A paraconsistent-preservationist approach to
a common confusion concerning predicate-extensions

Martin W. Allen*
Dept. of Philosophy
1001 CL, University of Pittsburgh
Pittsburgh, PA 15260

January 2, 2001

Abstract

The existence of multiple criteria for the introduction of a predicate may lead to confusion when the criteria diverge as to whether or not some object falls under the predicate. It can be difficult to represent the semantics of sentences featuring such a predicate-term, and it is not obvious how a person is supposed to employ such confused terms in the business of language and reasoning. I consider two approaches to the problem: dialethism, which allows both a sentence and its negation to be true at once; and disambiguation, which represents any such confused predicate in terms of other, distinct predicates. I show the equivalency of plausible formal treatments of these approaches, discuss reasons for this equivalency, and present an alternative approach—a preservationist one, which does not interpret the confused predicates but rather seeks to contain the confusion present. I argue that a meaningful, and useful, concept of inference is available, even where the semantics of certain predicate-terms remain confusing.

1 Interpretation and commitment

We talk, we listen, we think. We argue, and we (try to) persuade. All common enough, and all involving us in the business of belief and interpretation. The discussion of what we mean when we speak can lead to talk about what we believe—talk which does not necessarily proceed in terms of “contents” or “representations,” but at least in terms of inference, implication, and commitment. Beliefs can be controversial items; whether they exist as “things in the head” or not, and what sort of things they are, are questions (among others) which remain open. So, beliefs can be controversial, but I take it that commitments are less so. Practices of discussion and argumentation—making claims, drawing consequences, proposing refutations and reductions—all rely, at least in part, upon the idea that we can and will be held accountable. What we say and do leaves it open for others to hold us responsible—and our commitments consist in these expected responsibilities. We all draw inferences from the speech of others, and expect them to do much the same, as regards our speech, and their own. Or, if this is too strong, if it suggests (perhaps wrongly) that we perform

*Thanks go to Bryson Brown, whose work inspired much of this.
and expect actual inferences, we can still at least talk about “commitment” in terms of potential inference, of being ready to account, someday, for what we say and do now. This accountability, then, comprises no small part of what someone can be said to mean, in speech and action; on this way of understanding things, the interpretation of others consists largely in decisions as to those things to which they are to be held accountable.

So what role, if any, does logic play in all of this? A first thought is that logic deals with inference, with relations between sentences (or facts, or states of affairs, or beliefs, or . . . ), and with “movement” from some sentences to others. As such, a logical system can provide a model for the ways in which we expect a person’s commitments to fit together. The principles of such a system—its precise description of such as implication or entailment—can guide (or at least represent) our expectations as to how our further commitments should follow from our current ones. In our interpretations, as in our logical systems, instances of inference (actual or potential, observed or expected) give way to principles of inference, and so what one is committed to expands. For example (in terms which recall Brandom1), interpretations move from that to which someone explicitly commits, to that to which she implicitly commits, or at least to that to which we expect her to commit. Again, these expectations need not be thought of in terms of actual inference. Indeed, there will be times when we think that one should not actually that infer to which one is committed—for reasons, say, of redundancy, irrelevance, or incoherence. Nonetheless, the idea of commitment plays an important role, even here. It is precisely because we can sometimes say that one is committed to what one ought not infer, that we can argue for changes in what one says or believes; that is, changes in one’s commitments. Conversation and argument rely heavily upon the presupposition that what someone already claims or believes commits them to other things, salient in varying circumstances, under numerous present or future considerations—such things as purpose of conversation or goal of inquiry.

Opinions vary as to logic’s proper role here, and vary over whether there is such a thing. To take a popular, and interesting, example, the Davidsonian view has it that conversation is interpretation, where this consists in our coming up with a “logical picture” of what others say and believe. This “picture” consists in a fairly well-developed theory, a logical model of the truth-conditions (and so the content) for the claims made and for the beliefs and commitments which seem to lie behind those claims, all against a background system of logical-inferential connections and constraints.2 Some, Hacking among them, have found this sort

---

of theory-based model of conversational interpretation unrealistic, especially if the idea is taken literally, in terms of explicit, detailed, and consciously worked-out logical notions on the part of the interpreter (what I will call the “explicitist” reading). As Bob Brandom has it, “linguistic understanding depends on interpretation in this sense only in extraordinary situations—where different languages are involved, or where ordinary communication has broken down.” On his own view, which he means to develop thoughts from Frege and Wittgenstein (among others), interpretation consists not so much in explicitly formulating hypotheses and theories about others’ beliefs, but rather in practices involving the implicit adoption and attribution of commitments. As Brandom sees things, interpretation proceeds more often by way of our own actions towards others, than by way of explicitly formulated expectations concerning their beliefs and behaviours. Still, such a view leaves logic some role to play with respect to commitment, even if it is more likely descriptive of our practices—viewed as it were “from a distance”—than it is in fact something we self-consciously accept and employ, at first hand, in the business of interpretation and conversation.

So, in the context of linguistic interaction, logical principles models, can play at least two distinct kinds of role. Given what we take someone to believe, or to be committed to, based upon what they have already said and done, we ascribe further beliefs or commitments to that person. On the one hand, this procedure of further ascription may be overt. If we are “explicitist” about it, we will think that interpretation proceeds according to logical principles which we acknowledge and accept. In dealing with our interpretive subjects, we come up with partial models of their apparent current condition, and proceed to “fill in the blanks,” based upon principles of inference which we are able to formulate, and tend to endorse. Or, on the other hand, the procedure of interpretation may be covert. That is, it may be only implicit in our practice, a logical system serving only to represent the patterns into which our ascriptions happen to fall. Either way, logical models of inferential and semantic connections between sentences can describe connections and relationships holding between those things that we ascribe to our subjects. This sense of “interpretation”—systematic ascription of commitment or belief—comprises some part of meaningful conceptual and linguistic interaction, and when we study logic, at least sometimes, we study ways in which these interactions may proceed.

3 For Hacking’s criticism, see his “Parody of Conversation,” Philosophy of Donald Davidson: A Perspective on Inquiries into Truth and Interpretation, E. LePore, ed. (Oxford: Blackwell, 1986), 447–458. Brandom discusses this briefly at Brandom (1994), 508–509; see also 702n9, where other sources of criticism are mentioned.


5 Brandom (1994), chapter 8 especially.
2 A problem for interpretation

With the preceding remarks as background, I want to consider a particular problem which may arise with respect to interpretation. The general problem concerns inconsistent belief and commitments; in particular, puzzles concerning the further commitments of someone who is already marked by inconsistency. The connection between this problem and logical treatments of interpretation is here relatively straightforward. After all, it is well known that, in the presence of inconsistency, classical logic is unprincipled with respect to inference, entailment, and consequence—everything follows from an inconsistent set. So, for Davidson, consistency becomes a necessary condition of any interpretation. That is, consistency is part and parcel of the “constitutive ideal of rationality,” an essential feature of those things we call “beliefs.”

When interpreting anyone, Davidson argues, we labour under “a burden”: we must try to find whatever consistency we can. It is just a feature of what an interpretation is that a subject’s beliefs and commitments be found consistent; this is so, Davidson says, because “[w]e weaken the intelligibility of attributions of thoughts of any kind to the extent that we fail to uncover a consistent pattern of beliefs..., for it is only against a background of such a pattern that we can identify thoughts.” If the logic at work in interpretation is classical, and inconsistent sets yield up unprincipled consequences under such a logic, then hope fades for any “pattern” of belief or commitment once one has admitted inconsistency into the picture. Any inconsistent set of beliefs or commitments would seem to be the same as any other, since all would generate the same universal class of further commitments, and, for all practical purposes, would seem to have the same content: that is, none at all.

Though this story certainly has its problems, it still gets things at least partly right. We can learn something from the classical treatment of inconsistency, even if we do not subscribe to the “explicitist” treatment of Davidson’s ideas, and regard logical systems and models as only describing what is merely implicit in our practice. If nothing else, we can say that if our inferential and attributional practices were classical, then, for all intents and purposes, a person with inconsistent beliefs or commitments would be uninterpretable (and, we might also say, irrational). Whatever it means to say that some system of classical logic “represents” the principles of interpretation, it would mean even less to say so, but deny that inference from inconsistency was unprincipled—if this were not so, then the system at work simply would not be a

---

7 For discussion of the relation between inference and content, see Brandom (1994), passim, but especially chapter 2, and the discussion of inferential articulation, 168-172.
8 For one thing, I think, some (at least slight) inconsistency of belief and commitment is so universal among real-world speakers and reasoners as to make any insistence upon its absence or elimination wishful thinking at best. And notice also that there is really no such thing as “slight inconsistency” from a classical point of view—every inconsistent set is equivalent to every other one, with respect to classical inference, entailment, and satisfaction.
classical one.\textsuperscript{9}

So, such problems with inconsistency suggest either that the logical basis of interpretation should be non-classical, or else that such individuals as appear inconsistent must be re-interpreted. On the first view, some new logical framework is needed, so that the inconsistency is acknowledged, and yet something productive—something limiting—can be said concerning further commitments and beliefs. On the latter view, the inconsistent individual must be re-described, such that inconsistency vanishes, before any meaningful interpretation is even possible. This much, then, is clear—but the choice between these different options is not always going to be so obvious.

2.1 A particular case.

At this point, I want to further complicate the story, by pointing out the deep problems inherent in the very idea of attributing inconsistent beliefs and commitments. As has already been noted, a straightforward and seemingly sensible view has it that interpretation just is the principled and systematic ascription of commitments—actual, potential, and expected. From this perspective, it starts to look as if, absent any such principled or systematic treatment of a subject, we fail to attribute anything recognizable as “belief” or “commitment.” The presence of confusion in belief threatens the very existence of belief. This hardly seems right, however. We often talk about “confused beliefs,” and it certainly does not seem as if such things were simply impossible, nor non-existent.

Furthermore, the confusions that we identify in subjects of interpretation threaten to propagate. Faced with others who seem confused, we their interpreters risk also becoming confused. If we base our interpretive claims upon the way they think about things, we may find the confusion in their thought lapping over onto our talk. Consider an example, adapted from Gupta,\textsuperscript{10} of persons whose language or “conceptual scheme” contains a predicate, the term or concept “up,” which is, as it were, “fractured.” As it turns out, these individuals possess dual criteria for the introduction and application of the predicate. On the one hand, they have a “perceptual” criterion, pretty much the same as our own, a rough common standard for when one object looks “up” relative to other objects. On the other hand, believing the world to be flat, these persons think that a “conceptual” criterion also applies.\textsuperscript{11} On their “flatland” theory, then, a line can be drawn from any object to any others which appear to be “up” above it; further, all these (possible) lines are in

\textsuperscript{9}That is, some other sort of system would be at work in the case at hand, or else there is no real meaning at all to claims that any sort of logical system was “at work.”

\textsuperscript{10}Anil Gupta, “Meaning and Misconceptions,” unpublished.

\textsuperscript{11}The terms “perceptual” and “conceptual” are Gupta’s own.
fact parallel to one another. According to the flatlanders’ understanding of the world and their conceptual criterion, a direction can be designated the “standard up”—a canonical direction that, once determined at any one spot in the world, can be fixed and applied as objective standard for the introduction and application of the predicate “up,” anywhere else in the world.\textsuperscript{12}

We know better, however. As it happens, of course, the world is in fact round, and various lines drawn between objects and those others which look to be “up” above them are not parallel; no matter how they appear locally, these lines all in fact intersect at the world’s center. As far as the flatlanders are concerned, the two criteria for “up”-introduction are interchangeable: whatever perceptually looks “up” is also objectively or conceptually “up,” and vice-versa. On the other hand, we realize that the two criteria in fact pick out different properties and that the predicate is being introduced over distinct classes of objects.\textsuperscript{13} In most cases, then, flatlanders will apply the term “up” to any objects correctly as regards one criterion for “up”-introduction, and incorrectly as regards the other. And, since both criteria are always equally available, a person employing such a confused concept of “up” may well find herself both wanting to say that some object \(a\) both is and is not “up” with respect to some other object \(b\).

So, the user of this dual-criteria version of “up” risks falling into some difficulty. Of course, the flatlanders may not ever notice this divergence in their predicate-concept. Still, the problem of interpretation remains. From our point of view, once the divergent criteria become apparent, it can be hard to say just which beliefs and commitments, if any, to ascribe to someone possessed of such a confused (and now confusing) predicate-concept. Normally, the process of ascription might go something like this: having identified the circumstances under which our subject would normally want to apply the “up”-predicate, and having identified that one such criterion has in fact been met, we ascribe to the subject the corresponding belief or commitment. But what do we do in the case when one criterion has been met, but another, supposedly equivalent, has not? We might say that our subject is committed to believing or claiming that “\(a\) is above \(b\),” and we might not. The flatlander, recall, thinks that the two criteria apply indifferently—satisfaction of either one is sufficient for applying the predicate. So, we should ascribe the appropriate belief or commitment. At the same time, however, failing to satisfy either criterion supposed to be equivalent to satisfying none of them, and so is reason enough to say that the predicate does not apply; in which case, we should not make the ascription (and in fact ought to ascribe a belief in, or commitment to the negation of “\(a\) is above \(b\)”).

\textsuperscript{12}We may imagine, as Gupta does, that these persons possess some sort of surveying device which will allow them, if they choose, to determine the standard direction of “up,” and whether or not two objects do in fact lie on a line parallel to that direction.

\textsuperscript{13}These classes do not intersect, except perhaps immediately local to the “canonical conceptual ‘up,’” the place where the criteria in fact match up, where any object \(a\) which looks “up” above another object \(b\) lies on a line \(a\rightarrow b\), which is identical with, or parallel to, the canonical “up.”
The case is invented, and perhaps a little contrived, but the situation is readily imaginable, and analogous to more common and less fanciful cases. For one, the use of natural-kind terms springs to mind. Often, the perceptual character of something like water, gold, or jade is (or has been) thought to coincide with just one particular (scientifically and conceptually determined) chemical composition. Often enough, however, it turns out that the things we identify as “looking the kind” belong to a different (albeit often overlapping) class than those things which fit the technical characterization of the type. In such cases, as in Gupta’s, our concepts and predicate-terms have become somewhat confused. The original idea is of a single kind of thing, a single predicate (“up,” “gold,” “water,” “jade,” etc.) that can be applied according to criteria which are multiple, but indifferent, so far as they are supposed to pick out equivalent classes of objects; that is, equivalent predicate-extensions. When the criteria diverge, however, cases arise on which the predicate-concept or -term does apply according to one criterion, but does not according to the other, and confusion threatens.

2.1.1 A note.

I should mention here that my interests with respects to these sorts of cases differ from Gupta’s, even though it was his work which inspired me to think about them. As I see it, the question for the logician, or logically-minded interpreter, is whether (and if so, how) a person whose language or thought contains confusions and resultant inconsistencies can still be seen to be employing their ideas and their talk reasonably or rationally. The idea is to explain, so far as we can, how one could (and should) reason in the presence of such confusions, and to think about what further commitments we can ascribe to the confused individual.

For Gupta, on the other hand, the problem translates into concern with how confused concepts could possibly serve the needs of engaged and active agents in the world. That is, he thinks, we need and want to explain how we can use these sorts of concepts and predicates—among other things, accounting for how persons employ the concepts in reasoning, and the terms in speech, in order to direct one another’s actions, and to relay information about such things as “that clock up above the sink.” In aid of this, Gupta considers the idea of “frames,” context-bound conventions which decide, for users of the fractured term “up”, which criterion applies in which circumstances (even if users are unaware of those conventions). Others have suggested that no such story will be adequate, and that we should not expect to explain the employment of confused terms to convey useful information, unless we already possess an explanation of how the use of words enables us to ‘get along’ in normal circumstances. Furthermore, this line of thought goes, no such simple explanation of the pragmatic virtues of language is currently available, nor is it a simple matter to
say when instances of language-use are “normal” or not.\textsuperscript{14}

Intuitions will vary on this matter, of course, but there is little to be done about that. In any case, my interest here lies not so much with the pragmatics of language as with logical interpretation of belief and commitment. It is not obvious that logic alone should be expected to be able to determine all the practical uses to which persons can and cannot put bits of their linguistic and conceptual repertoire. In keeping with this thought, then, I will try to stick to what logic can tell us: about the operation of inference and reasoning in the presence of confused predicate-terms and -concepts.

\section{A role for logical methods.}

As noted, inference and argument play a role in determining what claims to make concerning someone’s further beliefs and commitments. That is, claims about what someone believes now often draw upon other, prior claims concerning what the person believes already, with inferential and argumentative principles acting as interpretive guides, taking the interpreter from subject’s prior belief to subject’s current belief. As such, the interpretive transition from prior to current belief often invokes, implicitly or explicitly, the forms and standards of reasoning that are in some wise represented by these sorts of argumentative-inferential principles. In at least some instances interpretation depends upon what (the interpreter thinks) it would be right for someone to infer or argue, based upon present commitments. Traditionally, logic has been taken to be—at least in part—a systematic treatment of just this idea. Among other things, that is, the logician means to have something to say concerning the process of sorting good arguments from bad, and about what one ought (or at least is permitted) to infer on a given occasion. From the usual, classical perspective, however, such cases of inconsistency and confusion as have been described play havoc with our notions of “proper” or “valid” inference and argument. When faced with instances of genuinely inconsistent predication—as when, given an interpretive subject’s understanding of the predicate “up,” it seems right to say of something that it is both up and not up at once—the classical point of view holds that there is really no sense to be made of the predicate-term in question. On this view, apparent contradictions signal a breakdown in reasoning and understanding, and the unprincipled nature of classical consequence in the face of inconsistency reflects this breakdown. From inconsistent premises, the story goes, there are no good arguments, and so nothing to say about what else one ought and ought not infer—what else to which one ought and ought not commit. Further, predicate-concepts which lead into contradiction are not the sorts of things from which

\textsuperscript{14}Joe Camp has suggested as much (as at least a hypothetical objection) in a recent seminar on confusion in language and interpretation, offered at the University of Pittsburgh, Fall, 1998.
consistent and interpretable commitments are ever to be made. If nothing else, then, it is clear that making any sense of what a confused individual should or should not believe, infer, or accept, will require either “classicalizing” the operative predicate-concepts to somehow repair their inconsistency-generating defects, or else the interpretive employment of other-than-classical inferential principles.

3.1 The possibility of propagation.

So far, I have been discussing the unprincipled nature of logical consequence in the presence of inconsistency as if it were a bad thing in and of itself. In some ways, it is true, inferential volatility (the tendency to explode inferentially) is an artifact of the classical theory of consequence—after all, nobody really thinks that the simply inconsistent person is actually committed to everything. On the other hand, the unfettered promulgation of consequence from inconsistency gets something right; namely, that some degree of confusion tends to lead to ever greater confusion. While it may be true that not every undesirable consequence follows from a single moment of inconsistency or confusion, some do. Perhaps, then it is worth thinking about confusion’s tendency to spread, and to consider efforts to deal with confusion in terms of attempts to undercut this tendency. Here again, the paraconsistent approach, which is all about limiting the consequences of inconsistency in some meaningful way, will be apposite to questions concerning the “logical thing to do” in the context of belief or argument—that is, the reasonable or rational thing to do.

So, just what would be reasonable to do in cases such as those Gupta describes? Of course, we might simply want to escape from confusion. We would probably prefer that, if at all possible, our concepts (or those of some interpretive subject) be made “clean,” and that the problem of divergent criteria be uncovered and eliminated. This does not seem the logician’s proper task, however. Although the use of logical tools may point out an incoherence in one’s concepts, logic alone cannot tell us how exactly we ought to repair the problem. No logical model of the flatlanders’ commitments will, in and of itself, provide a ready-made solution to what ails them; pointing out that their employment of fractured predicate-concepts can lead to potentially contradictory consequences will not tell them which, if any, of their conceptual criteria in fact picks out the “true ‘up’.” One might argue that logical accounts of consequence have nothing beyond diagnosis to offer confused individuals, and that it is only after they gets their concepts “back into line” that logicians can offer any useful thoughts on valid argumentation and appropriate inference.

Such a limiting conception of the logician’s role cannot be right, however. While that proper role may not be obvious, there is surely more to be said about what it might be. For one thing, it is quite clear that standards of good and bad argument still apply, even where concepts have become confused. After all, we
are already presuming in all this that confusion, and by extension anything that leads to confusion, is in fact a bad thing. Where we are confused, something has gone wrong, or could go better—that much follows from having presented the case as a difficulty for the logical approach to interpretation to overcome. Clearly then, the extent to which we are confused, if it can be measured, is some standard against which processes of reasoning may be judged. Where we are confused, we may well wish that we were less so; failing that, however, it is certainly reasonable that we wish to avoid becoming more confused than we are already. If we are in search of some notion of “good argument” or “reasonable inference,” then here is a suggestion: it is reasonable (or “right” or “rational”) to argue or infer something only if it does not involve our falling into any more confusion than that in which we are mired already. It is just this sort of project to which a logician may lend an expert hand.

Consider first how easily confusion propagates. Like the north to which the compass needle tends, confusion draws our thought and talk always to itself. On the one hand, the confusion of the subject threatens to infect the interpreter. Following normal manners of speech and what I understood of her concepts, I might well find myself wanting to say of one of Gupta’s flatlanders that, given the satisfaction of one of the criteria for “up”-introduction, “she believes the clock is up above the sink,” and also that, given the failure of the other criterion, “she does not believe the clock is up above the sink.” But if I were to do that, then I myself would be saying something inconsistent, and the subject’s confusion would become my own. Of course, the threat is not really serious, and the solution simple: disengagement from the subject seems to be the best bet. I can isolate myself from the inconsistency of beliefs and commitments simply by moving the problematic negation away from my own attributions of belief and into the content of those described beliefs—placing it inside such an intensional construction as a “that”-clause, for instance. I say of my subject only, “she believes that the clock is up above the sink,” and “she believes that the clock is not up above the sink.” Further, I can insist that the latter sentence as a whole does not imply the negation of the former, even if, I agree, the sentence corresponding to the content of the second “that”-clause is the negation of that which corresponds to the content of the first. By doing so, I thereby avoid appearing inconsistent and confused myself and, if pressed for the reason why I deny the implication from one of my claims to the negation of the other, I can, if I like, reply that the context is merely de dicto. 15

Interpreting in this manner, I only claim that the “belief” or “commitment” in question is no more than some sort of thing, associated with the contents of the intensional construction in some manner; furthermore, I can do so without committing myself to any particular story about how such things could be true or false,

15For a detailed account of the following analysis of de dicto interpretation and ascription, see Brandom (1994), chapter 8.
appropriate or not. To “interpret” this way is to ascribe “beliefs” or “commitments” which exist as it were on their own, in the sense that I can make interpretive claims about them without thereby detailing any of their particular consequences. In addition, I can talk about a subject’s “belief that . . . ,” without implying that it is directly analogous to a “belief that . . . ” on my own part. In the latter case, of course, the content of my beliefs is explicated in terms of further commitments: what my beliefs or commitments actually are gets explained by way of what follows from them, or by way of what commitments are precluded by my having them. In the case of the confused interpretive subject, however, no such content, and no such explication, shall be forthcoming; to say that they possess such a belief is not to say anything else at all about their further commitments, or at least not to say that their further commitments are the same as mine would be in a similar situation. So, I am isolated from inconsistency, refusing to engage in the same game—whatever that is—as my subject.

But notice how far I have traveled from interpretation and explanation of behavior. It is one thing to note that I would wind up falling into inconsistency if I myself engaged in a linguistic or conceptual practice like the ones favoured by Gupta’s flatlanders. It is quite another thing, however, to refuse to assign any meaningful beliefs to my subject, just so that I can avoid such confusion. De dicto claims are fine if they sterilize the interpreter against conceptual and linguistic infection, but they don’t shed much light on the details of what else, exactly, a confused interpretive subject—who certainly seems to believe something contradictory—ought to believe. I sense that there is something wrong with what my subject is doing, and so take a “hands off” approach, refusing to let my own principles of inference and argumentation enter into the interpretation, refusing to credit the other with such further beliefs and commitments as would be appropriate in my own case. Now, however, I cannot say what, exactly, my subject is doing, since I can no longer say how (some sub-class of) her beliefs and commitments work at all. This goes quite beyond error theory (the claim that her beliefs or commitments are simply false or inappropriate), and quite beyond a simple translational account of interpretation. I am not saying that the subject’s beliefs are false, and I am not explaining the content of her commitments by associating them with any statements in my own language. Rather, since I am unable to say anything definite about how her “up”-beliefs relate to one another inferentially or semantically, it looks as if I can say neither that they are true nor that they are false—nor can I say anything about how they “connect up” with claims in my language, which have their own inferential and semantic interrelations. At the same time, however, the disagreements that arise, or could arise, over the various applications of “up” certainly seem like disagreements over the possible truth of conflicting claims, claims that conflict just because certain of them seem to imply that certain others of them must be false.
Once again, however, there is more to be said. After all, some standards of “reasonable” inference and argument still apply. It is reasonable that we, the interpreters, should not want to become confused by our subjects’ confusion; at the same time, it is certainly reasonable to expect that they themselves should not want to become more confused than they are already. Notice, too, how easily such “internal” confusion might propagate. We often use predicate-concepts and -terms to define other predicate-concepts and -terms, making it possible that confusion over one bit of language can become confusion over several. Consider for example the confused flatlanders. Asked for a defining example of the colour to which they apply the term “vermilion,” they reply, “Oh, that’s the colour of the lamp-shade on the lamp up above the sink.” Two questioners, colour-novices both, respond according to the two distinct criteria for “up” and identify two different lamps, with two distinctly-coloured shades. And so, the extension of the predicate “vermilion” is fractured, too. Different members of the linguistic community now identify the colour differently, according to what they (variously) take to be a paradigm instance—the shade on the lamp “up” above the stove. Later, the question, “Which cat is the Siamese cat?” elicits the response, “The one with the vermilion collar,” and conceptual contagion spreads further. And so it goes. Confusion can easily spread, instance to instance, predicate to predicate—the sort of “plague” or “stain which can never be wiped away” that W. K. Clifford once warned would affect all our beliefs and knowledge, and that of all our community, once ever introduced.16

So the logician has a role to play, after all: to propose canons of inference and argumentation sufficient to prevent confusion about one predicate from spreading to others. Even if some confusion or inconsistency is already evident in our beliefs and commitments, all is not lost; we can try to keep ourselves from any further difficulty. Clearly, however, classical canons of inference—that is, the application of the usual first-order predicate calculus—will not meet our needs in this regard. The classical idea that everything follows from an inconsistent set of sentences amounts to the idea that any confusion whatsoever is equivalent to complete confusion (or, at least, inexorably leads to complete confusion).17 Non-classical logicians are not necessarily committed to this equivalence, however, nor to the inevitability of the downward spiral into total confusion. So, a plausible goal comes into view: a “reasonable” theory of consequence, which either provides some form of repair, through resolution of problematic confusions, or at least some sort of damage control, preventing confusion from increasing.

17It turns out that Clifford worried about this, as well. See Gale (1991), 356f.
3.2 Paraconsistency and preservationism.

I have so far merely suggested an interest in non-classical alternatives for the interpretation of predicate-confusions. At this point, I would do well to be more specific about the particular alternatives I have in mind. While I will be discussing a number of distinct logical treatments of propositional and first-order languages, I will be honouring the following more-or-less typical conventions in all cases:

Notation 1 (Formal language). Let the following notational conventions stand, with respect to linguistic variables:

1. Lower-case Greek, e.g. α, β, γ, etc., ranges over sentences.
2. Upper-case Greek, e.g. Σ, Γ, ∆, etc., ranges over sets of sentences.
3. Upper-case Roman, e.g. F, G, H, etc., ranges over predicate-letters.
4. Lower-case Roman, from the beginning of the alphabet, e.g. a, b, c, etc., ranges over constant-letters.
5. Lower-case Roman, from the end of the alphabet, e.g. x, y, z, etc., ranges over variable-letters.

As indicated by the title of this paper, and by my focus on the inconsistencies generated by dual-criteria predicates, I am most interested in paraconsistent logics, and will consider a variety of distinct ways in which such logics might treat the problems at hand. Of course, saying so commits me to the claim that there is in fact some general notion of “paraconsistency,” applicable equally to a range of particular cases. Indeed, although there are any number of different approaches to paraconsistent logic, and to problems of inconsistency, I take the formulation of Da Costa and Wolf to be sufficiently general to do the trick: “the study of those logics which formalize non-trivial inconsistent theories.”

Treating ourselves to the notion of an arbitrary absurdity ⊥, we can talk about inconsistency with respect to some system or other of logical consequence. That is, Σ is an inconsistent set in system S if and only if Σ has an absurdity among its consequences: Σ ⊢⊥S.

Generally, classical inference from inconsistent sets is trivial or unprincipled in the sense that the presence of absurdity among the consequences of a set leads to “inferential explosion”—the inferability of anything at all from the inconsistent set—by way of the usual structural rules governing


\[^{20}\text{Note that the ‘absurdity’ }⊥\text{ is just some form of “trouble,” whichever is appropriate to the system at hand. In some systems, }⊥\text{ stands in for any self-contradictory expression of the object-language (e. g. }p\&\neg p\text{); in others, it represents, or is represented by, the empty set, or by there being nothing at all in the consequence position (}\Sigma ⊢⊥\text{). For my purposes, as I will make clear, I concentrate on cases in which the presence of “trouble” indicates that all consequences follow; that is, I will not here look into systems in which nothing follows in such a case.} \]
inference- and consequence-relations. Such a notion of “principled inference” can be given an explicit definition.

**Definition 1 (Principled and unprincipled inference).** Let $\models$ be some consequence-relation. Let $\Phi$ be the set of all sentences of the logical language with which the consequence-relation deals. Let $\text{Cn}_{\models}(\Sigma)$ be the set $\{ \alpha \mid \Sigma \models \alpha \}$.

For any set of sentences $\Sigma$, $\models$ is *principled relative to* $\Sigma$ iff:

$$\text{Cn}_{\models}(\Sigma) \neq \Phi.$$ 

That is, iff:

$$\exists \alpha \in \Phi: \alpha \notin \text{Cn}_{\models}(\Sigma).$$

The consequence-relation $\models$ is *unprincipled relative to* $\Sigma$ iff it is not principled relative to $\Sigma$; that is, $\models$ is unprincipled relative to $\Sigma$ iff:

$$\text{Cn}_{\models}(\Sigma) = \Phi.$$ 

That is, iff:

$$\forall \alpha \in \Phi, \alpha \in \text{Cn}_{\models}(\Sigma).$$

So to say that a consequence-relation is *paraconsistent* is just to say that it is principled, in the sense of Definition 1, with respect to *at least some* sets that are inconsistent from the point of view of classical logics. (Of course, the consequence-relations associated with such logics are generally *unprincipled* with respect to *all* such sets.) Similarly, to say that some logical *system* is paraconsistent is just to say that it contains a consequence-relation that is principled with respect to some classically-inconsistent set, in just this same sense. Such consequence-relations *contain* or *restrict* inferential explosions. The resulting theory of consequence is *non-trivial*—it takes a classically inconsistent set $\Sigma$ to a set of consequences which is free from absurdity, and so something less than the set of all possible consequences.

---

21 For discussion of inferential explosions, both in terms of systems in which inconsistent sets are those leading to absurdities, and in those on which the inconsistent set leads to nothing at all, see Schotch and Jennings (1989), 307–308; also their “Inference and Necessity,” *Journal of Philosophical Logic* 9 (1980): 328–329.
Definition 2 (Paraconsistency). Any consequence-relation $\vdash$ is paraconsistent iff some set of sentences $\Sigma$ is classically-inconsistent, but $\vdash$ is principled with respect to $\Sigma$. That is, $\vdash$ is paraconsistent iff:

$$\exists \Sigma: \exists \alpha: \alpha \notin Cn_{\vdash}(\Sigma)$$

As remarked, this is a highly general characterization of paraconsistency, and it radically underdetermines the details of any particular paraconsistent system of logical consequence. That is to say, Definition 2 does not say anything about any of the ways in which a theory of consequence will in fact limit what follows from classically-inconsistent sets of sentences. Just such a general conception will serve my purposes, however, as it provides some guidance in the examination and evaluation of various approaches to the problem of predicate-confusions. When looking for a logical system to use when interpreting in the presence of such confusions, the primary goal will be to give an account of consequence-relations which constrains the consequences of a set containing the inconsistencies generated by confusion. Doing so will at one and the same time exclude certain other things from the consequences of confusion, and will thus provide some initial basis for claims about what one ought and ought not infer and argue. This, in turn, will constitute at least a start to solving the problems of reasonableness and interpretability.

I now want to discuss a number of possible approaches which I take to answer the demand just presented. In doing so, I begin with two apparently opposed views—dialethism and disambiguation—which may already have occurred to the reader because more familiar. I will argue that these manners of interpreting our case here, although indeed at odds in some important respects, share other common features. Both approaches turn out to be motivated by some similar concerns, and by similar ways of thinking about the problems at hand. It will turn out that certain problems arising in association with both approaches are (perhaps best) avoided if the issues are viewed from the second perspective introduced in my title: preservationism.

4 Dialethist semantics

First, I wish to consider a dialethic approach to the problem of predicate-dissolution. Or rather, I want to consider two distinct approaches that differ philosophically, even as they fall under the same rubric. The first theory, dialethism proper, holds that there are in fact true contradictions (i.e. dialethias) present in the world, or that such a world is genuinely possible, if not actual. On this view, certain propositions may be both true and false at the same time; reality may be such as to satisfy both a proposition and its negation.
at once.\textsuperscript{22} The properly dialethic point of view, held most notably perhaps by Graham Priest, is admittedly controversial; still, Priest has offered many interesting and extended arguments for its coherent possibility, and it will not do just to dismiss these out of hand.\textsuperscript{23} So, certain propositions will be given both truth-values (or, what is the same thing, a third, “combination” value), and formal models of the situation require a different semantics than the usual, bivalent variety. Sentences involving “up,” for instance, may come out both true and false on our models of the flat-earth world, and this reflects the fact that some things both \textit{are} and \textit{are not} “up,” rather than some confusion about how best to apply the predicate. That is to say, the dual criteria \textit{both} correctly identify “up”—it is just that the predicate in question is \textit{ineradicably ambiguous}, actually applying to more than one distinct class of objects at one and the same time.\textsuperscript{24}

A related but distinct view might be called “pseudo-dialethism.” Rather than holding that the world actually is the site of true contradictions (or that such a thing is in fact possible), the pseudo-dialethist holds only that we must sometimes act as if dialethism were the case. That is, certain subjects and situations can only be sensibly handled by way of formal representations which model satisfaction on a dialethic, three-valued “glut” scheme.\textsuperscript{25} The pseudo-dialethist holds to this three-valued semantics even if such a thing as a “third truth-value” is thought nonactual, or even impossible. Such a view need not commit to the real existence of true contradictions, rather only insisting upon their coherence within a formal framework, and their dialectical necessity with respect to our (imperfect) understanding of certain concepts or situations. Things themselves may not be possibly both “up” and not “up” at once, but representations of our \textit{understanding} of things may need to treat them that way, or as if we thought they were that way. So, the pseudo-dialethist can use the same sorts of formal mechanics as can the genuine dialethist, but without the associated metaphysical claims. In fact, because we are concerned with interpretation here, and since the formal treatments \textit{are} equivalent, I will simply talk about “dialethic” treatments of predicates, meaning to refer to either sort of account interchangeably.

\textsuperscript{22}Assuming, of course, that “reality” is such a thing as \textit{can} make such as “propositions” true. This is only a manner of speaking. The point is not at issue here, and the reader may substitute their own favourite semantical talk.


\textsuperscript{24}A couple of things ought to be noted here. First, Priest also considers cases in which the ineradicable ambiguity of certain predications results not so much from the nature of the \textit{predicates} as from the nature of the \textit{objects themselves}. That is, Priest also works with first-order schemes on which object-names may refer to more than one object correctly (the eponymous amoeba of his “Multiple Denotation, Ambiguity, and the Case of the Missing Amoeba,” \textit{Logique et Analyse} \textbf{150–151–152} (1995): 361–373.). My own interests are limited here to the predicate-case, however, and I will not consider this thought. (For more on differences between what Priest is up to, and my own project, see Section 4.2, below.) Note also that, in conversation, Priest has stressed that it would misrepresent his view to claim that he thinks that \textit{all} true contradictions arise only in the context of the semantics of ambiguous terms. Rather, he thinks, it is (as) often the case that perfectly clear language still allows for the expression of true contradictions.

\textsuperscript{25}That is, a scheme which assigns to certain sentences to be both true and false, as opposed to those schemes which assign “gaps” where some sentences take no truth-value at all.
4.1 Formal mechanics of dialethism.

Perhaps the simplest way to represent the dialethic account of ineradicable predicate-ambiguity is by way of a semantics which treats truth-functions as taking sentences not to single truth-values, but rather to sets of truth-values. The two truth-values, True and False (T and F), are but elements of the value-set assigned to each sentence; every sentence has an associated set containing at least one, and perhaps both, of the truth-values. This framework has been worked out for the propositional-logical case in the simple paraconsistent logic LP.\(^{26}\) We begin with the assumption that LP possesses a standard and simple propositional logical grammar, with the connectives for “not,” “and,” “or,” and “if . . . then” (¬, ∧, ∨, and →). The semantics for LP are presented as follows, in terms of a valuation function V taking each sentence to a set of truth-values, letting \(V(α)\) be the set to which the function V takes the sentence α, and writing, for instance, \(T \in V(α)\).

**Definition 3 (LP semantics).** The function V is an LP-valuation iff for any sentence α, V takes α to one of the sets of values in \(\text{TrV} = \{\{T\}, \{F\}, \{T, F\}\}\). For any atomic (i.e. connectiveless) sentence α, \(V(α) \subseteq \text{TrV}\) and V deals with the connectives as follows:

\[
\neg : \quad T \in V(\neg α) \iff F \in V(α) \\
F \in V(\neg α) \iff T \in V(α)
\]

\[
\land : \quad T \in V(α \land β) \iff T \in V(α) \text{ AND } T \in V(β) \\
F \in V(α \land β) \iff F \in V(α) \text{ OR } F \in V(β)
\]

\[
\lor : \quad T \in V(α \lor β) \iff T \in V(α) \text{ OR } T \in V(β) \\
F \in V(α \lor β) \iff F \in V(α) \text{ AND } F \in V(β)
\]

\[
\rightarrow : \quad T \in V(α \rightarrow β) \iff F \in V(α) \text{ OR } T \in V(β) \\
F \in V(α \rightarrow β) \iff T \in V(α) \text{ AND } F \in V(β)
\]

In order to deal with cases of predicate-confusion, however, I want to extend the propositional case to cover predication and quantification. So, I here present a semantics for a first-order extension of LP, called here QLP, for “quantificational LP.” QLP takes over the basic connective grammar and semantics of LP, and employs as well a usual first-order grammar, with predicate-letters (F, G, . . . ), constant-letters (a, b,
Sentences in the grammar of QLP consist of open or closed atoms $F_{ab}, G_{xyz}, F_{acx}, \ldots$ and the usual sorts of compound sentences, involving combinations of the truth-functional connectives with the quantifiers for “all” and “some” ($\forall$ and $\exists$). Furthermore, QLP comes equipped with semantic models which include a non-empty domain of discourse, and an assignment-function $\mu$ such that for any constant-letter $a$ or variable-letter $x$ the assignments $\mu(a)$ and $\mu(x)$ are individual elements of the domain. This is all as usual.

The main difference from classical logic is that predicate-letters are not assigned single sets of tuples taken from the domain, but rather ordered pairs of sets. The typical semantics for classical first-order logic has the assignment-function $\mu$ taking every predicate-letter $F$ to a single set of ordered sequences of objects from the domain that are interpreted as having the property or standing in the relation named by the predicate-letter. The semantics for QLP, on the other hand, can assign more than one such set of sequences to a predicate-letter. That is, $\mu(F) = \langle \Gamma, \Delta \rangle$, where $\Gamma$ and $\Delta$ may or may not be one and the same set of tuples taken from the domain. The semantics for atomic sentences is adjusted in accord with this new kind of “split” extension, so that $T \in \mathcal{V}(Fa_1 \ldots a_n)$ just in case either set in $\mu(F)$ contains the tuple of corresponding elements of the domain $\langle \mu(a_1), \ldots, \mu(a_n) \rangle$ and $F \in \mathcal{V}(Fa_1 \ldots a_n)$ just in case either set in $\mu(F)$ does not contain that tuple.

The three possible truth-values—true only, false only, and both true and false—fall out accordingly. That is, the valuation-function $\mathcal{V}$ is such that for any atomic sentence $Fa_1 \ldots a_n$, $\mathcal{V}(Fa_1 \ldots a_n) = \{ T \}$ if and only if the tuple $\langle \mu(a_1), \ldots, \mu(a_n) \rangle$ is in the intersection of the two element-sets of $\mu(F) = \langle \Delta, \Gamma \rangle$; that is, just in case the tuple is in both $\Gamma$ and $\Delta$. Further, $\mathcal{V}(Fa_1 \ldots a_n) = \{ F \}$ if and only if that tuple is in neither $\Gamma$ nor $\Delta$. Lastly, $\mathcal{V}(Fa_1 \ldots a_n) = \{ T, F \}$ if and only if the tuple is in the disjoint union of $\Gamma$ and $\Delta$. That is, an atomic sentence gets mapped to the set containing both truth-values just in case the corresponding tuple is in exactly one of the two sets to which the predicate-letter is assigned. Note, however, that nothing dictates that any predicate $F$ must take a “dissolute” extension. The assignment may take any or all of the predicates in the language to pairs $\langle \Delta, \Gamma \rangle$, such that $\Delta = \Gamma$. Such “classicalized” predicates will then behave just as usual, figuring only in atomic sentences that are mapped to one of the unit value-sets $\{ T \}$ or $\{ F \}$.

---

27 There are no operators on terms in QLP. This is an inessential fact about the language.

28 Intending “dissolute” here to mean an extension which is divided into two non-identical, although perhaps overlapping, parts.
Definition 4 (QLP semantics). \( \mathfrak{M} = \langle D, \mu \rangle \) is a QLP-model iff there exists \( D, \mu \) such that:

1. The domain \( D \) is a non-empty set of objects.

2. The assignment \( \mu \) is a function with the set of variable-, constant-, and predicate-letters as its domain, and such that:
   
   (a) for any variable-letter \( x \), \( \mu(x) \in D \);
   
   (b) for any constant-letter \( a \), \( \mu(a) \in D \);
   
   (c) for any predicate letter \( F \), \( \mu(F) = \langle \Gamma, \Delta \rangle \) where:
   
   \( \Gamma \subseteq D^n; \Delta \subseteq D^n \), for some value of \( n \), where:

   \[ D^n = \{ \langle d_1, \ldots , d_n \rangle | \forall d_i, \, d_i \in D \}. \]

The valuation \( V \) is a function taking any QLP-model \( \mathfrak{M} \) and any sentence \( \alpha \) into the set of sets of truth-values: \( \text{TrV} = \{ \{ T \}, \{ F \}, \{ T, F \} \} \). For any sentence \( \alpha \), any QLP-model \( \mathfrak{M} \), the value of \( \alpha \) on \( \mathfrak{M} \), written \( V_{\mathfrak{M}}(\alpha) \), is an element-set out of \( \text{TrV} \). Since \( V_{\mathfrak{M}}(\alpha) \) is a set, it makes sense to write, for instance, \( T \in V_{\mathfrak{M}}(\alpha) \).

For any QLP-model \( \mathfrak{M} = \langle D, \mu \rangle \), \( V \) behaves as follows:

\[ \begin{align*}
\text{[At]}: \quad &T \in V_{\mathfrak{M}}(Fa_1 \ldots a_n) \iff \mu(F) = \langle \Gamma, \Delta \rangle \text{ and:} \\
&\text{either } \langle \mu(a_1), \ldots , \mu(a_n) \rangle \in \Gamma \text{ or } \langle \mu(a_1), \ldots , \mu(a_n) \rangle \in \Delta. \\
\text{[TF]}: \quad &V_{\mathfrak{M}}(\alpha) \text{ is as Definition 3, for the truth-functional connectives.} \\
\text{[∀]}: \quad &T \in V_{\mathfrak{M}}((\forall x)\alpha) \iff \forall d \in \text{Dom}, \forall \mathfrak{M}^* = \langle \text{Dom}^*, \mu[d/x] \rangle, \, T \in V_{\mathfrak{M}^*}(\alpha), \text{ where:} \\
&(i) \quad (\text{Dom} \in \mathfrak{M}^*) = (\text{Dom} \in \mathfrak{M}). \\
&(ii) \quad \mu[d/x](x) = d. \\
&(iii) \quad \text{Otherwise, } \mu[d/x] \in \mathfrak{M}^* \text{ is identical to } \mu \in \mathfrak{M}.
\end{align*} \]

That is, \( \mu[d/x] \) assigns the variable-letter \( x \) to the element \( d \) from the domain, and is in all other respects identical to \( \mu \).
F ∈ V_{M}(∀x)α iff ∃d ∈ Dom: ∃M∗ = (Dom, μ[d/x]): F ∈ V_{M∗}(α).

[∃]: T ∈ V_{M}(∀x)α iff ∃d ∈ Dom: ∃M∗ = (Dom, μ[d/x]): T ∈ V_{M∗}(α).
F ∈ V_{M}(∀x)α iff ∀d ∈ Dom, ∀M∗ = (Dom, μ[d/x]): F ∈ V_{M∗}(α).

The semantics for QLP has the obvious related conceptions of semantic satisfaction and entailment. We will use “satisfies” and “entails” on their own, but remember that when we use these terms in speaking about models of QLP, we are actually talking about satisfaction in QLP and entailment in QLP; we include a subscript with our turnstile notation to remind ourselves. A QLP model M = ⟨D, μ⟩ satisfies a sentence α, written M |=_{QLP} α, if and only if T ∈ V_{M}(α). Furthermore, M satisfies a set of sentences Σ, written M |=_{QLP} Σ, if and only if M satisfies every sentence α in Σ. That is to say, M satisfies α if and only if the function V takes every sentence in Σ to either the value-set “true” {T} or the value “both” {T, F} on that model. Finally, we say that a set Σ entails a sentence β, written Σ |=_{QLP} β, if and only if every model which satisfies Σ also satisfies β:

**Definition 5 (QLP satisfaction).** A QLP model M satisfies a sentence α iff:

(5.1) \quad T ∈ V_{M}(α)

That is:

(\exists M |=_{QLP} α) ⇔ (V_{M}(α) = \{T\} or V_{M}(α) = \{T, F\})

A QLP model M satisfies a set of sentences Σ iff:

(5.2) \quad ∀α ∈ Σ, M |=_{QLP} α

**Definition 6 (QLP entailment).** A set of sentences Σ entails a sentence α iff for any QLP model M,

M |=_{QLP} Σ ⇒ M |=_{QLP} α
This is not the place, nor the time, to explore the resultant conception of QLP-consequence. Note, however, that the resulting consequence-relation blocks the trivial consequences of sets like

\[ \Sigma = \{ Fa, \neg Fa \} \]

or even those containing an explicit contradiction, such as

\[ \Sigma' = \{ Fa \land \neg Fa \} \]

Such sets are now satisfied by any dialethic first-order model of QLP on which the assignment-function \( \mu \) takes the predicate-letter \( F \) to two distinct sets from the domain \( D \), and the constant-letter \( a \) is assigned to an element of \( D \) that is in the union, but \textit{not the intersection} of those two sets. That is, both \( \Sigma \) and \( \Sigma' \) are satisfied by any \( \mathfrak{M} = \langle D, \mu \rangle \) such that:

1. \( \mu(F) = \langle \Gamma, \Delta \rangle \) (\( \Gamma \subseteq D \), \( \Delta \subseteq D \)).
2. \( \Gamma \neq \Delta \).
3. \( \mu(a) \in \Gamma \Leftrightarrow \langle \mu(a) \rangle \notin \Delta \).

It is evident from consideration of Definition 4 that any such \( \mathfrak{M} \) satisfies both \( \Sigma \) and \( \Sigma' \), since:

1. \( V_{\mathfrak{M}}(Fa) = \{ T, F \} \).
2. \( V_{\mathfrak{M}}(\neg Fa) = \{ T, F \} \).
3. \( V_{\mathfrak{M}}(Fa \land \neg Fa) = \{ T, F \} \).

By treating the predicate “up” as taking a dual extension, the dialethic scheme can make sense of the dual-criteria case for Gupta’s flatlanders. “Up” has both meanings and both extensions. So, sentences featuring “up” can be both true and false at once, and the interpretive problem works itself out. A flatlander committed to beliefs involving “up” need not be interpreted as being committed to \textit{just anything}. Rather, they can be interpreted in accordance with the notion of QLP-entailment, and we may ascribe to them only those further commitments which follow on any QLP-evaluation of their current ones. The “confusion” with which we began turns out not to be so much confusion at all, but instead a form of irresolvable ambiguity or “fuzziness” in the flatlander version of “up”, something which we can deal with without needing to resolve it away.
4.2 A connection with Priest’s work.

Although originally intended as merely a hypothetical suggestion for a dialethist solution to the problem of interpreting dissolute predicates, this first-order extension of LP turns out to be just the solution Graham Priest himself suggests in his paper, “Multiple Denotation, Ambiguity, and the Case of the Missing Amoeba,” which came to my attention after I had already done the basic work on QLP. Although Priest’s presentation differs a little in detail, since he assigns to certain “ambiguous” predicates both an extension and an “anti-extension,” and because he is concerned not only with cases where predicates are ineradicably ambiguous, but also with the possibility of singular terms with ineradicably ambiguous referents (like the amoeba of the title), it is evident that the two presentations are equivalent with respect to satisfaction and entailment, in the ambiguous-predicate case.

In particular, Priest allows the extension and anti-extension of a predicate to overlap, so that certain first-order sentences take both values, just as in my own treatment of QLP. So, we can interdefine my semantic clauses and his. For any predicate $F$, if a QLP model $\mathcal{M}$ yields $\mu(F) = \langle \Gamma, \Delta \rangle$, then the Priestian extension of $F$ is just $(\Gamma \cup \Delta)$, and the Priestian anti-extension is just $D - (\Gamma \cap \Delta)$, the complement of the intersection of $\Gamma$ and $\Delta$ in the domain $D$. If this equivalence is respected, then all semantic consequences are the same, in either version of QLP. I note this fact, but do not prove it (as it is relatively obvious upon consideration of the two versions of the semantics). One nice feature of my own approach is that it makes more obvious the nature of the predicates under analysis. Although Priest is also concerned with dual-criterial predicates, his semantic scheme does not make this feature obvious. The extension of a predicate on Priest’s scheme consists of both criteria “lumped together,” without distinguishing between them. Further, the anti-extension is a less familiar concept than usual in first-order logic, since it may overlap the extension to any extent (from none at all to completely), and is not just the usual notion, the classical complement of the extension. On my scheme, however, each criterion of application is identified by a separate set assigned to a predicate, and the classical complement of the extension is still where first-order sentences are just false. Furthermore, my scheme allows for obvious generalizations of the dual-criterial case. Dialethic models of the QLP kind can be generated just as easily for cases with three or more criteria for the application of predicates; that is, we can just as easily generate models such that:

---

\[ \mu(F) = \langle \Gamma, \Delta, \Theta \rangle \]

or:

\[ \mu(F) = \langle \Gamma, \Delta, \Theta, \Lambda \rangle \]

and further, for larger tuples of extension-sets.\(^{30}\)

5 Disambiguation-semantics.

I want to consider another account of the apparent dissolution of predicates, one seemingly at odds with dialethism. Rather than suggesting that a predicate like Gupta’s “up” is irresolvably ambiguous, one might think, instead, that the predicate-concept in question needs to be disambiguated; where confusing ambiguity arises, we need to re-draw the boundaries of the predicate-concept. We may want to extend or diminish its extension; in other cases, perhaps, the best thing to do will be to divide the extension into separate parts, and divide the predicate, too. That is, we will sometimes want to change our predicate-concept, replacing the original with two or more others, each with its own separate extension and application. Of course, characterizing the circumstances under which we will want to do such a thing, and the ways in which we go about it when we do, would be a difficult task. In all likelihood, no systematic account can be given of just how such messy conceptual re-adjustment proceeds.

All the same, someone resistant to dialethism is sure to think that situations like that of Gupta’s flat-landers call for some form of conceptual re-configuration. Barring the possibility of truly contradictory applications of a predicate-concept like “up,” use of that predicate-concept must change. Of course, one thing to do would be to dispense with either, or both, of the uses of such a fractured and confusing concept. That is as may be, but at least one other potential solution presents itself: noting that there are two separate uses, and taking these as distinguishing two separate predicate-concepts. Whether we still want to call one in particular, or neither, of these new applications “up” is less important than that we can disambiguate the original predicate, making it possible to satisfy previously contradictory sets of claims, beliefs, or commitments.

\(^{30}\)Note again that in any such scheme, if the assignment function \( \mu \) takes every predicate to some \( m \)-tuple of subsets of \( D^n \), then predicates can be represented with a number of criteria of application less than or equal to \( m \). So, some or all of the predicates on that scheme can behave in a less-dissolute, or even purely classical manner (just as any QLP model \( \mathcal{M} \) allows there to be a combination of purely classical and dual-criteria predicates).
To suggest how this sort of disambiguating impulse might proceed, I wish to present a first-order semantic scheme (called QLD for “quantificational logic of disambiguation”), which relies upon the idea that an assignment-function gives an “interpretation” or “reading” of a predicate, and that we can interpret an ambiguous predicate (dual-criterial or otherwise) by considering it under both its readings. Again, I will present a semantic scheme which is only meant to be hypothetical. On the way of looking at things which I have in mind, predicates around which ambiguity or confusion circulate need to be considered under various possible “aspects,” and the satisfaction of a set containing occurrences of such predicates may require that various instances of a single predicate be re-interpreted as instances of two distinct predicates.

5.1 Formal mechanics of disambiguation.

The thought here is that a first-order model $\mathcal{M}_{QLD}$ for the logic of disambiguation will require, in addition to the usual domain, two separate assignment-functions, $\mu_1(\ldots)$ and $\mu_2(\ldots)$, which must agree on all variable- and constant-letter assignments, but need not agree with respect to predicate-letter assignments, allowing a single predicate to be considered under two different assignments (as if, in fact, ambiguously standing in for two distinct predicates). The valuation-function for $\mathcal{M}_{QLD}$ will reflect this possibility, determining the truth-value of sentences bivalently, but under the understanding that some sentences may only be satisfiable under some disambiguation. More formally:

Definition 7 (QLD semantics). Again, let us have a first-order language:

1. Atomic sentences of the form ‘$Fa_1 \ldots a_n$’, ‘$Ga_1 \ldots a_n$’, etc.

2. Compound sentences of the usual forms.

Let $\mathcal{M}_{QLD} = \langle D, \mu_1, \mu_2, V \rangle$, where:

1. the domain $D$ is a non-empty set of objects:

2. the assignments $\mu_1$ and $\mu_2$ are as follows:

   (a) for any variable-letter ‘$x$’, $\mu_1(x) \in D$, $\mu_2(x) \in D$, $\mu_1(x) = \mu_2(x)$;

---

31 Unlike in the dialethic case, I know of no-one who has actually floated a scheme like the following—it is only meant, again, as a proposal which someone with these sorts of background sympathies might find amenable.

32 As throughout, I am only considering here cases in which the extension of a predicate-concept divides into two distinct sets. As for QLP, generalizations will be obvious for the disambiguating models which I present (see section 4.2). Note, however, that the possibility of predicates which divide into more than two separate extensions raises another interesting avenue for the preservationist: the preservation not only of the number of predicate-concepts around which confusion circulates, but also the “degree of fracture,” measured in terms of number of separate extensions to which a predicate-concept need be apportioned. No more can be said about this suggestion here, however, as it is only a suggestion for further thought, and not something I have worked out in any real detail.
(b) for any constant-letter ‘$a_i$’, $\mu_1(a_i) \in D, \mu_2(a_i) \in D, \mu_1(a_i) = \mu_2(a_i)$;

(c) for any predicate letter, ‘$F$’, $\mu_1(F) = \Gamma \subseteq D^n, \mu_2(F) = \Delta \subseteq D^n$, for some number $n$.

NB: For clause 2c, it is important that the number $n$ be the one and same in both cases.

3. the valuation $V$ is as follows:

$[\text{At}]$: $V(Fa_1 \ldots a_n) = T$ iff $\exists i: \exists \Lambda: \mu_i(F) = \Lambda$ and:

$\langle \mu_i(a_1), \ldots, \mu_i(a_n) \rangle \in \Lambda$;

$V(Fa_1 \ldots a_n) = F$ iff $\forall i, \forall \Lambda, \mu_i(F) = \Lambda$ then:

$\langle \mu_i(a_1), \ldots, \mu_i(a_n) \rangle /\notin \Lambda$;

$[\neg]$: $V(\neg \alpha) = T$ iff $\alpha$ is interpretable-false on $\mathcal{M}_{QLD}$ (see definition 8 below);

$V(\neg \alpha) = F$ iff $\alpha$ is not interpretable-false on $\mathcal{M}_{QLD}$.

$[\land]$: $V(\alpha \land \beta) = T$ iff $V(\alpha) = T$ AND $V(\beta) = T$;

$V(\alpha \land \beta) = F$ iff $V(\alpha) = F$ OR $V(\beta) = F$.

$[\lor]$: $V(\alpha \lor \beta) = T$ iff $V(\alpha) = T$ OR $V(\beta) = T$;

$V(\alpha \lor \beta) = F$ iff $V(\alpha) = F$ AND $V(\beta) = F$.

$[\rightarrow]$: $V(\alpha \rightarrow \beta) = T$ iff $\alpha$ is interpretable-false on $\mathcal{M}_{QLD}$ OR $V(\beta) = T$;

$V(\alpha \rightarrow \beta) = F$ iff $V(\alpha)$ is not interpretable-false on $\mathcal{M}_{QLD}$ AND $V(\beta) = F$.

$[\forall]$: $V(\forall x \alpha) = T$ iff $\forall a_i: V([x/a_i]\alpha) = T$;

$V(\forall x \alpha) = F$ iff $\exists a_i: V([x/a_i]\alpha) = F$.

$[\exists]$: $V(\exists x \alpha) = T$ iff $\exists a_i: V([x/a_i]\alpha) = T$.

$V(\exists x \alpha) = F$ iff $\forall a_i, V([x/a_i]\alpha) = F$.

So, on this semantics, which is bivalent, the clauses for $V(\ldots)$ are just the same as for the usual classical first-order logic, excepting the cases for atoms and negation. In the atomic case, the semantics makes an
atomic sentence true if and only if at least one of the assignments satisfies it. In the case of negation, we rely upon the notion ‘interpretable false’, about which more need be said. To say that a sentence α is interpretable false is not to say that it is in fact false on the present model. After all, that would just yield classical negation, and the set \{ Fa_1 \ldots a_n, ¬ Fa_1 \ldots a_n \} would never be satisfiable on the resulting semantic scheme (which is the whole point of providing models which can disambiguate the predicate-concept ‘F’ in the first place). What we need—and what the idea of ‘satisfiable under some disambiguation’ suggests—is a conception on which ‘¬ Fa_1 \ldots a_n’ can be true if and only if we manage to disambiguate ‘F’ in such a way that we can make ‘Fa_1 \ldots a_n’ false. Conversely, ‘¬ Fa_1 \ldots a_n’ should be false if and only if we do not disambiguate ‘F’ so to make ‘Fa_1 \ldots a_n’ even potentially false (i.e., ‘Fa_1 \ldots a_n’ is true under all our readings). That is, a model satisfies ‘Fa_1 \ldots a_n’ and ‘¬ Fa_1 \ldots a_n’ by disambiguating ‘F’ in different ways as it appears in the two sentences. To this end, I define the required idea:

**Definition 8 (Interpretable-F alse).** ∀α, α is interpretable-false on \( M_{QLD} = (D, \mu_1, \mu_2, V) \) (written \( I-F[α]_{M_{QLD}} \), suppressing the subscript where convenient) iff:

\[
\begin{align*}
[\text{At}]: & \quad α = Fa_1 \ldots a_n \text{ (some } F, \text{ some } (a_1, \ldots, a_n)):
  & \quad I-F[α] \text{ iff } \exists \mu_i: \exists Λ: \mu_i(F) = Λ & (μ_i(a_1), \ldots, μ_i(a_n)) \notin Λ. \\

[\neg]: & \quad α = ¬ β \text{ (some } β ):
  & \quad I-F[α] \text{ iff } V(β) = T. \\

[\land]: & \quad α = β \land γ \text{ (some } β, γ).
  & \quad I-F[α] \text{ iff } I-F[β] \text{ OR } I-F[γ]. \\

[\lor]: & \quad α = β \lor γ \text{ (some } β, γ).
  & \quad I-F[α] \text{ iff } I-F[β] \text{ AND } I-F[γ]. \\

[\rightarrow]: & \quad α = β \rightarrow γ \text{ (some } β, γ).
  & \quad I-F[α] \text{ iff } V(β) = T \text{ AND } I-F[γ]. \\

[\forall]: & \quad α = (\forall x)β \text{ (some } x, β).
  & \quad I-F[α] \text{ iff } ∃a_i: I-F[\{x/a_i\}\{β\}].
\end{align*}
\]
\[\exists: \quad \alpha = (\exists x) \beta \text{ (some } x, \beta\).
\]

1-F[\alpha] \iff \forall a_i, 1-F[x/a_i,\beta].

Upon consideration, two facts become evident. First, the resulting definition of negation for QLD gives us the familiar interdefinability between connectives and quantifiers such that, say, ‘(\forall x)\alpha’ could be defined as ‘\neg(\exists x)\neg\alpha’, or vice-versa. Second, genuine falsity on a model for QLD implies interpretable-falsity. That is:

**Fact 1 (Falsity implies interpretable-false).** If any sentence \(\alpha\) is actually false on some \(M_{QLD}\), then \(\alpha\) is interpretable-false on \(M_{QLD}\). That is:

\[
\forall \alpha, \forall M_{QLD} = \langle D, \mu_1, \mu_2, V \rangle, \\
V(\alpha) = F \Rightarrow i-F[\alpha]_{M_{QLD}}.
\]

**Proof.** The Fact follows by induction on the degree of sentences (the number of connectives and/or quantifiers occurring in a sentence), and is evident upon consideration of definitions.

For consider the case where our sentence \(\alpha\) is of degree-zero—that is, \(\alpha = F a_1 \ldots a_n\). Suppose \(\alpha\) is false on some \(M_{QLD}\). Then, by Definition 7,

For \(M_{QLD} = \langle D, \mu_1, \mu_2, V \rangle, \\
\forall i, \forall \Lambda, \mu_i(F) = \Lambda \text{ then:} \\
\langle \mu_i(a_1), \ldots, \mu_i(a_n) \rangle \notin \Lambda.
\]

So, clearly, for \(M_{QLD}\),

\[
\exists \mu_i: \exists \Delta: \mu_i(F) = \Delta \& (\mu_i(a_1), \ldots, \mu_i(a_n) \notin \Delta.
\]

And, by Definition 8, \(i-F[\alpha]_{M_{QLD}}\).

Now, for the induction, suppose that Fact 1 holds for all sentences with degree less than some \(k\). Suppose \(\alpha\) has degree equal to \(k\). There are a number of cases:

(1) \(\alpha = \neg \beta\). Suppose \(\alpha\) is false on \(M_{QLD}\). Then, by Definition 7, \(\text{not } i-F[\beta]_{M_{QLD}}\). But \(\beta\) is of degree less than \(k\), and by the hypothesis of the induction, Fact 1 is true of \(\beta\), and, by contraposition, we know that \(\beta\) is true on \(M_{QLD}\). That is, \(V(\beta) = T\). So, by Definition 8, \(i-F[\alpha]_{M_{QLD}}\).
(2) $\alpha = \beta \land \gamma$. Suppose $\alpha$ is false on $\mathcal{M}_{QLD}$. Then by Definition 7, either $\beta$ is false on $\mathcal{M}_{QLD}$, or $\gamma$ is false on $\mathcal{M}_{QLD}$. But $\beta$ and $\gamma$ are both of degree less than $k$, and by the hypothesis of the induction, Fact 1 is true of $\beta$ and $\gamma$. So, we know that $i-F[\beta]$ or $i-F[\gamma]$. So, by Definition 8, $i-F[\alpha]_{M_{QLD}}$.

(3)–(6) The cases for $\lor$, $\forall$, and $\exists$ are all much the same as for $\land$, and I leave them out. Similarly, the case for $\to$ resembles the case for $\neg$, and is left out.

The Fact follows by induction on cases.}

Like QLP, QLD defines semantic satisfaction and entailment pretty much as usual. A QLD model, $\mathcal{M}_{QLD} = \langle D, \mu_1, \mu_2, V \rangle$, QLD-satisfies a first-order sentence $\alpha$ if and only if $V(\alpha) = T$, and QLD-satisfies a set of sentences $\Sigma$ if and only if it satisfies every sentence $\alpha$ in $\Sigma$. Further, we say that a set $\Sigma$ QLD-entails a sentence $\beta$ if and only if every model which QLD-satisfies $\Sigma$ also QLD-satisfies $\beta$. Formally:

**Definition 9 (QLD satisfaction).** A QLD model $\mathcal{M}_{QLD}$ QLD-satisfies a sentence $\alpha$ ($\mathcal{M}_{QLD} \models_{QLD} \alpha$) iff:

$$(9.1) \quad \mathcal{M}_{QLD} = \langle D, \mu_1, \mu_2, V \rangle \text{ and } T \in V(\alpha)$$

A QLD model $\mathcal{M}_{QLD}$ QLD-satisfies a set of sentences $\Sigma$ ($\mathcal{M}_{QLD} \models_{QLD} \Sigma$) iff:

$$(9.2) \quad \forall \alpha \in \Sigma, \mathcal{M}_{QLD} \models_{QLD} \alpha$$

**Definition 10 (QLD entailment).** A set of sentences $\Sigma$ QLD-entails a sentence $\alpha$ ($\Sigma \models_{QLD} \alpha$) iff:

$$\forall \mathcal{M}_{QLD}, \mathcal{M}_{QLD} \models_{QLD} \Sigma \Rightarrow \mathcal{M}_{QLD} \models_{QLD} \alpha$$

Again, it is not my purpose to go into the details of QLD-consequence here. It is noteworthy, however, that the disambiguating semantics, like the dialethic scheme, can equally well satisfy both those sets which are contradictory in toto, like

$$\{ Fa_1 \ldots a_n, \neg Fa_1 \ldots a_n \},$$

and those sets which contain self-contradictory sentences, like

$$\{ Fa_1 \ldots a_n \land \neg Fa_1 \ldots a_n \}.$$
Furthermore, the two approaches are the same in that they can satisfy both sorts of sets with a single model—in this case, on any disambiguating first-order model \( \mathcal{M}_{QLD} = \langle D, \mu_1, \mu_2, V \rangle \), on which the assignments \( \mu_1(\ldots) \) and \( \mu_2(\ldots) \) take the predicate-letter ‘\( F \)’ to two distinct sets of tuples from the domain \( D \), with the tuple \( \langle \mu(a_1), \ldots, \mu(a_n) \rangle \) in one, but not both, of those two sets.\(^{34} \) That is:

1. \( \mu(F) = \Gamma \subseteq D^n; \quad \mu_2(F) = \Delta \subseteq D^n \);
2. \( \Gamma \neq \Delta \);
3. \( \langle \mu(a_1), \ldots, \mu(a_n) \rangle \in \Gamma \Leftrightarrow \langle \mu(a_1), \ldots, \mu(a_n) \rangle \notin \Delta \);
4. \( V(F_{a_1 \ldots a_n}) = T \),
   \( V(\neg F_{a_1 \ldots a_n}) = T \) (because \( 1\neg[F_{a_1 \ldots a_n}] \)),
   \( V(F_{a_1 \ldots a_n} \land \neg F_{a_1 \ldots a_n}) = T \).

So, on any such model, ‘\( F_{a_1 \ldots a_n} \)’ is satisfiable because the assignments take the tuple of constant-letters \( \langle a_1, \ldots, a_n \rangle \) to one of the disambiguations of ‘\( F \)’, and ‘\( \neg F_{a_1 \ldots a_n} \)’ is satisfiable because the assignments take the tuple outside of the other disambiguation of ‘\( F \)’. On such a model, obviously, the set containing both sentences will be satisfiable. Furthermore, the model will also satisfy the set containing the conjunction of the two sentences, since satisfaction for conjunctions in \( \text{QLD} \) is just as for classical logic—namely, conjunctions are satisfied just in case both conjuncts are satisfied.

### 5.2 A note on supervaluations.

The \( \text{QLD} \) scheme, which is similar to more-familiar supervaluational treatments of ambiguous and vague predicates, differs from these in one important respect. On some accounts of ambiguity, the semantics given involves a supervaluational model that evaluates any sentence against two or more separate classical first-order models (‘worlds’) and that makes a sentence true if and only if it is satisfied on at least one of these models. Such a treatment can satisfy and block trivial consequence from a set like \{ \( F_{a_1 \ldots a_n}, \neg F_{a_1 \ldots a_n} \) \} by providing two separate models, one validating ‘\( F_{a_1 \ldots a_n} \)’, and the other validating ‘\( \neg F_{a_1 \ldots a_n} \)’. But sets containing explicit contradictions, such as \{ \( F_{a_1 \ldots a_n} \land \neg F_{a_1 \ldots a_n} \) \} cannot be satisfied on such a supervaluational scheme, since no classical model satisfies any single self-contradictory sentence.

There is no reason that supervaluational treatments need to fail where \( \text{QLD} \) succeeds, however. That is, if \( \text{QLD} \) can satisfy sets containing explicit contradictions, then so can schemes on which sentences are evaluated

\(^{34}\)Suppressing subscripts for \( \mu(a_i) \), because of the fact that \( \forall i, \mu_1(a_i) = \mu_2(a_i) \).
against two distinct classical first-order models. Priest has criticized such as Hartry Field’s attempt\textsuperscript{35} to deal with ambiguity by way of a form of supervaluational semantics. Priest argues that, while it is true that his own, dialethist semantics “will make certain sentences true that are not true on any single disambiguation,” this feature is less a deficit than a benefit, since “even when ambiguity is generated by ambiguous senses, we are quite happy to accept a sentence as true when different tokens of the same word have to be disambiguated differently. Consider: When playing his shot, the cricketer missed the ball, but hit the bat that flew past him with his bat.”\textsuperscript{36} It is true that the single-sense-per-sentence approach to disambiguation fails to sort out pretty commonplace sentences, but this failure has less to do with any intrinsic feature of the supervaluational approach than it does a certain short-sightedness with respect to the semantic clauses operative on such models. Although I will not prove the two approaches are not strictly interdefinable here, it is evident upon a little inspection that for any $M\text{QLD} = \langle D, \mu_1, \mu_2, V \rangle$ an equivalent supervaluational model can be defined which satisfies exactly the same first-order sentences. To see this, start with the single model $M\text{QLD}$, with the single domain $D$ and the pair of assignment-functions $\mu_1(\ldots)$ and $\mu_2(\ldots)$, and construct two classical first-order models, $M_a = \langle D_a, \mu_a, V_a \rangle$ and $M_b = \langle D_b, \mu_b, V_b \rangle$, as follows:

(1) $D = D_a = D_b$;
(2) $\forall a_i, \mu_1(a_i) = \mu_2(a_i) = \mu_a(a_i) = \mu_b(a_i)$.
(3) $\forall x_i, \mu_1(x_i) = \mu_2(x_i) = \mu_a(x_i) = \mu_b(x_i)$.
(4) $\forall F, \mu_1(F) = \mu_a(F) \& \mu_2(F) = \mu_b(F)$.

On such defined models, as long as we allow our supervaluational scheme to employ a notion like ‘interpretable false’, and employ a definition of the valuation-function for the supervaluational logic which is analogous to the one given for QLD—adjusting the clause for atoms to track $\mu_a(\ldots)$ and $\mu_b(\ldots)$, rather than $\mu_1(\ldots)$ and $\mu_2(\ldots)$—it is evident that exactly the same theory of consequence and satisfaction is at work. So, predicate-letters can be multiply disambiguated within the same sentence, and even explicit contradictions can be interpreted and satisfied.

\textsuperscript{36}Priest (1995), 369.
6 Some equivalencies.

QLP and QLD are not only alike in that each happens to allow for models which satisfy both

\{ Fa, \neg Fa \}\n
and

\{ Fa \land \neg Fa \},

(see the discussion at pages 21 and 28 for QLP and QLD, respectively). In fact, for any model of one system, a corresponding model of the other system is available which satisfies just the same sentences. That is to say, despite their quite different motivations, and somewhat different presentations, the two systems are capable of satisfying all the same sets of sentences. From this, it follows that \( \vdash_{QLP} \) and \( \vdash_{QLD} \) are equivalent with respect to consequence. I will sketch the details of the proofs. As it turns out, the semantics for each system make the task relatively straightforward. Before showing that the two consequence-relations are extensionally equivalent, I prepare the way by means of a pair of lemmas.

**Lemma 1 (Model-equivalence One: QLP \Rightarrow QLD).** For any QLP-model, some QLD-model satisfies just the same sentences:

\[
\forall M = (D_{QLP}, \mu_{QLP}, V_{QLP}),
\exists M_{QLD} = (D_{QLD}, \mu_{QLD1}, \mu_{QLD2}, V_{QLD}):
\forall \alpha, M_{QLD} \models_{QLD} \alpha \Leftrightarrow M_{QLP} \models_{QLP} \alpha.
\]

**Proof.** I prove the lemma for arbitrary \( M_{QLP} \), and appropriate \( M_{QLD} \), by induction upon degree of \( \alpha \).\(^{37}\)

Consider an arbitrary \( M_{QLP} = (D_{QLP}, \mu_{QLP}, V_{QLP}) \).

Define a corresponding QLD-model: \( M_{QLD} = (D_{QLD}, \mu_{QLD1}, \mu_{QLD2}, V_{QLD}) \) such that:

1. \( D_{QLD} = D_{QLP} \).

2. the assignments \( \mu_{QLD1}(...) \) and \( \mu_{QLD2}(...) \) are as follows:

\[(a) \text{ for any variable-letter '} x_i \text{'}, \mu_{QLD1}(x_i) = \mu_{QLD2}(x_i) = \mu_{QLP}(x_i); \]

\(^{37}\text{Recall that the degree of the sentence is the number of connectives and/or quantifiers occurring in the sentence. Atoms have degree zero (0).}\)
(b) for any constant-letter ‘a’, \( \mu_{QLD1}(a_i) = \mu_{QLD2}(a_i) = \mu_{QLP}(a_i) \);

c) for any predicate-letter ‘F’, \( \mu_{QLP}(F) = \langle \Gamma, \Delta \rangle \Rightarrow (\mu_{QLD1}(F) = \Gamma & \mu_{QLD2}(F) = \Delta) \).

To prove:

\[ \forall \alpha, M_{QLD} \models_{QLD} \alpha \iff M_{QLP} \models_{QLP} \alpha. \]

That is, \( V_{QLD}(\alpha) = T \iff T \in V_{QLP}(\alpha) \).

**Basis Case**: \( \alpha = Fa_1 \ldots a_n \) (some \( F \), some \( \langle a_1, \ldots, a_n \rangle \)).

By Definition 4, \( T \in V_{QLP}(Fa_1 \ldots a_n) \) iff \( \mu_{QLP}(F) = \langle \Gamma, \Delta \rangle \) and:

(i) \( \langle \mu_{QLP}(a_1), \ldots, \mu_{QLP}(a_n) \rangle \in \Gamma \),

or:

(ii) \( \langle \mu_{QLP}(a_1), \ldots, \mu_{QLP}(a_n) \rangle \in \Delta \).

So, by definition of \( M_{QLD} \), \( \mu_{QLD1}(F) = \Gamma \) and \( \mu_{QLD2}(F) = \Delta \). Furthermore,

\[ \langle \mu_{QLP}(a_1), \ldots, \mu_{QLP}(a_n) \rangle = \langle \mu_{QLD1}(a_1), \ldots, \mu_{QLD2}(a_n) \rangle = \langle \mu_{QLD2}(a_1), \ldots, \mu_{QLD2}(a_n) \rangle. \]

So:

(i) \( \langle \mu_{QLD1}(a_1), \ldots, \mu_{QLD1}(a_n) \rangle \in \Gamma \),

or:

(ii) \( \langle \mu_{QLD2}(a_1), \ldots, \mu_{QLD2}(a_n) \rangle \in \Delta \).

By Definition 7, \( V_{QLD}(Fa_1 \ldots a_n) = T \) iff:

\[ \exists \Lambda: \exists \Lambda: \mu_{QLD1}(F) = \Lambda \text{ and } \langle \mu_{QLD1}(a_1), \ldots, \mu_{QLD1}(a_n) \rangle \in \Lambda. \]

Therefore, \( V_{QLD}(Fa_1 \ldots a_n) = T \iff T \in V_{QLP}(Fa_1 \ldots a_n) \).

Therefore, \( V_{QLD}(\alpha) = T \iff T \in V_{QLP}(\alpha) \). (Conclusion of the Basis.)

**Hypothesis of the Induction**: Assume that for all sentences \( \beta \) of degree less than some \( k \),

\[ V_{QLD}(\beta) = T \iff T \in V_{QLP}(\beta). \]

**Inductive Case**: Assume that \( \alpha \) has degree equal to \( k \). There are several subcases:
\[\neg \alpha = \neg \beta \text{ (some } \beta) \]. The case has several subcases as well:

**[Subcase (i) for \(\neg\)]** \(\alpha = \neg \beta\), where \(\beta = Fa_1 \ldots a_n \text{ (some } F, \text{ some } \langle a_1, \ldots, a_n \rangle)\).

By Definition 4, \(T \in V_{QLP}(\neg Fa_1 \ldots a_n)\) iff \(F \in V_{QLP}(Fa_1 \ldots a_n)\).

That is, \(T \in V_{QLP}(\neg Fa_1 \ldots a_n)\) iff \(\mu_{QLP}(F) = \langle \Gamma, \Delta \rangle\) and:

(i) \(\langle \mu_{QLP}(a_1), \ldots, \mu_{QLP}(a_n) \rangle \notin \Gamma\),

or:

(ii) \(\langle \mu_{QLP}(a_1), \ldots, \mu_{QLP}(a_n) \rangle \notin \Delta\).

So, by definition of \(M_{QLD}\), \(\mu_{QLD_1}(F) = \Gamma\) and \(\mu_{QLD_2}(F) = \Delta\). Furthermore,

\[\langle \mu_{QLP}(a_1), \ldots, \mu_{QLP}(a_n) \rangle = \langle \mu_{QLD_1}(a_1), \ldots, \mu_{QLD_1}(a_n) \rangle = \langle \mu_{QLD_2}(a_1), \ldots, \mu_{QLD_2}(a_n) \rangle.\]

So:

(i) \(\langle \mu_{QLD_1}(a_1), \ldots, \mu_{QLD_1}(a_n) \rangle \notin \Gamma\),

or:

(ii) \(\langle \mu_{QLD_2}(a_1), \ldots, \mu_{QLD_2}(a_n) \rangle \notin \Delta\).

By Definition 8, \(i-f[Fa_1 \ldots a_n]\) iff:

\[\exists i: \exists \Lambda: \mu_{QLD_i}(F) = \Lambda \text{ and } \langle \mu_{QLD_i}(a_1), \ldots, \mu_{QLD_i}(a_n) \rangle \notin \Lambda.\]

Therefore, \(i-f[Fa_1 \ldots a_n] \Leftrightarrow T \in V_{QLP}(\neg Fa_1 \ldots a_n)\).

And, by Definition 7, \(V_{QLD}(\neg Fa_1 \ldots a_n) = T \Leftrightarrow i-F[Fa_1 \ldots a_n]\).

So, \(V_{QLD}(\neg Fa_1 \ldots a_n) = T \Leftrightarrow T \in V_{QLP}(\neg Fa_1 \ldots a_n)\).

Therefore, \(V_{QLD}(\alpha) = T \Leftrightarrow T \in V_{QLP}(\alpha)\). \hspace{1cm} (Conclusion of Subcase (i) for negation.)

**[Subcase (ii) for \(\neg\)]** \(\alpha = \neg \beta\), where \(\beta = \neg \gamma \text{ (some } \gamma)\).

By Definition 4, \(T \in V_{QLP}(\neg \neg \gamma)\) iff \(F \in V_{QLP}(\neg \gamma)\).

That is, \(T \in V_{QLP}(\neg \neg \gamma)\) iff \(T \in V_{QLP}(\gamma)\).

Of course, \(\gamma\) is of degree less than \(k\). Therefore, by the Hypothesis of Induction,

\[V_{QLD}(\gamma) = T \Leftrightarrow T \in V_{QLP}(\gamma).\]
By Definition 8, $1F[\neg\gamma]$ iff $V_{QLD}(\gamma) = T$.

And, by Definition 7, $V_{QLD}(\neg\neg\gamma) = T$ iff $1F[\neg\gamma]$.

So, $V_{QLD}(\neg\neg\gamma) = T$ iff $\neg\neg\gamma$.

Therefore, $V_{QLD}(\alpha) = T$ iff $T \in V_{QLP}(\neg\neg\gamma)$.

(Conclusion of Subcase (ii) for negation.)

A subinduction. Unfortunately, the remaining subcases for negation require a slightly more complicated, and indirect, approach. Consider for instance the case in which $\alpha = \neg(\gamma \land \delta)$. If we assume that $T \in V_{QLP}(\neg(\gamma \land \delta))$, we get that $F \in V_{QLP}(\gamma)$ or $F \in V_{QLP}(\delta)$. But the Hypothesis of Induction does not then tell us that $V_{QLD}(\gamma) = F$ or $V_{QLD}(\delta) = F$. This is as it ought to be, really—since $QLP$ allows a sentence to take both ‘T’ and ‘F’ into its evaluation-set, the presence of ‘F’ is no guarantee that ‘T’ is not also present. To show that $V_{QLD}(\neg(\gamma \land \delta)) = T$, we must show that $(\gamma \land \delta)$ is interpretable-false on $M_{QLD}$.

And to show that, I must first prove the following:

$$\forall \beta, \text{if degree of } \beta \text{ is less than } k \Rightarrow (\text{NOT } 1F[\beta]_{M_{QLD}} \Leftrightarrow F \notin V_{QLP}(\beta)).$$

That is, if $\beta$ is not interpretable-false on $M_{QLD}$, then $V_{QLP}(\beta)$ does not contain ‘F’ at all ($V_{QLP}(\beta) = \{T\}$).

This subproof is properly a subinduction, as it relies in parts on the main Hypothesis of Induction (specifically, in the Subinductive Cases for $\neg$ and $\rightarrow$), and so must fall within its scope. Recall that $\beta$ is itself of degree less than $k$, since $\alpha = \neg\beta$, and the degree of $\alpha$ is $k$. It is this fact which ensures that the Main Hypothesis of Induction will apply to $\beta$ (and to its subformulae).

The subproof is also by induction on degree of $\beta$.

**Basis Case for Subinduction:** $\beta = Fa_1 \ldots a_n$ (some $F$, some $\langle a_1, \ldots, a_n \rangle$).

By Definition 8, NOT $1F[Fa_1 \ldots a_n]_{M_{QLD}}$ iff:

$$\forall i, \forall \Lambda, \mu_{QLD}(F) = \Lambda \Rightarrow \langle \mu_{QLD}(a_1), \ldots, \mu_{QLD}(a_n) \rangle \in \Lambda.$$

By definition of $M_{QLD}$ for the main proof, we know that

$$\mu_{QLP}(F) = \langle \mu_{QLD}(F), \mu_{QLD}(F) \rangle,$$

which is to say that

1. $\exists \Gamma: \mu_{QLD}(F) = \Gamma;$
2. ∃Δ: \( \mu_{QLD_1}(F) = \Delta \):

3. \( \mu_{QLP}(F) = (\Gamma, \Delta) \).

And we know that

\[
\langle \mu_{QLP}(a_1), \ldots, \mu_{QLP}(a_n) \rangle = \langle \mu_{QLD_1}(a_1), \ldots, \mu_{QLD_1}(a_n) \rangle = \langle \mu_{QLD_2}(a_1), \ldots, \mu_{QLD_2}(a_n) \rangle.
\]

So: (i) \( \langle \mu_{QLP}(a_1), \ldots, \mu_{QLP}(a_n) \rangle \in \Gamma \), AND:

(ii) \( \langle \mu_{QLP}(a_1), \ldots, \mu_{QLP}(a_n) \rangle \in \Delta \).

By Definition 4, \( F \in V_{QLP}(Fa_1 \ldots a_n) \) iff \( \mu_{QLP}(F) = (\Gamma, \Delta) \) and:

(i) \( \langle \mu_{QLP}(a_1), \ldots, \mu_{QLP}(a_n) \rangle \notin \Gamma \)

OR:

(ii) \( \langle \mu_{QLP}(a_1), \ldots, \mu_{QLP}(a_n) \rangle \notin \Delta \).

Therefore, NOT 1-F[\( Fa_1 \ldots a_n \)] \( M_{QLD} \) \( \iff \) \( F \notin V_{QLP}(Fa_1 \ldots a_n) \).

Therefore, NOT 1-F[\( \beta \)] \( M_{QLD} \) \( \iff \) \( F \notin V_{QLP}(\beta) \). \( \text{(Conclusion of the Basis for Subinduction.)} \)

**Hypothesis of the Subinduction:** Assume that for all sentences \( \gamma \) of degree less than some \( m \),

\( 1-F[\gamma]_{M_{QLD}} \iff F \notin V_{QLP}(\gamma) \).

**Subinductive Case:** Assume that \( \beta \) has degree equal to \( m \). There are several subcases:

[\( \neg \) \( \beta = \neg \gamma \) (some \( \gamma \)).]

By Definition 8, NOT 1-F[\( \neg \gamma \)] \( M_{QLD} \) iff \( V_{QLD}(\gamma) = F \).

Because the degree of \( \gamma \) is less than \( k \) (if the degree of \( \beta \) is less than \( k \), then \textit{a fortiori} the degree of \( G \) is less than \( k \)), the Main Hypothesis of Induction dictates that

\( T \notin V_{QLP}(\gamma) \iff V_{QLD}(\gamma) = F \).

By Definition 4, \( F \in V_{QLP}(\neg \gamma) \) iff \( T \in V_{QLP}(\gamma) \).

Therefore, NOT 1-F[\( \neg \gamma \)] \( M_{QLD} \) \( \iff \) \( F \notin V_{QLP}(\neg \gamma) \).
Therefore, NOT 1-F[β]_{MQLD} ⇔ F \notin V_{QLP}(β). \hfill (Conclusion of Negation-case for Subinduction.)

[∧] \quad β = γ \land δ \text{ (some } γ, \text{ some } δ). 

By Definition 8, NOT 1-F[γ \land δ]_{MQLD} \iff NOT 1-F[γ]_{MQLD} \text{ and NOT 1-F[δ]}_{MQLD}.

Because γ and δ are both of degree less than m, the Hypothesis of the Subinduction dictates that

\[ F \notin V_{QLP}(γ) \iff \text{NOT 1-F}[γ]_{MQLD} \]

AND:

\[ F \notin V_{QLP}(δ) \iff \text{NOT 1-F}[δ]_{MQLD}. \]

By Definition 4, F ∈ V_{QLP}(γ \land δ) \iff F ∈ V_{QLP}(γ) \text{ or } F ∈ V_{QLP}(δ).

Therefore, NOT 1-F[γ \land δ]_{MQLD} ⇔ F \notin V_{QLP}(γ \land δ).

Therefore, NOT 1-F[β]_{MQLD} ⇔ F \notin V_{QLP}(β).

(Conclusion of Conjunction-case for Subinduction.)

[∨] \quad β = γ \lor δ \text{ (some } γ, \text{ some } δ). 

By Definition 8, NOT 1-F[γ \lor δ]_{MQLD} \iff NOT 1-F[γ]_{MQLD} \text{ or NOT 1-F[δ]}_{MQLD}.

Because γ and δ are both of degree less than m, the Hypothesis of the Subinduction dictates that

\[ F \notin V_{QLP}(γ) \iff \text{NOT 1-F}[γ]_{MQLD} \]

AND:

\[ F \notin V_{QLP}(δ) \iff \text{NOT 1-F}[δ]_{MQLD}. \]

By Definition 4, F ∈ V_{QLP}(γ \lor δ) \iff F ∈ V_{QLP}(γ) \text{ and } F ∈ V_{QLP}(δ).

Therefore, NOT 1-F[γ \lor δ]_{MQLD} ⇔ F \notin V_{QLP}(γ \lor δ).

Therefore, NOT 1-F[β]_{MQLD} ⇔ F \notin V_{QLP}(β).

(Conclusion of Disjunction-case for Subinduction.)

[→] \quad β = γ → δ \text{ (some } γ, \text{ some } δ). 

By Definition 8, NOT 1-F[γ → δ]_{MQLD} \iff V_{QLD}(γ) = F \text{ or NOT 1-F}[δ]_{MQLD}. 

Then, because the degree of $\gamma$ is less than $k$, the Main Hypothesis of Induction dictates that

$$T \notin V_{QLP}(\gamma) \Leftrightarrow V_{QLD}(\gamma) = F.$$ 

And, because $\delta$ is of degree less than $m$, the Hypothesis of the Subinduction dictates that

$$F \notin V_{QLP}(\delta) \Leftrightarrow \not i-f[\delta]_{\mathcal{M}_{QLD}}.$$ 

By Definition 4, $F \in V_{QLP}(\gamma \rightarrow \delta)$ iff $T \in V_{QLP}(\gamma)$ and $F \in V_{QLP}(\delta)$. Therefore, $\not i-f[\gamma \rightarrow \delta]_{\mathcal{M}_{QLD}} \Leftrightarrow F \notin V_{QLP}(\gamma \rightarrow \delta)$. Therefore, $\not i-f[\beta]_{\mathcal{M}_{QLD}} \Leftrightarrow F \notin V_{QLP}(\beta)$.

(Conclusion of Arrow-case for Subinduction.)

$[\exists] \quad \beta = (\exists x)\gamma$ (some $x \gamma$).

By Definition 8, $\not i-f[\exists x]_{\mathcal{M}_{QLD}} \Leftrightarrow \exists a_i: \not i-f[[x/a_i]\gamma]_{\mathcal{M}_{QLD}}$.

Then, because the degree of $[x/a_i]\gamma$ is less than $m$, the Hypothesis of the Subinduction dictates that

$$F \notin V_{QLP}([x/a_i]\gamma) \Leftrightarrow \not i-f[[x/a_i]\gamma]_{\mathcal{M}_{QLD}}.$$ 

By Definition 4, $F \in V_{QLP}((\exists x)\gamma)$ iff $\forall a_i, F \in V_{QLP}([x/a_i]\gamma)$. Therefore, $\not i-f[(\exists x)\gamma]_{\mathcal{M}_{QLD}} \Leftrightarrow F \notin V_{QLP}((\exists x)\gamma)$. Therefore, $\not i-f[\beta]_{\mathcal{M}_{QLD}} \Leftrightarrow F \notin V_{QLP}(\beta)$.

(Conclusion of Existential-case for Subinduction.)

$[\forall] \quad \beta = (\forall x)\gamma$ (some $x$, some $\gamma$).

(The case is suitably similar to that for $\exists$, and is left out.)

Therefore: The conclusion of the subinduction follows by induction on cases:

$$\forall \beta, \text{if degree of } \beta \text{ is less than } k \Rightarrow (\not i-f[\beta]_{\mathcal{M}_{QLD}} \Leftrightarrow F \notin V_{QLP}(\beta)).$$

(Conclusion of the Subinduction.)
We can now return to the Main Induction.

[Subcase (iii) for \(\neg\)] \(\alpha = \neg\beta\), where \(\beta = \gamma \land \delta\) (some \(\gamma\), some \(\delta\)).

By Definition 4, \(T \in V_{QLP}(\neg(\gamma \land \delta))\) iff \(F \in V_{QLP}(\gamma \land \delta)\).

That is, \(T \in V_{QLP}(\neg(\gamma \land \delta))\) iff \(F \in V_{QLP}(\gamma)\) or \(F \in V_{QLP}(\delta)\).

Now, assume (for the converse) that \(\not{V_{QLD}(\neg(\gamma \land \delta))} = T\). (That is, that \(V_{QLD}(\neg(\gamma \land \delta)) = F\).)

By Definition 7, \(V_{QLD}(\neg(\gamma \land \delta)) = F\) iff \(\not{i-f}_{M_{QLD}[\gamma \land \delta]}\).

And, by Definition 8, \(\not{i-f}_{M_{QLD}[\gamma \land \delta]}\) iff \(\not{i-f}_{M_{QLD}[\gamma]}\) or \(\not{i-f}_{M_{QLD}[\delta]}\).

Since \(\gamma\) and \(\delta\) are both of degree less than \(k\), by the Subinduction, we know that

\[
F \notin V_{QLP}(\gamma) \iff \not{i-f}_{M_{QLD}[\gamma]}\]

AND:

\[
F \notin V_{QLP}(\delta) \iff \not{i-f}_{M_{QLD}[\delta]}\text{.}
\]

So, we have that

\[
V_{QLD}(\neg(\gamma \land \delta)) = F \iff (F \notin V_{QLP}(\gamma) \& F \notin V_{QLP}(\delta))\text{.}
\]

So, conversely, \(V_{QLD}(\neg(\gamma \land \delta)) = T \iff T \in V_{QLP}(\neg(\gamma \land \delta))\).

Therefore, \(V_{QLD}(\alpha) = T \iff T \in V_{QLP}(\alpha)\).  

(Conclusion of Subcase (iii) for negation.)

[Subcase (iv) for \(\neg\)] \(\alpha = \neg\beta\) where \(\beta = \gamma \lor \delta\) (some \(\gamma\), some \(\delta\)).

(The negation-subcase for \(\lor\) is much the same as for \(\land\), and I leave it out.)

[Subcase (v) for \(\neg\)] \(\alpha = \neg\beta\), where \(\beta = \gamma \rightarrow \delta\) (some \(\gamma\), some \(\delta\)).

By Definition 4, \(T \in V_{QLP}(\neg(\gamma \rightarrow \delta))\) iff \(F \in V_{QLP}(\gamma \rightarrow \delta)\).

That is, \(T \in V_{QLP}(\neg(\gamma \rightarrow \delta))\) iff \(T \in V_{QLP}(\gamma)\) and \(F \in V_{QLP}(\delta)\).

Now, assume (for the converse) that \(\not{V_{QLD}(\neg(\gamma \rightarrow \delta))} = T\). (That is, that \(V_{QLD}(\neg(\gamma \rightarrow \delta)) = F\).)

By Definition 7, \(V_{QLD}(\neg(\gamma \rightarrow \delta)) = F\) iff \(\not{i-f}_{M_{QLD}[\gamma \rightarrow \delta]}\).

And, by Definition 8, \(\not{i-f}_{M_{QLD}[\gamma \rightarrow \delta]}\) iff \(V_{QLD}(\gamma) = F\) or \(\not{i-f}_{M_{QLD}[\delta]}\).

Because \(\gamma\) is of degree less than \(k\), by the Main Hypothesis of Induction, we know that

\[
V_{QLD}(\gamma) = F \iff T \notin V_{QLP}(\gamma)\text{.}
\]
And, because the degree of $\delta$ is less than $k$, by the Subinduction, we know that

$$\text{NOT I-F}[\delta]_{\mathsf{MQLD}} \Leftrightarrow F \notin V_{\mathsf{QLP}}(\Delta).$$

So, we have that

$$V_{\mathsf{QLD}}(\neg(\gamma \rightarrow \delta)) = F \Leftrightarrow (T \notin V_{\mathsf{QLP}}(\gamma) \text{ OR } F \notin V_{\mathsf{QLP}}(\delta)).$$

So, conversely, $V_{\mathsf{QLD}}(\neg(\gamma \rightarrow \delta)) = T \Leftrightarrow T \in V_{\mathsf{QLP}}(\neg(\gamma \rightarrow \delta)).$

Therefore, $V_{\mathsf{QLD}}(\alpha) = T \Leftrightarrow T \in V_{\mathsf{QLP}}(\alpha).$ \hspace{1cm} (Conclusion of Subcase (v) for negation.)

[Subcases (vi–vii) for $\neg$] $\alpha = \neg \beta$ where $\beta = (\forall x)\gamma$ (some $x$, some $\gamma$); or $\alpha = \neg \beta$ where $\beta = (\exists x)\gamma$ (some $x$, some $\gamma$)

(The negation-subcases for $\forall$ and $\exists$ are much the same as for $\land$, and I leave them out.)

Therefore, by subcases [i]–[vii] for negation, it follows that

$$V_{\mathsf{QLD}}(\neg(\beta)) = T \Leftrightarrow T \in V_{\mathsf{QLP}}(\neg(\beta)).$$

Therefore, $V_{\mathsf{QLD}}(\alpha) = T \Leftrightarrow T \in V_{\mathsf{QLP}}(\alpha).$ \hspace{1cm} (Conclusion of Negation-case for Main Induction.)

[\land] $\alpha = \gamma \land \beta$ (some $\gamma$, some $\beta$).

By Definition 4, $T \in V_{\mathsf{QLP}}(\gamma \land \beta)$ iff $T \in V_{\mathsf{QLP}}(\gamma)$ and $T \in V_{\mathsf{QLP}}(\beta)$.

Because $\gamma$ and $\beta$ are both of degree less than $k$, by the Main Hypothesis of Induction we know that

$$T \in V_{\mathsf{QLP}}(\gamma) \Leftrightarrow V_{\mathsf{QLD}}(\gamma) = T$$

and:

$$T \in V_{\mathsf{QLP}}(\beta) \Leftrightarrow V_{\mathsf{QLD}}(\beta) = T.$$

By Definition 7, $V_{\mathsf{QLD}}(\gamma \land \beta) = T$ iff $V_{\mathsf{QLD}}(\gamma) = T$ and $V_{\mathsf{QLD}}(\beta) = T$.

So, $V_{\mathsf{QLD}}(\gamma \land \beta) = T \Leftrightarrow T \in V_{\mathsf{QLP}}(\gamma \land \beta)$.

Therefore, $V_{\mathsf{QLD}}(\alpha) = T \Leftrightarrow T \in V_{\mathsf{QLP}}(\alpha)$. 
\textit{(Conclusion of Conjunction-case for Main Induction.)}

\[ \lor \]
\[ \alpha = \gamma \lor \beta \text{ (some } \gamma, \text{ some } \beta). \]

\textit{(The case for } \lor \text{ is suitably similar to the case for } \land, \text{ and I leave it out here.)}

\[ \rightarrow \]
\[ \alpha = \gamma \rightarrow \beta \text{ (some } \gamma, \text{ some } \beta). \]

By Definition 4, \( T \in V_{QLP}(\gamma \rightarrow \beta) \text{ iff } F \in V_{QLP}(\gamma) \text{ or } T \in V_{QLP}(\beta). \)

Assume (for the converse) that \( \text{not } V_{QLD}(\gamma \rightarrow \beta) = T. \) (That is, that \( V_{QLD}(\gamma \rightarrow \beta) = F. \))

By Definition 7, \( V_{QLD}(\gamma \rightarrow \beta) = F \text{ iff not } 1-F[\gamma]_{MQLD} \text{ and } V_{QLD}(\beta) = F. \)

Since the degree of \( \gamma \) is less than \( k, \) by the Subinduction we know that

\[ \text{not } 1-F[\gamma]_{MQLD} \Leftrightarrow F \notin V_{QLP}(\gamma). \]

And, because \( \beta \) is of degree less than \( k, \) the Main Hypothesis of Induction dictates that

\[ V_{QLD}(\beta) = F \Leftrightarrow T \notin V_{QLP}(\beta). \]

So, we have that

\[ V_{QLD}(\gamma \rightarrow \delta) = F \Leftrightarrow T \notin V_{QLP}(\gamma \rightarrow \delta). \]

So, conversely, \( V_{QLD}(\gamma \rightarrow \beta) = T \Leftrightarrow T \in V_{QLP}(\gamma \rightarrow \beta). \)

Therefore, \( V_{QLD}(\alpha) = T \Leftrightarrow T \in V_{QLP}(\alpha). \)

\textit{(Conclusion of Arrow-case for Main Induction.)}

\[ \forall/\exists \]
\[ \alpha = (\forall x)\gamma \text{ (some } x, \text{ some } \gamma); \text{ or } \alpha = (\exists x)\gamma \text{ (some } x, \text{ some } \gamma); \]

\textit{(The cases for } \forall \text{ and } \exists \text{ are similar to the cases for } \land \text{ and } \lor, \text{ and I leave them out.)}

So, by induction on all the cases for the Main Inductive proof,

\[ \forall \alpha, M_{QLD} \vdash_{QLD} \alpha \Leftrightarrow M_{QLP} \vdash_{QLP} \alpha. \]

That is, \( V_{QLD}(\alpha) = (T) \Leftrightarrow T \in V_{QLP}(\alpha). \)
So, for $\mathcal{M}_{QLP} = \langle D_{QLP}, \mu_{QLP}, V_{QLP} \rangle$,

$$\exists M_{QLD} = \langle D, \mu_1, \mu_2, V \rangle;$$
$$\forall \alpha, M_{QLD} \models_{QLD} \alpha \iff M_{QLP} \models_{QLP} \alpha.$$ 

But $\mathcal{M}_{QLP}$ was an arbitrary $\mathcal{QLP}$-model. Therefore:

$$\forall \mathcal{M} = \langle D, \mu \rangle, \exists M_{QLD} = \langle D, \mu_1, \mu_2, V \rangle;$$
$$\forall \alpha, M_{QLD} \models_{QLD} \alpha \iff M_{QLP} \models_{QLP} \alpha.$$

(Conclusion of the Proof of Lemma 1.) $\square$

The second lemma is just the obvious counterpart of the first: I show that for any $\mathcal{QLD}$-model, there exists a $\mathcal{QLP}$-model satisfying just the same set of sentences. In fact, since the proof is relatively similar to the proof of Lemma 1, I dispense with proving the second Lemma, and merely show how to construct the appropriate $\mathcal{M}_{QLP}$, for any $\mathcal{M}_{QLD}$.

**Lemma 2 (Model-equivalence Two: $\mathcal{QLD} \Rightarrow \mathcal{QLP}$).** For any $\mathcal{QLD}$-model, some $\mathcal{QLP}$-model satisfies just the same sentences:

$$\forall M_{QLD} = \langle D, \mu_1, \mu_2, V \rangle,$$
$$\exists \mathcal{M} = \langle D, \mu \rangle;$$
$$\forall \alpha, M_{QLP} \models_{QLP} \alpha \iff M_{QLD} \models_{QLD} \alpha.$$ 

**Proof.** We prove Lemma 2 for arbitrary $\mathcal{M}_{QLD}$, and appropriate $\mathcal{M}_{QLP}$, by induction upon degree of $\alpha$.

Consider an arbitrary $\mathcal{M}_{QLD} = \langle D_{QLD}, \mu_{QLD_1}, \mu_{QLD_2}, V_{QLD} \rangle$.

Define a corresponding $\mathcal{QLD}$-model: $\mathcal{M}_{QLP} = \langle D_{QLP}, \mu_{QLP}, V_{QLP} \rangle$ such that:

1. $D_{QLP} = D_{QLD}$.

2. the assignment $\mu_{QLP}(\ldots)$ is as follows:
(a) for any variable-letter ‘\(x\)’, \(\mu_{QLP}(x_i) = \mu_{QLD_1}(x_i) = \mu_{QLD_2}(x_i)\);
(b) for any constant-letter ‘\(a\)’, \(\mu_{QLP}(a_i) = \mu_{QLD_1}(a_i) = \mu_{QLD_2}(a_i)\);
(c) for any predicate-letter ‘\(F\)’, \(\mu_{QLD_1}(F) = \Gamma \& \mu_{QLD_2}(F) = \Delta \Rightarrow \mu_{QLP}(F) = (\Gamma, \Delta)\).

To prove:

\[
\forall \alpha, M_{QLP} \models_{QLP} \alpha \iff M_{QLD} \models_{QLD} \alpha.
\]
That is, \(T \in V_{QLP}(\alpha) \iff V_{QLD}(\alpha) = (T)\).

The proof is by induction on degree of \(\alpha\), and is similar to the proof of Lemma 1.

*Proof left as an exercise.*

From these results, it clearly follows that, in fact, the consequence-sets we get from any set \(\Sigma\) are the same in both \(QLP\) and \(QLD\). That is, with respect to our first-order language here, we get just the same sets of consequences from the two semantic consequence-relations \(\models_{QLP}\) and \(\models_{QLD}\).

**Theorem 1 (Entailment equivalence (\(QLP \iff QLD\)).** Any entailment-claim true of \(\models_{QLP}\) is also true of \(\models_{QLD}\), and vice-versa:

\[
\forall \Sigma, \forall \alpha, \Sigma \models_{QLP} \alpha \iff \Sigma \models_{QLD} \alpha.
\]

*Proof.* I prove each direction of the biconditional separately.

\[
[\Rightarrow] \quad \forall \Sigma, \forall \alpha, \Sigma \models_{QLP} \alpha \Rightarrow \Sigma \models_{QLD} \alpha.
\]

Assume, for arbitrary \(\Sigma\), arbitrary \(\alpha\), that \(\Sigma \models_{QLP} \alpha\).

That is, \(\forall M_{QLP}, M_{QLP} \models_{QLP} \Sigma \Rightarrow M_{QLP} \models_{QLP} \alpha\).

Which is just to say that \(\forall M_{QLP}, M_{QLP} \models_{QLP} \Sigma \Rightarrow M_{QLP} \models_{QLP} (\Sigma \cup \alpha)\).

Assume (for *reductio*) that \(\Sigma \not\models_{QLD} \alpha\).

That is, \(\exists M_{QLD}: M_{QLD} \models_{QLD} \Sigma \& M_{QLD} \not\models_{QLD} \alpha\).

Which is just to say that \(\exists M_{QLD}: M_{QLD} \models_{QLD} \Sigma \& M_{QLD} \not\models_{QLD} (\Sigma \cup \alpha)\).

But, by Lemma 2, \(\forall \alpha, \forall M_{QLD}, \exists M_{QLP}: M_{QLP} \models_{QLP} \alpha \iff M_{QLD} \models_{QLD} \alpha\).

So, \(\exists M_{QLP}: M_{QLP} \models_{QLP} \Sigma \& M_{QLP} \not\models_{QLP} (\Sigma \cup \alpha)\), which is absurd.

Therefore, \(\Sigma \models_{QLD} \alpha\).

But \(\Sigma\) and \(\alpha\) were arbitrary.

Therefore \(\forall \Sigma, \forall \alpha, \Sigma \models_{QLP} \alpha \Rightarrow \Sigma \models_{QLD} \alpha\).
[⇒] \[∀Σ, ∀α, Σ \Sigma \models_{QLD} α \Rightarrow Σ \Sigma \models_{QLP} α.\]

Assume, for arbitrary Σ, arbitrary α, that Σ \models_{QLD} α.

That is, \[∀M_{QLD}, M_{QLD} \models_{QLD} Σ \Rightarrow M_{QLD} \models_{QLD} α.\]

Which is just to say that \[∀M_{QLD}, M_{QLD} \models_{QLD} Σ \Rightarrow M_{QLD} \models_{QLD} (Σ \cup α).\]

Assume (for reductio) that Σ \not\models_{QLP} α.

That is, \[∃M_{QLP}: M_{QLP} \models_{QLP} Σ & M_{QLP} \not\models_{QLP} α.\]

Which is just to say that \[∃M_{QLP}: M_{QLP} \models_{QLP} Σ & M_{QLP} \not\models_{QLP} (Σ \cup α).\]

But, by Lemma 1, ∀α, \[∀M_{QLP}, ∃M_{QLD}: M_{QLD} \models_{QLD} Σ \Leftrightarrow M_{QLP} \models_{QLP} Σ \Leftrightarrow M_{QLP} \models_{QLP} (Σ \cup α).\]

So, \[∃M_{QLD}: M_{QLD} \models_{QLD} Σ & M_{QLD} \not\models_{QLD} (Σ \cup α), which is absurd.\]

Therefore, \[Σ \models_{QLP} α.\]

But Σ and α were arbitrary.

Therefore \[∀Σ, ∀α, Σ \Sigma \models_{QLD} α \Rightarrow Σ \Sigma \models_{QLP} α.\]

(Conclusion of leftward direction.)

Therefore, by the two directions of the proof:

\[∀Σ, ∀α, Σ \Sigma \models_{QLP} α \Leftrightarrow Σ \Sigma \models_{QLD} α.\]

(Conclusion of the proof of Theorem 1.)

From all this, then, it follows that QLP and QLD are indifferent with respect to consequence. Not only are just the same sets of first-order sentences satisfiable (or not) on models for QLP and models for QLD, but in each case, we get just the same consequences from any set of first-order sentences Σ (even if the semantic interpretations rely on quite different background assumptions in each case, and stake different claims on the nature of ‘satisfaction’ and ‘entailment’).
7 An unsettled debate about truth.

Since the dialethist and disambiguating approaches are equivalent as regards consequence, it might seem as if the two should just be interchangeable with respect to the business of interpretation; however, this just goes to show that extensional notions of consequence underdetermine what we are willing to recognize as ‘appropriate’ inference. In the end, we want more than a conception of consequence which just happens to yield up some set of further belief-commitments, without any explanation why or how this set ‘follows from’ original ascriptions. In particular, we should not presume to call any such suggested consequence-relation ‘rational’ or ‘reasonable inference’; we cannot simply assume that this is what we are identifying, unless we have some reason to think that the resulting consequences are ones which it would make real sense to attribute to others, or to accept ourselves. It will not do simply to assign some set of consequent belief-commitments to an ascribed or accepted prior set, and then just insist that these are the ‘right ones to infer’. Instead, we need to provide a background to our account of consequence, explaining why it is that these consequences are the ones we ought to want.

In fact, the two divergent semantic schemes here, although equivalent in terms of satisfiability and entailment, diverge sharply when it comes to their background assumptions. For its part, the dialethic scheme has it that genuinely contradictory sets and sentences are in fact satisfiable—or at least that it makes best sense sometimes to treat them as if they were satisfiable. When a sentence and its negation are both satisfied on a dialethic model, by assignment of the value-set \{T, F\} (‘both’) to each, this is not to assign some special third value, remember, but to say of each sentence that it is in fact true.\footnote{And also, as it happens, that it is in fact false.} According to the dialethist, some true contradictions are facts about the world, and some predicates are ineradically dissolute, or ambiguous—there will be objects to which they truly apply, and truly do not apply, at one and the same time. Sets of sentences involving such contradictory predicates can be satisfied, and this, furthermore, is the same conception of satisfaction familiar from classical first-order logic: the truth of all the sentences in the set. ‘Satisfaction’ is still satisfaction, only now we can satisfy more than we could before. Too, ‘entailment’ is still entailment—every situation in which the entailing set is satisfied is a situation in which the entailed sentence is satisfied—but there is now a restricted class of consequences for sets which before entailed all sentences whatsoever.

On the other side, the disambiguating semantic treatment of ‘satisfaction’ and ‘entailment’ diverges somewhat from the classical. To say that a set is ‘satisfiable’ is to say that the set can be satisfied, once the
predicates of the language have been disambiguated in some way. We may very well say that a set containing occurrences of ambiguous predicates ‘is satisfiable’, but this is only a form of short-hand. What we really mean to say, from the point of view of the disambiguating scheme, is that the set is satisfiable if and only if occurrences of what seems a single predicate stand revealed for what they really are: occurrences of multiple distinct predicates with different extensions. This short-hand manner of speech is unobjectionable, but only so long as we keep its provisional status in mind. To say that we can satisfy a set appearing to contain sentences equivalent to some contradiction is not to agree with the dialethist—is not to admit the possibility that a sentence be both true and false at the same time. Rather, to say that such a seemingly contradictory set is satisfiable is to say that the set can be re-written, or re-interpreted, so that a new version of the set can be satisfied.

So, differences over what is and can be meant by such things as ‘satisfaction’ and ‘entailment’—rather than differences over the actual consequences of particular sets—lie at the heart of the real and substantial disagreements between the dialethist and those who insist upon disambiguation. In particular, disputes arise over the very possibility of true contradictions. Of course, a dialethist need not think that every predicate possesses truly contradictory applications, nor that every set containing seemingly contradictory predication must be satisfied in a dialethic manner. Dialethists can distinguish the satisfiable from the genuinely incoherent just as well as anybody else. Further, a commitment to dialethism in general, and to the existence of ineradically ambiguous predicate-concepts, is certainly compatible with the idea that some predicates, at least, can be disambiguated; all that is required to earn the right to call oneself a dialethist is the conviction that not all predicate-concepts can be disambiguated. A dialethist need only commit to the idea that some sets are to be satisfied on dialethic models, and can well be willing sometimes to use other sorts of models (disambiguating models for instance) or to combine semantic schemes in order, say, to treat some predicate-concepts in a truly dialethic fashion, while disambiguating others.

Not everyone is willing to be even this much a dialethist, however. Many philosophers and logicians find any semantic scheme which allows for the satisfaction of seemingly contradictory sets and sentences at best problematic, and at worst incoherent. Many—Quine among them—have argued for the conceptual necessity of a classical, bivalent semantics, as the only sort of picture of truth and satisfaction which can rightly lay claim to those notions.\(^{39}\) Others have simply argued that any interpretation of predicate-concepts which

\(^{39}\)For Quine on bivalence, see for instance W. V. Quine, *Philosophy of Logic*, 2nd ed. (1970; Cambridge, Mass. and London, England: Harvard UP, 1986), 83–85. Other arguments from Quine have been mustered, by B. H. Slater for instance (“Paraconsistent Logics?” *Journal of Philosophical Logic* 24 (1995): 451–454) to claim that dialethic-paraconsistent logics simply change the subject, and that the connectives employed in such logics have nothing important in common with their classical correspondents. For the roots of this view in Quine, see Quine (1970;1986), chapter 6; for a contrary point of view, see
treats some as ineradicably ambiguous suffers from what amounts to a failure of nerve. In David Lewis’ “Logic for Equivocators,” for instance, he puts off the use of multiply-valued semantics in relevant logics to undue pessimism about the possibilities for the disambiguation of sentences. On Lewis’ view, we can always possibly disambiguate the ambiguous, and all seemingly contradictory sets of sentences are either genuinely unsatisfiable, or else can be reconfigured in terms of a better, and clearer, conception of the terms in play. And so the debate goes. Opponents of Quine and Lewis insist that not every situation, not every interpretation, can be made sense of in terms of the familiar bivalent and unambiguous notions. For those committed to multi-valued or dialethic schemes (these are not at all the same thing), some sets of sentences can only be properly modeled in other-than-classical terms. Against Lewis’ charge of pessimism comes the counter-charge of undue optimism. Richard Routley has argued that Lewis’ insistence that there must always be a suitable disambiguation is just as unwarranted as the counter-claim—that some situations are such that no disambiguation is ever possible—which Lewis criticizes.

In some ways, this debate has been highly unsatisfying. On either side, claims are made concerning rather deep features of reality, or possible reality—claims which are hard to decisively refute, or support. This is not to say that there should be no relation between a conception of the semantics of sentences, and some understanding of the reality to which those sentences (somehow) relate. At the same time, however, one might hope that an account of the interpretation and negotiation of belief-commitments need not always found itself upon hard and fast principles to do with the essential nature of facts and things. It would be preferable if we could make sense of reasoning, inference, and argumentation without having to resolve all the semantic and metaphysical disputes about (resolvable or irresolvable) ambiguous predicates in advance.

8 Preservation, truth, and ‘good argument’.

It is all well and good to complain about this unsettled state of affairs, but my scruples are of little worth unless I can give some other account of interpretation and reasoning as concerns confused or ambiguous predicate-concepts. After all, the debate between dialethists and those who would have us disambiguate only arises in the first place because both parties are at least attempting to deal constructively with what I Brown (forthcoming).

41Richard Routley, "Relevant Logics and Their Semantics Remain Viable and Undamaged by Lewis’s Equivocation Charge," *Topoi* 2 (1983): 205–216. Priest has distinguished perhaps a third response here: that we will end up disambiguating our predicate-concepts, and discarding the irresolvably ambiguous ones, but that the still-ambiguous semantics will be correct at some time, in some context. I am not sure about the coherence of this alternative view, at least in this context—see Page 52, below.
myself have already admitted is an interesting problem: a theory of reasonable inference, consequence and commitment, operating in the presence of inconsistency and confusion. Further, both sides do try to make good sense of these notions by way of different logical suggestions concerning what would and would not be a good thing to argue in various situations. ‘Good argument’, in turn, has been traditionally explicated in terms of the preservation of truth—a good argument is just a valid argument, one which preserves the truth of its premises. From this point of view, it can at least seem as if any account of the appropriate thing to infer or accept in cases of predicate-ambiguity and -confusion must give some nod to a theory of truth, some account of how predicates are satisfied. That is, it seems as if the classically unsatisfiable—if it is going to make any contribution to reasoning, commitment, inference, or consequence—is either going to have to turn out satisfiable after all (as in dialethist models), or is going to have to be made over into something which is classically satisfiable.

It is just this presumption—that ‘logical argument’, ‘good argument’, and ‘valid argument’ all mean the same thing—that I mean to question. And it is in this connection that I want to consider a variety of familiar and not so familiar ways in which logic can be concerned with preservation. Saying this, I mean to suggest that the truth-preservational account of logic gets things at least half right. That is, ‘logical’ argument has to do with preserving properties of premise-sets, surely—it is just that truth is not the only thing we can (and might want to) preserve. The general idea, then, is that we can generate and analyze theories of logical consequence in terms of their ability to preserve any number of properties of sets of sentences. As these properties may vary widely, so may the theories of consequence concerned with their preservation; not every account of argument will necessarily be concerned with the preservation of truth of satisfiability in particular.

The basic notion can be given by the following definition:

Definition 11 (Preservationist consequence). Let \( \varphi \) be a property of sets of sentences.

Let ‘\( \varphi(\Sigma) \)’ be read: ‘the set \( \Sigma \) has the property \( \varphi \)’. Again, let \( \text{Cn}_S(\Gamma) \) be the set of all \( S \)-consequences of \( \Gamma \).

---

42The ‘good argument = validity’ assumption pops up both in the context of ‘monistic’ arguments that there is a single important notion of ‘Logic’ and ‘good argument’, and in more ‘pluralistic’ accounts which stress the availability of a number of ‘logics’, each of which is rightly a candidate theory of proper argumentation. The monistic assumption (and the distinction between capital-L ‘Logic’ and small-l ‘logics’) can be found in Graham Priest, “Logic: One or Many?” presented at Society for Exact Philosophy Conference (Lethbridge, Alberta, Canada, 1999). Greg Restall and J. C. Beall defend a form of ‘logical pluralism’ in their “Logical Pluralism,” also presented at the SEP Conference (1999). Despite their differences, these authors all identify the business of logic as the analysis of forms of argumentation and reasoning which preserve truth.
Then: The system $\text{PP}$ is $\varphi$-preserving (alternately, $\text{PP}$ preserves $\varphi$) iff:

\[ \forall \Sigma, \forall \alpha, \alpha \in \text{Cn}_{\text{pp}}(\Sigma) \Rightarrow (\varphi(\Sigma) \Rightarrow \varphi(\Sigma \cup \{\alpha\})) \]

That is, a system $\text{PP}$ preserves $\varphi$ just in case the sentence $\alpha$ is a $\text{PP}$-consequence of the set $\Sigma$ only if adding $\alpha$ to $\Sigma$ preserves $\varphi$. Note that this definition captures important features of truth-preservation, certainly, but that there is nothing in it dictating that the property ‘$\varphi$’ be truth. Note also that the ‘only if’ appears here simply because the preservation of the property $\varphi$ will not entirely characterize a theory of consequence. Just as not everything that is simply consistent with a classically consistent set $\Sigma$ is actually a classical-logical consequence of $\Sigma$, not everything that happens to preserve $\varphi(\Sigma)$ when added to $\Sigma$ will be a $\varphi$-preserving consequence of $\Sigma$. The definition captures a necessary condition for being a $\varphi$-preserving consequence, but not a sufficient condition.\(^{43}\)

Of course, this preservation condition can be vacuously fulfilled. When $\Sigma$ does not possess the property $\varphi$, then the conditional ($\varphi(\Sigma) \Rightarrow \varphi(\Sigma \cup \{\alpha\})$) is automatically true, and every sentence $\alpha$ may be a $\varphi$-preserving consequence of $\Sigma$.\(^{44}\) To see this, it is sufficient to realize that another way of looking at Definition 11 is that the $\varphi$-preserving consequences of any set are those things which do not in any way take away from the presence of the property $\varphi$ in $\Sigma$. So, there is another formal explanation of the unprincipled nature of classical consequence in the presence of inconsistency. The absence of overall consistency from any set $\Sigma$ means that anything at all, added to $\Sigma$, preserves what consistency there is; namely, none at all. Similarly, the unsatisfiability of $\Sigma$ in classical terms means that the addition of any sentence at all in no way takes away from $\Sigma$’s satisfiability; after all, it has none to take away.

However, this commonplace, that logical consequence just is the preservation of truth, only leads to problems in a case like Gupta’s. Where there is no overall truth to preserve, there seems to be no work for a truth-preserving (classical) consequence-relation to do. In essence, confusion is pernicious just because it tends to leave us in the dark about what we can and cannot say about the truth of certain sentences. Once we have come to realize that multiple criteria or other factors make the application of a predicate-concept difficult to determine, we may have real difficulty trying to decide whether sentences containing that predicate-concept are—or even can be—true after all. ‘Preservationism’, then, searches for other things.

\(^{43}\)To make more sense of what else is required to get the actual $\varphi$-preserving consequences of a set $\Sigma$, see the ‘General Preservationist Schema’, Definition 12, below.

\(^{44}\)Again, this is only ‘may be’, and will rely upon the full details of the theory of consequence at work. Note, however, that vacuity ensures that the ‘General Schema’ (Definition 12, again) will generate the universal consequence-set: the universal set of sentences will be the only maximal $\varphi$-preserving extension of any $\Sigma$ such that $\text{not } \varphi(\Sigma)$.\)
besides truth, that we might want to preserve about sets of sentences, confused or otherwise. Or rather, since possible candidate features of sets abound, and are hardly hard to find, preservationism is better characterized as the search for means of preserving other properties in addition to, or instead of truth. (To those who rankle immediately at my ‘instead of’, I would say that even if we always want to preserve at least truth, we might also want to preserve other things as well.) Where there seems to be no truth to preserve—as in the case of an apparently unsatisfiable set—might we not want to see if there were at least something else to be done? Recall here Jennings and Schotch, who caution that, while “[m]ost have wanted to reform our theory of implication so that not even contradictory formulae explode [that is, lead to unprincipled consequence] or to alter the notion of truth so as to make contradictory sets satisfiable,” we may rather have

to accept contradictory premise sets as an unpleasant fact of inferential life and to ask what is to be done with them. To the researcher, forced to work with contradictory data, it is unlikely to be of interest to be told what would be the case if all of the data were correct, if the researcher is ultimately after a theory in which that case cannot arise.45

In the dirty business of thought and talk, we may find ourselves wanting a conception of consequence which still works even where we cannot see any obvious way to satisfy a set of sentences as it stands, nor how to re-interpret that set in order that we could satisfy it.

### 8.1 A general preservationist schema.

In the hands of the logician, preservationism is the project of determining principles of inference capable of preserving seemingly worthwhile properties of sets of sentences, beliefs, or commitment-ascriptions, so that argument and reasoning may proceed from those basis-sets. What this requires, first of all, is a precise characterization of the property or properties in question. Just as we can only make sense of and determine the soundness of the canons of classical formal inference once we possess well-defined semantics for sentences of the formal language, we will need some determinate and determinable measure-function with which to operate over those representations as now make themselves available. That is to say, with respect to our confused flatlanders, we will need to determine precisely what we mean by ‘confusion’, and by ‘degree of confusion’, before we can begin to talk about ‘preserving a particular degree of confusion’.46 If it turns out that some such determinate measure of confusion is available, and that this measure picks out some property

---

46While the classical cases of consistency and satisfiability are ‘all or nothing’ matters—a set is consistent or not, satisfiable or not—we can also see our way to talking about sets having properties to some extent or degree. For treatments of consistency and inconsistency in terms of a ‘level of (in)consistency’, see Jennings and Schotch (1981), Schotch and Jennings (1980), (1989).
determinable for all sets of sentences in our formal representation of the flatlander’s beliefs and commitments, the initial specification of a confusion-preserving consequence relation turns out to be relatively simple.

As Bryson Brown has formulated it, there is a general recipe for generating the set of ‘level-preserving’ consequences of any set, where ‘level’ refers generally to the measured degree (or presence) of some property of that set.47 As Brown sees, if we can speak of level-preserving extensions, or maximal level-preserving extensions, of a set, then we can speak of level-preserving consequences of that same set. Consider an example: as is well known, we can specify the truth-preserving consequences of some set, Σ, by way of the set of all truth-preserving extensions of that set. A sentence, α, is a truth-preserving consequence of Σ if and only if it is a truth-preserving extension of every truth-preserving extension of Σ. That is, α is a semantical consequence of Σ if and only if the addition of α to any satisfiable extension of Σ yields a set which is itself satisfiable. Once we have the notion of maximally-satisfiable extensions in hand,48 the same point can be put this way: α semantically follows from Σ if and only if α is an element of every maximally-satisfiable extension of Σ. Similarly, systems of propositional logic often allow characterization of derivable consequence in terms of consistency: α is derivable from Σ if and only if α is consistent with every consistent extension of Σ, or, equivalently, if and only if α is an element of every maximally consistent extension of Σ. Again, note that the absence of satisfiability or consistency renders the consequence relations vacuous—everything follows.

This overall framework—the framework of preservation—allows us now to define other consequence-relations, dependent upon one or more of the wide variety of properties which some set, and its extensions, might possess. The general form of the idea is this:

**Definition 12 (General Preservationist Schema).** For any property ϕ of set Σ, we can define the set of ϕ-preserving consequences of Σ as the set:

\[
\{ \alpha \mid \alpha \text{ is a } \varphi\text{-preserving extension of every } \varphi\text{-preserving extension of } \Sigma \}\]

\((12.1)\)

OR:

\[
\{ \alpha \mid \alpha \text{ is in the intersection of all maximal-} \varphi\text{-preserving extensions of } \Sigma \}\]

\((12.2)\)

---

47Brown (forthcoming).

48Of course, if we are concerned with some property ϕ(Σ) other than satisfiability, then we will want those extensions of Σ which are maximal for this other ϕ.
OR:

\[(12.3) \{ \alpha \mid (\forall \Sigma^*)([\Sigma \subseteq \Sigma^* \& \varphi(\Sigma^*) \& (\forall \beta, \beta \notin \Sigma^* \Rightarrow \not \varphi(\Sigma^* \cup \{\beta\})] \Rightarrow \alpha \in \Sigma^*) \}\]

Recalling our definition of ‘Preservationist consequence’ (Definition 11), we can say that a system PP is \(\varphi\)-preserving if and only if the set of all PP-consequences \(C_{n_{PP}}(\Sigma)\) is just the set of all \(\alpha\) such that \(\alpha\) is in the intersection of all the maximal \(\varphi\)-preserving extensions of \(\Sigma\).

Note that the General Preservationist Schema (GPS) does not give us any of the details of the particular proof-theory or semantics of a \(\varphi\)-preserving logical system or theory of consequence. Rather, Definition 12 only tells us what the \(\varphi\)-preserving consequence-set happens to be. Indeed, because it makes reference to maximal extensions of sets, extensions which have the potential to be infinitely large, we may not always be able to use the GPS to tell just what the consequences of some set actually are. That is, while Definition 12 adequately captures the set of consequences, working out that set may be beyond our powers of determination or calculation. This is not to say that the General Schema is of no use, however. In some cases, particularly where sets are relatively small and properties clearly determinable, it will be obvious—or at least determinable—what the \(\varphi\)-preserving consequences are and are not. If nothing else, this schema provides us with a basic beginning in our search for a theory of consequence that preserves some property in which we have become interested.

9 Confusion preservation.

The goal now is to define a property \(\varphi = \text{‘confusion’}\), for our flatlander subjects, so that we can flesh out the general preservationist schema and determine some particular instances of confusion-preserving consequence. Some rough suggestions about what we might want to preserve have already been made—for instance, we do not want confusion over one predicate-extension to slop over onto others—but nothing precise has been presented.\(^{49}\) In fact, I think that there is one good candidate for preservation: the number of predicates about which one is confused. Further, I will argue that we can make good, or better, sense of the concept and measure of confusion if we do not attempt to solve the problem of the meaning or satisfaction-conditions of sentences containing confused predicates (at least not right away). What I propose, then, is a conception of logical consequence which does not actually require that we understand all the predicate-concepts currently in play, nor that we necessarily know just how to repair problems with those predicate-concepts. In some

\(^{49}\)See Section 3.1
contexts, we can make room for a conception of reasonable inference, or good argument, even where we really have no idea how to interpret certain predicate-concepts. That is, we can see those predicate-concepts as in some wise opaque; as if we were simply unable to make out what they mean, only that they are somehow problematic.

As already discussed, the dialethist of course thinks that we can make out the meaning of such terms, and that ‘confusion’ is really only a prejudicial label. However, it is hard to know just how to take this advice, if ‘advice’ it be. Many of the motivating examples of ambiguously-referring predicate-concepts upon which Priest has relied to make his case for dialethism tend to be more outré than Gupta’s ‘up’, and it is not so obvious that Priest means his account to carry over to commonly-employed terms and concepts. Further, even if we can see our way toward revising the bivalent conception of truth that, all things considered, tends to play a part in most of our everyday reasoning, it is not obvious that we will want to appeal to this notion at every juncture. If, as Priest remarks, not every dialetheia is occasioned by ambiguity, so too not every ambiguity is an occasion to appeal to dialethism. Finally, as an interpretive claim about how we ought to expect the confused flatlander to deal with the revelation of ‘up’-ambiguity, dialethism may be somewhat presumptuous. Though the flatlander might appeal to dialethic devices, and could well solve their ‘up’-related problems if they did so, there is no telling that this is what they will want to do. As Priest himself admits, we might just want to abandon irresolvably ambiguous predicate-concepts, once their true nature becomes apparent. We ought to admit, that is, that “if it is ever discovered that the criteria determine different answers, the old predicate is likely to be discarded in favour of new ones whose senses (and denotations) correspond to the different criteria.” Whether the transition from bivalent to dialethist semantics is so easy or not, the revelation that we appear to have misunderstood the world in our employment of a dual-criteria predicate-concept is not always going to turn out to be a case of having understood things after all (even if ‘in a different way than we thought’). Sometimes we do get things wrong, and sometimes we are confused; in such cases (maybe especially in such cases), we need to be able to continue to reason, even if we do not know how to interpret some predicate-concept or other. Barring any other solution, of course, it might well be that all we have left is a glut-valued dialethic semantic scheme for reasoning with such predicate-concepts; I only hesitate to embrace this conclusion precisely because, I think, we do have other options.

—

50 Even if, as R. E. Jennings has repeatedly insisted, we do not really understand what ‘truth’ is, it is clear enough just what conceptual shape our usual misunderstanding takes.

51 Priest (1995), 367. At the same time as he says this, Priest appends the remark, “However, for the old predicate, the multiple-denotation [dialethist] semantics are correct.” (Emphasis mine.) This further claim might make sense in contexts, but I cannot see how to make sense of the ‘correct’ semantics for a dual-criteria term whose unitary meaning we or our interpretive subjects have just been wrong about. After all, the flatlander is only confused about ‘up’ because they want to employ a predicate-concept with a classical semantics, but are doing things incorrectly.
Those who would have us always disambiguate also seem too sure of themselves, however. That disambiguation would solve some of our problems, I do not doubt, but I am not so certain that we must disambiguate in order to solve those problems. True, some of our present concepts may well be such that sentences featuring them are satisfiable only under disambiguation. At the same time, we are not always interested in ‘satisfaction under disambiguation’. Rather, there may be times when we think that certain predicate-concepts are simply ambiguous, or vague, or just ill-defined. That is, we can come to recognize our own confusions, and to reckon them *irresolvable*; we may think that certain imprecise predicate-concepts are only artificially made over into precise correspondents, and that we distort more than we illuminate if we treat the confused always as if it can readily be made clear.\(^{52}\) There may well be predicate-concepts which we are going to think *will not* ever be made clear, in which case we may not be overly moved by the realization that some sentences containing those predicates *could* be satisfied on the class of all disambiguations.\(^{53}\) In any case, the disambiguating approach, like the dialethist scheme, presumes perhaps too much, insisting as it does that the subject of interpretation will (even must) accept reasoning from disambiguated versions of their predicate-concepts in lieu of reasoning from the originals. The interpreter, if only out of modesty, may not always presume to insist upon or expect any favourite solution to a problem of thought or language.

So *what about* deeply and genuinely confused predicate-concepts—how does one go about reasoning with such things as these, things about which one has to admit real confusion, a bona fide *lacuna* in understanding? As already mentioned, a first thought may simply be to abandon such predicate-concepts, expunging them from the vocabulary, and refusing to employ them in any sentences at all; in cases of extreme confusion, perhaps, this sort of linguistic amputation is just what the doctor ordered. In logic, as in medicine, however, amputation is first and foremost a species of *failure*, a reflection of our inability to save something we otherwise really would have liked to keep. The confused predicate-concept is somehow damaged, true enough, and in the heat of battle we may not have the resources to repair that damage, nor any idea quite how to do so—when it comes to questions about what we ought to infer, and that to which we are committed *right now*, we may not have time to re-evaluate our language, and may not be qualified to decide on the new

\(^{52}\) For an analogous criticism, see Jerry A. Fodor and Ernest LePore, “What Cannot Be Evaluated Cannot Be Evaluated, and it Cannot Be Supervalue Either,” *Journal of Philosophy* **93** (1996): 516–535. In this paper, they try to argue that supervenational treatments of vagueness cannot possibly succeed, since such treatments take a vague predicate-concept, and interpret it in terms of semantics involving predicates which are in no wise vague. On their view, such ‘interpretations’, whatever else they are, are not interpretations of *vagueness*. In some ways, although they really do nothing to propose what the semantics of vagueness might actually be, Fodor and LePore are here analogous to the dialethist, who insists that certain predicate-concepts simply require ambiguous semantics. I suspect, however, that the vagueness/precision debate, like the dialethist/disambiguation debate, will turn out more a debate about metaphysics than semantics, and that any proposed ‘truly vague’ semantics is going to wind up being more or less interdefinable (with respect to consequence) with some ‘precise’ supervenational semantics. This is only a hypothesis, of course, and it is going to be hard to test in the absence of suggestions (from Fodor and LePore, or elsewhere) for a meaningful ‘truly vague’ semantics; it would be a worthwhile subject for further study.

\(^{53}\) Recall the quotation from Jennings and Schotch, page 49, above.
standards of its usage.

Fortunately, not every moment is marked by such dire conceptual and linguistic need, and there is surely sometimes a moment or two to sit down and think hard about one’s concepts, and use of them, in hopes of deciding what to make of them, and in order to judge their sufficiencies and deficiencies. In such reflective moments, perhaps, one may find oneself taking sides in the debate between dialethist and disambiguating approaches—or even just choosing to abandon the concept, after all, ceding to its hopeless confusion. While still in the conversational and interpretive trenches, however, one shall not want to count upon such time to reflect. Instead, the best approach might just be to cut one’s losses—to keep what ground has already been won, rather than trying to ensure that every step forward is exactly the right one. I am suggesting, then, that we should not always approach inference, argument, and commitment from the perspective of one who wants all aspects of language and thought to be perfect, or even perfectible. Further, I want to honour the original tenor of the problem: that we have become confused by some component of a language, ours or others’.

Having said all this, I want to suggest the following first account of the ‘confusion of a set Σ’: the level of confusion of Σ is equivalent to the number of predicates occurring in sentences in Σ which need to be replaced by ‘sterile’ predicates before Σ can be classically satisfied (using ‘sterile’ here to mean predicates which do not occur elsewhere in Σ). So, I help myself to classical satisfiability, since it is a well-defined and generally well-understood formal idea, and I rely upon it in order to measure the degree to which a set of sentences is confused—a measure which I will then seek to preserve, by way of a preservationist consequence-relation designed especially for that purpose. The basic idea is this: rather than try to determine, in advance, how a confused predicate-concept like ‘up’ must be satisfied, or how the flatlanders will want it satisfied, I will ‘leave the confusion in’. That is, I will not presume to resolve the confusion present; rather, I set myself the task of keeping that confusion fixed, keeping it from growing any more widespread.

9.1 The formal mechanics of confusion-preservation.

For this sort of damage control to even possibly be effective, I require a clear notion of what it is, exactly, that I am attempting to contain. In aid of this, I define confusion for any set Σ as the least number of predicates occurring in sentences in Σ that, if replaced at every occurrence by other predicates not already in Σ, would allow Σ to be classically satisfied after all. So, for a set which is already classically satisfiable, the confusion of that set will be equal to zero (0). Such a set does not require that any predicates be replaced; to put it another way, such a set is satisfiable even when none of its predicates are replaced. Furthermore,
a set like:

$$\Sigma = \{ Fa, \neg Fa \}$$

has confusion equal to one (1), since $\Sigma$ can be satisfied if the single predicate ‘$F$’ were to be replaced throughout by predicates (‘$G$’, ‘$H$’) not already occurring in $\Sigma$. Such replacement would yield a set like:

$$\Sigma' = \{ Ga, \neg Ha \}$$

which can be easily satisfied on an ordinary classical first-order model. Obviously, since the set:

$$\Delta = \{ Fa \land \neg Fa \}$$

can also be classically satisfied if the single predicate ‘$F$’ is replaced, transforming it into a set like:

$$\Delta' = \{ Ga \land \neg Ha \}$$

we say that the confusion of $\Delta$ is the same as the confusion of $\Sigma$ (one in either case).\(^{54}\)

Because I am interested in interpretation and rationality, I make certain charitable assumptions about the beliefs and commitments of others. That is to say, I am interested in minimizing confusion, looking always for the least number of predicates requiring replacement in any set $\Sigma$. So, a set like:

$$\Gamma = \{ Fa, Gb, Hc, \neg(Fa \land (Gb \lor Hc)) \}$$

which can be transformed into either of the classically-satisfiable sets:

$$\Gamma^1 = \{ Ea, Gb, Hc, \neg(Ja \land (Gb \lor Hc)) \}$$

\(^{54}\)I like this feature, but some may not. In the original theory of level-preservation, and developments of that theory, it is customary to distinguish sets which are contradictory only as a whole from those containing explicit self-contradictory sentences. It may be that there exists some simple way of working this distinction in for confusion-preservation, or it may turn out to be more difficult. In any case, I see no advantage to be had from working that out in the context of this paper, and Gupta’s flatlander example, and so I do not.
(which systematically replaces occurrences of ‘F’), or:

$$\Gamma^2 = \{ Fa, Eb, Je, \neg(Fa \land(Db \lor Kc)) \}$$

(which replaces both ‘G’ and ‘H’), will be treated as having confusion equal to one (1), and not two (2)—we are interested in transformations like $\Gamma^1$ rather than those like $\Gamma^2$, because the former treat only a single predicate as requiring replacement, whereas the latter go further than is required.\(^{55}\)

The goal, then, is to preserve this value for any set; that is, I want to ensure that my suggested theory of consequence (which I will call CP, for ‘confusion-preservation’) has the feature that, for any set $\Sigma$, if $\Sigma$ can be classically satisfied once some number of predicates are replaced, then the CP consequence-set $C_{n_{CP}}(\Sigma)$ can also be classically satisfied once the same number of predicates are replaced. The first step, then, is to make precise the notion of ‘systematically replacing predicates’, and ensure that it is well-defined for any set of first-order sentences whatsoever. Although the basic notion is relatively simple, precision will require a number of preliminary definitions, and some new notation.

### 9.1.1 The systematic replacement of predicates.

In order to make the procedure for systematically replacing predicate-letters precise, I will need to be able to pinpoint any and all occurrences of predicate-letters in sentences and sets. I begin by defining the set of subformulae:

**Definition 13 (Set of subformulae).** For any first-order sentence $\alpha$, the set of subformulae of $\alpha$ is:

$$\text{Subf}[\alpha] = \{ \beta \mid \beta \text{ is a subformula of } \alpha \}$$

(13.1)

For any set of first-order sentences $\Sigma$, the set of subformulae of $\Sigma$ is the union of all the sets of subformulae for the element-sentences of $\Sigma$. That is:

$$\text{Subf}[\Sigma] = \{ \beta \mid \exists \gamma: \gamma \in \Sigma \land \beta \text{ is a subformula of } \gamma \}$$

(13.2)

\(\text{OR:}\)

$$\text{Subf}[\Sigma] = \{ \beta \mid \exists \gamma: \gamma \in \Sigma \land \beta \in \text{Subf}[\gamma] \}$$

\(^{55}\)Note that the same holds for the other possible replacements, which I do not explicitly consider here: the set replacing ‘F’ and ‘G’; the set replacing ‘F’ and ‘H’; the set replacing ‘F’, ‘G’, and ‘H’.
This definition relies upon the customary notion of subformulae of first-order sentences (with which I will presume the reader is familiar); for example, for the sentences

\[ \neg(Fa \lor (Gb \land Hc)) \]

and

\[ (\forall x)(Fx \rightarrow (\exists y)(\neg Fy)) \]

we have the following:

\[
\text{Subf}[\neg(Fa \lor (Gb \land Hc))] = \\
\{ \neg(Fa \lor (Gb \land Hc)), Fa \lor (Gb \land Hc), Gb \land Hc, Fa, Gb, Hc \}
\]

and:

\[
\text{Subf}[(\forall x)(Fx \rightarrow (\exists y)(\neg Fy))] = \\
\{ (\forall x)(Fx \rightarrow (\exists y)(\neg Fy)), Fx \rightarrow (\exists y)(\neg Fy), (\exists y)(\neg Fy), \neg Fy, Fx, Fy \}
\]

Of course, the auxiliary notion, the set of atomic subformulae, can be defined as well, using the following for convenience:

**Notation 2 (F★).** For any predicate-letter, ‘F’, we use the notation:

\[
F★
\]

to stand in for ‘F’, followed by any sequence of constant- and/or variable-letters. That is:

\[
F★ = Fa_1 \ldots a_n
\]

or:

\[
F★ = Fx_1 \ldots x_n
\]

or:

\[
F★ = Fa_1 \ldots a_i x_1 \ldots x_j \ldots a_k \ldots a_m x_n \ldots x_p
\]

for some sequence \( \langle a_1, \ldots, a_n \rangle \) or \( \langle x_1, \ldots, x_n \rangle \) or \( \langle a_1, \ldots, a_i, x_1, \ldots, x_j, \ldots, a_k, \ldots, a_m, x_n, \ldots, x_p \rangle \).

And so we define:
Definition 14 (Set of atomic subformulae). For any first-order sentence \( \alpha \), the set of atomic subformulae of \( \alpha \),

\[
\text{At-subf}[\alpha] = \{ \beta \mid \beta \in \text{Subf}[\alpha] \land \exists F: \beta = F\star \}
\]

(14.1)

(That is, \( \beta \in \text{At-subf}[\alpha] \) iff \( \beta \) is a subformula of \( \alpha \), of the form \( Fa_1 \ldots a_n \) or \( Fx_1 \ldots x_n \).)

For any set of first-order sentences \( \Sigma \), the set of atomic subformulae of \( \Sigma \) is the union of all the sets of atomic subformulae for the element-sentences of \( \Sigma \). That is:

\[
\text{At-subf}[\Sigma] = \{ \beta \mid \exists \gamma: \gamma \in \Sigma \land \beta \text{ is an atomic subformula of } \gamma \}
\]

(14.2)

OR:

\[
\text{At-subf}[\Sigma] = \{ \beta \mid \exists \gamma: \gamma \in \Sigma \land \beta \in \text{At-subf}[\gamma] \}
\]

Definition 15 (Complete sequence of atomic subformulae). For any sentence \( \alpha \), \( \text{CSeq}[\alpha] \) is a complete sequence of atomic subformulae for \( \alpha \) iff:

\[
\text{CSeq}[\alpha] = \langle F_1\star_1, F_1\star_2, \ldots, F_2\star_1, F_2\star_2, \ldots, F_n\star_1, F_n\star_2, \ldots \rangle
\]

Where:

\[
\forall F_i\star_j, F_i\star_j \in \text{CSeq}[\alpha] \Leftrightarrow F_i\star_j \in \text{At-subf}[\alpha];
\]

and there is exactly one occurrence of \( F_i\star_j \) in \( \text{CSeq}[\alpha] \)

for each occurrence of \( F_i\star_j \) in \( \alpha \).

For any set of sentences \( \Sigma \), \( \text{CSeq}[\Sigma] \) is a complete sequence of atomic subformulae for \( \Sigma \) iff \( \text{CSeq}[\Sigma] \) is some concatenation of candidate \( \text{CSeq}[\alpha_k] \)'s, for all \( \alpha_k \) in \( \Sigma \):

\[
\text{CSeq}[\Sigma] = \langle F_1\star_1, F_1\star_2, \ldots, F_2\star_1, F_2\star_2, \ldots, F_n\star_1, F_n\star_2, \ldots \rangle
\]

Where:

\[
\forall F_i\star_j, F_i\star_j \in \text{CSeq}[\Sigma] \Leftrightarrow \exists \alpha \in \Sigma: \exists \text{CSeq}[\alpha]_n: F_i\star_j \in \text{CSeq}[\alpha]_n;
\]

AND: \( \forall \alpha \in \Sigma, \forall \text{CSeq}[\alpha]_n \),

there is exactly one occurrence of \( F_i\star_j \) in \( \text{CSeq}[\Sigma] \)

for each occurrence of \( F_i\star_j \) in \( \text{CSeq}[\alpha]_n \).
So, CSeq[α] is some enumerated sequence of each and every occurrence of atomic subformulae in the sentence α. Presuming that α has more than one atomic subformulae, there will be multiple possible candidates for the complete sequence, which may be quantified over (‘some CSeq[α] such that . . . ’ and ‘for all CSeq[α], . . . ’), or indexed by subscript, if necessary (CSeq[α]₁, CSeq[α]₂, . . . ). The extension to the case of sets, CSeq[Σ], is some enumerated sequence of each and every occurrence of atomic subformulae occurring in all the element-sentences of Σ. Again, there may be multiple candidates, depending upon the order in which we consider the sentences α ∈ Σ, and the order in which we ‘transfer’ each atom from any CSeq[α]ₙ to CSeq[Σ]. (Note that we can consider any CSeq[α]ₙ, since these are all identical with respect to their elements, and differ only in the order imposed over them.)

Having thus defined CSeq[α] and CSeq[Σ], we can define a special case of the complete sequence, consisting of just those atomic subformulae featuring some particular predicate-letter, ‘F’:

**Definition 16 (Complete sequence of atomic subformulae in ‘F’).** For any sentence α, CSeq[α]F is a complete sequence of atomic subformulae in ‘F’ for α iff CSeq[α]F is a complete sequence of all those atomic subformulae occurring in α which are of the form F★. That is:

\[
\text{CSeq}[\alpha]^F = \langle G★₁, G★₂, \ldots, G★ₙ \rangle
\]

Where:

\[
\forall G★_j, G★_j \in \text{CSeq}[\alpha]^F \iff G = F \& \exists \text{CSeq}[\alpha]_n: G★_j \in \text{CSeq}[\alpha]_n;
\]

and:

\[
\forall G★_j, \forall \text{CSeq}[\alpha]_n,
\]

there is exactly one occurrence of G★_j in CSeq[α]F for each occurrence of G★_j in CSeq[α]ₙ.

(16.1)

For any set of sentences Σ, CSeq[Σ]F is a complete sequence of atomic subformulae in ‘F’ for Σ iff CSeq[Σ]F is some concatenation of candidate CSeq[αₖ]F’s, for all αₖ in Σ:
\[ CSeq[\Sigma]^F = \langle G\star_1, G\star_2, \ldots, G\star_n \rangle \]

Where:

\[ \forall G\star_j, G\star_j \in CSeq[\Sigma]^F \iff G = F \& \exists \alpha \in \Sigma: \exists CSeq[\alpha]^F; G\star_j \in CSeq[\alpha]^F; \]

AND: \[ \forall \alpha \in \Sigma, \forall CSeq[\alpha]^F, \]

there is exactly one occurrence of \( G\star_j \in CSeq[\Sigma]^F \)

for each occurrence of \( G\star_j \in CSeq[\alpha]^F \).

(16.2)

So, \( CSeq[\alpha]^F \) takes any complete sequence of the atomic subformulae of \( \alpha \), \( CSeq[\alpha]^F \), and includes just those atoms featuring the predicate-letter \( 'F' \). Multiple candidates may arise, depending upon the order which we then impose over these atoms. The extension to sets is as expected. That is, \( CSeq[\Sigma]^F \) is some enumerated sequence of all the \( 'F' \)-atoms occurring in element-sentences of \( \Sigma \). Again, there may be multiple candidates.

We now want to define another simple notion, this time just the set of predicates occurring in a sentence, or in some set of sentences. We refer to this as the predicate-set for \( \alpha \) (or \( \Sigma \)).

**Definition 17 (Predicate-set).** For any sentence \( \alpha \), \( Pred[\alpha] \) is the predicate-set for \( \alpha \) iff \( Pred[\alpha] \) is just the set of all predicate-letters occurring in the atomic components of \( \alpha \). That is:

\[ Pred[\alpha] = \{ F_1, F_2, \ldots, F_n \} \]

(17.1)

Where:

\[ \forall F_i, F_i \in Pred[\alpha] \iff \exists \beta \in At-subf[\alpha]: \beta = F_i\star. \]

For any set of sentences \( \Sigma \), \( Pred[\Sigma] \) is the predicate-set for \( \Sigma \) iff \( Pred[\Sigma] \) is just the union of all the \( Pred[\alpha_i] \)'s, for \( \forall \alpha_i \in \Sigma \). That is:

\[ Pred[\Sigma] = \{ G_1, G_2, \ldots, G_n \} \]

(17.2)

Where:

\[ \forall G_i, G_i \in Pred[\Sigma] \iff \exists \alpha \in \Sigma: G_i \in Pred[\alpha]. \]

So, for any sentence, or set of sentences, I can identify all of the following:

1. *All* the (atomic) subformulae;
2. A complete sequence of every occurrence of atoms;

3. A complete sequence of every occurrence of $F$-atoms;

4. All the predicates occurring in subformulae.

I can now define a number of ways of replacing elements of sentences and sets.

**Definition 18 (Atom-replacement).** For any sentence $\alpha$ and any predicate-letter ‘$F$’, let:

$$\Lambda^F_\alpha = \{ \text{CSeq}[\alpha]^F_1, \text{CSeq}[\alpha]^F_2, \ldots, \text{CSeq}[\alpha]^F_n \}$$

(the full set of $\text{CSeq}[\alpha]^F$’s). Now, we define the $G$-replacement of $F_j$ in $\alpha$:

$$\forall \text{CSeq}[\alpha]^F_m \in \Lambda^F_\alpha, \text{CSeq}[\alpha]^F_m = \langle F^\star_1, F^\star_2, \ldots, F^\star_k \rangle,$$

$$\forall G, [F_j/G]^F_m \alpha = \text{ the sentence } \alpha \text{ with } F^\star_j \in \text{CSeq}[\alpha]^F_m \text{ replaced by } G^\star, \text{ with } \star \text{ the same in } F^\star_j \text{ and } G^\star.$$  

(18.1)

The definition is much the same for sets of sentences. For any set $\Sigma$ and any predicate-letter ‘$F$’, let:

$$\Lambda^F_\Sigma = \{ \text{CSeq}[\Sigma]^F_1, \text{CSeq}[\Sigma]^F_2, \ldots, \text{CSeq}[\Sigma]^F_n \}$$

(the full set of $\text{CSeq}[\Sigma]^F$’s). Now, we define the $G$-replacement of $F_j$ in $\Sigma$:

$$\forall \text{CSeq}[\Sigma]^F_m \in \Lambda^F_\Sigma, \text{CSeq}[\Sigma]^F_m = \langle F^\star_1, F^\star_2, \ldots, F^\star_k \rangle,$$

$$\forall G, [F_j/G]^F_m \Sigma = \text{ the set } \Sigma \text{ with } F^\star_j \in \text{CSeq}[\Sigma]^F_m \text{ replaced by } G^\star$$

(18.2)

(in whichever single sentence in which $F^\star_j$ occurs),

with $\star$ the same in $F^\star_j$ and $G^\star$.

That is, for any complete ordering of $F$-atoms in any sentence $\alpha$, we can define a companion-sentence which is identical to $\alpha$, but for the $j^{th}$ occurrence of ‘$F$’, which is replaced by ‘$G$’. For example, in the sentence:

$$\gamma = (\forall x Fx \rightarrow Fa)$$
there are two occurrences of ‘F’. So, there are 2! = 2 distinct candidates for $\text{CSeq}[\gamma]_m^F$, and the complete set of these is:

$$\Lambda^F_\gamma = \{ \langle Fx, Fa \rangle, \langle Fa, Fx \rangle \}.$$ 

We can call these:

$$\text{CSeq}[\gamma]_1^F = \langle Fx, Fa \rangle$$

and:

$$\text{CSeq}[\gamma]_2^F = \langle Fa, Fx \rangle.$$ 

Then we have:

$$[F_1/G]_1^F \gamma = (\forall x Gx \rightarrow Fa);$$

$$[F_2/G]_1^F \gamma = (\forall x Fx \rightarrow Ga);$$

$$[F_1/G]_2^F \gamma = (\forall x Fx \rightarrow Ga);$$

$$[F_2/G]_2^F \gamma = (\forall x Gx \rightarrow Fa).$$

Furthermore, just as for any individual sentence $\alpha$, we can take any complete ordering of $F$-atoms in a set of sentences $\Sigma$, and identify a companion-set, identical to $\Sigma$, but with the $j^{th}$ occurrence of ‘$F$’ replaced by ‘$G$’. The case may be somewhat more complicated to keep track of, so as to ensure that one is replacing the right occurrence of ‘$F$’ in the right sentence in $\Sigma$, but the details are the same.

**Some advice:** For convenience, and since there can be many $F\star$’s in more complex sentences and sets (some of which will be the same $F\star$, occurring in different places in a sentence, or in different sentences in a set), the default candidate $\text{CSeq}[\alpha]_m^F$, when one has a choice about which to consider, should probably be that sequence which preserves the ordering of $F\star$’s left-to-right in $\alpha$—or left-to-right within every sentence of $\Sigma$, those sentences themselves considered in the order in which they are written in the representation of $\Sigma$ at hand (since of course the set $\Sigma$ itself, as opposed to the way it is written out, has no intrinsic ordering of its elements). Following these principles would, say, help one track which ‘Fa’ is which in a simple example like the set:

$$\Sigma = \{ \forall x Fx \rightarrow Fa, Fa \}.$$ 

For this $\Sigma$, the ‘default’ left-right/left-right ordering (call it ‘1’) is:
\[ \text{CSeq}[\alpha]_1^F = \langle Fx, Fa, Fa \rangle. \]

Now, we can easily see that, for this default ordering:

\[
[F_1/G]_1^F \Sigma = \{ \forall xGx \rightarrow Fa, Fa \};
\]

\[
[F_2/G]_1^F \Sigma = \{ \forall xFx \rightarrow Ga, Fa \};
\]

and:

\[
[F_3/G]_1^F \Sigma = \{ \forall xFx \rightarrow Fa, Ga \}.
\]

Of course, we may want to replace more than one occurrence of some predicate-letter with other predicate-letters. So we can extend Definition 18 to sequences of atoms:

**Definition 19 (Atom-sequence replacement).** For any sentence \( \alpha \) and any predicate-letter ‘\( F \)’, let:

\[ \Lambda_\alpha^F = \{ \text{CSeq}[\alpha]_1^F, \text{CSeq}[\alpha]_2^F, \ldots, \text{CSeq}[\alpha]_n^F \} \]

(as before, the full set of \( \text{CSeq}[\alpha]^F \)'s). Now, we define the \( G \)-sequence-replacement of an \( F \)-sequence in \( \alpha \):

\[ \forall \text{CSeq}[\alpha]_m^F \in \Lambda_\alpha^F, \text{CSeq}[\alpha]_m^F = \langle F\star_1, F\star_2, \ldots, F\star_k \rangle, \]

\[ \forall \Gamma = \langle G_1, G_2, \ldots, G_j \rangle, \forall \Phi = \langle F\star_{i_1}, F\star_{i_2}, \ldots, F\star_{i_j} \rangle, \]

if \( \Gamma \) is a \( j \)-membered sequence of predicate-letters,

and \( \Phi \) is a \( j \)-membered subsequence of

\[ \text{CSeq}[\alpha]_m^F, \]

then:

\[ \langle \langle F_i, \ldots, F_j \rangle/(G_1, \ldots, G_j) \rangle_{\star}^F \alpha = \alpha \text{ with:} \]

\[ F\star_{i+1} \text{ replaced by } G_{1}\star (\star \text{ the same}); \]

\[ F\star_{i+2} \text{ replaced by } G_{2}\star (\star \text{ the same}); \]

\[ \vdots \]

\[ F\star_{i+j} \text{ replaced by } G_{j}\star (\star \text{ the same}). \]

Unsurprisingly, we extend the definition to sets. For any set \( \Sigma \) and any predicate-letter ‘\( F \)’, let:

\[ \Lambda_\Sigma^F = \{ \text{CSeq}[\Sigma]_1^F, \text{CSeq}[\Sigma]_2^F, \ldots, \text{CSeq}[\Sigma]_n^F \} \]
(as before, the full set of \( \text{CSeq}[\Sigma]^{F} \)). Now, we define the \textit{G-sequence-replacement of an F-sequence in } \Sigma:\n\forall \text{CSeq}[\Sigma]^{F}_{m} \in \Lambda^{F}_{\Sigma}, \text{CSeq}[\Sigma]^{F}_{m} = \langle F \star_{1}, F \star_{2}, \ldots, F \star_{k} \rangle,\n\forall \Gamma = \langle G_{1}, G_{2}, \ldots, G_{j} \rangle, \forall \Phi = \langle F \star_{i_{1}}, F \star_{i_{2}}, \ldots, F \star_{i_{j}} \rangle,\n\text{if } \Gamma \text{ is a } j\text{-membered sequence of predicate-letters,}\n\text{and } \Phi \text{ is a } j\text{-membered subsequence of }\n\text{CSeq}[\Sigma]^{F}_{m}, \text{then:}\n(19.2)\n\lfloor \langle F_{i_{1}}, \ldots, F_{i_{j}} \rangle/\langle G_{1}, \ldots, G_{j} \rangle \rfloor^{F}_{m} \Sigma = \Sigma \text{ with:}\nF \star_{i+1} \text{ replaced by } G_{1} \star (\star \text{ the same});\nF \star_{i+2} \text{ replaced by } G_{2} \star (\star \text{ the same});\n\vdots\nF \star_{i+j} \text{ replaced by } G_{j} \star (\star \text{ the same}).\n\There are two things to notice. First, Definition 18 is a \textit{special case} of Definition 19 (for the unit-sequences \( \Gamma = \langle G_{i} \rangle \), and \( \Phi = \langle F \star_{j} \rangle \)). Second, for any \n\Gamma = \langle G_{1}, G_{2}, \ldots, G_{j} \rangle,\nand any \n\text{CSeq}[\alpha]^{F}_{m} = \langle F \star_{1}, F \star_{2}, \ldots, F \star_{k} \rangle,\nif \( j = k \), then:\n\lfloor \langle F_{1}, \ldots, F_{k} \rangle/\langle G_{1}, \ldots, G_{j} \rangle \rfloor^{F}_{m} \Sigma \n\text{is a full replacement of } \text{‘}F\text{’} \text{ for } \Sigma; \text{ that is, the corresponding set is } \Sigma \text{ with every occurrence of } \text{‘}F\text{’} \text{ replaced by an occurrence of some } \text{‘}G_{i}\text{’}.\n\With all this in hand, then, we can go on to make exact the idea of \textit{systematically replacing} predicate-letters in some set of sentences—just the idea that I have called upon in aid of understanding how to preserve a minimal level of confusion with respect to predicate-concepts. For any set of predicate-letters \( \Pi, \)
the systematic replacement of $\Pi$ for $\Sigma$ is a set identical to $\Sigma$, but with every occurrence of predicate-letters appearing in $\Pi$ replaced in each case by some other predicate-letter not occurring elsewhere in $\Sigma$. The notion is defined as an orderly construction:

**Definition 20 (Systematic predicate-replacement).** Consider any set $\Sigma$. Consider any set of predicate-letters $\Pi = \{H_1, H_2, \ldots, H_p\}$. Define the systematic replacement of $\Pi$ for $\Sigma$:

$$\Sigma^{-\Pi}$$

as follows:

1. Order $\Pi$, yielding some

   $$\Pi^{\text{ord}} = \langle H_1, H_2, \ldots, H_p \rangle$$

   (The identity of subscripts reflects the fact that the set $\Pi$ and the sequence $\Pi^{\text{ord}}$ contain the same number of elements, but not that the ordering of $\Pi^{\text{ord}}$ reflects any order already in $\Pi$.)

2. Define a sequence of sets (with $p + 1$ members, for any $\Pi^{\text{ord}}$ with $p$ members):

   $$\langle \Sigma^{-\Pi^0}, \Sigma^{-\Pi^1}, \ldots, \Sigma^{-\Pi^p} \rangle$$

   as follows:

   **Step 0:**

   $$\Sigma^{-\Pi^0} = \Sigma.$$

   **Step 1:**

   For $H_1 \in \Pi^{\text{ord}}$, some $\text{CSeq}[\Sigma^{-\Pi^0}]_{H_1}$:

   IF: $\text{CSeq}[\Sigma^{-\Pi^0}]_{H_1} = \emptyset$:

   $$\Sigma^{-\Pi^1} = \Sigma^{-\Pi^0}$$

   IF: $\text{CSeq}[\Sigma^{-\Pi^0}]_{H_1} = \langle H_{1_1}, H_{1_2}, \ldots, H_{1_k} \rangle$:

   $$\Sigma^{-\Pi^1} = [\langle H_{1_1}, \ldots, H_{1_k} \rangle/(G_1, \ldots, G_k)]_{H_1} \Sigma^{-\Pi^0}$$

   WHERE:

   [a] $\forall G_i, G_j \in \langle G_1, \ldots, G_k \rangle, G_i \neq G_j$;

   [b] $\forall G_i \in \langle G_1, \ldots, G_k \rangle, G_i \notin \text{Pred}[^{-\Pi^0}]$. 
Step 2:

For $H_2 \in \Pi^{\text{ord}}$, some $\text{CSeq}_{m}^{\Pi^{\text{ord}}} H_2$:

IF: $\text{CSeq}_{m}^{\Pi^{\text{ord}}} H_2 = \emptyset$:

$\Sigma^{-\Pi^2} = \Sigma^{-\Pi^1}$

$\Sigma^{-\Pi^2} = [(H_{21}, \ldots, H_{2k})/(J_1, \ldots, J_k)]_{m}^{H_2} \Sigma^{-\Pi^1}$

WHERE:

[a] $\forall J_i, J_j \in \langle J_1, \ldots, J_k \rangle, J_i \neq J_j$;

[b] $\forall J_i \in \langle J_1, \ldots, J_k \rangle, J_i \notin \text{Pred}[\Sigma^{-\Pi^1}]$.

; ;

Step p:

For $H_p \in \Pi^{\text{ord}}$, some $\text{CSeq}_{m}^{\Pi^{\text{ord}}} H_p$:

IF: $\text{CSeq}_{m}^{\Pi^{\text{ord}}} H_p = \emptyset$:

$\Sigma^{-\Pi^p} = \Sigma^{-\Pi^{p-1}}$

$\Sigma^{-\Pi^p} = [(H_{p1}, \ldots, H_{pk})/(K_1, \ldots, K_k)]_{m}^{H_p} \Sigma^{-\Pi^{p-1}}$

WHERE:

[a] $\forall K_i, K_j \in \langle K_1, \ldots, K_k \rangle, K_i \neq K_j$;

[b] $\forall K_i \in \langle K_1, \ldots, K_k \rangle, K_i \notin \text{Pred}[\Sigma^{-\Pi^{p-1}}]$.

At each step $n$, this construction replaces each occurrence in $\Sigma$ (if there are any) of predicate-letter $H_n \in \Pi^{\text{ord}}$ with an occurrence of some other predicate-letter—and each replacing predicate-letter is one which occurs nowhere in the set resulting from the previous series of replacements, $\Sigma^{-\Pi^{n-1}}$.

C: Lastly, we define our target-set, the systematic replacement of $\Pi$ for $\Sigma$:

$$\Sigma^{-\Pi} = \Sigma^{-\Pi^p}$$

While Definition 20 is clearly intended to remove (potentially confused) occurrences of predicate-letters from sets of sentences, it may not be so obvious just what the resulting $\Sigma^{-\Pi}$ is supposed to look like. Upon closer inspection, the following becomes clear: for any set $\Sigma$, and for any set of predicate-letters $\Pi$, the set resulting from systematically replacing just those predicate-letters in $\Sigma$ (that is, $\Sigma^{-\Pi}$) will contain at most one occurrence of any predicate-letter in $\Pi$. That is, the following is true:

Lemma 3 (Minimal occurrence of predicate-letters). For any set of first-order sentences $\Sigma$, and any
set of predicate-letters $\Pi = \{H_1, \ldots, H_p\}$,

$$\forall H_i \in \Pi, \forall \text{CSeq}\left[\Sigma^{-\Pi}H_i\right], \text{Card}\left|\text{CSeq}\left[\Sigma^{-\Pi}H_i\right]\right| \leq 1.$$  

(That is, there is at most one occurrence of $H_i$ anywhere in $\Sigma^{-\Pi}$, for any $H_i \in \Pi$.)

Proof. I show that each step in the constructive definition of $\Sigma^{-\Pi}$ (Definition 20) either removes all occurrences of any $H_i$, or adds no more than one such occurrence:

By definition of $\Pi^{\text{ord}}$ (see Definition 20 A):

$$\exists \Pi^{\text{ord}} = \{H_1, \ldots, H_p\};$$

$$\forall H_i \in \Pi, \exists H_j: H_j \in \Pi^{\text{ord}} \& H_i = H_j.$$  

And, by Definition 20, $\Sigma^{-\Pi'}$ is $\Sigma^{-\Pi'^{-1}}$ with each occurrence of $H_j$ replaced by some occurrence of a distinct predicate-letter not occurring anywhere in $\Sigma^{-\Pi'^{-1}}$. So, either:

1. $H_j \notin \text{Pred}\left[\Sigma^{-\Pi'^{-1}}\right]$.

2. $H_j \in \text{Pred}\left[\Sigma^{-\Pi'^{-1}}\right]$.

On the one hand, if (1) is true, then any $\text{CSeq}\left[\Sigma^{-\Pi'^{-1}}H_j\right] = \emptyset$, and so $\Sigma^{-\Pi'} = \Sigma^{-\Pi'^{-1}}$; therefore, $\text{CSeq}\left[\Sigma^{-\Pi'}H_j\right] = \emptyset$. Obviously, since $\text{Card}\left|\emptyset\right| = 0$, and because $H_j = H_i$, $\text{Card}\left|\text{CSeq}\left[\Sigma^{-\Pi'}H_i\right]\right| = 0$.

On the other hand, if (2) is true, then some

$$\text{CSeq}\left[\Sigma^{-\Pi'^{-1}}\right]_{H_j} = \langle H_{j_1}, H_{j_2}, \ldots, H_{j_n} \rangle,$$

and so:

$$\Sigma^{-\Pi'} = \langle H_{j_1}, \ldots, H_{j_n}\rangle \langle J_1, \ldots, J_n \rangle_{H_i}^{H_j} \Sigma^{-\Pi'^{-1}},$$

where none of $\langle J_1, \ldots, J_n \rangle$ occur in $\Sigma^{-\Pi'^{-1}}$.

So, again, $\text{CSeq}\left[\Sigma^{-\Pi'}\right]_{H_j} = \emptyset$. Therefore, as above, $\text{Card}\left|\text{CSeq}\left[\Sigma^{-\Pi'}\right]_{H_i}\right| = 0$.

In either case, then, for it to turn out that

$$\text{Card}\left|\text{CSeq}\left[\Sigma^{-\Pi}\right]_{H_i}\right| > 0,$$
any additional occurrences of $H_i$ must have been introduced at some $\Sigma^{-\Pi^k}$ such that $k > j$ (that is, $\Sigma^{-\Pi^k}$ occurs later on in the construction of the ultimate $\Sigma^{-\Pi}$ than does $\Sigma^{-\Pi^j}$).

Now, it can be shown that any such later step in the construction of $\Sigma^{-\Pi}$, $\Sigma^{-\Pi^k}$ will add at most one occurrence of $H_i$ to $\Sigma^{-\Pi^k-1}$. Again, there will be two cases, depending upon whether or not the predicate-letter $H_k$, which is under consideration at step $k$, is in the set $\Sigma^{-\Pi^k-1}$:

1. $H_k \notin \text{Pred}[\Sigma^{-\Pi^k-1}]$.
2. $H_k \in \text{Pred}[\Sigma^{-\Pi^k-1}]$.

Again, if (1) is true, then any $\text{CSeq}[\Sigma^{-\Pi^k-1}|H_k = \emptyset$, and so $\Sigma^{-\Pi^k} = \Sigma^{-\Pi^k-1}$. Therefore, $\text{Pred}[\Sigma^{-\Pi^k}]$ contains nothing not already in $\text{Pred}[\Sigma^{-\Pi^k-1}]$; in particular, no occurrences of $H_j$ will be added to $\Sigma^{-\Pi^k-1}$ (keeping in mind that $H_j = H_i$).

Finally, if (2) is true, then there exists some complete sequence of $H_k$-atoms taken from $\Sigma^{-\Pi^k-1}$:

$$\text{CSeq}[\Sigma^{-\Pi^k-1}|H_k = (H_{k_1}, H_{k_2}, \ldots, H_{k_n}),$$

and some replacing sequence $\langle J_1, \ldots, J_n \rangle$ such that:

$$\Sigma^{-\Pi^k} = (\langle H_{k_1}, \ldots, H_{k_n} \rangle/\langle J_1, \ldots, J_n \rangle |_{m} H_k \Sigma^{-\Pi^k-1}.$$ 

Now, the replacing sequence $\langle J_1, \ldots, J_n \rangle$ will contain at most one occurrence of $H_j$, depending upon which of two possibilities obtains:

(a) $H_j \notin \text{Pred}[\Sigma^{-\Pi^k-1}]$.
(b) $H_j \in \text{Pred}[\Sigma^{-\Pi^k-1}]$.

If (a) is the case, then, the replacing sequence $\langle J_1, \ldots, J_n \rangle$ may contain an occurrence of $H_j$, but can contain no more than one, since, as Definition 20 makes clear

$$\forall J_i, J_j \in \langle J_1, \ldots, J_k \rangle, J_i \neq J_j.$$ 

Further, if (b) is the case, then the replacing sequence $\langle J_1, \ldots, J_n \rangle$ contains no occurrence of $H_j$, since
Definition 20 stipulates that

\[ H_j \not\in \langle J_1, \ldots, J_n \rangle. \]

The Lemma follows by induction.

I now consider some particular examples of systematic predicate-replacement, in accord with Definition 20. While making the replacement-scheme explicit and formally adequate might be somewhat complicated, the essential idea is not. The set:

\[ \Sigma = \{ Fa, Gb, Hc, \neg(Fa \land (Gb \lor Hc)) \} \]

obviously contains three predicate-letters, ‘F’, ‘G’, and ‘H’; that is,

\[ \text{Pred}[\Sigma] = \{ F, G, H \}, \]

and any subset of \text{Pred}[\Sigma] can be replaced systematically, resulting in (for example, and among others):

\[ \Sigma^{-\{F\}} = \{ Ea, Gb, Hc, \neg(Ja \land (Gb \lor Hc)) \} \]

(which systematically replaces ‘F’ alone), or:

\[ \Sigma^{-\{G, H\}} = \{ Fa, Eb, Jc, \neg(Fa \land (Db \lor Kc)) \} \]

(which replaces both ‘G’ and ‘H’), or:

\[ \Sigma^{-\{G\}} = \{ Fa, Eb, Hc, \neg(Fa \land (Jb \lor Hc)) \} \]

(which replaces ‘G’ alone). Now, it is clear by inspection that both \( \Sigma^{-\{F\}} \) and \( \Sigma^{-\{G, H\}} \) are classically-satisfiable sets, while \( \Sigma^{-\{G\}} \) is not. It is now possible to determine, for any set \( \Sigma \), and any subset of \text{Pred}[\Sigma], whether or not systematically replacing the predicate-letters in that subset will yield a version which is classically-interpretable. Further, it is now determinable that in some cases there will be more than one subset of \text{Pred}[\Sigma] yielding up a classically-satisfiable set under systematic replacement. The interpreter
looking to minimize the apparent confusion in sentences under interpretation can now look for the *smallest* \(\Pi \subseteq \text{Pred}[\Sigma]\) such that \(\Sigma^{-\Pi}\) is classically-satisfiable.

### 9.1.2 Defining predicate-confusion.

As tedious as the procedure for systematic predicate-replacement may be, it does allow me to make the requisite notion of ‘confusion’ precise. In particular, I can now define, for any set \(\Sigma\), the *level of predicate-confusion for \(\Sigma*": the size of the *smallest set* of predicate-letters such that systematically replacing all those predicate-letters (as for Definition 20) renders \(\Sigma\) classically satisfiable.\(^{56}\)

**Definition 21 (Predicate-confusion).** For any set \(\Sigma\), the *predicate-confusion* of \(\Sigma\) is equal to the size of the smallest set of predicate-letters \(\Pi\) such that systematically replacing \(\Pi\) in \(\Sigma\) yields a classically-satisfiable set. That is:

\[
\forall \Sigma, \forall \Pi = \{ F_1, F_2, \ldots, F_n \}, \quad \text{PConfu}[\Sigma] = \min_{\text{Card}[\Pi]} \text{for the set } \{ \Pi | \Sigma^{-\Pi} \text{ is classically-satisfiable. } \}
\]

or:

\[
\forall \Sigma, \forall \Pi = \{ F_1, F_2, \ldots, F_n \}, \Sigma^{-\Pi} \text{ is classically-satisfiable } \Rightarrow
\]

\[
(\forall \Lambda, \text{Card}[\Lambda] < \text{Card}[\Pi] \Rightarrow \Sigma^{-\Lambda} \text{ is NOT classically-satisfiable}) \Rightarrow
\]

\[
(\text{PConfu}[\Sigma] = \text{Card}[\Pi]).
\]

The idea, then, is that for any \(\Sigma\), if \(\text{PConfu}[\Sigma] = n\), then there exist some \(n\) predicate-letters, occurring in sentences of \(\Sigma\), that are in some wise mysterious. That is, \(\Sigma\) cannot be interpreted (i.e., classically-satisfied) so long as those predicates still occur in the set—but \(\Sigma\) can be interpreted once these predicate-letters are taken out of the picture, and replaced with ‘safe’ surrogates. So long as the formal language contains an infinite stock of predicate-letters, any finite set of first-order sentences, howsoever predicate-confused, will be satisfiable under some systematic replacement of predicate-letters (even if this involves replacing all the predicate-letters occurring in sentences of the set). On this scheme, if \(\Sigma\) is already classically-satisfiable, then \(\text{PConfu}[\Sigma] = 0\), because no predicates need be replaced.\(^{57}\)

\(^{56}\)The reason why I am referring to ‘predicate-confusion’, and not simply ‘confusion’, will become obvious below (see Section 11).

\(^{57}\)Note that if this is the case, the characterization in Definition 21.2 is still adequate, since \(\Sigma^{-\Pi}\) is classically-satisfiable and, since no set is of a size smaller than \(\emptyset\), the embedded conditional, \((\forall \Lambda, \text{Card}[\Lambda] < \text{Card}[\emptyset] \Rightarrow \Sigma^{-\Lambda} \text{ is NOT classically-satisfiable})\), is vacuously satisfied.
confusion does not distinguish between

\[ \Sigma = \{ Fa, \neg Fa \} \]

and

\[ \Sigma^* = \{ Fa \land \neg Fa \}, \]

since in either case the level of predicate-confusion is one (1). On the other hand, it does distinguish between

\[ \Gamma = \{ Fa \land \neg Fa \} \]

and

\[ \Gamma^* = \{ Fa \land \neg Fa, Ga \land \neg Ga \}. \]

Whereas, for classical purposes, \( \Gamma \) and \( \Gamma^* \) are indistinguishable with respect to satisfiability and consistency, \( \text{PConfu}[\Gamma] = 1 \) while \( \text{PConfu}[\Gamma^*] = 2 \). In some respects, that is, \( \Gamma^* \) is worse off than \( \Gamma \); more predicates must be treated as problematic if \( \Gamma^* \) is ever to be classically interpreted.

Now that the rather vague notion of ‘confusion’ has been converted to a precisely defined measure of predicate-confusion, it is clear that there are better and worse ways to go about extending one’s beliefs and commitments. On the one hand, \( \text{PConfu}[\Sigma] \) will never be reduced by adding sentences to the set, since the original \( \Sigma \), now a subset of its extension, will still be classically-unsatisfiable, and require just the same revision as before. On the other hand, the addition of sentences to a set can certainly increase the level of confusion, just as the addition of ‘Ga’ and ‘\neg Ga’ to \( \Gamma \) above yields \( \Gamma^* \), and a higher level of confusion. It is now possible to distinguish those extensions that preserve predicate confusion of some set of sentences from those that do not. Too, we can distinguish the most we can add to any set of sentences without increasing confusion.

**Definition 22 (PCP extensions).** For any set \( \Sigma \), any superset \( \Sigma^* \) is a *PCP extension of \( \Sigma \) (that is, a ‘predicate-confusion-preserving extension’ of \( \Sigma \)) if and only if \( \Sigma^* \) has the same level of predicate-confusion
as $\Sigma$ itself:

$$\forall \Sigma, \forall \Sigma^*, \Sigma^* \text{ is a PCP extension of } \Sigma \text{ iff:}$$

$$\Sigma \subseteq \Sigma^* \& \text{PConfu}[\Sigma^*] = \text{PConfu}[\Sigma].$$

(22.1)

For any set $\Sigma$, any superset $\Sigma^*$ is a maximal PCP extension of $\Sigma$ if and only if $\Sigma^*$ is a PCP extension of $\Sigma$ and no sentence not already in that extension can be added to it without increasing the level of predicate-confusion:

$$\forall \Sigma, \forall \Sigma^*, \Sigma^* \text{ is a maximal PCP extension of } \Sigma \text{ iff:}$$

$$[a] \quad \Sigma \subseteq \Sigma^* \& \text{PConfu}[\Sigma^*] = \text{PConfu}[\Sigma]$$

and:

$$[b] \quad \forall \alpha, \alpha \notin \Sigma^* \Rightarrow \text{PConfu}[(\Sigma^* \cup \{ \alpha \})] > \text{PConfu}[\Sigma].$$

(22.2)

Furthermore, confusion-preserving consequence-relations now have a clearly-defined goal: any such consequence-relation should never take one from a set with one particular level of confusion to a consequence-set with some other, higher level of confusion, and should never allow one to extend one’s set of beliefs or commitments in such a way as to become more confused. The General Preservationist Schema (Definition 12), in combination with the definition of Preservationist Consequence (Definition 11), give the desired consequence-set:

**Definition 23 (PCP consequence-set).** For any set of sentences $\Sigma$, the PCP consequence-set (`predicate-confusion-preserving consequence-set’, or $\text{Cn}_{\text{PCP}}(\Sigma)$) of $\Sigma$ is the set of sentences in the intersection of all the confusion-preserving extensions of $\Sigma$. That is:

$$\forall \Sigma, \text{Cn}_{\text{PCP}}(\Sigma) =$$

$$\{ \alpha \mid \alpha \text{ is a PCP extension of every PCP extension of } \Sigma \}$$

(23.1)

or:

$$\{ \alpha \mid \alpha \text{ is in the intersection of all maximal PCP extensions of } \Sigma \}$$

(23.2)
OR:

\[
\{ \alpha \mid (\forall \Sigma^*) ([\Sigma \subseteq \Sigma^* \& \text{PConfu}[\Sigma^*] = \text{PConfu}[\Sigma] \& (\forall \beta, \beta \notin \Sigma^* \Rightarrow \text{PConfu}[\Sigma^* \cup \{\beta\}] > \text{PConfu}[^\Sigma]) \Rightarrow \alpha \in \Sigma^*) \}
\]

(23.3)

That is, a sentence \( \alpha \) is a PCP-consequence of a set \( \Sigma \) if and only if every way of maximally extending \( \Sigma \) so as to keep its level of predicate-confusion from worsening includes \( \alpha \). Obviously, the addition of \( \alpha \) itself will not then increase \( \text{PConfu}[\Sigma] \); furthermore, the addition of \( \alpha \) to any other superset of \( \Sigma \) will not increase the predicate-confusion of that superset, either. The conception of consequence is then simply defined:

**Definition 24 (PCP-consequence).** A sentence \( \alpha \) is a predicate-confusion-preserving consequence—alternately, is a PCP-consequence—of a set of sentences \( \Sigma \) (\( \Sigma \models_{\text{PCP}} \alpha \)) iff:

\( \alpha \in Cn_{\text{PCP}}(\Sigma) \) (as per Definition 23.)

This conception of PCP-consequence, resulting so far only from the General Preservationist Schema, does have its shortcomings. Whereas the prior accounts of consequence-relations given for QLP (Definition 6) and QLD (Definition 10), draw on the semantics for the respective systems, there is no ‘system’ PCP, or at least no systematic semantic account. As already mentioned with respect to the General Schema, determination of the necessary and sufficient conditions for membership in \( Cn_{\text{PCP}}(\Sigma) \) says nothing about the proof theory of a system leading from \( \Sigma \) to that consequence-set. Furthermore, the General Schema gives no indication as to how interpretation of any of the sentences in \( \Sigma \) relates (if at all) to interpretation of its PCP-consequences—in the main, because no semantic interpretation is provided in cases of predicate-confusion. Admittedly, then, the consequence-relation spelled out here is not meant to provide a theory of satisfaction or entailment, even thought it happens to draw upon the idea of classically-satisfiability; that is, Definition 21 says nothing about how a confused set is to be satisfied, basing its determination of \( Cn_{\text{PCP}}(\Sigma) \) only upon the idea that \( \Sigma \) could be re-written in such a way as to make it into some other set, which can itself be satisfied.

This last sticking point, although generally true of consequence-relations defined by way of the General Preservationist Schema, may in fact turn out to be particularly troublesome with respect to PCP-consequence, since it is somewhat difficult to see how any semantics for the predicate-letters actually occurring in a confused set \( \Sigma \) could connect up in a straightforward fashion with a definition of consequence based upon the satisfiability of other sets, featuring other predicate-letters. In part, at least, this is just what I would have expected, since one of the motivating thoughts behind my interest in confusion-preservation
has been that sometimes confusion means that some predicate-concepts simply cannot be meaningfully interpreted. At the same time, however, this attitude, reflected in the replacement of apparently confused predicates—instead of some attempt to sort out just what they actually might mean—also has the drawback that it can make it look as if providing a semantics for genuinely confused predicate-terms may just be impossible.\footnote{But, again, see Section 11, below, on further work to be done with confusion-preservation.}

Lastly, the General Preservationist Schema, and PCP-consequence with it, relies upon consideration of all the maximal confusion-preserving (or other-property-preserving) extensions of a set $\Sigma$, and this can make it difficult, if not impossible, to determine precisely what the PCP-consequences of $\Sigma$ actually are, since these extensions may well be infinite in extent and variety.

None of this is to say that the PCP approach is entirely inadequate, however. Although not all the systematic properties of PCP-consequence can be known right now, some of them certainly can. For one thing, since $\models_{\text{PCP}}$ is defined in terms of the intersection of all maximal-PCP extensions—which are, after all, supersets of the original set of sentences—it is clear that $\models_{\text{PCP}}$ is reflexive. That is, every member of any set of sentences is itself among the PCP-consequences of that set:

$$\forall \Sigma, \forall \alpha \in \Sigma \Rightarrow \Sigma \models_{\text{PCP}} \alpha.$$ 

In some cases, too, where the set $\Sigma$ is relatively simple, Definition 23 will indeed allow one to determine some of $\Sigma$’s important consequences. For instance, for the set:

$$\Sigma = \{ Fa, Gb, Hc, \neg (Fa \land (Gb \lor Hc)) \} ,$$

the following are all evidently true:

1. $\text{PConfu}[\Sigma] = 1$, since the set cannot be classically-satisfied unless the single predicate-letter ‘$F$’ is systematically replaced.

2. $\neg F a \in \text{CnP}_{\text{PCP}}(\Sigma)$, since any extension of $\Sigma$ for which predicate-confusion is 1 will include $\neg F a$.

3. $\neg G a \notin \text{CnP}_{\text{PCP}}(\Sigma)$ and $\neg H c \notin \text{CnP}_{\text{PCP}}(\Sigma)$, since any extension of $\Sigma$ including either of those sentences would have predicate-confusion greater than 1 (because any such extension would require that not only ‘$G$’ or ‘$H$’ be replaced, but at least one other predicate-letter).

Furthermore, it is obvious that $\models_{\text{PCP}}$ is not monotonic: there will be supersets of any set which increase its
level of confusion, and these can eliminate some of the PCP-consequences of the original from consideration:

$$\exists \Sigma: \exists \Delta: \exists \alpha: \Sigma \models_{\text{PCP}} \alpha \land (\Sigma \cup \Delta) \not\models_{\text{PCP}} \alpha.$$ 

Whereas $\neg Fa \in \text{Cn}_{\text{PCP}}(\Sigma)$, the following extension of $\Sigma$ does not have that consequence:

$$\Sigma^+ = \{ Fa, Gb, Hc, \neg(Fa \land (Gb \vee Hc)), \neg((Fa \lor Hc) \land (Gb \lor Hc)) \}.$$ 

To see this, note that $\text{PConfu}[\Sigma^+] = 2$, and $\Sigma^+ - \{F\}$ is no longer classically-satisfiable. True, $\Sigma^+ - \{F,H\}$ can be satisfied, but so can $\Sigma^+ - \{G,H\}$. So, there will be PCP extensions of $\Sigma^+$ which include $\neg Gb$ and $\neg Hc$, and which will not preserve level of predicate-confusion under the addition of $\neg Fa$.

This last example brings out what may seem a curious feature of PCP-consequence. Because the PCP approach is dedicated to minimizing possible confusion, it looks for consequences in terms of smallest-possible sets of predicate-letters to be replaced. Sometimes, as in the example-set $\Sigma$ just now considered, it is clear that this set is unique. Because

$$\Sigma = \{ Fa, Gb, Hc, \neg(Fa \land (Gb \lor Hc)) \},$$

can be re-written in a form like

$$\Sigma^+ - \{F\} = \{ Ja, Gb, Hc, \neg(Ka \land (Gb \lor Hc)) \},$$

which can be classically-satisfied, $\text{PConfu}[\Sigma] = 1$. The systematic replacement of the single predicate-letter ‘$F$’ throughout $\Sigma$ transforms the set into one which is satisfiable, and the charitable interpretive imperative, to minimize apparent confusion, dictates that one treat that one predicate-letter as the one which is in fact confused. After all, no other single predicate-letter occurring in $\Sigma$ will allow for satisfaction under systematic replacement, and we need not consider any larger sets of predicate-letters, having identified a minimal one. However, not all sets of sentences will present a unique opportunity for revision and satisfaction. For instance, multiple distinct subsets of $\text{Pred}[\Sigma^+]$, each the same minimal size (for $\text{PConfu}[\Sigma^+] = 2$), can be replaced to yield satisfiability. That is,

$$\Sigma^+ = \{ Fa, Gb, Hc, \neg(Fa \land (Gb \lor Hc)), \neg((Fa \lor Hc) \land (Gb \lor Hc)) \}.$$
can be re-written either as something like

\[ \Sigma^{+ - \{F, H\}} = \{Ja, Gb, Kc, \neg(La \land (Gb \lor Mc)), \neg((Na \lor Pc) \land (Gb \lor Hc))\}, \]

or a set like:

\[ \Sigma^{+ - \{G, H\}} = \{Fa, Jb, Kc, \neg(Fa \land (Lb \lor Mc)), \neg((Fa \lor Nc) \land (Pb \lor Qc))\}. \]

The addition of ‘\neg Fa’ to \(\Sigma^{+ - \{F, H\}}\) causes no problem, but \(\text{PConfu}[\Sigma^{+ - \{G, H\}} \cup \{ \neg Fa \}] = 3\), an increase. Similarly, while the addition of ‘\neg Gb’ to \(\Sigma^{+ - \{G, H\}}\) is allowed, \(\text{PConfu}[\Sigma^{+ - \{F, H\}} \cup \{ \neg Gb \}] = 3\), again an increase. So, because in each case there will be some PCP extension that will not allow addition, neither ‘\neg Fa’, nor ‘\neg Gb’ are PCP-consequences of \(\Sigma^{+}\).

This feature could be eliminated, of course: it is always possible to simply stipulate that, in such a case one should just choose one of the candidate subsets of predicate-letters arbitrarily. Indeed, one might well believe this sort of thing a better representation of what goes on in interpretation. Be that as it may, I do not so stipulate here: the PCP approach will sometimes find itself weakened by such multiply-confused sets, but at least it allows one to identify possible candidates for confusion-minimization (which is not nothing, after all). Furthermore, the definition of PCP-consequence does reveal one sufficient condition for membership in \(\text{Cn}_{\text{PCP}}(\Sigma)\). Even if there will be sentences which are not PCP-consequences of \(\Sigma\) because they contain predicate-letters which are in some, but not all, of the sets which are candidates for replacement, it is still demonstrable that, if a sentence \(\alpha\) contains only predicate-letters which appear in every set which is a candidate for replacement in \(\Sigma\), then \(\alpha\) is a PCP-consequence of \(\Sigma\):

**Lemma 4 (Sufficient Condition for PCP-consequence).**

\[ \forall \Sigma, \forall \alpha, \forall \Pi: \text{Card } |\Pi| = \text{PConfu}[\Sigma], \]

\[ (\Sigma^{+ - \Pi} \text{ is classically-satisfiable } \Rightarrow \forall F_i \in \text{Pred}[\alpha], F_i \in \Pi) \Rightarrow \alpha \in \text{Cn}_{\text{PCP}}(\Sigma). \]

**Proof.** I prove the contrapositive.

Suppose, for some sentence \(\alpha\), and some set \(\Sigma, \alpha \notin \text{Cn}_{\text{PCP}}(\Sigma)\). To show:

\[ \exists \Pi: \text{Card } |\Pi| = \text{PConfu}[\Sigma]: \]

\[ (\Sigma^{+ - \Pi} \text{ is classically-satisfiable } \& \exists F_i \in \text{Pred}[\alpha]: F_i \notin \Pi) \]
Since $\alpha \notin \mathbf{Cn}_{\text{PCP}}(\Sigma)$, the definition of the PCP-consequence set of $\Sigma$ (Definition 23) dictates that $\exists \Sigma^*$ such that $\Sigma^*$ is a maximal PCP extension of $\Sigma$ and $\alpha \notin \Sigma^*$.

By the definition of maximal PCP extension (Definition 22.2), we know that

$$\text{PConfu}[\Sigma^*] = \text{PConfu}[\Sigma]$$

AND: $\text{PConfu}[(\Sigma^* \cup \{ \alpha \})] > \text{PConfu}[\Sigma]$.

From this it follows, by the definition of $\text{PConfu}[\Sigma]$ (Definition 21), that

$$\exists \Pi: \text{Card} |\Pi| = \text{PConfu}[\Sigma^*]$$

AND: $\Sigma^* - \Pi$ is classically-satisfiable,

BUT: $(\Sigma^* \cup \{ \alpha \})^{-\Pi}$ is not classically-satisfiable.

From this it follows that the introduction of $\alpha$ to $\Sigma^*$ must also introduce some predicate-letter which is in $\alpha$, but not in $\Pi$.

But, since $\Sigma^* - \Pi$ is classically-satisfiable, and since $\Sigma \subseteq \Sigma^*$, $\Sigma - \Pi$ is classically-satisfiable.

So, $\exists \Pi: \text{Card} |\Pi| = \text{PConfu}[\Sigma]; (\Sigma - \Pi$ is classically-satisfiable & $\exists F_i \in \text{Pred}[\alpha]; F_i \notin \Pi$) (which was to be shown).

*The Lemma follows by contraposition.*

As a last note, I would point out that while the PCP account is similar to what I have called (taking the term from Brandom) 'de dicto' ascription, but with an important difference: while I still make no commitment (implicit or explicit) to any manner in which a confused set can be satisfied, I do account for the relation between some confused and inconsistent sentences/sets and others. Whereas the purely intensional or de dicto treatment refuses to endorse any view of the semantic and inferential connections between inconsistent and confused beliefs and commitments, the PCP approach, while certainly not an 'interpretation' in the sense of a semantics, still provides something on the order of an account of inference and argumentation. While not a complete story yet, PCP gets one on the way, and can be seen as providing at least a beginning to an account of the 'meaning' of confused and inconsistent discourse. If nothing else, PCP-consequence allows one to see a confused individual as still in some sense 'reasonable', so far, anyway, as they are engaged in reasoning.

---

59 See page 10, above.
10 Comparing the various approaches.

Although it has not been my purpose, in this paper, to explore all the details concerning the three consequence-relations, $\models_{\text{QLP}}$, $\models_{\text{QLD}}$, and $\models_{\text{PCP}}$, some general observations are worth making.

**Fact 2 (PCP provides the strongest consequence-relation).** Of the three consequence-relations under consideration, $\models_{\text{PCP}}$ is stronger than either $\models_{\text{QLP}}$ or $\models_{\text{QLD}}$ (obviously, if it is stronger than one, it is stronger than the other, since the latter two are equivalent). That is:

\[
\forall \Sigma, \forall \alpha, \Sigma \models_{\text{QLP}} \alpha \Rightarrow \Sigma \models_{\text{PCP}} \alpha, \quad \text{but} \quad \exists \Sigma: \exists \alpha: \Sigma \models_{\text{PCP}} \alpha \& \neg \exists \Sigma \models_{\text{QLP}} \alpha.
\]

(And similarly for $\models_{\text{QLD}}$.)

**Proof.** **Proof of [b]** I begin by proving [b], since it is sufficient to note that the classically-unsatisfiable-set

\[\Sigma = \{ Fa, Gb, Hc, \neg(Fa \land (Gb \lor Hc)) \}\]

has different consequences under $\models_{\text{PCP}}$ and $\models_{\text{QLP}}$.

Clearly, $\text{PConfu}[\Sigma] = 1$, because the single predicate-letter ‘F’ can be replaced throughout in order to yield a classically-satisfiable set; furthermore, any maximal-$\text{PConfu}[\Sigma]$-preserving extension of $\Sigma$ will contain the sentence ‘$\neg Fa$’. If not, then there exists some maximal-$\text{PConfu}[\Sigma]$-preserving extension of $\Sigma$, call it $\Sigma^*$, such that

\[\text{PConfu}[(\Sigma^* \cup \{ \neg Fa \})] > 1.\]

But this is not possible. Any such extension must contain $\Sigma$ as a subset, and so, any classically-satisfiable revision of $\Sigma^*$ must contain a classically-satisfiable revision of $\Sigma$ as well, which will require either that ‘F’ be replaced throughout $\Sigma^*$ (thereby allowing ‘$\neg Fa$’ to be added without penalty), or that both ‘G’ and ‘H’ be replaced throughout (which would amount to $\text{PConfu}[\Sigma^*] > 1$, which is absurd). Therefore:

\[\neg Fa \in \text{Cn}_{\text{PCP}}(\Sigma).\]
But

$$\neg F a \notin \text{Cn}_{\text{QLP}}(\Sigma),$$

because there exists some $M_{\text{QLP}}$ such that $M_{\text{QLP}} \models_{\text{QLP}} \Sigma$, but $M_{\text{QLP}} \not\models_{\text{QLP}} \neg F a$. Such a model would be one such that:

1. $\mu(F) = \langle \Gamma^1, \Delta^1 \rangle$ such that: $\Gamma^1 = \Delta^1$, and: $\mu(a) \in \Gamma^1$;
2. $\mu(G) = \langle \Gamma^2, \Delta^2 \rangle$ and: $\mu(b) \in \Gamma^2 \leftrightarrow \mu(b) \not\in \Delta^2$;
3. $\mu(H) = \langle \Gamma^3, \Delta^3 \rangle$ and: $\mu(c) \in \Gamma^3 \leftrightarrow \mu(b) \not\in \Delta^3$;

Condition (1) makes it such that ‘$\neg F a$’ is not satisfied on this $M_{\text{QLP}}$, while all of (1)–(3) together ensure that $\Sigma$ is satisfied.

**Proof of [a]** I want to show that

$$\forall \Sigma, \text{Cn}_{\text{QLP}}(\Sigma) \subseteq \text{Cn}_{\text{PCP}}(\Sigma).$$

So, suppose some sentence $\alpha \in \text{Cn}_{\text{QLP}}(\Sigma)$, but $\alpha \not\in \text{Cn}_{\text{PCP}}(\Sigma)$. Because PCP-consequence is reflexive, we know that $\alpha \not\in \Sigma$; further, by Lemma 4, we know that

$$\exists \Pi: \text{Card} |\Pi| = \text{PConfu}[\Sigma]:$$

$\Sigma^{-\Pi}$ is classically-satisfiable, and

$$\exists F_i \in \text{Pred}[\alpha]: F_i \notin \Pi,$$

and that

$$(\Sigma \cup \{ \alpha \})^{-\Pi}$$

is not classically-satisfiable.

That is, we know that there is some predicate-letter appearing in $\alpha$ that is not in some minimal set of predicate-letters $\Pi$ such that systematically replacing $\Pi$ alone in $\Sigma$ renders the set classically-satisfiable. In other words, the predicate-letters in $\Pi$ cause an inconsistency in $\Sigma$, one which is repaired once those letters alone are replaced. Furthermore, we know that no replacement of the predicate-letters from $\Pi$ alone will make $(\Sigma \cup \{ \alpha \})$ classically-satisfiable.

But if this is the case, then we can satisfy $\Sigma$ if, for each predicate-letter ‘$H_j$’ in $\Pi$, we treat some
$H_j$-atom $\in \text{CSeq}[\Sigma]$ as ‘both True and False’. So, some $\mathcal{M}_{\text{QLP}}$ will be able to satisfy $\Sigma$ by assigning all predicate-letters in $\Pi$ to appropriately fragmented extensions. At the same time, such a $\mathcal{M}_{\text{QLP}}$ will be able to satisfy $\Sigma$ without assigning some $F_i \in \text{Pred}[\alpha]$ to a fragmented extension. Since $(\Sigma \cup \{ \alpha \})^{-\Pi}$ is still classically-unsatisfiable, however, no such $\mathcal{M}_{\text{QLP}}$ will $\text{QLP}$-satisfy $\alpha$.

So, $\alpha \notin \text{Cn}_{\text{QLP}}(\Sigma)$, which is absurd.

Therefore, $\alpha \in \text{Cn}_{\text{PCP}}(\Sigma)$, after all.

The lemma follows from both halves of the proof together.

The result for QLD follows by Theorem 1. $\square$

The preceding proof is admittedly somewhat inexact, but it establishes the point: $^{60}$ PCP-consequence provides for a larger class of consequences than does the $\text{QLP}$ or QLD approach. Not only that, but PCP-consequence is strong in an interesting way: it yields a larger set of consequences precisely because it builds in the requirement that interpretation minimize the difficulties at hand. While the PCP approach concentrates on the least number of predicate-concepts that must be treated as confused, $\text{QLP}$ considers all the possible ways of assigning dialethic interpretations to predicate-letters, and QLD considers all the possible ways of disambiguating sets of sentences. In a case like Gupta's, in which a subject appears to have a single predicate-concept, among all the ones she employs, about which she is confused, $\text{QLP}$ and QLD cannot reveal that fact, since they must also consider models on which every possible combination of predicate-letters is considered as potentially dialethic, or in need of disambiguation.

The strength of PCP-consequence is perhaps best revealed when we consider how the different approaches treat sets which are actually classically-satisfiable in the first place. For any such classically-satisfiable set $\Sigma$, the set of PCP-consequences is just the set of classical consequences of $\Sigma$. Because $\text{PConfu}[\Sigma] = 0$ for classically-satisfiable $\Sigma$, the set of maximal PCP extensions of $\Sigma$ is just the set of its maximally-classically-satisfiable extensions, and the intersection of those is just the set of classical semantic consequences of $\Sigma$. But even simple classically-satisfiable sets like

$$\Delta = \{ Fa, Fa \rightarrow Gb \}$$

$^{60}$To make the proof more precise, I should need to establish what is for now only a conjecture: namely that for every set $\Sigma$ with $\text{PConfu}[\Sigma] = n$, there exists some $\mathcal{M}_{\text{QLP}}$ such that $\mathcal{M}_{\text{QLP}} \models_\text{QLP} \Sigma$ and $\mathcal{M}_{\text{QLP}}$ assigns exactly $n$ predicate-letters to ‘dissolute’ extensions. A similar result would then hold for QLD, obviously. I do, however, think that this conjecture can be made good. For further discussion, and some other implications of the claim see Section 11, below.
will not take their classical consequences (particularly, the sentence ‘Gb’) under QLP or QLD. After all, some $\mathcal{M}_{QLP}$ which assigns ‘Fa’ to the value-set {T, F} and ‘Gb’ simply to {F} alone will satisfy $\Delta$, but (obviously) won’t satisfy ‘Gb’ (see Definition 3 in order to check this claim). By Lemma 1, some $\mathcal{M}_{QLD}$ will do the same, and neither system will sanction the inference from $\Delta$ to ‘Gb’. In fact, the strength of the PCP approach applies even to some classically-satisfiable subsets of sets which are otherwise classically-unsatisfiable. For any set $\Sigma$, if a subset $\Sigma'$ is properly classically-satisfiable, in the sense that:

1. $\Sigma'$ is classically-satisfiable; and

2. $\forall \Sigma''$, if $\Sigma''$ is classically-unsatisfiable, then: $(\Sigma' \cap \Sigma'') = \emptyset$,

then the classical-consequences of $\Sigma'$ can be added to $\Sigma$ without ever increasing $\text{PConfu}[\Sigma]$. Too, these classical-consequences of $\Sigma'$ will be members of every maximal PCP extension of $\Sigma$, and so will be among the PCP-consequences of $\Sigma$. The PCP approach will therefore allow us to retain any classical-consequences of the ‘uninfected’ members of any set of sentences, a notable improvement over QLP and QLD.

Lastly, I would argue that PCP-consequence better captures cases like Gupta’s than do QLP or QLD. Where mistakes have been made in the application of predicate-concepts, and these mistakes are not yet sorted out, it does not seem right either to insist that some sort of dialethic semantics is in fact correct for the predicate-concepts in question, nor that the right solution to the problem lies with disambiguating them. The dialethic approach may not be objectionable in and of itself, but there is no saying that we have given the ‘correct meaning’ of a confused term, dialethic or not—confusion, after all, is an occasion for thinking that we do not in fact really understand the meaning of our terms, or that our subjects may not understand theirs. Too, concepts may well turn out to be ambiguous sometimes, and we may in fact want to separate them off into two or more predicate-concepts by way of disambiguation. But admitting that this sometimes happens, or sometimes seems the right thing to do, is not the same thing as claiming that, where a mistake in application and understanding has been made, the reason is necessarily ambiguity. Where one predicate-term appears to pick out two disjoint properties, it may not be that there are two predicate-concepts at work—rather, there may not even be one of them. If Gupta’s flatlanders were to discover that their two criteria for ‘up’-introduction did not in fact match, they certainly might decide to separate ‘up’ off into two different predicate-concepts. At the same time, however, the discovery, all of a sudden, that nearly no object was in fact truly ‘up’ under both criteria might just as well lead to the abandonment of both criteria, and the elimination of the ‘up’-concept altogether. After all, if I were to discover that men and women were in fact separate species, I would just as likely abandon any and all ideas I had about ‘human beings’ as I would
attempt to sort out ‘which was which’. The realization of deep confusion may not be something one will want to deal with by simple-minded division of one concept into many.

Because of its greater strength as a consequence-relation, and because it seems better suited to cases in which interpretation has to deal with such genuine deficits in understanding as confusion, PCP-consequence is worth some further consideration, and more investigation. Even on its own merits, leaving aside any comparisons with dialethism or disambiguation, confusion-preservation provides for something perhaps unexpected, and certainly interesting: the ability to discern reasoning and ‘good argument’, even in contexts where predicate-terms and -concepts are ill-understood, or even apparently meaningless. This sort of preservationist approach allows one to interpret, and perhaps to reason, even where one lacks a semantic account of every term in one’s language, and even in contexts where things like validity and consistency seem unavailable. Preservationism allows one to employ logical methods where many of the hallmarks of Logic are absent. This suggests, at least to me, that the notion of logical argument, and of reasoning generally, is a more general notion than has been thought. If ‘preservation’ can mean more than truth-preservation, and if logic can as much keep us from further error as it can lead us from truth to truth, then there are any number of new approaches to problems of interpretation and argumentation which one might want to take.

11 What remains to be done.

I choose to end this paper on a forward-looking note. If confusion-preservation, as I have claimed, is an interesting and perhaps profitable enterprise, then there is much more that can be done, given what basis has been provided. The most obvious task, I think, will be to weaken the notion of PCP-consequence. Although it is true that the PCP approach restricts the class of consequences from a classically-unsatisfiable set, it does not restrict that class quite so much as one might like. For observe: defined in terms of the set of all maximal predicate-confusion preserving extensions of $\Sigma$, $\models_{PCP}$ eliminates unprincipled inference with respect to predicates other than those about which the set may express confusion, but allows such unprincipled inference with respect to the potentially confused members of $\text{Pred}[\Sigma]$. That is to say, from the set

$$\Sigma = \{ \text{Gb}, F_a, \neg F_a \}$$

61To borrow Priest (1999)’s capitalization of the ‘canonical employment’ of logic.
the definition of PCP-consequence yields a consequence-set like

$$\mathbf{C}_{n_{\text{PCP}}} (\Sigma) = \{ Gb, Fa, \neg Fa, Fb, \neg Fb, Fc, \neg Fc, \ldots \}.$$  

Because the addition of any ‘F’-atom will not increase $\mathbf{PConfu}[\Sigma]$, $\mathbf{C}_{n_{\text{PCP}}} (\Sigma)$ contains them all.

If one is definitely confused about the application of the predicate ‘F’ (i.e. $\forall \Pi, \Sigma^{-\Pi}$ is classically-satisfiable, then $F \in \Pi$). then PCP-consequence gives, as consequence, that ‘F’ both does and does not apply to every object in our domain. On the one hand, this may not be so unreasonable; after all, so far as I am in the dark about the meaning of ‘F’, I have as much reason to think that the predicate applies to everything as I do to just insist that it applies to only some few particular things. On the other hand, this version of PCP-consequence is really too permissive. If one begins with a predicate-concept taken to be confused based upon a single instance of an object over which introduction-criteria diverge, then there is no apparent reason to suppose, without further evidence, that these criteria diverge in all cases.\textsuperscript{62} I have argued that it is a virtue of the PCP approach that, while it relies upon the idea that certain confused predicate-concepts cannot presently be meaningfully interpreted, it does not commit unnecessarily to any specific future solution to confusion. In particular, PCP-consequence ought not commit to a solution that would have to treat any even partially-confused predicate-concept as if it were in fact universally confused. To do so would be equivalent to the dialethist interpreting every dialethic predicate-concept as if it had no normal applications whatsoever, or the disambiguating scheme treating every ambiguous predicate-concept as if it were universally ambiguous. In any case, such a response would certainly go beyond what was strictly necessary, painting a picture of a worse problem than originally evident.

In part, the problem just noticed is only a function of the dual nature of first-order sentences. Just as these sentences are composed not only of predicate-terms naming predicates, but also of object-terms naming things to which predicates apply, so a measure of confusion meant to capture meaningful features of sets of these sentences will require that we pay attention to both dimensions. A more reasonable measure of confusion, then, will require that its system index two features of a set of sentences: not only the number of predicates that appear confused, but also the number of objects in the domain ‘around which’ predicate-confusion centers. If confusion-preservation can limit consequence with respect to both these indices, not only will it be able to prevent unprincipled inference with respect to objects as well as predicates, but it will

\textsuperscript{62}Of course, Gupta’s case, as described, is one in which the various versions of the confused predicate, ‘up’, do diverge in nearly all cases. This is no good reason to make universal divergence a general-purpose and first-glance supposition, however. Furthermore, there is no reason to expect that the flatlanders would want to make such a supposition before they investigated the matter further, and so any interpretation of their beliefs and commitments should probably not make the supposition either.
also allow interpreters to further determine the extent to which subjects are confused, ‘ranking’ them more precisely with respect to degree. What one wants, that is, is to distinguish

\[ \Sigma = \{ Fa, \neg Fa \} \]

and

\[ \Sigma^* = \{ Fa, \neg Fa, Fb, \neg Fb \}, \]

since \( \Sigma^* \) involves confusion about the application of ‘\( F \)’ to the objects named by ‘a’ and ‘b’, whereas \( \Sigma \) is only confused about the application of ‘\( F \)’ to the object named by ‘a’ alone. At the same time, it would be desirable to separate a set like

\[ \Delta = \{ \forall x Fx, \exists x \neg Fx \}, \]

which ‘says’ that there is confusion about *at least one object* with respect to ‘\( F \)’, from one like

\[ \Delta^* = \{ \forall x Fx, \forall x \neg Fx \}, \]

which ‘says’ that there is confusion about *all objects* with respect to ‘\( F \)’. The presence of quantifiers makes the case complicated, but I have worked out the beginnings of an account which will factor the domain of objects into a measure of confusion just as well as PCP deal with predicates.

Secondly, I should like to, once I have developed a dual measure of confusion, prove the following conjectures:

**Conjecture 1 (Satisfaction-equivalence).** For any set of sentences \( \Sigma \), \( \text{PConfu}[\Sigma] = n \) iff there exists some \( M_{QLP} (M_{QLD}) \) such that \( M_{QLP} \models \Sigma (M_{QLD} \models \Sigma) \) and \( M_{QLP} (M_{QLD}) \) assigns exactly \( n \) predicate-letters to extensions which are non-classical.

Of course, if this were true (and I have reason to believe it is), then the following would be true as well:

**Conjecture 2 (Strong consequence-equivalence).** Both the QLP and QLD approaches can be strengthened, so that they concentrate on the subset of their models which assign a *minimal number* of predicate-letters to non-classical extensions. Once this is managed, then the three approaches would all be equivalent.
from the point of view of logical consequence.

Lastly, I think that it is also likely that we can prove this sort of equivalence in the opposite direction:

**Conjecture 3 (Weak consequence-equivalence).** Again, demonstration of the truth of Conjecture 1 would allow me to show that PCP-consequence can be weakened, so that it concentrates upon all the possible sets, $\Pi$ such that $\Sigma^{-\Pi}$ is classically-satisfiable. Once this is managed, then again the three approaches would all be equivalent from the point of view of logical consequence.

This last, a first-order version of a result from Bryson Brown (which deals with equivalencies between dialethic LP and a bivalent disambiguating approach)\(^{63}\), would suggest that the preservationist approach shares even more with the others than might first be thought. As such, then, confusion-preservation should be worthy of further examination. While its proof-theory is presently unguessed (although it would likely need to reflect the fact that at least predicate-confusion-preservation is classical in the face of classical sets), some work has been done on a semantics, so to allow confusion preservation the benefit of models in some ways resembling those of QLP and QLD. In any case, the preservationist project is interesting often just because it operates even where it lacks such features. Although some might want to deny confusion-preservation the name ‘logic’, I honestly see no good reason to do so, given that it deals with inferential connections and standards of argument. This suggests, to me at least, that what ‘logic’ is will turn out more various than has been generally thought, and may really only be bounded by limits of imagination.

**References**


\(^{63}\)Brown (forthcoming).


[18] —. Logic: One or Many?. Presented at Society for Exact Philosophy Conference (Lethbridge, Alberta, Canada, 1999).


