable analyticity. This second requirement is a pressing one because, once Kripke's account has been rejected and a naturalistic reduction has been excluded, there seems to be no obvious way to represent conceptual contents as determined, if we don't want to accept them as

given. These requirements will be met respectively in Chapters 3-4 and Chapter 5.

Chapter 2

Chalmers on modality, meaning and reason

2.1 Setting the stage

While searching for an account of the normativity of meaning we had to face what appeared to be a dilemma between the extensional analysis of metaphysical modality that sacrifies necessity of conceptual content, and the rejection of modality itself. It's worth noticing here that, from a formal point of view, this may be not a real dilemma. It's possible to keep the extensional framework of Kripke's modality while discharging modal rigidity of non-descriptive vocabulary. Indeed, here is where, for example, two dimensional semantics looks for a better analysis of modality. In what follows I am going to exploit David Chalmers's presentation of 2D-semantics[1], but I won't load the burden of a deep evaluation of his assumptions. While these seem to be the most controversial part of two dimensional framework, I prefer to assume them diligently myself and try to derive the consequences that are relevant to the present analysis of modality. It is the evaluation of these consequences I'm interested in.

Two-dimensional semantics can be generally qualified as a possible world semantics. It considers meanings as to be represented by truth functional intepretations over possible worlds. However, while standard mono-dimensional semantics deploys the possible worlds apparatus only in order to represent modal variability, 2D-semantics requires every expression to be variably interpreted also for *actual* worlds. The basic idea is not that new: as 2D-semanticists often acknowledge, for the inspiration of their works they are deeply in

[1]See, for instance, Chalmers [35].

debt to those analyses like Kaplan's account of indexicals in terms of the two dimensions of "character" and "content". Two kinds of *possibilities* are to be considered: *epistemic* possibilities, i.e. hypotheses about how the world might actually be, and *metaphysical* possibilities, i.e. hypotheses about how the world might have been. In this sense, for instance, while metaphysically impossible – for reasons Kripke taught us –, it is epistemically possible that Hesperus is not Phosphorus: our world might have turned out to be one in which Hesperus and Phosphorus are indeed two different objects. Thus for example "water" may have different referents in different possibly actual worlds (which are called *epistemic scenarios*), according to how these references were fixed, and different referents in any counterfactual world built upon these. Three kinds of intensions are considered: a *primary intension* (1-intension), assigns to an expression a reference in actual worlds (scenarios), $f : S \to E$[2], a *secondary intension* (2-intension) assigns to it an extension in counterfactual worlds, $f : W \to E$, and eventually a *two-dimensional intension* (2D-intension) assigns to it a secondary intension, $f : S \to (W \to E)$. Primary intensions and secondary intensions are functions from worlds to extensions, 2D-intensions can be also represented as functions from pairs of worlds to extensions. This suggests a matrix structure for the representation of 2D-intensions, for instance,

"water"	w_0	w_1	$\cdots$
s_0	H_2O	H_2O	H_2O
s_1	XYZ	XYZ	XYZ
$\vdots$	$\vdots$	$\vdots$	$\ddots$

Extensions, i.e. the value of the interpretation function when the same maximally consistent description is considered both as the actual scenario and as the counterfactual world we are in, can be reconstructed by diagonalizing 2D-intensions.

Notice that Kaplan's account of indexicals matches this framework: indexicals have *contents* that specify their extensions in any counterfactual situation (thus behaving like

[2]This is not exaclty Chalmers's definition of primary intensions. His original definition characterizes primary intensions as "functions over centered possible worlds", to the extent that for any world W, the primary intension of S is true at W iff "if W is actual then S is a priori". In this sense centered worlds Ws need to be *canonically described*, in *neutral vocabulary* (i.e. in a purely descriptive vocabulary that may prevent modal or epistemic rigidity). The point Chalmers takes care to highlight is that there's no guarantee that any epistemically possible world may be specifed as "centered". I'll have to skip overe these details in the following.

2-intensions), and *character* that determine their contents in any different context (thus behaving like 2D-intensions).

Now, the point is that by exploiting this double perspective 2D-semantics allows to reconstruct the standardly established relations among realms of modality, reason and meaning. I'll follow again Chalmers in summing up these relation in three theses, associated (maybe a little roughly) with three big names in philosophy and logic.

(Kant) A sentence S is necessary $\Leftrightarrow$ S is *a priori*

$$\text{NECESSARY} \longleftrightarrow \text{A PRIORI}$$

$$\text{ANALITIC}$$

This establishes a correspondence between modality and reason: everything that is necessary can be known a priori and everything a priori has to be necesarily true. On the other side, if something is contingent then it can't be known but a posteriori, and, respectively, everything that has to be known a posteriori is to be contingent.

(Frege) Two expressions "A" and "B" have the same sense $\Leftrightarrow$ "A≡B" is congnitively insignificant[3].

$$\text{NECESSARY} \qquad \text{A PRIORI}$$

$$\text{ANALITIC}$$

This establishes a correspondence between meaning and reason: meaning identities are known apriori, while a posteriori sentences deal with different meanings. Notice, however, that Kant refused ($Frege \Leftarrow$). Notice also that Frege himself refused ($Frege \Leftarrow$) for what concerns geometry.

(Carnap) "A" and "B" have the same intension $\Leftrightarrow$ "A≡B" is necessary

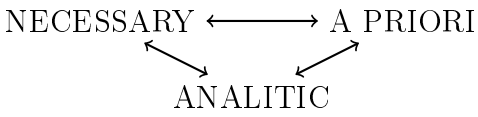
$$\text{NECESSARY} \qquad \text{A PRIORI}$$

$$\text{ANALITIC}$$

[3]Here in "A≡B", the relation "≡" represent equivalence between A and B, in the sense that A and B have the same extension. However, since A and B may represent different kinds of expressions with different semantic values, "≡" may indicate identity "=", biconditional "iff", etc.

At last (chronologically) here it comes Carnap's analysis of intensions that establishes a correspondence between modality and meaning. Notice that this is what really established, by implication, ($Frege \Leftarrow$).

The structure of relations pictured by these three theses is a tight triangular equivalence among the three domains of modality, reason and meaning; this is what Chalmers names the "golden triangle":



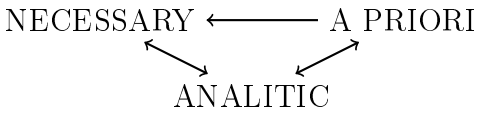
Unhopefully a kripkean snake came into this rationalistic Eden claiming the rigidity of referential behaviour of given pieces of vocabulary, with the consequence of subverting its balance.

Two theses, (K1) and (K2), sum up the consequences of Kripke's analysis.

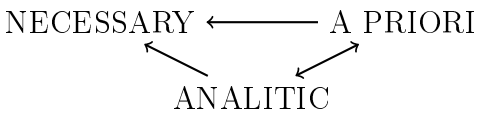
(K1) There are some necessary but non *a priori* sentences: $\Box p \wedge \neg \mathcal{A} p$

where "$\mathcal{A}$" means a priori.

For example "Hesperus is Phosphorus", while being an astronomical a posteriori discovery, is construed to be true in every possible world. Since the singular terms "Hesperus" and "Phosphorus" pick up the same object in the actual world, they pick up the same object in every metaphysical variation of it, no matter how we end up discover that. Thus ($Kant \Rightarrow$) fails.



But given ($Frege \Rightarrow$), then ($Carnap \Leftarrow$) must fail too. Thus



(K2) There are some *a priori* non necessary sentences: $\mathcal{A} p \wedge \neg \Box p$.

For instance, the wittgensteinian example "B is one meter long", where B is the platinum-iridim bar in Paris. Such a statement is a priori because it conventionally represent the definition of the length of one meter (it is an *initial baptism*), but is nonetheless contingent, since the legth of the bar might have been different. Thus ($Kant \Leftarrow$) fails.

But given ($Carnap \Rightarrow$), then ($Frege \Leftarrow$) must fail too. Thus

2.2 2D semantics at play

Now, 2D-semantics promises to reestablish the "golden triangle" of modality, reason and meaning, while acknowledging the rigid semantics of pieces of vocabulary. In fact, Chalmers maintains, 2D-semantics verifies a "Core Thesis":

(CT) S has a necessary 1-intension $\Leftrightarrow$ S is *a priori*

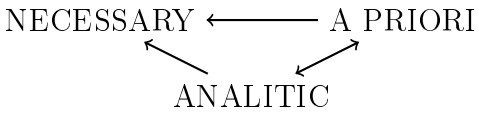
If (CT) is valid, kripkean counterexamples to the relations of the golden triangle have to be rejected.

Consider ($K2$) first. It says that there exists some sentence S such that S is a priori but S is not necessary. (CT) allows to claim that sentences Kripke had in mind, like "B is one meter long", are actually a posteriori (thus $\neg\mathcal{A}p$ after all). In fact "B" might have designated a different bar with a different length, i.e. our world could have turned out to be one in which "B" is not the platinum-iridium bar of Paris. Then "B is one meter long" has a contingent 1-intension, and, given (CT) it is not *a priori*.

However, while this is actually the argument Chalmers often seems to adopt in dealing with kripkean theses, it is not enough to prove $\neg(\mathcal{A}p \wedge \neg\Box p) \Longleftrightarrow \mathcal{A}p \rightarrow \Box p$. i.e. ($Kant \Leftarrow$) and the failure of ($K2$). In effect, I think, we really can do better, and establish ($Kant \Leftarrow$) directly because,

S has a necessary 1-intension $\Rightarrow$ S has a necessary 2-intension

In fact, if S doesn't contain rigid designators it can't be necessary, while if S exentially contains rigid designators like "B", the above is valid given rigidity of kripkean 2-intensions. Then



Notice that it can also be established that

S has a necessary 1-intension $\Rightarrow$ S has a necessary 2D-intension.

But consider now $(K1)$, that claims $\Box p \wedge \neg \mathcal{A}p$. We can try to apply the same reasoning and begin by asking how can "Hesperus is Phosphorus" not be necessary? There's obviously a sense in which it can't: given Kripke's account of 2-intensions, we know that $[\![\text{"Hesperus"}]\!]_2 = [\![\text{"Phosphorus"}]\!]_2 =$ Venus, thus S is necessary (*metaphysically* necessary). So we can't proceed as before.

Then, in order to reject $(K1)$ we must try to prove that "Hesperus is Phosphorus" is a priori after all. But this doesn't seem to be possible. Given 1-intension's epistemic variability, we also know how to represent the fact that S is not *a priori*, i.e. not *epistemically* necessary: there might be a scenario in which $[\![\text{"Hesperus"}]\!]_1 =$ Venus and $[\![\text{"Phosphorus"}]\!]_1$ = a satellite. We can't prevent this.

This is obviously because of the following

S has a necessary 2-intension $\not\Rightarrow$ S has a necessary 1-intension.

Does this lead the the attempt to reestablish $(Kant \Rightarrow)$ to a dead end?

In effect, *a fortiori*

S has a necessary 2-intension $\not\Rightarrow$ S has a necessary 2D-intension,

and, if necessary 2D-intensions are considered as representing analyticity, i.e. identity of meaning, then it seems that $(Carnap \Leftarrow)$ must also fail.

Is Chalmers completely mistaken?

More probably, something went wrong in our reconstruction of the argument, and it is indeed a crucial point. Let's look back for it.

How can "Hesperus is Phosphorus" not be necessary? There's obviously a sense in which it can't, and we saw it.

But there's indeed a sense in which it can. If "Hesperus" is *my* Hesperus, the object I learned to identify at celestial coordinates x, y at time t in the evening, and "Phosphorus" is *your* Phosphorus, the object you learned to identity at celestial coordinates x', y' at time t' in the morning, it may well be that they do not coincide in some possible world.

But how can that happen?

It happens because, as Chalmers pays a lot of efforts to explain, intensions apply to expression *tokens*, not to expression *types*.

Let me introduce the useful notation Robert Brandom developed to represent type-tokens relations: I'll use brackets to designate types, as in $< Phosphorus >$, and subscripted slashes to designate tokenings, as in$/Phosphorus/_i$.

Thus while

$$[\![< Hesperus >]\!]_2 = [\![< Phosphorus >]\!]_2$$

it may well be that

$$[\![/Hesperus/_i]\!]_2 \neq [\![/Phosphorus/_j]\!]_2.$$

Why? Because it may be that $[/Hesperus/_i]_1 \neq [/Phosphorus/_j]_1$, i.e. $/Hesperus/_i$ may be my tokening of $< Hesperus >$ identified in the *scenario* s_i where both Hesperus and Phosphorus refer to Venus, but $/Phosphorus/_j$ may be your tokening of $< Phosphorus >$ identified in the scenario s_j where Hesperus refers to Venus, but Phosphorus refers to a satellite.

In a sense, then, *epistemic necessity comes first.*

Notice that a different notion of necessity, 2D-intension necessity, has been introduced as an consequence of the new model adopted. And, it is obviously valid that

S has a necessary 2D-intension $\Rightarrow$ S has a necessary 1-intension,

which is what is needed to establish $\neg(\Box p \wedge \neg \mathcal{A} p)$.

Thus Kripke's attempt to sever the link between modality and reason is defused, and $(Kant \Rightarrow)$ reestablished together with $(Carnap \Leftarrow)$.

2.3 Good news and bad news

Time to take stock. Two dimensional semantics presents a matrix of intensional interpretations developed over the the two axes of epistemic possibility and metaphysical possibility. In order to have a smoother notation in what follows I'll use operator symbols to represent the relations at play here: "$\mathbb{A}p$" for p is analytic, $\mathcal{A}p$ for p is a priori and $\Box p$ for p is metaphysically necessary.

Variability over the metaphysical axis is represented by kripkean 2-intensions, which are rigid for non descriptive vocabulary. Necessary 2-intensions can be represented as $\Box p$ iff $\forall w$, $w \vDash p$.

Variability over the epistemic axis is to be represented by 1-intensions, which admit token variability, so that "Hesperus is Phosphorus" may be epistemically contingent. According to (CT) necessary 1-intensions represent a priori sentences, so that $\mathcal{A}p$ iff $\forall s$, $s \vDash p$.

While contingent a posteriori synthetic sentences have irregular matrices (any scenario combined with any possible world may identify a different extension), raws of the matrix are uniform for kripkean rigid designators, and columns are uniform for a priori sentences.

Analytic sentences are represented by necessary 2D-intensions, whose matrices are uniform: $\mathbb{A}p$ iff $\forall s, w$, $(s, w) \vDash p$.

Now, $\mathbb{A}p \rightarrow \mathcal{A}p$ and $\mathbb{A}p \rightarrow \Box p$. The point is that given (CT) and the behavior of 2-intensions, $\mathcal{A}p \rightarrow \mathbb{A}p$ is valid too.

Actually, (CT) may fail depending on how scenarios are defined: if some epistemic possibility is missed and doesn't fit within the set of scenarios, then necessary 1-intensions do not guarantee apriority anymore. But we may well concede a metaphysical adequacy to Chalmers's – or anyone else's – definition of scenarios that validates (CT), so that we can turn back to our initial puzzle about conceptual modality. There seem to be good news and bad news.

The good news is that 2D-semantics seem to provide a strong formal grasp on that link between the notions of apriority and analyticity Quine tried to discard: (CT) is an operative tool to part sentences into a priori and a posteriori ones according to their meanings. Dogmas of Empiricism are rehabilitated now as justified theses and we can trustfully go back to our scientific enterprises, with the guarantee we will be building meaningful and reliable analyses of the world we are in.

But there are also bad news, and they are really bad indeed.

Old cathegories seem to have been reestablished at the price of drastically tightening their admittance requirements. While we can be sure that necessary 1-intentions identify a priori sentences, these sort of sentences are now really hard to find. Well known old-fashioned candidates seem to fail in an embarassing regular way. We already rejected "Hesperus is Phosphorus", which had been downgraded to a contingent, synthetic a posteriori sentence, like "my laptop is grey". What about "Bachelors are unmarried men"? It's easy to realize that there are reason for it to be excluded as well, because there may well

be a scenario in which the expressions tokens it contains have different meaning, and the whole sentence means that moons are made of cheese. But things can go even worse. Let's consider a speaker, who has just learned that Hesperus is Phosphorus and thus has just changed his mind about the meaning of "Hesperus": he may well utter "Hesperus is Hesperus" as an a posteriori discovery – in the sense, "What I though to be Hesperus, an object different from Phosphours, is actually Hesperus, an object identical with Phosphorus" –, where the first occurrence of "Hesperus" is a token whose 1-intension picks up a satellite in a scenario, while the second occurrence of "Hesperus" is a token whose 1-intension picks up Phosphorus in another scenario.

In order to avoid these consequences Chalmers distinguishes two interpretations of 1-intension: a *contextual* understanding and an *epistemic* understanding.

In the contextual understanding, scenarios are considered as possible contexts of utterance. Thus in the contextual understanding 1-intension are helpless against the contingency of attribution of meanings to tokens. Since sentences like

> "My utterance of 'Hesperus' means Hesperus"

are a posteriori, every scenario will have its own 1-intension for the tokens of the orthographic type "Hesperus". There are no necessary 1-intensions at all and (CT) fails for the contextual understanding. Chalmers[4] considers some variations of the contextual definition of scenarios but the result of his analysis is that none of them is able to preserve *(CT)*. We can rely on his analysis here and move forward.

In the epistemic understanding scenarios are considered as epistemic possibilities, i.e. as possible worlds centered on a certain epistemic perspective[5]. Two ideas are at the core of this interpretation: first *plenitude*, the idea that epistemic possibilities are completely realized by possible worlds, in the sense that

> for all S, S is epistemically possible if and only if there is a scenario that verifies S[6];

and second *scrutability*, the idea that once we know enough about a scenario to determine its character – or, in other words, once we know which scenario is the actual one –, then we can determine the extensions of our expressions in it, thus

[4] *Ibid.*, pp.65-75

[5] Where the center is an ordered pair of an individual and a time reference. *Ibid.*, p.60

[6] *Ibid.* p.81.

> for any truth S there exists a truth D, s.t. D is independent of S and knowing that D is the case puts the speaker in a position to know (without further empirical information, on idealizes rational reflection) that S is the case[7].

The principles of plenitude and scrutability prescribe respectivley that scenarios display the whole space of epistemic possibility, and that any token-relative 1-intension is accessible and reidentifiable independently of scenarios. In order to obtain this result token-relativeness of 1-intensions has not to be related to expression tokens, whose meanings, as we saw, are determined opaquely within scenarios. Chalmers proposes to track it back to something like thought tokens[8]. Now, thoughts are objective enough to be reidentified through different epistemic scenarios and to make necessary 1-intensions available. In fact in every scenario one can recognize the identity of two tokens of the same thought even if such thought is part of the epistemic perspective at the center of another scenario. Thus Chalmers explicitly denies that "Hesperus is Hesperus" and "Bachelors are unmarried men" could be a posteriori in his epsitemic understanding. Each token's meaning can be tracked down and defined in terms of analyticity modality and apriority. "Hesperus is Hesperus" is a priori if both tokens are tokens of the same thought, i.e. if both occurrences of "Hesperus" are occurrences of the same linguistic unit.

This is what validates (CT) and prevents the triggering of Quine's indeterminacy of translation thesis: 2D-semantics has more than a strong grasp on meanings.

2.4 Scrutability and paradoxes

We can now appreciate the relevance of two dimensional semantics results for what concerns our topic.

As is was already said, it introduces a way to compose the gap between Quine's and Kripke's theses about modality. On the one side, Quine refuses any evaluation of possible particulars and thus acknowledges an irreducible indeterminacy of meanings. This characterizes inductively established generalizations as contingent in the humean sense of cognitively relevant regularities. On the other side Kripke's partition of language expressions in rigid designators and descriptive vocabulary allows the extensional representation of modality, but, within the framework of possible worlds, normativity can be characterized only in terms of a quantification over possible particulars. Thus, inductively established

[7] *Ibid.* p.90.
[8] *Ibid.*, pp.96-98.

generalizations may vindicate normative force only after being evaluated against every such possible particular (which is quite beyond reach for the actual linguistic and scientific practices). None of these paths then would lead to a substantial solution or explanation of the paradoxes threatening normativity of meanings, but some new hopes might be risen by the unified perspective 2D-semantics provides.

It's easy to see that the application of two dimensional framework to the analysis of normativity paradoxes like Kripke's and Goodman's mainly goes through the principle of *scrutability*. We've just seen why Chalmers establishes the so called scrutability of extensions with respect to epistemic scenarios, but it's worth coming back to it. The intuitive idea is that speakers can determine extensions for any given character the world may have:

> "If a subject possesses a concept and has unimpaired rational processes, then sufficient empirical information about the actual world puts a subject in a position to identify the concept's extension."[9]

From a formal point of view Chalmers's distinction between a metaphyscial and a descriptive conception of scenarios makes things a little more complicated, but we can skip over these difficulties here and go straight to Chalmers's choice of the descriptive option[10].

In the *descriptive* conception scenarios are construed as alternative canonical descriptions D, s.t. they are *complete* (for any S either $\models D \wedge S$ or $\models D \wedge \neg S$) and *neutral* (they contain just descriptive nonrigid vocabulary). In this sense *plenitude* amounts to the idea that

> for every hypothesis S there is a canonical description D such that
>
> $\models D \rightarrow S$.

In the descriptive conception the principle of *scrutability* is more clearly formulable:

> if D is a complete qualitative description of the world then for all S,

[9] Chalmers and Jackson [36], p.8.

[10] In the *metaphysical* conception scenarios are really the same sort of things as possible worlds, with the only difference that scenarios have a center, i.e. the couple of a speaker and a time reference for the utterance. In this metaphysical interpretation the principle of *plenitude* can be formalized as

> If S is epistemically possible then $\exists s, s \models S$.

More importantly, in this metaphysical interpretation the principle of *scrutability* prescribes that for any truth T there must be a witness scenario s_T, s.t. for any scenario s, if s is the witness scenario, then s verifies T. It's not obvious how this can be formalized, and Chalmers doesn't suggest it.

$$\models S \Rightarrow \models \mathcal{A}(D \rightarrow S).$$[11]

Notice, and this is the main point, that together these imply a form of *epistemic necessitation*, i.e.

for every hypothesis S there is a canonical description D s.t.

$$\models D \rightarrow S \Rightarrow \mathcal{A}(D \rightarrow S).$$

Now, let's introduce Goodman's paradox and consider

S: x looks green $\rightarrow$ x is green

and

S': x looks grue $\rightarrow$ x is grue.

Suppose we take for granted *plenitude* (in any of its forms) so that, since both S and S' are to be epistemically possible but together incompatible, then $\models D \rightarrow S$ and $\models D' \rightarrow S'$, while obviously $\neg\exists D, \models D \rightarrow (S \wedge S')$.

Now, *plenitude* and *scrutability* of truth say that there is a way to tell if it is the case that $\models S$ or $\models S'$ (or none): check if $\models D$ or $\models D'$.

Epistemic necessitation would then guarantee for the normativity of such a meaning relation.

This really looks like the solution of the puzzle.

Let me briefly recall what was on the table.

Is meaning normative? Are inductively established generalizations that define inferential relations between concepts modally robust or they simply reduce to contingent universal quantification over psychologically relevant domains? Carnap's original modal analysis of intensions provided an explanation of meaning statements in terms of possible worlds, but cognitive variability of extensions still represented an obstacle for semantics. Kripke, by separating descriptive variable and non descriptive rigid vocabulary, established stable foundations for the semantics of modal logic. However he – and the tradition flourished from his works – had no answer to the epistemic scruples about possible worlds framework and conceptual normativity. Quine, for example, refuted any evaluation of possible particulars and preferred to deny any modally robust relation among meanings. In this setting Goodman's paradox can be construed as denouncing the inability of these two perspectives

[11] Adapted from Chalmers [34], p.175. But see also Chalmers [35], p.91.

to account for the normative character of lawlike statements and meaning statements in general: possible particulars are not scrutable, and the scrutiny of actual particulars is not enough to part green-worlds from grue-worlds.

Two dimensional semantics proved to have a two steps solution: first, green-worlds and grue-worlds are univocally characterizable within the structure of epistemic scenarios by approriate canonical descriptions (*characters* are extensionally graspable); second, canonical descriptions necessarily imply every truth of the epistemic scenarios they identify (*contents* are scrutable given *characters*).

Thus, intuitively, how can it be established if emeralds are green or grue?

First, the *plenitude principle* guarantees that there exists a scenario canonically described by some D in which "x looks green $\rightarrow$ x is green" is true, and that there exists a scenario canonically described by some D' in which "x looks grue $\rightarrow$ x is grue" is true. Second, the principle of scrutability of truth guarantees that both the relation between D and "x looks green $\rightarrow$ x is green" and the relation between D' and "x looks grue $\rightarrow$ x is grue" are epistemically robust, i.e. *a priori*. What we have to do then is to check D and D' against actual particulars, in order to establish if the actual world is a D-world or a D'-world: that would adjudicate if the actual world is a green-world or a grue-world.

Notice that practically this may be an infinite task: we, as human, may not have enough time or resources to decide between D and D', so we might have to accept the probabilistic character of inductions. But this fact wouldn't jeopardize the normative character of the inductively established statements, as Goodman and Sellars clearly explain and Chalmers and Jackson conveniently repeat:

> "It might also be objected that no human could grasp [...] the truth of the relevant conditional. This is surely true, but it is no bar to the apriority of the conditional."[12]

2.5 Taking a step back

It's worth taking a step back to look at the picture 2D-semantics describe.

Indeed the "golden triangle" has been regained: the restored kantian thesis relates apriority and necessity, the fregean thesis relates meaning and apriority, thus the Carnapian thesis relates meaning and necessity.

[12] Chalmers and Jackson [36], p.18.

The matrix of 2D-intensions for analytic expressions is reassuringly uniform: Hesperus, for instance, is epistemically and metaphysically Hesperus in any scenario and possible world.

And yet, I think, there are reasons to be relevantly unhappy with this picture.

The main problems are: what sort of meanings are identified by these matrices? What is to be meaningful?

To begin with, notice again that

"Hesperus is Phosphorus"

may be contingent, as well as

"Hesperus is Hesperus",

and for the same reasons: it might be that

$$[\![/Hesperus/_i]\!]_1 \neq [\![/Hesperus/_j]\!]_1.$$

While the 2D-intensional matrices of $/Hesperus/_i$ and of $/Hesperus/_j$ are uniform, there seem to be not obvious way to relate them one another.

Indeterminacy of meaning here shows threatening metaphysical appearances.

Hopefully there can be found something easier to defuse.

In which sense $/Hesperus/_i$ and $/Hesperus/_j$ can be considered tokens of the same type?

First of all, let's improve our type-token notation and use indexed tokens to relate them to their type: thus $/a/_i^{<b>}$ indicates that the token $/a/_i$ is a tokening of type $< b >$.

Now, it's hard to see any grammatical sense in which two tokens with different meanings could be considered tokens of the same linguistic type, since grammatical rules are explicitly designed to avoid ambiguities and to represent different linguistic types as different grammatical types. Recall that Quine's argument against intensions hinges on the idea that no matter how hard we try, we won't have grammatic rules devoid of ambiguities. Here the point is that meanings can be independently distinguished, but then any grammar would have them represented as different types. For instance, in English, "bank" may mean both a financial institution, say $< bank_1 >$, and the sloping land near a river, say $< bank_2 >$. Thus we have occurences of "bank" as $/bank/_i^{<bank_1>}$ and as $/bank/_j^{<bank_2>}$. In the same way a linguistic analysis of this 2D-intensional picture would distinguish $/Hesperus/_i^{<Hesperus>}$ from $/Hesperus/_j^{<satellite>}$.

In the contextual interpretation of 1-intensions, it is suggested that orthography may tell which tokens are of the same type. But it is not clear how this could help. If orthographical-types have to be considered (and orthography has to make any sense), then "Hesperus is Hesperus" contains two tokens of the same type.

On the one side, if types have some role in the determination and communication of meanings, then "Hesperus is Hesperus" is necessary and, given Kripke's metaphysical rigidity of singular terms 2-intensions, "Hesperus is Phosphorus" is necessary too. Then, as we already saw, ($Kant \Rightarrow$) fails and the "golden triangle" is left uncomplete.

On the other side, if types have no relations with meanings, why talk about types at all? Everyone ends up with his own idiolect, which may happen to be congruent to the others'. What could prevent us to suppose that every tokening of "Hesperus" may have its own meaning even if tokened by the same person? Chalmers[13] responsibly stipulates that intensions don't really apply directly to tokens of expressions, but to tokens of thoughts, which are expressed by them and which are guardanteed to be intrinsically the entertaining of a content. But what could prevent an irresponsible skeptic to suppose we are schizophrenic enough to entertain different contents from one time to another?

Let's don't go thus far. Obviously the first moral of the above remarks is that one can't talk about tokens without talking about types: for something to be a token, it must be the token of some type (possibly the only token of its type). And the second moral is that for tokens to be meaningful they must be repeatable as type-tokenings.

In this sense, as Brandom points out[14], statements like

"Hesperus is Phosphorus"

can't but relate types, not tokens, of expressions. That is to say, they relate meanings, and this is a quite crucial point.

Chalmers remarks[15] that the determination of the actual meaning of an expression is not a real deal for two dimensional semantics: what really matters is to track analytic, rational and modal relations among contents. I think this is a sort of misunderstanding of what a definition of intensions should provide. Linguistic expressions are used with functional regularities and this is what makes of any $/Hesperus/_i$ and $/Phosphorus/_j$ tokenings of functional types, or meanings. But these can't be accounted merely in terms of isolated occurrences.

[13] Chalmers [35], p. 96.
[14] Brandom [14], pp.450-455.
[15] Chalmers [35], p. 104.

The whole metaphysical apparatus of possible worlds can be construed as purported to account for the features of these functional regularities. The so called "direct reference theory", which stipulates the semantical rigidity of several pieces of vocabulary, and the so called "causal reference theory", which stipulates how these rigid semantical relation have to be inherited within communicative practices, do not represent any metaphysically strong thesis, as Lewis provocatively put it. In other words, they don't stipulate anything at all, they rather highlight the linguistic behavior of those pieces of vocabulary: singular and thing-kind terms are used to anaphorically single out referents. They are the means by which languages assume representational dimension, and their modal invariance is a consequence of anaphoric mechanisms[16].

Now, Chalmers writes:

> "There is no clear analog of a *de re* modal intuition in the epistemic case. [...] [D]ifferent names for an individual are not generally a priori equivalent, so come apart in different scenarios, and there is no way in general to isolate a privileged class of epistemically equivalent designators here."[17]

In a sense, this is obviously correct: where pieces of vocabulary maintain their semantic interpretation in every possible world, any expression's interpretation varies along espistemic scenarios. Yet, this doesn't imply the irreducible epistemic variability of meanings. Where 2D-semantics tries to pin down isolated units of semantic analyis, there it can be recognized the tokening of a linguistic functional role, a meaning, that, while located within an epistemic perspective, can be anaphorically picked up and referred to in any other variated context (epistemic scenario or possile world). Two dimensional semantics seems to blur the distinction between the intensional variability proper of descriptive vocabulary, i.e. $[\![\text{"Hesperus"}]\!]_{w_i}$ = Hesperus and $[\![\text{"Hesperus"}]\!]_{w_j}$ = satellite, and token variability of non descriptive vocabulary, i.e. $/Hesperus/_i^{<Hesperus>}$ and $/Phosphorus/_i^{<Hesperus>}$: the first sort of variability doesn't allow token repeatability, while the second is embedded into anaphoric relations. There are indeed deep reasons behind this behavior of some pieces of vocabulary that possibly can be tracked back to the expressive resources of the language[18].

Kripke's theory of direct reference is often construed simply as an hypothesis that entitles him to distinguish between the modal metaphysical level and the epistemic one, but it is actually a semantic feature of that part of vocabulary that is charged with the

[16] Here I'm obviously endorsing Brandom's account of anaphora. See Brandom [14], ch.6-7.
[17] Chalmers [35], p. 102.
[18] Again, see Brandom [14], ch.6.

task of referring to objects in any descriptively variated situation. In fact, rather than being problematic, it makes no sense at all to ask if two tokens may be the same token, in the sense:

$$^{?}/Hesperus/_i = /Hesperus/_j.$$

This, however, is what would be crucially required for the whole apparatus of 2D-semantics to be meaningful. In fact, on the one side the very possibility to identify a priori sentences (and thus analytic and necessary sentences) hinges on the possibility for linguistic units to have the same 1-intension, on the other side if 1-intensions were type-relative they couldn't sidestep semantic rigidity of some pieces of vocabulary. But if 1-intensions must be token-relative, then for two expression tokens to have the same 1-intension they must be the same linguistic unit in some epistemically rigid sense. We saw above that this framework of token-relative 1-intensions doesn't provide any obvious reason to consider sentences like "Hesperus is Hesperus" or "Bachelors are unmarried men" as a priori.

Now, the principle of *scrutability* makes apriority available, but an analysis of scrutability easily suggests what implicitly lies behind his result. Scrutability forces tokens to correspond 1 to 1 to 1-intensions in any given scenario: if any token could have two different 1-intensions within the same scenario then speakers couldn't determine extensions. Token-relative 1-intensions vary from one epistemic scenario to another, but, while any scenario is considered as the actual one, token-relative 1-intensions are fixed to that scenario. This is the crucial advantage of epistemic understanding over the contextual understanding of scenarios.

For instance, while "Hesperus is Hesperus" may be a posteriori because the two occurrences of "Hesperus" may have to be interpreted into two different epistemic scenarios, nonetheless, given that *actually* "Hesperus is Hesperus", i.e. that both tokens are fixed to one and the same scenario, then "Hesperus is Hesperus" is *a priori*.

Notice that this is exactly the sort of rigidity required by Kripke's metaphysical modality to reidentify linguistic units while shuffling possible worlds. The same sort of rigidity is established by the principle of scrutability in order to make two dimensional semantics work. Notice as well that typical situations in which the epistemic scenario is guaranteed to be fixed, the sort of cases Chalmers usually presents as a priori sentences[19], are those same situation Kripke called "initial baptizms", where a priori coincide with metaphysical necessity, and referents can be picked up descriptively.

[19]See, for instance, Chalmers [35], p.95.

Thus, echoing Brandom again, we may say that the possibility to "navigate through doxastic perspectives" does not preclude, but actually presupposes the possibility to construct anaphoric relations among tokens of the same type.

To the contrary of what Chalmers maintains[20], and not surprisingly, the topic at issue here is really akin to what hides behind the problem of quantification inside modal or epistemic context. In fact, Quine uses his remarks against *de re* modality to argue for his thesis of the indeterminacy of translation. Brandom [14] proposes a neat answer to Quine's criticism, developed upon his analysis of anaphoric relations. These relations make sense of the identity claims that may happen to fail in some contexts where substitution of coreferential expressions seems to be prohibited. Rather than the diagnosis of meaning opaqueness, those contexts make explicit the perspectival nature of linguistic practices, i.e. epistemic variability of meaning. What substitution failures show is that different speakers may disagree on the identification of the anaphoric relation governing some tokenings, for example if an occurrence of "Ortcutt" is either an instance of $/Ortcutt/_i^{<the\ man\ with\ the\ brown\ hat>}$ or an instance of $/Ortcutt/_i^{<the\ man\ at\ the\ beach>}$, where $<the\ spy>$ and $<the\ man\ at\ the\ beach>$ are considered to be different types. Speakers can, nonetheless, pick up those types from others perspectives, and say something like

> Ralph believes that $/Ortcutt/_i^{<the\ man\ with\ the\ brown\ hat>}$ is a spy

and

> Ralph believes that $/Ortcutt/_i^{<the\ man\ at\ the\ beach>}$ is not a spy,

and even, maybe,

> Ralph believes that $/Ortcutt/_i^{<the\ man\ at\ the\ beach>}$ is not $/Ortcutt/_i^{<the\ man\ with\ the\ brown\ hat>}$,

where the speaker himself believe, while Ralph doesn't, that

> $<the\ man\ at\ the\ beach>=<the\ man\ with\ the\ brown\ hat>$[21].

[20] *Ibid.*, p.99 n15.

[21] As a consequence of this analysis, Brandom construes the difference between *de re* and *de dicto* beliefs as a difference in propositional ascriptions: in *de re* ascriptions speakers do not share ascripted beliefs about anaphoric chains. He also proposes a language regimentation to distinguish between the two sorts of ascriptions. While a *de dicto* ascription is to be expressed as

> Ralph believes *that* Ortcutt is a spy,

a *de re* ascription should be made explicit as

> Ralph believes *of* Ortcutt *that* he is a spy.

In other words what is at issue is the identity between two types of expressions that can be anaphorically tracked back to two different ways to identify an object (Ortcutt in this case), i.e. two different 1-intensions. Thus, notice also that exactly because of this reason, one *can't* say

$/Ortcutt/_i^{<the\ man\ at\ the\ beach>}$ is not
$/Ortcutt/_j^{<the\ man\ with\ the\ brown\ hat>}$

or

Ralph believes that $/Ortcutt/_i^{<the\ man\ at\ the\ beach>}$ is not
$/Ortcutt/_j^{<the\ man\ at\ the\ beach>}$.

To sum up, 2D-semantics clearly strikes a sound point: Kripke's rigid semantics for non descriptive vocabulary sacrifies differences of "cognitive value" which are instead an indispensable component of meanings. However, in order to rehabilitate this component, it enters the old well trodden path of truthfunctional analysis that leads to a duplication of Kripke's solution of an extensional grasp on intensions. This probably comes together with the inheritance of Kaplan, who actually aimed to a logic of indexicals by revealing their truthfunctional contribution to sentences. Unfortunately this attempt to project the structure of formal logic over contextual variability of meanings blurs good and bad points of kripkean analysis of modality. There are indeed two sorts of behaviour linguistic expressions may display with respect to reference: they may be contextually (in any broad, non technical sense) rigid, or they may vary. Quine's skepticism spun on this point till it crumbled down the very notion of meaning. Now, since extensional logic is truthfunctional, things have to be one accomodated so that these referential behaviors may be described by interpretation functions from a domain of lingusitic expressions to a codomain of objects. The meinongian option to adapt referents to reference variability has obviously never showed any real appeal, because it could not make real sense of the codomain for those functions. Kripke's powerful solution was to distinguish different roles within the vocabulary. Two dimensional semantics represent an attempt to reject this distinction while keeping truthfunctional analysis of meanings, but unfortunately this messes up the domain of interpretation functions: what is a linguistic unit of meaning? Brandom's analysis suggests another perspective: anaphoric relations among linguistic expressions, that Kripke described in terms of causal chains, account both for contextual variations and for rigidity of reference, because they show how to keep track of the referential purport expressions

have in the different contexts, epistemic or metaphysics, in which they are used[22].

Let me try to further clarify this point by considering other examples.

Chalmers admits[23] he finds unclear why Stalnaker, given his definition of an utterance's *diagonal proposition*, may maintain that a sentence like

"this bar is one meter long"

could not be false. If the proposition an utterance expresses varies in different possible worlds, why not conceive a world in which the expression "this bar is one meter long" may mean that the moon is made of cheese? Even without evaluating Stalnaker's proposal and its actual results, one can easily see what he had in mind. The reason why "this bar is one meter long" couldn't be false is that such an utterance represents the baptizm, i.e. the act of reference fixing, of the functional type $< one\ meter >$. Thus it can be analyzed as the act of using tokenings of $/length\ of\ this\ bar/_i^{<length\ of\ this\ bar>}$ and $/one\ meter/_1^{<one\ meter>}$ (notice that this would be the first tokening of the new type) in order to establish the equivalence

$$< length\ of\ this\ bar >=< one\ meter >,$$

so that tokens of "one meter" can be used as part of the anaphoric chain $< length\ of\ this\ bar >$, i.e. $/one\ meter/_i^{<length\ of\ this\ bar>}$. Anaphoric relations among tokens, once established, do not vary, so to say, at will: they do not depend on speakers' epistemic states, so they remain fixed in any model representation of them. Let me quote Brandom's remark about this:

> "There are two varieties of substitutional equivalence. There are *intra*term and *inter*term, or de jure and de facto equivalences of tokenings. The former are (taken to be) binding on all interlocutors; the latter vary from doxastic repertoire to doxastic repertoire, according to the particular substitutional commitments undertaken by or attributed to an individual. [...] Substitutional structure requires both sorts. They cannot be defined separately, apart from their role in such a structure; one cannot have the one sort of equivalence without the other. What intraterm equivalences are *for* is to be vehicles of interterm substitutional commitments. These latter in turn presuppose them, in that they could not otherwise have content."[24]

[22] Does this mean we have to change truthfunctionality for an analysis of linguistic behavior? The answer is no, but we have to renounce to monotonicity as we will see.

[23] Chalmers [35], p. 113.

[24] Brandom [14], p.452.

Indeed the main difference between 2D-semantics and other previous bidimensional approaches seem to be the abandoning of the idea that the referent is part of the content of singular terms, i.e. that singular terms constitute a *piece* of vocabulary whose function is to single out referents in a non descriptive way.

At last let's consider again

"Hesperus is Phosphorus".

Why it is to be considered as *a posteriori*? Because it is not an intrinsic feature of the types $< Hesperus >$ and $< Phosphorus >$ to be functionally equivalent, in the sense that

$$< Hesperus >=< Phosphorus >=< Venus >,$$

Once established, this implies that for any i, $/Hesperus/_i^{<Venus>}$ and for any j, $/Phosphorus/_j^{<Venus>}$: since these are instances of the same anaphoric chain, which represent the functional role of the linguistic type $< Venus >$, they keep this semantic role in any framework, even in counterfactual ones.

Why could it be not necessary? The reason is the same. Where the above equivalence of types is not established, what may occur in counterfactual situation are tokens of different types which are not part of the same anaphoric chain, i.e. $/Hesperus/_i^{<Hesperus>}$ and $/Phosphorus/_i^{<Phosphorus>}$.

2.6 Conclusions

There seems to be no reasonable way to construct a proof that

"Hesperus is Phosphorus",

if meaningful, may not be necessary. Again, as Brandom puts it, those statements represent type-identities of token repeatables which can be formally represented as

$$< Hesperus >=< Phosphorus >.$$

Putnam makes a similar point when he notices that

"water is not water"

is not a contradiction when the two instances of "water" are to be construed as two different words, in the sense that

$/water/_i^{<H_2O>}$ is not $/water/_j^{<XYZ>}$,

since

$< H_2O > \neq < XYZ >$,

i.e. Earth-water is not TwinEarth-water.

In order to redeem "conceptual content" from metaphysical necessity, two dimensional semantics has to reject the modal rigidification of these identities. However, some sort of epistemic rigidification has to be reintroduced in order to make the whole apparatus work. The result is that the new bidimensional perspective is compressed into the old kripkean framework with its distinction between semantically rigid and descriptive pieces of vocabulary.

Thus Kripke's criticism to ($Kant \Rightarrow$) has to be accepted and, as a consequence, the idea of the "golden triangle" has to be abandoned.

But there's something worse. In fact, while 2D-semantics may survive the failure to reestablish the "golden triangle", it can't stand without both *plenitude* and *scrutability*. And I think the above remarks may prove that either *scrutability* is false or the structure of epistemic scenarios doesn't represent every conceptual possibility, i.e. *plenitude* is false.

We saw that there are no expression tokens without expression types. Thus 1-intensions can't be token-relative without being type-relative.

Now, suppose each token of an expression could represent the only single instantiation of its type, then there wouldn't be any a priori sentence at all and the principle of *scrutability* would be false: there could be no way to reidentify 1-intensions through perspectives (notice that while CT would not be false, it would be just vacously true).

Alternatively, suppose the principle of *scrutability* is accepted and 1-intensions are required to be reidentifiable among scenarios. Thus, for instance, while

> "Neptune is the object that perturbs the orbit of Uranus"

may be true only in Leverrier's scenario,

> "if this is Leverrier's scenario, then Neptune is the object that perturbs the orbit of Uranus"

must be true in every scenario.

But consider proper names (and in general Kripke's rigid designators): while

"$/Hesperus/_i^{<Hesperus>}$ is $/Hesperus/_j^{<Hesperus>}$"

may be true only in my scenario,

"if this is my scenario, then $/Hesperus/_k^{<Hesperus>}$ is $/Hesperus/_l^{<Hesperus>}$"

must be true in every scenario.

Yet there is a sense in which *in my scenario*

"$/Hesperus/_i^{<Hesperus>}$ is $/Hesperus/_j^{<Phosphorus>}$"

might have been true. Thus

"if this is my scenario, then $/Hesperus/_k^{<Hesperus>}$ is $/Hesperus/_l^{<Hesperus>}$"

is not *a priori*.

Then, while *scrutability* is valid, *plenitude* fails and CT too.

Let me explain. What is crucial to notice above are the indexes i and j, and their occurrence together with the types $< Hesperus >$ and $< Phosphorus >$, to which each indexed token belongs. In the latter situation, tokens changed their type membership, or, in other words, the anaphoric chain they belong to. This is what the trick of the indexes represents. What happened is that while 1-intensions remained scrutable in the sense that every 1-intension points to a different extension in every scenario (thus, while *scrutability* remained valid), they had changed their relation to the token.

In other words, while anaphoric relations among tokens of the same type are rigid, i.e. remain stable across different metaphysical or epistemic perspectives, the membership of each token to any anaphoric chain is not rigid in any sense: this is the special condition of initial baptizms, where contingent unrepeatable tokenings are used to give origin to rigid anaphoric chains.

Again, to sum up, two dimensional semantics confuses the intensional variability of conceptual content with tokens variability of anaphoric chains.

Along this discussion we could establish a really promising conditional: if two dimensional semantics is correct, it provides a substantial solution to the problems of conceptual modality and normativity of meaning. Unfortunately its premise is wrong.

Chapter 3

Brandom's inferentialism

3.1 The inferential way to norms

In Chapter 1 a promissory note was signed and it's now time for it to be exstinguished. The analysis of the normativity of meaning highlighted some deep demands that standard extensional semantics can't satisfy. And the reasons for this incapacity were traced back to certain blindspots in the underlying framework that supports the deployment of those semantic tools. The first demand was to vindicate the intrinsical modal character of conceptual content against its reduction to the description of states of affairs. The second, consequent demand was to provide a sensible account for this modal character, alternative to the kripkean reduction to description of *possible* states of affairs. We seemed to find a promising practicable way to deal with this second demand in Lance and Hawthorne's idea of a revisable notion of analyticity, but in order to introduce that we had to take for granted the inferentialist semantics of Robert Brandom's analysis of language and his account of linguistic practices as instituting norms. In doing this we implicitly suggested that such an approach to semantics could also cope with the first demand. Thus, it's worth focusing on it.

Let me first try to set the stage. The term "inferentialism" is Brandom's own coinage for the thesis that conceptual content is to be inferentially *demarked*, in the sense that to be meaningful is to obey to inferential rules[1]. Brandom traces back its roots, through

[1]Thus generally put, the idea is not a newcomer on the stage of philosophical analysis of semantics. Similar views had already been discussed under the labels of "Inferential Role Semantics", Boghossian [10], or "Conceptual Role Semantics", Harman [68], Block [8]. For an overview see Greenberg and Harman [64]. What is radically different is the normative, as opposed to causal, interpretation of these inferential conceptual roles.

Dummett's influential work on Frege's philosophy of language[2], to a famous passage in the *Begriffsschrift*:

> «The contents of two judgments can differ in two ways: first, it may be the way that the consequences which can be derived from the first judgment combined with certain others can always be derived also from the second judgment combined with the same others; secondly, this may not be the case. The two propositions "At Plataea, the Greeks defeated the Persians" and "At Plataea, the Persians were defeated by the Greeks" differ in the first way. Even if one can perceive the slight difference in sense, the agreement still predominates. Now I call the part of the contents which is the same in both, the conceptual content.»[3]

This obviously clashes with the standard representationalist approach, which nonetheless sprang out of Frege's semantical analysis. If we want to proceed on this path we have to smooth out the clash and ground the inferential approach by providing a model for inferential rules.

Before we move on into all this, I have to declare here a due premise, motivated by the systematic structure of Brandom's thinking. In what follows I won't try to present Brandom's system with the strictness such an enterprise would deserve. And the reason why I discard such an enterprise – concerns about appropriateness and space apart – is because, I believe, there is a certain *proper* path one has to follow in approaching his system: it is the path that moves from the introduction of a normative pragmatics and an inferential semantics and ends in the definition of a social community of scorekeepers based on mutual recognition, it is the path that Brandom himself presents in his Brandom [14] and reuses in Brandom [15] when he needed to sum up his work in a more introductory form, it's the path that is worth being followed in all its length by any philosopher who is willing to understand Brandom. But it's not our path here. Our already declarated ending point will be just the definition of an inferential semantics. In order to provide such a definition we will have, first, to lay down some pragmatic basis as an account of the normative character of social linguistic practices, and, second, to show in which sense such basis can support our semantic interpretation. In the following sections, the reader will be forced to focus on this scenario and put into brackets any of her maybe legitimate perplexities about Brandom's theses – which I will mostly discard the responsibility to

[2]Dummett [41].
[3]Frege [50], § 3.

clarify here by referring to Brandom's own works. The reward for this effort will be, I hope, a better comprehension of the proposed semantic interpretation, indipendently, in a sense, from Brandom's own views.

3.2 We are creatures who say 'we'

Since this is not to be considered an extensive and complete exposition of Brandom's work, some preliminaries are especially in order.

We are going to introduce an inferential semantics, that is an account of linguistic contents in terms of inferential relations. Now, there's one main question that should bother those who encounter this sort of thing for the very first time: if it is true, as it is, that linguistic contents deal with how things are in the world, what do inferences have to do with that? Once the normative character of meaning is acknowledged however, this question actually comes in two parts. The first part asks how the normative character of the meaning of subsentential expressions can be accounted in terms of the normative character of the meaning of sentences (since inferences relate sentences). The second part asks how this normative character is made determinate (which, notice, is not equivalent to ask how it is defined, nor how it is specified). The present section is purported to answer this second part. I'll postpone the answer to the first part till section 3.4.

When Brandom wants to explain how the inferential approach can account for the representational purport of language, he often describes his path as "a social route from reasoning to representing"[4]. We need to catch a glimpse of this social character before we take such a route. In what follows I'll exploit Sellars's inferential account of observational knowledge in order to briefly introduce Brandom's analysis of normative stances. Three main theses will emerge as the focal points of this sketchy picture: a) we are rational beings in that we deploy concepts, b) any conceptual representation has normative character, c) the determination of such normative character is a social enterprise.

3.2.1 Sentience and sapience

In his *Essay Concerning Human Understanding*, at the very beginning of the third book, the book about language, John Locke wrote:

> «[Man can] make articulate sounds, which we call "words". But this was

[4]See Brandom [14, Chapter 8], and Brandom [15, Chapter 5].

> not enough to produce language, for parrots and some other birds can learn to make distinct enough articulate sounds, yet they are far from being capable of language.»[5]

This is a good place to start. Thus, to begin with, humans and animals may have similar organs that may produce similar phonetic performances. In fact, for instance, both a human and a parrot may be taught to regularly react to red stimulations with the utterance of the phonic design corresponding to "that's red". But, notice, this is a capacity animals and humans may share also with certain inanimate objects, e.g. a fancy measuring instrument made of a spectrometer connected with a speaker, or, more crudely, a chunk of iron that regularly reacts to the different degree of moisture in its environment by rusting. Indeed there's something more that humans and animals have in common and that distinguishes them from inanimate objects: an "exclusively biological phenomenon"[6] we may call *sentience*. *Sentience* is sensual awareness, the capacity for instance to *see* colors, *hear* sounds and *feel* pain. Now, this clumsy definition will rightly displease many cognitive scientists, but, for our purposes here, it's clear enough to qualify one of the characters of an old familiar play, that has been on philosophy's playbill at least since Plato and whose setting became broadly linguistic under Locke's direction. The other character of this play is *sapience*, the specifically rational capacity that distinguishes human beings from animals, or, we should say now, linguistic animals from non linguistic animals. As in any ancient tragedy, the plot is set into motion by the conflict between the absolute perspectives of the two protagonists. What does make the difference between mere *sentience* and *sapience*? What does make the difference between animal languagings and human linguistic performances?

Locke's answer is that

> «Brutes abstract not. [...] [T]he having of general ideas is that which puts a perfect distinction betwixt man and brutes.»[7]

Humans, as opposed to animals, are capable of *abstraction* and thus have general ideas which words are used to symbolize. In the absence of general ideas nothing can be communicated since particular ideas only represent particular objects whose experience might not be shared and thus don't support knowledge.

[5] Locke [88], 3.1.1.
[6] Brandom [15], p. 157.
[7] Locke [88], 2.11.10.

This was a quite innovative direction. Indeed, it sows the seeds of the empiricist tradition that will flourish till Rudolph Carnap, and some of whose outgrowths are still rotting. Let me ungratefully focus on these latter. If what distinguishes *sapience* from *sentience* is 'just' the capacity to master general terms, then we can venture to account for our rational faculties by providing an analysis of our mastering of these general terms. And our modern logical tools are powerful enough to make us quite confident that such an account could be actually provided. That was roughly but certainly the project of Carnap's *logische Aufbau*: to show how general terms and their use in scientific descriptions of the world could be logically constructed over particular terms. According to this conception, we are sapient creatures in that we can produce general knowledge about facts: we can represent those general features of things that allow us to classify them in sorts of *repeatables*, we can use general words to say how repeatables are *determinable*. Truthfunctional semantics suitably embodies this idea in terms of a formal representation of truth conditions: to understand what a sentence means is to understand how the world must look like in order for the sentence to be true.

And yet, if what distinguishes *sapience* from *sentience* is 'just' the capacity to master general terms, i.e. the capacity to have knowledge of determin*able* repeatables, then, in return, the capacity to have knowledge of determin*ate* repeatables must be somehow already given with *sentience*: we have, as merely biological creatures, an unaquired capacity to be aware of determinate repeatables, i.e. a capacity which has to be analyzed not in logical but in biological terms.

This is but one of the forms of the *Myth of the Given*[8], we already encountered in Section 1.2.2 in its more epistemological appearances. In order to unveil the myth it's worth focusing on a certain perplexity that is inevitably raised within the above conception: given that we have this capacity to be aware of determinate repeatables – e.g. that there is a red object over there – how can we be sure that what looks like a red object over there is really a red object over there? The answer, we are told by Descartes for instance, is that we just don't know: since our biological devices might fail, we can't even be sure that there is really something over there. We can only be sure that *it looks like* there is a red object over there, since we can't be mistaken about sense-datas themselves (there's no *looking as if it looked like* there is a red object other there). But this, Sellars denounces, is just the *Myth*[9]. The point is not the biological one that there are not *inner episodes* we can refer to as "impressions" or "sensations", but the logical one that these inner episodes can't

[8]See Sellars [147], §§26-29.
[9]See Sellars [147], §§10-23.

embody any basic virgin content uncontaminated by logical relations. In order for these episodes to have any content ϕ – in order for them to be impressions of ϕ, or ϕ-impressions, for instance – one must have reasons to deploy them in a story about ϕ-things. Thus, for instance, in order to deploy the general term "red", one must acknowledge something like

x *is* red $\equiv$ x would *look* red in *standard* conditions.

Notice, this is not to define the inferential use of "x is red" in terms of the reporting use of "x looks red", but to acknowledge certain logical properties of color words like "red", namely that there are certain conditions in which things *look* what they *are* and conditions in which they do not. When conditions are not standard, as in Sellars's parable of the tieshop, speakers deploy *looks*-talk not to report a more basic epistemic fact, but to withhold the endorsement from the report itself, since they don't know how to justify it, since they have no reason for it. This is to say that in order to deploy *looks*-talk one must already be able to deploy *is*-talk, since one must be able to know how to take an endorsement before he can know how to withhold one. Thus the above equivalence turns out to be a definition of these *standard* conditions as conditions in which content *should* be endorsed.

It can be concluded that the distinction between general knowledge, typical of sapient creatures, and particular knowledge, typical of sentient creatures, makes no sense, since there is no knowledge at all without justification. Sapient creatures express knowledge with their claims because they have reason for them: contents of knowledge dwell in the space of reasons, and we are sapient beings in that we handle reasons by justifying our claims.

This is a good starting point to talk about meaning.

3.2.2 A two-ply account of observation

So we took Brandom's own sellarsian path and it led us to the space of reasons. But before settling-down it's worth taking a look around to better understand where we are, since it's not Sellars's but Brandom's territory we want to inhabit. Let's try to get familiar with the enironment and, like tourists, not to feel ashamed to ask some naïve questions. For instance, why one should withhold the endorsement of a claim whatsoever in case she doesn't know how to justify it?

As it often happens, commonsense is akin to good philosophy and the answer to this apparently silly perplexity actually involves some important themes. So let's resume our

analysis. To endorse a claim is, in a sense, to believe it is worthy of credence. In this sense, Sellars says, claims have *authority* and this authority is what makes the difference between mere belief and knowledge. Now, since there is no knowledge without justification, there is no non-inferential knowledge, and thus it can be concluded that claims acquire authority only inferentially. This may obviously be true for analytic claims, but what about observational knowledge? To answer this question, first of all it has to be recognized that there is a sense in which both analytic and observational claims acquire authority by being *correctly* uttered according to *rules*: deductively established rules for analytic statements and inductively established rules for observational reports[10]. However, as we already began to see in Chapter 1, this won't do as it stands: indeed it's a crude form of the *Myth* to suppose that the authority of observational report, and thus observational knowledge, is parasitic on the inductive exploitation of a sample of behaviour whose qualification and classification is previously, pre-conceptually, given[11]. And yet, this acknowledgement lets us take one step forward in the analysis of our perplexity. If claims acquire authority in that they are performed, as actions, according to rules, then, if these linguistic rules are not mere uniformities in linguistic behaviour, the only way claims can be authoritative is to *be treated* as such. Here is how the difference between reliable differential responsive dispositions of inhanimate and sentient beings and knowledge of sapient beings becomes a normative one. Obviously, to repeat, this argument is sound only if the difference between regularities and rules is acknowledged. In this sense to endorse a claim is to treat it as performed according to reasons, to treat it as justifiable. Thus, for instance, my utterance of "This is red" can be treated as observational knowledge only if I also know other facts like "I'm a reliable reporter of red things" I can use to justify it. When I can't justify my claim,

[10]Some details can be usefully added here. Consider the distinction between sentence *types* and sentence *tokens*. Types of sentences are justified by semantic rules pertaining to their component expressions. This is true both for analytic sentences and for observational sentences. Now, analytic sentences do not contain expressions whose semantic rules depend on the context of their utterance and can be deductively established. Thus tokens of analytic sentences inherit justification directly from their types: any token of "$2 + 2 = 4$" is always justified by being a tokening of the type $\ulcorner 2 + 2 = 4 \urcorner$. Instead observational sentences contain those so called "token-reflexive" words, like indexicals for instance, whose semantic rules are sensitive to the context of utterance. Thus rules for observational sentences types have to be established inductively by correlating utterances with their contexts. However, once these rules are thus differently established, justification proceeds deductively as in the case of analytic sentences. As Brandom remarks in his *Study Guide* to *Empiricism and the Philosophy of Mind*, the unstable status of analitic sentences makes this distinction even slighter: if one construes analiticity as counterfactual robustness, as Sellars does, and construes inductive inferences as establishing normative entitlement to statements, as Sellars does as well, there remains just one conception of semantic rule and, consequently, just one way to justify statements.

[11]See Brandom's analysis of *reliabilism*: Brandom [14], pp. 206-216.

because for instance conditions are such that I'm no more reliable in my observational reports, I may prefer to withhold my endorsement. It's worth introducing Brandom's words[12] to explain this: by asserting a claim one commits to the responsibility to justify her entitlement to endorse it.

Let's propose a brandomian recap[13]. In his "two-ply account of observation" Sellars pursues a kantian strategy in describing the interaction of two sort of practical abilities, which fit the distinction between receptivity and spontaneity, *sentience* and *sapience*[14]: language-entry moves without language-language moves are blind, language-language moves without language-entry moves are empty[15]. Brandom then helps us to further distinguish two elements in the analysis of the space of reason: on the one side the inferential articulation of reasons, on the other side the correspondent normative representation of the practices of giving and asking for reasons.

Thus we can understand why one should withhold the endorsement of a claim: because he may not be able to vindicate the responsibility to justify it. Yet this may not completely wash away our perplexity. One might still wonder: who speakers are responsible to?

If possible, this second apparently naïve question digs even deeper than the previous one. I have to be content with reporting the *different* answers Sellars and Brandom gave to this answer. Brandom's answer is that responsibility is to be vindicated in a *social* dimension: the authority of claims has to be vindicated to other sapient beings, to whom speakers are responsible of their endorsements[16]. Sellars's answer instead is that responsibility is *internal*: speakers must be able to justify their claims on their own in order to be treated as sapient[17]. We'll better appreciate the consequences of these difference in section

[12]While this analysis is the pulsing heart of the normative account of Brandom [14], he came up with it much earlier, in Brandom [12]. We are going to deal with it in more detail below in section 3.3.

[13]Here I follow Brandom [19].

[14]In this sense C. I. Lewis's kantian pragmatism is probably one of the main silent interlocutors of Sellars in *Empiricism and the Philosophy of Mind*. See Lewis [85]. It's also worth noticing here that those who deny any citizenship to receptivity in the space of reasons, provocatively describe Sellars as wavering in front of an inevitable stark choice between this kantian pragmatism and the complete rejection of empiricism. See for instance McDowell [103].

[15]I owe this nice sellarsian paraphrase of Kant's famous adage to Brandom. See Brandom [19], p. 352.

[16]Notice however that this social dimension shouldn't be construed as a third-person perspective – what the community takes to be true or following according to the rules – overriding second-person perspective – what each member of the community takes to be true or following according to the rules. This is indeed a questionable point in Brandom's account. See Habermas [66].

[17]In order to have this difference clear it's worth reasoning on a discriminating example. Consider chicken-sexers who can sort hatchlings into male and female by applying *reliably* but *unconsciously* their differential responsive dispositions. They can tell the sex of the chicks they are inspecting, but they are unable to tell how, i.e. they can't justify their observational claims. To the contrary of Sellars, Brandom delcares these capacities as *sufficient* for observational knowledge to be attributed, since observational

3.3.5.

For those who are not already authonomously acquainted with Sellars, this disagreement might sound surprising. The fault is mine, since in reporting *Empiricism and the Philosophy of Mind* I deliberately adopted Brandom's interpretation of it: in particular, I took for granted the idea that claims have authority only in that they are *treated as* having authority. In Sellars [147], this is indeed a crucial exegetical point. But exegesis is not our task here[18]: recall we are just stepping through Brandom's account of normativity in order to grasp the underneath of his inferential semantics. Then I maintain my choice and move forward to propose an analysis of this clash within the perspective here adopted. The most insightful way I know to make this point clear is to pour it into the forms of Hegel's analysis of Kant's notion of *autonomy*, i.e. the ability of binding oneself to norms[19]. In this sense what we've learnt is that autonomy is the condition for the knowledge of determinate conceptual contents:

> «Kant was the first to think of representation in explicitly *normative* terms. To treat something as a representation is to take it to be subject to a distinctive kind of assessment of correctness. It is to grant a distinctive kind of *authority* with respect to such assessments to what one thereby takes to be represented by the item subject to those evaluations. To represent things as thus-and-so is to *bind* oneself, to make oneself *responsible* to the things for the correctness of one's representation.»[20]

knowledge is attributed by other sapients who can provide a suitable argument to justify those claims and *treat* them as expressing knowledge. See Brandom [15], pp. 102-106.

[18] I'll have to briefly come back to this in Chapter 5, but in the meanwhile I can't help laying down some hints about this theme that I hope may be useful here. Sellars certainly maintains that non-inferential knowledge is possible. Thus the exegetical problem is to explain how he construes statements that seem to acquire authority non-inferentially. If one doesn't take this hurdle as a *reductio* of Sellars's argument against *the Given*, then there are two main explanations, which have been put forward respectively by Brandom and McDowell. The first one says that, even if there is no knowledge without inferentially articulated conceptual contents, nonetheless, since knowledge is a normative status, it can be attributed to non-inferentially elicited statements privided that inferential justification can be supplied *ex post facto*. This first interpretation has to face one main problem: Sellars explicitly admits non-inferential reports of contentful inner episodes (Sellars [147], §60). The second one says that even if the space of reason is a normative space and reasons are inferentially articulated, inferential justification is just one way to handle contents: the other way is a primitive cognition of "pure *de re*" beliefs and it is this second sort of cognition that is at play in non-inferential reports. This second interpretation has to face one main problem: the determination of concetpual content. Both Brandom and McDowell hit on these themes pervasively in their works, but they are particularly put into focus in Brandom [14], pp. 215-221, Brandom [19] and McDowell [95, 97].

[19] This is explicit in Brandom [18], pp. 216-226.

[20] Brandom [25], Chapter 4, section II.

Now, Hegel criticized Kant's *internal* notion of autonomy and, unlike Kant, understood normative statuses as *social* statuses:

> «What I am calling [normative statuses'] 'social' character, Hegel takes to be synthesis by *reciprocal recognition*. This means that what such a status (paradigmatically, being *responsible* for something, as in judgment and action) is *in itself* (its essence) is a product of what it is *for* the one who undertakes it and what is *for* others, who attribute it. The underlying thought about normativity is what Hegel makes of Kant's connection of normative bindingness with autonomy [...]. This is the idea that one only counts as *normatively* bound or constrained by obligations, responsibilities, or commitments that one oneself acknowledges *as* binding. Seeing the contents of such normative statuses as instituted by a process of mutual recognition is Hegel's way of securing the *determinate contentfulness* of commitments in the context of using this autonomy condition as a principle of demarcation for the normative. For determinate contentfulness requires that *what* I have committed myself to not be up to me in the same sense as *that* I have committed myself to it (since where whatever seems right to me for that reason is right, there is not real question of right or wrong).»[21]

In his kantian approach Sellars recognizes the normative character of conceptual contents, but focuses only on the dimension of *authority* of normative statuses. To say that the one who endorse an observational report is the same who has to vindicate it inferentially, is to say that the one who has authority over a claim is the same whom such authority is responsible to. This account of normative character can work only if contents are – inductively – already completely determined before they are – deductively – applied (this is very roughly Kant's own view, as it emerges from his *Critique of Judgment*), so that it is the determinateness of contents that provides the objective perspective anyone should adopt in checking their application[22]. But if this complete determination of conceptual content

[21] Brandom [25], Chapter 3, section I.

[22] In fact, Sellars quite explicitly seems to share this analysis in §37 of Sellars [147], where he makes it clear that while no observational knowledge can be attributed to any speaker S *now* uttering "This is green" if she can't justify it *now* by showing knowledge of other facts like "utterances of 'This is green' are reliable indicators of the presence of green things", this doesn't mean that any *previous* utterance of "This is green" by S should have been considered as expressing observational knowledge, since the inductive process that led to the acquisition of knowledge of general facts about green thing might not have been completed. Let me add here a few words about the debate on Sellars [147]'s interpretation, since there is a certain question this remark should shed some light on: does Sellars admit reporting use for sentences

as independent from thier application is questioned, then this one-sided view doesn't put at play any real normative bindingness, and, consequently, it can't provide an analysis of knowledge of determinate contents. In his hegelian approach Brandom deploys both sides of normative statuses in a social articulation of authority and responsibility. This is why the one who endorse an observational report, i.e. the one who claims his authority over a certain content, is responsible for the justification of such authority not to himself, but to other sapient beings[23]. In this sense the normative status of one's claim does not depend only on his endorsement of them, but it also requires this endorsement to be attributed.

Now, even though this idealistic setting may look artificial as abruptly brung forward here in the meagre space of these preliminaries, it is instead the result of a long lasting effort of confrontation with sellarsian themes on Sellars's own terrains, which include a deep and participated analysis of the history of philosophy. I just want to note here one more observation we'll need to develop in Chapter 5. While acknowledging Sellars's normative characterization of conceptual contents, Brandom has always been dissatisfied with the criterion of counterfactual robustness used by Sellars to distinguish laws from regularities: his complaint is that any generalization supports particular ranges of counterfactual robustness, and the real problem is to determine these ranges[24].

Let's take stock. We've been describing rationality, as opposed to mere differential responsive dispositional capacities, as the ability to deploy conceptual contents. This in turn has been found to be manifested in the practical ability to endorse claims, that is to provide justification for them if required. As a practical ability, this so called "game of giving and asking for reasons" has a normative structure. This means that the rules of this game express, or make explicit, conceptual contents. We also found that if we don't require these rules are completely made explicit before the game starts, speakers must not

expressing knowledge, like "I know that p"? There are some like, for instance, McDowell who argue as follows: knowledge statements admit reporting use, then it must be admitted some sort of *immediate* – in the Hegelian sense –, non inferential but conceptual, knowledge these statements report. See McDowell [98, 100, 103]. Now, it is obviously true that Sellars admits a reporting use for knowledge statements: he does it in §59. And yet he also remarks that in recognizing this so acquired first-person authority over one's reports of her own conceptual knowledge, the interpreter should not forget the *intersubjective* process that led to such an acquisition: what can be treated as the report conceptual knowledge in the space of reasons is not guaranteed to be such at any previous point in the inductive process of acquisition of that knowledge, i.e. of learning how to place the relevant concepts in the inferential web of reasons.

[23]Here is where one should charge the responsibility to develop Sellars's *Meditations Hegeliènnes* a little further, but I can't venture to to that now. Hopefully Brandom actually took this responsibility. Thus I refer to Brandom [17, 18], to his *Woodbridge Lectures* given at the Columbia University in 2007, Brandom [29, 30, 31] and now published in Brandom [28, 1-3]), and to his yet unpublished work on Hegel's *Phenomenology* Brandom [25].

[24]See for instance Brandom [14], p. 634 and Brandom [23], p. 105.

have first-person authority over their application for these rules to be really binding. The normative structure is based on the second-person dimension of *reciprocity*: speakers have the authority to endorse any claim, but in doing that they become responsible of their endorsement towards other speakers who acquire, reciprocally, the authority to tell if the rule has been followed, i.e. if the concept has been correctly applied.

3.2.3 Normative phenomenalism

At this point one main demand needs to be satisfied before the question about a semantical representation of *this* normative structure can be asked: can norms be represented at all?

Indeed, analytical philosophy abounds with warnings against wrong ways to conceive this problem. Let me briefly sum these up in two main mistaken views. The first one, call it *regulism*, is the idea that norms are explicit rules that describe correct performances. Wittgenstein, in his *Philosophical Investigation*, taught us that the correct application of a rule can be ascertained only against another rule. The lesson we have to learn from the failure of *regulism* is that knowledge of norms is, *primarily*, not a theoretical "know-*that*" but a practical "know-*how*". The second one, call it *regularism*, is the idea that norms are implicit in practices in the form of regularities. The consequence of this idea is the reduction of normative vocabulary to non-normative one, but, again, Wittgenstein taught us that there is just nothing like *the correct* specification of a regularity.[25].

Brandom's way out of this dilemma is his "normative phenomenalism"[26]: «explaining having a certain normative status in effect as being *properly* taken to have it.»[27]

One good way to make sense of this idea is to start talking about intentional *stances*. In

[25]This is what Brandom calls the "gerrymandering" argument against *regularism*. See Brandom [14], pp. 26-30, pp. 206-212 and pp. 645-647. It is exactly what Quine showed for representational semantics: any specification of linguistic behaviour is undertermined with respect to its extensional interpretation.

[26]It should be noted that the interpretation of this solution is actually quite controversial. Some critics blame Brandom for creating an unresolved tension between two different phenomenalistic approaches to the problem of norms, which is convenient to baptize respectively "phenomenalism about norms" (henceforth PAN) and "normative phenomenalism" (henceforth NP). See Brandom [14], p. 628. According to PAN a performance is appropriate if it is taken to be appropriate: if you buy a ticket for the show they let you in. According to NP a performance is appropriate if it is appropriately taken to be appropriate: if you buy a ticket for the show they have to let you in. PAN reduces normative statuses – being appropriate – to normative attitudes – taking to be appropriate. In this sense PAN can fix the mistake of *regulism*, but since it doesn't prevent normative attitudes from being explained in non-normative terms, it fails the "gerrymandering" argument against *regularism*. In Brandom's intent NP should offer a non reductive account, but – this is the point of criticism – it would still beg the question for an explanation of original normativity, so that PAN would be the only really available position. See for example Gibbard [58], Rosen [136], Grönert [65], Wanderer [169] and Rödl [134].

[27]Brandom [14], p. 627.

fact, it seems we face a typical scenario for Dennett's[28] classical analysis of intentionality. Speakers are *treated* as rational if they can justify their claims, i.e. if they *properly* endorse claims. But endorsements are properly undertaken only if they are properly *attributed*. Brandom himself deploys the materials of this analysis to picture the question about the foundations of his structure of normative relations, and distinguishes between *simple intentional systems* and *interpreting intentional systems*[29]. In this sense classical problems rise. How to distinguish derived intentionality from original intentionality? How to avoid the collapse of normative statuses into normative attitudes given that intentionality is defined in terms of attributions of intentionality? It's worth digging a little more into this.

We saw above that *sentient* creatures are characterized by reliable differential responsive dispositions and that what distinguishes *sapient* creatures from them is the capacity to *treat* the performances produced according to those dispositions as having normative status. But what does this treating amounts to? Now, there is a sense in which intentional, rational, creatures may deploy this capacity non linguistically, for instance by applying sanctions. Yet it is crucial to realize that "non linguistically" here doesn't mean that these rational creatures are not linguistic creatures, it doesn't mean that they do not produce linguistic performances. They can *explicitly* acknowledge normative *statuses* by *saying* what they endorse. What they cannot do is to explicitly *attribute* normative statuses – suppose they lack suitable linguistic resources to *make explicit* normative *attitudes*. *Interpreting intentional systems* have at their disposal vocabularies to *make explicit* those normative attitudes, while *simple intentional systems* can *attribute* normative statuses only implicitly[30]. However if this distinction is not to be construed as a cognitive gap, an account is required of how these vocabularies may spring out from implicit cognitive practices. In fact, this is

[28]See Dennett [39].

[29]Brandom [14], p. 59.

[30]This is obviously the beating heart of Brandom [14]. Later he developed a specialized lexicon to picture this point, which can be useful to make it clear here. On the one side he names "Practice-Vocabulary sufficiency" the relation between a certain set of practices the engaging in which is sufficient for someone to be interpreted as deploying a certain vocabulary. So, for instance, the exercise of the ability to justify one's own observational claims is *PV-sufficient* to be interpreted as deploying observational vocabulary (and not merely differentially respond to stimulations). On the other side he names "Vocabulary-Practice sufficiency" the relation between a certain vocabulary and a certain set of practices when the vocabulary is sufficient to specify those practices. So, for instance, the intentional vocabulary deployed by *intepreting intentional systems* is *VP-sufficient* to specify what *simple intentional systems* ought to do in order to be treated as deploying observational vocabulary. This example is perspicuous enough to highlight another interesting relation, "*VV-sufficiency*", holding between vocabularies but mediated by practices: when a vocabulary V is *VP-sufficient* to specify a certain practice P which is *PV-sufficient* to deploy another vocabulary V', Brandom says that *V is a pragmatic metavocabulary for V'*, since it allows to say what one ought to do to be treated as deploying V'. See Brandom [23], pp. 9-14.

often the case according to Brandom: the abilities to engage in the practices that count as deploying these vocabularies can be algorithmically *elaborated* from more basic cognitive abilities, in the sense that those who implicitly engage in normative practices already know, *in principle*, how to do everything that is required to deploy vocabularies to specify those practices, although they have not actualized this potential[31]. These vocabularies that *make explicit* features of the linguistic practices from which they are *elaborated* are eminently *logical* vocabularies[32], because logical vocabularies elaborate-explicitate those basic discursive practices which are autonomous from, and thus often embedded in, all the others. In this sense logical vocabularies are much more widespread than what their partial formalization into logical language suggests: in particular, for what concerns us here, there is a set of *logical* locutions that *intepreting intentional systems* elaborated the ability to deploy and that allow them to make explicit those normative attitudes *simple intentional systems* already, but implicitly, undertake.

This obviously should make sense of the phenomenalistic account of normative statuses in both their derived and original sense. And yet here it rises the perplexity that mainly puzzles interpreters of *Making it Explicit*[33]: where do the norms come from? Or, in the other words we've just learnt, what is the logical meta-vocabulary required to make explicit what interpreters do in the usage of their logical vocabularies? Brandom's answer is that there's no such a thing, or, to be precise, that the logical vocabulary elaborating-explicitating the normative practices of applying concepts already contains everything that is required to make explicit the very interpretation of those practices. This condition, acquired by certain discursive practices, is what he calls *expressive completeness* and it is

[31]This presents another relation of the sort considered above, *PP-sufficiency*, holding this time between practices: when the abilities required to engage in a certain practice, P, are sufficient, in principle, to engage in another practice, P', Brandom says that P' can be *elaborated* from P. Thus, for instance children learn how to perform long divisions by learning an algorithm of multiplications and subtractions, which they already know how to do. See Brandom [23], pp. 26-28.

[32]Logical vocabularies stand in this relation of elaboration-explication (LX-relation) to autonomous discursive practices. An LX-relation holds between a vocabulary V and a practice P when P is *PP-sufficient* with respect to another practice P' which is *PV-sufficient* with respect to V. So, for instance, the logical vocabulary of conditionals is LX-related to the practice of drawing inferences, since the abilities to deploy conditionals can be algorithmically elaborated from the practice of drawing inferences. See Brandom [23], pp. 44-48.

[33]Such a perplexity actually assumed various appearances according to the own tenets of the several authors that diagnozed it. There are those who construed it as a decisive failure of Brandom's project (see for instance Rosen [136], or Hattiangadi [69]), there are those who construed it as the manifestation of some deep philosophical assumptions that led to unwelcomed consequences (see for instance Habermas [66]), there are those who, sailing on Brandom's same boat, take it as a proof of his inability to set the course (see for instance McDowell [101] and Dennett [40]), and there are those who simply try to fix it (see for instance Wanderer [169]).

what prevents the regress argument for intentionality to which Dennett proposed an evolutionistic solution: Brandom maintains that the members of a linguistic community who can deploy these logical vocabularies achieve an *interpretive equilibrium* in which interpretive stances are *recirpocal*, and, in the hegelian sense we put into play above, this this is to be considered that *social self-consciousness* required to make normative character determinate[34].

3.3 Linguistic roles

Now that we are off preliminaries it's time to face semantical issues[35]. In what follows I'll try to lay down our main coordinates: first sentence primacy will be displayed, second some misunderstandings that use to accompany standard *referential* semantics will be disclosed, third the idea of an inferential semantics will be presented, fourth Sellars's analysis of Truth will be deployed to introduce our semantical *explananda*, fifth a normative model will be provided for them.

3.3.1 Asserting

It's worth noting here a crucial point, in order to nip some anxieties in the bud and prevent them from infesting the reception of what follows.

[34]See Brandom [14], pp. 641-643. A bit too much of philosophical chauvinism could urge here the need for something more: while by adopting interpretive intentional stance interpretive equilibrium can be achieved, when it is really *correct* to adopt such a stance? Now, I think, it's worth taking seriously Brandom's "latitudinarian" answer that one ought to adopt it whenever she can. And, notice, to take this answer serioulsy doesn't mean to take it as an aswer to chauvinistic urges. Here the price for not adopting the interpretive intentional stance is not to be paid in terms of what me might know *mistakenly*, but in terms of what we might know, *period*. After all, if I can allow myself a provocation, why should the hegelian spirit become self-conscious? Now, this could be a proper conclusion for this piece of rortian reasoning, but I'd like to put forward as another possible finale the suggestion to take the hard path to investigate the concept of *person*. Sellars began to do that in his Sellars [152] and Sellars [160], and Brandom is doing it in his his work on Hegel. This, I think, is not far from Dennett's own suggestion, in his comments on Brandom (Dennett [40]), to put into play evolutionistic adaptation; but he misses this point probably because he keeps attributing to Brandom a kripkensteinian notion of "community".

[35]Our purposes however require to enter this is territory with care. Such a care would suggest to start from an analysis of our primitives and take some time here to deal with the notion of Truth. An argument should be provided wide enough to trace a clear picture of our approach in the framework of the state-of-the-art discussion, and, at the same time, straight enough to lead us quickly to a firm ground where we can start build Brandom's system. I frankly confess I refuse to squeeze such an argument in the tight space of this section. But I won't completely discard the responsibility: those same coordinates I'm going to lay down for our semantical analysis are the coordinates such an argument should follow. For a more extended presentation of an argument along these lines see Chapter 5 of Brandom [14].

As consequence of the normative approach described in section 3.2, *sapience* is to be construed as a capacity to properly produce assertions and to endorse claims: to the extent that knowledge depends on justification, and justification is (in a sense we still have to qualify) inferential, conceptual knowledge is primarily knowledge of *propositional* contents. This is highly controversial, but there is not enough space here to rehearse a complete argument to support this claim. Actually, we already catched a glimpse of this matter in Chapter 1 when we had to deal with some mistakes of designational semantics, but I think some more, though crude, remarks are still in order. To make a long story violently short, the roots of this thesis lie in the tendency, quite widespread nowadays and supported by the authority of Dummett's intepretation of Frege's philosophy of language[36], to construe the approach to the analysis of conceptual content as passing through three different phases along the history of logic: a) the classical aristotelian tradition, where concepts, as *ideas*, were conceived as the minimally acquainted contentful unit to be linguistically deployed in a subject-predicate structure; b) Kant's transcendental logic where concepts loose their contentfulness primacy in favour of the *act* of judgement, within which they become functions of unity of the sensory manifold; c) Frege's thesis that judgeable contents are the smallest units to which pragmatic force can be attached, which led to the inferential definition of conceptual content in §3 of his *Begriffsschrift* and whose theoretical consequences were bound to the fate of the "context principle" in his later works. Now, there are two main sorts of complaints against the idea of propositional primacy. On the one side one could wave the representational purport of conceptual content and imply that the representation of facts can only be accounted compositionally on the basis of the representation of objects. This implication is false: the representation of objects can be well compositionally accounted starting from the representation of facts since, as we will see, compositionality can be exploited in both directions, either from components to componds or from compounds to components. I think that underneath this idea there hides a confused construal of the very notion of representation, as I'll try to show in the next section. On the other side one could philologically disagree with Dummett on the interpretation of the results of Frege's analysis of predication and of the role the context principle had in his philosophy of language. Thereby one could imply that the rejection of the context principle is the best way to avoid *semantic holism* and thus the only way to prevent contents from quinean indeterminacy theses[37]. However things might be with the Dummett's work on Frege, this implication is false too: we're about to see that compositional analysis doesn't require

[36] See Dummett [41].
[37] Compare with Section 1.2.2.

unrelated basic "building blocks", and can work quite well in an holistic setting[38].

3.3.2 Naming and saying

While the above remarks merely intended to appease the discomfort sentence primacy might raise with respect to certain representationalist tradition, here I'd venture to catch a glimpse of what might lie in the underneath of such complaints. Let me start beating around the bushes.

This tradition adopts a designational approach to semantics, so that the representational relation is construed as holding between *things*[39], and the paradigm is *naming*. Originally, this view raised the well know puzzles concerning the sort of *things* named by *common nouns*. However, thankfully for modern logic, Frege swept away these basic ontological embarassments by showing how to understand predicates as functions which apply to objects in judgements[40]. But now, what sort of *things* are represented by judgements? Our traditional answer here, indeed traceable back to Aristotle, is that judgements represent *states of affairs*. What's hard to explain then is what sort of *things* states of affairs are to be represented as. Notice that the point is not to explain when the representation is *correct*: to account for such a correctness would be to account for *Truth*, which may possibly be an epistemological concern. The point is to tell what judgements deal with, what sort of *this*-nesses take part in a judgement and in which sense they represent a state of affair. Here is where our tradition really invites to bite the bullet of ontological debates.

Regrets. I think I've kicked up a fuss enough to make my prey run out of its lurk. I won't try to catch it though, and I'll rather be content with pointing it out[41].

Consider the statement, "the particle a has spin-s", and ask what does it 'represent'. It represents the particle a *as* having a certain spin. In a sense this means that it *says of*

[38]For the sake of completeness I'd like to add another minor sort of complaints which springs from the recognition of language complexity. The doubt rises that even if representational purport could be satisfied in the tiny model of propositional calculus as a language, the same may not be obviously true for complex natural languages as well. See for instance some minor objections in Fodor and Lepore [48], but also Stainton [163], and, less recently but somehow more insightfully, Belnap [4].

[39]See Brandom [14], pp. 84-85.

[40]And yet, one might ask, what sort of *things* functions are to be represented as? Probably it was *logicism* that urged Frege himself to fatally fall back to this question: squeezed between the rejection of *formalism* and the need to explain how to derive the concept "number" from purely logical notion, Frege needed to supplement his second order definion of numbers with identity criteria for concepts, which, after his distinction between *sense* and *reference*, he pinpointed in their extensions, i.e. courses of values. See Macbeth [89, Chapter 5].

[41]I'm actually picturing myself more courageous than I really am: the prey I'm pretending to hunt has already been caught and made docile. See Sellars [145], and particularly Sellars [151].

a that it has spin-s, which content, as we saw in Section 3.2, is defined within normative discursive practices. In another sense it means that it *pictures* a *complex object,* e.g. the *fact* that a has spin-s, in the same way as a plan represents a building (where a building is a complex object and a single brick is not)[42]. This sense is quite unconveniently supported by its first order language formalization, say "$S(a)$": $*S(a)*$, as a sign design, is a complex object as well, so that $*S(\ )*$ and $*a*$, as sign designs, are part of it. In this sense our statement is a complex name for the complex object it labels.

The confusion of these two senses leads to the confusion of asserting with picturing[43]. In particular this leads to think that representational purport of linguistic expression could be explained in terms of the picturing relation that holds between them and things they *name* in the world: indeed, while linguistic expressions do, in a sense, picture things in the world, it is this very fact that begs the question about how represent*ing*s can point beyond themselves to represent*ed*s.

By committing to this view I doubly owe to the traditional semanticist: I have to provide a new account of conceptual content and a new paradigm for semantical analysis.

3.3.3 Inferring

There's a nice and smooth way to accomplish the task to provide an alternative account of conceptual content. I just need to point at the well known Frege-Carnap inferential definition of conceptual content[44] and borrow Dummett's complaints against Frege's 'crush' for Truth in the 1890s, with the consequent abandoning of the inferential path[45].

But I'm not really content with that, for two main reasons: first, by skipping over Frege's own reasons to 'abandon' the inferential path we run the risk to simply take the wrong path ourselves and, second, the inferential path wouldn't lead us much far anyway without due qualifications. So, let's try to dig deeper under inferentialism.

First of all let me notice that this story about inferentialism can't be directly traced back to the story about propositional primacy sketched above in section 3.3.1: Kant maintained a *classificatory* view about concepts according to which judgements, though primary with respect to objects, consist in the subsumption of *a certain* particular (the non-conceptual

[42]This latter is the sense Wittgenstein puts it in the *Tractatus*, Wittgenstein [173], 2.0272.

[43]Sellars severally denounces this confusion: see Sellars [147], § 24, and Sellars [153], p. 199, and specifically [151, 154, 157].

[44]See respectively the already quoted Frege [50], §3, and Carnap [32], §14 (in the Routledge ed., pp. 41-42).

[45]See for instance Dummett [41], pp. 432-433.

materials passively collected by intuitions of sensibility) under *a certain* universal (the unificatory functions of the spontaneous activity of understanding). What Kant rejects is the idea that this classificatory activity, as carried out by sapient creatures, might be explained either in purely naturalistic or in purely rationalistic terms. Here is where, if the confusion highlighted in section 3.3.2 is not avoided, one would ask what more features a representation must have in order to be conceptually contentful and not merely the product of reliable differential responsive dispositions. Here is where, by acknowledging the normative character of judgements, we introduced a sellarsian account the concepts in terms of practical mastery of *justifications*. In this sense the primary structure of conceptual – propositional – content turns out to be *inference* rather than *predication*.

The second point concerns Frege. He progressively came clear of the risks of the confusion between picturing and asserting in the terms of the confusion between the role words have in predication and the role they have as expressing meaning. And he tried to avoid it precisely by introducing in the 1890s, next to the distinction between concept and object, the distinction between *sense* and *reference* as *a different sort of distinction*:

> «we can't say that an object is part of a thought as a proper name is part of the corresponding sentence.»[46]

This is why proper names are not disguised descriptions, concept words are not common nouns for objects, unasserted thoughts lack truth values and it is the "striving for Truth" that makes thoughts dealing with objects.[47]

At last there is a third crucial point to note. Brandom[48] echoes Sellars [146] in denouncing a *formalist* received dogma according to which validity of inferences is determined only by the content of logical expressions. As a consequence of this dogma any *material* rule of inference, like e.g. "If I release a piece of chalk, it will fall", is to be considered an enthymeme. Sellars's complaint is that material inferences are essential to meaning and

[46] Frege [51], p. 187.

[47] That Frege never maintained a 'pictorial' interpretation of Truth is not really a new attainment of philosophical research. Dummett puts this just with particular reference to the confusion between picturing and asserting. Yet, while complaining angainst the shift towards the notion of Truth he is not equally prompt to acknowledge Frege's abandonment of the notion of fact as occurring in the *Begriffsschrift*. See Dummett [41], pp. 442ff.. To say that "p is true" is not simply Frege's 1980s way to say "that p is a fact", but Frege's way to clarify the distinction between assertion and picture by renouncing to his previous account. What Dummett's complaints risk to hide then is the unitary path that led Frege to introduce his semantical distinctions of the 1890s just to part conceptual content from extensions (something which he eventually failed to do for numerals). See Macbeth [89, Chapter 4].

[48] Paradigmatically in Brandom [14], pp. 94-104, and Brandom [15], pp. 52-56.

thus have to be considered *valid* inferences. Brandom seizes upon this suggestion and develops it according to his account of the *expressive* role of logical vocabularies we sketched above in section 3.2.3[49]: it is because we can elaborate *logical* vocabularies to make explicit as inferential relations the normative relations of material contents implicit in normative practices, that inferentialism can support a semantic account of conceptual contents. In this sense material inferences turns out to be an essential component of the inferentialist alternative to representationalism. This is why it's worth pausing to elaborate this point. In fact, quite obviously, we can't really make sense of the relations between inferentialism and representationalism if we just stick our nose into semantical perplexities, and forget that our semantics is but one big result of the development of modern logic, which has been carried on along the double path of what today is known as *proof theory* and *model theory*. In this sense, we should realize that if we decline the question about meaning in terms of proof theory we readily obtain a well worn *inferentialist* answer, firmly established since Gentzen's definition of the content of logical connectives in terms of inferential rules for their introduction and their elimination. It is this sort of answer that Dummett generalized to linguistic content *per se*: in his analysis of Frege's philosophy of language he applied this inferential account to the fregean meaningful unit, the sentence, in order to make sense of the basic notion of fregean semantic analysis, Truth[50]: propositional content is determined in terms of inferentially sufficient conditions for its assertion and inferentially necessary consequences of its assertion. It is here that we have to cope with the "formalist dogma". In fact it may seem that this approach, however generalized, would nonetheless drive us into a very steep path: we want to account for conceptual *contents* in terms of inferential roles, but these roles have to be determined within *valid* inferences, and valid inferences are defined in terms of their logical *form*.[51] In other words, valid inferences could make explicit analytic content only. The objection then rises that the idea that this inferentialist approach could ever define the content of non-logical lexicon ultimately depends on the presence, in the underneath of Brandom's inferentialism, of an unexplained clot of analytical necessity and synthetic content[52]. In Chapter 1 we already began to see that it is not at clot at all but a thesis, and we will see in Chapter 5 that isn't unexplained either.

[49]See Brandom [14], pp. 104-116, and Brandom [15], pp. 56-61.

[50]Dummett [41], particularly pp. 417-422 and 432-435.

[51]In fact, the main hurdle to clear for this approach has usually been considered to be the definition of *internal* criteria of *consistency* in order to rule out unwelcomed instances like Prior's *tonk* connective. See Prior [119], Belnap [6]. This was also Dummett's idea, in his generalized approach: criteria of *harmony* among introduction and elimination rules, while hard to define, would have been necessary and sufficient to determine the content of expressions. See Dummett [41], p. 455.

[52]See, for instance, Fodor and Lepore 49.

Let me just note here, in the meanwhile, that to show why this objection is misplaced, it's enough to rehearse Brandom's construal of logical vocabularies as elaborating and explicitating normative practices. In this sense, the received dogma of formalism simply lies on the privilege acquired by certain vocabulary, namely the vocabulary of formal logic. To make this clear it's useful to realize that, in this sense, *formally* good inferences can just be defined in terms of *materially* good ones: a material inference can be treated as *formally* valid *with respect to a certan privileged vocabulary* if it can't be turned into a bad one by substituting non privileged vocabulary in it[53].

Once these points are clearly acknowledged, inferential roles can be introduced in fregean terms as the place each sentence has in a certain inferential structure (determined

[53]Notice this obviously has not to be understood as the claim that one can differently pick up pieces of vocabulary and sort out formally good inferences at will. What Brandom has in mind here is, again, the comparison between inferential rules like $\wedge(A, B) \vdash A$ and $A, B \vdash \wedge(A, B)$ on the one side and $German(a) \vdash Boche(a)$, $Boche(a) \vdash Cruel(a)$ on the other side. Now, there's a sense in which they are similar. Both have one piece of *logical* vocabulary, represented by "$\vdash$", that makes explicit the fact that the rule is treated as a good one. Both have a privileged piece of vocabulary, respectively $\wedge(\,)$ and $Boche(\,)$. Both have an unprivileged part of vocabulary. Thus, they are similar because, in the same sense in which the first couple of rules deals with the meaning of "$\wedge$", so the second couple of rules deals with the meaning of "*Boche*". But there is also a sense in which they are different. It is the sense in which "$\wedge$", to the contrary of "*Boche*", is a piece of *logical* vocabulary in Brandom's sense: it makes explicit properties of the very practice of drawing inferences. As Brandom notes, this expressive sense of logical vocabulary seems to have been already acknowledged by Frege. In this respect it's worth recalling for instance that he, to the contrary of Russell, didn't consider Euclid's proofs as enthymematic, in the sense that they missed some premises: what Frege blame Euclid for was that he didn't make explicit some of the inferential rules he actually used in his proofs. See Macbeth [89], pp. 10-25. If we switch to Frege's bidimensional notation this point should be easier to be grasped. Consider

$$\begin{array}{l} \text{—}\, Boche \\ \text{└}\, German \end{array} \quad \text{and} \quad \begin{array}{l} \text{—}\, Cruel \\ \text{└}\, Boche \end{array}$$

Consider then what the conjunction expression allows to *say*,

$$\begin{array}{l} \text{—}\, B \\ \text{└}\, A \end{array}$$

and its introduction rule,

$$\begin{array}{l} \text{—}\, B. \\ \text{└}\, A \\ \text{└}\, B \\ \text{└}\, A \end{array}$$

Thus, while both inferential rules make explicit normative practices, they do not belong to the same *expressive* level. For further details on this point see Macbeth [89], pp. 74-79. In this sense the *formalist* received dogma originates from a double confusion: on the one side the confusion between the description of valid rules of inference and the expression of those rules, on the other side of different levels of logical expression.

by any inferential relation $\vdash$), and can be defined either in fregean terms as the pair

Definition 1. $IR(\phi) =< \{\Gamma \mid \Gamma \vdash \phi\}, \{< \Gamma, \Delta, \psi > \mid \Gamma, \Delta, \phi \vdash \psi\} >$[54]

(where the first sets makes explict the sentences ϕ is inferable from and the second set makes explicit the sentences which are inferable from ϕ together with correlate premises), or generally as a set of inferential rules that uniquely identifies such a pair.

3.3.4 Truth as semantical assertibility

The paradigm of analysis for representational semantics is Tarski's *T-equivalences,*

"ϕ" is true (in L) $\equiv \phi$,

which have to follow from any definition of "is true", for it to be *materially* adequate. This has appeared to be hardly deniable for those who think that Truth can be and is worth being defined as a semantical notion[55]. What can be questioned from a philosophical point of view is a certain standard understanding of this material requirement that may originate in the confusion pointed out in section 3.3.2.

Let's begin by asking *what* are the terms of a *T-equivalence*. Tarski's answer[56] is that they are a sentence on the right and the application of the truth predicate to its name on the left. That it is the name of the sentence and not the sentence itself to occur on the left, is because if we want to *say* that the truth predicate applies to sentences, we need names to designate the sentences to which the predicate applies. So far so good, or good enough. But now, let's ask what does it mean, for a name, to name a sentence. We have basically three options within the standard representational framework[57].

First, if we don't want to treat sentences are designating expressions we run into the trouble to explain how can a sentence occur without quotes in a semantic statement: in fact we simply couldn't say anything like

"snow is white" (in L) *names* snow is white

[54] See Peregrin [114]. What I define here as inferential role Peregrin originally calls inferential potential and distinguishes it from the inferential roles of subsentential component expression. I prefer to adopt "inferential role" in both cases and distinguish between "sentential inferential roles" and "subsentential inferential roles".

[55] The most famous alternatives to Tarki's standard definition, i.e. Kripke's proposal, Kripke [78], and the *Revision Theory of Truth*, Belnap and Gupta [5], accept the *material* requirement but try to handle its paradoxical *formal* consequences in order to avoid the regress of metalanguages.

[56] See Tarski [166], p. 156, and Tarski [165], p. 343.

[57] The following argument is massively borrowed from Sellars [153].

since the second occurrence of the statement would not occur truth-functionally.

Second, we can try to avoid the above conclusion by considering the naming relation as a correspondence relation between expressions in different langauges, s.t.

"la neve è bianca" (in Italian) corresponds to "snow is white" in *our* language

but this would obviously not solve the problem of what names of sentences are name of. We certainly can stop here and decide to dissolve the problem together with the semantic one about Truth[58], but that, equally certanily, is to throw the baby out with the bath.

Third, we can treat sentences as designating expressions after all, and try to show what objects they designate. Unfortunately, if what we began to see in section 3.2.3 is right, this is a dead end.

Let me ape a sellarsian turn of phrase: what is the alternative then? The alternative, once the difference between picturing and saying is treasured, is to go back to the second option and see if we can save the baby.

In fact we can now talk of sentences as complex sign designs that *picture what it is* without implying that this should explain how they *say of what it is that it is*. We acknowledged that the explanation runs in the opposite direction: it is because we use certain complex sign designs according to certain rules to say of what it is that it is that they picture what it is. But this sheds some light on the correspondence relation between complex sign designs in different languages: to say that two complex sign designs in different languages correspond in this sense, is to say that they are used according to the same rules to say of what it is that it is, it is to say that they have the same *linguistic role*. If we adopt Sellars's dot-quotation[59] to single out these roles – i.e. if we rewrite $IR(\phi)$ as $\bullet\phi\bullet$ –, we can paraphrase the above correspondence statement as

"la neve è bianca" (in Italian) is a •snow is white•.

[58] See Rorty 135.

[59] Just to briefly sum up, dot-quotation applies to expressions in a given *familiar* language to build *distributive singular terms* that refer to any expression in any language that play the same linguistic role of the quoted expression. So, as the *distributive singular term*

the pawn

refers to any *piece* (however materially realized) that is subject to certain rules in a chess game, in the same way the *distributive singular term*

•triangular•

refers to any sign design (however linguistically realized) that is subject to certain rules in a language game. For further details see Sellars [154] and Sellars [159, part III], §52.

But what does it mean then for a sentence to be true? We can get there in two steps. Recall *T-equivalences*. First, as we've just seen, we have to reconstrue what's in the left-hand side as dealing with linguistic roles for sentential expressions, so,

> the $\bullet\phi\bullet$ is true $\equiv \phi$.

Second we have to ask what does it mean for a linguistic role to be true. The answer can already be found in sections 3.2.1 and 3.2.2 above. It means that one is inferentially justified to assert complex sign designs that correspond to such a linguistic role, and this allows us to reconstrue the right-hand side of *T-equivalences*:

> the $\bullet\phi\bullet$ is *semantically* assertible.

In short, *T-equivalences* are to be considered requirements for the material adequacy of semantical definitions of Truth, not because they relate linguistic entities on the left-hand side to non linguistic entities on the right-hand side, but because they make explicit what is to treat a sentence as true within a linguistic practice.

3.3.5 Scorekeeping

At the beginning of this chapter I declared my intention to extinguish a promissory note to extensional semantics, and the sellarsian inheritance we've taken possession of here should be enough to pay the debt: we can now explain in which sense Truth is a normative notion. Unfortunately we've not yet done, and the last step is a hard one indeed. Let me try to put it as an objection.

> The notion of semantic assertibility needs to be further qualified. Suppose one takes for granted Sellars's analysis of Truth, i.e. that for a $\bullet\phi\bullet$ to be true is for it to be assertible according to semantic rules. Then, presumably, these rules will specify conditions for assertibility, such as those Sellars put forward with relation to empirical knowledge. In this sense any claim, the endorsement of which a speaker can justify, is assertible, and thus, by transitivity in *T-equivalences* in right-to-left direction:
>
> > if one is *allowed* to assert $\bullet\phi\bullet \rightarrow \bullet\phi\bullet$ is true
>
> But this is absurd.

In other words *objectivity* of Truth has to be accounted for. Notice that the point of the objection is well taken since there's no obvious answer to it in the framework we've been describing. Not yet. One premise is in order before we proceed: as it was foretold

in section 3.2.2 above, regretfully we have to leave here Sellars's path and line up with Brandom's analysis of this objection, his criticism to Sellars and his solution[60]. This will let us introduce the last piece of our framework: the notion of *Incompatibility*. We're going to get there in two steps.

The first one is to stick to *inferentialism* as presented in section 3.3.3 above and provide an account of linguistic roles in terms of inferential roles. In this sense any specific role, $\bullet\phi\bullet$, is to be defined in terms of inferential rules that specify *sufficient conditions* to assert $\bullet\phi\bullet$s and *necessary consequences* of the assertion of $\bullet\phi\bullet$s. What is important to recall of the above discussion is that *both* conditions and consequence of assertion are required to specify an inferential role $\bullet\phi\bullet$, since this provides us with an explanation of why the semantic assertibilism we've just presented fails the analysis of linguistic content: to make explicit how the assertion of a $\bullet\phi\bullet$ *could* be justified is not enough to specify its semantic content, which is instead completely determined only by making it explicit both introduction and elimination rules for it. Any theory that neglects either of these sides can't but fall short of a definition of a semantic notion of Truth.

The second step is to ask what do these inferential rules for lingusitic roles make explicit, in the sense of section 3.2.3. As we have already remarked, the answer to this question is one of Brandom's oldies[61]: since «assertions are essentially performances that can both serve as and stand in need for reasons»[62], what inferential rules make explicit is the normative structure of linguistic practices which is articulated in terms both of *entitlements* and *commitments* to assertions. As in any other normative practices, as a baseball game for instance[63], in human linguistic practices speakers are subject to rules that endorse or sanction their performances according to their normative status. These statuses can be aptly represented in terms of a *score* of entitlements and commitments to assertions. Suppose one makes a move in the practice by asserting "This tie is green". Then, on the

[60]Particular regret comes out from the fact that a proper account of the different perspectives of the two authors on this, though crucial, point would run us out of space. Let me just note here that the clash would probably reveal to be not so severe as it might appear at first sight. In Sellars [159], the next chapter after "Truth" is "Picturing", which completes the threefold part entitled "The Conceptual and the Real". *Picturing*, for Sellars, is a relation holding between relational structures, through a method of projection: in this sense language pictures the world through the occurrent conceptual framework. This is what establishes correctness criteria. The point is that in the absence of an acknowledgement of the hegelian dynamics of *mediation*, the sort of objectivity gained by Sellars can't reach far beyond the phenomenical perspective of the conceptual framework. In private conversation Brandom reported Sellars's personal dissatisfaction with the idea that this word of his could be the last on this problem.

[61]See Brandom [12], p. 640.

[62]Brandom [15], p. 189.

[63]This is the original example provided by Lewis [87], to whom the original formulation of the idea of human communication as a scorekeeping practice is due.

one side one *ought* to add to her score the commitment to it as well as to its consequences, e.g. the commitment to "This tie is colored". Now, as it can be drawn from section 3.2.3, to be committed to the assertion of a linguistic content doesn't imply the acknowledgement of such commitment – to commit to A, is to be disposed to assert all A's consequences, but one might not actually have all the dispositions she ought to have –, and nonetheless practitioners *qua* rational beings have to take the responsibility of the claims they endorse. Thus, on the other side, the speaker who commits to "This tie is green", if questioned, has to show her *entitlement* to her assertion, e.g. by justifying it on the basis of her reliable dispositions to assert "This is green" in the presence of green things. However, one has not to commit to any content she is entitled. This is the so called practice of *scorekeeping*.

We are now ready to understand what *Incompatibility* is. As the two dimension of inferential definition, introduction and elimination rules, dynamically articulate linguistic content in *logic vocabulary*, so, what logic makes explicit is dynamically articulated by these two dimensions of normative status, commitment and entitlement, in *normative vocabulary*: «two assertible contents are *incompatible* in case *commitment* to the one precludes *entitlement* to the other.»[64]

Notice then that there are three basic ways to perform *scorekeeping*[65]. One can keep score of commitments through *commitment*-preserving inferences. Thus, e.g "This is green" *commitment*-entails "This is coloured", which commitment-entails "This is extended", and so on. This has been typically formalized into deductively valid rules of inference. One can keep score of entitlement through *entitlement*-preserving inferences. Thus, e.g. "There's smoke here" *entitlement*-entails "There's fire here". This has been typically formalized into inductively valid rules of inference. But it is by considering the dynamic interaction of these normative dimensions that one can make sense of conceptual *contents*. One can keep score of incompatibility relations through *incompatibility*-entailment. Thus, e.g. being a man *incompatibility*-entails being a mammal because everything which is incompatible with being a mammal is incompatible with being a man. Inferences of this sort support counterfactuals, and in fact they have been typically formalized as modally valid inferences.

It can be concluded that *incompatibility*-entailment represents Sellars's notion of material inference and provides the basis for the inferential definition of linguistic roles we were looking for.

[64]Brandom [15], p. 194.
[65]See Brandom [14], pp. 189-191; Brandom [15], pp. 194-195; Brandom [23], pp. 120-122.

3.4 Subsentential roles

It was promised in section 3.3.1 that sentential primacy wouldn't jeopardize the semantic analysis of subsentential expressions, but nothing has been yet really said about how to extend the inferential approach beyond the account of simple propositional contents. The task here seems to be either to show how to do semantics without compositionality or to reproduce in terms of subsentential linguistic roles what Sellars did for Truth. The criticism to Brandom on this point maintains that neither of these paths are practicable by waving results from linguistics[66]: on the one side the far out productivity of linguistic competence can only be explained by presupposing the possibility to compose basic grammar elements, on the other side the idea that these elements could be recovered from complexes clashes with empirical datas.

Notice that this reason is particularly pressing for us here in two senses: on the one sides it should hold apart from any complaint about the representational purport of languages; on the other side and more importantly, it bites where it really hurts: the whole idea of a "semantics that must answer pragmatics" (along with Brandom's *adagio*) is based on a certain representation of linguistic practices as making explicit performances of rational beings, but if the best analysis of linguistic practices we have, as it comes out from linguistics, doesn't fit the representation insomuch as it turns out not to represent human languages at all, the whole inferentialist building would fatally collapse.

Let me recall that our task is double: recover compositionality of subsentential expressions and put it into play in an holistic framework. The accomplishment of this latter will let us shed some light on certain points of semantical analysis where doubts against the inferential approach might still be lurking.

3.4.1 Substitution and the determination of repeatables

Frege's 1982 solution to the problem of adjusting semantic analysis to the productivity of language was his famous *principle of compositionality*: the semantic intepretant of a complex expressions must be function of the semantic interpetant of the component expressions. This principle is usually applied *bottom-up* in order to define the semantics of complex expressions in terms of the simple ones. But it can be well construed as an application of Frege's original *substitutional* interpretation of conceptual content: given a semantically authonomous grasp on the notion of good inference, two contents will be

[66] See for instance Fodor and Lepore [48].

equivalent if by substituting the one for the other never turns a good inference into a bad one. In this sense the principle of compositionality can be applied *top-down* as well. Since Brandom too has a pragmatic grasp on the notion of good inference, he can peacefully enough establish that «[a] pair of sentences can be said to have the same pragmatic potential if across the whole variety of possible contexts their utterances would be speech acts with the same pragmatic significance.»[67] In other words propositional linguistic roles form equivalence classes under substitution in good material inferences. Once this is established and inferential roles are defined, it's easy to realize that the same substitutional strategy can be applied to obtain, indirect subpropositional inferential roles: a pair of subsentential expressions can be said to have the same indirect inferential role if and only if they can be always substituted the one for the other while preserving the inferential role of the sentences in which they occur. These indirect inferential roles are implicitly acknowledged by endorsing *substitutional inference*, as e.g. the inference from

"Hesperus has just appeared"

to

"Phosphorus has just appeared".

Here the conclusion is reached by *substituting in* the premise the expression "Phosphorus" *for* the expression "Hesperus". The appropriateness of substitutional inferences depends on the normative status speakers undertake by deploying substituted-for expressions, which status Brandom calls *simple material substitutional inferential commitment*, or SMSIC. Thus, in the sense of previous section, the meaning of any subsentential expression is defined by the class of SMSICs that determine the appropriateness of the substitutional inferences in which it occurs. The vocabulary of identity provides the linguistic resources required to make SMICSs explicit, as e.g.

Phosphorus is (=) Hesperus.

Let's try to put it formally: first suppose that subsentential inferential roles are at least as many as sentential inferential roles[68], and then require that the composition of the same subsentential inferential roles originate the same sentential roles. Thus,

[67] Brandom [15], p. 129.

[68] One could object to this that inferentialism should admit that there are more sentential roles because it is committed to meaning holism. But the objection is misplaced because it takes the commitment to holism as the commitment to the idea that the meaning of 'the whole' is something more than 'the sum of the meaning of the parts'. But this is false: holism here is simply the idea that inferential roles have to be defined over all good material inferences. See section 3.4.3 below for further details.

Definition 2. For every subsentential expressions e and e',

1. if $IR(e_i) = IR(e'_i)$, $1 \leqslant i \leqslant n$, then for every component function O, both $O(e_1, \ldots, e_n)$ and $O(e'_1, \ldots, e'_n)$ are defined;

2. $IR(O(e_1, \ldots, e_n)) = IR(O(e'_1, \ldots, e'_n))$.

At this point an important objection should rise: how could all this work for intensional contexts? These identity statements that make explicit substitutional inferential commitments must be affected by intensional variability: one might simply not know that he can use "Phosphorus" for "Hesperus". A quick remark can introduce the reply. Recall that scorekeeping practices are perspectival: each partecipant keeps his own score together with the score of the others. In this sense, the whole substitutional definition of linguistic contents takes place in the, so to say, 'extensional' context of each scorekeeper perspective. But then – one could rush to ask back with Quine [126] – what about *objectivity*? This further question really blurs two perplexities[69]: the first one deals with the objectivity of the *definition* of subsentential contents, the second one deals with the objectivity of the representational purport of language. The former already finds an answer in section 3.3.5: you can't report rabbit flies instead of rabbits with "Gavagai!" if you are not entitled to infer of what you report that it flies. The latter points to the problem of tracking inferential roles inside the whole bunch of communicative performances: I have to be able to pick any tokening of "Hesperus" you deploy in order to keep the score of your SMSICs. Notice that this same problem lies implicitly in the sellarsian account of truth we accepted in section 3.3.4 where we took for granted there an intuitive grasp on the cotypicality of claims. However, as we began to see in Chapter 2, things seem to be much more complicated for subsentential expression because of stronger realist intuitions about the notion of reference. Now it's time to include eventually these intuitions into the inferential perspective adopted here.

3.4.2 Anaphora and token repeatability

When we got in touch with pragmatic linguistic roles above, we took for granted an obvious distinction between a pragmatic role $\bullet\phi\bullet$ and the linguistic material $*\phi*$ which instantiate expressions of role $\bullet\phi\bullet$. Linguistic materials 'belonging' to one same type may vary both interlinguistically – thus, both *la neve è bianca* and *snow is white* are •snow is white•s –

[69] See Brandom [22].

and intralinguistically – thus, as in Frege's example, both ∗the Greeks defeated the Persians∗ and ∗the Persians were defeated by the Greeks∗ are •the Greeks defeated the Persians•s. From the pragmatic perspective we've been adopting one could say that ∗ϕ∗ is the token of an expression and •ϕ• is its type. Yet things are slightly more complicated. Consider the italian proverb: "a brigante, brigante e mezzo". We have here two ∗brigante∗s. The conclusion is that ∗ϕ∗ is not a token but a token-class of linguistic sign designs[70]. At this point, in order to represent these tokenings we can introduce Brandom's notation we already deployed in Chapter 2: thus we will use subscripted slashes to indicate tokenings, e.g. /brigante/$_1$ and /brigante/$_2$. Notice however that Brandom's corresponding brackets notation for types, e.g. <brigante>, is purported to single out intralinguistically expressions ruled by the same SMSICs: in these sense $< \phi >$ corresponds not to ∗ϕ∗, but to •ϕ•. Thus, for instance /brigante/$_1$ and /brigante/$_2$ belong to the token-class ∗brigante∗ but are tokenings of type <brigante>, since they are •rogue•s.

This notational warm-up should smooth enough the conceptual effort to recognize what is really at play when we make explicit our SMSICs in identity claims as we did above. To begin with, it is types that are equated[71], that is to say pragmatic linguistic roles. And it should be easy to realize that lexical cotypicality is not sufficient nor necessary for pragmatic cotypicality: thus on the one side while two tokenings /Giacomo/$_i$ and /Giacomo /$_j$ are obviously both ∗Giacomo∗s, they might well belong to different types, e.g. respectively <the author of this book> and <the author of *Madama Butterfly*>; on the other side tokenings of different token-class might belong to the same pragmatic type, e.g.

> /Giacomo Puccini/$_i$ was born on the 23rd of December, 1858 in Lucca, Tuscany. /He/$_i$ was the son of a choirmaster and organist. His father died when /Giacomo/$_i$ was five years old, and /he/$_j$ was sent to study with his uncle. In 1872 /he/$_k$ did begin his career as a local church organist.

This is what in linguistics is called *anaphora*. And, in linguistics, it's been acknowledged enough that anaphoric structures permeate the referential mechanisms of natural languages much more deeply than what orginally Chomsky suggested with the rigidly bipartite picture he described in this syntatctic theory of *Government* and *Binding*[72]. Brandom's idea is

[70] To be quibbling one should also distinguish another sense of token-class of sign designs. Consider the italian tongue-twister: "trentatre Trentini andavano a Trento tutti e trentatre trotterellando". We have here 6 phonic tokenings of "tre". This usage doesn't have to do with expressions, i.e. well formed strings, thus we have to take it apart. For further details on pragmatic uses of quotation marks see Sellars [144].

[71] See Brandom [14], pp. 450-452.

[72] See e.g. Kamp and Reyle [76], Ariel [3], Recanati [133].

precisely that to keep the score of SMSICs it requires to treat speakers as committed to certain tokens being part of certain recurrence structures which behave as anaphoric chains: it is only because a certain tokening $/\phi/_i$ is taken as part of a certain anaphoric chain of type $< \phi >$ that it can be treated as governed by certain SMSICs. In other words *substitutional commitments* define the meaning of subsentential expressions as organized in a grid of *anaphoric commitments*.

Let's take one step forward. As it has been already suggested, these structures of anaphoric recurrence may be not only intrapersonal but also interpersonal, i.e. they may hold among tokenings by different speakers. It turns out, obviously, that this feature is crucial for the objective evaluation of intensional variability of reference: to treat a tokening of one's own as anaphorically dependent on a tokening of another speaker is to commit to treat it as governed by the same SMSICs that govern the antecedent and this is what lets different scores to interact. Anaphoric commitments among tokenings may be shared among speakers who endorse different substitutional commitments toward types. In fact anaphoric commitments are much less – possibly at all – influenced by the perspectival character of scorekeeping because they deal with those basic syntactic features that allow to recognize a certain content as the same again, which it is easy to understand once one realizes that they are made explicit, for instance, by the binding of variable with quantifiers. To repeat what we said in Chapter 2, the failure to acknowledge the combined action of these two sorts of commitments and the idea that they both would deal with intensional variability, rises the troubles of two dimensional semantics. In this sense intensional variability of meaning, as represented by the different sets of SMSICs different speaker might acknowledge, makes no hurdle to communication because anaphoric commitments let speakers pick up those different sets from others' perspectives and evaluate them in their own one against their own sets of SMSICs.

3.4.3 Holism

Inferential semantics is committed to holism because inferential roles, like $\bullet\phi\bullet$, have to be defined over *all* material inferences. Because of its holistic character, there are mainly two sorts of objections one could rise against the proposal to treat this account of linguistic roles as a definition of linguistic contents. The first one is technical and deals with *compositionality*. The second is linguistical and deals with *translation*. I want however to highlight a third philosophical perplexity which might lie in the underneath on the two others and deals with *representation*.

3.4.3.1 Compositionality, a technical point

Shortly, the objection is that language's productivity and learnability require compositionality, but inferential role do not compose. Now, section 3.4.1 would suggest an easy way to reply that this simply isn't so: provided that propositional structures can be singled out as the minimal meaningful units, Definition 2 *establishes* that their inferential roles are individuated by substitutional decomposition in materially good inferences, and so compositionality is a built-in feature of Brandom's semantics[73]. However this doesn't satisfy critics like Fodor and Lepore [49], so let's dig deeper[74]. Let me say in advance that the technical problem of composition will be almost completely removed away and a metaphysical root will be found below it.

Montague [109] established a well known formal representation of compositionality as the requirement that the algebra of meanings is to be homomorphic to the algebra of expressions. Thus, suppose $\mathbf{S} =< S, \langle O_\gamma \rangle_{\gamma \in \Gamma} >$ is such a partial algebra where S is the set of strings and each O_γ is a partial operation on S with a fixed finite arity. Consider then a meaning function $m : S \to M$, where M is the set of meanings.

Definition 3. Consider O, a k-ary syntactic operation on S. We say that m is *O-compositional* iff there is a k-ary partial function G on M s.t., whenever $O(s_1, \ldots, s_k)$ is defined

$$m(O(s_1, \ldots, s_k)) = G(m(s_1), \ldots, m(s_k))$$

We say that m is compositional iff it is *O-compositional* for all operations O_γ on S.

If m is *compositional* it induces the semantic algebra $\mathbf{M} =< M, \langle G_\iota \rangle_{\iota \in I} >$ which we didn't have before, and which is homomorphic to $\mathbf{S}$.

Let me now make two points. On the one side, suppose m is defined well enough to ascertain that an operation O_γ is always defined for lists of synonymous strings, i.e.

Fact 4. *if* $m(s_i) = m(s'_i)$, $1 \leqslant i \leqslant k$, *then*

$$O(s_1, \ldots, s_k) \textit{ is defined iff } O(s'_1, \ldots, s'_k) \textit{ is defined.}$$

Then it follows

[73] This may seem trivial, but it's important to properly set the debate: why should someone disagree? See Peregrin [114].

[74] In what follows I'm going to exploit Westerståhl [171].

Fact 5. *m is O-compositional iff whenever $m(s_i) = m(s'_i)$, $1 \leqslant i \leqslant k$, and $O(s_1, \ldots, s_k)$ is defined, we have*

$$m(O(s_1, \ldots, s_k)) = m(O(s'_1, \ldots, s'_k)).$$

Here obviously Fact 5 represents the *Principle of Substitution*: by substituting simple expressions with the same semantic interpretant, the semantic interpretant of the complex expressions in which they occur doesn't change.

In this sense any semantics that provides a way to define an interpretation function m that preserves synonymity in composition *is* compositional. Inferential semantics can certainly do that. However this might happen for the trivial reason that substitutional strategy would eventually isolate only singletons, i.e. no synonyms. What sort of representation of meaning would it be? To make the general point, consider the result of Zadrozny [177]:

Theorem 6. *Consider the partial algebra $\mathbf{S} =< S, \cdot >$, where S is a set of string and $\cdot$ a binary operation (such as concatenation), and $m : S \to M$. Then there exist a set of functions $M*$ and a unique map $\mu : S \to M*$ s.t. for all $s, s' \in S$*

1. $\mu(s \cdot s') = \mu(s)(\mu(s'))$
2. $\mu(s)(s) = m(s)$

Here *1* says that μ is compositional and *2* says that the original meanings can be retrieved. On the whole Theorem 6 establishes that any semantic algebra can always be functionally modified to behave compositionally if one accepts to be loose enough in the structures that represent meanings. The reasonable complaint is that it makes no sense to talk about semantics in the absence of a sensible notion of meaning.

On the other side, let's try to apply in this framework Brandom's idea that compositionality can be exploited both directions. Thus consider Definition 3 and suppose to start with a semantic algebra $\mathbf{M} =< M, \langle G_\iota \rangle_{\iota \in I} >$ rather than a syntactic one. Now ask how this latter syntactic algebra could be induced on a set S of strings. We simply need a function $\sigma : M \to S$.

Definition 7. Consider G, a k-ary semantic operation on M. We say that σ is *G-decompositional* iff there is a k-ary partial function O on S s.t., whenever $G(m_1, \ldots, m_k)$ is defined

$$\sigma(G(m_1, \ldots, m_k)) = O(\sigma(m_1), \ldots, \sigma(m_k))$$

We say that σ is decompositional iff it is *G-decompositional* for all operations G_ι on M.

Let's complete this trivial inversion of the compositional framework with the following requirement,

Fact 8. *If* $\sigma(m_i) = \sigma(m'_i)$, $1 \leqslant i \leqslant k$, *then*

$$G(m_1, \ldots, m_k) \text{ is defined iff } G(m'_1, \ldots, m'_k) \text{ is defined.}$$

From which it follows

Fact 9. σ *is G-decompositional iff whenever* $\sigma(m_i) = \sigma(m'_i)$, $1 \leqslant i \leqslant k$, *and* $G(m_1, \ldots, m_k)$ *is defined, we have*

$$\sigma(G(m_1, \ldots, m_k)) = \sigma(G(m'_1, \ldots, m'_k)).$$

Here Fact 9 represents Brandom's idea of susbtitutional strategy for the individuation of component expressions' meanings: as, given Fact 5, compositional functions allow to determine the semantic interpretant of complex expressions, so, given Fact 9, if a decompositional function σ is provided then equivalence classes $\sigma(m_i) = \sigma(m'_i)$ can be selected into semantically complex expressions whenever the substitution of simple expressions doesn't change the semantic interpretant of the complex ones.

Notice, again, that σ induces the syntactic algebra $\mathbf{S} =< S, \langle O_\gamma \rangle_{\gamma \in \Gamma} >$, so it makes no sense to ask if the *decompositional* function σ is the inverse of a *compositional* function m indipendently defined. What it makes sense, however, is to ask what sort of syntactic representation $\mathbf{S}$ would be. This second complaint is reasonable as well. Let's try to make the general point once again and consider the result of Janssen [73]:

Theorem 10. *Consider* $m : S \to M$, *where* S *is recursively enumerable. Then there is a partial algebra* $\mathbf{S} =< S, \langle O_\gamma \rangle_{\gamma \in \Gamma} >$ *a partial algebra* $\mathbf{M} =< M, \langle G_\iota \rangle_{\iota \in I} >$ *and a function* $h : S \to M$ *s.t.* $h(O(s_1, \ldots, s_k)) = G(h(s_1), \ldots, h(s_k))$ *and for all* $s \in S$, $h(s) = m(s)$.

The theorem shows that a recursively enumerable language can be always generated by any compositional grammar. The reasonable complaint is that it makes no sense to talk about compositionality in the absence of a sensible notion of syntactic structure[75].

These mathematical results of Zadrozny and Janssen's, however, had a relatively unproblematic reception in linguistics. What they proved is that *compositionality* is a quite trivial technical requirement to obtain. This is enough for us to reject the objection as it

[75] For a deeper discussion of these two points see Westerståhl [171], Janssen [74].

was stated in the argument above: compositionality *per se* is no technical hurdle to clear in order to obtain learnability and productivity of language. Compositionality is rather a methodological tool in the search for solutions to semantical problems. What the discussion of these two points shows is that the problem of compositionality rises when a certain account both of syntax and of meaning are *given*.

3.4.3.2 Translation, a linguistic point

Now that the techincal issues should have shown to miss the mark, we can more clearly see where holism itches: what still lies on the table is the problem to explain how holistically defined contents can be objectively translated. Thus the argument now goes like this: "if holism is true, then I can't understand any of your language unless I can understand practically all of it"[76] and there are cognitive reasons to suppose that speakers can't or simply do not need to do that in order to communicate each other, so a certain sort of compositionality is required for natural language to be communicable. In this sense, I think, the underlying demand is once again to secure a sufficiently objective notion of reference. But if this is the demand, Brandom has a clear answer to it. Notice, preliminary, that in order to communicate speakers do not need to *share* the same contents, but to *grasp* – as Frege put it – the same contents. To put it in terms of *scorekeeping*, if I kept your score correctly I can identify any of the contents you commit to as precisely as you do, but that doesn't mean I would commit to it myself. Or else, to put it in the words of section 3.3.2 above, if I don't share your ways to describe things in the world it doesn't mean that, therefore, I can say nothing *of* those things you describe in those ways.

Once this is clearly acknowledged there's just a part of the reply still to tell. It resumes from section 3.4.2. There we saw how scorekeepers can undertake interpersonal anaphoric commitments in order to 'navigate through different perspectives': scorekeepers can keep the score of SMSICs of other speakers and evaluate them according to their own set of commitments. If one accepts the notion of objectivity coming out from Brandom's normative account of meaning, what is left to explain is just how scorekeepers can make explicit the practice of giving and asking for reasons *interpersonally*: this can be done by an analysis of the vocabulary for the ascription of *propositional attitudes*. In English, paradigmatic pieces of this vocabulary are those operators like "S believes that ϕ", which allow to say that someone is committed to some claim. Now, to those who would complain that this field has already been dug with the ploughshare of holism and that nothing but skepticism

[76] Fodor and Lepore [47], p. 9.

sprouted, I reply that perhaps nothing but quinean seeds had been sown. Quine [125], notoriously, argued that no extensional account can be provided for propositional attitudes, which, combined with Quine [122]'s rejection of the notion of analyticity, amounts to say that they can't be explained at all: four years later indeterminacy of translation could be harvested. The point, to repeat, is that we know from section 3.4.2 how to deal with perspectives over referents and we now need to make the score explicit. Let's preliminary focus on the difference between the ascription of knowledge and the ascription of mere belief. Thus, according to the standard view that knowledge is *justified true belief*, when a scorekeeper ascribes to S the knowledge of ϕ he does three things: he attributes to S the commitment to ϕ (S believes that ϕ), he attributes to S entitlement to ϕ (the belief is justified), and, crucially, he himself acknowledges the commitment to ϕ (the belief is true). The latter lacks in the ascription of mere belief. Now we can go back to the question about extensionality. Quine distinguished the ascription of propositional attitudes in *de dicto* ascriptions, e.g. "S believes *that* $\phi(t)$", and *de re* ascriptions, "S believes *of* t that $\phi(it)$", and claimed that while it is the latters we use to talk about referents, these can't be accounted for in purely extensional, i.e. referential, terms. But, let's ask, as before, what does a scorekeeper do in using *de re* ascriptions. He certainly attributes both commitment and entitlement to $\phi(t)$. What else? Recall that, according to Quine, *de re* ascriptions mistakenly project the extensional framework in opaque contexts, where coreferential terms can't be substituted salva veritate. In terms of the scorkeeping practice this means that a scorekeeper while ascribing a content $\phi(t)$ to S can't attribute to S commitment to his own SMSICs about t. Hence, the difference between a *de dicto* and a *de re* ascription of $\phi(t)$ to S is that in the former but not in the latter the scorekeeper attribute to S his own substitutional commitments about t[77]. Thus in ascribing a content de re a scorekeeper makes explicit he doesn't share with the speaker the *representational meaning* of t, but this, crucially, doesn't prevent him to extract *information* from the speaker's claims about t: e.g. "the native believes *of* a rabbit – which I know to be an animal, while he believes to be a temporal undetached part of an animal – that it just run in front of us". This is how Brandom account for the possibility of communication in his holistic framework.

I want however to concede that, while their criticism is misplaced, Fodor and Lepore strike a sound point here: in an holistic framework the problem rises of the finitary cognitive resources speakers have at their disposal to determine and manage contents. From a techincal point of view this turns out to be the problem of identifying the computational

[77]Notice that this makes of what is usually considered a distinction between two *sorts of beliefs*, a distinction between two *sorts of ascriptions* of beliefs. For a deeper analysis see Turbanti [167].

complexity of the holistic semantic process. That might be beyond what humans reasoners can actually accomplish. As McCullagh [93] put it, it must be distinguished between "finististic compositionality" and "atomistic compositionality": the former but not the latter (which Fodor and Lepore argue for) is a requirement linguistic analysis should take seriously into account. Thus Peregrin [114] too claims the inferentialist has to provide an argument for a finite basis of the computation of syntactically complex meanings. I stick to this scruple so in what follows, wherever differently specified, I'll just consider logically finite languages, that is languages that contain only a finite set of primary sentences.

3.4.3.3 Compositional representation, a philosophical point

Let's start from scracth, once again: what is a singular term? A very common answer recites that *singular terms are expressions that refer to objects*. An alternative answer has been provided above, but not much has been said against this standard one. Indeed there are contexts in which, I think, such an answer is quite correct: consider for instance the teacher who tries to explain the formal language of first order logic. In general, I think this approach is quite correct when the syntax and the semantics are *given*: in the case of the teacher who explains first order predicate logic they are simply given by inductive definition. However, from a philsophical point of view, it is important to realize that to take this as given implies to take the *logical space of particulars* as given. Now, there are two main reasons to think this is a good thing: a *metaphysical* thesis about what sort of things are there in the world, or an *epistemological* thesis about how we come to know them. Both these perspectives simply take the structure of first order language as shaped on either metaphysical or epistemological structures. Both can be easily discarded if only one catches a glimpse of Sellars [147]. Unfortunately to reject these as premises is not enough to avoid the temptation simply to assume the very *logical structure* as given. Quine is a notorious example, as we've just seen. While trying to reject any metaphysical and epistemological assumption, he maintained that quantified variables are the sterile scalpel with which theories carve their ontological partitions, but he took the structure of quantification to be given in advance to the determination of conceptual contents. And this inevitably rised his puzzlement about underdetermination of referents with respect to meanings, because he end up with thinking that the layout of one partition could be detached and overimposed on another one in order to confront the parts they carve out: no surprises they would be found incomparable[78]. In fact, to take the logical space of

[78]See Brandom [14, pp. 409-412.], but also Sellars [147, §§ 30-31] and Sellars [148, §§ 104-105].

particulars as given is to confuse, so to say, a logical and a metaphysical point: on the one side there's the logical structure of the representational linguistic device, on the other side there's the temptation to take this as the metaphysical anatomy of the world which a semantics simply correlates to an ontological model.

In this sense Frege, in the 1890s, felt the urge to trace a clear distinction on this terrain by adding to his previous logical distinction between *object* and *concept*, the semantical distinction between *reference* and *sense*. Let me try to dig the distinction again[79]. In Frege [50], § 9, he declares that the *argument / function* structure, with which he intends to replace the *subject-predicate* one, is not simply built in the content of a statement, but is given only in relation to an analysis. That is to say, for instance, that one can actually consider the judgement "$2^4 = 16$" either as dealing with the object 2 as an argument the function "4^{th} *root of* 16" – i.e. as expressing the content "2 is a fourth root of 16" – or as dealing with the object 4 as an argument of the function "*logarithm of* 16 *to the base* 2" – i.e. as expressing the content "4 is a logarithm of 16 to the base 2"[80]. Brandom's substitutional analysis of subsentential expressions treasures this insight. When the substitutional strategy is applied within a *substituted-in* expression it is possible to distinguish the expressions *substituted-for*, i.e. the expressions that are substituted, and the *substitutional frame*, i.e. the remainder of the substitution: thus in the example of section 3.4.1 "Hesperus" and "Phosphorus" are substituted-for expressions while "( ) has just appeared" is the substitutional frame. Now the susbtitutional inferences that govern the SMSICs of substituted-for expressions are symmetric, reflexive and transitive, i.e. substituted-for expressions form equivalence classes under substitution: this, according to Brandom, is what characterizes singular terms[81]. To the contrary, the substitutional inferences that govern the SMSICs of substitutional frames are antisymmetric: these inferences, according to Brandom, make sense of the relation among predicates that are made explicit by universally quantified conditionals, e.g. $(x)(x$ has just appeared$) \rightarrow (x$ is now observable$)$. What is crucial to notice is that, as it was for Frege, expressions might acquire these substitutional roles only with relation to an application of the substitutional machinery. So, the problem of the referential purport of singular terms rises only once these logical roles have been singled out, and it would be a doubtful idea to complain that many of the inferences we draw *about* the

[79] I'm going to develop this according to the analysis in Macbeth [89], Chapter 2. A similar point was stressed in Ramsey [130], with relation to the problem of universals.

[80] The example is one of Frege's himself. See Frege [51], 16-17.

[81] Not only Brandom, obviously. This is a way to make sense of Leibniz' principle of *identity of indiscernibles* and of its uses in recent analytic philosophy, from Frege [52]'s definition of natural number to Quine [121]'s criterium of identity.

referents of singular terms do not follow this analysis of the logical role of singular terms[82].

And things may get even messier when natural languages are considered. The idea of sentential primacy comes to the representationalist as the idea of the theoretical priority of a certain complex object (sentence), whose complexity (logical form) has to be explained and analyzed as the output of the syntactic processing. Thus, the fact that the substitutional strategy, applied to the structure of this complex object, fails the syntactic analysis as described in grammars by linguistics is construed as a failure of the inferentialist semantic approach that takes sentential meanings as priority[83]. The semantic *desideratum* would be instead a correlation from elements of grammatical cathegories to elements of ontological cathegories so that the syntactic generation would match the metaphysical one. The valuable corollary of this idea is the lining up of epistemology, privided that grammars would *in principle* represent functional structures of the mind. This ecumenical prospect, unfortunately, leans on a second misunderstanding that piles up to the first one between logic and metaphysics: the above argument is supported by the idea that grammatical categories play for the syntactically complex natural languages the same role as the logical structure expressed by the simple syntaxes of formal languages. The doubful outcome of this view is that generative grammars would provide the metaphysical structure of the world so that it would make sense to ask what in the world – what elements of what categories of what ontological models – corresponds, for instance, to a morphema that expresses aspectuality. Although this is just tauntingly put, the idea is perhaps unconsciously but remarkably spread enough.

3.4.4 Concluding remarks

Eventually, let's come back to the criticism we opened with: inferentialism fails linguistics. I think it has been shown enough why this criticism is misplaced: once the distinction between the logical and metaphysical level and a normative account of objectivity have been anckowledged, holism is no threat for inferential semantics. As the last bit of my argument here I want, provocatively, to turn that same criticsm against representationalism.

What linguistic datas from post-gricean communication theory make it clear is that 'sentence', as a complex syntactic object, is an *underdetermined* representation of propositional meaning: the semantical account of the composition of basic representational ele-

[82]See for instance Graham [63], McCullagh [94] and Fodor and Lepore [48, 49], and Brandom's replies to Fodor and Lepore in Brandom [22].

[83]Fodor and Lepore [48], for instance, try to apply to Brandom the same arguments Chomsky [37] applied to taxonomic linguistics.

ments falls short of an account of '*what is said*'. One straightforward solution is to accept pragmatic intrusions to provide what's missing. The overall picture of linguistic interpretation would be that of a production line with three stations: syntax assembles intentional raw materials, the output of syntax provides the input to semantics, semantics hooks up word-pieces to world-pieces, the output of semantics moves to pragmatics where it gets sumberged into contextual filler till meaning is ready. Unfortunately the impact of contextual information in the determination of propositional meaning appears to be much more widespread than the mere fixing the holes left by some indexical elements, and its interaction with semantical and syntactical analysis more pervasive.

What some linguists ask then about compositionality is "what exactly is it, if anything, that is compositional?"[84] And they look for it in representations of utterance processing[85]. At this level the elements that *compose* representations entertain world-meaning relations which are much more complex than extensional correlation.

Indeed, something goes missing in shifting from representationalism to normative inferentialism, but, I think, it all belongs to bad philosophy[86].

[84] Jaszczolt [75], p. 70.

[85] See for instance Zeevat [178], Recanati [132], Jaszczolt [75].

[86] Or, at least, they belong to a bad mongrel of philosophy and cognitive sciences. See Brandom [26].

Chapter 4

Brandom's Incompatibility Semantics

4.1 Introduction

The task of presenting Brandom's *Incompatibility Semantics* (*IS*) is relatively easy from a logical point of view. My main concern in this Chapter will be with properly lighting up the stage for the presentation itself. Let me thus say first something about the very enterprise of defining a formal semantics out of the inferentialist approach[1]. Then I'll move on to the discussion of the very definitions of *IS* as they are presented in the Appendixes of Chapter V of Brandom [23], which henceforth I'll refer to as B&A (2008)[2].

4.1.1 Why formal semantics?

Why do *we* need a formal semantics, in first place? A formal semantics for a given language $\mathcal{L}$ is a mathematical structure that provides a function to map strings of $\mathcal{L}$ to sets of possible worlds. There are two main reasons to believe that it would be a bad idea for the inferentialist approach to flow into a formal semantics.

The first one is a broadly rortian concern[3]. Once the battle with representationalism is won, once the epistemic framework of world representation has been substituted with the social framework of perspectival justification into the spotlight of human experience, once it has been acknowledge that justification has to do with relations internal to vocabularies rather than with external relations between minds and world, then the demand of a formal

[1] I will basically hit on the same point Peregring more extensively argumented for in Peregrin [110, 111, 112, 113, 115], since it virtuously fits with the purposes of Brandom's formal enterprise.

[2] This reflects the fact that these Appendixes are the product of the conjoined work of Robert Brandom and Alp Aker.

[3] In what follows, I borrow Brandom [16]'s analysis of Rorty's rejection of representationalism.

representation for meaning might sound highly suspiciuos. In fact, on the one side, if the demand is internal to a certain linguistic game, then it simply amounts to the trivial need for a disquotationable vocabulary of *Truth* that may allow to make explicit the provisional authority attributed to certain justificatory patterns. But on the other side, if it is manifested as the need for a general description of any linguistic game, it may hide a dangerous backslide into representationalism. Formal semantics establish rules relating expressions to their interpretants. In the sellarsian inferentialist framework we borrowed from Brandom these interpretants seem to be linguistic roles, thus an inferentialist formal semantics presumably would establish rules to associate any expression to its linguistic role. Now it would be a short leap to construe these roles as pieces of certain sorts of normative facts performing the old task of truthmakers. But the point such a formal semantics would hide is that there is no normativity at all, outside vocabularies themselves, to be represented. Wouldn't it be better to drop any representationalist talk at all? This is a theme worthy of a more extensive analysis which can't be pursued here. I will be content with two remarks which, I hope, will be sufficient to keep our path open. The first one is that if the complaint is about the imposition of a unique all-embracing metavocabulary, then formal semantics has nothing to do with that: it aims to provide a logical vocabulary to make explicit the rules of the space of reasons. The second remark is that to blame formal semantics for the metaphysical implications that philosophers draw from them is to throw the baby with the bath: this will be the theme of Section 4.1.2.

The second reason could be put as an advice: if inference is your concern, perhaps you'd better do proof theory. The force of this reason measures the gap between two main areas of modern logic: proof theory and model theory. There's a sense in which the gap seems not to be bridgeable: this sense is traditionally established on the one side by the famous Gödel's incompatibility result that blocked Hilbert's program and on the other side by the success of Tarski's analysis of logical consequence as truth-preservation. Obviously, the more is made of this view the more inferentialism and representationalism are conceived as cutting language at the joints of syntax and semantics. As a consequence, the whole normative inferentialism's approach to meaning would be construed, from a bird's eye view, as the revengeful enterprise of teaming up syntax with pragmatics in order to surround and defeat semantics, which has retreated in the unsafe and already damaged fortress of representationalism. I really don't think this picture makes sense, and, hopefully, I don't need to build a bridge from proof to model theory to show that[4]. As Brandom puts it with

[4]It's worth noticing, however, that inferentialism is revealing to be a nice spur in the direction of

relation to the language of arithmetic:

> «The concepts of arithmetic cannot be fully specified by finitely stateable rules of inference. Nevertheless, we do grasp those concepts. But this is just to say that we do in fact understand their inferential significance. To make explicit the inferences that articulate the concepts of arithmetic, we must employ model-theoretic metalanguages. This fact in no way impugns the inferential conception of conceptual content; it merely shows that traditional proof-theoretic metalanguages are not sufficiently expressively powerful to make such inferential roles explicit.»[5]

The fact that proof theory metalanguages fall short making certain conceptual contents explicit doesn't mean either that we can't grasp them nor that our grasp of them won't be inferential. If one only focuses on proof theory and lets everything spin around the problem of the explicit definition of proper finite sets of inferential rules, a centrifugal force is generated that crushes inferentialism and representationalism respectively onto the two extremes of syntax and model theory. In this sense, the above perspective on Brandom's inferentialism turns out to be quite narrow as soon as one realizes that our semantic theories of conceptual content do not have to share the same fate of representationalism. Once again, we'd better try to save the baby. Let's see how then.

4.1.2 Labelling and modelling

In the last Chapter, we often happened to focus on the distinction between the logical point of recognizing different elements in a structured representation of language and the semantic point of taking the representation itself as meaningful. When the problem was addressed in Section 3.3.2 a certain designational approach to semantics was presented as traditional of representationalism. My main concern there was with compositionality issues, but it's worth resuming the topic here in order to sweep away the perplexities it uses to rise by interacting with certain metaphysical interpretation of formal semantics. The idea that words have meaning in that they stick as labels to objects digs deeply in the ancient classificatory interpretation of concepts, which scholastic metaphysics developed and delivered to modern philosophy. And still after the so called 'linguistic turn', the more accurate the linguistic analysis of concepts becomes, the more difficult is to answer to the

investigating the gap. See, for instance, Peregrin [113], Prawitz [118].

[5] Brandom [14], p. 667n58.

same pressing question: "what in the world does this piece of language stick to?". Thus, as soon as the formal tools were developed to represent certain components of meaning recalcitrant to extensional treatment, in spite of Kripke's careful warnings not to treat possible worlds as "distant planets discovered by powerful telescopes", a new modal population immediately began to crowd formal ontologies. And it is quite revealing that the reactions to this ontological invasion, such as Quine's, charged it for the formal tools themselves rather than for the representational analysis of the semantic framework. Once again I just bounce the offer to engage in metaphysical debates. Please notice that I'm not implying they are pointless: as David Lewis' scrupulous work should have taught us, it's better to finish climbing the ladder before throwing it away – to use a wittgensteinian metaphor. The reason why I poked this fire here is just to draw the proper boundaries in my analysis.

The first preliminary point to make clear is that the rejection of the representational approach is not to be taken as the substitution of factual truthmakers with pragmatic ones. Let me make the monstrous offspring of this misunderstanding explicit. Suppose one has been convinced that meaning is normative and accepted the idea of Section 3.3.4 that *Truth* is not a relation between linguistic and non linguistic elements. Suppose she is so convinced by the inferential account of meaning that she is even willing to endorse the top-down approach to subsentential expressions. Well, she might still be so imbued with the metaphysics of naming to look for objects in the normative linguistic practice to stick labels on, and thus she could be puzzled by queries of the sort: "are there any normative facts?"[6]. What should be realized is that formal semantics doesn't explain meanings by providing a metaphysics of entities over which to stick them.

But then, the second consequent point to make clear is in which sense formal semantics explains meanings. Now that the field is clean of the ontological snares that, we've just seen, would make it collapse to representationalism, I can try to take a step in the good direction by considering what it is to have a *theory* of a certain phenomenon. In general – but linguistic – terms, it means to have a theoretical language which can be correlated to the observational language we use to describe the phenomenon through *correspondence rules*: then it can be said that the theory explains the phenomenon because the theoretical language is perspicuous enough to express laws explicitly[7]. In other words, it explains because it *models*. A model is a good one not because it correctly pictures what it models,

[6]See Hattiangadi [71] for an analysis of this issues from different point of view and outcomes.

[7]This doesn't commit either to an ontological nor to a methodological *derivation* of one language from the other. See Sellars [150].

but because it makes explicit the content of what it models[8].

Thus, when the metaphysical representationalist praises model-theory for providing a picture of the ontological commitments of language so perspicuous that *Truth* blows up by mere inspection, she misses the point that it is just because laws of *Truth*-preservation are made explicit that the model-theretic picture is a good model. In this sense the inferentialist too may enjoy the virtues of model-theoretic metalanguages as a model of inferential roles. As Peregrin puts it:

> «We may then say that the model is adequate if the stipulated inferential rules explicate the inferential rules [implicitly] constitutive of the real language. And we may say that semantics is simply our way of seeing the inferential roles as distributed among individual expressions – as the expressions being mapped on their contributions to the inferential patterns they support.»[9]

4.1.3 Making norms explicit

If we take stock it's easy to realize that there are two main demands a formal semantics must satisfy as a proper development of Brandom's normative inferentialism: first, it must be a model for inferential roles; second, it must not introduce an ontology of normative entities as representational correlates of inferential roles. We've just seen how these two demands are compatible, by showing how, *in general*, formal semantics explains meaning by modelling rather than labelling meanings with the formal apparatus of possible worlds. But it's important to acknowledge, *in particular*, what it is to model the normative structure of linguistic roles. In point of fact, one of the conspicuous values of Brandom's normative approach is to provide an operating alternative to the metaphysical burden of certain interpretations of model-theory that may sharpen the consciousness with which nowadays modal vocabulary is employed as a semantic lever to unhinge classical empiricist questions[10]. Thus, there are two points to make clear: the reasons for the importance of

[8]Let me note down two passing remarks. First, it is not obvious at all that this necessary condition is also sufficient for a model to be a good one: there might be some structural features that constrain the picturing relations of the model – see Sellars [159], pp. 116-150. Second, and in some sense consequently, models should be investigated in their structural features in order to appreciate their ability to make explicit the content they model – and this is also why metaphysics is better not be discarded too early.

[9]Peregrin [115], p. 4.

[10]Chapter IV of Brandom [23] interestingly looks back to the controversial reception of modality in the history of analitical philosophy facing its traditional empiricist rejection. With consideration of this story it can't be enough to point at formal results in order to understand the reasons of the pervasiveness of modal vocabulary in current philosophical talk.

modal vocabulary and its relation to normative vocabulary.

As Brandom uses to introduce the theme, the importance of modal vocabulary for the analysis of traditional empiricist questions is brightly summed up by the title of one of Sellars's early essays: "concepts as involving laws and inconceivable without them"[11]. As Sellars explicitly puts it in Sellars [148] one of the consequences of the *Myth of the Given* for *traditional* Empiricism is to consider the practice of deploying the phenomenal vocabulary of thing-kinds and properties as immediate in the linguistic representation of a world of facts[12]. One crucial element in this immediacy is, in particular, the autonomy from the nomological dimension of explanation, which is thought instead as established by induction: e.g. it would be by induction over immediate applications of "this match is scratched" and "this match lights" that the lawlike generalization "Every scratched match lights" is established. But the rejection of the *Myth* overturns this picture: if concepts are primarily placed in the space of reasons, their descriptive content is defined by those inferential relation which they are embedded in and which are then expressed by lawlike generalizations. This is what Brandom calls the "Kant-Sellars modal thesis":

> «in using ordinary empirical vocabulary, one already knows how to do everything one needs to know how to do in order to introduce and deploy modal vocabulary.»[13]

As to the second point, it should be intuitively obvious now to whoever read Chapter 3 what is the relation between modal and normative vocabulary. To repeat, it is by taking part in the practice of giving and asking for reasons that speakers acquire normative statuses defined by the rules of the practice, and these rules, implicit in the practices, are made explicit by developing and deploying suitable metavocabularies. In this sense both modal vocabulary and normative vocabulary are pragmatically mediated metavocabularies that make explicit the rules of the practice of giving and asking for reasons: modal vocabulary

[11] See Sellars [143]. However Sellars's argument there is hard-metaphysical, so that Brandom once admitted in private conversation that he couldn't really use it as it stands.

[12] The analysis of modal vocabulary is one of the main themes of Sellars's early essays, but it seems to fade away from the spotlight of sellarsian thought after his masterpiece Sellars [147]. This is often construed as a progressive drift towards the kantian themes of Sellars [159]. Despite this, Brandom [27] convincingly argues that the theme of modality is but one of the battlefields over which Sellars conducted his many-sided attack against *traditional* Empiricism, where Sellars [147] represents his more glorious and famous results. Brandom organizes the structure of this enterprise into three theses and conjectures that it was just by bracketing the analysis of modality that Sellars could develop the other two sides and eventually end up writing Sellars [147]. Sellars would attempt to deal with the third side, modality, in the other big essay of those years, i.e. Sellars [148].

[13] Brandom [23], p. 98.

makes them explicit, for instance, in terms of a model for semantic content; normative vocabulary makes them explicit, for instance, in terms of rules for scorekeeping.

To have what we need is enough to notice that this short-circuits: modal vocabulary is a semantic metavocabulary of norms, i.e. formal semantics provides a model for the contents involved in scorkeeping practices.

4.2 Definitions of Incompatibility Semantics

This Section provides a relatively informal definition of the main technical notions of *Incompatibility Semantics*, which is exactly an attempt to provide a formal semantics to represent normative inferential meaning. Now, Brandom obviously already told himself such a story in his own presentation of *IS*, so one may wonder why not just refer to B&A (2008) and move on? The answer is that I need here to highlight some features of these notions that might easily slip out the consideration of the reader rightly eager to start playing around with *IS*. I'll briefly consider the choices that led to these features and the reasons for those choices. I need to do that because I plan, later on, to test the theoretical basis of *IS* against some alternative choices. This is also why, in this Section, I'll have to considerably narrow my point of view and focus on some technical issues that might not evenly fit, at first sight, with the broader perspective of this work, according to which *IS* is primarily just a semantic metavocabulary to make explicit the normative dimension of meaningful linguistic practices. At any rate, I suggest not to skip this Section even to those who already have enough acquaintance with the details of *IS*, since I'll have to solve some issues they could still find interesting.

In what follows, first the formal settings of incompatibility frames will be introduced and a notion of logical consequence will be defined as a way to represent the dynamic articulation induced onto conceptual contents by incompatibility relations. Then logical connectives and operators will be defined to make explicit privileged inferential patterns in this articulation. Notice that the introduction of a logical connective in *IS* doesn't consist in establishing its semantic interpretant but in making explicit how to compute it: thus directly inferential definitions will be provided. This, crucially, is what allows us to introduce modal operators at the same way.

Eventually, let me remark that it can be proved that these definition satisfy Theorem 6 of Section 3.4.3.1, which should be the last word about the formal issue of compositionality.

4.2.1 Incompatibility frames

So the formal semantics for a given language $\mathcal{L}$ must be a model of the inferential role of $\mathcal{L}$'s well formed expressions. According to what we said in Section 3.3.3 this should mean that for any expression $\phi \in \mathcal{L}$ it must make explicit how to compute $IR(\phi) =< \{\Gamma \mid \Gamma \vdash \phi\}, \{< \Gamma, \Delta, \psi > \mid \Gamma, \Delta, \phi \vdash \psi\} >$. Notice however that, until now, we've been talking of inferential roles in relatively loose terms: in particular we required our practices to sort out inferences into materially good and bad, but we haven't put forward anything like a set of structural rules to define "$\vdash$" in order to make that possible in a proof-theoretic approach. Thus we can't now simply look for a model to represent it. Neither we can simply inherit model theoretic characterization of good inferences in terms of the tarskian notion of *logical consequence* because we rejected *Truth* as a primitive. We have to trace another route and look for another primitive. Brandom proposes to take as a primitive semantic notion the relation of incompatibility, which he can independently explain in terms of his analysis of normativity. This proposal resonates to a certain way to deploy model-theoretic semantics in order to represent the information conveyed by propositional contents which is quite standard in information theory. The idea is that to know that p is the case is to narrow the space of possibilities, of what else could be the case, by excluding what p rules out: then the information conveyed by the propositional content p equals to those other propositions that are incompatible with it. Thus, formally, the proposal is to take the set of sentences in $\mathcal{L}$ that are incompatible with p as the semantical interpretant of p.

The next question is how to pick up incompatibile sentences. To answer this, one has preliminary to get clear about what it should be, formally, for two sentences to be incompatible. B&A (2008) take a two step path. First, they define a relation of *incoherence*, Inc, over a given language $\mathcal{L}$ in order to obtain an *incoherence frame* $< \mathcal{L}, Inc >$. They require for an incoherence frame to be *standard* the following properties:

Persistence $X \subseteq Y \Rightarrow X \in Inc \Rightarrow Y \in Inc$

Symmetry $X, Y \in Inc \Rightarrow Y, X \in Inc$[14]

Then they define an *incompatibility function* $I : \wp(\mathcal{L}) \rightarrow \wp(\wp(\mathcal{L}))$ as their semantic interpretation function, and let it be related to Inc as follows:

Partition $X, Y \in Inc \Leftrightarrow X \in I(Y)$

[14]This is redundant here since sets are not ordered, but I want to make everything as explicit as I can so that the analysis of those results can properly evaluate what each definitory choice is responsible for.

Thus they define a *standard incompatibility frame* the set of sentences $\mathcal{L}$ and the subset of its powerset I closed to supersets $< \mathcal{L}, I >$.

4.2.2 Entailment

There is a certain hasty suggestion – I'd waver to call it 'interpretation' – about the notion of *following logically*, which were already highlighted and criticized by Carroll [33] but rather die-hard as we saw in Chapter 1, according to which a logical inference e.g. from p to q amounts to a *prescription* of asserting q once p is asserted (or, at least, to believe q once p is believed). This is obviously nonsense, as the tortoise tried to explain to Achilles. Logical inference is much better construed as establishing *constraints*: not what one ought to do, but what she ought not[15]. It is these constraints that ancient logicians tried to single out in what we identify as their theories of conditionals. Thus, as Sextus Empiricus handed down, Philo of Megara put it in terms of truth:

> «a conditional is false only when it begins with a truth and ends with a falsehood.»

To the same demand Chrysippus proposed another solution:

> «a conditional is sound when the contradictory of its consequent is *incompatible* with its antecedent.»[16]

These are considered the archetypes of two different interpretation of the conditional, but in both cases the constraint established by "if p then q" is the impossibility to deny q once p is asserted[17].

Now, if we are precluded from the truthfunctional path of Philo, it seems straighforward to put into play in Brandom's framework Chrysippus' idea of incompatibility. Thus the inferential constraint can be provisionally put in terms of scorekeeping practices by saying that *commitment to p prevents entitlement to everything which is materially incompatible with q*. But since any two claims are material incompatibility if and only if the commitment to the one prevents the entitlement to the other, this amounts to

[15] See also Peregrin [113], p. 3.

[16] Both quotations from Sextus Empiricus are cited as in William and Martha [172] (emphasis is mine).

[17] There might be those willing to complain that this is but a philological heresy. I have no philological answer for them here. I can just trace back essentially to Ryle [141] this way to think about conditionals in analytical philosophy. I obviously have theoretical reasons, as I have been trying to argue with relation to representationalism.

> «*p incompatibility-entails q* if and only if *everything incompatible with q is incompatible with p.*»[18]

Thus, for a well-worn instance, "Socrates is a man" incompatibility-entails "Socrates is an animal" because everything which is incompatible with being an animal is incompatible with being a man.

Some straightforward remarks. First, notice that incompatibility relations are established between sets but in general do not distribute to their elements: so, for instance, being adult is incompatible with being human and being five years old, but it is not incompatible with either taken individually. Second, incompatibility so construed is an implicitly *modal* notion: intuitively put, if two properties are incompatible then it is *impossible* for something to have both. As a consequence incompatibility-entailments are counterfactually robust. This, as we've already seen in Section 4.2.1, suggests an intensional reading of sets of incompatibles as representing propositional content. I'll have to come back on both of these points below. Third, this definition of incompatibility-entailment contains an universal quantification over the sentences of the language. This produces the expected holistic character of semantic content and implies that the propositional content of every sentence functionally depends on all the sentences expressible in a language, and thus it depends on the expressive richness of the language. As a consequence every extension, contraction or revision of the language might potentially reverberate on every single propositional content.

Some specifics. First, notice that, given a set of sentences $Y = \{y_1, ..., y_n\}$ there are two main ways to evaluate the set of sentences incompatible with it: either one may take the sentences that are incompatible with *all* of the y_i, thus $I(Y) = \{s : \forall y(y \in Y \Rightarrow s \in I(\{y\}))\}$, or one may take the sentences that are incompatible with *at least one* of the y_i, thus $I(Y) = \{s : \exists y(y \in Y \wedge s \in I(\{y\}))\}$. Brandom adopts the second way, thus interpreting $X \models_I Y$ as $X \models_I y_1$ *or ... or* y_n, and requiring Y to be finite. This is no real novelty: consider Gentzen's reading of a sequent $A_1, \ldots, A_n \vdash B_1, \ldots, B_k$ as an assertion that whenever *all* A_i are true, *at least one* B_i will also be true. The complete definition adopted by B&A (2008) is thus the following:

Definition 11. ($\models_I$) $X \models_I Y$ iff $\bigcap_{p \in Y} I(\{p\}) \subseteq I(X)$

Second, as we've just seen, Brandom establishes two requirements on incompatibility relation: *Symmetry*, if being a man is incompatible with being a chair then being a chair

[18] Brandom [23], pp. 120-121.

is incompatible with being a man, and what he calls *Persistence*, if being a man is incompatible with being a chair then being man is incompatible with with being a chair and being white. Now, although Brandom claims he wants to keep the definition as general as possible, these requirements applied to the above definition of consequence lead strictly to *classical* logical consequence. Recall classical consequence is characterized by three conditions:

Reflexivity: $A \subseteq Cn(A)$;

Cumulative Transitivity: $A \subseteq B \subseteq Cn(A) \Rightarrow Cn(B) \subseteq Cn(A)$;

Monotony: $A \subseteq B \Rightarrow Cn(A) \subseteq Cn(B)$[19].

Now, *Reflexivity* is given for free with Definition 11 of entailment, since $I(A) \subseteq I(A)$. The same for *Cumulative Transitivity*, for suppose $I(B) \subseteq I(A)$ then, given that $\subseteq$ is transitive, for any $I(p) \subseteq I(B) \Rightarrow I(p) \subseteq I(A)$. But *Monotony* also follows from *Reflexivity* and *Cumulative Transitivity* given Definition 11 and *Persistence*: in fact $I(A) \subseteq I(A, B)$, then for any $I(p) \subseteq I(A) \Rightarrow I(p) \subseteq I(A, B)$.

Before moving on, let's pause to make our point on the map: is this incompatibility-entailment really satisfying as consequence relation? First, one might complain that what primarily characterizes Tarski's formalization of the notion of "following logically" as a consequence relation is the idea of truth-preservation: thus, for some *connexion* to be a *consequence* relation it must primarily establish the preservation of a certain interesting status – if you don't like *Truth*, call it *justification*, *normative status* or whatever –, but incompatibility-entailment as normatively defined in terms of commitment and entitlement, rather than preserving something is just the conditional disavowal of an entitlement. Against this, it's easy to notice that, in the inferentially classical framework of *IS*, there is indeed something which is preserved by incompatibility-entailment and it simply is *compatibility*: to say that X incompatibility-entails p if and only if everything which is incompatible with p is incompatible with X is to say that everything which is compatible with X is compatible with p. This should not surprise us, for it puts into play Sellars's interpretation of *Truth* as semantic assertibility, and yet it must be taken carefully since, as we're about to see, compatibility and incompatibility usually do not complement with

[19] I adopt here Makinson's characterization, in consideration of the fact that *Cumulative Transitivity* is equivalent to general *Transitivity* given *Reflexivity* and *Monotony*. This will turn out to be usefull when we'll have to consider nonmonotonic relations in Chapter 5. See Makinson [90, 91].

respect to negation[20]. Second, suppose one has not really been convinced by the exploitation of Chrysippus' conditional and prefers to stick to *Truth* as a primitive. Now, let me note parenthetically that the choice of primitives is a delicate matter and it would be a bad idea to settle it just in terms of preferences. At any rate, this sort of objection can be better coped with by populating the undergrowth of Brandom's thesis with references rather than with arguments: in fact, this allows me to highlight the crucial contribution that Brandom's reflection on Hegel's work has on this point. As it was already noticed the idea of determination of content through exclusion is currently standard in information theory, and it's worth adding here that it is no less high-born than Truth. Let me recall how Brandom explicitly attaches his notion of incompatibility to Hegel's "determinate negation":

> «Hegel endorses the Spinozist principle "Omnis determinatio est negatio". For him, *determinateness* of content – whether of judgments and concepts on the subjective side of certainty, or of facts and properties on the objective side of truth – is always a matter of *exclusive* [ausschließend] *contrast* with, the ruling out of, other possibilities. These fundamental relations of material incompatibility, what he calls "determinate negation", in turn give rise to material *inferential* relations among the contents they articulate: what he calls "mediation".»[21]

This would turn out to be the inferential articulation of incompatibility entailment we've just seen: being a man entails being an animal in that everything incompatible with being an animal is incompatible with being a man. Contents are determined because the application of the one precludes the application of the other, and it is these relations of exclusion that establishes the *mediation* which is Hegel's hallmark for concepts and which Brandom construes as an implicit inferential structure[22].

4.2.3 Negation

So, the semantic interpretant of a sentence p is the set of its incompatibles. But then, what is it to be incompatible with *not* p?

[20] As we'll see in Section 4.3, Peregrin [115] showed how to make formal sense of the idea that *compatibility* in *IS* plays the same role *Truth* plays in model-theoretic semantics. By the way such an observation had already been put forward by Alp Aker in the third appendix of Lecture V of [23].

[21] Brandom [21].

[22] If however one would like to take it as an argument she could find in Pippin [116] reasons to reject it. I suggest however, for completeness, to take a look also at Brandom's responses in Brandom [20].

Brandom clears the issue in a few words:

> «Incompatible sentences are Aristotelian *contraries*. A sentence ad its negation are *contradictories*. [...] 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.»[23]

However, I think it's worth digging it out a little deeper and I'll try to do that step by step through examples. The first thing to notice is that the definition of negation is not so trivial as in a truthfunctional framework, since obviously we can't simply go for the complement of the set of sentences incompatible with p. In fact the set of incompatibilities of $\neg p$ would contain p and much more, indeed $I(\neg p)$ would easily risk to contain too much and end up with being self incompatible.

Example 12. Consider the language of primary colors $\mathcal{L}_C = \{Red,\ Blue,\ Yellow\}$ and suppose to have the intuitive pairwise incompatibility frame. Thus suppose $I(Red) = \{Blue,\ Yellow\}$, $I(Blue) = \{Red,\ Yellow\}$, $I(Yellow) = \{Blue,\ Red\}$. Notice, though, that this is correct but incomplete, since the whole powerset $\wp(\mathcal{L}_C)$ must be considered, while respecting *Symmetry* and *Persistence* of incompatibility relation. A perspicuous way to obtain that is to put values in a table[24]:

	R	B	Y	R,B	R,Y	B,Y	R,B,Y
R	+	-	-	-	-	-	-
B	-	+	-	-	-	-	-
Y	-	-	+	-	-	-	-
R,B	-	-	-	-	-	-	-
R,Y	-	-	-	-	-	-	-
B,Y	-	-	-	-	-	-	-
R,B,Y	-	-	-	-	-	-	-

Figure 4.2.1

If we take, for instance, $p = \{Red\}$, then $I(p) = \{\{Blue\}, \{Yellow\}, \{Red,\ Blue\}, \{Red,\ Yellow\}, \{Blue,\ Yellow\}, \{Red,\ Blue,\ Yellow\}\}$. Now, if we define $I(\neg p) = \overline{I(p)}$ then $I(\neg p) = \{Red\}$, as expected.

[23] Brandom [23], p. 126.
[24] I follow here Göcke et al. [59].

But now, suppose to extend $\mathcal{L}_C$ to $\mathcal{L}_S = \mathcal{L}_C \cup \{Triangle,\ Square\}$ and, other incompatibility sets being equal, establish $I(Triangular) = \{Square\}$ and $I(Square) = \{Triangular\}$. Thus, if this is extended to the powerset the following table is obtained[25]:

	R	B	Y	T	S	R,T	R,S	B,T	B,S	Y,T	Y,S	R,B	⋯
R	+	-	-	+	+	+	+	-	-	-	-	-	
B	-	+	-	+	+	-	-	+	+	-	-	-	
Y	-	-	+	+	+	-	-	-	-	+	+	-	
T	+	+	+	+	-	+	-	+	-	+	-	-	
S	+	+	+	-	+	-	+	-	+	-	+	-	
R,T	+	-	-	+	-	+	-	-	-	-	-	-	
R,S	+	-	-	-	+	-	+	-	-	-	-	-	
B,T	-	+	-	+	-	-	-	+	-	-	-	-	
B,S	-	+	-	-	+	-	-	-	+	-	-	-	
Y,T	-		+	+	-	-	-	-	-	+	-	-	
Y,S	-	-	+	-	+	-	-	-	-	-	+	-	
R,B	-	-	-	-	-	-	-	-	-	-	-	-	
⋮													⋱

Figure 4.2.2

This time $I(\neg p)$ would contain both $\{Triangular\}$ and $\{Square\}$ and thus it would be self-incompatible.

What this example shows and what is important to realize is that the semantical role of *not* p is not the semantical role of the set of p's incompatibles. This is because incompatibility relations do not part sentences into sets of *contradictories*, but into sets of *contraries*: if two sentences are incompatible one can't commit to both, but can commit to neither. In the language $\mathcal{L}_S$ of Example 12, one can commit to $Triangular$ without committing neither to Red nor to $Blue$ nor to $Yellow$. Thus the question is how to define contradictories in terms of contraries. Hopefully, that has been common knowledge among Aristotle's scholars. Peter of Spain in his *Tractatus* already provided a notorious lawlike formulation of Aristotle's theory of contraries:

> «The law of contraries is such that if one is true, the other is false, but not conversely.» (T I.14 (7)),

[25]Here and below, when the diagonal turns negative it means that sets are self-incompatible: after that the table will be uniformly negative.

For instance, *Red* implies *not Blue*, but *not Blue* doesn't imply *Red*. How to obtain the converse direction as well? The answer is obviously to take the disjunction of the other contraries, thus *not Blue* if and only if *Red or Yellow*. Negation is the disjunction of all the contraries.

Now we have the techincal resources to make this common knowledge more precise. So, consider a lattice algebra $< S, \wedge, \vee >$, let $a \leqslant b$ iff $a \vee b = b$ and define an incompatibility relation Inc over S, so that $a, b \in Inc$ means that a is incompatible with b. We want a complementation operation $x \mapsto \neg x$ for negation s.t. $a \leqslant \neg b$ iff $b \leqslant \neg a$. Now, provided that disjunction is lattice's join operator $\vee$, a's negation is just the least upper bound of all b s.t. $a, b \in Inc$. Thus we define $\neg a$ in terms of such a l.u.b. as

$$x \leqslant \neg a \text{ iff } x, a \in Inc$$

Here the same idea is applied to identify the the contradictory as "the *minimal* contrary", as we saw above. In this sense, $\neg p$ can be defined as

Definition 13. ($\neg$I) *Something is incompatible with $\neg p$ if and only if it entails p, i.e.*

$$X \in I(\neg p) \Leftrightarrow X \models p$$

Brandom and Aker take this as the axiom controlling the behavior of negation in *IS*, and that makes the negation to behave standardly: *Contraposition* and *Double Negation* will be guaranteed[26]. This is quite a standard definition in complemented lattice, but it is

[26]As Peter Smith pointed out to me, I'm skipping some important steps here. Notice in fact that, according to the algebraic reasoning, Brandom's definition for the introduction of negation should have been

Definition. ($\neg$I') $X \in I(p) \Leftrightarrow X \models \neg p$

Now, from a purely algebraic point of view, this would be enough to differentiate intuitionistic negation from classical negation. In fact ($\neg$I') would give us only *constructive contraposition*.

Proposition. $p \models \neg q \Leftrightarrow q \models \neg p$

Proof.

1. $p \models \neg q$
2. $p \in I(q)$ [1 ($\neg$I')]
3. $q \in I(p)$ [2 *Symmetry* of incompatibility]
4. $q \models \neg p$ [3 ($\neg$I')]

□

Instead ($\neg$I) gives us both *contructive contraposition* and *classical contraposition*.

Proposition. $p \models \neg q \Leftrightarrow q \models \neg p$

important to be confident with it since we'll need to consider modifications of this plain algebraic structure.

4.2.4 Conjunction

The definition for conjunction is smoother, but it involves some important details.

What is it to be incompatible with *a and b*? Since we are in a set-theoretic framework the answer should readily jump to one's mind that to be incompatible with both a and b is to be incompatible with the set $\{a, b\}$. Thus

Definition 14. **($\wedge$I)** $X \cup \{a \wedge b\} \in Inc$ iff $X \cup \{a, b\} \in Inc$

This seems to be just appropriate: on the one side *Persistence* guarantees that $a \wedge b \vDash a$ and $a \wedge b \vDash b$, because $I(a) \subseteq I(a, b)$ and $I(b) \subseteq I(a, b)$, and on the other side Definition 14 'immediately' establishes that $\{a, b\} \vDash a \wedge b$. But there's an important point worth noticing about this immediacy. Apparently, the notion of incompatibility is not distributive over *conjunction*: something can be incompatible with a and b while being singularly compatible with a and with b. Thus, to quote Brandom's example, a blackberry can't be both red and ripe, while it can be red and it can be ripe.

Proof. ($\Rightarrow$) Suppose $I(\neg q) \subseteq I(p)$. Given $q \vDash q$, by ($\neg$I) we have $q \in I(\neg q)$. But then $q \in I(p)$.

Now suppose for some Z, $Z \in I(\neg p)$. Then by ($\neg$I) we have $Z \vDash p$, i.e. $I(p) \subseteq I(Z)$. But then $q \in I(Z)$, and, by *Symmetry* of incompatibility, $Z \in I(q)$. Thus, in general, $I(\neg p) \subseteq I(q)$.

($\Leftarrow$) Similarly. □

Proposition. $\neg p \vDash q \Leftrightarrow \neg q \vDash p$

Proof. As for the proof of constructive contraposition with ($\neg$I'). □

However, in *Incompatibility Semantics* there's another preliminary reason that makes this choice here irrelevant. In fact Definition 11 forces us to consider *multi-conclusion* inferences, thus

Definition. **($\neg$I')** $X, p \vDash Y \Leftrightarrow X \vDash \neg p, Y$

And that, as Peregrin [115] notices, is enough to establish classical double negation.

Proposition. $\neg\neg p \vDash p$

Proof.

1.	$p \vDash p$	[Ref]
2.	$\vDash \neg p, p$	[1 ($\neg$I')]
3.	$\neg\neg p \vDash \neg\neg p$	[Ref]
4.	$\neg\neg p, \neg p \vDash$	[3 ($\neg$I')]
5.	$\neg\neg p \vDash p$	[2,4 Cut]

□

Consider now the familiar principle for the introduction of conjunction

if $X \models A$ and $X \models B$ then $X \models A \wedge B$.

Notice that while it is obviously valid that

if$I(A) \subseteq I(X)$ and $I(B) \subseteq I(X)$ then $I(A) \cup I(B) \subseteq I(X)$,

there seem to be no obvious way to guarantee that

if $I(A) \subseteq I(X)$ and $I(B) \subseteq I(X)$ then $I(A, B) \subseteq I(X)$.

The key point to observe is that $I(A, B)$ as $I(A \wedge B)$ on the right, can't be construed as $I(A \cup B)$[27]. In other words, when it is on the left, the comma in $I(A, B)$ has to be construed as union, when it is on the right instead the comma has to be construed as intersection. So, why $\{a, b\} \models a \wedge b$ is valid then? It is valid because, by *Peristence*, $I(A \cap B) \subseteq I(A \cup B)$.

Thus the definition of entailment in *IS* guarantees conjunction's familiar meet-ish behavior in the lattice of entailment, which is very good for manageability of the formal system. Let me add that, as in the case of negation, conjunction can be introduced fully recursively from basic vocabulary.

And yet, one might protest, doesn't this fly in the face of the intuition supporting the blackberry example? Honestly, it's hard to deny that. Something more about this will be said below in Section 4.2.6, but in the meanwhile I think two points have to be already highlighted here: (i) Brandom wants the conjunction connective not to be extensionally compositional since he claims that what is incompatible with a conjunction might be not incompatible with either conjunct, and nonetheless (ii) the interaction between a *certain* conjunction and a *certain* incompatibility-entailment validates *Weakening*. Recall, once again, that our aim here is to notice where the joints of the analysis lie.

4.2.5 Modality

4.2.5.1 A first failure

What it is to be incompatible with *necessarily* p? I turns out it's not so obvious to express that in *IS*. Rather than simply stating a definition, what I'm going to tell in this Section

[27] The obvious reason is that it would originate a *tonk* connective: introduction rule of disjunction and elimination rule of conjunction.

Notice the difference with the principle of weakening according to which

if $X \models Y$ then $X \models Y, V$.

That is valid since $\bigcap_{p \in Y \cup V} I(p) \subseteq \bigcap_{p \in Y} I(p)$. Here, crucially, the "," stands for disjunction and is represented as set theoretic union. See Lemma 2.1, B&A (2008) p. 143.

is a story about how it can be establishes. Then there are mainly two reasonings one can follow.

The fist one starts from necessary cases and moves forward. Thus, to begin with,

(i) everything that is self-incompatible is incompatible with *necessarily* p,

but also

(ii) everything that is incompatible with p is incompatible with *necessarily* p.

What else? An idea that is tempting to borrow from common knowledge about modality is that *not* p rules out *necessarily* p. Thus, given the definition of negation in *IS*, the suggestion that might jump to one's mind idea is that something is incompatible with *necessarily* p if it *doesn't entail* p. That, given $(\models_I)$ turns out to mean that

(iii) everything that is compatible with something incompatible with p is incompatible with *necessarily* p.

To put it formally,

Definition 15. **($\boxed{1}$ I)** $X \in I(\boxed{1}\, p)$ iff $\exists Y(X \cup Y \notin Inc \wedge Y \in I(p))$.

Unfortunately this is a wrong suggestion, but it is instructive to consider it in order to improve our comprehension of incompatibility relations[28]. The technical reason why this definition can't be accepted as sufficient is that it would validate both the **S5** axiom $\Diamond p \rightarrow \boxed{1}\, \Diamond p$ and the converses of the Browerian axioms $\boxed{1}\, \Diamond p \rightarrow \Diamond \boxed{1}\, p$: this situation, as it is well known, produces the collapse of modality in the sense that $p \equiv \boxed{1}\, p$[29]. This result sheds some light on what is going wrong with this definition. In the standard framework of possible world, the only models that satisfy both **S5** axiom and the converses of the Browerian axioms without being incoherent are those containing just one single world. That gives representational sense to the collapse of modality. Here the situation is similar, even if, in this case, it is the very semantical definition of necessity that picks up the collapsed case by simply ignoring what differentiate it from the others. To understand why, begin with recalling from Section 4.2.3 that, in *IS*, incompatibilities come into families of contraries. This conversely generates inside a language *families of compatible sentences*

[28] As it often happens one might learn a lot from mistakes. This is especially true this time since it is a mistake of Brandom's: he admitted in Brandom [24], p. 139, he came out at first with this mistaken definition of necessity and spent a lot of time working out an alternative.

[29] See the *Appendix* to this Chapter. This is sometimes equated by Brandom to the situation of Fitch's paradox, but they are slightly different issues. See Section 4.4.1 below.

which do not rule out each other, in the sense that they can be in principle be asserted all together. In a sense that will be made formally precise in Section 4.3 below, these families of compatibles can be construed just as possible worlds. So, to define what is incompatible with *necessarily* p as what is compatible with *not* p is to narrow the application of modal vocabulary within one single family of compatibles, or one single possible world: that makes modal vocabulary superfluous. We know from Section 4.2.3 that *not* to admit *not* p, i.e. to be incompatible with *not* p, is to entail p. In this sense to be incompatible with *necessarily* p would be not to entail p. But this is obviously wrong for, in terms of Kripke's relational semantics, other accessible worlds have to be taken into account. Let me picture a paradigmatic case.

Example 16. Consider the frame in Figure 4.2.3, where straight lines represent incompatibility relations and circled areas represent families of compatibles.

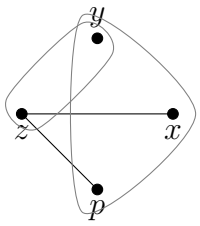


Figure 4.2.3

In this frame x entails p and there's nothing incompatible with p that is also incompatible with x. In this frame x is not incompatible with $\boxed{1}\, p$.

But what to do then? The solution at this point is to try to go beyond the boundaries of one single family of compatibles. An obvious way to do that is to require for something to be incompatible with *necessarily* p, not only that it *doesn't entail* p, but that it is compatible with something that *doesn't entail* p. As Brandom puts it intuitively:

> «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.»[30]

Thus, formally

[30] Brandom [23], p. 129.

Definition 17. (**$\boxed{2}$ I**) $X \in I(\boxed{2}\, p)$ iff $X \in Inc$ or $\exists Y(Y \cup X \notin Inc \wedge Y \nvDash p)$.

This is the definition eventually adopted by Brandom and Aker for the modal operator. So I'll drop the index inside the box when no confusion might arise.

Here is where it is important to consider the second reasoning.

It starts from sufficient cases and moves backwards. According to the intuitive notion of necessity, something is necessary if nothing can prevent it, if nothing is incompatible with it. Thus something is incompatible with *necessarily* p simply if something is incompatible with p. Here thus we meet the suggestion that would lead to a definition that makes modality collapse. We've already analyzed it and we know how to avoid it: it is not enough to take something that is compatible with what is incompatible with p, we have to consider something that is compatible with what doesn't imply p, since the defeasor of p might be in another *family* of compatibles. This is enough to establish the definition

Definition 18. $X \in I(\Box p) \Leftrightarrow X \in Inc$ or $\exists Y(Y \notin Inc \wedge Y \nvDash p)$.

Time to take stock.

We followed two reasonings that led us to two different definitions for the introduction of the necessity operator. Now the crucial question is: in which sense are they different? In fact, B&A (2008) prove that they are equivalent, and this becomes what they call "the basic observation about modal formulae":

Proposition 19. $X, \Box p \vDash \varnothing \Leftrightarrow X \vDash \varnothing$ *or* $\Box p \vDash \varnothing$[31].

It basically says that what is incompatible with p, i.e. what establishes that $\Box p \vDash \varnothing$, has nothing to do with X: in particular it has not to be compatible with X. This is what establishes the simplest kind of necessity as represented by **S5** system.

It's worth pausing to take a deeper look at the proof of this theorem. All the trick is in the ($\Rightarrow$) direction. It says that if something (self-compatible) doesn't imply p then $\Box p$ is self-incompatible. It does that by establishing

$$X \cup \Box p \in Inc \Rightarrow \nvDash p \Rightarrow \Box p \in Inc.$$

The X simply disappears. The proof applies one main observation[32]:

[31] Prop. 3.3 B&A (2008), p. 144.

[32] Apparently the observations that support Proposition 19 are two:

(i) $X \cup Y \notin Inc \Rightarrow Y \notin Inc$

(ii) $\nvDash p \Leftrightarrow \Box p \vDash \varnothing$

$$X \cup Y \notin Inc \Rightarrow X \notin Inc.$$

In fact, this is why $\exists Y(X, Y \not\models \varnothing \wedge Y \not\models p)$ implies $\exists Y(Y \not\models \varnothing \wedge Y \not\models p)$, which is equivalent to $\not\models p$[33].

But the real magic is in the ($\Leftarrow$) direction which 'simply' follows by *Persistence*.

Notice that *Persistence* amounts to the contrapositive of the above principle:

$$X \in Inc \Rightarrow X \cup Y \in Inc.$$

The crucial point is that once X has vanished, *Persistence* makes it re-appear *for free*. In this sense what Proposition 19 shows is that the particular families of compatibiles are *irrelevant* for the meaning of modal operators, because the defeasor of $\models p$ may be anyone. Thus, *a fortiori*, it doesn't matter if what invalidates $\models p$ is somehow indirectly, i.e. transitively, compatible with X, and this is why *S4*-axiom can't fail.

4.2.5.2 To persist is diabolical

In the previous section I suggested that the problem with *Persistence* as a problem of *relevance*. This remark helped me to qualify the problem but now I have to admit I used it also as a bait. Those who might have swallen it, probably resonate to a certain way to construe the logical representation of necessity which has been put forward by Relevant logicians. At the opening of Anderson and Belnap [2] the reasons motivating the whole enterprise of *Relevance Logic* are presented as in a par with C.I. Lewis's[34] complaints for the so called "paradoxes of material implication" in Russell's *Principia Mathematica*: in particular, $p \rightarrow q \rightarrow p$.

But in fact (ii) holds because of (i):

$$\begin{aligned} \not\models p &\Leftrightarrow \exists Y(Y \notin Inc \wedge Y \not\models p) \\ &\Leftrightarrow \exists Y(Y \cup \varnothing \notin Inc \wedge Y \not\models p) \\ &\Leftrightarrow \Box p \models \varnothing \end{aligned}$$

[33] See Prop. 3.2, B&A (2008), p. 143.

[34] The wary reader may still wonder why, if my aim was to compare Brandom's entailment with Lewis's strict implication, I had to step through *Relevance Logic*. Indeed, there are both historical and theoretical reasons, the evaluation of whose force I leave to her. Firstly, Nuel Belnap presently works and teaches at the Department of Philosophy of the University of Pittsburgh, which makes of Brandom his colleague and made of Aker his student: such a brighful logical intellect couldn't but spread some precious seeds whose growth I could admire and testify during the seminar on *IS* to which I attended in Pittsburgh in Fall 2008. Secondly, the collapse of modality in *IS* depends on the purely extensional representation of entailment which, *also* because of the axiom of *Persistence*, can't be avoided: both strict and relevant entailment are ways to limit and amend such a purely extensional framework, but this common path can be better appreciated from the later standpoint of *Relevance Logic*.

Here material implication only represents purely extensional relations between propositional contents contents and this makes any other relation *irrelevant* for the implication of a true proposition[35]. This is what Lewis avoided with his strict implication:

> «In terms of *material* implication, if $pq. \supset .r$ and p is true then $q \supset r$, since $pq. \supset .r :=: p. \supset .q \supset r$. But in terms of *strict* implication, if two premises, p and q, together imply r, and p is true, it does not follow in general that $q \dashv\!\!\supset r$; since $pq. \dashv\!\!\supset .r$ is not equivalent to $p. \dashv\!\!\supset .q \dashv\!\!\supset r$.»[36]

Now, it takes but a moment to realize that Lewis and Brandom work with very akin intuitions on entailment, for compare Lewis's definition of strict implication upon the binary operator "$\circ$" for consistency,

$$p \dashv\!\!\supset b =_{Def} \neg(p \circ \neg q),$$

with Brandom's Definition 11, which, given Definition 13 for negation, is equivalent to

$$p \models_I q \Leftrightarrow p \in I(\neg q).$$

And yet there is a macroscopic difference in the two approaches. While Lewis construes strict implication as the proper representation of the necessary character of entailment and then procedes to define material implication in a different way in order to distinguish them, Brandom accepts his definition as of *the* only notion of implication in his system and then procedes to define modal operators.

The first crucial point to notice, in this sense, is that the idea of incompatibility as a directly modal notion translates into a system of material implication[37].

In fact it is trivial to prove in *IS* $\models_I p \rightarrow q \rightarrow p$, but it is important to see why. Now $\models_I p \rightarrow q \rightarrow p$ follows from $p, q \models_I p$, which is valid in *IS*, by two applications of Theorem 3.3 in B&A (2008), which is nothing but *Deduction Theorem*. This is a typical situation you want to avoid if you care about the issue of *relevance*, and the temptation to see it as a stark choice between two obvious principles of implication: *Deduction Theorem* and *Reflexivity*. But such a temptation should be resisted since there is more than meets the eye. A quick look to algebras for substructural logics could help[38]. Let me borrow just the essential to

[35] See Lewis and Langford [86], p. 85, and Anderson and Belnap [2], pp. 3-5.

[36] Lewis and Langford [86], p. 165 (where Lewis's horseshoe for material implication corresponds to our "$\rightarrow$").

[37] In what follows I rely on B&A (2008) representation theorem for *IS*.

[38] I suggest Dunn [42], Dunn and Hardegree [46], which are directly connected with the topic.

make my point. Consider a partially ordered groupoid $\langle S, \leqslant, \circ \rangle$ and introduce a binary operation "$\rightarrow$" such that it satisfies the following property usually named *Residuation*[39]:

$$a \circ b \leqslant c \text{ iff } b \leqslant a \rightarrow c.$$

In general, given a poset $\langle S, \leqslant \rangle$ and a pair of function (f, g) on it, (f, g) is a *residuated pair* if for any $A, B \in S$, $f(A) \leqslant B$ iff $A \leqslant g(B)$. For instance, modal operators form a residuate pair $(\Diamond, \Box)$ in $\mathbf{B}$ because $\Diamond a \leqslant b \Leftrightarrow a \leqslant \Box b$[40].

Now this property is important precisely because it shows the relations holding between operations and in general functions on algebras. For what concerns us here it enables us to see that deduction theorem does nothing but display in an implicational lattice the relation between "$\rightarrow$" and that particular sort of conjunction which is ",":

$$a, b \vDash c \text{ iff } b \vDash a \rightarrow c.$$

In other words, material implication residuates extensional conjunction. Notice there's nothing wrong with this. What we want to avoid is that material implication residuates also intensional conjunction, or fusion, "$\circ$". That would force us to accept *Augmentation*, i.e. $p \circ q \vDash p$, which is unwanted for fusion – compare with "if p is compatible with q then p is true".

To sum up, behind the troubling choice we seemed to be forced above there's the teanable position of requiring both that strict implication doesn't residuate extensional conjunction and that intensional conjunction doesn't validate *left lower bound*, i.e. $p \circ q \leqslant p$. This is what both the systems of strict implication and system $\mathbf{R}$ of *Relevance Logic* require, by imposing fusion not to be idempotent[41]. But Brandom does not prevent that in *IS*.

---

[39] I assume here that fusion is *commutative* so I don't have to distinguish between left and right residuation.

[40] It's easy to see that with a relational semantics. In any frame (W, R) we have $\Diamond_{\downarrow} A \subseteq B$ iff $A \subseteq \Box B$. In fact $\{w \mid \exists a(aRw \wedge a \in A)\} \subseteq \{w \mid w \in B\}$ iff $\{w \mid w \in A\} \subseteq \{w \mid \forall a(wRa \Rightarrow a \in A)\}$, and since R is symmetric we have $\{w \mid \exists a \in A \text{ and } R(a, w)\} = \{w \mid \exists a \in A \text{ and } R(w, a)\}$, that is$\Diamond_{\downarrow} A = \Diamond A$.

[41] This maybe requires more few words. Probably neither Lewis nor Belnap and Anderson ever wrote that fusion is not idempotent, and yet to require that is enough both for strict implication and for system $\mathbf{R}$ to avoid the collapse into material implication *as described above*. Since I refer to his work below, I have to note here that Read [131], p. 128, explicitly contrasts $\mathbf{R}$'s fusion with Lewis's consistency operator and claims that the latter validates *Augmentation*. This however, as far as I can see, needs some clarification. Lewis only accepts the *consistency postulate*: $\Diamond(p \wedge q) \rightarrow \Diamond p$. But this is $p \circ q \dashv p \circ p$, which doesn't imply $p \circ q \dashv p$ – to be distinguished from $p, q \dashv p$ – if the consistency operator is not idempotent. And in fact the consistency operator is not idempotent, because while $p \dashv p \circ p$ is valid (Lewis and Langford [86, Th. 17.6]), $p \circ p \dashv p$ is not valid. What is valid is $\neg(p \circ p) \dashv \neg(p \circ \neg q)$, which implies $\neg\Diamond p \dashv p \dashv q$ (Lewis and Langford [86, Th. 19.74]): this might be bad for relevance logic but doesn't affect modal logic.

4.2.5.3 Towards full kripkean modality

If all the problems come from the axiom of *Persistence*, why don't we just drop it?

Indeed there are encouraging reasons to believe that would be a good idea. Among these, there's the quite promising fact that all the characterizing formulae of normal modal systems would be easily provable anyway, from rule of *Necessitation* to K, and so on. But what's best is that the validity of each single theorem would depend on the expected properties of compatibility relations: T-axiom would be valid for compatibility is *reflexive*, B-axiom would be valid for compatibility is *symmetric*, $S4$-axiom would be valid if and only if compatibility were *transitive* (which, presumably, is not)[42].

But there are some discouraging facts as well. The first one has nothing to do with modality, but it is crucial for the point about compositionality we defended in Section 3.4. The axiom of *Persistence* supports one key step in the proof of the series of Lemmas with which B&A (2008) establish that even for *infinite* languages the semantic interpretation of complex formulae is recursively computable from that of simpler ones. In Section 4.2.6 below I'll add some comment about this. What is enough to remark now is that in the absence of such a proof our defense seriously risks to fall down. This obviously doesn't mean that the same result couldn't be proved without *Persistence*. The second fact is possibly even more discouraging. Suppose in fact we could really drop the axiom of *Persistence* without any unacceptable lost. Well, the bad news is that that wouldn't be enough to avoid its effects in *IS*. Consider the most insecapable and apparently innocuous principle of an entailment relation, *Reflexivity*. The fact is that where sets of sentences are considered, as in *IS*, it becomes: $X \models a$ iff $a \in X$. This immediately gives a form of *Augmentation* since it couldn't be denied that $X, a \models a$. But things are even worse. Once this is acknowledged, *Weakening* on the left can be re-established in its full generality:

Lemma. $X \models p \Rightarrow X, Y \models p$[43]

Proof. Assume $X \models p$. We show $X, Y \models p$ for arbitrary Y.

Suppose $Z \in I(p)$. But, as a consequence of *Reflexivity*, $I(p) \subseteq I(Y, p)$.

Thus $Z \in I(Y, p)$.

Then by *Partition*, $Z \cup Y \in I(p)$. Then $Z \cup Y \in I(X)$ by *Entailment*.

Then $Z \in I(X, Y)$ by *Partition* again. Thus $X, Y \models p$. □

[42] For further details see below in the Appendix, Section 4.4.3.

[43] A correspondent proof was originally provided by Alp Aker.

The moral to be drawn is that *IS* is too deeply entrenched in set theoretic extensional fremework. Again, this doesn't mean that it can't be amended so to represent *directly* an intensional conjunction: it means that modifications have to be deep.

While I do really believe that the path of a non set-theoretic incompatibility semantics is practicable, here I take this two facts as a knockout combination. Limits of knowledge, space and courage strongly suggest me to give up on this. What's more important, I don't really need to open that path in order to say all I want to say in this work. I can point at another way out of the *empasse* with modality in *IS*, which actually bends towards more standard intuitions and formal machineries. In a sense, it consists in the choice to embrace Lewis's idea of a strict notion of implication to represent necessity of entailment besides material implication in *IS*. Thus, while taking the entailment of *IS* as representing material implication, introduce a further level of compatibility, call it provisionally "compossibility", through an appropriate binary operator, say e.g. the fusion operator "$\circ$" itself. What matters, in this sense, is that it is not residuated by standard *IS*'s entailment and that *Persistence* doesn't apply to it. Once we have this operator for compossibility we define a new notion of entailment that may behave strictly as Lewis did. Then we can simply follow the stream of Lewis's proofs of theorems for his systems. Notice we shouldn't bother about proving the validity of the characterizing axioms for the different systems, since we would be just assuming them to obtain different theorems representing features of the modality expressed by our new strict implication. However, in this sense, an intutive interpretation of *compossibility* is lacking. That we'll be provided below in Section 4.3. In fact, in another sense, this wayout consists in the introduction of the notion of possible world in *IS*.

4.2.6 Some provisional summary

4.2.6.1 Compositionality again

As I announced at the beginning of this section I hope I can now convincingly wipe off all the representational dust that has (been) attached to this cognitively crucial but formally innocuous problem. Let's begin by noticing that, as I took care to point out, all logical connectives of *IS* seem quite obviously to fail compositionality. The semantic interpretant of any sentence p is the set of sentences incompatible with it, i.e. $I(p)$. For what concers *negation*, in order to to compute the semantic interpretant of $\neg p$, as we saw, we can't simply take the complement of $I(p)$ but we have to consider what is implied by all the elements of $I(p)$. For what concers *conjunction*, as the example of blackberries testifies,

something can be incompatible with $p \wedge q$ without being incompatible with p nor with q – even if *Persistence* prevents the contrary. For what concerns *necessity*, in order to determine the semantic interpretant of $\Box p$ it is required to take into account everything that is to be compatible with something that doesn't entail p. Now, what all this *does* testify *for* is the *holistic* character of *IS*: the semantical interpretation of compound expressions *can't* be recursively determined *only* from the semantical interpretations of the component expressions. What all this *does not* testify *against* is the *compositional* character of *IS*: the semantical interpretation of compound expressions can be *recursively determined* from the semantical interpretations of other expressions. There's a precise sense to say this and it has been already introduced as Theorem 6 in Section 3.4.3.1: given any partial algebra of strings and a semantical interpretation, there exists a set of functions that recursively compute meanings of complex strings which (i) are compositional and (ii) allow to recover the simpler semantic interpretations. This, it was said, is a very weak result since it doesn't commit to any sensible formal representation of meanings: one thing is to establish that there exists such a set of functions, quite another thing is to find an instance that satisfies the condition we may want to impose on a semantical interpretation. The fact is the logical operators of *IS* are such an instance: Aker proved that given a language $\mathcal{L}$ and an incompatibility frame Inc, there exists and it is unique an incompatibility frame Inc' for a language $\mathcal{L}'$ such that $\mathcal{L}'$ *recursively extends* $\mathcal{L}$ and Inc' is *inferentially conservative* w.r.t. Inc, in the sense of $X \models_{Inc} Y \Leftrightarrow X \models_{Inc'} Y$[44]. In Section 4.4.4 I'll prove an equivalent result also for an extended version of *IS*.

4.2.6.2 Main features of *IS*

At last let me provide a very brief summary of the choices made by Brandom and Aker in the definition of *IS* and of the consequences they have. Let's start with considering the general framework: *IS* is based on the semantic primitive of incompatibility and a technical notion of consequence relation is defined in terms of incompatibility relations. Then let's get on considering the choices of Brandom and Aker.

First, incompatibility relations are characterized as *Symmetric* and *Persistent*. This – especially *Persistence* – induces a *classical* behavior on the consequence relation.

Second, incompatibles are contraries, thus contradictories are defined as those contraries that are entailed by all the rest. This, as we saw, is equivalent to say that to be compatible with both a sentence and its negation is to be self-incompatible.

[44] See B&A (2008), pp. 147-155.

Third, conjunction is supposed to be intensional, but entailment is defined in such a way that incompatibilities are checked distributively w.r.t. the *set* of consequents[45]. This is enough to establish the standard extensional behavior of conjunction. We are now in the position to add some more words about this. The use of extensional set theoretical operations of union and intersection in the definition of entailment rules out the intensional character of incompatibility. Extensional conjunction is strictly stronger that intensional conjunction,

$$p \wedge q \Rightarrow p \circ q.$$

In particular, while extensional conjunction is *augmentable*,

$$\neg(p \circ q) \Rightarrow \neg(p \circ q \wedge x),$$

intensional conjunction is not,

$$\neg(p \circ q) \not\Rightarrow \neg(p \circ q \circ x)^{46}.$$

In this sense the complaints of those who would to stick to the intuition of the blackberry example were not out of place.

Fourth, B&A (2008) have a proof that their definition of modal operators is enough to obtain **S5** modality for *IS*'s entailment. But that definition in *IS* establishes that in the evaluation of the incompatibility sets of modal formulae families of compatibles are irrelevant. As a consequence, even if compatibility relations are *not transitive* there's be no way to ascertain the failure of such transitivity in the semantic evaluation of modal formulae. Together with *Symmetry* of compatibility, this explains why nothing less than **S5** is obtained. In this sense their result looks less surprising. A first lewisian way out to this situation has been proposed, and we're about to see how it can be employed to improve Brandom's expressive project of norms in modal vocabulary.

4.3 Possible worlds in *IS*

4.3.1 Definition

So far I've been talking loosely and intuitively about 'families of contraries' and correspondent 'families of compatibles'. Now, intuition is a precious advisor but a dreadful leader.

[45] Consider also Peregrin's remark about multi-conclusions entailment and classical logic cited in Note 26 above.

[46] See Lewis and Langford [86], p. 155, but also Read [131], pp. 31-34.

Let me thus formally specify my loose talk. Fortunately I don't have to look far: if there is an idea deeply entrenched in the whole modern reflection on modality since Leibniz himself is the notion of *compossibility*. As Leibniz himself explains to Bourguet:

> "[N]ot all possibles are compossible. Thus, the universe is only a certain collection of compossibles, and the actual universe is the collection of all existing possibles, that is to say, those which form the richest composite. And since there are different combinations of possibilities, some of them better than others, there are many possible universes, each collection of compossibles making up one of them." (GP III 573/L 662)

This led to the standard definition of possible worlds as *maximally consistent sets of propositions*. This idea can be easily adopted in Brandom's framework to make sense of families of compatibles:

Definition 20. (Possible World)

$$PW_{Inc} =_{Def} \{S \mid S \notin Inc \text{ and } \forall X \subseteq \mathcal{L}(X \cup S \notin Inc \Rightarrow X \subseteq S)\}$$

Thus defined possible worlds are maximally *compatible* sets of propositions, in the sense that any added proposition would make the sets self-incompatible.

Peregrin [115] notices that a useful remark immediately follows. As you may recall from Section 4.2.2, one of the reason of discontent with the notion of incompatibility entailment was that it seemed to drop, together with the notion of *Truth*, also the idea that one main purpose of a consequence relation is to represent the preservation of a certain semantically relevant status. But now, consider what it means to be true in a possible world in the framework just defined. Given the definition of possible worlds as maximally coherent set of propositions, for a proposition to be true in a possible world is for it to be part of that world, in the standard sense that it is compatible with that world-story. Notice then that it is equivalent to say that a sentence p is true in a possible world $w \in PW_{Inc}$, that $p \in w$, that everything compatible with w is compatible with p and that $w \models_I p$.

It's worth following up here with an interesting observation. In standard truth functional semantics it is possible to define the set $AT^{\mathcal{L}}$ of *atoms* of $\mathcal{L}$:

Definition. An atom of $\mathcal{L}$ is a sentence of the form

$$\alpha = p_1^{\epsilon_1} \wedge \ldots \wedge p_n^{\epsilon_n}$$

where $p_1, \ldots, p_n \in \mathcal{L}$ and truth values $\epsilon_i \in \{0, 1\}$.

These atoms are nothing but Carnap's state descriptions. It is common knowledge that *atoms* $AT^{\mathcal{L}}$ of a language $\mathcal{L}$, correspond one-to-one to *valuation functions* V on $\mathcal{L}$:

Remark. For any atom $\alpha \in AT^{\mathcal{L}}$ there is only one valuation $v_\alpha \in V$ s.t. for any p_i in α, $v_\alpha(p_i) = e_i$. For any valuation $v \in V$ there is only one atom $\alpha \in AT^{\mathcal{L}}$ s.t. $\alpha = p_1^{v(p_1)} \wedge \ldots \wedge p_n^{v(p_n)}$.

In other words, one possible world one valuation, one valuation one possible world. This equivalence crucially interacts with the intuitive representation of models employed in the tarskian definition of logical consequence. In fact the models for a certain sentence s can be represented as

$$M_s = \{\alpha \in AT^{\mathcal{L}} \mid \alpha \models s\} = \{\alpha \in AT^{\mathcal{L}} \mid v_\alpha(s) = 1\}.$$

Now, nothing like this is the case in *IS*. In fact, in *IS* it is possible to define models for s in terms of possible worlds as

$$M_s = \{w \in PW_{Inc} \mid w \models_{Inc} s\},$$

But there are no valuations in *IS* and it makes no sense to select such a set in terms of different *incoherence relations*. This is because there's no way to characterize a specific set of sentences (for instance those one commits to) in terms of a *frame*: one incoherence frame many possible worlds, many possible worlds one incoherence frame[47]. It is hard to say how much important this observation is for the proper evaluation of modality in *IS* and in general for Brandom's very account of normativity[48]. For the moment it's enough to see what light it sheds on Brandom's claim that incompatibility is a *directly* modal notion:

[47]Notice that this has nothing to do with the absence of the notion of *Truth*: we've just learnt that *compatibility* in a frame plays the role of *Truth* in a model. As Aker points out, this corresponds to the notion of *satisfiability* in *IS*. In fact he also shows how to define an *incompatibility frame* in order for an *atom* to be compatible in it.

Remark. For any atom $\alpha \in AT^{\mathcal{L}}$ there is an incompatibility frame Inc_α s.t. $X \in Inc_\alpha$ iff for some finite $X' \subseteq X$, $\bigwedge X'$ is not in α.

See Theorem 1.1 in Appendix 3 of B&A (2008), p. 162. This is obviously a basic prerequisite in order to obtain a completeness result.

However, and here's my point, it's not possible conversely to define a single atom in terms of an incompatibility frame: as it's easy to see if $p \cup q \in Inc_\alpha$ then either p but not q, or q but not p can be pushed into α. That is enough to break the correlation between *incompatibility frames* and models.

[48]I'll come back to this in Section 5.5.2.2 below and in the conclusion in Chapter 6.

the whole articulation of possible worlds lies within a single incompatibility frame.

4.3.2 How to Kripke *IS*

Now that a definition of possible world and a notion of truth have been derived, it is obviously tempting to try to do better than Brandom by following Kripke's well-trodden path. The idea is to adopt incompatibility relations as a primitive and to redefine all the notions of relational semantics for modality, in order to reproduce the standard kripkean framework inside Brandom's *IS*. In other words the idea is to define a tranlsation manual from *IS* to Kripke's relational semantics for possible worlds. That this can be done has already been showed by Göcke et al. [59] and Peregrin [115].

To begin with, recall that, metaphysical issues apart, the main problem with the reception of the idea of possible worlds as maximally coherent sets of sentences within the standard truth-functional semantics was that to treat compossibility as consistency in a strictly bipartite evaluation of semantic contents simply is to crush necessity on logical validity. Kripke's relational semantics, by pivoting on the primitive notion of *accessibility*, disentangled modal possibility from logical possibility and opened the doors to the modern analysis of modality. Our next goal then is to define the relation of *accessibility* with the resources of *IS* and the standard definition of possible worlds. Fortunately, the trick to obtain accessibility is common knowledge. Suppose you have a space of possibilities already defined in terms of possible worlds, then, by reversing the basic intuition about necessity as truth in any accessible world, a binary accessibility relation R between worlds can be defined by taking w_1 to be accessible from w_2 if everything which is necessary in w_1 is true in w_2:

$$w_2 R w_1 \text{ iff } \{p \mid \Box p \in w_1\} \subseteq w_2.$$

In terms of Brandom's definition of the necessity operator, that is to say that for any $w_1, w_2 \in PW_{Inc}$, w_1 is accessible by w_2 if and only if for any $p \in w_2$ there is a subset $X \subseteq w_1$ such that $X \cup p \notin Inc$[49]. Formally,

Definition 21. (Compossibility)

$$w_2 R w_1 \text{ iff } \forall p(w_2 \vDash p \Rightarrow \exists X(X \subseteq w_1 \wedge X \cup p \notin Inc))$$

[49]Notice that version (iii) of the introduction of necessity is adopted here. This is acceptable now since with the accessibility relation we gain another parameter to play with in order to evaluate compatibility and avoid the collapse of modality.

As Peregrin notices, in *IS* this amounts to treat *accessibility* as a second-level weaker compatibility: while any two possible worlds are incompatible as a whole, it might well be that any piece of the one is compatible with *some* piece of the other. This definition of accessibility very aptly fits with Brandom's own treatment of modal operators. Recall that the basic idea is to make explicit what it means to be incompatible with *necessarily* p by showing what doesn't prevent *not* p. Thus, we can simply adapt this idea here by saying that something is incompatible with *necessarily* p if and only if any possible world which contains p is *compossible* with a possible world which doesn't contain p. Formally,

Definition 22. (PW-□I)

$$X \in I(\Box p) \text{ iff } \forall w_1(w_1 \vDash X \Rightarrow \exists w_2(w_2 R w_1 \wedge w_2 \nvDash p))$$

By mimicking Peregrin [115]'s labels, I'll call this semantics which implants kripkean framework within Brandom's *IS*, *Extended Incompatibility Semantics* (*EIS*). We can now conclude this Section with an analysis of the results of *EIS*.

Notice that compossibility inherits all the properties of compatibility, in particular, for what concerns us here, it is *reflexive* and *symmetrical*. Thus, it is easy to show that *IS* with (PW-□I) validates axioms **T** and **B**[50]. Quite surprisingly, however, **S4** axiom fails. The reason is that accessibility relation as defined above may be *not transitive*. A counterexample has been provided already by Göcke et al. [59], but I propose here another one which I find possibily easier to grasp.

Example 23. So consider three possible worlds: $w_1 = \{p, a\}$, $w_2 = \{b, p\}$, $w_3 = \{b, c\}$. According to Definition 21, any two possible worlds are compossible if they share at least one element. Let me put it graphically:

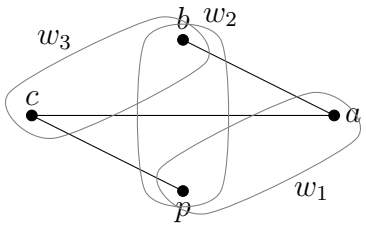


Figure 4.3.1

As before, in Figure 4.3.1 circled areas represent possible worlds and straight lines represent incompatibilities.

[50]For the detailed proofs I refer to Göcke et al. [59].

Now **S4** axiom establishes that $\Box p \models \Box\Box p$. So, to show that **S4** axiom fails in this model for *IS* with (PW-$\Box$I) we have to find a sentence p s.t. $\exists X(X \notin I(\Box p) \wedge X \in I(\Box\Box p))$. Thus take p and $\{x\}$. Now, according to (PW-$\Box$I) $\{x\} \notin I(\Box p)$ since it is true in w_1 and all the worlds compossible with w_1, namely w_2 and $w_1$itself, contain p. But $\{x\}$ is incompatible with $\Box\Box p$ because w_3, which is compossible with w_2, doesn't contain p, so $w_2 \in I(\Box p)$.

4.3.3 What it means to Kripke *IS*

From the failure of **S4**, both Göcke et al. [59] and Peregrin [115] draw the conclusion that by reproducing Kripke's relational semantics inside *IS* with *EIS* Brandom could partially overcome the limitation of his basic semantics which could only represent the logical modality of **S5**. I have to admit I'm not really content with this interpretation the authors give of their own work. In this last Section I want to claim that the application of Kripke's relational semantics to *IS* doesn't pull any rabbit out of a hat, rather it simply makes explicit some features of modality that remain implicit in *IS* after all. Does that mean that Kripke's framework provides a better semantic metavocabulary for incompatibility? Let's see.

Before we even begin with the analysis, it is crucial to ask whether, even with this kripkean implant, modality collapses in *EIS*. The answer is "no". First, if *compossibility* is not *transitive* then *EIS* doesn't verify *S5*-axiom, and that is enough to prevent the collapse even if they would verify the converses of the browerian axiom. Second, as I prove in Section 4.4.2 in the Appendix *EIS* in fact doesn't verify the converses of the browerian axioms even. These results were expected: the relation of *compossibility* produces that second-level compatibility that doesn't verify *Persistence*. In other words, *compossibility* in Definition 22 'gently' forces compatibility to behave as an intensional conjunction with relation to modally robust implication: it corresponds to the introduction of an operator in the boolean algebra of material implication, where the residuated pair $(\Diamond, \Box)$ is obtained[51]. This is common knowledge about modality, but what makes it relevant for our analysis here is that it might raise serious worries about the fulfillment of Brandom's purposes. In fact Brandom declares that with his incompatibility semantics he aims to

> «explore the relations between normative and modal vocabulary [...], showing how normative vocabulary can serve both as a *pragmatic metavocabulary* for modal vocabulary and as the basis for a *directly* modal formal semantics for

[51] See Dunn [44].

> ordinary empirical vocabulary that does not appeal in any way to a notion of truth.»[52]

Thus one may wonder whether the implant of possible worlds, while conveniently required from a logical point of view, is a step back from the expressive results of *IS* itself. The intended benefit of *IS* would have been the possibility to deploy a directly modal notion of entailment and substitute the metaphysically laden semantic primitive of *truth in a world* with the pragmatically entrenched one of incompatibility. If instead it would be shown that a further operator is nonetheless required to obtain the same expressive results of Kripkean modal logic, then *IS* would need two primitives rather than the one of the standard kripkean framework and its value would quickly get lost. In other words one may wonder whether the indirect path of *compossibility* amounts after all to declare the failure of the Brandom's project in formal semantics. But the answer, again, is "no". To see why, it is enough to get clear of what "directly" means there. In fact there are at least two senses in which the notion of *incompatibility* can be said to *directly* provide a modal semantics. In a first more technical sense, incompatibility would directly provide a modal semantics if incompatibility entailment is itself modally robust like Lewis's strict implication. In a second apparently less sharp sense incompatibility would directly provide a modal semantics if nothing prevents modal vocabulary to be defined in terms of incompatibility semantics. It is only this latter sense which is relevant for Brandom's project:

> "by [directly modal semantics] I mean one that does not approach modality by beginning with a more basic semantic notion of *truth*."[53]

And in this second sense *EIS* does satisfy Brandom's claim. The relation of *compossibility* and the modal operators in *EIS* are defined only in terms of *compatibility*, so to make explicit the relations between sets of sentences belonging to different families of compatibles. Compare with kripkean relational semantics: first the notion of *Truth* is appealed to in order to define entailment, and *then* the relation of *accessibility* is independently introduced with the apparatus of possible worlds in order to define modal operators. In other words the gap between *IS* and *EIS* is not notional but expressive. But then this second sense in which the notion of *incompatibility* can be said to *directly* provide a modal semantics turns out to be no vague at all. On the one side, it is required that the abilities sufficient to deploy modal vocabulary can be elaborated from those of non modal vocabulary, in

[52] Brandom [23], p. 116 (my emphasis).
[53] Brandom [23], p. 124.

the distinguished sense of Brandom [23] that by knowing how to deploy the vocabulary of incompatibility entailment one already knows all she needs to know in order to deploy *Incompatibility Semantics*. On the other side it is required that the modal vocabulary of *Incompatibility Semantics* is *inferentially conservative*[54] with respect to the syntactically simpler vocabulary of incompatibility-entailment it extends. The first requirement guarantees that no cognitive limit prevents the semantics of syntactically complex sentences to be computed in an holistic semantics, the second requirement guarantees that this computation doesn't introduce any modification of previously computed meanings, together they guarantee the recursiveness of the semantics. Alp Aker proved how both requirements can be met by *IS*.[55]. But, as a matter of fact, such a proof is easily adaptable to *EIS*[56], and that guarantees the existence of an algoritm to elaborate *EIS* as an inferentially conservative extension from languages containing no modal expression.

Now we can go back to our comparison between standard kripkean semantics and *EIS*. The task of a formal semantics for incompatibility is to discharge all the burden of the semantic analysis from the notion of *Truth*, which has been proved to lie against the fallacious grounds of representationalism, and move it into the normative analysis of linguistic practices. *EIS* provides the bridge to connect the modal vocabulary of semantics with normative vocabulary of pragmatics and does that by pivoting only on the notion of *incompatibility*. In Brandom's terms, *EIS* shows how the vocabulary of norms is a pragmatic metavocabulary for the modal vocabulary of semantics. In this sense the expressive advantage of *EIS* over Kripke's relational semantics is patent. Let me try to illustrate with an example. In the previous Section it was claimed that *EIS*'s sort of modality is the browerian one of system **B**. Does that mean that according to *EIS*, or in general according to Brandom, system **B** represent the *real* modality? This question might be tricky but the answer negative. Recall from Section 4.1.1 that *EIS*, as a formal semantics, is a metavocabulary to make explicit normative contents implicit in linguistic practices. Kripke's relational semantics for modal logic is a similar modal metavocabulary. The decisive advantage of *EIS* over Kripke's relational semantics is that it is based on an independent normative analysis of linguistic practices, which provides the pragmatic metavocabulary to express it. In fact, even if *EIS* can't vindicate *incompatibility* as a *technically direct* modal notion, the expressively direct connection with such a normative analysis remains. This advantage pays back not only because it cuts off metaphysical issues about possible

[54]See Belnap [6].
[55]See Section 5 of Appendix I in B&A (2008), Brandom [23], pp. 147-155.
[56]See Section 4.4.4 in the Appendix for a few more comments on this.

worlds – which is not a faint result, bt the way –, but also because it puts some normative flesh on the algebraic bones of the accessibility relation. So, is **B** the *real* modality? In Kripke's relational semantics the answer is: "Well, let me check if *accessibility* is reflexive and symmetric but not transitive." How can you tell that? In *EIS* the answer is: "Well, let me check if *compossibility* is reflexive and symmetric but not transitive." How can you tell that? Look at normative linguistic practices.

4.4 Appendix

The purpose of this Appendix is simply to provide some technical details to support what was claimed in the main text about the collapse of modality in *IS* and related systems. First, in Section 4.4.1, I'll explain what the collapse of modality is and how it generates. Then in Section 4.4.2, I'll prove that *EIS* doesn't suffer from such a collapse. In Section 4.4.3, I'll collect the proofs of some theorems that suggest how a system of incompatibility semantics without *Persistence* would behave in the representation of different systems of modality. Eventually in Section 4.4.4 I'll prove a theorem esteblishing the semantic recursiveness of *EIS*. Please notice that the index "I" of the incompatibility frame in the turnstile will be omitted where obvious.

4.4.1 The collapse of modality

As we saw, the semantic system of *EIS* aboundantly lies against the results of Peregrin [115] and Göcke et al. [59]. Now, in his replies to Göcke et al. [59], Brandom warns them against the risk of obtaining a system of degenerate modality. He wonders whether their definition of necessity might validate the converses of the browerian axioms, thus pulling down modality. As I argued above it is not so, but no proof has been provided yet. What Brandom has in mind is an adaptation of the so called Fitch's Paradox to general modal logic: any definition of incompatibility with $\Box p$ that would validate the converse of B-axiom, i.e. $p \vDash \Diamond\Box p$, would make modality collapse. Let me illustrate the well known result.

Theorem 24. *Fitch Paradox*

Proof. Suppose that not everything is necessary: $\exists p(p \wedge \neg\Box p)$.

But $p \wedge \neg\Box p \Rightarrow \Diamond\Box(p \wedge \neg\Box p) \Rightarrow \Diamond(\Box p \wedge \Box\neg\Box p) \Rightarrow \Diamond(\Box p \wedge \neg\Box p) \Rightarrow \bot$.

Thus $\forall p(p \rightarrow \Box p)$. □

This obviously leads to the collapse of modality in any system that validates T-axiom.

However some clarifications are in order. Fitch's paradox was originally formulated with relation to epistemic logic as a paradox of knowability: it establishes that, if every truth is know*able* – our B converse, $p \rightarrow \Diamond Kp$ – then the existence of any unknown truth implies, contradictorily, that every truth is in fact known. Now, in a wider modal framework this *per se* doesn't imply the collapse of modal operator in the sense of $p \equiv \Box p$. Indeed, one of the classical wayouts to Fitch's paradox in epistemic logic was to reject quantified modal logic.

But there is a straighter way to obtain the collapse of modality to PC in the sense that $p \equiv \Box p$: any system that validates both **S5** axiom and the converses of the browerian axioms **M** (I name them after Hughes and Cresswell [72]) produces such a collapse[57]. In his replies to Göcke et al. [59] Brandom admits he originally ended up with a definition for the introduction of necessity operator that produced the collapse of modality. As we saw from Section 4.2.5.1 above, what Brandom' original collapsing definition was:

($\boxed{1}$ **I**) $X \in I(\Box p) \Leftrightarrow \exists Y(Y \notin I(X) \wedge Y \in I(p))$.

I'm now going to show how this definition produces the collapse of modality. First I'll prove it is enough to validate $S5$-axiom. In order to do that I'll simply reproduce for (iii) the proofs already provided in B&A (2008). Then I'll prove it also validates the converses of the browerian axioms, **M**. These together imply the collapse of modality.

Let me thus prove a couple of lemmas first.

The first one establishes the standard behavior of necessity on both sides of the turnstile:

Lemma 25.
a) $\Box p \vDash Y \Leftrightarrow \vDash Y$ *or* $\nvDash p$
b) $X \vDash \Box p \Leftrightarrow X \vDash$ *or* $\vDash p$

Proof. Part a:

$$\begin{aligned} \Box p \vDash Y &\Leftrightarrow \forall Z(Z \in I(Y) \Rightarrow Z \in I(\Box p)) \\ &\Leftrightarrow \forall Z(Z \in I(Y) \Rightarrow Z \nvDash p) \\ &\Leftrightarrow \forall Z(Z \notin I(Y) \vee Z \nvDash p) \end{aligned}$$

Thus $\vDash Y$ or $\nvDash p$.

Part b:

$$\begin{aligned} X \vDash \Box p &\Leftrightarrow \forall Z(Z \in I(\Box p) \Rightarrow Z \in I(X)) \\ &\Leftrightarrow \forall Z(Z \nvDash p \Rightarrow Z \in I(X)) \\ &\Leftrightarrow \forall Z(Z \vDash p \vee Z \in I(X)) \end{aligned}$$

[57] See Hughes and Cresswell [72], p. 131.

Thus $X \models$ or $\models p$. □

Now Aker's proof of *S5*-axiom can be applied[58] also to Definition ($\boxed{1}$ I).

The second Lemma describes the behavior of modal operator in a way that shows the role of families of compatibles as possible worlds.

Lemma 26. $X \models \Box p \Leftrightarrow \forall Y(Y \cup X \notin Inc \Rightarrow Y \models p)$

Proof. Just apply (iii):

$$\begin{aligned} X \models \Box p &\Leftrightarrow \forall Y(Y \in I(\Box p) \Rightarrow Y \in I(X)) \\ &\Leftrightarrow \forall Y(Y \models \varnothing \text{ or } \exists Z(Z \cup Y \notin Inc \wedge Y \in I(p)) \Rightarrow Y \in I(X)) \end{aligned}$$

If $Y \models \varnothing$ then $Y \in I(p)$.

Otherwise, $\forall Y(\exists Z(Z \cup Y \notin Inc \wedge Y \in I(p)) \Rightarrow Y \in I(X))$. Which amounts to $\forall Y(Y \not\models p \Rightarrow Y \in I(X))$, i.e. $\forall Y(Y \cup X \notin Inc \Rightarrow Y \models p)$. □

So far so good, but then, unfortunately, we can prove also the converses of browerian axioms, **M**.

Proposition. (M) $\Box\Diamond p \models \Diamond\Box p$

Proof. By definition $\Diamond p \Leftrightarrow \neg\Box\neg p$, thus what we want to prove is $\Box\neg\Box\neg\Box p \models \neg\Box\neg\Box\Box p$. Thus,

1. $\Box\neg\Box\neg\Box p \models \neg\Box\neg\Box\Box p$;
2. $\forall X(X \in I(\neg\Box\neg\Box\Box p) \Rightarrow X \in I(\Box\neg\Box\neg\Box p))$;
3. $\forall X(X \models \Box\neg\Box\Box p \Rightarrow X \not\models \neg\Box\neg\Box p)$;
4. $\forall X(\forall Y(Y \cup X \notin Inc \Rightarrow Y \models \neg\Box\Box p) \Rightarrow X \models \Box\neg\Box p)$;
5. $\forall X(\forall Y(Y \cup X \notin Inc \Rightarrow Y \in I(\Box\Box p)) \Rightarrow \forall Z(Z \cup X \notin Inc \Rightarrow Z \models \neg\Box p))$;
6. $\forall X(\forall Y(Y \cup X \notin Inc \Rightarrow Y \not\models \Box p) \Rightarrow \forall Z(Z \cup X \notin Inc \Rightarrow Z \not\models \Box p))$.

□

This, as it is well known[59], is enough to make modal operators collapse, in the sense that $p \equiv \Box p$ turns out to be provable.

[58] See Proposition 4.4, B&A (2008), p. 161.

[59] See for instance Hughes and Cresswell [72], p. 131.

Theorem 27. $p \equiv \Box p$

Proof. ($\Leftarrow$). It's just T.

($\Rightarrow$). A couple of Lemmas first.

Lemma 28. $\Box\Diamond p \Leftrightarrow \Diamond p$

Proof. ($\Rightarrow$) ($\Diamond p/p$+T). ($\Leftarrow$) (S5). □

Lemma 29. $\Diamond\Box p \Leftrightarrow \Box p$

Proof. ($\Rightarrow$) $\Diamond\neg p \vDash_I \Box\Diamond\neg p$ ($\neg p/p$+S5)

$\neg\Box p \vDash_I \neg\Diamond\Box p$ (Def.)

$\Diamond\Box p \vDash_I \Box p$ (Contraposition)

($\Leftarrow$) ($\Box p/p$+T1). □

Now the desired result easily follows:

$\Diamond p \vDash_I \Box p$ (M+Lemma 28+Lemma 29)

$p \vDash_I \Box p$ (T1+Cut) □

4.4.2 The modal stability of *EIS*

So, a modal system collapses if it validates both B-axiom and its converse together with *S4*-axiom: any S5 system + $p \vDash \Diamond\Box p$ collapse into *Propositional Calculus*. Now, since the system of Göcke et al. [59] fails *S4*-axiom, it doesn't really run the risk. Anyway, in order to answer directly to Brandom's concern, it is actually possible to reject $p \vDash \Diamond\Box p$ in the authors' system. Let me put forward some trivial results first.

To begin with, I need to define what it means to be incompatible with *possibly* p. Here I simply mimick the definitional strategy of Göcke et al. [59].

Definition. (PW-$\Diamond$I) $X \in I(\Diamond p)$

$$\begin{aligned} X \in I(\Diamond p) &\Leftrightarrow \neg\exists w(X \text{ true in } w \wedge \Diamond p \text{ true in } w) \\ &\Leftrightarrow \forall w(X \text{ true in } w \Rightarrow \Diamond p \text{ false in } w) \\ &\Leftrightarrow \forall w(X \text{ false in } w) \vee \exists w(X \text{ true in } w \wedge \Diamond p \text{ false in } w)) \\ &\Leftrightarrow \forall w_1(X \text{ false in } w_1) \vee \exists w_1(X \text{ true in } w_1 \wedge \forall w_2(w_2 R w_1 \Rightarrow w_2 \vDash \neg p))) \end{aligned}$$

thus

$$\begin{aligned} X \in I(\Diamond p) \Leftrightarrow \quad & X \in Inc \text{ or} \\ & \exists w_1(w_1 \vDash X \wedge \forall w_2(w_2 R w_1 \Rightarrow w_2 \vDash \neg p))) \end{aligned}$$

Then I have to prove this definition is adequate by showing it is the converse of (PW-$\Box$I).

Theorem 30. *$X \in I(\Diamond p) \Longleftrightarrow X \in I(\neg\Box\neg p)$ and $X \in I(\Box p) \Longleftrightarrow X \in I(\neg\Diamond\neg p)$*

Proof. Respectively,

$$\begin{aligned} X \in I(\Diamond p) &\Longleftrightarrow X \in Inc \vee \exists w_1(w_1 \vDash X \wedge \forall w_2(w_2 R w_1 \Rightarrow w_2 \vDash \neg p))) \\ &\Longleftrightarrow X \in Inc \vee \neg(\forall w_1(w_1 \vDash X \Rightarrow \forall w_2(w_2 R w_1 \Rightarrow w_2 \vDash \neg p))) \\ &\Longleftrightarrow X \in Inc \vee \neg(\forall w_1(w_1 \vDash X \Rightarrow \exists w_2(w_2 R w_1 \wedge w_2 \nvDash \neg p))) \\ &\Longleftrightarrow X \in I(\neg\Box\neg p). \end{aligned}$$

$$\begin{aligned} X \in I(\Box p) &\Longleftrightarrow X \in Inc \vee \forall w_1(w_1 \vDash X \Rightarrow \exists w_2(w_2 R w_1 \wedge w_2 \vDash \neg p)) \\ &\Longleftrightarrow X \in Inc \vee \neg(\exists w_1(w_1 \vDash X \wedge \neg\exists w_2(w_2 R w_1 \wedge w_2 \vDash \neg p))) \\ &\Longleftrightarrow X \in Inc \vee \neg(\exists w_1(w_1 \vDash X \wedge \forall w_2(w_2 R w_1 \Rightarrow w_2 \nvDash \neg p))) \\ &\Longleftrightarrow X \in I(\neg\Diamond\neg p). \end{aligned}$$

□

Let me notice that (PW-$\Diamond$I) can be translated into the Kripke-style framework of Göcke et al. [59].

Theorem 31. (KI$\Diamond$) *$w_1 \vDash \Diamond p \Longleftrightarrow \exists w_2(w_2 R w_1 \wedge w_2 \vDash p)$*

Proof. ($\Rightarrow$)

By definition of maximal coherence: $w_1 \vDash \Diamond p \Rightarrow w_1 \cup \{\Diamond p\} \notin Inc$.

By definition of incompatibility,

$$\begin{aligned} w_1 \cup \{\Diamond p\} \notin Inc &\Longleftrightarrow \neg(w_1 \in Inc \vee \exists w_3(w_3 \vDash w_1 \wedge \forall w_4(w_4 R w_3 \Rightarrow w_4 \vDash \neg p) \\ &\Longleftrightarrow w_1 \notin Inc \wedge \neg(\exists w_3(w_3 \vDash w_1 \wedge \forall w_4(w_4 R w_3 \Rightarrow w_4 \vDash \neg p)) \\ &\Longleftrightarrow w_1 \notin Inc \wedge \forall w_3(w_3 \vDash w_1 \Rightarrow \exists w_4(w_4 R w_3 \wedge w_4 \vDash p)) \end{aligned}$$

But $w_1$is maximally coherent, thus if it is implied by w_3 which is maximally coherent it is equal to it

> ("since w is itself maximally coherent then the only maximally coherent set it is entailed by is itself.")

thus,

$$w_1 \notin Inc \wedge \exists w_4(w_4 R w_1 \wedge w_4 \vDash p)).$$

But w_1 is maximally coherent, thus it reduces to $w_1 \vDash \Diamond p \Rightarrow \exists w_4(w_4 R w_1 \wedge w_4 \vDash p))$.

($\Leftarrow$)

Let $w_1$maximally coherent and (i) $\exists w_2(w_2 R w_1 \wedge w_2 \vDash p)$. To show $w_1 \vDash \Diamond p$, i.e. $\forall X(X \in I(\Diamond p) \Rightarrow X \in I(w_1))$. Suppose not.

Let also (ii) $X \in I(\Diamond p) \Rightarrow X \in Inc$ or $\exists w_3(w_3 \vDash X \wedge \forall w_4(w_4 R w_3 \Rightarrow w_4 \vDash \neg p))$.

If $X \in Inc$ then $X \in I(w_1)$. $\bot$.

If $X \notin Inc$, suppose $X \notin I(w_1)$, i.e. $X \cup w_1 \notin Inc$.

Then by maximal coherence of w_1, $w_1 \vDash X$. Together with (ii) this implies $\forall w_4(w_4 R w_1 \Rightarrow w_4 \vDash \neg p)$.

This contradicts (i). $\bot$. $\square$

Eventually it can be proved that $p \vDash \Diamond\Box p$ can be rejected.

Theorem 32. $p \nvDash \Diamond\Box p$

Proof. Suppose $\forall q(q \in I(\Diamond\Box p) \Rightarrow q \in I(p))$

$$
\begin{aligned}
X \in I(\Diamond\Box p) \Longleftrightarrow \quad & X \in Inc \text{ or} \\
& \exists w_1(w_1 \vDash X \wedge \forall w_2(w_2 R w_1 \Rightarrow w_2 \vDash \neg\Box p) \Longleftrightarrow \\
& \exists w_1(w_1 \vDash X \wedge \forall w_2(w_2 R w_1 \Rightarrow w_2 \vDash \Diamond\neg p) \Longleftrightarrow \\
& \exists w_1(w_1 \vDash X \wedge \forall w_2(w_2 R w_1 \Rightarrow \exists w_3(w_3 R w_2 \wedge w_3 \vDash \neg p)))
\end{aligned}
$$

A counterexample would be any $X \notin Inc$ and s.t. $\exists w_1(w_1 \vDash X \wedge \forall w_2(w_2 R w_1 \Rightarrow \exists w_3(w_3 R w_2 \wedge w_3 \vDash \neg p)))$ but $X \cup p \notin Inc$.

Thus consider:

	x	y	z	p	x,y	x,z	x,p	y,z	y,p	z,p	x,y,z	x,y,p	x,z,p	y,z,p	x,y,z,p
x	+	+	-	+	+	-	+	-	+	-	-	+	-	-	-
y	+	+	+	+	+	-	+	+	+	-	-	+	-	-	-
z	-	+	+	-	-	-	-	+	-	-	-	-	-	-	-
p	+	+	-	+	+	-	+	-	+	-	-	+	-	-	-
x,y	+	+	-	+	+	-	+	-	+	-	-	+	-	-	-
x,z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
x,p	+	+	-	+	+	-	+	-	+	-	-	+	-	-	-
y,z	-	+	+	-	-	-	-	+	-	-	-	-	-	-	-
y,p	+	+	-	+	+	-	+	-	+	-	-	+	-	-	-
z,p	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
x,y,z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
x,y,p	+	+	-	+	+	-	+	-	+	-	-	+	-	-	-
x,z,p	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
y,z,p	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
x,y,z,p	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Figure 4.4.1: Counterexample to $p \vDash \Diamond\Box p$

This model satisfies *Persistence*, *Reflexivity* and *Symmetry*. Here we have only two possible worlds: $w_i = \{x, y, p\}$ and $w_j = \{y, z\}$. Both are accessible the one from the other.

Now let $X = \{x\}$ and $p = \{p\}$. Then there is just one possible world, i.e. w_i, s.t. $w_i \vDash X$. There are two worlds accessible from w_i: w_j and w_i itself. And there is a world accessible from both of them, i.e. w_j s.t. $w_j \nvDash p$. And yet $w_i \vDash p$. □

4.4.3 Modal properties without *Persistence*

In this Section of this Appendix I simply collect some results about the modal behaviour of Brandom's modal operator in an hypothetical system of *Incompatibility Semantics* without *Persistence* of incompatibility. Although no such a system has been defined yet and these theorems just float in a logically unacceptable unconstrained space, nonetheless I think it is important to lay them down for two reasons: (i) to illustrate Brandom's idea of a direct model theoretic transposition of the language of norms and (ii) to entice some good logician to endorse the task to develop these suggestion into a well defined system. In this sense the most important thing to notice is that incompatibility relations (and thus compatibility relations) here are still treated in extensional set theoretic terms.

Lemma 33. (Kripkean equivalency)

$X \vDash \Box p \Leftrightarrow \forall Y(Y \cup X \notin Inc \Rightarrow \forall Z(Z \cup Y \notin Inc \Rightarrow Z \vDash p))$

Proof. Expand $X \models \Box p$:

$$\begin{aligned} X \models \Box p &\Leftrightarrow \forall Y(Y \in I(\Box p) \Rightarrow Y \in I(X)) \\ &\Leftrightarrow \forall Y(Y \models \varnothing \text{ or } \exists Z(Z \cup Y \notin Inc \wedge Z \not\models p) \Rightarrow Y \in I(X)). \end{aligned}$$

If $Y \models \varnothing$ then $Y \in I(p)$. Otherwise, $\forall Y(\exists Z(Z \cup Y \notin Inc \wedge Z \not\models p) \Rightarrow Y \in I(X))$. Which amounts to $\forall Y(Y \cup X \notin Inc \Rightarrow \forall Z(Z \cup Y \notin Inc \Rightarrow Z \models p))$. □

Theorem 34. (Necessitation)

$\models p \Rightarrow \models \Box p$

Proof. Suppose $\models p$ and $\not\models \Box p$

1. $\models p$ and $\exists X(X \not\models \varnothing \wedge X \not\models \Box p)$.
2. $\forall Z(Z \not\models \varnothing \Rightarrow Z \models p)$ and $\exists X(X \not\models \varnothing \wedge \exists Y(Y \cup X \notin Inc \wedge Y \in I(\Box p))$.
3. $\forall Z(Z \not\models \varnothing \Rightarrow Z \models p)$ and $\exists X(X \not\models \varnothing \wedge \exists Y(Y \cup X \notin Inc \wedge \exists Z(Z \cup Y \notin Inc \wedge Z \not\models p))$.

Since compatibility is *reflexive*, we $\forall Z(Z \not\models \varnothing \Rightarrow Z \models p)$ and $\exists X(X \not\models \varnothing \wedge X \not\models p))$. $\perp$. □

Theorem 35. (K)

$\Box(p \to q) \models \Box p \to \Box q$

Proof. Define $p \to q \equiv \neg(\neg p \wedge \neg q)$.

Thus what is to prove is $\Box\neg(\neg p \wedge \neg q) \models \neg(\neg\Box p \wedge \neg\Box q)$.

1. $\Box\neg(\neg p \wedge \neg q), \neg\Box p, \neg\Box q \models \varnothing$.
2. $\Box\neg(\neg p \wedge \neg q) \models \Box p, \neg\Box q$.
3. Suppose $\exists X(X \in I(\Box p, \Box q) \wedge X \notin I(\Box\neg(\neg p \wedge \neg q)))$, i.e.
$$\exists X(\begin{cases} \exists Y(Y \cup X \notin Inc \wedge Y \not\models p) \\ \qquad\qquad \wedge \\ \exists Z(Z \cup X \notin Inc \wedge Z \not\models q) \end{cases} \wedge X \not\models \varnothing \Rightarrow \forall K(X \cup K \notin Inc \Rightarrow K \models \neg(\neg p \wedge \neg q))).$$
4. $\exists X(\begin{cases} \exists Y(Y \cup X \notin Inc \wedge Y \not\models p) \\ \qquad\qquad \wedge \\ \exists Z(Z \cup X \notin Inc \wedge Z \not\models q) \end{cases} \wedge X \not\models \varnothing \Rightarrow \forall K(X \cup K \notin Inc \Rightarrow K, \neg p, \neg q \models \varnothing)).$

5. $\exists X(\begin{cases}\exists Y(Y \cup X \notin Inc \wedge Y \not\models p) \\ \wedge \\ \exists Z(Z \cup X \notin Inc \wedge Z \not\models q)\end{cases} \wedge X \not\models \varnothing \Rightarrow \forall K(X \cup K \notin Inc \Rightarrow \begin{cases}K \models p \\ \wedge \\ K \models q\end{cases})).$

If $\varnothing \models \varnothing$ the result follows. Otherwise instantiate appropriately to obtain a contradiction. □

Theorem 36. (T)

$\Box p \models p$

Proof. We have to show that $\forall X(X \in I(p) \Rightarrow X \in I(\Box p))$.

If $X \models \varnothing$ then $X \in I(\Box p)$.

If $X \not\models \varnothing$, suppose $X \in I(p) \wedge X \notin I(\Box p)$. Then $X \in I(p) \wedge \forall Y(Y \cup X \notin Inc \Rightarrow Y \models p)$.

But then, assuming that compatibility is *reflexive*, this implies $X \models p$, i.e. $X \notin I(p)$. $\perp$. □

Theorem 37. (B)

$p \models \Box\Diamond p$

Proof. Since $\Diamond p \equiv \neg\Box\neg p$, what is to prove is $p \models \Box\neg\Box\neg p$.

1. $\forall X(X \in I(\Box\neg\Box\neg p) \Rightarrow X \in I(p))$;
2. $\forall X(X \models \varnothing$ or $\exists Y(Y \cup X \notin Inc \wedge Y \not\models \neg\Box\neg p) \Rightarrow X \in I(p))$.

If $X \models \varnothing$ then $X \in I(p)$. Otherwise,

3. $\forall X(\exists Y(Y \cup X \notin Inc \wedge Y \models \Box\neg p) \Rightarrow X \in I(p))$;
4. $\forall X(\exists Y(Y \cup X \notin Inc \wedge \forall Z(Z \cup Y \notin Inc \Rightarrow \forall K(K \cup Z \notin Inc \Rightarrow K \models \neg p))) \Rightarrow X \in I(p))$.

But then, assuming that compatibility is *reflexive* and *symmetric*, this implies $X \in I(p) \Rightarrow X \in I(p)$. □

Theorem 38. (S4)

$\Box p \models \Box\Box p$

Proof. We have to prove $\forall X(X \in I(\Box\Box p) \Rightarrow X \in I(\Box p))$.

1. $\forall X(X \models \varnothing$ or $\exists Y(Y \cup X \notin Inc \wedge Y \not\models \Box p) \Rightarrow X \models \varnothing$ or $\exists Z(Z \cup X \notin Inc \wedge Z \not\models p))$.

If $X \models \varnothing$ then $X \in I(\Box p)$. Otherwise,

2. $\forall X(\exists Y(Y \cup X \notin Inc \wedge \exists K(K \cup Y \notin Inc \wedge K \in I(\Box p)) \Rightarrow \exists Z(Z \cup X \notin Inc \wedge Z \not\models p))$;

3. $\forall X(\exists Y(Y \cup X \notin Inc \wedge \exists K(K \cup Y \notin Inc \wedge K \models \varnothing \text{ or } \exists S(S \cup K \notin Inc \wedge S \not\models p))) \Rightarrow \exists Z(Z \cup X \notin Inc \wedge Z \not\models p))$.

If $K \models \varnothing$ then $X \models \varnothing$, and then $X \in I(\Box p)$. Otherwise, assuming that compatibility is transitive,

$$\forall X(\exists Y(Y \cup X \notin Inc \wedge Y \not\models p))) \Rightarrow \exists Z(Z \cup X \notin Inc \wedge Z \not\models p)).$$

Instantiate Y and Z at the same way to obtain a tautology. □

4.4.4 Semantic recursiveness for *EIS*

4.4.4.1 The problem

As Brandom conceived it, *Incompatibility Semantics* should be a modal semantic metavocabulary for the discursive practice of giving and asking for reasons. In which sense Kripke's *relational semantics* is not an equivalent modal semantic metavocabulary? Kripke obtains his success in the extensional approach at the price of the introduction of possible worlds as primitives. So the score is: one primitive for *IS*, incompatibility, two primitives for Kripke's semantics, truth and possible worlds. But *EIS* has possible worlds too: does that even the score? Now, one could argue that *EIS* actually doesn't take possible worlds as primitives, but to show that is really to explain why the question can't be settled just by counting the number of primitives. The key element to consider are the relations with the normative vocabulary of incompatibility, which is a pragmatic metavocabulary for the very same practices of giving and asking for reasons: normative metavocabulary of incompatibility and semantic metavocabulary of incompatibility are two *corresponding* ways to make *directly* explicit the meaning of these practices. In Brandom [23] these relations are made explicit by the so called meaning-use diagrams which here I allow myself to drastically simplify in order to show just what it takes to make my point.

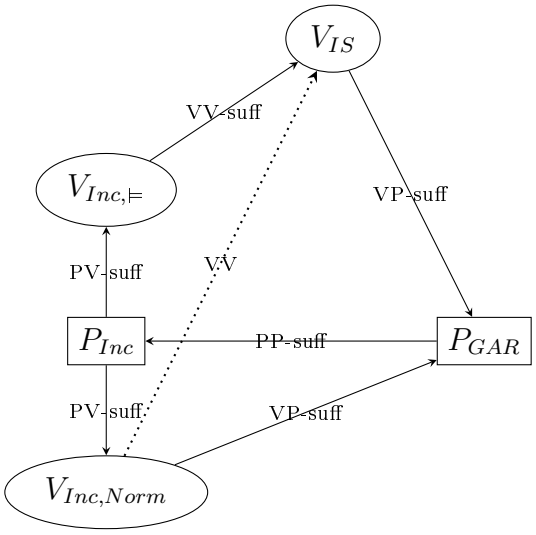


Figure 4.4.2

A few comments may help reading Figure 4.4.2. In engaging in the *Practices of giving and asking for reasons* one knows all that is needed to engage in the *Practices of treating contents as incompatible*s. These in turn are sufficent to support the use of a semantic vocabulary and a pragmatic vocabulary of incompatibility, in the sense that one can be construed as deploying those vocabularies because of the very fact that he engages in the practices of treating contents as incompatibles. Eventually the vocabulary of incompatibility entailment is sufficient to characterize the modal vocabulary of *Incompatibility Semantics*, that is to say the latter is recursively *reducible* to the former[60]. Brandom is particularly interested in the resultant VV relation between normative vocabulary and modal semantic vocabulary which techinically represents Sellars's claim that "the language of modalities is a 'transposed' language of norms"[61].

Notice that the abilities to engage in the practices which are sufficient to deploy the modal vocabulary of *Incompatibility Semantics* can be *directly*, i.e. algorithmically, elaborated from those of incompatibility entailment. This is what Kripke's relational semantics can't do: there, the external step through possible worlds is required.

I claim that *EIS*'s vocabulary can take the place of *IS*'s vocabulary in Figure 4.4.2, so to say that the abilities to engage in the practices which are sufficient to deploy *EIS* can be algorithmically elaborated at the same way. The only proviso – which *EIS* however satisfies – is that *compossibility* is *reflexive*. And such a claim can be proved.

[60]See Definition 40 below and compare with Section 5, Appendix I, B&A (2008).
[61]Sellars [146], p. 280.

4.4.4.2 The strategy

The strategy is inherited from the conservativity results for relevance logic in Meyer [104]. It's applied by Alp Aker in a series of Lemmas[62], whose crucial role is to establish that the meanings of complex sentences in *IS* are *holistically* but *recursively* computable on those of syntactically simpler ones. These Lemmas provide an algorithm to extend *incoherence frames* for languages of progressively more complex syntax while keeping *meanings* stable, in the following sense.

Definition 39. Let $L \subseteq L'$ and let Inc be an *incoherence frame* for L. Then a frame Inc' for L' is *inferentially conservative (IC) with respect to Inc* if, for $X, Y \subseteq L$, $X \models_{Inc} Y \Leftrightarrow X \models_{Inc'} Y$.

We require L' to be a *proper extension* of L in the sense that for any formula $p \in L'$, if q is a subformula of p then $q \in L'$ and all atomic formulae in L' are contained in L. The overall result of these Lemmas is to prove that given any language L and its frame Inc, for any (possibly infinitary) extension L' of L there exists a frame Inc' for L' that contains only subsets of L whose incoherence can be shown to follow from Inc by applications of reduction rules for propositional operators. These rules establish that entailments containing syntactically complex formulae are equivalent to (a boolean combination of) entailments mentioning fewer connnectives. Thus Inc' will be the smallest frame for L' which is *inferentially conservative* with respect to Inc.

This result can be adapted to *EIS*'s framework. In what follows I just give the sketch of the proof integrating Aker's original one and the strategy is to require that

> (i) the semantic value of syntactically complex sentences is *recursively* computable on those of syntactically simpler ones;
>
> (ii) the introduction of operators produces an inferentially conservative extension.

Requirement (i) would ensure the technical advantages of compositionality in the holistic setting of *Incompatibility Semantics*. Requirement (ii) would ensure that the introduction of these complex formulae do not change other meanings in any way, thus defusing the threats of holisms.

Notice in (ii) we need the inferentially conservative extension

[62] See Lemmas 5.2.1-5.2.8, Brandom [23] pp. 147-155.

(ii.i) to exist;

(ii.ii) to be unique.

Let me put down a quick remark on uniqueness: for infinite languages it can't be taken for granted that there will be just only one conservative extension. In fact it might well be the case that many different incoherence frames for an infinite language might match all the incompatibility sets of the frame they extend and yet diverge the one from the other by containing other incompatibilities. But *IS*'s framework provides a trick to cope with this. Consider, among these frames, the minimal one, the one that contains the minimal amount of new incompatibilities. Now *Persistence* establishes that such minimal frame is contained in all incoherence frames for inferentailly conservative extensions. In this sense notice that if such a minimal incoherence frame exists then by definition it is unique: in other words it is enough to show (ii.i) in order to have (ii.ii) as well. Since we are dealing just with finite languagues we can just smooth over this, even if the same results could be obtained for infinite languages as well.

It is useful to sum all this up with a definition.

Definition 40. An incoherence frame $< L', Inc' >$ is *recursively reducible* to another incoherence frame $< L, Inc >$ if

($\varnothing$) for any $X \in Inc'$ there is some $X' \subseteq X$ s.t. $X' - L$ is finite and $X' \in Inc'$.

(1) there is a computable (boolean) function F s.t. for any $X', Y' \in L'$ and $X_i, Y_i \in L$,

$$X' \models_{Inc} Y' \text{ iff } F(X_1 \models_{Inc} Y_1; ...; X_n \models_{Inc} Y_n);$$

(2) Inc' is *inferentially conservative* (IC) w.r.t. Inc.

4.4.4.3 The proof

I take for granted Aker's proof for non modal operators, so I'll just deal with the extensions realized by the introduction rules of modal operators in *EIS*. To begin with reduction rules are needed for the modal operator in *EIS* (they correspond to those of B&A (2008), namely Prop. 4.2.1 and Prop. 4.2.2). This is the crucial step that **T** modality allows.

Lemma 41. *If compossibility is reflexive,*

$$X, \Box p \models Y \textit{ iff } X \models Y \textit{ or } X \not\models p$$

Proof. ($\Rightarrow$)

$$X, \Box p \vDash Y \quad \Leftrightarrow \forall k(k \in I(Y) \Rightarrow k \in I(\Box p, X))$$
$$\Leftrightarrow \forall k(k \in I(Y) \Rightarrow \forall w_1(w_1 \vDash k \wedge X \Rightarrow \exists w_2(w_2 R w_1 \wedge w_2 \nvDash p)))$$

Thus either (i) $\neg\exists w_i(w_i \vDash k \wedge X)$, i.e. $k \in I(X)$ and so $X \vDash Y$,

or (ii) w_2 testifies for $X \nvDash p$.

($\Leftarrow$)

Suppose $X \vDash Y$, then by $X, \Box p \vDash Y$ by *Persistence*.

Suppose $X \nvDash p$, then, if *compossibility* is reflexive, $X, \Box p \vDash \varnothing$. Thus $X, \Box p \vDash Y$. □

Lemma 42. *If compossibility is reflexive,*

$$X \vDash Y, \Box p \text{ iff } X \vDash Y \text{ or } X \vDash p$$

Proof. ($\Rightarrow$)

$$X \vDash Y, \Box p \quad \Leftrightarrow \forall k(k \in I(Y, \Box p) \Rightarrow k \in I(X))$$
$$\Leftrightarrow \forall k(\forall w_1(w_1 \vDash k \wedge Y \Rightarrow \exists w_2(w_2 R w_1 \wedge w_2 \nvDash p)) \Rightarrow k \in I(X))$$

Thus either (i) $\neg\exists w_i(w_i \vDash k \wedge Y)$, i.e. $k \in I(Y)$ and so $X \vDash Y$,

or (ii), if *compossibility* is reflexive, $k \in I(p)$, thus $X \vDash p$.

($\Leftarrow$)

Suppose $X \vDash Y$. Then $X \vDash Y, \Box p$ by *Weakening* (recall from Section 4.2.2 that sentences are disjuncts on the right).

Suppose $X \vDash p$. If *compossibility* is reflexive, $X \vDash \Box p$. Thus $X \vDash Y, \Box p$ by *Weakening*. □

These reduction rules allow to describe a recursive procedure to evaluate the interpretation of formulae containing the box operator. This is enough to meet requirement (1) in Definition 40. Once this is acknowledged it must be shown that this procedure produces inferentially conservative extensions of the incoherence frame for a non-modal language.

Lemma 43. *Given a language L and a frame Inc, let $L_\varnothing$ be a fragment of L s.t. that for any $X, Y \subseteq L$, $X - L_\varnothing$ and $Y - L_\varnothing$ are finite. Then there is a boolean function F s.t. for $X_i, Y_i \subseteq L_\varnothing$*

$$X \vDash_{Inc} Y \text{ iff } F(X_1 \vDash_{Inc} Y_1; ...; X_n \vDash_{Inc} Y_n)$$

Proof. (Sketch)

By induction on the number of connectives in X_i, Y_i. For non modal vocabulary see the proof of Lemma 5.2.1 in B&A (2008).

For modal vocabulary, suppose

$$F(X_1 \vDash_{Inc} Y_1; ...; X_n \vDash_{Inc} Y_n) \text{ iff } F(X_1 \vDash_{Inc} Y_1; ...; \Box p, Z_i \vDash_{Inc} Y_i; ...; X_n \vDash_{Inc} Y_n).$$

Apply Lemma 41 to obtain

$$F(X_1 \vDash_{Inc} Y_1; ...; X_n \vDash_{Inc} Y_n) \text{ iff } F(X_1 \vDash_{Inc} Y_1; ...; Z_i \vDash_{Inc} Y_i \text{ or } Z_i \nvDash_{Inc} p; ...; X_n \vDash_{Inc} Y_n).$$

Do the same for modal operator on the right by applying Lemma 42. □

Now it must be shown that the *inferentially conservative* frame Inc' for any proper extension L' of L always exists[63]. In order to do that, an algorithm must be provided to actually build such a frame: this is what Lemmas 5.2.3-5.2.7 in B&A (2008) do. To do the same here we have to show that the frame Inc for L, obtained as the *IC* extension of the frame $Inc_{\varnothing}$ for the non modal fragment $L_{\varnothing}$, verifies all the axioms of *EIS*. This may look tricky since possible worlds are involved: these, in fact, are defined as maximally compatible sets of sentences, so they might be affected by modifications of the frame when new operators are added to the language. But if we take a look at these modifications from our point of view, i.e. with relation to the problem of recursively reducing L to $L_{\varnothing}$, we see that they amount to moves from possible worlds in L

$$PW_{Inc'} = \{S \mid S \notin Inc' \text{ and } \forall X(X \cup S \notin Inc' \Rightarrow X \subseteq S)\}$$

to possible worlds in $L_{\varnothing}$

$$PW_{Inc} = \{S \mid S \notin Inc \text{ and } \forall X(F(X) \cup S \notin Inc \Rightarrow F(X) \subseteq S)\}.$$

What is crucial to realize is that, even if these worlds may be composed of different sets of sentences, they relate equivalently with F-correspondent sets of sentences in L and L': when you move from $PW_{Inc'}$ to PW_{Inc} (and *viceversa*) by applying F you still obtain maximally compatible sets of sentences.

This also means that *compossibility* w.r.t. to Inc' corresponds, through F, to *compossibility* w.r.t. to Inc:

Proposition. $w_j R w_i \Leftrightarrow F(w_j) R F(w_i)$.

[63]Notice that for infinite languages a preliminary step is required: it must be shown how to identify *the* inferentially conservative frame Inc' for any proper extension L' of L. That is the purpose of Aker's Lemma 5.2.2 in B&A (2008).

Proof. Consider $w_j, w_i \in PW_{Inc'}$ s.t. $w_j R w_i$. Thus

$$\forall p(p \cup w_j \notin Inc' \Rightarrow \exists X(X \cup w_i \notin Inc' \wedge X \cup p \notin Inc')).$$

If we instantiate X with some set W and apply our equivalence between incoherence frames Inc' and Inc under function F we obtain

$$\forall p(p \cup F(w_j) \notin Inc \Rightarrow F(W) \cup F(w_i) \notin Inc \wedge F(W) \cup p \notin Inc).$$

Then

$$\forall p(p \cup F(w_j) \notin Inc \Rightarrow \exists X(X \cup F(w_i) \notin Inc \wedge X \cup p \notin Inc).$$

But since $F(w_j)$ and $F(w_i)$ belong to PW_{Inc} then we have what we need: $F(w_j)RF(w_i)$. □

That is enough to establish the required results. In fact, once L has been extended to L' through (an) F above, in order to verify the axiom (PW-□I) for *EIS* in L' one can simply argue as follows.

Lemma 44. *Let Inc be a frame for L. Let $L' = L \cup \{\Box p\}$ for some $p \in L$. We want $F(X) \in Inc$ iff $X \in Inc'$, thus let*

$$X \in Inc' \text{ iff } \begin{cases} X \nvDash_{Inc} p & \text{if } X' = X \cup \{\Box p\} \\ X \in Inc & \text{otherwise} \end{cases}$$

Then (i) Inc' is a frame for L' and (ii) Inc' is IC w.r.t. Inc.

Proof. In order to prove (i) we verify the axioms of *EIS*.

For *Persistence*, (¬I) and (∧I) as in the proof of Lemma 5.2.3 in B&A (2008).

For (PW-□I), we want to show

$$X \cup \{\Box r\} \in Inc' \Leftrightarrow \forall w_1(w_1 \vDash_{Inc'} X \Rightarrow \exists w_2(w_2 R w_1 \wedge w_2 \nvDash_{Inc'} r))$$

(⇒)

Suppose $X \notin Inc'$, then $F(X) \notin Inc$.

Also, $X \cup \{\Box r\} \in Inc' \Leftrightarrow F(X) \cup \{\Box r\} \in Inc$. So,

$$\begin{aligned} F(X) \cup \{\Box r\} \in Inc &\Leftrightarrow \forall w_1(w_1 \vDash_{Inc} F(X) \Rightarrow \exists w_2(w_2 R w_1 \wedge w_2 \nvDash_{Inc} r)) \\ &\Leftrightarrow \forall w_1(F(X) \cup w_1 \notin Inc \Rightarrow \exists w_2(w_2 R w_1 \wedge r \cup w_2 \in Inc)) \end{aligned}$$

This last step follows by the *Either-Or Lemma*

Lemma. *(Either-or Lemma) [Göcke et al. [59]]*

$$\forall w_i \in PW_{Inc}, X \subseteq L(W \models_{Inc} X \Leftrightarrow X \cup w_i \notin Inc).$$

Now, since L doesn't contain the new operator we have $w_i = F(w_i)$.
But then $\forall w_1(w_1 \models_{Inc'} X \Rightarrow \exists w_2(w_2 R w_1 \wedge w_2 \not\models_{Inc'} r)$.
($\Leftarrow$)
By applying the *Either-Or Lemma* to

$$\forall w_1 \in PW_{Inc'}(w_1 \models_{Inc'} X \Rightarrow \exists w_2 \in PW_{Inc'}(w_2 R w_1 \wedge w_2 \not\models_{Inc'} r))$$

we obtain

$$\forall w_1 \in PW_{Inc'}(X \cup w_1 \notin Inc' \Rightarrow \exists w_2 \in PW_{Inc'}(w_2 R w_1 \wedge r \cup w_2 \in Inc')).$$

Instantiate to obtain

$$X \cup W_1 \notin Inc' \Rightarrow W_2 R W_1 \wedge r \cup W_2 \in Inc'.$$

Now apply function F

$$F(X) \cup F(W_1) \notin Inc \Rightarrow F(W_2) R F(W_1) \wedge r \cup F(W_2) \in Inc.$$

But then, since $F(W_1)$ and $F(W_2)$ are maximally compatible sets of sentences w.r.t. Inc we can generalize as

$$\forall w_1 \in PW_{Inc}(w_1 \models_{Inc} X \Rightarrow \exists w_2 \in PW_{Inc}(w_2 R w_1 \wedge w_2 \not\models_{Inc} r).$$

Thus $F(X) \cup \{\Box r\} \in Inc \Leftrightarrow X \cup \{\Box r\} \in Inc'$.
For (ii) the proof again as in Lemma 5.2.3 in B&A (2008). □

Eventually we can prove the following [64].

Theorem 45. *The language of EIS is recursively reducible to the language of incompatibility entailment.*

[64] Notice with respect to Aker's original proof, requirement (∅) of Definition 40 comes for free since we do not have to consider extensions by several connectives.

Proof. Apply Lemma 43 and Lemma 44. □

This concludes my sketchy adaptation of Aker's proofs for *EIS*. We can now say that the vocabulary of incompatibility entailment is *VV-sufficient* for the vocabulary of *EIS* and begin to deploy such an modal semantics as an extensional representation of normative meanings.

Chapter 5

Towards Causal Modalities, a Nonmonotonic Path

5.1 Introduction

The last two Chapters were dedicated to the presentation and the deployment of Brandom's analysis of the normativity of meaning. That let the seeds we planted in Chapter 1 grow into a full blown modal representation of semantic contents. The precious fruit we could pluck is *EIS*, a system of formal semantics directly set up on the pragmatic notion of *incompatibility*, which is the drive shaft in Brandom's account of the normative structure of *sapient* beings' linguistic practices. This approach has already been compared with the standard extensional account of semantics in several respects, so it should be clear enough, by now, the sense in which Brandom's inferential and normative model let us cope with the first of the blindspots diagnosed in Section 1.4.1. Let me however briefly recap. According to a certain intuitive but still deeply rooted augustinian view, the basic move in the semantic game is sticking labels on entities. This idea crucially presupposes that the overall structure of the world's furniture is already given before the game can kick off. That is to say that the structure of repeatables is treated as given independently from the structure of their determinations. This idea has been put under philosopical attack at least since Kant's idealism, but lately, through *semantic atomism*, it found a safe shelter into the kernel of Russell's extensional semantics. From that fortress it still resists these days and inevitably spreads its influence on the analyses of conceptual content. As we could see, Brandom counter this view with an inferentialist approach, which goes back to Frege's pre-extensional intuitions in the *Begriffschrift* and which aims to represent meanings as rules

for reasoning. This is not a brand new enterprise in logic, but, historically, the important results of Gödel and Tarski drove it to a separated track, where proof theory flourished as an analysis of syntax. Now, Brandom's *normative* inferentialism shows how this syntactic flavour can be dispelled as soon as a suitable pragmatic analysis of linguistic practices is provided as a normative metavocabulary for the logical vocabulary that makes explict conceptual rules.

And yet, this promising path we've taken still presents an apparently overwhelming obstacle: all that we've introduced and discussed till now still provides no answer to the second of the blindspots of standard semantics diagnosed in Section 1.4.2. Let me recall what the problem is. *EIS*, as a possible worlds semantics, inherits from Kripke's approach the so called *metaphysical* interpretation of modality, and consequently it forces upon us a metaphysical representation of *normativity*. Norms have to be out there applying to every possible entity, and when entities do not conform to norms it's because we are just wrong in representing norms from our epistemically limited perspective. This is a problem for normative inferentialism because, if we don't know whether the norms we take as expressing our conceptual contents are the right ones, we can't simply know if we are applying our concepts *right*. That would make the whole process of making contents explicit quite iffy.

In this framework, Brandom's pragmatic metavocabulary would eventually fare no better than the standard extensional one.

One could suggest, with Lance and Hawthorne, to re-establish our grasp on conceptual contents by providing a pragmatic analysis of what it is for us, *sapient* beings, to *treat* contents *as analytic*: their proposal is that propositional contents are analytic if they are *de jure* unchallengeable. Unfortunately, if our representation of the norms that define what is de jure and what is not is metaphysical, the the only dimension in which de jure unchallengeability can be distinguished from de facto unchallengeability is the epistemic one.

In other words, we still face a semantic question: how to distinguish proper norms from accidental but universal constant correlations?

And that drags us back to where we started in Chapter 1, that is to the analysis of counterfactuals. Were we too hasty in discarding anti-normativist complaints? It may seem so unless we explain how to avoid the metaphysical perspective. This is what I indend to do in this Chapter. While the solution of the first blindspot was basically rooted in Wilfrid Sellars's famous thoughs in *Empiricism and the Philosophy of Mind*, here I move from another paper of Sellars's which is temporally and theorically very close to Sellars

[147] but, unfortunately, credited with a far less widespread influence: *Counterfactuals, Dispositions and the Causal Modalities*. In Section 5.2 I rehearse some main parts of Sellars's arguments in Sellars [148] for the criticism of Goodman's analysis of counterfactual conditionals. That will provide us with suitable theoretical tools to eventually approach a resolution of the second blindspot diagnosed inSection 1.4.2: how to express the distinction between contingently costant conjunction and necessary but fallible normative laws. That is what I'll try to do in Section 5.3. Then in Section 5.4, I'll try to apply this approach to the theme of "causal modalities" in order to remove some of the obstacles such a notion found on its path and to build a bridge towards Brandom's pragmatist project. At last, in the two concluding sections of this Chapter I'll develop the logical tools required to turn this analysis into a full fledged definition of a normative semantics suitable to represent these causal entailments in terms of a nonmonotonic consequence relation. An Appendix will be dedicated to technical details.

5.2 Sellars on counterfactuals

The analysis of modal vocabulary is one of the main themes of Sellars's early essays, but it has been noticed (not without regret) how this topic seems to fade away from the spotlight of sellarsian thought after his masterpiece *Empiricism and the Philosophy of Mind* of 1956. This is often construed as a progressive drift towards the kantian themes of *Science and Metaphysics* of 1968. Despite this, Brandom [27] convincingly argues that the theme of modality is but one of the battlefields over which Sellars conducted his many-sided attack against *traditional* empiricism, where Sellars [147] represents his more glorious and famous results. Brandom organizes the structure of this enterprise into three theses and conjectures that it was just by bracketing the analysis of modality that Sellars could develop the other two sides and eventually end up writing *EPM*. Sellars would attempt to deal with the third side, modality, in the other big essay of those years, i.e. *Counterfactual, Dispositions and the Causal Modalities*:

> «I conjecture that what made it possible for Sellars finally to write 'Empiricism and the Philosophy of Mind' was figuring out a way to articulate the considerations he advances there without having also at the same time to explore the issues raised by empiricism's difficulties with modal concepts.»[1]

[1] Brandom [27], p. 6.

5.2.1 A three-fold attack to empiricism.

Let me begin by rehearsing Brandom's argument in order to show how relevant Sellars [148] is for the step we're trying to take. So, Brandom recognizes three main theses in the criticism Sellars developed against *traditional* Empiricism.

The first one is a form of *inferentialism*, according to which semantics has to be primarily accounted for not in terms of word-world correspondence relations, but in terms of the functional role expressions have in language-entry, intra-language and language-exit transitions.

The second one is a form of *pragmatism*, according to which these functional roles must be defined in terms of the practices a speaker must engage in and the abilities she must deploy in order for her use of expressions to be treated as meaningful[2].

Notice that these are the well known pillars supporting the building of Brandom [14]. Notice also that the *pragmatist thesis* corresponds to what in Brandom [23] he calls the "pragmatist challenge" to analytical philosophy, i.e. the attempt to explain the meaning of a vocabulary in terms of what one must do in order to use it.

The third one is a *normative* account of semantic vocabulary, according to which the practice of reporting experience is not only the practice of describing the world as accurately as possible, but also, essentially, the practice of deploying counterfactual reasoning, the practice of saying what might be the case if things were such and such. In this sense, it is a direct answer to humean and quinean criticisms against the role of modal vocabulary in an epistemological framework and a refusal to explain modality away in naturalistic terms.

Brandom maintains that while the first and, though only implicitly, the second thesis are the main weapons deployed against Empiricism in Sellars [147], Sellars had to remove

[2]Brandom admits that this second thesis is only implicit in Sellars [147], but he dedicates part of the essay to show that this pragmatist element is required for Sellars's argument to go through. I won't try to evaluate here the details of his argument. It would surely make Sellars's strategy impressively adjoining to the theses of Brandom [14]. Anyway, it's worth noticing that such an argument clearly hinges on Brandom's overall interpretation of the rejection of the *Given*, as he described it in his *Study Guide to Empiricism and the Philosophy of Mind* and particularly in Brandom [19]. His point is that the functional role of concepts is to be vindicated as a full-fledged inferential role, and that this inferential role has to be accounted for as the result of normative practices of giving and asking for reasons. For a critique of this idea of a primacy of inference in Sellars's account of knowledge see for example McDowell [103]. Let me just add that in a somehow more charitable analysis, this very pragmatist thesis is to be construed as Brandom's way to explain the nominalistic account of inner episodes that occupy the last sections of Sellars [147]. In his *Study Guide* Brandom is quite clear about this. Where he and McDowell do disagree is about Brandom's 'two-ply' interpretation of Sellars's account of observational knowledge. So, if one would undermine such an interpretation one should better try to argue straightforwardly that this pragmatist thesis, as Brandom states it, is not a sellarsian one. I don't think this is a safe move, though.

the third thesis from that argument line and dedicated Sellars [148] to it. But is this third thesis really missing in Sellars [147]? As a matter of fact it's quite easy to detect the third thesis in Sellars [147] too.

In introducing the very argument against the independence of observational language Sellars considers the statement

x is red = x would look red to standard observers in standard conditions[3]

and states that while it can't be a definition, it is nonetheless a necessary truth. In the following famous account of 'looks'-talk Sellars explains that *looking-red* can't be considered an epistemic fact logically independent from *being-red*. So the above identity can't be considered a definition, if identity of meaning is to be construed as analytic identity, but it fits Sellars's analysis of meaning-talk, in the sense that "x is red" has the same role of "x would look red to standard observers in standard conditions". He later would equip himself with the technical notation to make this explicit:

•x is red• = •x would look red to standard observers in standard conditions•[4]

Thus he states that the functional roles of the two concepts of *being-red* and *looking-red in standard condition* coincide. Following Brandom's analysis one can paraphrase this by saying that what one must be able to do in order to use the phenomenalist vocabulary of looking-red-in-standard-conditions coincides with what one must be able to do in order to use the objective vocabulary of being-red.

But why do the practices of these vocabularies *necessarily* coincide? Sellars explains that it's necessary because "standard conditions" means "conditions in which things look what they are"[5], and (in a note added in 1963) that the apparent triviality of this definition disappears if one *makes explicit* the conditions, C, "in which color words have their primary perceptual use".

Let's focus on this set of conditions C. Consider the following counterfactual conditional, that makes explicit the dependence of phenomenalist vocabulary from the objective one:

if x look*ed* red to standard observers in standard conditions C, x *would be* red

which obviously equals to

if x look*ed* red and C, x *would be* red.

[3]Sellars [147], §12.
[4]Please recall Sellars's usage of *dot-quotation* as we already described it in Section 3.3.4.
[5]Sellars [147], §18.

What does it mean then to establish that C must be *standard*? Sellars's remark in the note added in 1963 can be interpreted as saying that the definition of those conditions specifies the range of counterfactual robustness of such a material inference. This is why, for example, if we say that the standard condition is daylight we are specifying the use of color concepts and thus their meaning.

But Sellars explicitly addresses this topic just in Sellars [148]. There he argues against a purely descriptive interpretation of those generalizations: his conclusion is that modality – the *causal* modality he is trying to introduce – is necessary for an account of the content of empirical knowledge. Our main concern here will be with the logic of this nowadays unusual sort of modality, that hardly fits in the standard picture of modal logic developed upon the Kripke's metaphysical modality. Even if it would be tempting, we won't try to evaluate the possible consequences of these reflections for an analysis of inductive logic. We'll rather let us be led by Sellars's analysis of Goodman [60]'s account of counterfactuals in Sellars [148].

5.2.2 The criticism against Goodman's analysis of counterfactuals

Let's try however to follow Sellars consciously, and, first, ask: why Goodman? This should sound as a trivial question with relation to our enterprise. As we saw in Section 1.2.1.3, Goodman's "new riddle of induction", the problem of distinguishing projectible from non projectible predicates is nothing but the problem of distinguishing between modally robust and merely accidental regularities. But there's something more to this choice with relation to Sellars specifically. Let me put forward a quotation as a hint. Just after the introduction of his solution to the problem of induction, hinging on the notion of "projection", Goodman explains:

> «the dispositional predicate 'is orange' is a projection of the manifest predicate 'looks orange', even though not everything that looks orange (e.g. under yellow light) is orange.»[6]

Although Sellars never quotes this passage, it strikingly represents at the same time both the similar setting of the two enterprises and the great divergence in the conclusions obtained: they both clearly maintain that, once the status of inductive generalizations will be clarified, relations between 'looks'-talk and 'is'-talk will be clarified too, but they construe these relations in opposite ways. Thus Sellars is not concerned with an evaluation of the

[6] Goodman [60], p. 56n.

notion of "projection" because his disagreement with Goodman is much more basic: even if Goodman recognizes the need to overcome the syntactical approach to the analysis of induction and to take into consideration the content of the hypotheses to be inductively confirmed, he still looks for an extensional description of the differences between these hypotheses[7]. Sellars instead looks for a modal account in order to explain how inductively established generalizations express the norms according to which we apply concepts. One of his main points, again, is that descriptive vocabulary can't stand by itself without a structure of inferences that account for the meaning of the descriptive expressions it contains.[8]

Goodman's analysis of counterfactuals goes like this:

> «[A] counterfactual is true if and only if there is some set S of true sentences such that S is compatible with C and with $-C$, and such that $A{\cdot}S$ is self-compatible and leads by law to C; while there is no set S' compatible with C and with $-C$, and such that $A{\cdot}S'$ is self-compatible and leads by law to $-C$.»[9]

Let me go *all-in* right away. I construe this as a statement of *monotonicit*y: a valid entailment is not to be turned into a non valid one just by adding premises.

Thus if we indicate with "$>$" the counterfactual conditional, it would simply mean

$$A > C \text{ iff } \forall S(A \cdot S \rightarrow C).$$

But monotonicity obviously fails for counterfactuals, so Goodman tries to fix his analysis by introducing the so called requirement of *Cotenability* of A and S:

> «The requirement that $A \cdot S$ be self-compatible is not strong enough; for S might comprise true sentences that although *compatible with* A, were such that they *would not be true if* A *were true.*»[10]

He wants to define the set S s.t. the implication in the above generalization can't fail and his proposal is to exclude those sentences that could be false if A were true (i.e. the sentences non-cotenable with A) from the set of conditions which are relevant for the

[7] And yet, when we asks what is the logic of "projection" Goodman introduces inferential practice as a crucial element for the definition of such description: he asks what is the logic of the process of content determination and his seminal answer is to look at the practices of establishing incompatibilities between counterfactual inferences. I'll come back to this below in Section 6.2.2.

[8] See Sellars [148], §108.

[9] Goodman [60], p. 13.

[10] *Ibid.*, pp. 13-14.

evaluation of the counterfactual. However this condition generates a regress because the only way to check for *Cotenability* is to to consider another counterfactual:

$A > C$ iff $\forall S(A \cdot S \rightarrow C)$

but

$\forall S(A \cdot S \rightarrow C)$ iff $\forall S'(A \cdot S' \rightarrow S)$

and so on.

This is actually an expected feature of nonmonotonic reasoning. You can't re-establish the monotonicity of a conditional simply by injecting premises untill you extensionally exhaust *ceteris paribus* clauses: what you obtain this way is not a monotonic conditional, but a *trivial* conditional[11]. However this is not Goodman's point. As a matter of fact, Goodman's very argument here is not so straight, so let me pause to disentangle it a bit. He says that if we don't require *Cotenability* we are forced to consider false counterfactuals as true. So he is trying a *reductio ad absurdum*, but he blurs the line of the argument a bit. Thus, *given*

i) If match M had been scratched (A), it would have lighted (C)

his example of a false counterfactual which ends up to be true is:

ii) If match M had been scratched (A), it would not have been dry ($\neg S$).

In fact, if we consider among relevant conditions also "the match doesn't light", we obtain

(iii) If match M had been scratched (A) and it doesn't light ($\neg C$), it would not have been dry ($\neg S$).

Goodman says (iii) is true given (i). The reason is obviously that, once we add ceteris *paribus clauses* S to the counterfactual conditional in (i), (iii) turns out to be simply the *contrapositive* of (i). The strategy of Goodman's argument could be presented as follows:

Suppose Cotenability is not required.

Assume (i) $A \cdot S > C$, then (iii) $A \cdot \neg C > \neg S$, by *Contraposition.*

But (iii) is false, then (i) must be false. Contradiction.

Then *Cotenability* is required (to prevent the second step).

[11]See Brandom [23], p. 107.

Here is where Sellars shows his point. In Sellars [148] §§24-26, he maintains that Goodman, in this argument, confuses the roles of sets A and S and their logics: in fact (i), if true, is meant to be supported by a lawlike generalization which makes *impossible* the conjunction of A and $\neg C$. According to Sellars, Goodman fails to distinguish between natural laws like

1) If match M is dry then if it were scratched it would light

and

2) If match M doesn't light then if it were scratched it would not be dry

on the one side, and implications ("general hypotheticals") like

3) If match M were scratched *and* it were dry it would light

and

4) If match M were not dry *and* were scratched it would not light.

Here (1) and (2) have the form

a) $A > S > C$,

while (3) and (4) are of the form

b) $A \cdot S \supset C$.

Only in the mistaken interpretation of counterfactuals as established by principles in the form (b) a further principle would be required to avoid $A \cdot \neg S$ and rule out the contrapositive. To the contrary, on the modally robust interpretation (a) the one principle (1) rules out the other (2) because it is *incompatible* with it. Sellars thus introduces the idea that the generalization proposed by Goodman as an analysis of counterfactual conditionals expresses a *causal* law, which has to be formalized by some *stronger* conditional.

But then I can *see* Sellars's point here. In Goodman's argument the crucial step follows by *Contraposition*, so, from a logical point of view, the main problem is to show how to prevent it. Now, the failure of *Contraposition* is a well known common feature of nonmonotonic consequence relations, thus one could refute Goodman's argument simply by arguing that the counterfactual conditional "$>$" is *nonmonotonic*.

5.2.3 Objections to the nonmonotonic interpretation

I've presented Sellars's analysis of Goodman's account of counterfactuals as leading to a nonmonotonic interpretation of these conditionals, but maybe I ran too quickly and aggressively all the way down. Let me now consider some objections in order to clarify one minor and one major point.

Minor point first. I said that Sellars claims for a *stronger* conditional to express natural laws. But – one may wonder – is nonmonotonic conditional *stronger* or *weaker* than classical extensional one? David Makinson convincingly confronted this question in Makinson [91]. Very briefly, his answer is "both": nonmonotonic consequences are *weaker* since the rules they validate are less in number; nonmonotonic cosequences are *stronger* because the classical consequences of any given set of sentences is a subset of its nonmonotonic consequences.

Let's move to the major point then. In fact, more importantly, one might also complain that while my interpretation is not completely out of place[12], it has nothing to do with what Sellars is trying to explain here. And she might point at Sellars [148], §26, where he remarks that the correct logical form for (i) is (a) because ">" has to be interpreted as a *strict* implication. Now, I can obvioulsy agree that Sellars had modal logic in mind, since modal entailment in the '50s was probably the most clear example of that sort of nonmonotonic behavior which is relevant for his argument. This is no matter of concern for my point, but I'm interested in the reasons one could put forward to support this complaint because they introduce the major points I want to deal with.

Thus, first one might argue that my interpretation is ruled out by §§20-21 where Sellars compares

> "Matches *wouldn't* be dry, if they do not light when scratched"

with

> "Matches *cannot* be dry, if they do not light when scratched",

and explains why the latter, but not the former, is acceptable:

> «Beating about the bushes for other asymmetries pertaining to our familiar generalization about matches, we notice that while it tells us that scratching

[12]Nonmonotonicity of counterfactual conditionals is a well-trodden path. For an account of counterfactuals centered on the idea of revision see Gärdenfors [56]. For the full blown theory of belief revision see Alchourrón et al. [1]. For of a proof of the relation of this approach and nonmonotonic logic see Gärdenfors and Makinson [57].

> matches causes them to light, *it doesn't tell us the cause of matches not being dry*; and that while it enables us to explain the fact that a match lighted on a certain occasion by pointing out that it was scratched and was dry, *it doesn't enable us to explain the fact that a match was not dry by pointing out that it was scratched without lighting.* On the other hand, the generalization does enable us to explain *how we know* that a given match was not dry.»

Then, the objection would conclude, what I'm taking as a point about nonmonotonicity is actually the difference between the epistemic level of generalizations and the modal level of norms as already highlighted by Kripke. Notice how this echoes the problems with Lance and Hawthorne's point of view, analyzed in Section 1.3.2.2 above: norms are metaphysically determined out there where our epistemic generalizations cannot reach, so all we can do is to (acknowledge this and) account for what it is to treat our generalizations as explaining how we know what we know. In this sense, I would be confusing two sorts of *explanations*: on the one side a *semantic* explanation which deals with the relations between conceptual content and the occurrence of words in analytic sentences, on the other side an *epistemological* explanation which deals with the relation between conceptual content and the empirical experience of the world. While it's true that these use to interact profitably, nonetheless I would be overimposing the former on the latter by waving this idea of revisable content. As Sellars's argument goes on, in §24-29 he provides other examples of subjunctive conditionals as contrapositives of (i), namely,

> "*Since M did not light*, if it has *also* been the case that it was scratched, it would have been the case that it was not dry".

Now, I acknowledge that it is surely *these* sort of conditionals which put into play epistemic considerations. As Sellars notices, what they do is exactly to translate the issue about *causes* into an issue about *co-occurrences*. Notice however that they don't loose their modal strength, provided that *all* the *relevant* occurrences are taken into account in the premises as several *also*s. So, here is where the problem rises of the *knowledge* of the facts that *necessarily co-occur*: how can we know how to implement the *also*s' list? The only wayout Sellars suggests is to require that the context must provide the resources, in terms of presuppositions, to fill it in. Now, if this were Sellars's last word on counterfactual conditionals, the objection would be completely right. Doesn't this translation look like Kripke's 'extensionalization' of intensions? If this were Sellars's horizon, then to look for a *non-metaphysical* modality would be to try to search for something Sellars himself didn't find in his *causal* one.

Here is where the objector could try to hang another piece of reasoning to argue that this is *in fact* Sellars's horizon. So, in §25 Sellars claims that conditionals like Goodman's problematic example (2) above are false because in their form (a) they imply, for instance, that what one comes to know by them is that "the match is dry" can be inferred *without qualifications* from "the match is scratched". But what does Sellars mean with "without qualifications"? The objector would point at §38 and §62 to claim that Sellars can't but refer, at least in part, to those *contextual implications* that specify the other premises of the conditional and to the fact that, since the list of these premises might not always be complete, the validity of the conditional has to be established on *inductive* grounds. Then, to sum up, here is the picture of counterfactual conditionals described by the objector: they are *extensional* conditionals of form (b), their full set of premises is pragmatically presupposed and has to be integrated by *contextual information*, they are treated as – as opposed to just "are" – modally robust (in metaphysical sense) because their validity is established by *induction*.

Now, this view has certainly some misleading strength due to the fact that it is close to the letter of what Sellars actually writes in these paragraphs. However, as a matter of fact, it hinges on a misinterpretation. As Sellars explains in the last part of his paper, he construes these notions he deploys as lying on a pragmatic rather than epistemic plane. This will become completely clear when we'll deal with Sellars's interpretation of induction in the next section. For the moment let me lay down some minor observations that might begin to shake the apparent solidity of this picture.

To begin with, while Sellars dedicates several paragraphs to the analysis of the extensional translations of counterfactual conditionals it can't be oversighted – as I mischievously did above for the sake of the argument – that he explicitly considers those subjunctive conditionals just as "almost" acceptable[13]. The reason why I've become picky all of the sudden, is that I want to stress that the extensional settings these translations introduce – notice his examples: events that had been the case and numbers – are exactly those that, by formatting counterfactuals as according to form (b), mask the *qualifications* required for their proper understanding. And this is precisely what makes those contexts useful to make them stand out. Now, let's ask again: what are these qualifications? They are indeed the more premise(s) S required in a conditional of the form (b) to make the consequent C follow from A. But they are required *in this sense* only if counterfactual conditionals are formatted as (b), which, as he repeats in Sellars [148], §§29-30, is simply incorrect. The

[13]See Sellars [148], §24 and §26.

reason is that in a counterfactual conditional premises do not have all the same status and their logical relations can't be represented by simple extensional conjunction: modally robust counterfactual conditional make explicit conceptual content by specifying how the sorts of things to which those contents apply are disposed to behave given certain standing conditions: some of these conditions *relevantly* interact with what sort of thing the contents are about and what dispositions they exhibit. So, *in another sense*, when counterfactuals are properly formatted as (a), qualifications are required in order to specify this complex structure of relations. And indeed it's true, these qualifications are induced by the fact that counterfactual conditionals are established by *induction*, since induction according to Sellars involves a practical argument for the acceptability of a sentence given certain circumstances, as he explains in Sellars [156] and we are going to show in a moment.

5.3 Necessary and contingent constant conjunctions

The sellarsian path we took in the previous Section drove us to a clear cut distinction between two interpretative options in the analysis of counterfactual conditionals: either they have to be construed as nonmonotonic entailments or they have to be construed as inductively established generalizations. Notice how the choice between these two options has been chasing us since Chapter 1, where we dealt with the problem of the normative character of meaning. However, given what we learnt from Chapters 3 and 4 about the relations between normative vocabulary and modal vocabulary, it should not be surprising that this choice eventually presents us in terms of the analysis of the inferential structure of conceptual contents as represented by necessary laws or by contingent constant conjunctions.

5.3.1 Semantics vs. Epistemology

The interpretation of the inferences involving causal modalities is the crucial point in the fictional debate between Mr. C(onstant Conjunction) and Mr. E(ntailment) in the second half of Sellars [148]: the former maintains the classical humean position according to which lawlike statements are simply universal generalizations, the latter, who embodies Sellars's own position, argues for an analysis of lawlike statements in terms of the expression of the rules we use in reasoning. Sellars's strategy, in §74, is to represent these different interpretations of the content of lawlike generalization by making explicit the different interpretation of the inferences they are involved in. In this sense, as a confirm of Brandom's analysis,

Sellars's strategy is genuinely *inferential*[14]. Mr. C tends to unpack inferential rules that introduce and eliminate lawlike generalization according to a standard extensional schema,

$I.$ ...
So (in all probability) $(\forall x)(Ax \supset Bx)$
$II.$ $(\forall x)(Ax \supset Bx)$, Ax_1
So, (of logical necessity) Bx_1

To the contrary, Mr. E directly introduces counterfactual conditionals in order to represent lawlike generalizations, thus

$I'.$ ...
So (in all probability) if anything were A it would be B
$II'.$ Ax_1
So (of physical necessity) Bx_1

Now this pinpoints the difference but still doesn't explain it clearly enough. The conclusions of argument II and II' were expected to be different: they express the different position of Mr. C and Mr. E which is the theme of the debate. Sellars tells us that this divergence originates in the difference between conclusions of argument I and I': they express the different interpretation of inductively established lawlike generalizations. But we still miss the relation between the two groups of different conclusions that give reason to such a divergence.

One may think the two arguments could be rephrased respectively as

$I.$ $a)$ ...
$b)$ So (in all probability) if anything *is* A it *is* B
$II.$ $a)$ Ax_1
$b)$ So (of *logical* necessity) Bx_1,

and

$I'.$ $a)$ ...
$b)$ So (in all probability) if anything *were* A it *would be* B
$II'.$ $a)$ Ax_1
$b)$ So (of *physical* necessity) Bx_1.

[14]As noticed by Brandom, the difference Sellars states here between what is explicitly said and what is contextually implied by an assertion corresponds to the difference he stated in Sellars [148], p. 333, between what one says by making a statement and what one thereby conveys. However, even if Sellars doesn't seem to concern much with a technical use of the notion of contextual implications, Brandom considers it a crucial step for the argument of Sellars [148]. See Brandom [27], p. 15. I'll come back to this below.

In this sense the difference would amount to a distinction between *indicative* and *subjunctive* conditionals. Mr C. and the traditional empiricist would argue for a reduction of the modal vocabulary of subjunctive conditionals to the descriptive vocabulary of indicative conditionals. According to Mr. E and Sellars instead, the real point would be to show that it's not subjunctive conditionals that can't be explained in terms of the indicative ones but the other way around[15]. Thus the moral we should learn is to rehabilitate the status of modal vocabulary against traditional empiricist criticism. And one could feel comfort in this idea by noticing that Sellars introduces the defeat of Mr. C, in the last bit of the debate, just by proving he fails in producing a modal logical form for lawlike generalizations suitable to explain the difference between necessary and contingent constant conjunctions[16].

But I don't think this would be enough to settle the notion of causal modalities. Indeed, if one focuses on modal vocabulary, as modern modal logic teaches us to apply it, and considers its role in expressing the meaning of counterfactual conditionals, it seems to be unclear where Sellars wants to lead us with the whole debate between Mr. C and Mr. E. Let me review some of the options.

One can construe Mr. C as rejecting causal modality and arguing that probability statements are about a property of a state of affairs (i.e. descriptions of particular facts of the form $prob(p, e)$). But then Sellars shows that that is a wrong interpretation of probability statements[17].

One can construe Mr. C as rejecting causal modality and arguing that probability statements establish generalizations of the form $(\forall x)(Ax \supset Bx)$. But then Mr. C couldn't explain the difference between accidental and necessary constant conjunctions.

But if one concedes that a slightly smarter Mr. C could accept modalities but still not recognize material inferences then it suddenly gets hard to describe a plausible position for Mr. E. In fact, while Mr. C would maintain that there's just one kind of necessity and an epistemic process to discover it, Mr. E instead would maintain that there's another kind of necessity, a physical one, whose modal laws could be false in a non epistemic sense (i.e. not just liable to be discovered false).

[15] In Brandom's words this becomes the Kant-Sellars thesis. But notice again how the *pragmatist thesis* works in his interpretation: it's because inductively established lawlike generalizations express practices of entitlement attribution which have to be expressed by modal vocabulary that indicative conditionals are showed not to be autonomous from subjunctive conditionals.

[16] Sellars [148], §§69-72.

[17] *Ibid.*, §§60-61.

I suggest then that the real apple of discord is the interpretation of modality itself. In other words, in the above arguments we would have

$$(I.b)\ (\Box)(\forall x)(Ax \supset Bx)$$

and

$$(I'.b)\ (\boxed{c}\ )(\forall x)(Ax \supset Bx)^{18}$$

where still the difference between "$\Box$" and "$\boxed{c}$" (this latter being *causal* modality) has to be explained.

Now, those who maintain, with Mr. C, that there's no relevant difference between necessary and accidental generalizations can simply reject the idea of a middle-way modality. An entailment can be true or false, and if it's proved to be true it's necessary, as stated by the *Necessitation* rule for modal logic: from a logical point of view it's hard to make sense of something like a *physically* necessary statement that could be false. In fact, Mr. C is completely at home in the standard modal framework, and accuses instead Mr. E to postulate new weird modalities. Quite surprisingly then, there's seem to be no easy way to save Mr. E's position thus construed.

But there's a way indeed, according to Sellars: so let's take it from the beginning. Consider the particular sort of subjunctive conditionals which Sellars calls "subjunctive identicals"[19], for example:

[18] This formula requires some comment, since it hides a difficulty about the quantification under $\boxed{c}$, in the very sense we began to see in Section 5.2.3. Let me first say that, since Sellars [148] writes it simply as $(\boxed{c}\ )(\forall x)(Ax \supset Bx)$, I could have easily smoothed over this difficulty. But I think indeed that to raise the issue of the well-formedness of this formula is to try directly to penetrate the kernel of the issue about the logic of the normative character of conceptual content. What is it to *legitimately infer* from something being A that *it* is B? In modern times the most famous answer to this question was attempted by Frege in his *conceptual notation*. The problem deals with the generality that makes such an inference legitimate. Frege had *two* ways to account for such generality:

$$\vdash \begin{array}{l} A(x) \\ \lfloor\, B(x) \end{array} \qquad \text{and} \qquad \vdash_{\mathfrak{a}} \begin{array}{l} A(\mathfrak{a}). \\ \lfloor\, B(\mathfrak{a}) \end{array}$$

From the way Russell's theory of quantification has *mis*-taken Frege's original insight many of the problems originate of standard extensional framework we've been dealing with. See Macbeth [89]. It's interesting to notice how this very problem reappears in the literature on nonmonotonic logics when examples are provided to illustrate defeasible reasoning in terms of notions like "normality": the large majority of these examples deal with normal individuals having some properties but the standard apparatus of nonmonotonic logics works on propositional calculus. And indeed, at the moment, there's no well established theory of first order nonmonotonic logic. Given all this difficulties, in what follows I'll try to stick to propositional formalizations wherever I can. Where I can't, I'll rely on Frege's own original usage of latin letters to express that sort of generality.

[19] *Ibid.*, §98.

> "If anything were identical with one of the things which are (were, will be) ϕ, it would be ψ".

Given Leibniz's principle $x = y \Leftrightarrow (\forall\phi)(\phi x \leftrightarrow \phi y)$, they are clearly authorized by generalizations of the form $(\forall x)(\phi x \supset \psi x)$.

Consider now an extensional semantics in which a domain of objects $D = \{x_1, ..., x_n\}$ is *given* and concepts ϕ, ψ are interpreted as functions whose ranges are collections K of objects, i.e.

$$x \in K =_{Def} x = x_1 \vee x = x_2 \vee ... \vee x = x_n$$

and predication is represented as

$$\phi x =_{Def} x \in K.$$

Consider eventually the counterfactual

> "If anything were ϕ it would be ψ"

and its Stalnaker-Lewis analysis in the standard modal semantics of possible worlds. Clearly $(\forall x)(\phi x \supset \psi x)$ isn't enough to authorize it, because there could be a possible world w_i in which $(\exists x)(\phi x \wedge \neg\psi x)$. The reason, obviously, is that in such a world w_i the extensions of ϕ and ψ could be different from their the extensions in the real world w_r. In possible worlds semantics quantification works intra-worlds.

Coherently with this extensional framework Mr. C makes his point in terms of the range of the quantifier of $(\forall x)(\phi x \supset \psi x)$.[20] He sees just two possibilities: either natural laws have to be distinguished from logical laws (because, for example, they are not to be considered analytic) and thus quantifiers range just over *actual* particulars (everything that was, is and will be), or they are (at least, tend to be, in a regulative picture of epistemology) equivalent to logical laws and thus quantifiers range over *possible* particulars (everything that could be in any possible world). No other option is available[21].

[20] This point is very clearly taken in Sellars [143].

[21] Notice that the same point can be made in inferential terms. If induction has to be considered as an inferential process which could be in principle axiomatized in a system *In*, then inductively established laws have the form $\vdash_{In} (\forall x)(\phi x \supset \psi x)$. In this sense Mr. C could easily concede that $(\forall x)(\phi x \supset \psi x)$ express an accidental constant conjunction and that it wouldn't authorize any counterfactual. But what is a *necessary* constant conjunction then? Mr. C's answer could be that it is simply a constant conjunction that is valid in every possible world: inductively established generalizations express those necessary constant conjunctions because they have the form of $\vdash_{In} (\forall x)(\phi x \supset \psi x)$, and the *Necessitation* rule of modal logic, in fact, states that $\vdash_{In} (\forall x)(\phi x \supset \psi x) \Rightarrow \vdash_{In} (\Box)(\forall x)(\phi x \supset \psi x)$.

But, according to Sellars, if the notion of law of nature has to make some sense this position is untenable. The reason is an incoherence he already diagnosed in Sellars [143], where he characterized it in three theses:

a) Laws of nature are not analytic;
b) Laws of nature are not in their scope restricted to actual happenings;
c) Propositions about all possible particulars must be analytic if true.

Clearly one can't maintain all these coherently, thus some room must be made for a notion of necessity which doesn't coincide with logical analiticity[22].

In Sellars [148], the first step on the way out of this blind alley is to refute the idea that the *discrimen* between contingent and necessary constant conjunctions has to be researched in a higher degree of generality of the scope of their quantifiers: natural laws are not descriptive generalizations at all[23]. The argument here involves two sellarsian themes. The first one is the topic of Sellars [146], i.e. the idea that *material* inferences are *not enthymematic* inferences, but proper rules of inference which allow the usage of a vocabulary for counterfactuals. We've already dealt with this in Section 3.3.3. The second theme, explicitly presented in Sellars [156], is the idea that the main feature of the inductive process is to provide "inference tickets". Thus an induction establishing a generalization of the form $(\forall x)(Ax \supset Bx)$ is analyzed as a practical meta-argument establishing the acceptability of Bx given Ax[24]:

[22]In Sellars [143] room is actually made for physical necessity in the standard picture, but the argument that gives *reason for* this introduction is less satisfying. It is constructed to make a point in a metaphysical debate, and it takes for granted a certain methodological realism of logical particulars. It begins by asking in virtue of what a *universal* is the universal it is. The answer is that the distinctive property of a universal concerns its exemplification in all possible *histories* (possible worlds) in which it is exemplified. Thus possible particulars and universals go hand in hand. But now, if these histories were all logically possible histories then every universal would be exemplified in the same way, in the sense that there wouldn't be any precluded exemplification. So the question becomes: "How can the number of possible histories be fewer than the number of conceivable histories?" The answer to this latter question is that *families of histories* must be carved within the space of possible histories. These *families* correspond to invariances in the exemplification of universals. The key feature of this picture is that no universal can be considered outside the family to which it owns its identification. While the overall picture may seem to be convincing, two questions are left unanswered: (a) how these families are determined, and (b) how different histories can be compared. Question (a) deals with the very distinction of lawful and non-lawful generalizations: what makes the difference between "Every coin made of the metal whose atomic number is 29 melts at 1083 °C" and "Every coin in my pocket melts at 1083 °C" if there are no rules to carve out the family of the universal "copper"? Question (b) deals with the problem of theory change: if universals can't be identified outside their family of histories, how can their meaning change in the epistemic progress? I hope I can hint at a possible answer for this sort of questions by the end of this Chapter.

[23]See Sellars [148], §98.

[24]This is an instance of the difficulties I anticipated in Note 18 above: if we were to introduce extensional quantification here we would immediately loose the normative character of the inductively established

$$\begin{array}{c} x \text{ is } A \\ \vdots \\ \text{That } x \text{ is } B \text{ is acceptable} \end{array}$$

so that in the presence of an x which is A one is *entitled* to infer that it is B.

Notice how these two points interact the one with the other: the practical meta-argument is not a premise to be inserted into an enthymematic argument from Ax to Bx, but the expression of an inferential pattern that makes it reasonable, when asked "Why do you claim that x is B?", to answer "Because *it* is A"[25]. In fact, Sellars's decisive move to introduce Mr E's victory involves his interpretation of induction:

> «the statement 'Being A physically entails being B' [...] *contextually implies* that the speaker feels himself entitled to infer that something is B, given that it is A.»[26]

The idea, again, is that inductive inferences *entitle* speakers to assertions. Mr. C maintains that inductive inferences establish laws in the form $(\forall x)(Ax \supset Bx)$ and considers this a required step for the validity of counterfactual conditionals: induction has to establish major premises of otherwise enthymematic inferences like "if you scratch this match it will light". On the contrary Mr. E believes that the inductive process supports principles according to which we reason.

But – the smarter Mr. C might complain – what is the cash value of all this? No matter how much normativity you want to dress necessity with, it won't change the fact that if extensional generalizations like $(\forall x)(Ax \supset Bx)$ are part of the meaning of the causal implications expressed by counterfactuals, then $\Diamond\!\!\!\!-\!\!\!\!\supset (\exists x)(Ax \wedge \neg Bx)$ is simply ruled out.

Now, as it's explained in Sellars [148], §§ 92-100, this is just not true: as a matter of fact a universal generalization is necessarily true if and only if it is true for any possible entity and nothing like that can be established by induction. The reason is not epistemic uncertainty, but the fact that if we could check all possible entities we *couldn't* inductively establish anything. Lawlike statements, Sellars concludes, simply do not have the form of universal generalizations, "with or without contextual implications". Quantification belongs to an idealized, expressively complete, representation of language, and it presupposes

inferential pattern. The usage of x here has to be better construed as the representation of a sort of generality akin to that of anaphoric relations: pick a certain x and say *it* is A, then say *it* is B. See Chapter 7 of Brandom [14].

[25] See Brandom [13].

[26] Sellars [148], 75 (my emphasis).

rather than explain the modal features which express linguistic norms already at play in material inferences[27]. The extensional definition of concepts represents an idealized phase of linguistic practices in which conceptual resources are completely and definitively made explicit. Instead, the key feature of these principles is that they are *fallible*, and in this sense the statements they express, though normatively robust, might *go wrong*. To repeat, this has nothing to do with our epistemic uncertainty about their validity: in this sense "Emeralds are green" is on the same boat with "Birds fly". They just have possible defeasors. The point is, in Sellarsian terms, that we know how to apply these concepts because we know what follows from them, or in terms of *IS*, we know how to define these ineferential contents because we know what they are incompatible with.

So the difference between the two formalizations of $(I.b)$ and $(I'.b)$ depends on the different entailment considered: $(I.b)$ include classical consequence relation "$\models$" and the modality it express is *logical*, $(I'.b)$ include nonmonotonic consequence relation "$\mid\!\sim$" and the modality it express is *physical*. The difference would be between two kinds of entailment:

i) $A \models C$

and

ii) $A \mid\!\sim C$

where (i) implies $A \wedge S \models C$, but (ii) doesn't imply, *without qualifications*, $A \wedge S \mid\!\sim C$.

5.3.2 The logic of thing-kinds

It's time to take over the argument against anti-normativism just where we left it. In Chapter 1 it was claimed that the intuitive basis of an account of conceptual content lies in the relation

Fs do f,

where "f" is what counterfactually characterizes things of kind F. While that was enough to highlight normative features of conceptual content, it was admittedly very rough. We are now in the position to dig deeper. Let's begin by adapting our relation to the present discussion on counterfactual conditionals, and improve it, with Sellars [148] §31, by specifying "f" as

[27] See *ibid.*, §105.

ϕ-ing Fs causes them to ψ.

Sellars himself gives us a piece of an account of conceptual content which this representation would fit in:

> «Suppose [...] we have reason to believe of a particular thing of kind K, call it x_1, which is in C, that
>
> x_1 would ψ, if it were ϕ-ed.
>
> And if it were ϕ-ed and did ψ, and we were asked "Why did it ψ?" we would answer, "Because it was ϕ-ed"; and if we were then asked, "Why did it ψ when ϕ-ed?" we would answer "Because it is a K". If it were then pointed out that Ks don't always ψ when ϕ-ed, we should counter with "They do if they are in C, as this one was".»[28]

Doesn't this story remind you about John's dialogue about green things with his friend in the tie shop? Now, the cunning reader may already have perceived here the basis for a thesis about the logic of conceptual content, but it's worth securing these foundations a little before we try to build anything over them.

Notice, with Sellars, that it wouldn't do to treat this formulation as a straight definition of conceptual contents, i.e. as

Definition 46. $K(x) =_{Def} \phi(x) \mathrel{|\!\sim} \psi(x)$[29].

But why? Isn't it a proper way to express the idea that Ks are those things that *normally* ψ when ϕ-ed? Consider our "green things" for instance: aren't they those things that normally are found to be green when inspected? So emeralds, italian flags in certain parts, leaves for certain periods, traffic-lights at certain intervals, etc. But that doesn't seem completely right. Wouldn't it be better to say that green things are normal instances of that green stuff which is always in emeralds, in part in italian flags, for certain periods in leaves, at certain intervals in traffic-lights, etc.? Enough. This was just to show how easy is to slip back into the extensional ravine. To be sure the idea that

[28] Sellars [148], §31.

[29] To be precise this remark introduces the analysis of two connected but different points in the last paragraphs of Part II of Sellars [148]. However, the first one, in §§34-43, is once again the inappropriateness of the *monotonic* character of *strict* entailment to the representation of *nonmonotonic causal* entailment. Sellars takes care to stress the crucially different role played by the pieces of his formula in the definition of thing-kinds, since that is the only way he has to impose defeasibility on strict implication. We are not in that position however, since we dispose of nonmonotonic entailment. I'll come back to this in Section 5.6. Here we will directly focus on the second point, which Sellars deals with in §§43-45.

"x is F"

is to be construed as

"There is an F-thing in x which is responsible for its being F"

is a common pattern of reasoning at least since Aristotle's theory of predication, and it's hard to avoid it. But after all what we said about inferentialism and against extensional semantics, we just can't run the risk of adopting an analysis of conceptual contents that may ingender this mistake.

So, what did go wrong? To understand that, as Sellars suggests by referring to his Sellars [149], one has to directly tackle Artistotle on this point. Reasons of space force me to be very rough in summarizing the point, so I'll try to dribble any question about the metaphysics of thing-kinds and focus on their logic. The question to ask is obviously what is x (*ti esti*)? According to *Categories* 2b4ff, there are two ways to predicate what x is, namely e.g.

x is green

and

x is *a* green.

In fact, supposing that x is a proper subject of predication, i.e. a primary substance, the predicate "green", as every-'thing', might either be "present in" or "said of" it. Among the things which are "present in" a primary substance, only the *name* sometimes can be "said of", but never can the *definition*. So it is legitimate to pick Goodman's emerald and say "this is green" but not to say "this is *a* green" or worse "this is *a* colour". Here is were the path gets slippery, for one might be tempted, as Aristotle was, to ask why it is legitimate to say of Goodman's emerald that it is green while it is not *a* green, and to answer, as Aristotle did, that it must be because there's another *thing* which is *a* green and is "present in" this emerald. But this move is uncompelling: *we* have inherited from Chapter 3 a better inferential explanation, so we can easily avoid the pitfall.

Then, going back in topic, we can't accept Definition 46 because thing-kinds, which are properly "said of" have to be distinguished from mere properties, which are merely "present in". But how to do that? Notice this question has two senses. In a first one, it amounts to the request for criteria to list down all thing-kinds: that is something Aristotle might have taken as an issue worth worrying about since he probably didn't conceive evolution

neither of kinds themselves nor of our *epistheme* of them. However, and even if we don't share Sellars's interpretation of the progress in sciences, we can safely treat this as an epistemic issue. But in a second sense that is a genuinely logical issue, that is: what is the proper way to formally represent the distinction between kinds and mere properties? Notice this is a crucial question because it is only by making this distinction explicit that our formalization may aspire to represent the revisability of conceptual contents: as Sellars notices,

> «It is the regulative connection of thing-kind words with the schema
>
> Ks ψ when ϕ-ed, in appropriate circumstances
>
> which guides them through the vicissitudes of empirical knowledge.»[30]

Thus, we logically distinguish kinds from characterizing inferential properties by distinguishing their behavior with relation to nonmonotonic entailment. So, what is x? X is *an* F, where, following both Sellars and (his interpretation of) Aristotle, F is defined as

Definition 47. $F(x)$ iff $K(x), C(x) \mathrel{|\!\sim} \phi(x) \supset \psi(x)$

5.4 Towards a semantics for causal modalities

The next step is to put some contentful flesh on the logical bones of causal modalities we have been describing. In order to do that I have to take a step back and focus on the main theme around which Brandom [27]'s interpretation of the sellarsian path we've been following hinges.

Let me start by noticing that causal modalities here seem to play the role of *unexplained explainers*: they are required to account for the meaning and the logic of counterfactuals, but Sellars seems much more concerned with defending their role against the reductionism of traditional empiricism than with explaining what their content is[31]. Now, this might seem a silly expectation: what else but the empirical enterprise of natural sciences could ever tell that? That's quite obvious and it's not my point. What I want to ask is how to put causal modalities into play, given that we know they are required for us to play the empirical game. This is a less silly demand. The problem is that even if we accept that causal modalities are what actually supports our representation of natural laws and

[30] Sellars [148], §48.

[31] Please notice that I intentionally avoided to write, more plainly, "what they are". I did that in order to avoid ontological misunderstandings: I'm not calling for *causal things* out there, so don't submit any.

in general our claims about thing-kinds, there seems to be no obvious way for Sellars to explain how we could *express* them without falling prey of the smarter Mr. C's criticism: if we try to say that we feel entitled to assert that matches light if scratched we have to *say* that a certain inference is valid, but if such an inference has to be modally robust it will be *metaphysically* modally robust (given our present logical resources). We learnt a solution to this problem which was not in Sellars's tool-box: we can try to express it with a nonmonotonic inference. Now I want to make it clear what it means to do that.

Remember how Brandom [27] describes Sellars's criticism against traditional empiricism as a complex argument line based on three main theses: an inferentialist thesis, a pragmatist thesis and a modal thesis. The *inferentialist thesis* is patent in what we've just said. Now, it's quite obvious that the *pragmatist thesis* would entitle Sellars to argue straightforwardly that the causal modalities make explicit the underlying normativity of linguistic practices of descriptive vocabulary. In effect this is exactly what Brandom thinks to be the case. He wants to parallel the argument against the primacy of observational vocabulary in Sellars [147] with the argument against the primacy of purely descriptive vocabulary in Sellars [148]: he wants to say that Sellars proves that non-observational vocabulary is *VV-necessary* for observational vocabulary in the same way as modal vocabulary is *VV-necessary* for descriptive vocabulary, *because* the practices which are *PV-sufficient* for the first vocabulary in each couple are *PP-necessary* for the practices *PV-sufficient* for the second one[32]. So, he pinpoints Sellars's use of the notion of "contextual implication"[33] as a way to introduce a sort of pragmatic inference, in order to account not only for the content of counterfactual conditionals, but also for what one is *doing* in asserting them.

Let's try then to apply the *pragmatist thesis* in the context of Sellars [148]. In this sense, Sellars would maintain that to assert, for example, the inductively established *lawlike* generalization expressed by the subjunctive conditional

"If the match M were scratched it would light"

is to implicitly endorse the the normative licence to treat scratched matches as inflammable. Such an endorsement implicit in practice could be made explicit by the notion of "physical entailment" which we formalized as nonmonotonic entailment:

"Scratching the dry match M *physically entails* it to light".

What is crucial to realize is that this is not yet to make explicit the inferential license, which would have the standard modal form

[32]For a quick recap of the meaning of these relations just look back at Section 3.2.3.
[33]See the quotation in Section 5.3.1.

"If the dry match M is scratched then *necessarily* it lights",

but to *contextually imply* that the speaker commits to such an inference.

This is Brandom's interpretation. Now, personally I don't think that the notion of "contextual implication", as Sellars uses it, should be considered in itself the herald of the *pragmatic thesis* in Sellars [148], because it really seems to be more akin to those pragmatically determined aspects of linguistic content which are usually designated with "conversational implicatures". These don't have much to do with *pragmatist thesis* as the idea that semantic vocabularies make explicit what speakers do in their linguistic practices. But I *do* think Brandom is right, from his perspective, in insisting on the role of the *pragmatist thesis* here. Indeed, however things are with contextual implications, there are other reasons to dispel any doubt about the presence of the *pragmatist thesis* in this piece of sellarsian reasoning. Among these there's the fact that the seeds of a pragmatic interpretation of inductive inferences soughed here will grow into a full blown theory of induction in his Sellars [156]. There, in dealing with the meaning of inductive inferences, he will distinguish among the "proximate" the "practical" and the "terminal" outcome of inductive reasoning (where "outcome" here is explicitly used to broaden the context in which the consequences of inductive inferences have to be considered): while the *proximate outcome* corresponds to the logical conclusion, the *practical* an the *terminal outcome* are respectively the conclusion of a practical inference regarding the acceptability of the *proximate outcome* and the state of actually accepting such a conclusion. Thus, according to him, it is the pragmatic result of endorsing the content of a statement as a result of an inductive inference that unveils the path for a proper account of the modal entailment expressed in counterfactual conditionals. Again, one of the merits of Brandom's analysis is to point out the pivotal role of the *pragmatist thesis* in the sellarsian line of argument, but, I think, this is only part of the story. The *pragmatist thesis* would grant that counterfactual conditionals acquire their modal force from the underneath normative practice that inductive inference make explicit, but it can't explain the logic of such a modality: the *pragmatist thesis* by itself wouldn't make any difference between $\Box$ and $\boxed{c}$. If we rehearse the position of the 'smarter' Mr. C described in section 3, we see again that this opponent could just endorse the *pragmatist thesis* and accept the modal force of lawlike generalizations as derived from normative practice, while still denying any sense to the notion of *accidental* but *universal* generalization.

We already saw that the key feature of these practices, for what concerns our discussion here, is that they are *fallible*. The point is to understand what it means. Mr. C

could simply recognize that this fallibility mirrors the epistemic uncertainty of scientific enterprises: lawlike generalizations are established with a degree on uncertainty which is properly accounted for by inductive logic, and modality (the only metaphysical one) simply represents the ideal phase of the scientific enterprise in which probabilities have converged to 1. This is a heavy claim indeed, so massive that it inevitably attracts any other position gravitating nearby and makes it collapse. We saw that that was the case of Lance and Hawthorne's attempt to lift a stable ground over analyticity with the pragmatic lever of a normative interpretation of meaning. The point is, that is not what it means for normative practices to be fallible: commitments and entitlements do not admit degrees, rather it is the scorekeeping practice to be perspectival. This perspectival nature accounts for a defeasibility which would not rule out exceptions to modal conditionals. This kind of defeasibility is what we were looking for as a feature of laws of nature.

Some aid and comfort to this complaint come just from what Sellars says about statistical induction at the end of Sellars [148]. Let me quickly recap. It seems that the following couple is unteanable:

(i) In all probability $(\forall x)(Ax \supset Bx)$;

(ii) $(\diamondsuit)(\exists x)(Ax \wedge \neg Bx)$.

This is because (i) amount to $prob(Ax \supset Bx) = 1$, and any evenience of $Ax \wedge \neg Bx$ would irreparably lower that number, in the sense that it would be $0 \leqslant prob(Ax \supset Bx) < 1$ and nothing could take it back. Now, clearly (i) might *turn out to be false*. Yet, according to probability theory no statement which might turn out to be false would have probability 1. In fact no inductively established generalization has probability 1. But then no inductively esablished generalization would express the entitlement to assert of something that it is B given that it is A. To repeat, there's no commitment to p s.t. $0 \leqslant C(p) \leqslant 1$ since commitments (in this sense) do not come in degrees.

Brandom [23] showed how an inferentialist semantics, intended to formalize the norms underlying linguistic practices with the notion of incompatibility could represent modal logic. Unfortunately it doesn't deal with the perspectival nature of those norms. Thus while *EIS* has been proven to represent the kripkean metaphysical modality, it is not suitable to make explicit this sort of *causal* modalities.

Again, both parts of the story have to be considered. The lesson to be learnt from the case of Lance and Hawthorne in this sense is paradigmatic. You can't apply the pragmatic lever to lift analyticity alone because, as long as modal semantics is your vocabulary to

talk about meanings, metaphysical modality would drag it down: you have to put your pragmatic lever under modal vocabulary too. If we accept that there are no words to tell this as a unique whole story, we can just be content to wave our hands in the direction of such a defeasible analyticity anytime we try to make explicit the normative content of our descriptive assertions. That is somehow what Brandom draws from Sellars [148]. And yet it really would be a frustrating outcome for this whole project. I want to *say* what I mean when I say that scratching dry matches causally necessitates them to light. And I'll look for the words to say it in a *nonmonotonic incompatibility semantics* .

5.5 How to go nonmonotonic

Reasons have been put forward to look for a better comprehension of the normativity of meaning in the nonmonotonic character of material inferences: the inference from p to q is materially good even if the inference from p and r to q might not be. In the sellarsian account these are not enthymemes but fully valid inferences to be made explicit as modally robust yet fallible entailments. They are *safe* enough to support norms. But how to turn all this in cash value?

Let me be basic. *Monotonicity* is a property of consequence relations. How are consequence relations things that have properties? Consider first those formal representations of language which logicians work with: a fomal language is a set $\mathcal{L}$ of formulae picked off the possible combinations of a certain vocabulary because they are well formed according to certain criteria. A consequence relation is a relation "$\mathrel{|\!\sim}$" between sets of well formed formulae in $\mathcal{L}$. In Chapter 4 we saw how a relation of this sort is defined for *IS*. A relation has certain properties in the sense that it satisfies certain rules. For instance the *classical* consequence relation "$\models$", as defined by Tarski [164], satisfies the following properties:

(Reflexivity) $A \models x$ iff $x \in A$
(Transitivity) if $A \models b$ for all $b \in B$ and $B \models x$ then $A \models X$
(Monotonicity) if $A \models x$ and $A \subseteq B$ then $B \models x$

Nonmonotonic consequence relations do not satisfy *Monotonicity*. Nonmonotonic logics represent valid inferences in terms of nonmonotonic consequence relations. This means that, in these logics, a formula $x \in \mathcal{L}$ is accepted as an inferentially valid consequence of a set of premises A even if x is not, in general, an inferentially valid consequence of a superset $A \cup B$ of A. Thus nonmonotonic logics are suitable to represent fallible reasoning. However,

as Makinson uses to notice, "nonmonotonic logics" is an unfortunate label for these systems which do not just drop *Monotonicity* but apply weaker forms of such a principle.

Notice that *Monotonicity* is not equivalent to the rule of *strengthening the premises*: $a \models x \Rightarrow a \wedge b \models x$. This is a rule for the connective of conjunction. However, as we already saw in Section 4.2.5, the two are equivalent in those systems, like *IS*, in which conjunction is interpreted as the *intersection* of *sets*.

5.5.1 Nonmonotonicity and relevant reasoning

5.5.1.1 Ranges of counterfactual robustness

Let me go back to Brandom's interpretation of Sellars's causal modalities. We've just seen that his account is compatible with the downplaying thesis of our smarter Mr. C. As a matter of fact some of his claims seem to resonate to this idea. For instance, in Brandom [23] he rejects the idea that counterfactual robustness could be the criterion to distinguish lawlike from accidental generalizations. Thus

> "All samples of copper melt at 1083.4 °C"

should be lawlike because it supports the counterfactual

> "If this silver coin would be made of copper it would melt at 1083.4 °C",

while

> "All the coins in my pocket are made of copper"

should be accidental because it fails to support the counterfactual

> "If I were to put this silver coin in my pocket it would not be made of copper";

but such an accidental generalization actually does support counterfactuals like

> "If I were to choose a coin at random from my pocket it would be made of copper"[34].

According to Brandom *any* material inference which makes explicit the content of observational vocabulary is counterfactually robust to the extent that "If p then q" is incompatible with "It is possible that p and *not* q": in order to commit to the content of an inference

[34] See Brandom [23], pp. 104-106.

one has to realize which are the circumstances that would possibly defeat it. That is the same idea that motivates the determination of semantic interpretants in *IS* and leads to monotonic modal entailment of *EIS*.

However, in that very piece of reasoning Brandom claims that material inference is *nonmonotonic*. This clearly strikes an incongruous note[35], but it's easy to detect what Brandom has in mind if we just follow his reasoning. So, he admits that, although any material inference has to be associated with a set of possible defeasors in order to be contentful, and although the task of modal vocabulary is to make such a content explicit by identifying all of them, the actual cognitive resources speakers deploy in linguistic practices are way poorer that what would be required for the expletation of such an explicitating task. Speakers, as cognitively finite reasoners, can't but face the problem of updating their holistic system of beliefs by deploying defaulting strategies: what speakers do then is to associate to any inference just the set of possible defeasors which are *relevant* enough to jeopardize the goodness of the inference in the given occasion.

Consider again our counterfactually robust inference about lighting matches: "If this match were scratched it would light". There are clearly some circumstances that might defeat it: the match being wet, the match being under a strong magnetic field, there not being enough oxygen in the air, and so on. But there are also a lot of other circumstances which are irrelevant for the inference to go through: the match being scratched at noon, me thinking about this chapter, a butterfly taking off from a flower in another corner of the earth, and so on. Thus speakers committed to the lighting of the match should consider a revision of their inferences in case they happen to learn something new about the oxygen but not about the butterfly, since the former but not the latter is a relevant possible defeasor for the inference.

5.5.1.2 Semantics for *Relevance Logic*

This sort of nonmonotonic behavior is typical of the semantics for *relevant* logics. Urquhart's operational semantics (Urquhart [168]) is particularly clarifying in this sense. Operational semantics is a set theoretical semantics based on two notions: a *frame* K for a language $\mathcal{L}$, a set whose members are construed as pieces of information, and a *model* on that frame, i.e. a relation $\Vdash$ that associates pieces informations with formulae in $\mathcal{L}$. Thus $a \Vdash A$ is construed as "A holds according to the piece of information a". Urquhart's basic move to

[35] However, let me smooth over this here as if it were a merely technical flaw. I'll come back on this point below in Section 5.6, where I'll be able to rise a more articulated complaint.

obtain relevance is to establish the following condition for implication:

(→I) $x \Vdash A \to B$ iff $\forall y \in K\ y \Vdash A \Rightarrow x \sqcup y \Vdash B$.

Here the operation "$\sqcup$" represents the combination of pieces of information and it is characterized by the following axioms

(Identity)	$\varnothing \sqcup x = x$
(Commutativity)	$x \sqcup y = y \sqcup x$
(Associativity)	$x \sqcup (y \sqcup z) = (x \sqcup y) \sqcup z$
(Idempotence)	$x \sqcup x = x$

Crucially, it doesn't satisfy Kripkes's *Hereditary Condition*:

(HC) If $x \Vdash A$ then $x \sqcup y \Vdash A$.

It's easy to see how this semantics fits with the relevantist's complaints about the paradoxes of implication. The failure of (HC) and the commutativity of $\sqcup$ guarantee the failure both of $A \vdash B \to A$ and of $A \vdash B \to B$:

$$A \nVdash B \to A \text{ iff } \forall x(x \Vdash A \Rightarrow \exists y(y \Vdash B \text{ and } x \sqcup y \nVdash A))$$

$$A \nVdash B \to B \text{ iff } \forall x(x \Vdash A \Rightarrow \exists y(y \Vdash B \text{ and } x \sqcup y \nVdash B))$$

In fact, even if a piece of information x determines A, it might well be that when combined with another piece of information y that determines B, x won't determine A or y won't determine B anymore.

Notice that Urquhart's semantics defines a semilattice $(K, \sqcup, \varnothing)$. Notice also that in this structure the lattice zero, i.e. the null information, can be construed as representing *validity*, in the sense that for instance if $x \Vdash A$ then $\varnothing \sqcup x \Vdash A$, thus $\varnothing \Vdash A \to A$.

This solution hinges exactly on the idea that the process of updating the information required for the justification of one's beliefs may likely yeld to their revision. The leading idea here, as in modal logic, is that the very addition of any new information might just jeopardize the goodness of an inference, thus the nonmonotonic behavior is obtained by blocking *weakening on the left*, i.e. $x \circ y \nvdash x$. As we saw in Section 4.2.5.2, in terms of properties of a consequence relation this means to reject full *Reflexivity* $\Delta, x \vdash x$ and accept it just in the very weak form $x \vdash x$. In this sense

Let's go back to Brandom now. He claims that it would be incongruous to the limited cognitive resources of human reasoners to require that every such potentially harmful

information y must be ruled out before any inference can be drawn from a set of premises containing x. Thus, in this logical perspective, the proposal would be to weaken the request for relevance to a cognitively manageable set of possible defeasors: when y is not in such a class just let $x \circ y \vdash x$ go through. As we've just seen, a clear example of this situation is the case of null information.

5.5.1.3 Problems with *IS*

Now, in the halls of the Cathedral of Learning in Pittsburgh, where Nuel Belnap has been working and teaching for years, *Relevance logic* would probably be happily welcomed as the proper tool to develop this point. In fact it has been already used to represent Brandom's notion of committive inference[36], and, as a matter of fact, the whole system of *IS* with its definitions and proofs is deeply entrenched in this tradition.

Unfortunately, as it should be clear after Chapter 4, this door is shut in *IS*. Let me just sketch the obvious required connections. Consider first the introduction of intensional conjunction as a way to make explicit the underlying combination of pieces information:

($\circ$I) $x \Vdash A \circ B$ iff, where $x = y \sqcup z$, $y \Vdash A$ and $z \Vdash B$[37].

Then consider the implicational lattice defined by this model on the formulae in $\mathcal{L}$, and notice that (i) *right-residuation* holds between "$\rightarrow$" and "$\circ$", and that (ii) "$\circ$" doesn't validate *lower bound*[38].

This is exactly the same situation we encountered above in Section 4. There we improved our expressive resources by recurring to relational semantics, because that let us

[36] See Lance [81].

[37] Notice here fusion $p \circ q$ is not just $\neg(p \rightarrow_R \neg q)$: in that case $(\circ, \rightarrow_R)$ wouldn't be a *residuated pair*.

[38] Let me briefly illustrate.

Fact. *$A \circ B \vdash C$ iff $B \vdash A \rightarrow C$*

Proof. ($\Rightarrow$) Suppose $\forall x$ s.t. $x = y \sqcup z$ and $y \Vdash B$ and $z \Vdash A$, $x \Vdash C$. Suppose then, to the contrary, that for some $k \Vdash B$, but $\exists m(m \Vdash A \wedge k \sqcup m \nVdash C)$. But this means that there is some $x = k \sqcup m$ s.t. $k \Vdash B$ and $m \Vdash A$ but $x \nVdash C$. Contradiction.

($\Leftarrow$) Suppose $\forall x$ if $x \Vdash B$, then $\forall y(y \Vdash A \Rightarrow x \sqcup y \Vdash C)$. Suppose then, to the contrary, that for some z s.t. $z = k \sqcup m$, $k \Vdash B$ and $m \Vdash A$ but $z \nVdash C$. But this means that there is some $x = k$ s.t. $x \Vdash B$ and there is an $y = m$ s.t. $y \Vdash B$ but $x \sqcup y \nVdash C$. Contradiction. □

Fact. *$A \circ B \nvdash A$*

Proof. According to ($\circ$I), $x \Vdash A \circ B$ iff, where $x = y \sqcup z$, $y \Vdash A$ and $z \Vdash B$. But there's no guarantee that $z \Vdash A$ too, and so, since (HC) fails, there's no guarantee that $x \Vdash A$. □

force intensional behavior on compatibility. So, one could suggest, why don't we take the relational way again? As a matter fo fact such a way is well trodden: it was opened by Routley-Meyer semantics for relevant logic in Routley and Meyer [138, 139, 140] and developed by Mares [92]. This semantics is based on the notion of a *relational frame* $(K, R, 0)$, where K is again construed as a set of pieces of information (or "set-ups", as Meyer called them), R is a ternary relation on K and $0 \in K$. The relation R mimics Urquart's combination of pieces of information, so that $Rabc$ if $a \sqcup b = c$, and the axioms for R are purported to represent the behavior of Urquhart's semilattice. The definition for relevant implication is consequently:

($\rightarrow$I) $x \vDash A \rightarrow B$ iff $\forall x, b, c \in K$(if $Rxbc$ and $b \vDash A$ then $c \vDash B$)

Indeed the path might seem promising. Nonetheless, while Peregrin was able to provide a neat interpretation of accessibility within the basic resources of *IS*, I can't figure it out what the ternary relation required for relevant semantics would correspond to in terms of incompatibility relations[39]. Now, before any diligent philosopher comes to my rescue by putting forward some clever interpretation, let me suggest why such an enterprise might not be worthy of the effort.

A technical point first. Suppose we had a technically successful and philosophically satisfying *relevant relational incompatibility semantics*. Then, admittedly, the ternary accessibility relation would give us a lot to play around with (and things would get even more exciting with negation). But still the overall picture wouldn't change in the only detail that concerns us here: inference would be unregimentedly nonmonotonic everywhere but in the one single fully-monotonic case of logical validity with null information. Thus, to the contrary of what happens for modal logic, relational semantics here wouldn't solve our problem.

Let me also put forward a second theoretical point. To repeat, in Brandom's picture nonmonotonicity of material inferences depends on a restriction of the class of possible defeasors in order to obtain a cognitively acceptable balance between *monotonicity* and *relevance*. He has reasons to push in that direction: he wants to challenge Sellars on the criteria to identify material inferences. These criteria, Brandom claims, can't just be counterfactual robustness, since even accidental generalizations do support certain counterfactuals. So, for instance "all the coins in my pocket are copper" supports "if I were to pick a coin at random from my pocket it would be copper". Brandom concludes it must

[39]Nor Brandom and Aker see any practicable path: see [23], pp. 173-175.

be pragmatically determined criteria of restricted relevance that specify *certain ranges* of counterfactual robustness. Now, recall that Sellars takes into account these generalizations, whose accidental validity depends on a proper restriction of the domain class, and names them "subjunctive identicals": he considers these modally robust statements as a wrong way to represent causal modalities just because their extensional character blurs the difference between the defeasible modality he's chasing and the standard metaphysical one. The analysis of nonmonotonicity as it shows up in relevant reasoning suggests that Sellars (or me, interpreting him) might be right in considering extensional restriction of possible circumstances the wrong way to go in a set theoretical framework. But that is also just to say that we need a better grasp on these criteria for *restricted monotonicity*.

5.5.2 Nonmonotonicity and defeasible reasoning

5.5.2.1 A different approach to irrelevance

Normativity of meaning taught us to be careful about composition of relevant information while drawing our inferences. In ideal conditions one should take into account any relevant information in order to be entitled to draw an inference. Yet there are compelling reasons, e.g. Brandom's scruples about cognitive resources of human agents, for considering purely relevant reasoning a too tight representation of inferential practices. That raised an issue about nonmonotonicity.

Let me introduce a little twist in this picture by quoting Karl Schelchta's formulation of the "question about irrelevant information":

> «Given T and a possible conclusion ϕ, what information can be added to or subtracted from T without changing the fact that $T \mid\!\sim \phi$ (or $T \mid\!\not\sim \phi$)? Less, abstractly, given a (large) database T, and a query ϕ, is ther a method to single out a (considerably smaller) subset $T' \subseteq T$, such that the information contained in T' suffices to answer the query ϕ? Can we perhaps give a generic procedure, which for a query of type x singles out an appropriate $T_x \subseteq T$?»[40]

So, here's the twist: a "generic procedure" is needed to prune the whole set of information from all those contents which can be *treated* as *irrelevant* for the goodness of a *certain* inference. From a formal point of view this procedure can't but be represented as a

[40] Schlechta [142], p. 6.

selection or *default function* $\delta : \wp(\mathcal{L}) \to \wp(\mathcal{L})$[41]. However a default function is not enough: we need the selection to be *appropriate* to the specific inference. This is what puts the spin on the nonmonotonic approach to defeasible reasoning.

Now, there's a flourishing vegetation of theories which represent different interpretations of this appropriateness with different nonmonotonic logics, and their botanization is a valuable and intriguing task in itself[42]. But I won't enter the path of a comparison among these logics in order to find the best match with *Incompatibility Semantics*. I'll rather try to build what I need directly from the raw materials this field generously provide (refinements will be considered only later when required). In this way I'll obtain pretty standard results, at the obvious price of roughly smoothing over the analysis of the features that make them standard among the others.

Let's ask then: what does it mean for a relevant selection of contents to be appropriate to a specific inference?

We know that in order to represent our nonmonotonic inference $T \mathrel{|\!\sim} x$, our selection $\delta(T) \subseteq T$ must single out relevant content s.t. $\delta(T) \vdash x$ even if it might be that $T \nvdash x$ for some irrelevant $X \subseteq T - \delta(T)$. So the question turns out to be how to make sense of this situation. Recall from the previous Section 5.5.1 our example about matches, oxygen and butterflies: we conceded that a human agent S could reasonably ignore information about the butterfly in the other corner of the earth while drawing his inferences about lighting matches. Thus $\delta(T) \vdash x$. And yet, suppose that the wing beat of that butterfly generates the proverbial storm in this side of the world all around our match preventing it from lighting if scratched. Thus $T \nvdash x$. What we want to say is that S is nonetheless entitled to 'jump to his conclusion' since that counterfactual situation is very unlikely, or not *normal*. Thus $T \mathrel{|\!\sim} x$. The trail of "normality" is well trodden in the intuitive presentation of these logics of defeasible reasoning, so let's follow it. What does it mean to reason and draw inferences in normal situations? Quite trivially, it means to reason and draw inferences *as if* from a perspective in which only certain states of affairs are taken into account. That is, for instance, the perspective of *another* scorekeeper S who reasons

[41] Please notice that "default function" here is just an expositive device. I have to admit that it turns out to be quite a luxury permission in the context of nonmonotonic logics for defeasible reasoning, where the notion of "choice function" has been abundantly deployed as a precise technical device. See for instance Rott [137], Lehmann [83], Lehmann et al. [84]. Do not confuse my expositive device with those technical devices.

[42] Makinson [91] proposed to collect all these species under three main genres: default assumptions approach, default valuations approach, default rules approach. In this classification for instance his MAK models belong to the first approach together with Poole systems, Circumscription and KLM models belong to the second approach, Reiter's Default Logic belong to the third approach.

and draws inferences based on his set $C_S \subseteq \mathcal{L}$ of commitments: in this sense C_S represents conditions on normality.

5.5.2.2 A twofold path in formalization

Now, intuitively, there are two main ways to simulate an inference from X to Y in the perspective of a scorekeeper S within *Incompatibility Semantics*: either to expand the premises of the inference as to include those subsets of C_S which are maximally compatible with X, or to restrict the evaluation of the inference to those possible worlds $w_i \in PW$ which are minimally incompatible with C_S. Both ways force the inference to respect C_S, and thus *normality*, as much as possible. Let's make all this more precise.

The first way basically consists in a generalization of Poole [117]'s system without constraints, and in the literature is usually named *default-assumptions approach* after Makinson [91]. The idea is to treat the sentences of a specific default set as background assumptions to be added to the premises in any inference: since the resulting union set of premises might be self-incompatible, each time we maximize the subset of default assumptions that can be added while preserving compatibility. In this sense, given an inference with a set of premises X, for any C' maximally X-compatible subset of C_S, our default function δ will select the union of X and C': thus $\delta(X) = X \cup C'$. Let me parenthetically clear any doubts about the notion of maximally X-compatible subsets of C_S: a set $C' \subseteq C_S$ is maximally X-compatible if $C' \cup X \notin Inc$ and there's no $C'' \subseteq C_S$ s.t. $C' \subset C''$ and $C'' \cup X \notin Inc$. Now it's easy to define a notion of nonmonotonic entailment in IS as follows:

Definition 48. $X \mid\!\sim_{C_S} Y$ iff $\forall C' \subseteq C_S$ maximally X-compatible, $\bigcap_{p \in Y} I(\{p\}) \subseteq I(C' \cup X)$.

In order to illustrate the failure of monotonicity consider the following example.

Example. Let $\mathcal{L} = \{p, q, r, s, x\}$ and

$p \cup r \in Inc$,

$q \cup s \in Inc$,

$x \cup s \in Inc$.

Suppose $C_S = \{p, q\}$.

Then $r \mid\!\sim_{C_S} x$. In fact there is just one maximal $C' \subseteq C_S$ that is compatible with $\{r\}$, thus $C' = \{q\}$. And it's easy to check that $I(\{x\}) \subseteq I(\{q, r\})$.

But $r, s \mid\!\not\sim_{C_S} x$. In fact there is no maximal $C' \subseteq C_S$ that is compatible with $\{r, s\}$, thus $C' = \varnothing$. And it's easy to check that $s \in I(\{x\})$ but $s \notin I(\{r, s\})$.

The second way amounts to an application of *Preferential Semantics*, which has become one of the industrial standards in the logical analysis for defeasible reasonings thanks to Kraus et al. [77]. Now, a particularly effective way to introduce *Preferential Semantics* is to talk about semantic valuation functions: rather than by adding assumptions the selection of default conditions is obtained by restricting the valuations allowed. Thus, given the set V of possible valuations over a language $\mathcal{L}$, if we consider a subset $W \subseteq V$ and an order relation $\prec$ on W, then the poset $(W, \prec)$ is a *preferential model* representing the degree of normality of any given semantic evaluation from the perspective of the agent who establishes the order. Then for any set of premises X we can define $X \mid\!\sim_{\prec} Y$ if and only if Y follows from X in all minimal models for X according to $\prec$. This way to present preferential models highlights the semantic character of the whole operation: we are restricting the models we use in the evaluation of logical consequence. However there's not really anything like valuation functions in *IS*: recall from Section 4.3.1 that a whole articulation of possible worlds correspond to any incoherence frame. In this sense to establish an order on incoherence relations, i.e. $\overrightarrow{Inc}: Inc_1, Inc_2, \ldots, Inc_n$, simply wouldn't do because it is not *Inc*s that specify models but possible worlds. So, what we have to do in order to open this second way for *IS* is to define an order on models. More specifically, given a subset $W \subseteq PW_{Inc}$ and a set of premises X, let $|X|_{W_{Inc}}$ be the set of models for X in W_{Inc}, i.e. $|X|_{W_{Inc}} = \{w \in W_{Inc} \mid w \vDash_{Inc} X\}$, and let $min_{\prec_{W_{Inc}}} |X|_{W_{Inc}}$ the set of minimal models for X in $\prec$. In this sense, that is the set selected by our default function δ, thus $\delta(X) = min_{\prec_{W_{Inc}}} |X|_{W_{Inc}}$. Now it's easy again to define another notion of nonmonotonic incompatibility entailment:

Definition 49. $X \mid\!\sim_{\prec_{W_{Inc}}} Y$ iff $min_{\prec_{W_{Inc}}} |X|_{W_{Inc}} \subseteq |Y|_{W_{Inc}}$

In order to illustrate the failure of monotonicity consider the following example.

Example. Let $\mathcal{L} = \{p, q, x\}$ and $q \cup x \in Inc$. Then there are just two possible worlds in $W \subseteq PW_{Inc}$:

$w_1 = \{p, x\}$;

$w_2 = \{p, q\}$.

Suppose $w_1 \prec w_2$.

Then $p \mid\!\sim_{\prec_{W_{Inc}}} x$. In fact there is just one minimal $w \in PW_{Inc}$ s.t. $w \vDash_{Inc} p$, i.e. $min_{\prec_{W_{Inc}}} |p|_{W_{Inc}} = w_1$, and $x \in w_1$.

But $p, q \nvdash_{\prec_{W_{Inc}}} x$. In fact there is just one minimal $w \in PW_{Inc}$ s.t. $w \vDash_{Inc} \{p, q\}$, i.e. $min_{\prec_{W_{Inc}}} |p, q|_{W_{Inc}} = w_2$, and $x \notin w_2$.

So, we've described two paths that lead to the formalization of a nonmonotonic consequence relation for *Incompatibility Semantics*. But how are they really different? It's interesting to notice that, given some provisos which are particularly reasonable in our framework, these two paths perfectly overlap[43]. This has been actually proved, for finite languages, by Freund [54]. Freund had also provided a representation theorem for these injective models in Freund [53].

I shall rely on this result and take Definition 49 as a sufficiently stable cornerstone to support the nonmonotonic semantics I need to make explicit Sellars's *causal modalities*: I'll allow myself to call it *Preferential Incompatibility Semantics* (*PIS*)[44].

5.5.2.3 Inference from a perspective

I want now to compare these results with Brandom's original approach to irrelevance. Is there any way to account for the differences between the two representations of nonmonotonic entailment we've been dealing with, i.e. relevant reasoning and defeasible reasoning?

If we look at proof theory we find a macroscopic clue: while they both reject *Weakening on the Left*, defeasible consequence relations only accepts a weakened form of *Cut*,

$$\frac{\Gamma \vdash \phi \qquad \Gamma, \phi \vdash \psi}{\Gamma \vdash \psi}$$

Thus once again, as in Section 4.2.5.2, one could be tempted to rush to the conclusions: *Restricted Cut* is the price defeasible reasoning has to pay for maintaining a set theoretical approach that imposes a certain form of weakening,

$$\frac{\phi \vdash \phi}{\Gamma, \phi \vdash \phi}$$

But once again there's more than meets the eye, and the best way to focalize it is still the algebraic point of view. Recall that a *residuated pair* is a couple of functions (f, g) defined over a poset $(S, \leqslant)$ which obey to the law of *Residuation*. In Section 4.2.5.2 we saw how to apply this to groupoids $(S, \leqslant, \circ)$, where we construe $\leqslant$ as consequence and the binary operator $\circ$ as grouping premises: the residuated pair $(\circ, \rightarrow)$ displays the implications

[43] For details see Section 5.7.4 of the Appendix below.
[44] See Section 5.7.3 of the Appendix.

as residuals of conjunctions, thus

$$a \circ b \leqslant c \text{ iff } b \leqslant a \to c$$

This is what the *Deduction Theorem* says. But what does it mean? As the law of *Residuation* highlights for residuate grupoids, it means that there is a perfect match between the semantic representation of inference in terms of consequence relation and the syntactic representation of inference in terms of implication. In algebras for relevance logic this is particularly clear in the 'logic' case, where *null information* is put into play. In fact if we consider a costant e for null information s.t. $e \circ a = a$ we obtain a rule to display inferences in the object language of implication and *viceversa*. Dunn [43] imaginatively calls this the *push-'n-pop* rule:

$$a \leqslant b \text{ iff } e \leqslant a \to b.$$

But what if one puts this push-'n-pop rule into doubt? After all there are some logics which reject the converse of the *Deduction Theorem*, among which nonmonotonic logics for defeasible reasoning are just a significant instance. As we'll see in the Appendix (Section 5.7.3.3 below), these logics have technical reasons to do that, but we can't simply take them for granted here because we are interested precisely in what lies below those technical results. Nonetheless they offer a precious lead for our enquire. Suppose thus, heuristically, to adopt the choice of these logics and drop the *'pop'*-part of the rule. That means to adopt an inegalitarian account of the two ways to represent inferences on the left and on the right of the turnstile: what we take as good reasons on the left we take as good reasons on the right but, possibly, *not viceversa*. As we're going to explain in Section 5.7.5 of the Appendix, we are saying that by breaking the unit of the residuated pair $(\circ, \to)$ two perspectives on inference representation are created which are not reciprocal. This obviously doesn't mean that they could not just match: indeed that is what happens if every implication on the right is the result of making explicit consequences on the left by *pushing* premises on the other side of the turnstile. The real point is that, even if they match, they can't be smoothly coordinated and integrated into a unique representation of inferences. In other words we might have a problem with monotonicity. Consider in fact the following argument:

$$\frac{a \leqslant c}{a \circ b \leqslant c}$$

We would like to conclude just by *Augmentation* but we can't. Why?

Notice first that constant e is a red herring. One could try to argue that it is just the special properties of null information that short-circuit the two perspectives so that the push-'n-pop rule can apply: there's no other point x in the semilattice s.t. $x \circ b = b$. That is the perspective of relevant reasoning: "no surprises, the problem is *Reflexivity*!" But in fact we face another sort of problem: if we drop right-to-left direction of the law of *Residuation* for the pair $(\circ, \rightarrow)$, when we 'pop' premises on the left we can't just group them as the others with $\circ$ and we'd better introduce another operator, say $\star$, to display $\rightarrow$. So the actual logic of push-'n-pop rule turns into something like

$$e \star a \leqslant b \text{ iff } e \leqslant a \rightarrow b$$

Thus even if we had *Augmentation* for both operators for grouping premises, thus $a \star b \leqslant a$ and $a \circ b \leqslant a$, still we couldn't verify *Monotonicity*. What we need are provisos to grant that, at least in some cases, $\star$ and $\circ$ can be coordinated. That is exactly what Gabbay [55] proposed with his rule for *Cautious Monotonicity*:

$$\frac{a \leqslant b \qquad a \leqslant c}{a \circ b \leqslant c}$$

This allows us group premises with $\circ$ when we can already derive them on the right. That is to say that we can drop the difference between $\star$ and $\circ$ for those inferential contents which are already available in perspectival reasoning. Intuitively, what we want is to be sure we can 'pop' something from the right at least when it can be slipped into a cluster of premises which relevantly implies it: that would be enough to annihilate its harmful potential because such a cluster may be safely "$\circ$"-conjuncted.

Crucially, the counterpart of this weakening of *Monotonicity* are provisos on *Cut*. In fact if we can't coordinate inferences made from different perspectives we might very likely have a problem with transitivity too. Consider the following argument:

$$\frac{a \leqslant b \qquad d \circ b \leqslant c}{d \circ a \leqslant c}$$

Here the problem is that if we infer in perspective we can't be sure that a still implies b when "$\circ$"-conjuncted with some other premise. Provisos must be added to grant that b can be cut anyway: *Cumulative Transitivity* obtains that by requiring conclusions to follow from b "$\circ$"-conjuncted with some other premise in the presence of a as well[45], thus

[45]See Section 5.7.5 in the Appendix below for further details.

$$\frac{a \leqslant b \qquad d \circ a \star b \leqslant c}{d \circ a \leqslant c}\ (CT)$$

By way of concluding, notice that this is a quite different way to construe nonmonotonicity compared to relevant reasoning. The purpose of relevant entailment is to allow inferences to be drawn only when all information in the premises is relevant for the conclusion: as a consequence any information added to the premises is to be considered potentially harmful. The purpose of defeasible entailment is to represent inferring from a limited perspective: as a consequence provisos are put forward to single out conditions safe enough to draw conclusions in that situation. Let me just put forward one last remark. There are several ways to construe these perspectival limits, but Chapters 3 and 4 give us a solid hint: limits come from implicitness. According to Brandom, meaningful linguistic practices embed normative implicit contents which are made explicit in terms of material inferences. While implicit, these contents are completely defined by the normative relations entertained by partecipants to the linguistic practices and can be thus singled out by appropriate semantic representations like *IS*. Consequence relations defined inside these semantics represent normative relations. But the process of making these normative relations explicit in appropriate logical languages is part of the very practice of deploying them, because every speaker can't but deploy his own perspective on the process. The practice of *Scorekeeping* is perspectival and objectivity in content's definition the purpose rather than the prerequisite of explicitation.

5.6 The logic of causal modalities

After a long run we took a flying leap in Section 5.4 on the demand for appropriate expressive logical tools. It's time to ask: how far did we get? So let me rehearse our example about scratched matches one last time:

"If this match were scratched it would light."

We've learnt that we can't simply construe it as the 'subjunctive identical'

"If this match were one of the K things then if it is scratched it lights"

and let it be supported by the extensional generalization

"All K things, if they are also scratched things then they are lighting things"

where K ranges over all *possible* particulars.

That is because material inferences, while analytic because purported to make explicit conceptual content, are defeasible. Thus we concluded that, in order to directly express those conceptual contents, we need a directly nonmonotonic entailment. But here we should drop our expressivist demand if we were to be content with Brandom's pragmatic interpretation of Sellars [148]: by endorsing the material inference "if this match were scratched it would light" one doesn't *say* "all scratched matches light" but only *conveys* the information that she is entitled to an inferential pattern that goes from "this match is scratched" to "this match lights". The question seems to be: is Brandom right? If he is, then we might have gone through the whole Section 5.5 pointlessly. Truth is Brandom is right indeed. If I were to say the contrary, I would go against the whole last precious section of Sellars [148]. There Sellars explains how causal modalities testify against the "tendency to assimilate of all discourse to describing"[46] because the inductively established lawlike statements that make them explicit do not represent restricted or unrestricted universally quantified generalizations but rules of inference[47]. In this sense lawlike statements, as rules, go wrong in a pragmatic rather than epistemic way: the fact that one may mistakenly or correctly perform the inference from "x is a match and it is scratched" to "x lights" does not depend on her knowledge of a sufficient number of facts of the form "a is a match and a is scratched and a lights". Rather, the direction of dependency goes the other way around: it is the descriptive vocabulary of particulars and quantification that makes explicit the contents of the inferential rules that specify the normative perspective of a certain scorekeeper. Notice that the output of this expressive process will not contain any residual possibility to be pragmatically mistaken, since such an evaluation can only pertain to the perspective of another scorekeeper. This is why Sellars claims that

> «the logic of variables and quantification involves not only the *momentary* crystallized content of the language at a cross section of its history, but also its character as admitting – indeed demanding – modification, revision, in short, development, in accordance with rational procedures.»[48]

[46] Sellars [148], §103.

[47] This – Sellars hurries to add – doesn't mean that one can't represent rules of inference in terms of universally quantified statements: it just means that one can do that only at the price of an expressive impoverishment. Such an impoverishment is exactly of the sort that according to Macbeth [89] affects the russellian interpretation of Frege's *Begriffsschrift*'s quantification, in which the difference between concavity notation with german letters and latin italic letters disappears. See also Note 18 above.

[48] Sellars [148], §105.

Here's my point then. If I want to develop proper logical tools to deploy causal modalities in my object language, I shouldn't look for a representation of this descriptive framework which is the product of their expression, because from within such a framework there's no way to regain the normative character of material inferences except by a descriptive 'all'-statement. Unfortunately, that is what Brandom's *IS* does. Please notice this is not just a vague complaint. As we've shown in Chapter 3 and here above in Section 5.5, Brandom's *Incompatibility Semantics* has one main adequacy flaw with respect to material inferences: even in its modally expressive extension *EIS*, it is monotonic. So, while Brandom is correct in the analysis of Sellars's account, he doesn't provide us with a suitable representation of the logic of material inferences. And, crucially, there's nothing that rules out the possibility to provide such a representation in Sellars [148]. Notice, to the contrary, that Sellars himself writes down a few hints on how a provisional logical formalization of a causal entailment should look like[49]. We already saw this in Definition 47, but let me repeat it here:

> Let K be an expression for thing-kinds, C be an expression for circumstances, ϕ be an expression for something done to a thing and ψ an expression for a thing's reaction.
>
> Then a causal entailment can be represented as
>
> $$K(x), C(x) \mathrel{|\!\sim} \phi(x) \supset \psi(x)$$

What's crucial for Sellars in this proposal are the provisos specifying the different role of the different premises, for they are the only defense against Goodman's paradox of cotenability. In fact, if we were to confuse the role of conditions C, for instance, with properties ϕ, we could easily slip from

$$K, C \mathrel{|\!\sim} \phi \to \psi$$

to

$$K, C, \phi \mathrel{|\!\sim} \psi,$$

e.g. from

> "If this match is dry, then if it were scratched then it would light"

to

[49] See Sellars [148], §37.

"If this match is dry and it is scratched, then it would light".

As a consequence we would have to accept by *Contraposition*

$K, \neg\psi \mathrel{|\!\sim} \phi \rightarrow \neg C$,

e.g. our problematic

"If this match doesn't light then if it were scratched it would not be dry"[50].

But then, as we've explained in Section 5.5.2.3, this distinction is exactly what nonmonotonic consequence relation $\mathrel{|\!\sim}$, as defined in *PIS*, allows to do. More specifically, it prevents premises occurring in the perspectivally specified context on the right from crossing the turnstile and be grouped on the left among the very information that specify the context. Nonmonotonicity thus springs out from the fact that by modifying premises on the left the context of application of the inference on the right gets changed, and that might defeat the goodness of the inference. As we've just seen, in Brandom's framework this shift of perspective turns out to be an expressive gap: it can be construed either interpersonally as the point of view of a scorekeeper who evaluates another speaker's deontic score, or intrapersonally as the discrepancy between the full explicitation of commitments and the point reached by a speaker in the explicitationg process. I'll try to make all this clear with a couple of examples.

Suppose a speaker S endorses the inferential pattern "If this match is dry, then if it were scratched it would light" as making explicit (part of) the content of the thing-kind word "match". That means that S's deontic status of commitments and entitlements defines a preferential model $(W, \prec)$ such that

$$Match(x), Dry(x) \mathrel{|\!\sim}_{\prec_W} scratched(x) \rightarrow light(x).$$

That is, according to all those maximally compatible sets of commitments that better correspond to S's deontic status in which matches are dry, one is entitled to infer that they light if scratched. But that *doesn't* imply that according to all those maximally compatible sets of commitments that better correspond to S's deontic status in which

[50]Notice however with Sellars [148], §20, that we can obtain $K \mathrel{|\!\sim} \phi \wedge \neg\psi \rightarrow \neg C$, e.g. "If this match was scratched and didn't light then it wasn't dry". Here the switch to the indicative mode is legitimate. Now we can see why: *Contraposition* is applied only to the right. Notice in fact we can also obtain $K, C \mathrel{|\!\sim} \neg\psi \rightarrow \neg\phi$, e.g. "If this match is dry, then if it didn't light then it wasn't scratched". See Section 5.7.3.3 in the Appendix below for a few more words on *Contraposition*.

matches are dry and scratched one is entitled to assert that they light. Let's mimic the jargon of defeasible reasoning's logicians and just call these *maximally compatible sets of commitments which better correspond to S's deontic status* just "normal" (according to S) conditions. Thus, in more plain words, the above means that in *normal conditions for dry matches*, one is entitled to infer that they light if scratched, but it might well be that instead, in *normal conditions for scratched dry matches*, one is simply not entitled to assert that they light. The reason, intuitively, is not that it might be that in normal conditions for dry matches these are not scratched, but that normal conditions for scratched dry matches might force to consider other contents, such as the amount of oxygen in the air, that might rule out matches' lighting[51]. In general the nonmonotonic character of $\vdash_{<_W}$ prevents the unregimented addition of premises on the left, because these might interact in a new context as to block the inference to what is on the right.

This concludes my *showdown*. The value of my expressivist point should be clear enough by now, but I'm not so sure that everyone would concede that I won the pot. In fact, the structure of the argument of Brandom [23] might support a particular complain. So, eventually, one could rise and ask: ok, but Brandom's *IS* represents modality and where's modality in all this? After all my efforts at clarification that would be a discouraging remark, but it's worth spending some more words on this point rather than just referring the questioner back to Section 4.2.5 and Section 5.5. However, since all cards are on the table by now, I'll allow myself to be a little extreme, accepting the risk to be misunderstood. Thus, I claim that modality here is a red herring. The whole point about modality has to

[51] Consider the following.

Example. Let $\mathcal{L}$ = {Match, Dry, Scratched, Light, Oxygen, *not*Light} and

{Match, Dry} $\cup$ {Scratched, *not*Light} $\in$ *Inc*,
{Oxygen} $\cup$ {Light} $\in$ *Inc*,
{Oxygen} $\cup$ {*not*Light} $\in$ *Inc*,
{Light} $\cup$ {*not*Light} $\in$ *Inc*.

We have

w_1 = {Match, Dry, Oxygen},
w_2 = {Match, Dry, *not*Light},
w_3 = {Match, Dry, Scratched, Oxygen},
w_4 = {Match, Dry, Scratched, Light}.

Suppose $w_1 \prec w_2 \prec w_3 \prec w_4$.
Then {Match, Dry} $\vdash_{<_W}$ {Scratched} $\rightarrow$ {Light}, because {Scratched, *not*Light} $\nsubseteq w_1$.
But {Match, Dry, Scratched} $\nvdash_{<_W}$ {Light}, because {Light} $\nsubseteq w_3$.

do with the representation of normativity. And from a formal point of view the correct representation of normativity hinges on the failure of the law of *Residuation* with respect to intensional conjunction and material implication. The reason why we tend not to recognize modality in this feature is that we mainly have learnt to deal with modality directly in the descriptive framework of possible worlds. But this whole argument purports to shake such a presupposition. In other words, when Sellars claims that the language of modality is a transposed language of norms he's proposing an interpretation of modality rather than an interpretation of norms[52].

5.7 Appendix

In this appendix I'll complete the presentation of the nonmonotonic consequence relation for *IS* with the required technical details which have been skipped in the previous part. Since I'll just adapt results which are already well established for nonmonotonic logics, I won't really provide any technically interesting proof. The main result will be a modest one: the soundness of *Preferential Incompatibility Semantics* with respect to the standard set of rules for cumulative consequence relations. I'll also show some rules that fail, the most important of which is *Deduction Theorem*. Then I'll be able to introduce a representation theorem for these consequence relations. Eventually I'll put forward some remarks about the algebraic structure of these systems, which will help characterizing the differences between the 'relevant' and the 'defeasible' approach to nonmonotonicity.

5.7.1 Makinson's dilemma

As correctly noted by Makinson [91], nonmonotonic consequence relation defined through default-assumptions are subject to a dilemma:

(a) If C_S is not closed under classical consequence, the determination of the consequence of A, $\{x \mid A \mid\!\sim_{C_S} x\}$, is syntax-dependent.

(b) If C_S is closed under classical consequence, $\mid\!\sim_{C_S}$ is *the* classical consequence relation.

Let me give some aid to the reader's intuitions. So, on the one side consider this example of Makinson's with relation to point (a):

[52] See Sellars [146], p. 332.

Example 50. Suppose two speakers S and S' have two logically equivalent but syntactically different sets of commitments $C_S = \{p \to (q \wedge r),\ r \to \neg q\}$ and $C_{S'} = \{p \to q,\ p \to r,\ r \to \neg q\}$. It's easy to check that $p \not\vdash_{C_S} q$ while $p \vdash_{C_{S'}} q$.

On the other side, intuitions should be more at ease with point (b): if no information is missing in C_S, then no unconsidered information can prevent the conclusions to follow.

This dilemma is the principal reason for Makinson's discontent with this sort of nonmonotonic consequence relation. In fact it flies in the face of another very basic intuition in logic:

(c) inferences are valid in virtue of their logical forms.

Now, (c) is usually considered so close to the heart of logic that it might seems awkward to treat it as a third horn beside (a) and (b). But in Brandom's framework, as we saw in Chapter 3, the rejection of intuition (c), the so called "formalist dogma", plays an important role in the reevaluation of material inferences. And once the horn (c) of the trilemma is dropped, the incompatibility of (a) and (b) simply becomes an expected rather than a disappointing consequence of the expressivist approach: it is the practice of making explicit contents by means of logical vocabularies that make them available to be put into play in the inferential game of giving and asking for reasons. All in all, syntax-dependence is one of the reasons why this sort of consequence relations are nonmonotonic: if all commitments were explicit, then there would be no perspective at all, but just epistemic uncertainty.

5.7.2 An important simplification

In Section 5.5.2.2 it was claimed that default-assumption approach and preferential approach can be treated as equivalent with some simplifications.

While these simplifications are reasonable, they deal with crucially important formal issues. The first important simplification has to do with one the inferential rules that mostly characterizes the notion of nonmonotonicity in defeasible reasoning, i.e. *Cautious Monotonicity*:

(CM)

$$\frac{\Gamma \mathrel{|\!\sim} \phi \qquad \Gamma \mathrel{|\!\sim} \psi}{\Gamma, \psi \mathrel{|\!\sim} \phi}$$

This is a form of restricted monotonicity that controls what irrelevant information is to be treated as innocuous for the conclusion to follow. It says that informations that can be

already inferred can be safely added to the premises. It is this sort of rule, first introduced by Gabbay [55], that characterize the shift of perspective on nonmonotonic inference from the logic for relevant reasoning to the logic for defeasible reasoning.

Now, contrary to default-assumption consequence relations, preferential consequence relation might fail (CM).

Example 51. (Makinson [90]) Let a preferential model $(W, \prec)$ be s.t. $W = \{w_i \mid i < \omega\}$ and $w_i \prec w_j$ iff $i < j$.

Suppose for some p, q and r, $w_i \vDash p$ for all i, $w_i \vDash q$ for no i and $w_i \vDash r$ for $i = 1$.

Thus $p \mid\sim_{\prec} q$ and $p \mid\sim_{\prec} r$ since $min_{\prec} |p|_W = \varnothing$. But $p, r \not\mid\sim_{\prec} q$ since $min_{\prec} |p, r| = w_1$ and $w_1 \nvDash r$.

In general (CM) fails when the preferential order of models allows bottomless descending chains. For this reason Kraus et al. [77] established the following condition on the preferential relation on models:

Definition 52. Let $\prec$ be a binary relation on a set W. We shall say that $\prec$ is *asymmetric* iff $\forall w_i, w_j \in W$ s. t. $w_i \prec w_j$, we have $w_j \not\prec w_i$.

Let $V \subseteq W$ and $\prec$ a binary relation on V. We shall say that $w_i \in V$ is *minimal* in V iff $\forall w_j \in V$, $w_j \not\prec w_i$.

Let $V \subseteq W$ and $\prec$ a binary relation on V. We shall say that V is *smooth* iff $\forall w_j \in V$, either $\exists w_j$ minimal in V, s. t. $w_j \prec w_i$ or w_i is itself minimal in V.

It is possible to show that a preferential consequence relation verifies (CM) if and only if the set of models is smooth.

But now notice that *Smoothness* is always satisfied by finite sets of models. And since we are working only with finite languages our sets of models will be finite. As we will see, this remarkably simplifies our semantic machinery and let us achieve a crucial result like (CM) with relative ease.

5.7.3 Preferential Incompatibility Semantics

5.7.3.1 Definitions for *PIS*

In the following I'll basically adapt the definition of preferential models from Kraus et al. [77] in order to define a *Preferential Incompatibility Semantics* (*PIS*).

Let me put forward a quick preliminary remark. As it is well known such a definition envisages the possibility of multiple copies of possible worlds to be indexed, or 'labelled',

differently: that possibility is required for the proof of a certain representation theorem. However such a technical device, already counterintuitive, turns out to be definitely unaceptable in inferential semantics: there are just no copies of sets of sentences. Thus I'll just drop this requirement and accept to complicate a bit the representation theorem in next section.

Definition 53. A *preferential frame* PF is a couple $< W_{Inc}, <>$ where W is a set of possible worlds $W \subseteq PW_{Inc}$, and $\prec$ is a binary relation on W_{Inc} satisfying the following *Smoothness condition*: $\forall X \in \mathcal{L}$, the set of models for X, $|X|_{W_{Inc}} \stackrel{def}{=} \{w \in W_{Inc} \mid w \models_{Inc} X\}$, is smooth.

Since we are considering only finite languages we have *Smoothness* for free. So we can directly define the semantical consequence relation:

Definition 54. Given a preferential frame $PF =< W_{Inc}, <>$ and $X, Y \in \mathcal{L}$,
$X \models_{<_{W_{Inc}}} Y$ iff $\forall w_i \in W_{Inc}$ minimal in $|X|_W$, $Y \subseteq w_i$.

In what follows I'll adopt some notational simplifications. First, I'll simply write $X \models_{<_{Inc}} Y$ for $X \models_{<_{W_{Inc}}} Y$. Second, as in Definition 49, I'll write $min_{\prec} |X|_{W_{Inc}}$ for the set of possible worlds which are minimal in $|X|_{W_{Inc}}$. Third, I'll omit the reference to a given incoherence frame Inc when obvious.

5.7.3.2 Soundness

We are going to test *PIS* for soundness against the syntactic characterization of cumulative reasoning fixed as standard in Gabbay [55], Makinson [90]. However we also need a non-horn rule for the proof of representation theorem as we'll see in Section 5.7.4 below.

Theorem 55. *Given a preferential incompatibility frame* $(W, \prec)$*, the following conditions are verified.*

(1)	*(Reflexivity)*	$A \mid\sim_{\prec} A$;
(2)	*(Left Logical Equivalence)*	*if* $\models A \leftrightarrow B$ *and* $A \mid\sim_{\prec} C$ *then* $B \mid\sim_{\prec} C$;
(3)	*(Right Weakening)*	*if* $A \mid\sim_{\prec} B$ *and* $B \models C$ *then* $A \mid\sim_{\prec} C$;
(4)	*(Cautious Monotonicity)*	*if* $A \mid\sim_{\prec} C$ *and* $A \mid\sim_{\prec} B$ *then* $A, B \mid\sim_{\prec} C$;
(5)	*(Cumulative Transitivity)*	*if* $A, B \mid\sim_{\prec} C$ *and* $A \mid\sim_{\prec} B$ *then* $A \mid\sim_{\prec} C$;
(6)	*(AND)*	*if* $A \mid\sim_{\prec} B$ *and* $A \mid\sim_{\prec} C$ *then* $A \mid\sim_{\prec} B \wedge C$;
(7)	*(OR)*	*if* $A \mid\sim_{\prec} C$ *and* $B \mid\sim_{\prec} C$ *then* $A \vee B \mid\sim_{\prec} C$;
(8)	*(Disjunctive Rationality)*	*if* $A \vee B \mid\sim C$ *and* $B \not\mid\sim C$ *then* $A \mid\sim C$.

Proof. (Reflexivity). Suppose $A \not\sim_{\prec} A$. Then $\exists w_i \in min_{\prec} |A|$ s.t. $A \nsubseteq w_i$. Contradiction.

(Left Logical Equivalence). Suppose to the contrary that $\models A \leftrightarrow B$ and $A \sim_{\prec} C$ but $B \not\sim_{\prec} C$. Then $|A|_W = |B|_W$, and so $min_{\prec} |A|_W = min_{\prec} |B|_W$. But then if $min_{\prec} |A|_W \subseteq |C|_W$, then also $min_{\prec} |B|_W \subseteq |C|_W$. Contradiction.

(Right Weakening). Suppose to the contrary that $A \sim_{\prec} B$ and $B \models C$ but $A \not\sim_{\prec} C$. Then $min_{\prec} |A|_W \subseteq |B|_W$ and $|B|_W \subseteq |C|_W$. But then by transitivity of "$\subseteq$" $min_{\prec} |A|_W \subseteq |C|_W$. Contradiction.

(Cautious Monotonicity). Suppose to the contrary that $A \sim_{\prec} C$ and $A \sim_{\prec} B$ but $A, B \not\sim_{\prec} C$. Then $\forall w_i \in min_{\prec} |A|_W$, $B \subseteq w_i$ and $C \subseteq w_i$ but $\exists w_j \in min_{\prec} |A, B|_W$ s.t. $C \nsubseteq w_j$. Now, by definition $X \in min_{\prec} |X|_W$. Thus it must be that $min_{\prec} |A|_W$ is empty while $min_{\prec} |A, B|_W$ is not. As we saw in Example 51 above, this is possible only if W is not smooth. But since $\mathcal{L}$ is finite W is smooth. Contradiction.

(Cumulative Transitivity). Suppose $A \sim_{\prec} B$ and $A, B \sim_{\prec} C$ but $A \not\sim_{\prec} C$. Then $\forall w_i \in min_{\prec} |A|_W$, $B \subseteq w_i$ and $\forall w_j \in min_{\prec} |A, B|_W$, $C \in w_j$, but $\exists w_k \in min_{\prec} |A|_W$ s.t. $C \nsubseteq w_j$. Then $\exists w_k (w_i \in min_{\prec} |A|_W \Rightarrow w_i \in min_{\prec} |A, B|_W)$. But since $\forall w_i \in min_{\prec} |A|_W$, $B \subseteq w_i$ then $\forall w_k (w_i \in min_{\prec} |A|_W \Rightarrow w_i \in min_{\prec} |A, B|_W)$. Contradiction.

(AND). Suppose to the contrary that $A \sim_{\prec} B$ and $A \sim_{\prec} C$ but $A \not\sim_{\prec} B \wedge C$. Then $min_{\prec} |A|_W \subseteq |B, C|_W$. But $|B, C|_W \Leftrightarrow |B \wedge C|_W$. Contradiction.

(OR). Let me first put forward an obvious lemma.

Lemma 56. $|A \vee B|_W = |A|_W \cup |B|_W$

Proof. $|A \vee B|_W = \{w \in W \mid w \models A \vee B\} = \{w \in W \mid w \models A\} \cup \{w \in W \mid w \models B\} = |A|_W \cup |B|_W$ □

Now we prove (OR). Suppose to the contrary that $A \sim_{\prec} C$ and $B \sim_{\prec} C$ but $A \vee B \not\sim_{\prec} C$. Thus $\forall w_i \in min_{\prec} |A|_W$, $w_i \models C$ and $\forall w_j \in min_{\prec} |B|_W$, $w_j \models C$ but $\exists w_k \in min_{\prec} |A \vee B|_W$ s.t. $w_k \not\models C$.

Now, by Lemma 56, $\forall w_k (w_k \in min_{\prec} |A \vee B|_W \Rightarrow w_k \in min_{\prec} |A|_W \cup |B|_W)$. Then $\forall w_k (w_k \in min_{\prec} |A \vee B|_W \Rightarrow w_k \in min_{\prec} |A|_W)$ and $\forall w_k (w_k \in min_{\prec} |A \vee B|_W \Rightarrow w_k \in min_{\prec} |B|_W)$. Contradiction.

(Disjunctive Rationality). Thus $\forall w_i \in min_{\prec} |A \vee B|$, $C \subseteq w_i$ and $\exists w_j \in min_{\prec} |B|$ s.t. $C \nsubseteq w_j$. We want to show that $\forall w_i \in min_{\prec} |A|$, $C \subseteq w_i$.

Now $|A \vee B| = \{w \in W \mid w \models A \vee B\} = \{w \in W \mid w \models A, B\}$. Thus $|A| \cup |B| = |A \vee B|$. But it must be that $\exists w (w \in min_{\prec} |B|$ and $w \notin min_{\prec} |A \vee B|)$, thus $\forall w (w \in min_{\prec} |A \vee B| \Rightarrow w \in min_{\prec} |A|)$. So $\forall w_i \in min_{\prec} |A|$, $C \subseteq w_i$. □

5.7.3.3 Some worth noticing failures

This sort of restricted monotonicity is the result of a delicate balance. It's easy to fall back into monotonicity as soon as we try to strengthen some other rule. Here are some important examples.

The most characterizing one, as we said in Section 5.5.2.3, is that we can't have a full *Deduction Theorem* in *PIS*.

Theorem 57. $X \mid\sim_{\prec} p \to q \nRightarrow X, p \models\mid\sim_{\prec} q$

Proof. For a counterexample, consider $\mathcal{L} = \{a, b, c\}$ and let $b \cup c \in Inc$.

There are two possible worlds $w_1 = \{a, b\}$ and $w_2 = \{a, c\}$. Consider a preferential incompatibility frame $(W, \prec)$ s.t. $W = \{w_1, w_2\}$ and suppose $w_2 \prec w_1$.

Now, $a \mid\sim_{\prec} b \to c$ is valid in $(W, \prec)$. We have to check $\forall w_i \in min_{\prec} |a|$, $w_i \models b \to c$, but since $min_{\prec} |a| = w_2$, it's easy to see that $w_2 \models b \to c$ since $b \notin w_2$ and $c \in w_2$.

But $a, b \mid\sim_{\prec} c$ is not valid in $(W, \prec)$. In fact $min_{\prec} |a, b| = w_1$ and $c \notin w_1$. □

This means that in general defeasible reasoning is not represented by residuated lattices.

Second, while we have an apparently boolean negation, *Contraposition* fails in *PIS*.

Theorem 58. $p \mid\sim_{\prec} q \nRightarrow \neg q \mid\sim_{\prec} \neg p$

Proof. For a counterexample, consider a language $\mathcal{L} = \{a, b, c, d\}$ and let $a \cup b \in Inc$ and $c \cup d \in Inc$.

There are three possible worlds $w_1 = \{a, c\}$, $w_2 = \{b, c\}$ and $w_3 = \{d, a\}$. Consider a preferential incompatibility frame $(W, \prec)$ s.t. $W = \{w_1, w_2, w_3\}$ and suppose $w_1 \prec w_2 \prec w_3$.

Now, $a \mid\sim_{\prec} c$ is valid in $(W, \prec)$. In fact $min_{\prec} |a| = w_1$, and $c \in w_1$.

But $\neg c \mid\sim_{\prec} \neg a$ is not valid in $(W, \prec)$. In fact $min_{\prec} |\neg c| = w_3$ and $a \in w_3$. □

Third, we can't have *full Transitivity* in *PIS*.

Theorem 59. $p \mid\sim_{\prec} q$ *and* $q \mid\sim_{\prec} r \nRightarrow p \mid\sim_{\prec} r$

Proof. For a counterexample, consider $\mathcal{L} = \{a, b, c\}$ and let $a \cup c \in Inc$.

There are two possible worlds $w_1 = \{a, b\}$ and $w_2 = \{b, c\}$. Consider a preferential incompatibility frame $(W, \prec)$ s.t. $W = \{w_1, w_2\}$ and suppose $w_1 \prec w_2$.

Now, $a \mid\sim_{\prec} b$ is valid in $(W, \prec)$. In fact $min_{\prec} |a| = w_1$, and $b \in w_1$.

And $b \mid\sim_{\prec} c$ is valid in $(W, \prec)$. In fact $min_{\prec} |b| = w_2$, and $c \in w_2$.

But $a \mid\sim_{\prec} c$ is not valid in $(W, \prec)$. In fact $c \notin w_1$. □

Notice that, as shown by Kraus et al. [77], any semantics that would verify any of these rules would collapse into monotonicity.

5.7.4 Notes about the representation theorem

Which are the consequence relations represented by *PIS*? Which properties does a consequence relation defined by *PIS* satisfy? As a matter of fact, in the toolbox of nonmonotonic logics there's a readymade representation theorem suitable to answer this question. So we could just pick the following

Theorem 60. *(Freund [53, Theorem 4.13]) Let $\vdash$ be a consequence relation and W_{Inc} the set of possible worlds in the incoherence frame $< \mathcal{L}, Inc >$. Then $\vdash$ is represented by an injective model iff there is a smooth strict order $\prec$ over W_{Inc} such that*

$$X \vdash Y \Leftrightarrow min_{\prec} |X|_W \subseteq |Y|_W$$

Nonetheless, I think it's worth pausing here and show a very general fact that unifies the technical machinery we've deployed to analyze nonmonotonicity in defeasible reasoning. In fact it can be shown that the consequence relation defined according to the so called *default assumptions* approach is equivalent, given certain conditions, to the consequence relation defined according to the *preferential* approach that we used to build our nonmonotonic incompatibility semantics. It is particularly important for us to realize that the whole preferential order, *in our framework*, is nothing but a representation of reasoning from the perspective of another's speaker's commitments. So let's take this little turnaround.

First step. Consider the intersection of any possible world w_i with C_S. The size of such an intersection represents, so to say, the degree of congruity of w_i with the deontic position of the speaker S: the wider the intersection the more accurate the congruity. This naturally suggests the possibility to establish an order of possible worlds to measure this accuracy:

$$w_i \prec_{C_S} w_j \text{ iff } w_i \cap C_S \supset w_j \cap C_S.$$

Second step. Freund [54, Theorem 14] proved that the consequence relation $\vdash_{\prec_{C_S}}$ defined by the model $(PW, \prec_{C_S})$, where $\prec_{C_S}$ is an *injective* function, is such that *for finite languages* $\mathcal{L}$, for every $X, Y \in \mathcal{L}$

$$X \vdash_{C_S} Y \text{ iff } X \vdash_{\prec_{C_S}} Y.$$

Third step. What properties do these consequence relations have? Kraus et al. [77] proved a famous representation theorem for preferential consequence relations satisfying Gabbay-Makinson's conditions for cumulative reasoning, i.e. (1)-(7) in Section 5.7.3.2. This is a very neat result from a logical point of view, but it requires some unintuitive complications in the structure of models. In *Incompatibility Semantics* we have a very neat representation of models we don't want to loose. What we need is a representation theorem for those injective models singled out in the previos step. But Freund [54, Theorem 4.13] proved that preferential consequence relation represented by injective models are preferential consequence relations that verify (1)-(7) and also the following non horn-rule, *Weak Disjunctive Rationality*:

$$(9) \quad \text{(WDS)} \quad C(X \vee Y) = Cn(C(X) \cup C(Y))$$

where $C(A) = \{x \mid A \mathrel{|\!\sim} x\}$ and $Cn(A) = \{x \mid A \models x\}$.

Fourth step. Now, as it's easy to see, (WDS) is a weakened form of the rule of (Disjunctive Rationality) that we proved above in Section 5.7.3.2. Since *PIS* verifies (Disjunctive Rationality) it also verifies (WDS). So, eventually we can apply Freund's representation theorem, cited here above as Theorem 60.

5.7.5 Cumulative transitivity: *Cut* in perspective

In this last section I want to complete, with some technical detail, the characterization of monotonicity in defeasible reasoning provided in Section 5.5.2.3 above.

Let me begin by introducing a generalization in the definition of *residuated pairs*.

Definition 61. Consider two posets $\mathcal{A} = (A, \leq)$ and $\mathcal{B} = (B, \sqsubseteq)$ with maps $f : A \to B$ and $g : B \to A$. The pair (f, g) is called *residuated* if for all $a \in A$ and $b \in B$

$$f(a) \sqsubseteq b \text{ iff } a \leq g(b).$$

As shown in Dunn [42], this generalization highlights the relations between residuated pairs and *Galois Connections*. I'll try not to abuse terminology and, with relation to the above definition, I'll refer to order preserving – monotonic – maps as *residuated pairs* and to order reversing – i.e. antitonic – maps as *Galois Connections*. Notice that in this sense a *residuated pair* makes explicit correlative maps between two posets $\mathcal{A}$ and $\mathcal{B}$. If we take $\mathcal{A} = \mathcal{B}$ and $(\circ, \to)$ to be a residuated pair, we can define the residuated tonoid $\mathcal{M} = (S, \leqslant, \circ, \to)$, which is a simplified instance of Dunn [43]'s *partial gaggles*. As Dunn

showed these structures are particularly suitable to provide algebraic representations of substructural logics, since many structural rules can be obtained just by proper tuning of the properties of $\circ$.

For what concerns us here, just notice that the binary operator $\circ$ is *isotonic* both on the left and on the right, thus

Isotonicity $a \leqslant b \Rightarrow c \circ a \leqslant c \circ b$
$a \leqslant b \Rightarrow a \circ c \leqslant b \circ c$

Obviously there's no need to distinguish left and right isotonicity if $\circ$ is required to be commutative. We are interested in isotonicity *because*, as it's easy to see, it is the property on which *Cut* rule depends:

Cut

$$\frac{\Gamma \vdash \phi \qquad \Delta, \phi \vdash \psi}{\Gamma, \Delta \vdash \psi}$$

So far so good. But now notice what happens if we break the correlation of the residuated pair. Suppose thus we only accept the left-to-right direction:

$$a \circ b \sqsubseteq c \Rightarrow b \preceq a \rightarrow c.$$

Notice here it's crucial to drop also the identification of the two posets. Other things being equal, what we are left with is a unique carrier set with two orders and a (so to say) semi-residuated pair. As a consequence we have to distinguish between two ways to apply isotonicity: now $\circ$ is isotonic with w.r.t. $\sqsubseteq$ but it is not isotonic w.r.t. $\preceq$. And this obviously is enough to falsify *Cut* on $\preceq$[53]:

$$\frac{a \preceq b \qquad d \circ b \preceq c}{d \circ a \preceq c}$$

It's hard to maintain the possibility to 'infer in perspective' in the absence of something like a *Cut* rule to connect the results of different inferences in a unique argument. What to do then? We'll get to the solution in three steps.

First. As it's easy to predict, this weird halved algebraic structure we've been considering for the sake of the argument can be completed and turned into something more regular. In fact Dunn [43] proved the following

[53]Notice however that single-premise *Cut* still goes through, since $\preceq$ is transitive: this issue is all about the operator for grouping premises "$\circ$".

Theorem 62. *(Embedding Theorem for Implication Posets) Every implicational poset* $(S, \leqslant, \rightarrow)$ *can be embedded in a right residuated p.o. groupoid* $(S', \leqslant, \circ, \rightarrow)$.

So, let $(S, \preceq, \star, \rightarrow)$ embed our implicational poset. Obviously, in this way, it is possible to reestablish a residuated pair $(\star, \rightarrow)$ s.t.

$$a \star b \preceq c \text{ iff } b \preceq a \rightarrow c.$$

But it is as much easy to notice that $\star$ might be completely different from $\circ$. Suppose however they both behave like standard meets with respect to the two different orders:

$$a \preceq b \text{ iff } a \star b = a \qquad \text{and} \qquad a \sqsubseteq b \text{ iff } a \circ b = a.$$

Second. Notice that the only direction of *Residuation* we've decided to accept guarantees that

$$a \circ b \sqsubseteq c \Rightarrow a \star b \preceq c.$$

Thus, in general we have $a \sqsubseteq b \Rightarrow a \preceq b$.

And yet, what we are interested in is a way to express an inference $a \preceq b$ in terms of $\sqsubseteq$ in order to exploit isotonicity of $\circ$ to verify *Cut*. But if we accept the above definitions of the two orders in terms of $\star$ and $\circ$, this is indeed trivial algebraic manipulation:

$$\begin{aligned} a \preceq b \quad &\Leftrightarrow a \star b = a \\ &\Leftrightarrow a \circ a \star b = a \circ a \\ &\Leftrightarrow a \circ a \star b = a \\ &\Leftrightarrow a \sqsubseteq a \star b \end{aligned}$$

Third. We have now at our disposal a way to apply isotonicity of $\circ$:

$$a \preceq b \Rightarrow c \circ a \sqsubseteq c \circ a \star b$$

This is enough to verify the following rule for a restricted *Cut*:

$$\frac{a \preceq b \qquad d \circ a \star b \preceq c}{d \circ a \preceq c}$$

If we now drop all the encumbrances of this admittedly unhappy algebraic notation it's easy to see such a rule is nothing but *Cumulative Transitivity*:

$$\frac{\Gamma \mathrel{|\!\sim} \phi \qquad \Gamma, \phi \mathrel{|\!\sim} \psi}{\Gamma \mathrel{|\!\sim} \psi}$$

Notice that here "$\circ$" represents nonmonotonic addition of premises which is conceptually different from the "," connecting Γ and ψ in the second premis: in fact "," can verify *Augmentation*, $a \preceq b \Rightarrow a, c \preceq b$, and still $\mid\!\sim$ be non monotonic. In this sense $\mid\!\sim$ as a consequence relation represents the inferences as drawn in a *limited* perspective, where limits are given by the unary direction of *Residuation*. Unfortunately, this remark of mine is very weak, since I'm not providing anything like a representation theorem to support it.

Let me conclude with some important remarks about *negation*. Here more than a few words were worth being spent, but I'll be content with suggesting to the reader at least the main point. So, Dunn [45] showed that, for the analysis of substructural logics, negation can be successfully treated in algebraic semantics as a *Galois Connection*. This can be done by introducing operators for negation as in Birkhoff [7]: given two posets $(A, \leq)$ and $(B, \sqsubseteq)$, and given two functions $\sim$: $A \to B$ and $\neg : B \to A$, the pair $(\sim, \neg)$ is a *Galois Connection* if and only if

$$a \leq b \text{ iff } \sim b \sqsubseteq \sim a \qquad \text{and} \qquad a \sqsubseteq b \text{ iff } \neg b \leq \neg a.$$

Now, the crucial point for us to notice is that if we want to keep the distinction between the two posets we can't just drop one direction of the law of *Residuation* between $\circ$ and $\to$ and allow at the same time a *Galois Connection* $(\neg, \sim)$: that in fact would be sufficient to reestablish the equivalence when negations are combined with the other operators. As a consequence, in order to keep the distinction in a system with negation we have also to break the *Galois Connection* $(\sim, \neg)$. Notice this will be obviously enough to falsify *Contraposition*. However, as we did before with $\star$, it is possible to introduce another operator in order to establish another *Galois Connection* $(-, \neg)$. Than will guarantee all the other properties of negation while 'inferring in perspective'.

Chapter 6

Conclusions

6.1 What's left to say

It is custumary to dedicate an essay's concluding part to wide recapitulations. Hence, I should now turn back to contemplate all the path on which I brought along the reader, help her to collect herself and perhaps try to convince her that the journey was worthy of the effort. That I take to be a hard task indeed. So, rather than wrapping up my theses and trying to sell them, in this last Chapter I want to put forward a few more general remarks about the extent of this work. In this sense though, I can't evade from providing a short summary after all. In Section 6.2 I'll draw a map of the path we've taken through normativity. But then, instead of rhetorically asking the initial question – "Is meaning normative?" – and gloriously concluding by stating the answer I argued for – "Yes, it is!" –, I start over with another question: is this all there is to say about normativity of meaning? All things considered I can't but face the scruple not to suggest wrong answers to this question. This motivates the last two sections. In Section 6.3 I'll tackle the temptation to construe the bounds of the brandomian account I've defended as a *reductio* of the normative approach *tout court*, and I'll try to provide a proper diagnosis of the demands that exceed those bounds. Eventually, in Section 6.4, I'll suggest a desirable therapy against them.

6.2 A path through normativity

In Chapter 1 I described the structure of an analysis of the normativity of meaning. Basically, all the rest of this work has consisted in putting some flesh on those bare bones.

Let's take a look at the finished product.

6.2.1 The path

We began, as anyone, with puzzles about the 'meaning relation'

"..." means ⋯.

What are the relatas? How do they relate? How is the relation established? Here we met a traditional pattern of answers according to which "means" relates linguistic expressions with things in the world, by describing the representational purport of language, which established by certain facts about language's use. Here we also met the main problem of this solution: no collection of facts suitably justifies any such relation.

There we pinned down our first benchmark: normativity of meaning doesn't *primarily* consist in prescriptions for speakers to apply linguistic expressions to things in the world in given situations. Thus, in order to explain what else normativity of meaning is, with Chapter 3 we plugged in Robert Brandom's account of conceptual content. That let us treasure a crucial sellarsian inheritance: "means" doesn't *primarily* relate linguistic items with things in the world. And this led to quite a huge subversion of the traditional view in semantics. We learnt from Brandom our way through a semantics which provides an interpretation of linguistic expressions in terms of their inferential role. We were told that these inferential roles make explicit the deontic statuses that speaker acquire as a consequence of their engaging in rational linguistic practices. While performing assertions, speakers acquire commitments and entitlements to further performances. But certain commitments can rule out certain entitlements. Thus by keeping the 'score' of these deontic statuses, speakers can reciprocally evaluate the correctness of their deployment of conceptual contents.

In spite of my efforts to smooth out the difficulties along the way, this was probabily the first significant effort I imposed to the reader. But I hope it paid the reader back as soon as she had to confront with the question: how can these norms be made explicit then? In fact, from the point of view just reached it was easy to see not only that normative contents must be made explicit through modally robust conditionals, but also, consequently, that this deployment of modal vocabulary is a non dispensable feature of any representation of conceptual content. This is what we labelled "the first blindspot" of traditional semantics: to the contrary of what empiricists like Hume and Quine thought, the possibility to represent content of descriptive obsevational vocabulary is parasitic on the possibility to

deploy modal vocabulary. This claim corresponds to the Kant-Sellars thesis in Brandom [23]: modal vocabulary is VV-Necessary with respect to empirical vocabulary.

Boosted by this explanatory success, in Chapter 4, we continued to follow Brandom up to the definition of his *Incompatibility Semantics*. As we saw, that is a semantic system directly based on a set theoretical representation of the very normative notion of incompatibility between contents. From the foothills of the mountain we were approaching, this appeared as a workable alternative to traditional semantics, suitable to do justice of those normative features of meaning we had just highlighted. However, the climbing soon became harder than we thought. In its original formulation Brandom's *IS* presented serious obstacles in the formalization of modality, which we could overcome only with the help of Göcke et al. [59], Peregrin [115] and an extended version of our semantics, *EIS*. That is how we could reach the end of the path opened by Brandom's *Incompatibility Semantics*: a normative representation of Kripke's relational possible worlds semantics.

There we could pin down another benchmark and meet the experience of Lance and O'Leary-Hawthorne [82]. They had follow a different path. Rather than stepping through modality, they directly questioned the notion of analyticity from a normative point of view and asked: what does it mean to *treat* a steatment as analytic? To this question they gave a brandomian answer which led them very close to us. However we noticed that, when confronted with Kripke's modal analysis of analyticity, their normative notion of *de jure* analyticity could hardly be distinguished from epistemic uncertainty. The bad news was that we couldn't do better. This is what we labelled "the second blindspot": since conceptual content, *qua* normative, is revisable, it shouldn't be represented in the static terms of our basic extensional formal tools. As we said in Chapter 5, we can't just put a normative lever under analyticity, we have to lift our representation of modality as well.

Then, Brandom himself pointed at a further goal: nonmonotonic material inferences. Henceforward the path would have been really steep, so it was worth asking whether this second blindspot was not just mine after all. Thus in Chapter 2 two-dimensional semantics were evaluated and their extensional resources were tested against these representational demands. But the results were unsatisfying: the multiplication of dimensions for the evaluation of intensional contents hardly patches the problems of undertermination at the price of dangerously blurring intuitions about language.

Thus eventually in Chapter 5 we had to undertake the last bit of our journey. We did that by following the precious directions of Wilfrid Sellars who walked this trail in Sellars [148] while confronting with Goodman's analysis of counterfactuals and induction. But

that happened in 1957 and very few people had trodden it since then, so that the path was overgrown and difficult to follow (and doubts could be raised whether we followed it properly). In was just by deploying the powerful tools recently developed to handle nonmonotonic inferences that we could walk our way to our last result in this work: *Preferential Incompatibility Semantics*. That let us develop Brandom's semantic treatment of the normative notion of incompatibility into a formal system suitable to represent material inferences in terms of a nonmonotonic consequence relation.

6.2.2 The boundaries

This concludes my recollection of the path through normativity that led us to the end of this work. But, as I anticipated, it is exactly here that I want to ask a troubling question: is this all there is to say about normativity of meaning? Or, at least, is this all there is to say about normativity along this brandomian path?

In order to understand, first, why it is important to raise this question now, recall how we began our analysis of Sellars's criticism to Goodman: we stepped back from the question about the logic of projection and focused (with Sellars) on the preliminary question about the content of counterfactual conditionals. Now that we've developed an answer to this latter question, it's time to see what we might have missed by neglecting the former. It's easy and frightening to realize that our results apparently can't still solve Goodman's "new riddle of induction". We can, so to say, represent projectibile contents, but we can't tell which have to be projected. We know what it means for a speaker to be committed and entitled to "emeralds are green" and we know what it means to be committed and entitled to "emeralds are grue". We know why one could be entitled to both although they are incompatible. We know what is incompatible with "emeralds are green" and what is incompatible with "emeralds are grue". And we know whether a speaker *has to* be committed either to "emeralds are green" or to "emeralds are grue". Eventually, we know how to represent all that in formal semantics. But, it figures, we can't tell whether emeralds *are* green or grue?

6.3 A few last words on representing and quantification

Here is where a semanticist would be willing to hang a pretty standard argument for the rejection of the normative approach to meaning:

> So, at last we have to admit that there is something about meanings which the normative approach can't account for. And it's not a secondary detail. It's just the problem we started with: meanings are underdetermined with respect to our best description of them. So here's the recurring process we're embarking into at least since Kripkenstein: we have a problem with the description of meanings, we plug in normativity, we still have the same problem. Why don't we just drop this idea?

My rejoinder to this objection is 'simply' that it misconstrues the whole point. And this can be shown by pointing out a mistake in its very formulation. The question about the logic of projection is not a question about the *correct* description of meanings. It can't be because there is no ground for one to stand and ask such a question. Below the nonmonotonic perspectival expression of material inferences there's nothing to appeal to in order to justify the correctness of a perspective rather than another. There's no magic trick to turn "$\mathrel{|\!\sim}$" into "$\models$", there's only the wittgensteinian bedrock. If what is asked is whether there is a logic of what goes on below the bedrock, then the question just doesn't make any sense. McDowell bravely stands there on the edge of the abyss of the *Myth of the Given* warning us to step back: below the bedrock there's no justification, hence there's no conceptual content.

This is something I already said here and there along the way. But I have to say something more for those who may still feel this need to jumble the table of normativity in order to preserve the correctness of semantic representation. Let me precipitate the whole point to its basics. It is important to realize that all what I've said doesn't subvert an intuitive realistic thesis according to which words picture things in the world. This I take to be an obvious common ground. However, since I'm afraid of the abyss, I not only reject the 'Fido'-Fido theory, but also any giveness of the intentionality or aboutness of contents. Here I follow Brandom in construing this intentionality as realized by practices of resolving incompatibilities in the inferential articulation of linguistic patterns. Is all this incompatible with realism? Now, the fact that a semantic metavocabulary establishes certain relations between words and linguistic patterns is trivial. It's also legitimate to point at these relations as *picturing* relations and claim that words represent facts through linguistic patterns, i.e. uniform correlations between 'languagings' and things. These correlations should guarantee the grounding of the representation. But – one might ask – the representation of *what*? These scruples are legitimate and they can be satisfied: we get to the answer in two steps.

The first step consists in understanding what it means to provide a representation of objects in terms of semantical particulars. Just recall two points we've already explained. First, subsentential contents can be defined in an inferential semantics through the substitutional strategy described in Section 3.4.1[1]. Second, intentional practices are "thick" – as Brandom puts it – in the sense that they can't be semantically specified without involving the representation of *what* they are about.

The second step is much more important to me. Indeed it would require a much more technically detailed treatment than what I can put forward here, but I'm confident that my account will be explicative enough for the present argument. This step is triggered by a residual ontological demand: how can we seriously talk about representations, if we don't have at our disposal a domain of objects to quantify over and turn our subsentential commitments into ontological cash value? The short reply is to reject the conditional itself: why should we need a domain of particulars as *given* in order to represent *objects*? This might cope with the objection but obviously it wouldn't satisfy the demand. Thus a long answer is needed, and it comes in two parts. The first part consists in recognizing that the standard quantification of *Principles of Mathematics* is a logical device explicitly designed to express the idea that an inference is good if it is valid for *all cases*[2]. This is a legitimate choice, which however involves (as probably anything in logic) certain assumptions on the sort of reasoning which has to be formalized: for what concers us here, the main assumption is that the inferential articulation of conceptual contents has to be completely, permanently deployed. Notice that the domain of obects, by itself, is neutral with respect to any ontological determination. But it is obvious (at least since Quine has noticed it) that, in such assumed conditions, it would be possible to reverse engineer through quantification the ontology represented by conceptual contents. This should provide enough reasons to reject the conditional: we can't use quantification as a *criterium* for the acceptability of different representations of objects, because it can only work by taking for granted any of those representations. Then, the second part of the answer must provide other reasons to take seriously these representations of objects. Thus the question is: granted that subsentential commitments equip inferential semantics with a representational purport, how can an objective dimension emerge from the perspectival character of *Scorekeeping*? In order to answer this question we just have to look at how we ascribe commitments to

[1]Notice this is not a thesis one can just refuse to buy. Theorem 45 establishes that the substitutional strategy works in the non-compositional framework of *Incompatibility Semantics*. In order to reject this thesis then one has to show there's something wrong in the proof of such theorem.

[2]Let me refer once again to Chapters 2 and 3 of Macbeth [89].

other speakers. As we already saw in Chapter 2 there are basically two ways to do that: either *de dicto*, e.g.

Pierre believes *that* London is pretty

or *de re*, e.g.

Pierre believes *of* London *that* it is pretty.

Obviously the grammatical difference between these two examples is merely a linguistic regimentation, but it makes explicit a deep feature of normative linguistic practices: the ability of scorekeepers to *navigate across perspectives*. Thus while in *de dicto* ascription the ascriber endorses the SMSICs of the speaker she's attributing the propositional commitment, in a *de re* ascription she distinguishes between her substitutional commitments and those of the other speaker. And once two different SMSICs are identified, e.g. types $< London >_1$ and $< London >_2$, anaphoric commitments allow to track any occurrence of subsentential expressions, e.g. $/London/_i$, to its correct type. Since anaphoric commitments do not suffer from perspectival evaluation the tracking procedure is objective[3]. This is how the representational content is defined.[4].

6.4 Idealism

In the previous section we've seen what we shouldn't say about our residual questions. But what should we say then? Let's go back to McDowell: there's no conceptual content below the bedrock. Brandom agrees, but crucially adds: there's no *explicit* conceptual content below the bedrock. And that is enough to switch on the whole process of semantics we have described here.

Notice, however, that everything in this work and in particular the whole *Preferential Incompatibility Semantics* is placed *above the bedrock*. Notice also that the logical tools we've developed provide us with a surprisingly neat way to specify that: we just need to focus on the role played by *incompatibility frames*. In *Incompatibility Semantics* the expressions of a given language $\mathcal{L}$ are interpreted in terms of incompatibility sets, which are

[3]This is one of the most precious treasures one can find in Brandom [14], and I'm aware and truly ashamed of the fact that this was just an extremely rough presentation of it. Thus I strongly recommend to read it in its original source in Brandom [14], Chapters 7-8. In Turbanti [167] I tried to display some of its explanatory power in linguistics.

[4]For those who still feel like their cognitive possibilities are disabled by the lack of a quantificational machinery, I suggest to look for one in Sellars [162].

determined according to a given *incompatibility frame*, i.e. according to an *incompatibility relation* Inc over $\mathcal{L}$. But now recall what we said in Section 4.3.1: possible worlds are defined only *inside* an *incompatibility frame*. Thus *incompatibility frames* do not really correspond to the models of standard modal semantics, rather they really contain complete articulations of models[5]. What's crucial to notice then is that the whole preferential structure that triggers the definition of nonmonotonic consequence relation lies within a single *incompatibility frame*. In other words, the frame that ultimately establishes semantic contents is not revisable. In this sense, if we accept to construe nonmonotonicity as due to implicitness, then we should identify Inc with the implicit definition of conceptual contents that linguistic practices make explicit through logical vocabulary. Hence, to ask whether Inc is the correct representation of contents doesn't make any sense, neither semantically nor epistemically: it just is the representation of contents.

But then, what can we say about Inc? What should we reply to those Goodmans who ask what is the *real* logic of the *Scorekeeping*? Let's begin with saying what we know, i.e. that this question may have two senses. In a sense it asks what is the logic of the deontic score that represents the content of normative linguistic practices. In another sense it asks what is the logic of the process of keeping the score. In the first sense the question is answerd by *Preferential Incompatibility Semantics*. In the second sense instead, it echoes the question which haunts every single page of Chapter 9 of *Making it Explicit*: where do the norms come from?

As it's well known, Brandom has an hegelian answer to such a question: look at the *history* of "the process by which the commitments undertaken by members of a discursive recognitive community – and with them the concepts that articulate and constrain what counts as successfully intergrating them – change and develop over time"[6]. I won't try to elaborate this now[7]. What I want to focus on briefly are the complaints raised by this sort of answer, which, notoriously, culminated into the charge of *objective idealism*[8]. Here, obviously, I'm concerned with the extent of my own work, rather than with defending Brandom who can take care of himself and really doesn't need my support. So, is objective idealism the only outcome of this approach to normativity? Let me try, with the help of

[5] In this sense, perhaps with a little twisting, *incompatibility frames* could be better compared to those families of possible worlds Sellars talked about in Sellars [143].

[6] Brandom [31], §2.

[7] This might seem an unacceptably rough way to shrug the problem off, but the truth is that for years Brandom himself has collected and chosen with great care his words, so I just won't venture to hastily throw down here any of mine. I just refer, as always, to Brandom [25] and also to Brandom [17, 18, 29, 30, 31].

[8] See Habermas [67].

Sellars, to shed some light on this problem from another perspective:

> "That languagings are *evoked* (in contexts) by happenings of certain kinds is a *causal* fact which is nevertheless essential to their conceptual character. This causal aspect of perceptual takings, introspective awarnesses, inferences, and volitions accounts for the selecting of *one* world story *rather than another* and connects the 'is' of this selecting with the rule-governed or 'ought to be' character of language."[9]

Languaging, Sellars recalls, belong both to the causal order and to the order of reasons. This "janus-faced" nature of languagings can't be swept away, not even by idealism. Thus, an analysis of language as a natural phenomenon produced by human animals who deploy naturally selected abilities is the required complement to Brandom's hegelian path. That is the sort of work Ruth Garreth Millikan, for instance, as a philosopher, has burdened herself with[10]. Let me just notice what's crucial for the proper interpretation of the doubt that motivated this last chapter: in both cases, in Brandom's historical path and in Millikan's naturalistic path, what should be looked for are not justifications, but, so to say, *narrations*. But I can't provide any such narration now: I'm not able to, yet. With a huge dose of self-importance, I could say that this work, as *Making it Explicit*, is just the first half of a book on normativity, but, to the contrary of Brandom, I haven't written the second part yet.

[9] Sellars [162, After Meanings], § 63.

[10] As I did for Brandom above, so I won't try to squeeze Millikan's theory either in these few words of my conclusions. Let me just pin down, with no claim of completeness, some benchmarks: Millikan [105, 106, 107, 108].

Bibliography

[1] Carlos E. Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50(2):510–530, 1985.

[2] Alan Ross Anderson and Nuel Belnap. *Entailment: The Logic of Relevance and Necessity*, volume I. Princeton University Press, Princeton, 1975.

[3] Mira Ariel. Interpreting anaphoric expressions: a cognitive versus a pragmatic approach. *Journal of Linguistics*, 30(1):3–42, 1994.

[4] Nuel Belnap. Declaratives are not enough. *Philosophical Studies*, 59(1):1–30, 1989.

[5] Nuel Belnap and Anil Gupta. *The Revision Theory of Truth*. MIT Press, Cambridge, MA, 1993.

[6] Nuel D. Belnap. Tonk, plonk and plink. *Analysis*, 22(6):130–134, 1962.

[7] Garrett Birkhoff. *Lattice theory*. American Mathematical Society, Providence, 1940/1948/1967.

[8] Ned Block. Advertisement for a semantics for psychology. *Midwest Studies in Philosophy*, 10:615–78, 1986.

[9] Paul Boghossian. The rule-following considerations. *Mind*, 98(October):507–49, 1989.

[10] Paul Boghossian. Does an inferential role semantics rest upon a mistake? *Mind and Language*, 8(1):27–40, 1993.

[11] Paul Boghossian. Is meaning normative? In Christian Nimtz and Ansgar Beckermann, editors, *Philosophy - Science - Scientific Philosophy. Main Lectures and Colloquia of GAP.5, Fifth International Congress of the Society for Analytical Philosophy*, pages 205–218. Mentis, 2005.

[12] Robert Brandom. Asserting. *Noûs*, 17(4):637–650, 1983.

[13] Robert Brandom. Inference, expression, and induction. *Philosophical Studies*, 54(2): 257–285, 1988.

[14] Robert Brandom. *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Harvard University Press, 1994.

[15] Robert Brandom. *Articulating Reasons: An Introduction to Inferentialism*. Harvard University Press, 2000.

[16] Robert Brandom. Vocabularies of pragmatism: Synthesizing naturalism and historicism. In Robert Brandom, editor, *Rorty and His Critics*. Blackwell Publishers, 2000.

[17] Robert Brandom. Holism and idealism in hegel's *Phenomenology*. In *Tales of the Mighty Dead: Historical Essays in the Metaphysics of Intentionality*, pages 178–209. Harvard University Press, 2002.

[18] Robert Brandom. Some pragmatist themes in hegel's idealism. In *Tales of the Mighty Dead: Historical Essays in the Metaphysics of Intentionality*, pages 210–234. Harvard University Press, 2002.

[19] Robert Brandom. The centrality of sellars' two-ply account of observation to the arguments of "empiricism and the philosophy of mind". In *Tales of the Mighty Dead: Historical Essays in the Metaphysics of Intentionality*, pages 348–368. Harvard University Press, 2002.

[20] Robert Brandom. Responses to pippin, macbeth and haugeland. *European Journal of Philosophy*, 13(3):429–441, 2005.

[21] Robert Brandom. Sketch of a program for a critical reading of hegel. comparing empirical and logical concepts. *Internationales Jahrbuch des Deutschen Idealismus*, 3:131–161, 2005.

[22] Robert Brandom. Inferentialism and some of its challenges. *Philosophy and Phenomenological Research*, 74(3):651–676, 2007.

[23] Robert Brandom. *Between Saying and Doing: Towards an Analytic Pragmatism*. Oxford University Press, 2008.

[24] Robert Brandom. Replies. In Bernd Prien and David P. Schweikard, editors, *Robert Brandom. Analytic Pragmatist*. Ontos, 2008.

[25] Robert Brandom. A Spirit of Trust. 2008.

[26] Robert Brandom. How analytic philosophy has failed cognitive science. In *Reason in Philosophy: Animating Ideas*. Harvard University Press, 2009.

[27] Robert Brandom. Pragmatism, inferentialism, and modality in sellars's arguments against empiricism. In Willem A. DeVries, editor, *Empiricism, Perceptual Knowledge, Normativity, and Realism: Essays on Wilfrid Sellars*, pages 33–62. Oxford University Press, 2009.

[28] Robert Brandom. *Reason in Philosophy: Animating Ideas*. Harvard University Press, 2009.

[29] Robert Brandom. Norms, selves and concepts. In *Reason in Philosophy: Animating Ideas*. Harvard University Press, 2009.

[30] Robert Brandom. Autonomy, community and freedom. In *Reason in Philosophy: Animating Ideas*. Harvard University Press, 2009.

[31] Robert Brandom. History, reason and reality. In *Reason in Philosophy: Animating Ideas*. Harvard University Press, 2009.

[32] Rudolf Carnap. *Logische Syntax der Sprache*. Springer, Wien, 1934. En. transl. by A. Smeaton, The Logical Syntax of Language, London: Kegan, Paul, Trench Teubner & Cie, 1937. Reprinted in Paterson, NJ: Littelefield, Adams & Co., 1959.

[33] Lewis Carroll. What the tortoise said to achilles. *Mind*, 14:278–280, April 1895.

[34] David Chalmers. Does conceivability entail possibility? In T. Gendler and J. Hawthorne, editors, *Conceivability and Possibility*, pages 145–200. Oxford University Press, 2002.

[35] David Chalmers. Foundations of two dimensional semantics. In M. Garcia-Carpintero and J. Macia, editors, *Two Dimensional Semantics*, pages 55–140. Oxford University Press, 2006.

[36] David Chalmers and Frank Jackson. Conceptual analysis and reductive explanation. *Philosophical Review*, (110):315–361, 2001.

[37] Noam Chomsky. *Aspects of the Theory of Syntax*. MIT Press, 1965.

[38] Donald Davidson. Truth and meaning. *Synthese*, 17(1):304–23, 1967.

[39] Daniel Dennett. Intentional systems. *Journal of Philosophy*, 68(February):87–106, 1971.

[40] Daniel Dennett. The evolution of "why?". In Bernhard Weiss and Jeremy Wanderer, editors, *Reading Brandom: Making it Explicit*. Routledge, 2010.

[41] Michael Dummett. *Frege: Philosophy of Language*. Duckworth, 1973.

[42] Michael J. Dunn. Gaggle theory: An abstraction of galois connections and residuation with applications to negation and various logical operations. *Logics in AI, Proceedings European Workshop JELIA 1990, Lecture notes in Computer Science.*, 476, 1991.

[43] Michael J. Dunn. Partial gaggles applied to logics with restricted structural rules. In Peter Schroeder-Heister and Kosta Došen, editors, *Substructural Logics*, pages 63–108. Claredon Press, Oxford, 1993.

[44] Michael J. Dunn. Gaggle theory applied to modal, intuitionistic, and relevance logics. In I. Max and W. Stelzner, editors, *Logik und Mathematik: Frege-Kolloquium Jena*, pages 335–368. de Gruyter, 1993.

[45] Michael J. Dunn. Star and perp: Two treatments of negation. In James E. Tomberlin, editor, *Philosophical Perspectives*, volume 7, Language and Logic, pages 331–357. Ridgeview Publishing Company, Atascadero, California, 1993.

[46] Michael J. Dunn and Gary Hardegree. *Algebraic Methods in Philosophical Logic*, volume 41 of *Oxford Logic Guides*. Clarendon Press, Oxford, August 2001.

[47] Jerry Fodor and Ernest Lepore. *Holism: A Shopper's Guide*. Blackwell, 1992.

[48] Jerry Fodor and Ernest Lepore. Brandom's burdens: Compositionality and inferentialism. *Philosophy and Phenomenological Research*, 63(2):465–481, 2001.

[49] Jerry Fodor and Ernest Lepore. Brandom beleaguered. *Philosophy and Phenomenological Research*, 74(3):677–691, 2007. Reprinted in Weiss and Wanderer [170].

[50] Gottlob Frege. *Begriffsschrift: Eine Der Arithmetische Nachgebildete Formelsprache des Reinen Denkens*. L. Nebert, 1879.

[51] Gottlob Frege. *Posthumous Writings*. Basil Blackwell, Oxford, 1979.

[52] Gottlob Frege. *The Foundations of Arithmetic: A Logico-Mathematical Enquiry Into the Concept of Number*. Northwestern University Press, 1980.

[53] Michael Freund. Injective models and disjunctive relations. *Journal of Logic and Computation*, 3(3):231–247, 1993.

[54] Michael Freund. Preferential reasoning in the perspective of poole default logic. *Artificial Intelligence*, 98(1-2):209–235, 1998.

[55] Dov M. Gabbay. Theoretical foundations for nonmonotonic reasoning in expert systems. In K. R. Apt, editor, *Logics and Models of Concurrent Systems*. Springer-Verlag, Berlin, 1985.

[56] Peter Gärdenfors. Conditionals and changes of belief. In I. Niiniluoto and R. Tuomela, editors, *The Logic of Scientific Change*, pages 381–404. North Holland, Amsterdam, 1978.

[57] Peter Gärdenfors and David Makinson. Revisions of knowledge systems and epistemic entrenchment. In M. Vardi, editor, *Proceedings of the Second Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 83–95. Morgan Kaufmann, Los Altos, 1988.

[58] Allan F. Gibbard. Thoughts, norms, and discursive practices: Commentary on brandom. *Philosophy and Phenomenological Research*, 56(3):699–717, 1996.

[59] Benedikt Paul Göcke, Martin Pleitz, and Hanno von Wulfen. How to kripke brandom's notion of necessity. In Bernd Prien and David P. Schweikard, editors, *Robert Brandom. Analytic Pragmatist*. Ontos, 2008.

[60] Nelson Goodman. *Fact, Fiction & Forecast*. University of London, London, 1954.

[61] Nelson Goodman. *Problems and Projects*. Bobbs-Merrill, Indianapolis, 1972.

[62] Nelson Goodman and Willard V. Quine. Steps toward a constructive nominalism. *Journal of Symbolic Logic*, 12(4):105–122, 1947. Reprinted in Goodman [61], pages 173-198.

[63] Peter J. Graham. Brandom on singular terms. *Philosophical Studies*, 93(3), 1999.

[64] Mark Greenberg and Gilbert Harman. Conceptual role semantics. In Ernest Lepore and Barry Smith, editors, *Oxford Handbook of Philosophy of Language*. Oxford University Press, 2007.

[65] Peter Grönert. Brandom's solution to the objectivity problem. *Pragmatics & Cognition*, 13(1):161–175, 2005. Reprinted in Pirmin Stekeler-Weithofer, editor. *The Pragmatics of Making it Explicit*, John Benjamins Publishing, 2008.

[66] Jürgen Habermas. From kant to hegel and back again – the move towards decentralization. *European Journal of Philosophy*, 7(2):129–157, 1999.

[67] Jürgen Habermas. From kant to hegel: On robert brandom's pragmatic philosophy of language. *European Journal of Philosophy*, 8(3):322–355, 2000.

[68] Gilbert Harman. Conceptual role semantics. *Notre Dame Journal of Formal Logic*, 28(April):242–56, 1982.

[69] Anandi Hattiangadi. Making it implicit: Brandom on rule-following. *Philosophy and Phenomenological Research*, 66(2):419–31, 2003.

[70] Anandi Hattiangadi. Is meaning normative? *Mind and Language*, 21(2):220–240, 2006.

[71] Anandi Hattiangadi. *Oughts and Thoughts: Rule-Following and the Normativity of Content*. Oxford University Press, 2007.

[72] G. E. Hughes and M. J. Cresswell. *A New Introduction to Modal Logic*. Routledge, 1996.

[73] Theo Janssen. *Foundations and Applications of Montague Grammar*. Centrum voor Wiskunde en Informatica, Amsterdam, 1986. ISBN 90-6096-292-7.

[74] Theo Janssen. Compositionality. In Johan van Benthem and Alice Meulen, editors, *Handbook of Logic and Language*, pages 417–473. MIT Press, Cambridge, MA, 1997. ISBN 0262220539.

[75] Katarzyna M. Jaszczolt. *Default Semantics: Foundations of a Compositional Theory of Acts of Communication*. Oxford University Press, Oxford, 2005.

[76] Hans Kamp and Uwe Reyle. *From Discourse to Logic*. Kluwer Academic Publishers, 1993.

[77] Sarit Kraus, Daniel J. Lehmann, and Menachem Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44(1-2):167–207, 1990.

[78] Saul Kripke. Outline of a theory of truth. *Journal of Philosophy*, 72(19):690–716, 1975.

[79] Saul Kripke. *Naming and Necessity*. Harvard University Press, Cambridge, MA, 1980.

[80] Saul Kripke. *Wittgenstein on Rules and Private Language*. Cambridge: Harvard University Press, 1982.

[81] Mark Lance. Two concepts of entailment. *The Journal of Philosophical Research*, 20:113–137, 1995.

[82] Mark Lance and John O'Leary-Hawthorne. *The Grammar of Meaning*. Cambridge University Press, 1997.

[83] Daniel Lehmann. Nonmonotonic logics and semantics. *Journal of Logic and Computation*, 11:229–256, 2001.

[84] Daniel J. Lehmann, Menachem Magidor, and Karl Schlechta. Distance semantics for belief revision. *Journal of Symbolic Logic*, 66(1):295–317, 2001.

[85] Clarence Irving Lewis. *Mind and the World-Order*. Dover Publication, New York, 1956.

[86] Clarence Irving Lewis and Cooper Harold Langford. *Symbolic Logic*. Century Company, New York, 1932. Reprinted by Dover Publications (New York) in a 2nd edition, 1959, with a new Appendix III by Lewis, "Final Note on System S2".

[87] David Lewis. Scorekeeping in a language game. *Journal of Philosophical Logic*, 8(1): 339–59, 1979.

[88] John Locke. *Essay Concerning Human Understanding*. Claredon Press, Oxford, 1689/1975.

[89] Danielle Macbeth. *Frege's Logic*. Harvard University Press, 2005.

[90] David Makinson. General patterns in nonmonotonic reasoning. In *Handbook of logic in artificial intelligence and logic programming (vol. 3): nonmonotonic reasoning and uncertain reasoning*, pages 35–110. Oxford University Press, Inc., New York, NY, USA, 1994. ISBN 0-19-853747-6.

[91] David Makinson. *Bridges from Classical to Nonmonotonic Logic*. London: College Publications. Series: Texts in Computing, vol 5, 2005.

[92] Edwin D. Mares. *Relevant Logic: A Philosophical Interpretation*. Cambridge University Press, Cambridge, 2004.

[93] Mark McCullagh. Do inferential roles compose? *Dialectica*, 57:431–438, 2003.

[94] Mark McCullagh. Inferentialism and singular reference. *Canadian Journal of Philosophy*, 35(2):183–220, 2005.

[95] John McDowell. De re senses. *Philosophical Quarterly*, 34(136):283–294, 1984. Reprinted in McDowell [99].

[96] John McDowell. Wittgenstein on following a rule. *Synthese*, 58(March):325–364, 1984.

[97] John McDowell. Intentionality "de re". In Ernest Lepore and Robert Van Gulick, editors, *John Searle and His Critics*. Blackwell, Cambridge, 1991. Reprinted in McDowell [99].

[98] John McDowell. *Mind and World*. Harvard University Press, Cambridge, MA, 1994.

[99] John McDowell. *Meaning, Knowledge, and Reality*. Harvard University Press, 1998.

[100] John McDowell. Having the world in view: Sellars, kant, and intentionality. *Journal of Philosophy*, 95(9):431–492, 1998. Reprinted in McDowell [102, Ch. 1-3].

[101] John McDowell. Motivating inferentialism: Comments on Making It Explicit (ch. 2). *Pragmatics and Cognition*, 13(1):121–140, 2005.

[102] John McDowell. *Having the World in View: Essays on Kant, Hegel, and Sellars*. Harvard University Press, 2009.

[103] John McDowell. Why is sellars's essay called "empiricism and the philosophy of mind"? In *Having the World in View: Essays on Kant, Hegel, and Sellars* McDowell [102], pages 221–238.

[104] Robert K. Meyer. Conservative extension in relevant implication. *Studia Logica*, 31: 39–48, 1972.

[105] Ruth Millikan. *Language, Thought and Other Biological Categories.* MIT Press, 1984.

[106] Ruth Millikan. *On Clear and Confused Ideas: An Essay About Substance Concepts.* Cambridge University Press, 2000.

[107] Ruth Millikan. *Varieties of Meaning.* MIT Press, 2004. The 2002 Jean Nicod Lectures.

[108] Ruth Millikan. *Language: A Biological Model.* Clarendon Press, Oxford, 2005.

[109] Richard Montague. Universal grammar. *Theoria*, 36(3):373–398, 1970.

[110] Jaroslav Peregrin. *Meaning and Structure.* Ashgate Publishing, Aldershot, 2001.

[111] Jaroslav Peregrin. Pragmatismo e semantica. In Jr. P. Ghirardelli and J. Zabatiero, editors, *Filosofia Americana.* CEFA & FTSA, 2003. English original available at *http://jarda.peregrin.cz/mybibl/PDFTxt/482.pdf.*

[112] Jaroslav Peregrin. Developing sellars' semantic legacy: meaning as a role. pages 257–274, 2006.

[113] Jaroslav Peregrin. Meaning as an inferential role. *Erkenntnis*, 64(1):1–35, 2006.

[114] Jaroslav Peregrin. Inferentialism and the compositionality of meaning. *International Review of Pragmatics*, 1(1):154–181, 2009.

[115] Jaroslav Peregrin. Brandom's incompatibility semantics. *Philosophical Topics*, 2(36), 2009.

[116] Robert B. Pippin. Brandom's hegel. *European Journal of Philosophy*, 13(3):381–408, 2005.

[117] David Poole. A logical framework for default reasoning. *Artificial Intelligence*, 36 (1):27–47, 1988. ISSN 0004-3702.

[118] Dag Prawitz. Meaning approached via proofs. *Synthese*, 148(3), 2006.

[119] Arthur N. Prior. The runabout inference-ticket. *Analysis*, 21(2):38–39, 1960.

[120] Willard V. Quine. Nominalism. In Dean W. Zimmerman, editor, *Oxford Studies in Metaphysics*, volume 4, pages 3–21. Oxford University Press, Oxford, 1946. *Nominalism* was presented at the Harvard Philosophical Colloquium in Emerson B, Cambridge, Massachusetts on March 11, 1946. The 86 unpublished 3x5 note cards were transcribed by Paolo Mancosu. Reprinted in Quine [129].

[121] Willard V. Quine. On what there is. *Review of Metaphysics*, 2, 1948. Reprinted in Quine [123].

[122] Willard V. Quine. Two dogmas of empiricism. *Philosophical Review*, 60(1):20–43, 1951. Reprinted in Quine [123].

[123] Willard V. Quine. *From a Logical Point of View*. Harvard University Press, 1953.

[124] Willard V. Quine. Reference and modality. In Willard Orman Quinvane, editor, *From a Logical Point of View*. Harvard University Press, 1953.

[125] Willard V. Quine. Quantifiers and propositional attitudes. *Journal of Philosophy*, 53(5):177–187, 1956.

[126] Willard V. Quine. *Word and Object*. MIT Press, 1960.

[127] Willard V. Quine. Epistemology naturalized. In *Ontological Relativity and Other Essays*, pages 69–90. Columbia University Press, 1969.

[128] Willard V. Quine. *Pursuit of Truth*. Harvard University Press, 1992.

[129] Willard V. Quine. *Confessions of a Confirmed Extensionalist: And Other Essays*. Harvard University Press, 2008.

[130] Frank P. Ramsey. Universals. *Mind*, 34(136):401–417, 1925.

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

[132] François Recanati. *Literal Meaning*. Cambridge University Press, 2004.

[133] François Recanati. Deixis and anaphora. pages 286–316. Oxford University Press, Oxford, 2005.

[134] Sebastian Rödl. Normativity of mind versus philosophy as explanation. In Bernhard Weiss and Jeremy Wanderer, editors, *Reading Brandom: Making it Explicit*. Routledge, 2010.

[135] Richard Rorty. Representation, social practise, and truth. *Philosophical Studies*, 54 (2):215–228, 1988.

[136] Gideon Rosen. Who makes the rules around here? *Philosophy and Phenomenological Research*, 57(1):163–171, 1997.

[137] Hans Rott. *Change, Choice and Inference: A Study of Belief Revision and Nonmonotonic Reasoning*, volume 43 of *Oxford Logic Guides*. Oxford University Press, 2001.

[138] Richard Routley and Robert K. Meyer. The semantics of entailment ii. *Journal of Philosophical Logic*, 1(1):53–73, 1972.

[139] Richard Routley and Robert K. Meyer. The semantics of entailment iii. *Journal of Philosophical Logic*, 1(2):192–208, 1972.

[140] Richard Routley and Robert K. Meyer. The semantics of entailment. In Hughes Leblanc, editor, *Truth, Syntax, and Modality*, pages 199–243. North Holland, 1973.

[141] Gilbert Ryle. 'if', 'so' and 'because'. In Max Black, editor, *Philosophical Analysis*. Englewood Cliffs, Nj: Prentice Hall, 1950.

[142] Karl Schlechta. *Nonmonotonic Logics, Basic Concepts, Results, and Techniques*, volume 1187 of *Lecture Notes in Computer Science*. Springer, 1997. ISBN 3-540-62482-1.

[143] Wilfrid Sellars. Concepts as involving laws and inconceivable without them. *Philosophy of Science*, 15(October):287–313, 1948.

[144] Wilfrid Sellars. Quotation marks, sentences, and propositions. *Philosophy and Phenomenological Research*, 10(4):515–525, 1950.

[145] Wilfrid Sellars. Particulars. *Philosophy and Phenomenological Research*, 13(2):184–199, 1952.

[146] Wilfrid Sellars. Inference and meaning. *Mind*, 62(247):313–338, 1953.

[147] Wilfrid Sellars. Empiricism and the philosophy of mind. *Minnesota Studies in the Philosophy of Science*, 1:253–329, 1956. Reprinted in Sellars [155, Ch. 5], pp. 127-96; republished as a book with an introduction by Richard Rorty, and a study guide by Robert Brandom. Cambridge, Mass.: Harvard University Press, 1997; it.: Empirismo e filosofia della mente. *Iride*. 4/5:159-222, 1990.

[148] Wilfrid Sellars. Counterfactuals, dispositions, and the causal modalities. In Michael Scriven Herbert Feigl and Grover Maxwell, editors, *Minnesota Studies in The Philosophy of Science, Vol. II*, pages 225–308. University of Minnesota Press, Minneapolis, 1957. Reprinted in part in Ernest Sosa, editor. *Causation and Conditionals*. Oxford University Press, 1975.

[149] Wilfrid Sellars. Substance and form in aristotle. *Journal of Philosophy*, 54(22): 688–699, 1957.

[150] Wilfrid Sellars. The language of theories. In Herbert Feigl and Grover Maxwell, editors, *Current Issues in The Philosophy of Science*, pages 57–77. Holt, Rinehart, and Winston, New York, 1961. Reprinted in Sellars [155].

[151] Wilfrid Sellars. Naming and saying. *Philosophy of Science*, 29(1):7–26, 1962.

[152] Wilfrid Sellars. Philosophy and the scientific image of man. In Robert Colodny, editor, *Frontiers of Science and Philosophy*, pages 35–78. University of Pittsburgh Press, Pittsburgh, 1962. Reprinted in Sellars [155], pp. 35-78.

[153] Wilfrid Sellars. Truth and "correspondence". *Journal of Philosophy*, 59(2):29–56, 1962. Reprinted as Sellars [155, Ch. 6].

[154] Wilfrid Sellars. Abstract entities. *Review of Metapysics*, 16:627–71, 1963. Reprinted in Sellars [158].

[155] Wilfrid Sellars. *Science, Perception and Reality*. Routledge and Kegan Paul, London, 1963. Republished by Ridgeview Publishing Company, Atascadero, California, 1991.

[156] Wilfrid Sellars. Induction as vindication. *Philosophy of Science*, 31(3):197–231, 1964.

[157] Wilfrid Sellars. Aristotle's metaphysics: An interpretation. In *Philosophical Perspectives*, pages 73–124. C.C. Thomas, Springfield, ILL, 1967.

[158] Wilfrid Sellars. *Philosophical Perspectives.* C.C. Thomas, Springfield, ILL, 1967. reprinted in two volumes by Ridgeview Publishing Co.: Part. 1: Hisory of Philosophy; Part 2: Metaphysics and Epistemology.

[159] Wilfrid Sellars. *Science and Metaphysics: Variations on Kantian Themes.* Humanities Press, New York, 1968.

[160] Wilfrid Sellars. Metaphysics and the concept of a person. In Karel Lambert, editor, *The Logical Way of Doing Thing*, pages 219–52. Yale University Press, New Haven, 1969. Reprinted in Sellars [161].

[161] Wilfrid Sellars. *Essays in Philosophy and its History.* Reidel, 1974.

[162] Wilfrid Sellars. *Naturalism and Ontology.* Ridgeview Publishing Co., Reseda, California, 1980. The John Dewey Lectures for 1973-4.

[163] Robert Stainton. *Words and Thoughts: Subsentences, Ellipsis, and the Philosophy of Language.* Oxford University Press, 2006.

[164] Alfred Tarski. über eine fundamentale begriffe der metamathemamik. *Comptes Rendus des Seances de la Société des Sciences et des Lettres de Varsovie*, 23(III):22–29, 1930. En.: On some fundamental concepts of methamatematics. In *Logic,Semantics, Metamathematics*, Claredon Press: Oxford, 1956.

[165] Alfred Tarski. The semantic conception of truth: And the foundations of semantics. *Philosophy and Phenomenological Research*, 4(3):341–376, 1944.

[166] Alfred Tarski. The concept of truth in formalized languages. In J.H. Woodger, editor, *Logic, Semantics, Metamathematics*, pages 152–278. Oxford Claredon Press, 1956.

[167] Giacomo Turbanti. Belief reports: Defaults, intentions and scorekeeping. In Piotr Stalmaszczyk, editor, *Philosophy of Language and Linguistics*, volume I. Ontos Verlag, 2010.

[168] Alasdair Urquhart. Semantics for relevant logics. *The Journal of Symbolic Logic*, 37 (1):159–169, 1972.

[169] Jeremy Wanderer. *Robert Brandom*. Acumen Publishing, 2008.

[170] Bernhard Weiss and Jeremy Wanderer. *Reading Brandom: Making it Explicit*. Routledge, 2010.

[171] Dag Westerståhl. On mathematical proofs of the vacuity of compositionality. *Linguistics and Philosophy*, 21(6), 1998.

[172] Kneale William and Kneale Martha. *The Development of Logic*. Oxford University Press, 1962.

[173] Ludwig Wittgenstein. *Tractatus Logico-Philosophicus*. Dover Publications, 1922.

[174] Ludwig Wittgenstein. *Philosophical Investigations*. Macmillan, New York, 1953.

[175] Ludwig Wittgenstein. *On Certainty*. Arion Press, 1969.

[176] Ludwig Wittgenstein. *Remarks on the Foundations of Mathematics*. Basil Blackwell, 1978.

[177] Wlodek Zadrozny. From compositional to systematic semantics. *Linguistics and Philosophy*, 17(4), 1994.

[178] Henk Zeevat. The asymmetry of optimality theoretic syntax and semantics. *Journal of Semantics*, 17:243–262, 2000.