Necessity and Propositions

Tristan Grøtvedt Haze

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Faculty of Arts and Social Sciences
University of Sydney

This is to certify that to the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

Tristan Grøtvedt Haze
Acknowledgements

Thanks to my supervisor David Macarthur and my associate supervisor N.J.J. Smith for their assistance.

Special thanks to Libby McLean and my parents Gubbie and Jon for their love and support.
## Introduction

1. Subjunctive Necessity De Dicto

1.1. What is Subjunctive Necessity De Dicto?
1.2. What is Our Task?
1.3. Assumptions and Guiding Ideas
1.4. Conclusion

2. Some Existing Accounts of Necessity De Dicto

2.1. Pre-Kripkean Analyticity Approaches
2.2. Ersatz Possible Worlds Accounts
2.3. Modal Fictionalism
2.4. Kit Fine’s Essence-Based Account
2.5. Kment’s Counterfactual Account
2.6. Primitivism

3. Modal Realism

3.1. Exposition
3.2. The Ontological Objection
3.3. The Epistemic Objection
3.4. The Humphrey Objection
3.5. The Semantic Objection
3.6. The Contingent Totality Objection
3.7. The Motivation for Modal Realism
3.8. Conclusion

4. Sider’s Quasi-Conventionalism

4.1. Exposition
4.2. Five Objections
4.3. Conclusion

5. An Account of Subjunctive Necessity De Dicto

5.1. The Account Introduced
5.2. Some Applications of the Account
5.3. Inherent Counterfactual Invariance Further Clarified
5.4. Implication and its Role in the Account
5.5. The Account Reviewed
5.6. A Fallback Position
5.7. Objections and Replies
5.8. Conclusion

Appendix 1: Arguments for the Coextensiveness of Necessity and Apriority
Appendix 2: De Re Modality and Quantifying In

6. Propositions and Meaning

6.1. Preliminary Explanation
6.2. Being More Specific
6.3. Internal and External Meaning
6.4. Internal Meanings as Roles in Language Systems
6.5. External Meaning Further Explained
6.6. Names
6.7. Flexible Granularity
6.8. How This Account Fits With That of the Previous Chapter

7. Conclusion

7.1. Stepping Back
7.2. Indicative Necessity, Apriority and Analyticity
Introduction

Some propositions are not only true, but could not have been otherwise.

This thesis is about modality and the philosophy of language. Its centrepiece is a new account of the conditions under which a proposition is necessarily true in the above sense.

The primary motivation for this account has been my hunch that semantic considerations ought to come into the explanation of why a given necessary truth is necessary rather than contingent (more on this in Section 1.3. below), together with the fact that no existing, viable account does justice to this hunch.

The plan for the thesis is as follows. In Chapter 1, I specify the topic - the notion of subjunctive or metaphysical necessity de dicto, as isolated by Kripke - more fully and carefully than in the first sentence above, and also specify in more detail the problem or task with respect to the topic. I also make explicit some of my working assumptions and guiding ideas.

Then I turn to consider some other accounts of this notion. In Chapter 2, I briefly discuss six types of accounts which I will not be considering in great detail, indicating my reasons for not adopting them. These are: (i) pre-Kripkean analyticity approaches, (ii) ersatz possible worlds accounts, (iii) modal fictionalism, (iv) Kit Fine’s essence-based account, (v) Boris Kment’s counterfactual account, and (vi) primitivism. By explaining briefly why I do not favour these accounts I hope to help motivate my own approach.

In Chapters 3 and 4, I consider in more detail two accounts which I have judged to be the most important ones: David Lewis’s modal realism (defended principally in Lewis (1986)) and Theodore Sider’s quasi-conventionalism (defended principally in Sider (2011)).

These two accounts are important from the point of view of this thesis in different ways. Modal realism is important because it is so formidable and central to philosophical debate surrounding my topic, as well as being profoundly wrong. Sider’s account is important because it is onto something important but also makes a big mistake. My account will be seen to share a structure with Sider’s, but with a new illuminating notion in the place of the big mistake.

In Chapter 5, I turn finally to presenting, explaining and defending my account of subjunctive necessity de dicto. After having given the account and defended it from close to, I turn in Chapter 6 to a more extended aspect of its defence, namely the development of an approach to certain questions in the philosophy of language which can underpin it. My account of subjunctive necessity de dicto is an account of a property of propositions, and one of the key notions it employs is that of inherent counterfactual invariance, also a property of propositions. Thus, for my account to be worthwhile, it had better be the case that the idea of these things called propositions is a legitimate one, and it had better be the case that they are the sorts of things which could bear the posited property of inherent counterfactual invariance. By sketching an independently attractive account of propositions, names and meaning which fits well with the idea that propositions can have the special property of inherent counterfactual invariance, I hope to make a good case that these requirements can be fulfilled. Some may want to combine my account of subjunctive necessity de dicto with an account of propositions and meaning which differs on some points from mine. My aim is to show by example that there is at least one reasonable way of going.

I conclude in Chapter 7 by stepping back and saying something about the significance of the contents of this thesis, and offering some brief suggestions about indicative necessity, a priority and analyticity which could be taken up in further work.
That is the plan of the thesis. Before continuing to Chapter 1, a word on methodology is in order. The desiderata for the account of subjunctive necessity being sought in this thesis are specified in a general way at the outset (Chapter 1). The desiderata become more specific as I go on to consider, and set aside, other approaches to the target notion (Chapters 2 through 4). (A methodological approach on which theoretical desiderata are fully specified at the outset is not taken here, since getting a more detailed view of these is a substantial part of the work undertaken in this thesis.) The positive desiderata we begin with are twofold (and explained further in Chapter 1): first, the account should be illuminating and get under the hood of the target notion. Second, the account should do justice to the ‘semantic hunch’ explained in Section 1.3. It is also emphasized in Chapter 1 that we are not assuming at the outset that a good answer to our leading question ought to be given in non-modal terms, i.e. that it should be reductive.

Finally, while the main topic of this thesis is subjunctive necessity conceived as an attribute of propositions, debates about the metaphysics of propositions are largely eschewed. (Such debates could only be discussed at the expense of detail with respect to the main topic.) For the bulk of the thesis, the notion of ‘proposition’ in play is left intuitive; propositions say things about how things are, and are true or false. In response to potential worries about whether there are such things as propositions - or if there are, whether they are the sorts of things that could bear the property of subjunctive necessity and the property of inherent counterfactual invariance explained in Chapter 5 and appealed to in my account - I turn in Chapter 6 to the task of sketching a philosophical account of propositions on which these worries are allayed. By sketching this account I aim to show by example that there is at least one viable way of going here - i.e., that the account of subjunctive necessity which is the main focus of this thesis does not rely on leaving the notion of a proposition intuitive and unexamined. Further debate concerning our theoretical understanding of propositions must be left for another occasion.
1. Subjunctive Necessity De Dicto

1.1. What is Subjunctive Necessity De Dicto?

Our main topic is subjunctive necessity considered as an attribute of propositions. The purpose of this section is firstly to specify this topic in some more detail. The notion in question of course looms large in contemporary analytic philosophy, but it will serve us well and keep us grounded to rehearse to ourselves in as clear a way as possible a basic characterization of it. Following that, I will turn to specifying our problem or task with respect to the topic. Following that, I will state some assumptions and guiding ideas. I will finish with a few words in justification of pursuing the problem.

The key source for the notion of subjunctive necessity de dicto is of course Kripke's Naming and Necessity. It was there that our topic was (to the best of my knowledge) first clearly isolated and characterized. Priority aside, Kripke's characterization is not easily improved upon and has been very influential. (Regarding the notion itself, not its characterization: it is a very interesting historical question to what extent this notion was present in earlier thinking. Or to what extent similar notions were, and how they may relate to the present notion. I will make no attempt here to answer this.)

Kripke's starting-point in characterizing the notion is to remark that, while many (at the time he was speaking) seem not to differentiate between apriority and necessity, he certainly will not use 'a priori and 'necessity' in the same way (Kripke (1980), p. 34). He then, after emphasizing that the notion of apriority is an epistemological one and mentioning some issues which might arise with that notion, gives the following characterisation of necessity:

The second concept which is in question is that of necessity. Sometimes this is used in an epistemological way and might then just mean a priori. And of course, sometimes it is used in a physical way when people distinguish between physical and logical necessity. But what I am concerned with here is a notion which is not a notion of epistemology but of metaphysics in some (I hope) nonpejorative sense. We ask whether something might have been true, or might have been false. Well, if something is false, it's obviously not necessarily true. If it is true, might it have been otherwise? Is it possible that, in this respect, the world should have been different from the way it is? If the answer is 'no', then this fact about the world is a necessary one. If the answer is 'yes', then this fact about the world is a contingent one. (Kripke (1980), pp. 35 - 36.)

This should go a long way to giving us an acceptable grasp of the notion of subjunctive necessity de dicto. Kripke also says some things about the extension of the notion which may be of further help to this end. Before proceeding to that, however, I want to tighten up Kripke's characterization in a couple of ways, as well as emphasizing and de-emphasizing certain parts of it.

For one thing, note that Kripke moves freely here between talking of 'facts about the world' as well as things which can be called true or false, as the bearers of necessity. Later, he speaks also of 'states of affairs' and 'statements'. This is fine, but I want to make it clear that the topic of this thesis is the notion of necessity as it applies to things which can be called true or false: statements - or as I say, propositions. This is what I mean by 'de dicto' in 'subjunctive necessity de dicto'.¹ To be still

¹ I should emphasize that that is all I mean by 'de dicto' in 'necessity de dicto'. The term 'de dicto', and the contrasting term 'de re', are used in various ways in philosophy. It is especially important to realize that I count all attributions of necessity to propositions as attributions of necessity de dicto, even when those propositions are “singular propositions” about individuals - i.e., propositions attributions of necessity to which David Lewis would deploy counterpart theory to understand. (This aspect of Lewis's views is explained in Chapter 3.)
more precise about what propositions are - for a start, whether they are or involve sentences themselves, or just their meanings - will not be necessary until later.

Regarding the 'subjunctive' in 'subjunctive necessity de dicto': something I want to emphasize in Kripke's characterization is the way it cashes out necessity in terms of counterfactual scenarios - to use the language of some two-dimensional semanticists, scenarios considered as counterfactual, rather than scenarios considered as actual. That is, our topic is a notion which has to do with what could have been the case, rather than what could actually be the case. This could also be emphasized by calling our topic 'counterfactual necessity de dicto', but I use 'subjunctive', which comes from use of subjunctive constructions ('could have been') in specifying the notion in question, as this is better established in the literature.

Note that, since I have to name my topic often, and since 'subjunctive necessity de dicto' is quite a long name, I will often drop 'subjunctive'. (I will also, in the context of talking about propositions, drop the 'de dicto', i.e. I will not talk of a proposition being 'necessary de dicto', since the 'de dicto' is redundant here.) Dropping 'subjunctive' is something I do with some reluctance, as I am sympathetic to Chalmers' idea of the tyranny of the subjunctive - roughly, the idea that there is another philosophical notion equally deserving of the word 'necessity'. I would rather not contribute to the tyranny by calling my main topic just 'necessity de dicto', but I will do it anyway.

Something I want to de-emphasize in Kripke's characterization, on the other hand, is the way he classifies the notion of necessity he wants to talk about as a notion belonging to metaphysics. I do not think this is essential to grasping the notion in question: that can be done without any recourse to a notion of metaphysics. Kripke's use of the category of metaphysics here may be slightly useful in emphasizing that necessity de dicto is not an epistemological notion, but that point can be emphasized without a notion of metaphysics. Since we can easily get by here without invoking a notion of metaphysics, I think we ought to avoid invoking one. I will not argue the point at length here, but I suspect that invoking a notion of metaphysics may lead to some unhelpful prejudice about how necessity de dicto is best to be understood and analyzed (if it is to be analyzed) - or more to the point, how it is not to be analyzed. In particular, I worry that it may cause prejudice against accounts which crucially involve semantic considerations, by promoting a vague idea that necessity de dicto is "all about" how things are in the world, as opposed to having anything to do with language and thought.

Finally, Kripke's characterization should be supplemented with something about the sense of 'necessary' being unrestricted. To see this, consider an utterance like 'It is true that I stayed home yesterday. This couldn't have been otherwise, as I had to be there to let the electrician in'. This utterance may be true, but in that case the 'couldn't have been otherwise' part is not about necessity de dicto in the sense I am interested in - we are dealing with a contextually restricted range of ways things could have been. For instance, we are probably ignoring ways things could have been in which I never made the appointment with the electrician, or in which the appointment was on a different day, or in which I stop caring about having electricity. This supplementation of the Kripkean characterization has become customary. Witness Timothy Williamson in an interview:

Something is metaphysically necessary if it couldn't have been otherwise, in the most unrestricted sense. (Williamson & Antonsen (2010), p. 18.)

---

2 This terminology was introduced by Davies & Humberstone (1981). I make use of it later, in Section 5.3., when laying out part of my account of necessity de dicto.

3 You may think that this is the same as the point that subjunctive necessity de dicto is not to be understood epistemologically, but I'm dubious about that. See Section 7.2.

4 David Chalmers is perhaps the most prominent writer using 'subjunctive' in this way, e.g. in his (2006) and (2009).

Or Daniel Stoljar, referring to:

(... the completely unrestricted sense of possibility that philosophers sometimes call "logical" or "metaphysical" possibility (... (Stoljar (2006), p. 34.)

Or this terminological stipulation made by van Inwagen:

Modal terms will be used in their "metaphysical" or "unrestricted" sense (...). (van Inwagen (2015), p. 35.)

There is a wrinkle here, however. For some things philosophers say may seem to go against the propriety of characterizing our topic in this way. On the way of speaking I have in mind, there are necessities in the sense of our topic which are not necessary in some other sense - 'logically' or 'mathematically' or 'epistemically' for example. See, for instance, this passage in Salmon (2005):

Metaphysical modality is definitely not an unrestricted limiting case. There are more modalities in Plato's heaven than are dreamt of in my critics' philosophy, and some of these are even less restrictive than metaphysical modality. One less restrictive type of modality is provided by mathematical necessity and mathematical possibility. [...] Another type of modality less restrictive than metaphysical modality is provided by what is sometimes called 'logical necessity' and 'logical possibility', to be distinguished from genuinely metaphysical necessity and possibility, or necessity and possibility tout court. A proposition is logically necessary if its truth is required on logical grounds alone [...]. Although there is a way things logically could be according to which I am a credit card account, there is no way things metaphysically might have been according to which I am a credit card account. (Salmon (2005), p. 136)

But notice the contrast at the end of this passage between 'could be' and 'might have been'. Salmon is concerned here with what he calls 'the confusion between the generic notion of a way for things to be and the modal notion of a way things might have been'. According to Salmon, this confusion

is very probably the primary source of the idea that metaphysical modality is the limiting case of restricted modalities, that metaphysical necessity and possibility is the unrestricted, and hence the least restricted, type of necessity and possibility. For metaphysical necessity is indeed truth in all ways things might have been (modal, not generic), and metaphysical possibility is indeed truth in at least one way things might have been (modal, not generic). (p. 136.)

So, since we are explicitly talking about ways things might have been, it seems that Salmon would have no real disagreement after all with Williamson's succinct characterization of our topic, quoted above (except perhaps for some pragmatic disagreement about what to emphasize, or how best to use language to avoid potential confusions).

In any case, one thing that should be clear is that we are not dealing with a notion where certain contextually relevant matters of fact may be held fixed, as in the electrician example above. We are

---

6 I will not use the terms 'logical possibility', 'logical necessity', 'logical modality' and the like in this thesis. They are used in a bewildering variety of ways by different philosophers.

7 It seems like the usage of 'logical possibility' Salmon has in mind here is different from that alluded to by Stoljar above. It is perhaps telling that neither writer really uses the term himself, both instead alluding to how it is 'sometimes' used; the term seems to be a bit worn out at present. Again, I will not use it in this thesis.
dealing with a notion which applies to truths which 'could not have been otherwise' in a broad, unrestricted sense.

So much for the intensional characterization of the notion of necessity *de dicto*. Another thing which may help us grasp the notion is consideration of its extension - cases, and what types of cases there are. Most instructive in this way are cases lying outside the overlap of necessity and apriority. After giving his intensional characterization of the notion, Kripke goes on to say that he will be arguing that, in addition to being conceptually different, the categories of necessity and apriority are extensionally different: 'I will argue below that in fact they are not even coextensive—that necessary *a posteriori* truths, and probably contingent *a priori* truths, both exist.' (Kripke (1980), p.38.)

An aspect of the character of the notion of necessity *de dicto* is captured vividly in some of Kripke's intuitive appeals regarding the necessary *a posteriori*, in particular with the use of the phrase 'given that', and similar language. For instance, if I think some object I have encountered empirically, *a*, is the same object as I have encountered empirically in other situations, *b*, then - while I might conceivably turn out to be wrong, i.e. while it might turn out to be the case that *a* is distinct from *b* - *given that a* is indeed *b*, then *a* couldn't have been distinct from *b*; *'a = b'* is necessary.

Regarding the contingent *a priori*, perhaps the most straightforward and instructive type of case occurs when a name is stipulated to refer to whatever object satisfies some description, where the description is of a sort where an object satisfying it could have failed to satisfy it. So if I stipulate (following Evans (1979)) that 'Julius' is to refer to the inventor of the zip (if there was an inventor of the zip), then the proposition 'Julius invented the zip, if anyone did' is *a priori*: in virtue of the way I have set 'Julius' up to work, it just can't turn out empirically that Julius exists and yet didn't invent the zip after all. Now suppose that there is an inventor of the zip. In that case, the proposition 'Julius invented the zip, if anyone did', while *a priori*, is contingent: someone else could have invented the zip.

We will consider these cases again along with other sorts of cases (in Section 1.3. below and later in the thesis). For now, the point was just to highlight two sorts of cases which may help us grasp the notion of necessity *de dicto* and keep it distinct from notions it may be confused with.

### 1.2. What is Our Task?

We have now characterized our topic, first intensionally, by taking and modifying slightly Kripke's famous characterization, and then extensionally, by pointing to some striking cases. The next question we must address is 'What is the problem or task in relation to the topic, and why is it worth pursuing?'. My preferred way of articulating the problem with respect to the topic is this: under what conditions is a proposition necessarily true?

What this calls for, I take it, is an illuminating, or non-trivial, statement of what those conditions are. Failing that, a case must be made that no such statement of what those conditions are can be given - that is, a case must be made that necessity *de dicto* be taken as primitive.\(^8\)

One caveat regarding the requirement that the statement of conditions be 'illuminating, or non-trivial': I mean illuminating or non-trivial *given the present state of play in philosophy*. If my account, or some other account fulfilling this requirement is highly successful, then perhaps it will become so bound up with our very understanding of the notion of necessity *de dicto* that it will

---

\(^8\) Another option here would be to take a skeptical attitude to the very notion of necessity *de dicto*. While I hope that my account, and what I have to say in connection with it, may ultimately help to answer such skepticism, I begin with the assumption that there is a legitimate notion here. See Section 1.3. below for a fuller statement of this assumption.
come to seem trivial in some sense. Far from being unwelcome, I would consider that a great success for that account.

This is a pretty good start at articulating the philosophical question central to this thesis. It seems to be a necessary feature of a solution to our problem that it either gives an illuminating statement of conditions, or makes a case for taking the notion as primitive. But that may not be sufficient. What I am thinking of is the possibility of statements of conditions - 'if and only if' statements about when a proposition is necessary - which are true, and illuminating in the sense that they may be instructive, and yet do not satisfy us with respect to our question. For instance, consider:

(CF) A proposition P is necessary iff, for all propositions Q, if it had been the case that Q, it would have been the case that P.

This in fact is the guiding idea behind counterfactual accounts such as Kment's, which I consider briefly in Chapter 2. To be sure, Kment's account is ambitious and does not stop there - he goes on to attempt, among other things, a non-modal analysis of counterfactuals - but even as it stands, this statement is arguably true and non-trivial. It shows us, we might say, a connection between counterfactual conditionals and the universal quantifier on the one hand, and the notion of necessity de dicto on the other. And yet, taken by itself, this statement does not satisfy me as a solution to the problem I am trying to articulate. It seems to me that more can be done - that a more penetrating statement of conditions is still perhaps available, one which so to speak gets more under the hood of the notion of necessity de dicto - or again, failing that, some story about why no such thing is available.

One way of remedying this - one way of supplementing the description of our task - would be to employ some general notion which will distinguish the sort of statement of conditions we hanker for from those which will not satisfy. Some notion, for instance, of 'a good analysis'. That way, faced with a plausible and non-trivial statement of conditions under which a proposition is necessarily true, we can say 'Very well, but that's not a good analysis.' I think this is legitimate as far as it goes, but it raises difficult questions about what analysis is all about and what makes for a good analysis. Beyond intuitive talk about shedding light on the concept, penetrating the concept, or getting under the hood, I have nothing much of a positive nature to offer on this score - still, maybe this intuitive talk is nothing to be sneezed at.

Another way of supplementing the statement of our task would be to impose the requirement that our statement of conditions, if there is to be one, must not employ modal notions. That is, we must either reduce the notion of necessity de dicto to non-modal notions, or give some story about why this isn't possible. But this won't do. It is at once too stringent and not stringent enough; with respect to the first disjunct it is too stringent, while with respect to the disjunction as a whole, it is not stringent enough. Let me explain.

Requiring that our statement of conditions be non-modal is too stringent, inasmuch as a modality-involving statement may be satisfying and seem to get under the hood of the notion. I see no reason for ruling that out in advance. Furthermore, a non-modal statement may not be possible (or, perhaps clearer but more controversial, the whole idea of a non-modal statement of the conditions under which a proposition is necessary may be a confused one). It is true that those who are trying for a reductive analysis of modality have often taken necessity de dicto as a starting point, with the result that we see projects which might look like attempts to tackle just our problem, but which are actually aimed primarily at a distinct problem, that of reducing modality (or that of either reducing modality or making a case against trying to do so). But surely we may attempt to shed light on necessity de dicto without trying for a reductive analysis of modality. That we should be wary of conflating the problem of giving an account of necessity de dicto with the problem of reducing modality is an important theme of the present thesis and will be touched on again.
Requiring either a non-modal statement of conditions or a case for not seeking one is not stringent enough in the following sense. Suppose that, as I will argue, a statement of conditions can be given which has a modal element and constitutes a good account of necessity de dicto, a good solution to the problem I am trying to articulate. Suppose further that no non-modal reduction of the notion is to be hoped for. In that case, the task of either providing a non-modal reduction or a story about why none is to be hoped for may be completed - by way of the second disjunct - while passing over the good but modality-involving account. In that case, the problem I am trying to articulate and this other problem call for quite different solutions. (I would be very interested in hearing something insightful about why no non-modal reduction of necessity de dicto is to be hoped for, but I have not tasked myself with that in this thesis. I just want to motivate, present and defend my modality-involving account.)

So much for that requirement. Another requirement which naturally suggests itself is that our statement of conditions, if there is to be one, should avoid invoking the concept of necessity de dicto at any point. I am dubious about this. For one thing, there may be statements of conditions which are not satisfying, but which nonetheless fulfill this requirement - perhaps the counterfactual-based example above is of this sort. For another thing, there may be an account which violates the requirement but which is satisfying, at least to some extent - particularly if the propositions whose necessity is appealed to on the right hand side of the analysis form a restricted subset of all the necessary propositions. The account I will propose is, as I understand it, not circular or recursive - but I do offer, as a fallback position, a version of it which is.

So, in addition to an intuitive appeal to the idea of a good analysis, we have considered two possible general requirements with which we might supplement the statement of our problem, emphatically rejecting the first and casting doubt on the second. I also want to flag a general worry that urging some general requirement may not be quite truthful, and may be masking a more specific requirement lying behind our desire for an account.

A more specific requirement, one which has particularly motivated me, is that the account either do justice to, or convincingly discredit, what I call 'the semantic hunch' below in Section 1.3. - the idea that semantic considerations should come into our account.

For now, let us leave our characterization at that. It may not be perfectly clear and precise, but the question 'Under what conditions is a proposition necessarily true?', supplemented with:

- the intuitive idea of a good analysis (shedding light, penetrating, getting under the hood), and
- the negative points that, at the present stage of investigation at least, it does not seem that our answer need be modally reductive or non-recursive, and perhaps also
- the more specific requirement just alluded to regarding the semantic hunch,

will turn out, I think, to be enough with which to work fruitfully.

Before moving on to consider critically some existing accounts of necessity de dicto, I want to end this preliminary chapter by indicating some of the assumptions and guiding ideas of this thesis.

1.3. Assumptions and Guiding Ideas

Legitimacy and Interest. Probably the most important and most continuously operative assumption I am making in this thesis is that the notion of necessity de dicto is a coherent, legitimate notion. Note that this is not an assumption about how important or useful the notion is, only that it makes sense. None of the other assumptions and guiding ideas seem to be such that their failure would
render my account of necessity *de dicto* completely worthless, but this one seems non-negotiable.\(^9\)
While it is non-negotiable, and is assumed as a starting point, this is not an assumption I leave completely unsupported. My account itself, for instance, may help to allay the worries of someone who might otherwise have been a skeptic about the legitimacy of the notion. Another of my guiding ideas to be discussed shortly - that we shouldn't get too hung up on unclear or borderline cases - may also help with skeptical worries, as I will make clear.

It could be said that I am also assuming, in a practical sense, that the notion of necessity *de dicto* is interesting or important. Otherwise, why bother? I am indeed inclined to think it is extremely interesting, but perhaps that's not as essential as one might think. Perhaps it is enough that the philosophical community currently finds it interesting. If the present work were to show, by revealing more about it, that it is in some ways not as interesting or important as it seemed, then that would also be something.\(^10\)

Another assumption I have been working under is that the notion of necessity *de dicto* is non-empty: there are some necessarily true propositions. However, a communication from David Ripley in 2013 made me realize that this assumption too is not as essential as one might think; Ripley expressed tentative approval of my proposal as a good account of what it would take for a proposition to be necessary, but also indicated that he suspects that no proposition makes the cut. So it seems that even if my working assumption of non-emptiness is wrong, that does not automatically make my account wrong or uninteresting.

Cases. In light of Kripke's work, I take it as given that there are necessary *a posteriori* propositions and contingent *a priori* ones. The denial that one or both of these compound categories has instances does not conflict with the account I will put forward in any very direct way, as far as I can tell, but the idea that such propositions exist has been a major guiding principle in this work. My account was developed with their accommodation being regarded as an important requirement.

I will now summarize my guiding ideas with respect to various important types of cases.

The clearest kind of case of the necessary *a posteriori* is that of identity statements involving proper names, such as 'Hesperus is Phosphorus' (or these made conditional on existence, as in 'If Hesperus exists, it is Phosphorus'). That Hesperus is Phosphorus cannot be known *a priori*, but given that it is, Hesperus could not have failed to be Phosphorus.

Other central cases of the necessary *a posteriori* include statements about the underlying natures of what philosophers have called 'natural kinds' - for example, 'Cats are animals' and 'Water is H\textsubscript{2}O'. With these cases, however, many are inclined to think that things are not so simple - for instance, that these terms ('cats', 'water') may naturally be understood in such a way that those sentences come out contingent, and even that sort of understanding is more faithful to ordinary usage - cf. Putnam (1990) and Wikforss (2013). My attitude to such worries is relaxed. It doesn't really matter to me in this work whether the terms in question are ordinarily meant in such a way that the sentences come out necessary, or which cases they are and are not so meant - such empirical linguistic issues may be left to one side. It is enough that we can, in thinking about these matters, seize on particular understandings of these terms on which these sentences *do* come out necessary *a posteriori*.

\(^9\) However, as I will emphasize, the treatments of certain questions in the philosophy of language sketched late in this thesis should be able to survive the failure of this assumption.

\(^10\) Compare Wittgenstein's remark at the end of the preface to the *Tractatus*: 'I am, therefore, of the opinion that the problems have in essentials been finally solved. And if I am not mistaken in this, then the value of this work secondly consists in the fact that it shows how little has been done when these problems have been solved.' While, as I have said, I do think my topic is interesting and important, I also suspect there may be an element of this in play as well, especially with respect to some of the more metaphysical ideas surrounding my topic.
Regarding cases of the contingent *a priori*: as already mentioned above in Section 1.1., the clearest cases are perhaps those where it is stipulated that a name is to refer to the actual referent of some definite description. For example, the case in Evans (1979) of `Julius invented the zip` (or perhaps `Julius invented the zip, if anyone did`), where the name `Julius` has been introduced with the stipulation that it is to refer to the inventor of the zip, if there is one.

Another prominent case is that of the metre stick (discussed by Wittgenstein (1953) and then Kripke in *Naming and Necessity*): if we call it `S` and regard it as the standard of measurement for metres, we can say that `S` is one metre long, or perhaps `S` is one metre long at time `t`, where time `t` is some particular time, or perhaps `S`, if it exists at time `t`, is one metre long then, is contingent *a priori*. It must be said that this case seems to involve a pretty drastic simplification of the real linguistic practice of fixing units of measurement. But in the spirit of our relaxed attitude to worries about the empirical veracity of the claim that the sentences `Cats are animals` and `Water is H₂O` are in their ordinary meanings necessary *a posteriori*, we may happily grant that there is a type of case here, since it seems relatively clear that we *could* have a linguistic practice which really does work in the required way - where the length of some stick is regarded as playing a role as the sole standard for certain statements (which we might call *statements of length*, by analogy with our actual practises). We might imagine the length of the stick at a certain time `t` as being the sole standard, or perhaps less problematically in some ways (avoiding problems about memory and knowledge of the past), we might imagine the practice working such that it is always the then-present length of the stick which we work with in testing such statements. We could also imagine an analogous practise centring on a colour sample.

Another sort of apparent case of the contingent *a priori* worth mentioning is `I exist`. As Kripke (2011, p. 304) says (albeit along with some doubts), this sort of case has a different flavour from those touched on above.

Regarding cases where the necessary/contingent and the *a priori/a posteriori* line up: prominent sorts of cases of the necessary *a priori* include arithmetical propositions like `1 + 1 = 2`, propositions established by *a priori* reasoning such as `First-order logic is undecidable`, propositions of colour exclusion such as `A colour patch cannot be both blue and red`, and propositions like `All bachelors are unmarried` (here again our relaxed attitude to empirical complications must be remembered; we may be simplifying our idea of how `bachelor` really works here, but that doesn't matter since we clearly could have a word which by definition applies to all and only unmarried men). Prominent sorts of cases of the contingent *a posteriori* include statements about the locations of particular things or people, phenomenological statements about what can be seen (or heard, or felt, etc.), and generalizations like `Everyone who came on time today had blue jumpers on`.

**Vagueness or Indefiniteness.** I am emphatically not assuming that the notion of necessity *de dicto* is non-vague, or that it is clear for every proposition - even if we know that it is true and know the basis of its truth - whether it is necessary or not. Indeed, I strongly suspect it is vague. Kripke (1980, p. 36) remarks of the notions of apriority and necessity that `[b]oth concepts may be vague. That may be another problem.` So I am in good company here in not assuming that the notion of necessity *de dicto* isn't vague. I would suggest further that we needn't even think of this as a problem. Sure, vagueness itself can be philosophically puzzling, but that this notion is vague need not itself seem like any kind of problem or surprise - we may instead regard it as completely natural and unsurprising.

**Unclear Cases.** A methodological remark related to the above: when trying to get grip on the notion of necessity *de dicto*, and when trying to analyze it, we should be wary of getting bogged down in worries about unclear cases, and try instead to focus on clear ones. We should be especially wary
of trying to force particular answers on unclear cases. Compare: when trying to give someone a grasp of the notion of tallness, it is better to work with examples of people who are definitely tall, or definitely not tall. To start insisting on certain judgements about more borderline cases is not to the point, and may make the whole business seem dubious.

Taking a wrong attitude to unclear cases, I want to suggest, may mislead us about what sort of account we should look to give of necessity de dicto. Also, such an attitude (which is, unfortunately, not warned against by Kripke, and possibly even encouraged) may be contributing to skepticism about the legitimacy or coherence of the notion.

Even some of the canonical cases Kripke adduces in Naming and Necessity strike me as not particularly clear cases. That is, they strike me as borderline or disputable cases. The sorts of cases I have in mind are those of the table - could it have been made of ice? (Kripke intuits that it couldn't.) And the Queen: could she have been born of different parents? (Kripke intuits that she couldn't.) You might think these are clear cases (whether or not you agree with Kripke's verdict that they are necessary; you could also think they're cases where he's clearly wrong). Still, whatever you think the unclear cases are, I suggest not worrying about them too much too early in our inquiries into necessity de dicto.

It is worth remarking that the notion of intuition plays a curious role here. It is one thing to have an "intuition" about a clear case - a judgement you're inclined to make that just seems correct and isn't theoretically derived. I want to suggest that the intuiting can take on a different, more dubious character, when we come to unclear cases. In the clear cases, you might say, we proceed confidently and coolly. With the unclear cases, it can seem like a kind of hearkening or special receptivity is supposed to be needed. This may seem, so to speak, occult. And if it doesn't put us off the notion of necessity de dicto altogether, this may yet mislead us about what sort of account we should look to give of it.

Names and Meaning. Now for two guiding ideas which fall more squarely in the philosophy of language, having to do with names. Firstly, I accept Kripke's contention that ordinary proper names are rigid designators - that they designate the same object in all possible worlds in which that object exists. I accept this on the basis of the arguments in Naming and Necessity, now standardly classified into three categories: modal, epistemic and semantic. This hangs together with my acceptance of the thesis that identity statements involving names are necessary (and my consequent acceptance that, since such statements are often a posteriori, such cases provide especially clear instances of the necessary a posteriori); that such statements are necessary can be argued for from the assumption that names are rigid (see Schwarz (2006)).

Secondly, I reject Millianism, whether it is understood as the thesis that names have no meaning, or as the thesis that the meaning of a name is its referent (if it has one). Kripke's flirtation with

---

11 Kripke (1980), pp. 113 - 114.
13 These questions, strictly speaking, are de re and not about the status of propositions. However, there are corresponding questions about propositions: are 'This table is not made of ice' and 'The Queen was born of George VI and Elizabeth Bowes-Lyon' necessary?
14 One thing about the Queen case which has always bothered me is what I call the fish argument. This argument works by iterating the supposed necessity of origin; if the Queen is necessarily the child of her actual parents, and they are necessarily the children of their parents, then we seem to be forced to conclude that the Queen is necessarily the descendant of some fish which she is in fact descended from - call him Colin. That is, there is no possible world involving the Queen where Colin isn't also around. This seems dubious to me. (Baumann (2012) shares essentially this worry.)
15 In Chapter 5, f.n. 14, I will suggest an alternative approach to defining rigidity, based on the notion of a genuine counterfactual scenario description developed in Chapter 5.
16 As far as I can tell, the classification originated in Salmon (1982).
Millianism\textsuperscript{17} has long seemed to me an unfortunate aspect of his work, and it is important to note that Millianism in no way follows from the assumption that names are rigid designators (as Kripke is well aware). In Chapter 6, when I develop an alternative account of names, I will also try to say some diagnostic things about why Millianism has attracted philosophers. For now, I will just say that my most basic reasons for rejecting Millianism include:

- A desire to be able to affirm, in a straightforward way, the intuitively very plausible view that 'Hesperus is Phosphorus' has a different meaning from 'Hesperus is Hesperus', and to be able to lay this difference at the door of the semantic contribution of the names, and
- a conviction that Millianism makes it very difficult to give a philosophically satisfying treatment of the problems surrounding singular existence statements and their negations.

Note that none of this means I hold to a form of descriptivism about names, whether it be the simplest sort held by Frege and Russell (or at least attributed to them by Kripke (1980)), a cluster type view like Searle's in his (1958), wide scope descriptivism (see Dummett (1973, p. 110)), rigidified descriptivism or a two-dimensionalist approach (Jackson (1998), Nelson (2002), Chalmers (2006)), or something else.\textsuperscript{18} Rather, I seek a middle way between descriptivism and Millianism.

These - that names are rigid and that Millianism is false - are my most definite, relevant guiding ideas as regards the philosophy of language. But I want to emphasize that these are not in any direct way mandated by my account of necessity \textit{de dicto}. I would be happy if a philosopher accepted this account even if they persisted in being a Millian or had problems with the claim that names are rigid. I think such a philosopher's views would, on account of their Millianism or their non-recognition of the rigidity of names, be subject to serious challenges, but it doesn't seem to me that accepting my account of necessity would exacerbate their difficulties very much, if at all. It is when I turn from elaborating my account of necessity \textit{de dicto} (Chapter 5), to the development of an approach to the philosophy of language which can (not must!) underpin my account (Chapter 6), that these guiding ideas come into their own and accounts upholding them are sketched.

\textit{The Semantic Hunch}. I do not expect all readers to share this hunch, especially in view of more indirect and impressionistic aspects of the influence of Kripke's work, but I hope some will share it. In any case, it plays a key role in the present work.

The hunch, roughly speaking, is that considerations of meaning should get into the act at some point in the explanation of what makes some true propositions necessary, and some contingent. (A more specific version of the hunch is that the meaning or nature of the propositions themselves ought to have something to do with their having the modal status they have.)\textsuperscript{19}

\textsuperscript{17} According to Lycan (2008, p. 49), however, Kripke does not believe that names are Millian.

\textsuperscript{18} That said, the untenability of any form of descriptivism is something I am much less confident about than the untenability of Millianism, since 'descriptivism about names' as I am understanding it here is a somewhat open-ended category, and I have not attempted a detailed study of all its extant forms. This is especially true for the case of two-dimensionalism, which I will not consider in this thesis. (I do not see my proposals about meaning as ruling out two-dimensionalism.)

\textsuperscript{19} I say 'meaning or nature' because I am trying to be as theory-neutral as possible about propositions at this point. In Chapter 6, I will propose an account of propositions which takes them to be, basically, sentences together with their meanings. On this picture, it makes sense to talk of the meaning of a proposition, but - unlike with sentences considered in abstraction from any particular use or interpretation of them - this meaning will be something inherent to the proposition. Assuming that framework, I may also talk of the nature of the proposition here; it doesn't make much difference. On the other hand, if 'proposition' is used for meanings of sentences, then the hunch may be expressed by mentioning the natures of propositions, or even just the propositions themselves. In any case, the hunch remains expressible in a rough way as: meaning gets into the act in the story of why a necessary proposition is necessary rather than contingent. (Of course, if we regard propositions as meanings, certain conceptions of propositions will fit poorly with this hunch, e.g. a Russellian one on which the proposition that John is tall is just something like a complex of
Of course, old-fashioned analyticity-based accounts (to be discussed in Chapter 2) satisfy this requirement, but I think we should be firmly convinced by Kripkean considerations that such accounts are non-starters with respect to metaphysical or subjunctive necessity. (I consider such accounts, despite their being non-starters, on the grounds that it is worthwhile to see how they fail.)

Since Kripke, many have simply abandoned the hunch. Lewis’s modal realism, for instance, explains a proposition’s being necessary entirely in terms of what goes on outside the proposition in reality, so to speak. Other post-Kripkean accounts, such as Sidelle’s (1989) conventionalism and Sider’s (2011) quasi-conventionalism, seem to aim at doing some justice to the hunch, but (as I will argue at some length in the case of Sider) they face serious problems. With still other accounts, such as modal fictionalism (discussed briefly in Chapter 2), it is not immediately clear whether they do justice to the hunch or not.

Abandoning the hunch is of course an option, but it seems to have been very widely had by pre-Kripkean philosophers - at least, insofar as they can be said to have had a hunch about a notion which hadn’t been clearly isolated yet; perhaps we should say the hunch concerned a family of notions of which the notion of subjunctive necessity de dicto is a member. Furthermore, it arguably has a lot of naïve or pre-theoretical appeal. So I think philosophers who want to abandon the hunch, or who want to propose accounts which do no justice to it, have some explaining to do. Why, if they are right, were so many philosophers attracted to the old-fashioned accounts, and why, now that Kripke has clearly isolated subjunctive necessity de dicto from quite differently-behaved notions (such as apriority and, if this is a different thing, indicative necessity\(^2\)), do philosophers continue to develop accounts which aim to do justice to the hunch, i.e. which involve semantic considerations at some crucial point?

1.4. Conclusion

Let us sum up what has been done in this chapter. First we have delineated our topic, subjunctive necessity de dicto, as a property which attaches to propositions when they are true and could not have been otherwise, in an unrestricted sense. We have articulated (with the caveat that the articulation may not be perfectly precise) our task with respect to the topic: to analyze the notion of necessity de dicto by giving an illuminating account, which gets under the hood of the notion, of the conditions under which a proposition is necessary de dicto. We have emphasized that this account need not reduce necessity de dicto to non-modal notions. Finally, we have touched on some of the assumptions and guiding ideas of the thesis, principally: that the notion of necessity de dicto is legitimate and non-empty, that the necessary a posteriori and contingent a priori exist, that the notion of necessity de dicto is probably vague and that’s OK, that we shouldn’t get hung up on unclear cases, that names are rigid and Millianism is false, and the semantic hunch: that considerations of meaning should get into the act at some point in our story about what makes a proposition necessary.

In the following three chapters we will look critically at some existing accounts of necessity de dicto.

\[^2\text{In Chapter 7 I briefly consider apriority and indicative necessity, and suggest that the two be conceptually distinguished, with indicative necessity being understood semantically (in terms of the notion of internal meaning developed in Chapter 6) and used to explain apriority understood epistemologically.}\]
Introduction and Chapter 1 References


2. Some Existing Accounts of Necessity De Dicto

In this chapter, I will criticize some existing accounts of necessity de dicto. I will be somewhat brief in places here, the point being more to orient the reader and give a sense of where I am coming from than to comprehensively evaluate these accounts. For more comprehensive treatment I have chosen Lewis's modal realism or Sider's quasi-conventionalism, which I consider in the next two chapters.

2.1. Pre-Kripkean Analyticity Approaches

Analyticity is - or at least, is supposed to be - a property propositions possess when they are, in some sense, true in virtue of meaning. Analyticity approaches to explaining the modal statuses of propositions are important due to how widespread and appealing they were, and because it is important to see how they fail with respect to subjunctive necessity de dicto, not least of all because even when applied to that notion, they contain what I consider to be a grain of truth: semantic considerations should come into our account of what makes a proposition subjunctively necessary. (This is the semantic hunch introduced in Section 1.3.)

First of all, it must be said that these approaches were probably not actually supposed to apply to the notion of subjunctive necessity de dicto which is our topic. As Sider says in his (2003, p. 202):

Note that it is probably inaccurate to describe these philosophers as identifying what I have been calling metaphysical necessity with analyticity. It would be better to say that these philosophers rejected, or did not possess, the contemporary concept of metaphysical necessity; their claim was that analytic necessity is the only sensible sort of necessity in the neighborhood.

Accordingly, my discussion here is not to be thought of as a critical assessment of an actual contender. Analyticity approaches to our topic are, I think, non-starters. But it is worth seeing how these approaches fail when applied to subjunctive necessity de dicto.

What about the 'only sensible sort of necessity' claim mentioned by Sider? While it is understandable that some might have held such a view in the past, before Kripke had clearly pointed out the notion of subjunctive necessity de dicto, it would be much harder to maintain now.

Two major historical examples of analyticity approaches to modality are Carnap (1947) and Ayer (1936). Carnap's account is more technical than Ayer's, and embodies a special conception of philosophical analysis, but in their fundamental ideas about modality they are both quite explicit and certainly closely akin. Both philosophers regarded analyticity as more philosophically tractable on its own terms than both necessity and apriority, and they explicate both of these latter in terms of analyticity (or, in Carnap's case, an 'exact' surrogate). Let us now take a brief look at these approaches, in order to extract the core claim of analyticity approaches.

Carnap begins by proposing that modal concepts should be clarified by being 'correlated' with semantic concepts:

Various systems of modal logic have been proposed by various authors. It seems to me, however, that it is not possible to construct a satisfactory system before the meanings of the modalities are sufficiently clarified. I further believe that this clarification can best be achieved by correlating each of the modal concepts with a corresponding semantical concept (for example, necessity with L-truth). (Carnap (1947), Preface, v.)

Carnap defines L-truth in terms of his notion of a 'state description'. Going into the details of
Carnap’s definition of L-truth is not necessary for our purposes. We may focus on Carnap’s preliminary ‘Convention 2-1’, which reads as follows (using ’p’ in place of Carnap's choice of sentential variable):

2-1. Convention. A sentence \( p \) is L-true in a semantical system S if and only if \( p \) is true in S in such a way that its truth can be established on the basis of the semantical rules of the system S alone, without any reference to (extra-linguistic) facts. (Carnap (1947), p. 10.)

Immediately after this Carnap comments:

This is not yet a definition of L-truth. It is an informal formulation of a condition which any proposed definition of L-truth must fulfil in order to be adequate as an explication for our explicandum. (Carnap (1947), p. 10.)

In view of this, showing that the concept of necessity \textit{de dicto} fails to line up with this informal formulation suffices to show that, either, it also fails to line up with Carnap’s full, official notion of L-truth, or that notion is inadequate by Carnap’s own lights.

Now for a word on Ayer’s account. Ayer does not sharply distinguish the notions of apriority and necessity. As Sider (2003, p. 199) bemusedly notes, the index entry in \textit{Language, Truth and Logic} for ’necessary propositions' reads simply: ’See A priori propositions'. When Ayer uses the word ’necessary' he often follows it with ’certain', and he often talks as though the mark of a necessary proposition is that it 'cannot be confuted by experience' (Ayer (1936), p. 80). This makes it doubtful that he meant by 'necessary proposition' what I mean in this thesis (though perhaps in some moods he did - perhaps his usage was inherently confused, although to be sure he could not be said to be uniquely guilty of this). \footnote{Later on, in Section 7.2., I will make some suggestions about analyticity, apriority and what might be called ’indicative necessity’ (recall our touching on Chalmers’ idea of the tyranny of the subjunctive in Chapter 1). What Ayer says about ’necessity’ and apriority will be seen to agree in large part with what I suggest - and what he says about analyticity will be disambiguated so that it may be right on one reading, but a narrower notion of analyticity enabling us to call some propositions ’synthetic a priori’ will be suggested as well. For now, however, we are - unfairly in a way - just treating Ayer’s account of what he calls necessity, and of apriority, as an account of subjunctive necessity \textit{de dicto}.} In any case, his account of a priori propositions and what he called necessary propositions was that these were 'tautologies'. He explains, ’I use the word “tautology” in such a way that a proposition can be said to be a tautology if it is analytic; and I hold that a proposition is analytic if it is true solely in virtue of the meaning of its constituent symbols' (Ayer (1936), p. 185).

So much for Carnap and Ayer. I take the core claim of analyticity approaches as applied to subjunctive necessity \textit{de dicto} to be: a proposition is necessary iff it is analytic. \footnote{One caveat: some philosophers, such as Carnap himself, who should be thought of as holding analyticity approaches may not want to make serious use of the concept of necessity, or even the concept of analyticity, on the grounds that these are insufficiently clear. Carnap, as we have seen, had a conception of philosophical analysis on which problematic concepts are to be replaced by ’exact’ ones - ’explanations’. For this reason, discrediting the claim that a proposition is necessary iff it is analytic does not automatically discredit Carnap’s explication regarded as an explication of the concept of subjunctive necessity \textit{de dicto}. But the explication faces a dilemma: either it is to be regarded as simply discounting the behaviour of the concept of necessity which shows it to fail to line up with that of analyticity (or an ’exact’ surrogate for it), in which case it may be rejected as unacceptably revisionary, or it is to be regarded as attempting to do justice to the behaviour, in which case it fails.}

I want to distinguish three sorts of objections which may be levelled against analyticity approaches regarded as accounts of subjunctive necessity \textit{de dicto}. Firstly, some analyticity approaches involve the notion of truth by convention - there are certainly many hints of this in Ayer. This notion may be attacked - and was, most famously by Quine. Secondly, the notion of analyticity itself may
be called into question, quite apart from anything to do with convention (Quine is also the main figure here). Thirdly, the notion of analyticity may be granted, and seen not to line up with that of necessity de dicto. I will argue that the third sort of objection is the most powerful.

**Approach 1: Attacking Truth by Convention**

In his celebrated (1936) paper ‘Truth by Convention’, Quine attacks the very idea that some things are true solely in virtue of linguistic conventions, or in other words, true by definition. He wants to suggest that the idea does not survive scrutiny, i.e. that it is incoherent:

( . . . ) [D]evelopments of the last few decades have led to a widespread conviction that logic and mathematics are purely analytic or conventional. It is less the purpose of the present inquiry to question the validity of this contrast than to question its sense. (Quine (1936), p. 250.)

His main argument is a regress argument. Definitions, Quine says, are to be thought of as replacement rules for expressions. So definitions enable us to ‘transmit truth’ (Quine (1936), p. 250): a true statement, together with a definition of one of its terms, yields another true statement: that which we obtain by substituting the *definiens* for the *definiendum* in the original statement. In a relative sense, we may call the resulting statement true by definition - but the original statement's truth is an essential factor here. In connection with this idea of truth by definition *relative* to some truth or body of truths, Quine mentions the case of mathematics: it seems, from the work of Frege, Russell and Whitehead, and others, that there is a case to be made that the truths of mathematics are true by definition relative to the truths of logic. But then what of logic itself? For an analyticity approach to analyzing necessity to work, it needs to be the case that *all* necessary truths, including logical truths, can be held to be true in virtue of meaning. So, is there a notion of truth by convention available which is not relative in the above sense, and which applies to logic itself? At this point Quine moves to consider a form of convention which differs from definition (in his sense of ‘definition’):

[...][I]f we are to construe logic also as true by convention, we must rest logic ultimately upon some manner of convention other than definition: for it was noted earlier that definitions are available only for transforming truths, not for founding them. (Quine (1936), p. 259.)

The form of convention Quine then considers works by stipulating a truth-value for all sentences conforming to some specification. For example, taking the schema:

(1) If $p$ then $q$ then if $q$ then $r$ then if $p$ then $r$. (Quine (1936), p. 262.)

We can introduce the convention:

(I) Let all results of putting a statement for '$p$', a statement for '$q$', and a statement for '$r$' in (1) be true. (Quine (1936), p. 262.)

The problem with this, according to Quine, is that these stipulations are *general*, and that we need to use logic itself to derive particular consequences from them.

In the adoption of the very conventions (I)-(III) etc. whereby logic itself is set up, however, a difficulty remains to be faced. Each of these conventions is general, announcing the truth of every one of an infinity of statements conforming to a certain description; derivation of the truth of any specific statement from the general convention thus requires a logical inference, and this involves us in an infinite regress. (Quine (1936), p. 270.)
Now, perhaps Quine's argument here does show something (whether it does is not very important for our purposes here). As for whether it casts doubt on the very idea of truth by convention, it may be criticized for failing to be exhaustive of the types of convention which might make things true. For instance, other understandings of the nature of definitions may be available, on which they are not or not simply - stipulated replacement rules, and Quine's argument may not apply if they are employed. Perhaps more worryingly, it seems there are linguistic conventions which are neither definitions nor general stipulations about truth-values. 

However, let us not worry too much about criticizing Quine's particular argument that idea of truth by convention is suspect; perhaps there are further problems with the idea. The most serious problem with Quine's argument, or any argument against the idea of truth by convention, as an objection to analyticity approaches to analyzing necessity de dicto is that the target - the notion of truth by convention - is not essential to analyticity approaches. Analyticity approaches need not make any claims about convention or definition.

One way to bring this out is to reflect that meanings on the one hand, and definitions and linguistic conventions on the other, are quite different things. The latter may be thought of as things which link at least some expressions to their meanings (perhaps indirectly, by linking them to other expressions) - and so, not as meanings themselves. In that case, the core claim of analyticity approaches - that necessary truths are in some sense true in virtue of meaning - is quite different from the claim that they are true by (or in virtue of) convention or definition.

In light of all this, I think objections to analyticity approaches to necessity de dicto which work by attacking the notions of truth by convention or truth by definition are relatively weak and should not be relied upon. Analyticity approaches - whatever the case may be about certain historically prominent instances - are not wedded to these notions. (There is sometimes a tendency, in contemporary discussions of pre-Kripkean thinking about modality (for example Sider (2003)), to overlook this and focus too narrowly on particular analyticity approaches which happen to incorporate these notions.)

The same cannot be said for the second sort of objection to analyticity approaches, which we will now consider.

Approach 2: Attack Analyticity Itself

Analyticity approaches are, on this sort of objection, attacked at their very root: the idea of analyticity - truth in virtue of meaning - itself. The classic attack on analyticity is of course Quine's (1951). This is well-trodden ground, so I will be very brief in summarizing Quine's argument and my response, both to it and to other possible attacks on the very notion of analyticity.

Quine argues that the notion of analyticity is suspect by considering some natural ways of explaining or defining it, and arguing that these just lead us around in a circle (or something like a circle⁴), through notions that are no better understood than analyticity. One potential problem with this sort of argument is that perhaps there are other ways of defining analyticity which Quine hasn't been able to think of. In my judgement though, the more fundamental problem with this sort of

---

³ To see this, consider indefinable or semantically primitive expressions. (I am supposing that these exist.) Suppose that 'N' is such an expression in some language. 'N', we will want to say, has a meaning. And it seems hard to deny that there is at least an element of convention behind 'N' having the meaning that it does rather than some other meaning, and the fact that 'N', rather than some other sign, is used to express this meaning.

⁴ What Quine actually says is: 'Our argument is not flatly circular, but something like it. It has the form, figuratively speaking, of a closed curve in space'. (Quine (1951), p. 29.)
argument is that, even if it does consider all the best prospects for defining the notion, it fails to rule out the possibility that the notion is in good order but cannot be non-circularly defined.

On this matter I share Kripke's sentiment as expressed in f.n. 17 of his (2008):

In other contexts (not this one), I have heard some people object to an example on the grounds that “it depends on the analytic-synthetic distinction” (really the a priori-a posteriori distinction), which Quine supposedly refuted. (...) Anyone who says that Quine or anyone else showed that such distinctions make no sense is simply incredible.

For a recent positive, extended treatment of the analytic/synthetic distinction see Russell (2008).

**Approach 3: Exhibit Necessary A Posteriori Truths as Counterexamples**

The really effective way to discredit analyticity approaches, I maintain, is simply by cases. Once the stage has been set, so that the notion of subjunctive necessary truth is clearly isolated (see Chapter 1), the cases do the rest. Cases of the necessary a posteriori are the crucial discreditors. They discredit analyticity approaches, both directly and via the epistemological contrast between a priori and a posteriori. They do it directly, since it is obvious (on reflection, if not immediately) that propositions like 'Hesperus is Phosphorus' and 'Cats are animals', or whatever examples of the necessary a posteriori you prefer, are not true in virtue of meaning. They also do it via the epistemological notions: there are necessary a posteriori truths, and no a posteriori truth could be true in virtue of meaning.

To sum up: the first approach - attacking the notion of truth by convention - fails to be sufficiently general, because this notion is inessential to analyticity approaches. The second - attacking the notion of analyticity itself - is completely on target, but has not been carried out convincingly. The third, counterexample-based approach is much stronger. It is, I maintain, the simple intuitive compellingness of cases of the necessary a posteriori, not any sophisticated abstract argument, which shows that analyticity approaches are non-starters with respect to our topic.

**2.2. Ersatz Possible Worlds Accounts**

By 'ersatz possible worlds accounts', I meant accounts which, like Lewis's modal realism which we will consider in Chapter 3, proffer Leibnizian biconditionals as analyses of necessity and possibility de dicto, but which, unlike Lewis's modal realism, appeal to abstract objects to play at least something of the role that concrete worlds play in modal realism. Prominent developers of the theory of ersatz worlds include Adams (1974), Plantinga (1992), Roy (1995) and Stalnaker (1976). (Some sophisticated accounts which involve ersatz worlds, such as Kment's, which I consider briefly below, I do not count under this heading.)

The rationale for such accounts is easy to understand: Leibnizian biconditionals seem like they could offer us a way of accounting for facts about whether propositions could have been true, in terms of whether they are true with respect to something which in fact exists. Lewis's account holds out this promise, but flies in the face of common sense while also generating puzzles and theoretical problems. Perhaps we can avoid these pitfalls by appealing to entities which we have less trouble believing in, which seem more like things we could know about, and which do not seem as irrelevant to modal facts. Ersatzism also seems fit to avoid the contingent totality objection I will raise against modal realism (see Section 3.6.): as the ersatz players of the world-role are not the same sorts of things as the actual world, an ersatz can happily accept that reality as a whole could have been different - the actual world could have been the way one of the ersatz “worlds” represents things to be.
Lewis's (1986) is the classic source for criticism of ersatz approaches to modality. One of the highlights of Lewis's critique is the objection (found on Lewis (1986), pp. 150 - 151) that linguistic ersatsizm - which is perhaps the leading kind - has to appeal to a modal notion to filter out "impossible worlds"; for a set of sentences to be, or represent, a possible world, it had better be the case that all of its members taken singly could be the case, and it had also better be the case that any two members of such a set could both be true together. (Lewis also objects that, insofar as the ersatzist needs to appeal to, not just the sentences that make up their sets, but to what they jointly imply, and since this implication seems to be best understood modally, modality creeps in in a second way. I am less impressed by that. I discuss the corresponding issue for my account, which also appeals to implication, in Section 5.4.)

The version of this objection that I advocate is different from Lewis's, and it is worth spelling out how. For Lewis, the objection is simply that modal notions are required here. Aiming as he is for a reductive account of modality, which happens to take de dicto necessity and possibility as its starting points, this is all there is to it. But this can't be my objection to ersatzism construed as an account of necessity de dicto - construed, that is, as an attempt at the task specified in Chapter 1, rather than as an attempt at a reductive account of modality. For involving modal notions is no objection to an account construed as such - indeed, the account I will put forward involves modality as well.

My version of the objection is partly that the modal notions which seem to be needed for the ersatzer to rule out impossible "worlds" are too similar to the one being analyzed; the relevant notion of consistency seems to amount to something like the conjunction of the propositions in question being possible. But there is more to it than that: you could, in principle, have an account which uses even the very notion being analyzed - a circular or recursive account - which is still informative and constitutive of a theoretical advance. So the real point, which appears when we realize that the ersatzist has to appeal to modal notions to have an account, is that linguistic ersatzism fails to illuminate or penetrate the concept necessity de dicto.\(^5\) It just takes us circuitously through these big, idealized, abstract entities, to no explanatory purpose - or so I say. (I think a similar point may apply to non-linguistic forms of ersatzism, but I will not argue for that here.)

2.3. Modal Fictionalism

Like ersatzism, modal fictionalism can be understood as beginning with Leibnizian biconditionals and modal realism, and trying to have the apparent benefits of these things without the difficulties which attend modal realism. The basic idea in this case is that a proposition is possible iff, according to the fiction of modal realism, it is true at some possible world, and that a proposition is necessary iff, according to the fiction of modal realism, it is true at all possible worlds. Rosen (starting in his (1990) and continuing in his (1993) and (1995)) is a pioneer of the approach. It has been developed more recently by Woodward (in his (2008) and (2011)). For a survey, see Nolan (2016).

One of the puzzling things about modal fictionalism is understanding how it differs from linguistic ersatzism. The simplest and most standard versions do differ in one obvious way: they deal with a fiction - some kind of representation - according to which there are many worlds, whereas ersatzism deals with many representations, each one representing a single world. (See Armstrong\(^5\))

\(^{5}\) Here it may be asked: what constitutes illumination, or penetration, if not reduction? I have no general answer, but I think the case of recursive definitions in logic and mathematics shows that an account can be illuminating or penetrating without being reductive, and thus that being illuminating or being penetrating is not the same thing as being reductive. Furthermore, even if you think that in the case of necessity de dicto any illuminating or penetrating account would have to be non-recursive, there is still room for an account which is reductive in the sense that the very notion of necessity de dicto does not appear in the account, but which is non-reductive in the sense that some other modal notion or notions do appear in the account. The account I give in Chapter 5 is of this kind.
(1989) for a version which brings fictionalism closer to ersatzism in this respect - on this version, there are two levels of fiction: a series of smaller fictions and one big fiction.) But this still leaves it obscure to what purpose fictionalism differs from ersatzism, and it still leaves it obscure what role the concept of fiction is playing: why might it seem helpful to talk of fictions instead of just propositions which aren't (all) true? Furthermore, if there is an important difference, then difficult questions about the nature of fiction may arise.

The difference just mentioned between linguistic ersatzism and at least the simplest and most standard versions of fictionalism also gives rise to a serious technical difficulty, known as the Brock-Rosen objection (see Brock (1993) and Rosen (1993)). Fictionalism, in its canonical form, says that a proposition is necessary iff, according to the fiction of modal realism, it is true at all possible worlds. Consider now the proposition 'There are many concrete worlds'. According to the fiction of modal realism, there are many concrete worlds, and this will, according to modal realism, be true at all those worlds. Thus fictionalism seems to deliver the verdict that this proposition is necessarily true. But one of the primary virtues of fictionalism was supposed to be that it does not require, as the modal realist does, a plurality of concrete worlds. Thus modal fictionalism, at least in its canonical form, seems to undermine its own rationale. This objection has given rise to complex and ongoing debate: it has been argued by Noonan (1994) that this objection fails even for the canonical form of fictionalism, so long as we adhere strictly to the method of reconstructing modal claims given by Lewis (1968). Others have granted that the canonical form of fictionalism should be abandoned, but sought to amend it. These amendments have in turn been argued to fail, giving rise to analogous objections. For a good overview of the situation see Nolan (2016, 3.1).

So, modal fictionalism faces difficulties surrounding what exactly it comes to, or how it should best be understood, as well as a serious technical difficulty in the form of the Brock-Rosen objection. But what seems to me the biggest mark against it has yet to be mentioned, and this is a fairly simple, intuitive thing. It is that, intuitively, it just doesn't seem right that talk about the modal status of propositions is tacitly about, or tacitly goes via, any sort of fiction. Note that this isn't the same as saying that ascriptions of modal status to propositions do not seem to be fictional - it would be a misunderstanding of modal fictionalism to think it implied that they were fictional. Rather, the point is that it doesn't seem like fictions are getting into the act at all. The very concept of fiction seems irrelevant. In this connection it is also worth noting that, for this to be a good criticism, it doesn't have to be the case that all possible formulations of modal fictionalism are false. There may be a way of getting true fictionalist biconditionals, but I would suggest that any such true biconditionals fail to constitute good analyses of the concepts of possibility or necessity de dicto - fail to shed any real light on these concepts.

### 2.4. Kit Fine's Essence-Based Account

The starting point for Kit Fine's account of when a proposition is necessary is his celebrated (1994) argument that essential properties are not all and only the necessarily possessed ones: there are necessarily possessed properties which are not intuitively not essential to their possessor, the primary example being Socrates's being a member of his singleton set, the set {Socrates}. Another example Fine gives is Socrates's being distinct from the Eiffel Tower. Both properties - being a member of {Socrates} and being distinct from the Eiffel Tower - are necessarily possessed by Socrates, but intuitively they are not essential to him. They don't have enough to do with Socrates as he is in himself to count as essential to him, we might say.

Then Fine, on the basis of such examples, argues that we should not try to explain essence in terms of necessity. He then makes the further suggestion that we should try to do things the other way around and explain necessity in terms of essence. In particular, he suggests that a proposition

---

6 See (Fine (1994), pp. 4 - 5).
7 See (Fine (1994), p. 5).
is necessary iff it is ‘true in virtue of the nature of all objects whatever’:

Indeed, it seems to me that far from viewing essence as a special case of metaphysical necessity, we should view metaphysical necessity as a special case of essence. For each class of objects, be they concepts or individuals or entities of some other kind, will give rise to its own domain of necessary truths, the truths which flow from the nature of the objects in question. The metaphysically necessary truths can then be identified with the propositions which are true in virtue of the nature of all objects whatever. (Fine (1994), p. 9.)

(Fine seems to take natures and essences to be the same, or at least intimately related.)

As far as I know, Fine has not publicly developed this particular claim much, although he has developed surrounding ideas concerning varieties of necessity (in his (2002)) and ‘in virtue of’ (in his (2012) and other papers). Despite this, Cameron in his influential (2010) survey paper on accounts of the ‘grounds of necessity’ considers this account of Fine’s as one of three main contenders (the others are modal realism and conventionalist approaches).

I will now try to state briefly why I do not find this account attractive. (Some of this may carry over to powers theories such as those of Jacobs (2010) and Pruss (2011), which I am not considering in this thesis.) First of all, note that Fine makes three moves:

1. Arguing that the essential properties are not just the necessarily possessed ones.
2. Suggesting that essence should not be accounted for in terms of necessity.
3. Suggesting that we should go the other way round.

I am very sympathetic to (1) and suspect that it constitutes a lasting and enviable contribution. Regarding (2) I have serious doubts: I am attracted to the idea that what is missing from the necessary but not essential properties is that they are in some sense not intrinsic to their bearers, and think there are good prospects for an account that says, roughly, that essentiality = necessity + intrinsicality. This idea has been developed, impressively in my view, by Denby in his (2014). Denby considers some objections to this way of going, and his responses are formidable (although I think that in some cases other, possibly better responses are available, but I will not go into that here).

Accordingly, regarding (3) I am unenthusiastic to begin with, since I suspect (2) - which was part of the journey to (3) - to be misguided. Furthermore, I do not find the account to be very clear or illuminating. One reason for this is that the notion of essence seems to call out for explanation just as much as the notion of necessity (cf. in this connection Hofweber’s (2009) criticisms of what he calls ‘esoteric metaphysics’). But it’s not just that: the analysis as a whole seems to me to be less clear than any of its parts taken singly - it seems to me that much of the unclarity of the account comes, not from the ideas composing it taken separately (‘essence’/‘nature’, ‘in virtue of’, ‘all objects whatever’), but from the way they are combined.8

---

8 To explain a bit further: you may be skeptical about some of the uses of these ideas in philosophy, but it is hard to deny that all three of them - ‘essence/nature’, ‘in virtue of’ and ‘all objects whatever’ - have some kind of meaning, some kind of life. For all of them, there are some clear, or at least compelling, cases where they appear. But, I want to say, Fine's proposal here is not one of these clear or compelling cases for any of them. I am reminded here of Wittgenstein’s criticism of Ewing’s proposal that ‘Good is what it is right to admire’ (although what follows may not be identical with the criticism of Fine’s proposal that I have just made):

The definition throws no light. There are three concepts, all of them vague. Imagine three solid pieces of stone. You pick them up, fit them together and you get now a ball. What you’ve now got tells you something about the three shapes. Now consider you have three balls of or lumps of soft mud or putty - formless. Now you put the three together and mold out of them a ball. Ewing makes a soft ball out of three pieces of mud. (Wittgenstein (1986), p. 42.)
Beyond worries about clarity and explanatoriness, I am moved by two further (related) worries. The first might be called 'the excess natures objection': for many or even most necessities, to say that they are true in virtue of the natures of all objects whatsoever is overkill. Hesperus is Phosphorus, and necessarily so. But do we need the natures of all objects in place in order for Hesperus to be Phosphorus? Why does the nature of my laptop have to get into the act of making Hesperus identical to Phosphorus? Aren't the natures of everything minus my laptop sufficient for making Hesperus identical to Phosphorus? And if so, surely we can run the argument again and again, deleting objects from the essentialist base for that proposition. And this seems to be quite widely true of necessities - the example of 'Hesperus is Phosphorus' isn't particularly special. Going along with the idea that metaphysical necessity is grounded in natures, it seems that different metaphysical necessities are grounded in different (collections of) natures. Now, what follows regarding Fine's account? A strong conclusion we might draw is that Fine's account apparently doesn't allow for this, since it says that the ground of all metaphysical necessities is the nature of all objects. Therefore it's false. A weaker conclusion we might draw is that, since Fine's account gives no information about which natures get in on the act for which necessities, it can't be the full story about the conditions under which a proposition is necessary. It's incomplete, or not a good analysis.8

The other worry is more conditional or defeasible: it is that the account does not do justice to the semantic hunch. If we suspect, as I do, that a story about what makes necessary propositions necessary can be told which appeals at a crucial point to the meanings or natures of propositions, then Fine's account should make us suspicious for neither doing this, nor giving us reason to think it can't be done.

Additional objections to Fine's approach are suggested in Cameron's aforementioned (2010).

2.5. Kment's Counterfactual Account

Kment's (2006) offers a new and impressively detailed analysis of the concept of necessity construed as an attribute of propositions. One of the striking things about the account is its multifacetedness; counterfactual conditionals play a role - the guiding idea being that, if something is necessarily true, then it would have been true no matter what else had been the case. Ersatz worlds - in particular, maximal consistent sets of sentences - play a role, in cashing out what makes the counterfactuals true. And Finean ideas about essences or natures play a role, in separating the metaphysically possible ersatz worlds from the impossible ones. So it is at once a counterfactual account, an ersatzist account, and an essentialist account.

The amount of detail is impressive, and the multifacetedness of the account can, I think, be said to be rhetorically advantageous given the contemporary state of play. We have had certain detailed but fundamentally bold and simple - that is, less multifaceted - accounts around for decades now, and all face formidable problems. The idea that, if we are to obtain a good account of necessity, a number of fairly separate considerations or leading ideas need to be woven together in a new, subtle way, is looking more plausible than ever.

One of Kment's leading ideas is that a proposition is necessary iff it would have been true no matter what had been the case. In symbols, □p ↔ ∀Q (Q □→ p). Kment notes that this is not new:

8 Jacob Archambault has suggested to me that this objection should not lead us to the conclusion that Fine's suggestion is false, since Fine's understanding of 'in virtue of' is such that it obeys weakening - i.e. that from 'A in virtue of B' we can conclude 'A in virtue of B and C', for any true 'C'. In that case, my objection is just that the task we have set in Chapter 1 with respect to the notion of necessity de dicto does not seem to be met by Fine's suggestion, true as that suggestion may be given a weakening-obeying understanding of 'in virtue of'.
The core idea of the account I will propound has been foreshadowed (though not
developed) by Davis Lewis in the 1970's in his book on counterfactuals, and in a
posthumously published paper by Ian McFetridge. In the last couple of years, Marc Lange,
Timothy Williamson, Christopher Hill, and I have worked out similar ideas independently of
each other. (Kment (2006), p. 8.)

Before continuing, I will now briefly explain why I am discussing Kment's account here, rather than
Lange's, Williamon's or Hill's. Lange's (2005) account focuses on logical truth and the question of
what unifies different varieties of necessity (also a concern of Kment's), rather than on subjunctive
or metaphysical necessity in particular. Williamson's (2005) does not go as far as Kment's into the
analysis of the counterfactuals which are used to define necessity. This is in keeping with
Williamson's focus in the paper in question being more on the epistemology of modality than the
analysis of modal concepts or an account of the metaphysics of modality. Hill's (2006) too is
primarily concerned with epistemology, but he also maintains that 'it is possible to explain the
metaphysical modalities reductively in terms of the subjunctive conditional' (Hill (2006, p. 219).
However, he does not go on to account for the latter. He works with Lewis's well-known (1973)
account - because, he says, he thinks it is largely correct - but says that 'there are other theories of
subjunctives that would serve my purposes almost as well' (Hill (2006), p. 220).

Kment goes further than these three authors, in that he advances from the insight about necessity
being definable in terms of counterfactuals to an account of the latter which employs linguistic
ersatz "worlds" both impossible and possible. Then, to filter out the undesired worlds, i.e. the
intuitively impossible ones, he uses the picture familiar from Lewis's (1973) of a nested system of
spheres containing worlds which, as you move out to more inclusive spheres, get more and more
different from the actual world, and invokes a Finean essence-based conception, supplemented by
the idea of a mathematical truth, to pick out the right sphere. The supplementation is due to a point
of difference between Fine and Kment. Kment writes (after giving an argument which I will not
rehearse):

I think that not all necessary truths owe their modal status to the fact that they flow from the
essential truths about things. Many mathematical truths, such as the proposition that 2
exists, do not. The most natural thing to say about them is that their modal status is simply
grounded in the fact that they are mathematical truths. It seems that there is more than one
non-modal property that can ground the special modal status of a metaphysical necessity: some propositions are metaphysically necessary because they flow from the essential
properties of things, others are necessary because they flow from the mathematical facts.
This suggests the following, revised version of Fine's account: any necessary truth owes its
special modal status to the fact that it is underwritten by the mathematical and/or the
conditioned essential facts about the world. I will use this view as my working account of
the features that ground the modal status of metaphysically necessary propositions. (Kment
(2006), pp. 267 - 268.)

The above is far from a full exposition of Kment's account, but for space reasons I will leave it
there.\footnote{In a more recent and even more ambitious work (Kment (2014)), aimed at explaining the origins and role of
modal thought as well as analyzing the notion of necessity, Kment favours a notion of 'metaphysical laws' to
do the work of picking out the sphere of metaphysically possible worlds.}

A major problem with Kment's account is that, drawing as it does on Finean essentialist ideas, it
inherits the problems of Fine's account (some of which were gone into above).

Furthermore, the special appeal to the notion of mathematical truth, while the need for it might be
justifiable *assuming* a broadly Finean approach, is intuitively unappealing. Intuitively, I want to say, the notion of necessity is more unified and does not *involve* a notion of mathematical truth. Of course, you might have a broadly Kmentian account which does not disagree with Fine here as Kment does, but such an account is in trouble if Kment is right that mathematical truths about the existence of numbers make trouble for the Finean approach.

Finally, although the multifacetedness of the account is impressive in a way, and perhaps rhetorically advantageous, I suspect that ultimately the account fails for being too baroque and not being explanatory with respect to the semantic hunch outlined in Section 1.3.\(^1\) Something less baroque, which if possible does justice to the semantic hunch, should be preferred, at least from the point of view of wanting an illuminating account of necessity *de dicto* which doesn't necessarily have to reduce it to non-modal notions. That is what I seek to provide in Chapter 5. (Of course, there is a big difference in aim between my account and Kment's, namely that Kment *does* want to avoid modal notions in his analysis. On the other hand, there is also a big difference between Kment's attempt to avoid modal notions and what we might call more ideologically austere attempts, such as Lewis's modal realism (to be discussed in Chapter 3) and Sider's quasi-conventionalism (Chapter 4). These latter, we might say, try to reconstruct modal notions out of materials which do not seem modality-like at all - a critic might express this by saying that they try to pull off a philosophical magic trick. Kment on the other hand draws on a Finean notion of essence, which does seem modality-like.)

2.6. Primitivism

Primitivism about necessity *de dicto* is the view that no non-trivial analysis of the concept of necessity *de dicto* is possible - or, in a more pragmatic version, the view that we should not try for such an analysis.

It is very important to distinguish primitivism about necessity *de dicto* from other primitivist views concerning modality, for example:

(1) The view that at least some modal concept needs to be taken as primitive - i.e. not analyzed, or

(2) The view that no modal concept can be reduced to non-modal concepts.

(1), or rather views that plausibly imply (1), have been defended by deRossett (2005) and Wang (2013).\(^2\) (1) is compatible with giving, as I do, an analysis of necessity *de dicto*: perhaps it is some other modal concept, or class thereof, which needs to be taken as primitive. So is (2): we can give, as I do, an analysis of necessity *de dicto* that involves a modal component.

While I will not argue for them in this thesis, I am very sympathetic to (1) and (2). But I am dead against primitivism about necessity *de dicto*, because it precludes an analysis which seems to me to be true and illuminating. (Before I had the ideas which led to my analysis, I was for a while more sympathetic to it.) I am not aware of an author explicitly endorsing primitivism about necessity *de dicto* specifically, although I would not be surprised to see this. Obviously many authors take necessity *de dicto* as primitive in practise, in the sense of using the notion without a non-trivial

---

\(^1\) Not being explanatory with respect to the semantic hunch is of course a failing by my lights rather than Kment's; I do not mean to suggest that being explanatory with respect to the semantic hunch was among his desiderata.

\(^2\) These authors prefer more metaphysical formulations, or versions, of modal primitivism. deRosset glosses the modal primitivism he defends as the view that 'necessity and possibility are part of the fundamental structure of the universe, and some modal claims would appear in any basic, overall description of that structure' (deRosset (2005), p. 1). Wang glosses the modal primitivism she defends as 'the view that metaphysical modality cannot be reduced to something entirely non-modal' (Wang (2013), Abstract).
definition or analysis, but that is of course not the same as holding the view that no such thing is possible.

What can be said against primitivism about necessity de dicto? One thing which may help is simply to point out, as I have done above, that its denial is perfectly compatible with (1) and (2). People may, with good reasons or just good instincts, be attracted to (1) or (2) without reflecting that the denial of primitivism about necessity de dicto is compatible with these.

Another point I want to make is that we should think of primitivism about necessity de dicto as a methodological last resort. It may not be mistaken in any way we can see short of having an analysis in hand or on the horizon, but I propose that we should not adopt it with confidence until we have considered plenty of types of analysis. My account, which has not been on the scene historically, should for example be given a hearing.

So, the main and most fundamental reason why I reject primitivism is just that I have an analysis which I think is correct. I can also argue that it is wrong without wheeling in my whole analysis, in the following way: surely, all necessary truths are true, and so perhaps we can give the analysis:

A proposition is necessary iff it is both true and has P,

where P is cashed out as the property which distinguishes necessary from contingent truths. Of course, it may be replied that this talk of P is not something we have much of a grasp of. Perhaps that's fair enough, but in Chapter 4, when I consider Sider's quasi-conventionalism, I will consider a schema which Sider's account and my account share, a schema which (I will urge) constitutes an important step forward:

A proposition is necessary iff it is, or is implied by, a proposition which is both true and C.

If we can motivate this schema, can't we then say that it shows primitivism about necessity de dicto to be false?

As I hope to make clear in Chapters 4 and 5, it would not be easy to maintain that we can't get a grasp on condition C. (Of course, just calling it 'C' might leave us feeling that something remains to be articulated - but that's good news for me and bad news for the primitivist.) I think that, with a bit of practise perhaps, anyone who has a grasp of necessity de dicto can attain to a working grasp of condition C - they can judge cases, and what's more, these judgements follow their own distinctive pattern, different from, but systematically related to, that of judgements of necessity. Furthermore, this minimal schematic analysis seems to shed light on why, with the previous proposal, the condition P might not seem very graspable or unitary.

I think these considerations suggest that primitivism about necessity de dicto itself is in bad shape. But these considerations leave primitivism about condition C untouched - and that I also want to reject; I think we can give an account of it. And whether we have an account in hand already or not, surely primitivism about condition C should, like primitivism about necessity de dicto, be regarded as a methodological last resort.

To sum up this section: I suggest that we should treat both primitivism about necessity de dicto and about condition C as methodological last resorts. Furthermore, I suggest that primitivism about necessity de dicto is even worse than primitivism about condition C, since the former precludes even the theoretical advance (shared by Sider's and my account) embodied in the 'C'-involving

---

13 Thanks to N.J.J. Smith for this phrase.
14 More-or-less; a couple of qualifications to the claim that Sider's account embodies (Schema) are given in Section 4.1.
analysis above.

Chapter 2 References


3. Modal Realism

3.1. Exposition

Central to modal realism are the Leibnizian biconditionals,

- A proposition is necessary iff it is true in all possible worlds.
- A proposition is possible iff it is true at some possible world.

These tie attributions of necessity and possibility to quantificational statements about possible worlds. Different philosophical accounts which use these sentences differ over what sorts of things possible worlds are taken to be, and over the role given to the Leibnizian biconditionals. (With typical forms of modal fictionalism, the biconditionals are typically augmented with an 'According to F' operator, where 'F' names a fiction.) The distinctive marks of modal realism, setting it apart from other philosophical uses of the Leibnizian biconditionals, are that it takes possible worlds to be of the same kind as the actual, concrete world we live in, that it takes the Leibnizian biconditionals to be true all by themselves (no fiction operator required), and that it takes these to constitute analyses of the modal notions appearing on the left hand sides.

The chief proponent and developer of modal realism, David Lewis, intends it to be a reductive account of modality - so his theory of possible worlds must be spelled out non-modally. Accordingly, the 'possible' in 'possible world(s)' on the right hand sides of the biconditionals is not supposed to be taken as anything more than part of a conventional, historically familiar way of referring to the worlds which do the work in his account.

Such is the theory of modal realism in broad outline. Its characteristic commitments may be summed up in one sentence as 'There are other worlds, and every way our world might have been is a way some world is' (Lewis (1986), p. 2).

Before moving on to consider objections, let us consider in a preliminary way three finer points about the theory.

One finer point concerns the individuation of worlds. As Lewis phrases the question, 'What makes two things worldmates? How are the worlds demarcated one from another? Why don't all the possibilia comprise one big world? Or, at the other extreme, why isn't each possible neutrino a little world of its own?' (Lewis (1986), p. 70). Lewis's answer to this is: spatiotemporal relatedness. 'Whenever two possible individuals are spatiotemporally related, they are worldmates. If there is any distance between them - be it great or small, spatial or temporal - they are parts of one single world.' (This gives rise to an objection - the island universes objection - based on the idea that we should not in our analysis of modality rule out the possibility of a world with multiple spatiotemporally unrelated "universes". I will not consider this objection at length, but see Lewis (1986, p. 71), Bricker (2001) and Vacek (2013).)

The second finer point concerns the treatment of propositions about particular individuals, and how they are to be evaluated with respect to worlds other than our own (or more generally, worlds other than the one from which the propositions in question are being evaluated). To begin with, note that general statements pose no corresponding difficulty. Going along with the modal realist's doctrine that there are other worlds, a question like 'Is "All swans are white" true at all worlds?' seems to have a straightforward meaning (at least given the familiar point that we want to hold fixed the meaning of the sentence in question when evaluating the proposition with respect to other worlds). But if we ask 'Is "John is white" true at all worlds?', where John is some actual swan named 'John', the question arises: does John himself exist at any of the other worlds?
The two different answers we might give to this question correspond to different forms of modal realism. If we answer in the affirmative, we get what is called modal realism with overlap. If we answer in the negative, get what is called modal realism without overlap. The canonical form of modal realism, David Lewis's as developed in his (1986), is of the latter sort.

In order to enable us to evaluate propositions about particular individuals with respect to other worlds in the framework of modal realism without overlap, Lewis developed a theory of counterparts. To evaluate 'John is white' at some world W, we as it were look at that world and select the closest counterpart to our this-worldly swan John, and then consider whether that swan is white. If so, we say that 'John is white' is true at W. This approach has been felt to be damagingly counterintuitive, giving rise to an objection originated by Saul Kripke called the Humphrey objection, which we consider below in Section 3.4.

The third and final finer point I want to note concerns the issue of what, if anything, modal realism has to say about the extent or range of the worlds - what worlds are there, and what are they like? As Lewis saw the matter, it was incumbent on him to provide principles which so to speak "generate" sufficient worlds, so that there is one for every possibility. To this end he proposed a principle of recombination, but he admitted that this was inadequate (Lewis (1986), p. 92). More recently, it has been questioned whether any such principles are needed for the theory qua analysis of modality (see Cameron (2012)).

Note that modal realism is obviously free of the chief defects of analyticity approaches - the modal realist analysis does not push us toward the conclusions, implausible ever since Kripke, that necessary truths are true in virtue of meaning, or that they are all a priori. This is one of the things which, together with the boldness and clearness (at least in a certain sense) of the theory, makes it such a serious contender given the present state of play.

We will now move on to consider a list of objections to modal realism, some of which we have just alluded to. This is far from an exhaustive list - modal realism has occupied a central place in philosophical debate about the analysis of modality, and a huge amount has been written about it - but it contains what I have judged to be the most important objections. My ultimate conclusion will be that the most serious objections are very serious indeed, and devastating when taken together.

3.2. The Ontological Objection

The ontological objection to modal realism is simply the objection that it doesn't seem plausible that all those other worlds exist - i.e., one for every single way things might have been. It is easy to think of this as the primary problem for modal realism. I think it is safe to say that it is the most glaring problem.

I do think the ontological objection has considerable power. However, I want to caution against putting too much weight on it. To do so would be to miss a philosophical opportunity. I think it is certainly wrong, for instance, to think of it as the intuitive problem with modal realism, all other possible criticisms being "merely technical" by comparison. To be sure, we do not ordinarily think there are lots of other worlds. But we should beware of letting this feature of modal realism, that it implies that there are (at least, in conjunction with plausible assumptions about the extent of possibility - cf. Cameron (2012)), obscure the fact that, even if we did believe in lots of other worlds, we may still have a hard time believing modal realism as an analysis of modality (or more narrowly, as an analysis of necessity de dicto). In my view, the semantic objection and the contingent totality objection, considered below, are the more fundamental intuitive problems with the theory.
So, my basic position is that the ontological objection is strong but should not monopolize our attention. I offer four reasons for thinking this:

(1) **Missed opportunity.** The issue of the ontological implications of the theory is, following Lewis (who was following Quine - see Lewis (1986), p.4), often put in terms of cost and benefit. On this construal, the ontological objection is the objection that the ontological cost of modal realism is **too high.** To work with this construal (imperfect as it arguably is), the following analogy may help to illustrate why we should not put all the stress on the ontological objection. Suppose we are interested in building and driving cars. Someone outlines a radical new design for a car, and says that they can build it for one billion dollars. We could never spend one billion dollars on a car, and so we refuse. We have thus avoided going broke trying to pay for something we can't afford, but we haven't learned anything. The whole affair will not help us one iota to build and drive cars. It would serve us well to look at the plans, and see whether we could learn something - something to imitate (compare modal fictionalism), or something to avoid.

(2) **Potential distortion via bias.** Somewhat more speculatively, I suspect that overemphasizing the ontological objection may not just be a missed opportunity, but may actually work, in conjunction with a natural bias, to harm our thinking about modality. The bias I am thinking of is that of taking cost - or to distance ourselves from the metaphor a bit, difficulty - as an indicator of value. That is, we may fall into thinking that, since modal realism makes such big demands on us ontologically, it must have something going for it. This is foolish in the present case, but quite understandable and forgivable, and clearly an instance of what is sometimes a good heuristic (only here, as in plenty of other cases, the heuristic fails). Thinking this way about modal realism threatens to distort our thinking about modality and the analysis of necessity **de dicto.**

(3) **Presumptuousness about what there isn't.** The ontological objection could be said to require a certain presumptuousness on our part concerning things we may just not know much about. 'What do we know about whether there are other worlds?', we might think, when in an open-minded mood. I suggest that we might know more about what we mean by modal claims than we do about **that,** in which case perhaps the semantic objection below is stronger. (Similarly for the contingent totality objection: it seems to me we might know more about whether or not reality as a whole could have been different than we do about whether or not there are lots of other worlds.)

(4) **Rhetorical weakness.** Another reason for not getting too hung up on the ontological objection is more rhetorical and has to do with the history of the dialectic. Lewis, the leading proponent of modal realism, was perfectly upfront in acknowledging the seriousness of the ontological objection. It would be a hard sell to do otherwise, and I am not aware of any subsequent philosophers who have denied that the ontological objection has some weight. The argument from modal realists has, it seems, always been that this cost is offset by benefits. Because of this, focusing narrowly on the ontological objection - saying, as it were, 'I don't care what you say, I just won't accept that there are all those other worlds' - while it might not be an unrespectable way of thinking for one's own sake, is suboptimal both from a rhetorical as well as a purely philosophical point of view. Focusing on the purported benefits of modal realism might lead to something more powerful.

I think there are good prospects for this. Ultimately, I think we can argue roughly as follows: let's grant for the sake of argument that accepting the ontology of modal realism could be worth doing if we got, say, a satisfying theory about the meaning of modal expressions, or a satisfying account of how we have modal knowledge, and if the theory were to respect at least our most central modal judgements. However, we get none of this: the semantic theory is highly problematic and counter-intuitive in various ways, the epistemology of modality becomes more difficult (going, we might say, from merely puzzling in a vague way to acutely paradoxical), and the theory predicts what seem to be the wrong truth-values for certain statements which are central to our modal beliefs. In the sections which follow I will try to build such a case.
To sum up: although this section has in large part been devoted to de-emphasizing the ontological objection, let me conclude it by reiterating that I do think it is pretty compelling, and adding something about the way it compels (for me at least): there is something noteworthy about where the ontological objection seems strongest - namely, when it comes to absurd or ridiculous possibilities. Regarding the idea of other worlds in general, or the idea of other worlds similar to ours, or worlds like ours but lacking certain forms of order (e.g., lacking life, or solar systems), I am, at least in some moods, inclined to think 'What do I know?'. But when I consider, say, worlds with talking donkeys, or worlds where there is only one person, who is just like me, but he says hello non-stop for five billion years and then turns into a cat, I get a strong feeling that reality almost certainly doesn't contain anything like that. I am not prepared to give an underlying reason for this feeling, but that doesn't make it any less strong.¹

3.3. The Epistemic Objection

Modal realism is epistemologically problematic. If the facts about what is necessary and what is contingent are to be understood in terms of the goings on in disconnected universes, how can we know such facts? Canonical sources for this objection are Richards (1975) and Skyrms (1976).

I think it is helpful, in assessing how much of a problem modal realism has on the epistemic front, to distinguish three versions of the epistemic objection:

1. The Benacerraf-style puzzle: how, if modal realism is true, do we know the modal statuses of propositions, if we do not have causal contact with the other worlds?
2. The minimal intuitive version: it seems like the goings on at other worlds aren't the sort of thing we know about. But we seem to know the modal statuses of some propositions. This seems incompatible with modal realism, so we should reject modal realism.
3. The theoretical challenge: how do we account for the epistemology of modality if we embrace modal realism?

Lewis, after floating the basic objection in a loose way, goes on to say that it 'echoes Benacerraf's famous dilemma for the philosophy of mathematics' (Lewis (1986), p. 108). According to this dilemma (see Benacerraf (1973)), we either interpret mathematics in a way which parallels our typical way of interpreting the rest of language, where the truth conditions of mathematical statements involving what look like referring terms indeed deal with objects referred to by those terms. In that case, we face the problem of how we know any mathematical truths, since the objects in question are ones which it would seem we have no causal contact with. Or, we give a different interpretation to mathematical statements which doesn't assign truth conditions which require us to know about things we have no causal contact with, assigning other truth conditions instead. But if we do this, says Benacerraf, 'we do so at the expense of failing to connect these conditions with any analysis of the sentences which shows how the assigned conditions are conditions of their truth' (Benacerraf (1973), p. 662). Simply put, the dilemma seems to be: have a clear semantics and an epistemological mystery, or have a clear epistemology and a semantic mystery.

¹ A modal realist could concede that maybe reality doesn't contain anything so ridiculous, if they are prepared to say that maybe such ridiculous scenarios are not possible after all. My reply is that, given that we are talking about unrestricted subjective possibility, this response just doesn't do justice to the phenomena; a sentence saying that one of those ridiculous scenarios doesn't obtain just doesn't seem to be a necessary truth in the sense of 'necessary', delineated in Chapter 1, that I am interested in in this thesis. Of course, saying that the sorts of ridiculous scenarios I have in mind 'are not possible' can sound very plausible, but I think that that relies on hearing 'possible' in a sense different from the one relevant for us here.
Lewis's response to this, for the mathematical case, is forthright and compelling. To take the second horn of the dilemma and give mathematics some special interpretation would be to reform mathematics, and to do this for the sake of saving some epistemological theory according to which causal connection is necessary for knowledge of objects, would be ridiculous, as mathematical knowledge is much more secure than such an epistemological theory. 'It's too bad for epistemologists if mathematics in its present form baffles them, but it 'would be hubris to take that as any reason to reform mathematics.' (Lewis (1986), p. 109.)

Thus Lewis takes mathematics as a 'precedent for knowledge beyond the reach of our causal acquaintance' (Lewis (1986), p. 109). If the objection is simply that Lewis cannot account for knowledge of goings on in other worlds because they are beyond the reach of our causal acquaintance and we can't know about such things, the case of mathematics diffuses the objection.

To this, however, it may be replied that mathematical knowledge is importantly different in kind from knowledge of goings on at other worlds. So, while it may not be true that causal acquaintance with objects of knowledge is always required, it is required for some kinds of knowledge. One natural candidate for drawing the line between knowledge for which causal acquaintance is required and knowledge for which it is not is the distinction between the concrete and the abstract. If we draw the line that way, since knowledge of goings on at other worlds surely counts as concrete, we still have a problem, notwithstanding our ability to know mathematics.

Lewis's response to this enables him to agree that some knowledge indeed does require causal acquaintance. He accepts this, but denies that the concrete/abstract distinction is the way to draw the line. The contingent/necessary distinction is (Lewis (1986), p. 111).

Is there a non-

-ad-hoc reason to believe that? Lewis offers the following considerations about counterfactual dependence:

If I know by seeing, for instance, my visual experience depends on the scene before my eyes; if the scene had been different, within limits, my experience and my subsequent belief would have been correspondingly different. Likewise other channels of causal acquaintance set up patterns of counterfactual dependence whereby we can know what is going on around us. But nothing can depend counterfactually on non-contingent matters. For instance nothing can depend counterfactually on what mathematical objects there are, or on what possibilities there are. (Lewis (1986), p. 111.)

On the basis of this, Lewis says:

So we have the desired boundary between knowledge that does and that doesn't require causal contact with the subject matter. It is a principled boundary, though motivated within the very modal realism that is in dispute. (I am mounting a defensive operation, and will be content with a standoff.) (Lewis (1986), p. 111.)

There are at least two difficulties we could raise about Lewis's attempt to make this a principled rather than an ad hoc response, the second of which compounds the first. Firstly, it seems like he is assuming, not modal realism, or not just modal realism, but also a substantive and controversial thesis about counterfactuals: that there are no non-vacuous counterfactuals with impossible antecedents. For if there were such, then things would sometimes 'depend counterfactually on non-contingent matters'. Secondly, to make matters worse, assuming modal realism makes this assumption even less plausible: if all the goings on at all the worlds were different in some such way (e.g. so that there were no tall men at any worlds), then they would also differ in some further way (there would be no tall men with hats). 'If there had been no tall men at any worlds, there
would have been no tall men with hats' seems non-vacuously, if trivially, true. And if you delete the 'no' you get something false. (This is closely related to the contingent totality objection, to which I devote Section 3.6. below. But here the point is, not that the totality of goings on at all worlds is intuitively contingent, but that if we follow modal realism and hold it necessary, then it seems we get a whole host of apparently non-vacuous counterfactuals with impossible antecedents. ²)

Still, it must be admitted that Lewis does have a response here, whether or not it is independently motivated. But is it correct? Is contingency the characteristic thing about causal-acquaintance-requiring knowledge? Consideration of paradigm Kripkean cases of the necessary a posteriori suggests otherwise. Consider, for example, 'Hesperus is Phosphorus', 'Cats are animals', or 'Water is H₂O' (it does not matter if you think some of these cases are spurious, as long as you think there are some cases of this sort). Since these are empirical propositions about how things are in the external world, don't we need causal acquaintance - or testimony from someone who does have such acquaintance - with their subject matter in order to know them? If so, then Lewis's move cannot be right. So it seems his options are as follows: deny that these propositions (as well as any other appropriate candidates) are necessary, which would be highly revisionary in this post-Kripkean world, or deny that they possess the causal acquaintance requirement. This latter road then forks again: Lewis could either maintain that these cases are actually a priori, or he could maintain that they are, as they seem, empirical, but that they can somehow be known without causal acquaintance with their subject matter. None of these responses seems very palatable.

So much for (1), the Benacerraf-style objection. Now for (2), a less theory-laden objection, which may be even more difficult for the modal realist to respond to. It is hard to deny that there is considerable intuitive force to the idea that we can't know about goings on at other worlds spatiotemporally disconnected from ours. And we don't have to spell this intuition out in terms of some causal acquaintance requirement.

What can the modal realist say to this? Lewis's work suggests two types of response: one broad, one narrow.

The broad response is to accept that there may be some intuitive difficulty here, but that the theoretical power of modal realism is so great that we should put it to one side. I will not try to argue against that move at this point, as it would take us too far afield, but see Section 3.7. below. However, I hope this objection together with the others begins to suggest that modal realism is not, on closer inspection, so theoretically attractive after all.

The narrow response comes in the form of some positive suggestions about how we get modal knowledge - or at least, how we come by modal opinions (more on this difference below). Again drawing a parallel with mathematics, Lewis suggests that with modality as with mathematics, we arrive at our modal opinions 'largely by reasoning from general principles that we already accept'. According to Lewis,

(...) our everyday modal opinions are, in large measure, consequences of a principle of recombination - something along the lines discussed in section 1.8, though doubtless there is room to improve my formulation of it, but in fact our reasoning is more likely to take the form of imaginative experiments. (Lewis (1986), p. 113.)

(The principle of recombination as formulated by Lewis runs as follows: 'Roughly speaking, the principle is that anything can coexist with anything else, at least provided they occupy distinct spatiotemporal positions. Likewise, anything can fail to coexist with anything else.' (Lewis (1986),

² Lewis famously held that counterfactuals with impossible antecedents are vacuously true (see Lewis (1973)). I do not know if this tension in his views has been remarked upon.
Lewis then goes on to discuss some difficulties and complications which aren't relevant for our purposes here.)

Lewis allows that there may be other principles at work:

For more far-fetched possibilities, recombination is less useful. But there are other principles that we can apply. A rejection of arbitrary-seeming limits on the plenitude of worlds, for instance, might lead us to conclude that if any worlds have seventeen dimensions then others have eighteen; (...) (Lewis (1986), p. 114.)

Still, it seems fair to say that the principle of recombination plays a major role in this picture of how we come by modal opinions.

Now, as an account of how we come by our modal opinions these suggestions may have merit. Indeed, they may even have some merit as a partial account of how we in fact come by our modal knowledge. But, I want to point out, as a modal realist response to how we come by modal knowledge, the principle of recombination suggestion just passes the buck, since it raises the equally difficult question of how we know the principle of recombination if modal realism is true.

This is easy to miss, since some sort of principle or recombination governing what is possible both seems intuitively correct and has an a priori character. But we must consider what this principle amounts to given modal realism. Once we reflect that, in a modal realist context, this principle has abundant and substantial consequences for what there is in concrete reality, but spatiotemporally disconnected from us, the epistemological question comes up again with all its original intuitive force: how could we know such a thing?

In fairness to Lewis, he explicitly presents his suggestions as an account of how we come by our modal opinions, in answer to an imagined request for a 'naturalized epistemology'. But then that just means that, beyond the broad response we considered above - which effectively just discounts the relevant intuitions as insufficiently worrying, in view of modal realism's supposed virtues - Lewis just doesn't have an answer to the minimal intuitive version of the epistemic objection (which, unlike the imagined demand for a naturalized epistemology, really is asking how, given modal realism, we can have modal knowledge).

Finally, let us consider version (3) of the epistemic objection: how do we account for the epistemology of modality if we embrace modal realism? Here, we see a similar failure on Lewis's part to respond to the challenge in its most acute form, as we did above with the minimal intuitive version.

Lewis considers three ways of taking the question 'If we don't know by causal interaction that other worlds and their donkeys exist, how do we know?', all of which fall under this third heading of ours as being versions of the theoretical challenge. (Lewis (1986), p. 113.) The first way is to construe the question as the request for a 'fully general analysis' of knowledge. Lewis's response: 'That is a fair request, and I regret that I cannot deliver the goods. But I don't see that this is especially my problem. It is a problem for everyone (certain sceptics and conventionalists excepted) and it is not worsened by a modal realist construal of the content of our modal knowledge' (Lewis (1986), p. 113).

The second way is to construe the question as a request for a 'naturalistic epistemology' (Lewis (1986), p. 113), an account of how we in fact come by our modal opinions, questions of knowledge aside, and this is where Lewis makes his suggestion (introduced above) that we come by our modal opinions by way of the principle of combination and perhaps some other principles.
The third and final way is to construe the question as a sceptical request for justification: 'put this knowledge on a firm foundation, show that it is derived by an infallible method' (Lewis (1986), p. 114). Lewis has some interesting things to say here, but for our purposes all we need to know is that he rejects the question so construed as fundamentally misguided. I take it most of us would take this as the obvious, sensible response.

What makes Lewis's overall argument here weak is that this list of three construals is not exhaustive, and passes over what is arguably a more challenging construal. He turns, from the request for a fully general theory of knowledge, to the request for a 'naturalized epistemology'. But what about the request for a theory of modal knowledge? A theory which, like an answer to the first construal, really would explain how modal knowledge is knowledge but which, like an answer to the second construal, could be restricted to the matter of modality. Taken that way, we have a theoretical challenge which Lewis has much less to say in response to. It is less plausible for him to say, as he does in the face of a request for a fully general analysis of knowledge, that this is expecting too much. The contention that it is not made harder by adopting modal realism is also suspect - why should we agree with that? And his positive material, about the principle of recombination, fails for the same reason as it fails as a response to the minimal intuitive objection considered above: it just raises the question of how, given modal realism, we could know that the principle of combination holds. So, the theoretical challenge understood this way constitutes a fairly solid objection. However, Lewis can still respond - not totally implausibly - that giving a theory of modal knowledge is hard for everyone, so if he does not meet that task it shouldn't be held too much against him. Thus it seems that the theoretical challenge version of the epistemic objection has a drawback which doesn't affect the minimal intuitive version, and is without any distinctive strengths of its own which may offset that.

To sum up, the Benacerraf-style puzzle by itself is not a serious threat to modal realism, since that puzzle applies to mathematics too, but we clearly have mathematical knowledge. This puzzle augmented with the consideration that goings on at other worlds are the sort of thing we should need causal contact with to know about has some bite, despite what Lewis says, since his contention that contingency rather than concreteness is the mark of causal-contact-requiring knowledge may be criticized.

Stronger is the minimal intuitive version of the epistemic objection, since it does not presuppose anything about our having to have causal contact with that which we know about. Apart from the broad strategy of simply dismissing the intuition that there is a problem here on the questionable grounds that modal realism's other virtues outweigh it, the only other response suggested by (but strictly speaking not contained in) Lewis's work - a positive suggestion about how we know, according to which our knowledge of goings on in other worlds is derived from our knowledge of something like the recombination principle - just raises the equally problematic question of how we could know the recombination principle given modal realism.

The third sort of epistemic objection, which consists in challenging the modal realist to provide an epistemological theory of some kind, or a refutation of skepticism, is not as strong, since such challenges are difficult across the board. It may not be devoid of force, but it would seem that this sort of objection has no advantage over the minimal intuitive version.

Overall then, modal realism faces considerable difficulty on the epistemic front, perhaps most acutely in the humble form of the minimal intuitive version of the epistemic objection.

3.4. The Humphrey Objection

The Humphrey objection, due to Kripke, centres on counterpart theory, and alleges that this assigns counterintuitive truth-conditions to modal statements about individuals. It is historically
important, as Kripke made the objection in his influential *Naming and Necessity* lectures, long before Lewis published his full defence of modal realism in 1986.

Kripke put the objection as follows:

Thus if we say "Humphrey might have won the election (if only he had done such-and-such)"*, we are not talking about something that might have happened to Humphrey but to someone else, a "counterpart". Probably, however, Humphrey could not care less whether someone else, no matter how much resembling him, would have been victorious in another possible world. Thus, Lewis's view seems to me even more bizarre than the usual notions of transworld identification that it replaces. (Kripke (1980), p. 45, f.n. 13.)

It is by now, I think, pretty widely accepted that Kripke, while he may have been on to something here, did not put the point optimally. To the objection put this way, there is a cogent response: it is not correct to say that, according to counterpart-theoretic modal realism, we 'are not talking about something that might have happened to Humphrey'. What counterpart-theoretic says is that talk about 'what might have happened to Humphrey' is to be analyzed in terms of what does happen to his counterparts in other worlds. So according to the counterpart-theoretic modal realist, when we say 'Humphrey might have won the election', we are indeed talking about what might have happened to Humphrey. Their characteristic claim is to add that this thing we're talking about is to be analyzed in terms of what happens to counterparts. (Lewis drives this point home on (1986), pp. 195 - 196.)

Similarly, the second part of Kripke's objection - that 'Humphrey could not care less whether someone else, no matter how much resembling him, would have been victorious in another possible world' - can be convincingly argued to miss the mark. The counterpart-theoretic modal realist can agree that Humphrey could not care less about this. For the way they analyze talk about whether 'someone else' (a counterpart) 'would have been victorious in another possible world' is in terms of counterparts of that someone else - counterparts of Humphrey's counterparts. And it is compatible with Humphrey not being interested in what happens to the counterparts of some one of his counterparts, that he be interested in something which - upon analysis - turns out to be a question of what happens to his own counterparts.

This last point may be a bit pedantic, however. What if we simply reform the second part of Kripke's objection by changing the 'would have been' to an 'is'? This yields: 'Humphrey could not care less whether someone else, no matter how much resembling him, is victorious in another possible world?'.

This is better, but there is still a strong reply. As Sider says in his (2006), this is 'just the paradox of analysis':

A reasonable person can care about a property under one description ("possibly winning") while not caring about the same property under another description ("having a counterpart who wins"), provided it is not obvious that the descriptions pick out the same property. Correct analyses need not be obvious to competent language users. Obviousness may count for something, but theoretical virtues are important as well in determining which analyses we ought to accept (Sider (2006), p. 2.)

I endorse this as a response to the version of the Humphrey objection just considered. However, I want presently to register a difference with Sider about whether this response also works for another version of the objection.
This other version puts aside what Humphrey cares about, and appeals directly to our intuitions. Sider puts this version of the objection as follows: 'Look, it is just obvious that possibly winning is not the same as having a counterpart who wins'. (Sider (2006), pp. 1 - 2.) And the response quoted above is put forward by Sider as a response to both the previously considered version and this one. (He explicitly prefaces the passage with 'Reply to ii) and iii' (Sider (2006), p. 2.).

Does Sider's response apply here too? On reflection, I think clearly not. The response makes the point that a correct analysis need not be obvious (while granting that obviousness may count for something). But the present version of the objection is alleging, not that it isn't obvious, but that it is obviously not the case. Sider, in putting the passage in question forward as a response to this, is sliding from '(~p) is obvious' to '~(p is obvious)' and thus failing to address the objection.

So we seem to have a version of the Humphrey objection which is stronger than the others so far considered. But perhaps we can improve it further, from a rhetorical point of view, by getting away from obviousness altogether. Saying that possibly winning is obviously not the same as having a winning counterpart risks seeming dogmatic. Someone may respond 'Well, it isn't obvious to me, and things that have seemed obvious to people have turned out not to be the case'. Perhaps the rhetorically wise thing is to tone it down and simply enter a plea that it doesn't intuitively seem that possibly winning is the same as having a winning counterpart. Or putting the point semantically: the truth-condition Lewis assigns to 'Humphrey could have won' is counterintuitive.

So, despite the availability of strong responses to the original and certain subsequent versions of the Humphrey objection, the core point remains that the truth-condition assigned by Lewis is counterintuitive.

(Incidentally, Lewis suggested that forms of ersatzism are no better on this score: in that case, what "gets into the act" is not another person, but 'some abstract whatnot' (Lewis (1986), p. 194). This isn't a strong reply to the objection, of course, as ersatzism is far from the only other game in town when it comes to the semantics of modal attributions such as 'Humphrey might have won'. Nevertheless and for what it's worth: perhaps an abstract whatnot getting into the act is, from an intuitive point of view, not quite as bad as another person getting into the act. Bringing in another person, it seems to me, feels more like crowding out Humphrey, more like putting something - someone - in his place.)

So, there is a version of the Humphrey objection which has some force. However, modal realism with overlap, in contrast to counterpart-theoretic modal realism à la Lewis, is immune to the Humphrey objection. Lewis wasn't swayed by this, since he had reasons to think modal realism with overlap unpalatable. Since then, advocates of overlap have emerged (most notably McDaniel (2004)).

It may be that the considerations against overlap are quite compelling, in which case these together with the Humphrey objection (once it is freed from its initial faulty formulation) have significant force against modal realism in general. However, I do not want to get deep into comparing the relative merits of counterpart-theoretic modal realism and modal realism with overlap, and would prefer to have an objection along similar lines which applies to both. Therefore, I advocate that we take the Humphrey objection, not just as a self-sufficient objection which affects the dominant form of modal realism but not modal realism with overlap, but also as a clue: modal realism - in both flavours - may be counterintuitive on the semantic front, and this may be a good reason to reject it. Since the Humphrey objection itself fails to apply to modal realism with overlap, we should set it aside and go on to try for a more general semantic objection. I turn to this now.

3.5. The Semantic Objection
This objection makes an appeal to our better judgement about the meaning of modal talk. Perhaps it doesn't give us much rhetorical leverage against the modal realist. It doesn't, for instance, take something the modal realist should qua modal realist endorse, and turn that against modal realism. It is, however, very strong in the sense that, for a right thinking person, this objection is probably one of the main reasons why they do not embrace modal realism.

The objection, simply put, is just that modal talk seems not to be about the goings on at other worlds. When we say something is possible, that just doesn't mean that it happens at some other world.

We find a strident expression of this objection in Nathan Salmon's (1988) review of Lewis (1986):

> It is tempting to conclude from the theory and its defense that Lewis officially endorses an extremely implausible cosmological theory and, believing this theory to be relevant to the content of modal discourse, does not understand what it means (in English) to say "it might have been that such-and-such." This reviewer conjectures that Lewis's highly eccentric views concerning alternative universes, counterpart relations, and their alleged role in modal discourse have their ultimate source in a conceptual confusion between the modal proposition that x might have been such-and-such (where x is a possible individual) and the nonmodal proposition that x is in fact such-and-such (and in some "world"). (Salmon (1988), p. 240.)

And at the very end of the review, Salmon states his position categorically:

> Whichever is the case [Salmon is here referring to a range of interpretative options he has laid out with respect to Lewis's theory], Lewis seriously misunderstands what "might have" means. (Salmon (1988), p. 244.)

One issue with this objection the way Salmon has put it is that he is attributing to Lewis himself a failure to understand a part of English. One thing that may make that claim problematic is that, presumably, outside of the context of philosophical theorizing, Lewis has a good grasp of "might have" and like expressions, and does understand them. But this is a relatively pedantic quibble; clearly, the core of the objection is really that Lewis's theory gets the semantics wrong. I share Salmon's opinion about this, as evidently do many others. This point seems to lie behind several other, more multifarious, complaints made against modal realism in the literature. For example, Roy's (1993) complaint that if some of Lewis's worlds were 'annihilated' this intuitively would not change the modal facts. Or Williamson's complaint (quoted at greater length below in Section 3.6. on the contingent totality objection, which the complaint also suggests) that:

> [e]ven if there are mutually disconnected spatiotemporal systems such as Lewis postulates, they are not the distinctive subject matter of modal discourse. They are simply more of what there is, about which we can ask genuinely modal questions. (Williamson (2002), p. 239 - 240.)

A potential problem for this line of objection is that it is not completely clear that Lewis intends modal realism to provide a semantic analysis of exactly what modal expressions currently mean. We may adopt a reading of Lewis's proposal on which it is more conceptually revisionary: modal talk may not currently amount to talk about possible worlds, but we should abandon that modal talk in favour of talk which does amount to talk about possible worlds.

What should we say about this, from our point of view? Two counter-replies to this reply to the semantic objection to modal realism suggest themselves.
The first counter-reply is that, if Lewis is not trying to give an analysis of our actual concept of necessity *de dicto*, but rather to replace it with some other concept which he thinks we should be using instead (for some philosophical purposes, at least), then he simply isn't addressing our problem, which is to understand better and if possible analyze the concept of necessity *de dicto* - the concept we actually have. This is an interesting problem - at least I think it is, and I think I can count on plenty of contemporary philosophers agreeing with me on this - and the project of coming up with a replacement, in pursuit of some general ideological or metaphysical preferences, just isn't directly relevant.

The second counter-reply is that, in any case, it is far from clear that a Lewisian replacement for our modal concepts has any real hope whatever. It is one thing to talk in very abstract terms about simplifying our ideology, or our total theory, by replacing our actual modal concepts with ones which concern Lewisian possible worlds. It is quite another to actually carry through the replacement. The easy, breezy way we philosophers have of talking about such potential changes should not be allowed to mask their radicalness. Our actual modal concepts are plausibly very deeply rooted in our thinking - it may be, practically speaking, psychologically impossible for us to replace them. And if that's true, how are we really to assess the prospects for doing so in principle (i.e. allowing that perhaps we're, due to unfortunate human frailty, unable to make the advantageous change)?

Casting doubt in this way on both the practical viability (i.e., viability for us) and in principle viability (i.e. viability for some thinking being) of the change advocated by Lewis read as a conceptual revisionist also strengthens the claim that our topic is interesting in its own right; if it were more feasible to replace existing modal concepts with Lewisian ones, it could be more plausibly claimed that our existing modal concepts are not particularly worth philosophizing about. There would still be plenty of room to reply that, no, they're still very interesting in their own right, but if they are very hard or impossible to replace, their claim to our interest is clearly stronger still.

So, overall we have a dilemma here: Lewis's account may be construed as descriptive, or revisionary, when it comes to our notion of necessity *de dicto*. If descriptive: it seems like Lewis's account does not do justice to the real meaning of 'is necessary': intuitively, it just doesn't seem that such ascriptions of modal status are about other concrete worlds and the goings-on at them. If revisionary: this is just changing the subject, and furthermore the revisionary project may be far less tenable than Lewis would have us think - which makes it all the less of a threat to the interest of our actual topic: the existing notion of necessity *de dicto*.

### 3.6. The Contingent Totality Objection

Intuitively, reality as a whole could have been different. But according to modal realism, reality as a whole - that is, the totality of the posited worlds - is necessarily the way it is. Lewis is very upfront about this. Witness:

> There is but one totality of worlds; it is not a world; it could not have been different. (Lewis (1986), p. 80.)

Thus modal realism appears to give the wrong answer about this matter, and should therefore be suspected of falsity.

---

3 It may superficially seem as though I am flirting with some dubious form of psychologism here, but I don't think so; the point isn't that if it is difficult or impossible for us to think something (or think in some way), then it isn't true (or isn't a good way to think). The point is rather that really following through on modal realism in our thinking may be less feasible than we tend to realize, and thus positive evaluations, by us, of its performance with respect to theoretical virtues like simplicity may be more superficial and inaccurate than we tend to realize.
Somewhat surprisingly, it is hard to find published instances of exactly this objection, put in terms of the contingency of the way reality is as a whole not being predicted by modal realism. (Perhaps this is because it is such a basic point, and Lewis has already bit the bullet here.) But there are objections in the literature which touch on it, for example this passage from Williamson (2002):

Even if there are mutually disconnected spatiotemporal systems such as Lewis postulates, they are not the distinctive subject matter of modal discourse. They are simply more of what there is, about which we can ask genuinely modal questions: for instance, whether there could have been more or fewer spatiotemporal systems than there actually are. To put the point another way, the modal realist claims that one can fully specify how things are in an extensional language without modal operators, restricted quantifiers or other expressions indexed to worlds. Yet, still according to modal realism, nothing stated in that language is contingent. Thus the view implies that it is not genuinely contingent how things are. Of course, the view also implies that one may truly say 'It is contingent that there are no talking donkeys'; that shows that it is also wrong about the truth-conditions of modal statements. Lewis misidentifies contingency as a special kind of indexicality, just as Berkeley misidentified material objects as special groups of sense impressions. (Williamson (2002), pp. 239 - 240.)

There is a lot going on here (what I treat separately as the semantic objection makes an appearance, for instance), but the contingent totality objection appears quite clearly with the third-last sentence: 'Thus the view implies that it is not genuinely contingent how things are.'

We can also give a closely related counterfactual conditional variant on the contingent totality objection; the following seems true: 'If I had behaved differently this morning from the way I actually behaved, reality as a whole would have been different from how it actually is.' Reality may not have been very different in such a case - or maybe we should think it would have been, in light of the butterfly effect - but in any case, it would have been different. I suspect modal realism's failure to allow this lies at the heart of the much-discussed ethical objections to modal realism. (I do not consider the ethical objections in their own right in this thesis. I think that they involve too many disparate philosophical difficulties, especially in metaethics, to be powerful - or at least, to be made powerfully within the confines of this thesis. (For discussion, see Adams (1979, p. 195), Lewis (1986, p. 123), and Heller (2003).)

I think this simple objection is powerful. I know of no formidable response to it. Before moving on, I will consider and argue against some potential responses.

One response of course would be to deny that reality as a whole could have been different. But I think it must be admitted that such a denial is theoretically costly, because it is so counterintuitive. Even if it is rarely made explicit, the belief that reality as a whole, not just parts of it, could have been different, seems quite basic to our intuitive thought about modality. Indeed, it seems to be part of our intuitive thinking about modality that we think that, if any part of reality had been different from the way it actually is, then reality as a whole would have been different from the way it actually is - and modal realism seems incapable of upholding that implication too.

Another sort of response would be to deny that modal realism implies that reality as a whole is necessarily the way it is. And, in an unfortunate complication, there are some features of modal realism as developed by Lewis which, on certain natural interpretations at least, may make it seem as though this response is viable. I suggest that this is a red herring; these very features give rise to disastrous consequences, which have been discussed in the literature under the heading of 'advanced modalizing' (see Divers (1999), Parsons (2012), and Jago (2016)).
The features and behaviour of modal realism which are behind this complication have to do with the notion of being true at a world. Lewis’s analysis says that a proposition is possible iff it is true at some possible world. Now, according to the modal realist, there are many worlds. So, since what is the case is possible, the modal realist ought to say that the proposition ‘there are many worlds’ is possible. But their analysis of this attribution of possibility makes it equivalent to: at some possible world, there are many worlds. This can be read in such a way that it seems false - and if the reading intended by the modal realist is like this, then the modal realist is driven towards the disastrous consequence, that their own theory, by its own lights, is not possibly true. Let us call any such understanding of truth at a world ‘world-bound’.

Now it could be argued that the world-bound way of understanding truth at a world - or one of them, if there are multiple such ways - yields the result that at some worlds, reality as a whole is different from the way it actually is. So if the modal realist sticks to this way of understanding truth at a world, they can escape the counter-intuitive apparent consequence of their theory that the way reality is as a whole is not contingent.

I have two responses here: firstly, the objection is explicitly about reality as a whole: surely the ‘as a whole’ in the objection, understood in the way it is intended, should cancel any world-boundness in the understanding of truth at a world. Secondly: even if the first response can somehow be resisted by the modal realist, it still remains the case that this way of understanding truth at a world leads to the disastrous consequence that their theory is not possibly true.

Thus it seems we have a dilemma for the modal realist: either accept that modal realism doesn’t allow for reality as a whole being the way it is contingently - not an attractive option - or appeal to a strongly “world-bound” notion of truth at a world and face the problems of advanced modalizing, such as the consequence that their theory is not possibly true.  

3.7. The Motivation for Modal Realism

This is not a sharp objection like those above, but rather an attempt to say something about what really motivates modal realism, and an invitation to consider these motivations critically.

It is now a commonplace in discussions of Lewis’s philosophy that he took up his teacher W.V.O. Quine’s attitudes regarding how language works - or rather, how “serious scientific language” should work - but revised Quine’s skepticism about modal language by providing a theory which promises to analyze modal language into a form which satisfies Quinean strictures: modal realism with counterpart theory. As Williamson writes:

[Quine’s] standard of intelligibility in logic was austere: first-order non-modal logic, roughly, that of the logical constants ‘not’, ‘and’, ‘or’, ‘everything’, ‘something’, and ‘is’. For Quine, logic is first-order non-modal logic. Lewis assumed modal realism because it permits the reduction by translation of a quantified modal language to a first-order non-modal language in which one talks about worlds and individuals in those worlds. Crucially, Lewis’s modal

---

4 A more radical sort of response to this objection, which I would not expect the typical modal realist to find congenial, would be to take a skeptical attitude to the very idea of reality as a whole. Serious consideration of such a radical idea is beyond the scope of this thesis, and I have no pithy objection to this response. I am mostly content to simply put this response to one side and hope that anyone attracted to it is still moved by my other objections to reject modal realism. However, here is one parting speculation: perhaps the objection could be modified so that skepticism about reality as a whole does not affect it, by substituting, in the place of the idea of reality as a whole, some reference to a large aggregate of things, an aggregate that cannot acceptably be counterpart-theoretically related to any distinct entity. Or we could forget talking about a single "thing", reality as a whole, and even forget talking about absolutely every thing, and talk instead of, say, absolutely all hats. Surely things could have been such that there are less hats in existence than there actually are - but of course the objector may counter this with a radical rejection of this sort of quantification.
realism gave him a way of informally explaining what a possible world is in non-modal terms: roughly, a spatiotemporal system; the individuals in such a system are spatiotemporally connected to each other and to nothing outside the system. Lewis thereby aimed to make quantified modal logic intelligible by his teacher’s standards. (Williamson (2014), p. 10.)

Do not be misled by Williamson's focus on logic here. Lewis’s attempt to make quantified modal logic intelligible by his teacher’s standards also amounts to an attempt to make modal discourse, possibly carried out in natural language, so intelligible. One might begin by translating such discourse into quantified modal logic, and then follow Lewisian procedures for translating the result into first-order logic. In many ways though, just the theoretical option for doing this is enough; if that is vouchsafed, we can relax and use modal talk in natural language directly, without worrying that what we are saying cannot in principle be put into the forms of “serious scientific language”.

This is not the place to enter upon a detailed analysis of exactly what the Quinean strictures are, but one of the chief linguistic desiderata for Quine was extensionality - roughly, the property a sentence has if and only if replacing referring terms with other terms which refer to the same things does not alter the truth-value of the sentence. (This definition then runs into the question 'Which are the referring terms?', and different answers yield different notions.) As the Williamson passage above suggests, another touchstone for Quine was translatability into first-order logic.

This sort of attitude to language - the Quine-Lewis attitude to language - is thus properly seen as a major root, if not the major root, of modal realism. It follows that if we want to combat modal realism, a very effective way of going would, if possible, be to combat the Quine-Lewis attitude to language. If we leave that intact, then all our strong objections above (and make no mistake, they are strong) might still fail to topple modal realism. Someone wedded to the Quine-Lewis attitude might respond to all these as follows: yes, there are serious difficulties here, but we cannot follow Quine in his skepticism - that just isn't a serious option any longer - and Lewis’s proposal, for all its grave problems, is still the best thing on the market. It is the only non-deflationary, non-skeptical, thoroughly reductive option which satisfies Quine-Lewis linguistic desiderata.

I hope that the objections given above have - modus tollens style - cast doubt on the Quine-Lewis linguistic desiderata. Also, I hope that merely pointing out that these linguistic desiderata are playing a large motivating role has some effect. This is not the place to get deep into criticizing directly what I have called the Quine-Lewis attitude to language, so at this point I will just suggest that we do try to consider it critically.5

Another important strand of Lewis’s motivations for modal realism, one which in his (1986) he emphasized more than the Quine-Lewis linguistic desiderata, is the desire to systematize and simplify what Quine and Lewis called ‘total theory’: the totality of everything we believe. Relatedly, Lewis had a metaphysical vision of reality on which

   (…) all there is to the world is a vast mosaic of local matters of particular fact, just one little thing and then another. (Lewis (1986b), ix.)

These elements are woven together in the following description from Soames (2014):

   Here we see what appears to have been the enduring influence of David’s teacher, Quine, the great champion of naturalism and extensionalism, and the uncompromising scourge of the modalities. The underlying philosophical purpose of modal realism and counterpart

5 A further, more controversial, suggestion on how to go about this: think of the Quine-Lewis attitude as a special instance of what Wittgenstein criticized - in his (1953) and other work - as the tendency of philosophers to treat all language along the same lines.
theory was [to] reduce an intensional object-language to a purely extensional semantic metalanguage, in the service of an antecedently desired conception of reality. Whereas Quine taught that vindicating naturalism and extensionalism required eliminating intensional facts and rejecting intensional constructions, his student, David Lewis, tried to show that intensional facts are just a species of extensional facts, and that intensional constructions in language are no threat to the integrity of an austere, naturalistic vision of reality. (Soames (2014), p. 83.)

Separating them, on the other hand, Williamson notes that Lewis over time placed greater emphasis on the metaphysical and 'simplifying total theory' aspects, and less on semantics:

Lewis’s case for his modal realism itself evolved over time. In the original 1968 paper, the emphasis is on the relation between modal and non-modal languages, and the clarity to be achieved in modal logic by translating the modal language into the non-modal language of Lewis’s counterpart theory (the precursor of his modal realism). (…) By the time he wrote what became the canonical case for modal realism, his book On the Plurality of Worlds (Lewis 1986), based largely on his 1984 John Locke lectures at Oxford, Lewis’s perspective had changed. He talks much less about linguistic matters, and much more about the abductive advantages of modal realism as a theoretical framework for explaining a variety of phenomena, many of them non-linguistic. (Williamson (2014), p. 11.)

Again, these are deep waters, and so our aims here must be limited, but I have some critical remarks to offer regarding this latter aspect of Lewis's motivations.

As the saying goes, 'Make things as simple as possible, but no simpler.' It is one thing to prefer a simpler theory if that theory really works, and makes sense of things. But in the case of modal realism, the proposal is more just something that, considered very abstractly, would simplify things in a certain way if it worked. We have to be on our guard at letting this ideal of simplicity, so wholesome in its proper place, lead us up the garden path when we are puzzled and doing philosophy. Modal realism, for all the technical brilliance of its developer, is at bottom a crazy speculation. And it's crazy in a peculiar way - not just factually crazy, but logically crazy so to speak. (Seen as such, however, there is something wonderful about it. There is something uncanny about the fact that humans have so much as thought such a thing.) So, the worry here is something like this: the type of simplicity offered is very specific and peculiar, and it is far from clear that the theory offering it is at all viable - indeed, if the previous objections are anything to go by, it may well be that it is not.

Another criticism we may level at the attempt to motivate modal realism in terms of simplicity is more practical-minded: things haven't gotten much simpler since modal realism came on the scene. Modal realism may by its own lights "simplify" things in a peculiar sense, purporting to offer us a way to answer or dispense with certain philosophical questions about modality, but it raises all sorts of difficulties and strange questions which philosophers pursue further each year. So far, its practical effect on our collective conceptual apparatus has if anything been to expand and complicate it. (This may plausibly be counted as a gain, but it would seem to undermine the simplicity motivation for modal realism.)

Of course, this objection could be countered in a number of ways. Perhaps it's a case of things having to get worse before they can get better. Or perhaps the importance, for relevant philosophical purposes, of the practical viability of incorporating a theory into our thinking can be downplayed. The objection may even seem foolish. I don't want to lay too much stress on it, but I think it may have a certain sobering, grounding effect, calling us back from a philosophical dream.

---

6 The saying is often attributed to Einstein, but this is doubtful. A blog post by O'Toole (2011) investigates the issue, with further information added by one of the commentators there.
Finally, we may question whether Lewis's movement, as charted by Williamson, from motivating modal realism with semantic considerations towards motivating it with metaphysical considerations, is really just a movement from one kind of motivation to another kind. Perhaps the old Quine-Lewis attitude to language is still there at the root, but is just harder to get at now that it is wrapped up in thorny metaphysical ideas. Thus we may interpret Lewis's later emphasis on metaphysics as opposed to semantics as a defensive manoeuvre. In that case, the deepest way to critique modal realism may still be to critique the Quine-Lewis attitude to language.

3.8. Conclusion

In this chapter I have outlined and given an extended criticism of one of the most prominent and most fully developed answers to the question 'Under what conditions is a proposition necessarily true?'. There is something slightly odd about this, since modal realism has probably never been in danger of becoming widely accepted. However, an incredulous stare is not an argument, and here I have tried to show that our incredulity can be (a) refined and articulated, and (b) backed up with detailed and compelling arguments.

Chapter 3 References


4. Sider’s Quasi-Conventionalism

4.1. Exposition

Starting in his (2003), Theodore Sider has been defending a theory of necessity de dicto called quasi-conventionalism. The most up-to-date version can be found in his (2011) and his (2013) replies to a symposium on that work. It states necessary and sufficient conditions for a proposition to be necessary, but as we will see, one of the key concepts involved has been left open-ended, so the account is to be regarded as partial. The account is supposed to reduce necessity de dicto to non-modal notions, and to be extendable to de re modality.

What makes Sider’s account so worthy of discussion in the present thesis is that it takes what I believe to be an important step forward with respect to the task of giving an account of necessity de dicto. The step forward is that it embodies a certain structure, which the account I will present shares. Abstracting from the details of Sider’s account and simplifying a bit, the shared structure can be expressed in the form of a schematic analysis as follows:

(Schema) A proposition is necessary iff it is, or is implied by, a proposition which is both true and meets a certain condition C.

(Sider, as we shall see, does not quite use his schema, but his analysis can easily be re-expressed in almost this form.)

So, Sider’s view takes, as I will argue, an important step forward. But it also has grave defects. Considering Sider’s view, then - seeing that it instantiates the suggestive and appealing (Schema), and seeing what is wrong with it - offers a nice way of leading up to and motivating my own account, which I will give in the next chapter. (This is not the way I actually arrived at my account, but it could have been.)

Two of Sider’s main starting points seem to be: (i) the desire to find a way of reducing modal concepts to non-modal concepts, and (ii) a hunch that conventionalist theories according to which necessities are true by convention were on to something: roughly, that convention should play some key role in the analysis of necessity. Regarding (i) and the underlying motivation for it, there are two interrelated strands here: one is a relatively theory-neutral feeling that modality is mysterious, or cries out for explanation, but this then plays into the second strand, which is emphasized in his (2011): a desire for an account of the “fundamental nature of reality”, “reality’s fundamental structure” - an account that “carves nature at the joints”.

Sider argues that it was always a mistake to try to account for necessary truth by means of the idea of truth by convention: the idea is of dubious coherence, and in view of necessary a posteriori truths especially, would not seem to line up with the idea of necessary truth anyway. But that doesn’t mean convention can’t play a crucial role in, not truth-making, but in making ‘necessary’ apply to the propositions it does apply to. (A similar point applies in my account, but with the meanings or natures of propositions rather than conventions; necessary truths are not, in general, true in virtue of meaning, but that they are necessary...
rather than contingent can be explained in broadly semantic terms.) Part of the motivation and attraction of truth by convention theories of necessity, Sider allows, was their promise of shedding light on the epistemology of logic and mathematics. The theory of modality Sider is offering makes no claim to do that. (But so what? Who says the place to look for insight into the epistemology of logic and mathematics is in a theory of modality?)

In his (2011), Sider labels his account ‘Humean’. (I prefer his earlier label ‘quasi-conventionalism’ for being more descriptive.) Here is his first pass at expressing his account there:

To say that a proposition is necessary, according to the Humean, is to say that the proposition is i) true; and ii) of a certain sort. A crude Humean view, for example, would say that a proposition is necessary iff it is either a logical or mathematical truth. What determines the “certain sort” of propositions? Nothing “metaphysically deep”. For the Humean, necessity does not carve at the joints. There are many candidate meanings for ‘necessary’, corresponding to different “certain sorts” our linguistic community might choose. (Sider (2011), p. 269.)

The role of convention in Sider’s account, then, lies in distinguishing this “certain sort” - or these “certain sorts” (Sider switches as this point to the plural):

Perhaps the choice of the “certain sorts” is conventional. Convention can do this without purporting to make true the statements of logic or mathematics (or, for that matter, statements to the effect that these truths are necessary), for the choice of the certain sorts is just a choice about what to mean by ‘necessary’. Or perhaps the choice is partly subjective/projective rather than purely conventional. (Sider (2011), p. 270.)

As can be seen at the end of this last quote, Sider has some uncertainty about whether the choice of the “certain sorts” is ‘purely conventional’. We will not get deeply into Sider’s ideas of ‘conventional’ and ‘subjective/projective’ here. It is enough for our purposes that the “certain sorts” are, for Sider, ‘not objectively distinguished’ (Sider (2011), p. 270). Or again in different words:

The core idea of the Humean account, then, is that necessary truths are truths of certain more or less arbitrarily selected kinds. (Sider (2011), p. 271.)

At this point Sider introduces a refinement, and it is this that will allow us to see how Sider’s account embodies (Schema) above:

More carefully: begin with a set of modal axioms and a set of modal rules. Modal axioms are simply certain chosen true sentences; modal rules are certain chosen truth-preserving relations between sets of sentences and sentences. To any chosen modal axioms and rules there corresponds a set of modal theorems: the closure of the set of modal axioms under the rules.[f.n. omitted] Any choice of modal axioms and modal rules, and thus of modal theorems, results in a version of Humeanism: to
be necessary is to be a modal theorem thus understood.[f.n. omitted] (“Modal” axioms, rules, and theorems are so-called because of their role in the Humean theory of modality, but the goal is to characterize them nonmodally; otherwise the theory would fail to be reductive. […] (Sider (2011), p. 271.)

Then, getting more specific with a preliminary proposal:

A simple version of Humeanism to begin with: the sole modal rule is first-order logical consequence, and the modal axioms are the mathematical truths. (Logical truths are logical consequences of any propositions whatsoever, and so do not need to be included as modal axioms.) (Sider (2011), pp. 271 - 272.)

At this point, we can see how Sider’s account, or type of account, will embody (Schema); his notion of ‘modal axiom’ combines the requirements of truth and being of a “certain sort” (or one of “certain sorts”), and the main point of the ‘modal rules’ seems clearly to be to draw out implications of the axioms. So, separating the truth and “certain sort” requirements again, we can with little or no distortion put Sider’s preferred type of account into (Schema), or more accurately, (Schema) augmented with a parenthesis about conjunctions:

(SiderSchema) A proposition is necessary iff it is, or is implied by, a proposition which is both true and is of a more or less arbitrarily selected “certain sort” (or a conjunction of such propositions). ¹

The presence in the account of implication (or something like it) is in my view an important and laudable feature. It is perhaps not sufficiently emphasized by Sider, and has been glossed over in subsequent discussion of his view. For instance, Merricks (2013, p. 732) glosses Sider’s account as saying that ‘Sider reduces a proposition \( p \)'s being necessarily true to: \( p \) is true-and-mathematical or true-and-logical or true-and-metaphysical or…’. The importance of implication (or something like it) in (Schema) and views embodying it will be made clear in Chapter 5 when I put forward my own account.

After giving his preliminary version of Humeanism (or quasi-conventionalism), Sider goes on to consider a series of worries, responding with ‘a combination of refinement and argument’ (Sider (2011), p. 272). He never arrives at a definitive proposal, but aims to develop his strategy sufficiently to justify his general approach.

I think we have already gotten a pretty good sense of Sider’s approach, but I want more or less to complete the exposition of Sider’s approach before moving on to objections, none of which are among the worries Sider considers. So before moving on to objections, I will now

¹ The parenthesis is called for because, since Sider’s view is that to be a necessary truth is to be in the closure (under the ‘modal rules’) of the set of propositions of a “certain sort”, or one of “certain sorts”, you might get propositions in that closure which only follow by the modal rules from a conjunction of propositions each of which is of the relevant “certain sort”, or one of the relevant “sorts”, but where the conjunction is not itself of the relevant “certain sort” or one of the relevant “sorts”. With my own account, I argue, the (Schema)-style formulation and the closure-of-a-set formulation are equivalent (see Section 5.4.). I prefer the former for being easier to grasp.
briefly convey six further points which emerge in Sider’s responses to the worries he considers.

(1) **Logical consequence must be non-modal**: Sider wants his account to avoid modal notions, so modal characterizations of logical consequence are out. Remaining options include primitivism about logical consequence, something Sider calls the “best system” account of logical truth (which he describes in Section 10.3. of his (2011)) extended to an account of logical consequence, and model-theoretic approaches.

(2) **Analytic truths added as axioms**: Sider holds that analytic truths should come out necessary, and proposes to that end that each analytic truth be added as a modal axiom (Sider (2011), p. 274). This move is unapologetically *ad hoc*. (You might worry, as I do, that some examples of the contingent *a priori* should count as analytic, in which case not all analyticities are necessities. But perhaps there are different notions of analyticity which may give different results here. In any case let’s set this aside.)

(3) **“Metaphysical” statements added as axioms**: again, modulo some fuzziness about what it takes for a statement to count as metaphysical - the gloss Sider uses is ‘truths about fundamental and abstract matters’ (Sider (2011), p. 275) - true metaphysical statements are to be added as axioms. Again this is unapologetically *ad hoc*, or treated as a brute fact: ‘What justifies their status as modal axioms? This is just how the concept of necessity works. Such propositions have no further feature that explains their inclusion as modal axioms.’ (Sider (2011), p. 275)

(4) **A new class of ‘natural kind axioms’**: another unapologetically *ad hoc* addition, this time to accommodate the necessity of natural kind type examples of the necessary *a posteriori*, e.g. ‘Water is H₂O’. I refer the reader to (Sider (2011), pp. 282 - 283) for details.

(5) **Contextual variation of the “outer modality”:** it is conventional wisdom that modal talk in the wild should be understood as being about a contextually variable space of possibilities. This is often combined with a picture of an outer, unrestricted space of possibilities which does not vary. Sider suggests, as ‘more attractive’ (Sider (2011), p. 281), that even the outer space is contextually variable - that ‘there can be contextual variation both in the modal axioms and the modal rules’ (Sider (2011), p. 281).

(6) **Family resemblances (maybe)**: on (Sider (2011), p. 288) Sider rehearses his (by now familiar) attitude to necessity thus: ‘Why are logical (or mathematical, or analytic, or …) truths necessary? The Humean’s answer is that this is just how our concept of necessity works.’ But then he turns around and suggests (Sider (2011), pp. 288 - 289) that ‘a Humean need not be quite so flatfooted. [...] [The Humean] resists the idea that there is a single necessary and sufficient condition for being a modal axiom. Nevertheless, she is free to exhibit similarities between various modal axioms, just as one might exhibit similarities between things that fall under our concept of a game, to use Wittgenstein’s example. Doing this would help to show that the Humean concept of necessity is not utterly arbitrary or heterogeneous.’ This no doubt helps the plausibility of Sider’s account in a way, but may also play into the hands of an objector, as we shall soon see.
This completes our exposition of Sider’s theory. We will now turn to objections.

4.2. Five Objections

None of these take the form of counterexamples. As Merricks (2013) observes:

[...] Sider’s general approach—as opposed to specific instances of that approach—is immune to counterexample. For suppose that Sider lists the “certain sorts.” You then come up with an absolutely compelling example of a proposition that is necessarily true and not of a sort on the list. Sider need not abandon his overall approach to reducing necessity. Instead, he could just add a new sort to the list to accommodate that example. Or suppose you come up with an absolutely compelling example of a true proposition that is not necessarily true and is of a sort on the list. Sider could just expunge that sort from the list. (Merricks (2013), p. 732.)

1. Necessity does not seem disjunctive or arbitrary (at least, not to this extent).

This is an objection centering on our intuitive grasp of the concept of necessity de dicto. It seems like this is a notion we can grasp, with the help of Kripke’s characterizations as supplemented in Chapter 1. Now, when we grasp this idea, it seems we are grasping a single, unified concept: necessary truths could not have been otherwise, no matter how things had turned out. This just doesn’t seem like a disjunctive matter. Nor does it seem as though, when we call a proposition necessary, we are applying a predicate whose meaning is arbitrary - although of course there are unclear or borderline cases, which we may perhaps make stipulations about to some extent.

This is not a knock-down objection, of course. Sometimes philosophy can reveal things to be other than they might seem. But I think it is hard to deny, if we are willing and able to grasp the concept of necessity de dicto and careful to hold in abeyance any of our pet theoretical proclivities which may suggest otherwise, that the notion does seem more unitary and less arbitrary than Sider’s theory would have us believe. And I propose that that should count as a mark against Sider’s theory.

Furthermore, insofar as appearance really is different from what Sider says the reality is when it comes to necessity, there is some explanatory work for Sider, or more generally the would-be quasi-conventionalist, to do here: why the discrepancy? As far as I know, no answer has yet been given.

2. The ersatz substitute worry.

A starting point for this worry is the unapologetically ad hoc nature of Sider’s successive extensions of the toy version of his approach that he begins with (where the “certain sort” of propositions he takes as “modal axioms” are just the mathematical truths). This process seems to be one of going back and forth between a growing list of types of propositions - the list at the heart of an increasingly disjunctive account - and our grasp of the real modal
notion of necessity. This gives rise to the worry that all we are doing is building an ersatz substitute for the real notion, by looking at the extensional behaviour of the latter and stipulating this behaviour into the account. No matter how far we pursue this strategy, the disjunctive notion we are building will remain fundamentally different in character from the notion whose behaviour we are imitating with it. Supposing that what we want from an ‘if and only if’ style of account of necessity de dicto is not some substitute for that notion, but a biconditional which gives us insight into the notion itself, Sider’s approach will never satisfy.

Something of this worry is even suggested by what Sider says about family resemblances, rehearsed above as point (6). The quasi-conventionalist could simply insist that each of the items on their list of the types of propositions which count as modal axioms is there as a brute fact - that’s just how the notion of necessity works. But, Sider says, the quasi-conventionalist ‘need not be quite so flat-footed’, and is ‘free to exhibit similarities between various modal axioms, just as one might exhibit similarities between things that fall under our concept of a game, to use Wittgenstein’s example’ (Sider (2011), pp. 288 - 289). This move, offered as an optional extra for the quasi-conventionalist, is plausibly in tension with the way Sider’s successively extended accounts are formulated. Just as the concept of a game - allowing for the sake of argument that it is a family resemblance concept - is plausibly not actually captured by any particular disjunction, but is as we might say inherently open-ended, it is also plausible that we should admit that the real “certain sort” or “modal axiom” notion doing the all-important work in Sider’s account - allowing for the sake of argument that it is a family resemblance concept - is not captured by any particular disjunction either.

This of course suggests a variant of Sider’s approach, where it is held that the “modal axiom” notion is a family resemblance concept, and admitted that any definite, disjunctive list of types of propositions could only yield, when plugged into the overall account, an ersatz substitute for the notion of necessity de dicto. This variant is not, or at any rate is less, vulnerable to the the ersatz substitute worry. But it is not clear whether it could really satisfy a philosopher who wants insight into the notion of necessity de dicto, let alone a philosopher with Sider’s motivations. For instance, can it really claim to be modally reductive? Perhaps, on the contrary, the family resemblance notion in question should be counted as modal. Furthermore, it may seem to yield an account which is insufficiently insightful - essentially all we are now getting is (Schema) itself, together with the pronouncement that the condition C is given by a family resemblance concept. Is there nothing more which can be said?

Relatedly, the question now arises: is it after all true that the notion in question is a family resemblance concept? What reason have we to believe that? (I will suggest, somewhat ironically given that I am on the whole much more admiring of Wittgenstein’s philosophy than Sider is, that it isn’t true. The notion playing this ‘condition C’ role, i.e. the notion which when combined with the notion of truth yields a notion playing Sider’s “modal axiom” role, can be defined in terms of a single necessary and sufficient condition.)

3. No iteration?

When Sider says early on in the modality chapter of his (2011) that the account he offers will be partial, there is a footnote to this remark which runs as follows:
(16) For example, the account defines a property of propositions that do not themselves concern modality, and thus is insufficient to interpret iterable modal operators.

This raises the question: how come, faced with this failure of coverage, Sider doesn’t simply make the same move with modal statements as he does with analyticities, “metaphysical” statements, and natural kind statements - namely, include them expressly in the account?

Perhaps the answer is that this would threaten the account’s claim of reductiveness. For it seems that in order to include modal statements on the list, we need the concept ‘modal’. The question then becomes: is ‘modal’ modal? If it is, Sider’s account is in serious trouble: it cannot, as a matter of principle, handle iterated modality. For remember, it is supposed to be modally reductive. And if iterated modality is a real, legitimate thing, then what use is a theory which gives us - by design - some extensionally correct answers but cannot handle this whole class of cases? It seems such a theory could give us an ersatz substitute for modality at best (to recall the previous objection). Its failure, if it is a failure, to be extendable to a salient class of cases should perhaps suggest to us that it is on the wrong track.

So, is ‘modal’ modal? It is an interesting question, and suggests interesting analogous questions about other kinds of concepts. One reason to think it is, is that we don’t seem to have a general way of saying what ‘modal’ means which doesn’t work by way of example. We seem to need examples of modal notions - necessity, or contingency, or possibility, or impossibility, or some combination of them - to do the job. To be sure, we could be said to be mentioning rather than using these notions in our explanations of ‘modal’, but is that any help? Don’t we need to use them in some broad sense in order to mention them in the appropriate way?

Another way out which may occur to the reader is to somehow delineate the modal statements using notions which are distinct from ‘modal’ and the like, but which fortuitously give the right extension. I am pessimistic about this. For a start, I can’t think of any good candidate notions. Furthermore, even if there were notions around which could do the job, wouldn’t using them for this purpose play further into the ersatz substitute worry described above? In particular, it seems like this strategy, while it may help Sider’s account deliver extensionally correct answers, would take the account (even) further from the real meaning of modal expressions, or the real nature of modal notions.

One possible strategy remains to be considered: accepting that ‘modal’ is modal and simply giving up the claim to full modal reductivity. From one angle, this seems not unreasonable; the way that ‘modal’ introduces modality, assuming it does, into the account, seems quite special and different from the way modality would be introduced if a notion of possibility or necessity were directly used. So perhaps there is room to claim that a broadly Siderian quasi-conventionalist account involving the notion of ‘modal’ as an unreduced modal element could still constitute a theoretical advance in some way. I have no knockdown objection to this, but I do want to suggest that once this concession is made, other objections, such as the first two considered here - (i) that necessity does not seem as
disjunctive or arbitrary as quasi-conventionalism would have us believe and (ii) the ersatz substitute worry - become all the more acute; I am not sympathetic with the following sort of move, but you might try to argue that biting those bullets is worth it if we get in return a complete reduction of modality, with its attendant payoff in eliminating puzzlement and vindicating certain sorts of metaphysical visions. You can’t do that anymore under the present strategy. Indeed, the whole spirit of the quasi-conventionalist approach seems to be in tension with allowing such a modal element into the mix.

In sum, there is reason to suspect that iterated modality, and the failure of any existing version of Sider’s approach to cover it, poses a serious threat to Sider’s approach in general.

4. Reductivity a bug, not a feature.

Essentially this objection is raised against Sider’s theory by Merricks (2013, p. 732). The objection is simply that, if we have reason to think that a modal notion like that of necessity de dicto cannot be reduced to non-modal notions, or if we just intuitively feel that to be right, then we should on that score alone be suspicious of Sider’s theory, since it purports to give a reduction. In making this objection, Merricks cites an argument he gives elsewhere (namely in Chapter 5 of his (2007)) for the conclusion that such modal notions indeed cannot be reduced to non-modal ones.

5. Questionable motivation.

As we said at the outset, Sider’s account is partly motivated by general puzzlement about modality\(^2\), and partly by a metaphysical vision. Both these facets of the motivation can be made the focus of criticism. The following is not supposed to constitute a sharp, incisive objection, but rather to cast some doubt on these general features of Sider’s approach.

Regarding general puzzlement: yes, modality is puzzling to philosophers. But perhaps this puzzlement is not to be treated exclusively by means of reduction (or, for that matter, by ‘if and only if’ analysis whether modally reductive or not). Indeed, pursuing reduction can even be seen as pursuing an easy way out - albeit one which may be impossible in principle. Perhaps the only real way forward, with parts of our puzzlement at least, is, rather than trying to reduce modality to non-modal terms, to work on our way of looking at modal concepts themselves, using philosophical methods other than reductive analysis. (One method which comes to mind is the method, due to Wittgenstein, of imagining simplified language games and comparing and contrasting them to ours. In the Brown Book some

\(^2\)Sider expresses this puzzlement as follows in his (2003, p. 184):

I can see that this colored thing is extended, and indeed that all colored things I have examined are extended, but where is the necessity, that colored things must be extended? Part of the puzzlement here is of course epistemic, and epistemic reasons for reductionism have already been mentioned. But there is a particularly metaphysical puzzlement here as well. In metaphysics one seeks an account of the world in intelligible terms, and there is something elusive about modal notions. Whether something is a certain way seems unproblematic, but that things might be otherwise, or must be as they are, seems to call out for explanation.
steps are taken towards doing this with modality, but only cursorily. I mention this to give a particularly concrete and well-known example of a possibly helpful method, but this is just one among many - I do not mean to suggest it could suffice all by itself.)

Non-modally-reductive accounts of necessity de dicto such as mine do not face this criticism, since they do not aim to clear up all of our puzzlement about modality, or even just some core of it, by means of an ‘if and only if’ style analysis. Nor to they aim even to point the way to such a clearing up. By being less ambitious on that front, they offer a more realistic hope of genuine theoretical progress on our understanding of de dicto modal notions - and of how they relate to other notions both modal and non-modal.

Regarding the metaphysical vision: it is beyond the scope of this thesis to criticize Sider’s Hume-influenced, Lewis-influenced metaphysical vision head-on. But we may note that, insofar as there may be grave problems with this sort of metaphysics for all we know given the present state of philosophical inquiry - nothing of the sort may be tenable, ultimately - there may also be problems with a highly ambitious approach to modality which is in service of this sort of metaphysics. More generally, perhaps there is reason to be dubious of any approach to modality based upon a metaphysical vision. One reason may be that the vision is, so to speak, too antecedent to modal considerations: perhaps one should let modal considerations shape one’s approach to metaphysical questions, rather than trying to explain modality (away, if you like) in terms of an approach to metaphysical questions which had its appeal quite apart from, or even in spite of, modal considerations. Another reason may be that the best way to make theoretical progress on the notion of necessity de dicto is to keep sweeping metaphysical visions out of it. We may do better to just pose and try to answer the question ‘Under what conditions is a proposition necessarily true?’ without restricting ourselves to notions which comport straightforwardly with some metaphysical vision. One way this may help is that it might free us up to throw a wider variety of conceptual resources at the problem - for instance, semantic notions or modal notions which may seem problematic against some special metaphysical backdrop but are actually quite in order.

That concludes our list of objections or worries. For two further objections, see Merricks (2013).

4.3. Conclusion

I think the cumulative effect of the objections canvassed above should be for us to regard Sider’s theory as highly problematic. But note that none of these objections threaten (Schema). This raises the question: what if these were a more soberly motivated, more theoretically satisfying (Schema)-embowering account available? Some other candidate for the condition C in (Schema) which avoids these objections?

---

4 Very briefly, these are that Sider’s account falsely implies that the debate about the modal status of natural laws is non-substantive, and that it falsely implies that ‘arguments from possibility’ are always defective, simply in virtue of their being arguments from possibility. (These do not speak much to my philosophical concerns, but they may be just the thing for others.)
You might wonder why I think this is a question worth following up. All we have seen so far is a bad account which embodies (Schema) - so why think there might be a good one embodying it too? The reason is that condition C seems like a promising point of application for semantic ideas, offering the prospect of an account of necessity de dicto that does justice to the ‘semantic hunch’ outlined in Section 1.3. - the hunch that broadly semantic considerations ought to come into the explanation of why a necessary proposition is necessary rather than contingent. Recall, the chief problem with analyticity approaches is that they preclude necessary a posteriori truths. But if we use (Schema) and tell a semantic story about condition C - perhaps a story on which it would seem to be a priori whether or not a given proposition meets condition C - this will not preclude necessary a posteriori truths; (Schema) explicitly and separately requires necessary propositions to be true, and nothing in (Schema) stops it being the case that whether or not certain necessary propositions are true is an a posteriori matter. Whether a given proposition meets condition C may be an a priori matter, but some condition C propositions may be true, and others false, and whether they are true or false may sometimes be an a posteriori matter.

In the next chapter, I will develop just such an account of necessity de dicto - one which embodies (Schema) and provides a broadly semantic account of condition C.

**Chapter 4 References**


5. An Account of Subjunctive Necessity De Dicto

5.1. The Account Introduced

So far in this thesis, we have defined our task - to give an illuminating account of the conditions under which a proposition is necessarily true, or failing that, reason to think this cannot be done - and considered some existing proposals at varying levels of detail. We have given especially detailed consideration to Lewis’s modal realism and Sider’s quasi-conventionalism. Finally, we ended our discussion of the latter with the suggestion that quasi-conventionalism be rejected in light of the problems it faces, but that further consideration be given to the following schema which it can be seen to embody (once you tease apart truth and ‘being of a certain sort’ in Sider’s notion of a ‘modal axiom’ and augment the schema with the parenthesis ‘(or a conjunction of such propositions)’, as discussed in Section 4.1.):

(Schema) A proposition is necessary iff it is, or is implied¹ by, a proposition which is both true and meets a certain condition C.

The problems with Sider’s account centre largely on his approach to understanding the condition C. That some correct filling out of the schema is possible has not been called into doubt. And, as I noted at the end of the previous chapter, condition C seems like a promising point of application for semantic ideas, offering the possibility of an account of necessity de dicto which does justice to the semantic hunch outlined in Section 1.3.

Now let us take a closer look at how this condition C might behave. This will serve two purposes. Firstly, it will help us get a better sense of (Schema)’s plausibility and theoretical attractiveness. Secondly, it will lead us to a proposal for how to understand condition C - and therewith to a new account of the conditions under which a proposition is necessarily true, an account which I believe holds out great promise of being both true and illuminating. One of the key features of this new account is that the property which will play the condition C role is broadly semantic - a matter of the nature or meaning² of the propositions which possess it.

We could regard condition C as being fulfilled by a lot of necessary truths - for example ‘Hesperus is Phosphorus’, ‘2 + 2 = 4’, ‘First-order logic is undecidable’ and ‘Cats are animals’. For these, the ‘or is implied by’ bit of (Schema) will not come into play in the verdict that they are necessary. With other necessary truths, such as ‘Either Hesperus is Phosphorus, or my hat is on the table’ (and it doesn’t matter here whether my hat is on the table - what matters is that this second disjunct is, if true, contingently so), we could say that they do not fulfill condition C, but are classed by (Schema) as necessary because they are

¹ Let me flag in advance that in Section 5.4. I will address circularity worries, and other worries, about the appeal to implication here.
² I put it this way (‘nature or meaning’) to remain as neutral as possible at this point about what propositions are. If they are meanings, ‘nature’ is best here. If they have meanings, ‘meaning’ applies. (If, as in the account of propositions I sketch in Chapter 6, a given proposition has whatever meaning it has intrinsically - i.e. in virtue of being the very proposition that it is - then either term works.)
implied by a proposition which is true and does fulfill C. In the present case, by ‘Hesperus is Phosphorus’.

(Schema), separately from the condition C part, requires necessities to be true, since it explicitly requires them to be, or be implied by, *true* propositions meeting condition C (and since anything implied by a truth is itself true). As a result, we are free to regard condition C as being fulfilled by a lot of false propositions as well - propositions which would, so to speak, be necessary if only they were true. That is, regarding condition C as being fulfilled by a lot of false propositions as well as true ones will not lead to counterexamples. We might for instance regard ‘Hesperus is not Phosphorus’, ‘Hesperus is Mars’, ‘2 + 2 = 5’, ‘First-order logic is decidable’ and ‘Cats are not animals’ as fulfilling condition C.

Let us just go along with the idea that condition C behaves as described so far, and ask ourselves: what is it that is distinctive about all these propositions which fulfill it? They aren’t all necessarily true. They aren’t all true. But it seems like they are all such that, if they are true, they are necessarily true. (Contrast the disjunction ‘Either Hesperus is Phosphorus, or my hat is on the table’; if we don’t know the truth-values of the disjuncts, but are just told that the disjunction is true, we can’t infer that it’s necessarily true: it could, for all we know, be that Hesperus isn’t Phosphorus, and the disjunction is true only because my hat is on the table, in which case it would be a contingent truth.)

Of course, here we are invoking the concept of necessary truth. That doesn’t mean we haven’t said anything interesting - think of recursive definitions in logic and mathematics\(^3\) - but still, perhaps we can go deeper. Now consider: when we think a proposition is necessarily true, what is characteristic about how we operate with it? Well, these are the propositions we say *couldn’t have been otherwise*. In other words, when we describe counterfactual scenarios, these propositions do not vary. This, I think, is the clue to how we might characterize the propositions fulfilling condition C without invoking the concept of necessary truth. What if we say something like the following: a proposition fulfills condition C iff, when we believe it, it does not vary when we describe counterfactual scenarios? No apparent invocation of the notion of necessary truth there.

You might worry that, while the notion of necessary truth is not being invoked here, the suggestion is just a roundabout way of saying that a proposition is true at all possible worlds, or in all counterfactual scenarios (where counterfactual scenarios are understood to be all and only subjunctively possible ones). But this is not what I mean. If we believe things like ‘Hesperus is distinct from Phosphorus’ we will be prepared to produce counterfactual scenario descriptions according to which Hesperus is distinct from Phosphorus, even though this is necessarily false, i.e. even though the description does not correspond to a subjunctively possible world. Thus counterfactual scenario descriptions do not, in the account I am going to put forward, play the role of possible worlds (ersatz or otherwise). (I will return to this point again in Section 5.3., when the account has been further explained.)

\(^3\)I.e., this initial suggestion is plausibly recursive rather than circular: if you understand what it is for the propositions fulfilling condition C to be necessary, you can apply this suggestion and work out that further propositions - those implied by true C-fulfilling ones - are necessary. So here the C-fulfilling ones play the role of base cases.
I think this is on the right track, but there are some things not quite right about the proposal. Perhaps the most glaring worry, at this point, is the reference to what ‘we’ do, which may make it look as though we make given propositions necessary rather than contingent by behaving in one way rather than another. The account will be modified to avoid this worry at the end of this section, and then further in Section 5.3., but first I want to address another worry: the appearance of the concept of belief may feel a bit suspect here. As though it is bringing in irrelevancies. Reflection also suggests that this proposal may give the wrong answers: what if we actually believe the negation of some proposition P, but are supposing that P is true (say, for the sake of argument)? In that case we might produce counterfactual scenario descriptions featuring P, even if both it and its negation are the sorts of things we should regard as fulfilling condition C. For example, suppose we actually believe that cats are not robots. We might still say to someone: ‘OK, grant that cats are robots. In that case, they couldn’t have been otherwise, but things could have been such that cats were robots and my hat wasn’t on the table’. Here it looks like we are producing counterfactual scenario descriptions according to which cats are robots, even though we don’t actually believe they are. So according to our current proposal, perhaps ‘Cats are robots’ fails to come out, as desired, as fulfilling condition C.

What I think this suggests is that we need here, not the notion of belief, but the notion of supposing or granting or holding true. Believing and holding true often go together, but they can also come apart - for instance when we suppose something we don’t believe for the sake of argument.\(^4\) I hope also that this notion feels less out of place in the present context than did the notion of belief. (Holding true seems as it were to be a more contained, narrower affair. This strikes me as a good thing.) Following this suggestion gives us the following proposal:

A proposition fulfills condition C iff, when we hold it true, it does not vary across counterfactual scenario descriptions.

Plugging this into (Schema) we get:

A proposition is necessary iff it is, or is implied by, a proposition which is both true and such that, when we hold it true, it does not vary across counterfactual scenario descriptions.

That is a bit of a mouthful, though. Let us call condition C so understood inherent counterfactual invariance. To make it clearer what ‘inherent’ is doing in this piece of terminology: we might say that ‘Either Hesperus is Phosphorus, or my hat is on the table’ is counterfactually invariant when held true by holding it true that Hesperus is Phosphorus (and perhaps the other disjunct too), but counterfactually variable if it is being held true by holding Hesperus not to be Phosphorus but my hat to be on the table. ‘Hesperus is Phosphorus’ itself, on the other hand, is counterfactually invariant however it is held true, so we call it ‘inherently counterfactually invariant’.

---

\(^4\) I leave this notion of holding true intuitive here. I mean it to cover both cases where we merely grant or suppose something for the sake of argument without believing it and cases where we do believe the thing in question. For discussions of these notions, see Cargile (1995) and Green (2000).
We have now arrived at my preferred formulation of my account of necessity *de dicto*:

A proposition is necessary iff it is, or is implied by, a proposition which is both true and inherently counterfactually invariant.

While this is the formulation I will settle on in this thesis, we still have a way to go in spelling out what it means. The main notions involved seem to be: the notion of a proposition, the notion of implication, the notion of truth and the notion of inherent counterfactual invariance. Of these, the invocation of the notion of truth is the only element I propose to leave as is and not discuss further. The notion of truth seems so basic and so serviceable already that we can just let it do its work here - furthermore, problematizing it would take us far afield.\(^5\) The notion of a proposition will be given extended attention in the next chapter, so for now let us just stick with an intuitive version of it - propositions are the sorts of things that are true and false. The notion of implication, and its role in this account, will be discussed in Section 5.4. below, but talk of implication or following logically is familiar enough and serviceable enough that we can for now just leave it intuitive while still making good progress.

The notion of inherent counterfactual invariance on the other hand is new, and still not very clear. I propose to proceed as follows. Presently, I will make what I think is at this point the most urgent clarification regarding this notion. Then, with the hope that the basic idea of inherent counterfactual invariance has been conveyed, I will in Section 5.2. apply the account (understood, albeit, in a somewhat provisional, rough-and-ready way) to a series of cases, to give a better sense of how it works, and *that* it works. Following that, I will turn in Section 5.3. to a more detailed consideration of inherent counterfactual invariance and try to account for it more precisely. In Section 5.4. I will turn to consider the notion of implication and its role in the account. The main exposition of the account having been given, I will then review the account in Section 5.5. In Section 5.6. I will briefly outline a fallback position which is interesting in its own right. In Section 5.7. I will consider and reply to some possible objections, and conclude the chapter in Section 5.8. with some remarks highlighting the attractiveness of the account.

The most urgent clarification of the notion of inherent counterfactual invariance is as follows. Our provisional spelling out of what it is for a proposition to be inherently counterfactually invariant - namely, that when we hold it true, it does not vary across counterfactual scenario descriptions - suffers from the defect of making it appear that which propositions are inherently counterfactually invariant is subject to the behaviour we adopt in our practise of describing counterfactual scenarios. It may even seem like this is really a relational property: a proposition is inherently counterfactually invariant in relation to one way of doing things, but the same proposition is not inherently counterfactually invariant in relation to another. All this I want to deny.\(^6\) Rather, the idea is that inherent counterfactual invariance is a matter of

---

5 For a classic statement of this sort of attitude to the notion of truth, see Davidson (1996).

6 You could have an account where ‘inherently counterfactually invariant’ is understood more in this way. Such an account may be attractive to those like Sider who are keen to say that there is something arbitrary or conventional about necessity *de dicto* (see the previous chapter) - but this is not the account I advocate and is alien to my way of thinking about these matters.
the meaning or nature of the propositions which possess it. In order to capture this, we might supplement our provisional spelling out as follows: a proposition is inherently counterfactually invariant iff, when we hold it true, then no matter what we do, it does not vary across counterfactual scenario descriptions. (To anticipate somewhat, in case this sounds mistaken: my account ends up relying on an unanalyzed distinction between genuine and non-genuine counterfactual scenario descriptions (introduced in Section 5.3.), and here I mean just the genuine ones.) This as it were throws the focus back onto the proposition itself.

Our explanation of inherent counterfactual invariance, as it stands now, still leaves plenty to be desired. We will come back to that soon, but I hope that the basic idea has now been conveyed to some extent. Let us now turn to some applications of the proposed account of necessity de dicto to get a more concrete idea of its functioning. As we walk through the following examples, some of the claims made regarding how parts of the analysis turn out with respect to them may seem open to question. Much of this potential dubiousness should be cleared up in due course, as the notion of inherent counterfactual invariance is made clearer, but for now I will just be a bit dogmatic. In this connection, the point foreshadowed in the parenthesis immediately above is important to remember: I will end up invoking an unanalyzed distinction between genuine and non-genuine counterfactual scenario descriptions.

5.2. Some Applications of the Account

Consider the propositions ‘$1 + 1 = 2$’, ‘First-order logic is undecidable’, ‘Hesperus is Phosphorus’ and ‘Cats are animals’. They are true, and they are inherently counterfactually invariant: if we hold them to be true, then when describing counterfactual scenarios, they remain fixed - we don’t produce counterfactual scenario descriptions according to which $1 + 1$ doesn’t equal 2, or first-order logic is decidable, Hesperus is not Phosphorus, or cats aren’t animals.’ Thus my account gives the right answer: these propositions are necessary.

Consider the propositions ‘$1 + 1 = 3$’, ‘First-order logic is decidable’, ‘Hesperus is Mars’ and ‘Cats are robots’. These are false, but they are inherently counterfactually invariant just like the examples of the previous paragraph. The account filters these out, so to speak, from being classified as necessary by means of the truth requirement. So again, we get the right answer: these propositions are not necessary.

Consider the propositions ‘Either Hesperus is Phosphorus, or my hat is on the table’ and ‘Everything is either such that it is either not a cat or is an animal, or such that it is either less than 100 kilograms in weight or not in my room’. They are both true, irrespective of whether my hat is on the table, or whether everything in my room weighs less than 100 kilograms: Hesperus is Phosphorus, and all cats are animals. They aren’t inherently counterfactually invariant: you could hold the first one true while holding it false that Hesperus is Phosphorus but holding it true that my hat is on the table. In that case, you would be prepared to produce

---

7 If you are tempted to object here, keep in mind (i) that I will be making crucial use of a distinction between genuine and non-genuine counterfactual scenario descriptions and (ii) that I am not wedded to the cat case as a claim about ordinary language claims in current English.
counterfactual scenario descriptions according to which neither is Hesperus Phosphorus, nor
is my hat on the table. Likewise, you could hold the second one true by holding it false that
everything is such that it is either not a cat or is an animal (that is, holding it false that all cats
are animals - or holding it true that there are non-animal cats), but holding it true that
everything is either less than 100 kilograms in weight or not in my room (that is, holding it true
that everything in my room weighs less than 100 kilograms). In that case, you would be
prepared to produce counterfactual scenario descriptions according to which not everything
is either such that it is either not a cat or is an animal, or such that it is either less than 100
kilograms or not in my room; these descriptions might for example have it that someone has
assembled a car in my room. While they aren’t inherently counterfactually invariant, both of
them are implied by a proposition which is both true and inherently counterfactually invariant;
in the first case, by 'Hesperus is Phosphorus', and in the second case, by 'Everything is such
that it is either not a cat or is an animal'. Thus the account gives the right answer: both
propositions are necessary.

Now that we have seen how the account functions with respect to some examples, let us
return to the all-important notion of inherent counterfactual invariance and try to make it
clearer.

5.3. Inherent Counterfactual Invariance Further Clarified

For brevity, I will use ‘ICI’ as an abbreviation for both ‘inherent counterfactual invariance’ and
‘inherently counterfactually invariant’.

Above, we left off our discussion of ICI with the following provisional definition: a proposition
is ICI iff, when we hold it true, then no matter what we do, it does not vary across
counterfactual scenario descriptions. We will now try to improve upon this. Some of these
improvements will be motivated simply by a desire for greater perspicuity and clarity, but
later in this section we will consider a more pointed worry about the definition letting too
much in, and there our response will be motivated by a desire to deliver the right verdicts. So
we are in a sense leaving the most important work till later in the section - the hope is that
making the less crucial clarifications first will ease the task.

To start off with, let us try to make more precise what is meant by ‘does not vary across
counterfactual scenario descriptions’. One thought which may suggests itself is that, if a
proposition does not vary across counterfactual scenario descriptions, then it appears in, or
is true according to, every counterfactual scenario description (‘CSD’). But this is not what
we should say: CSDs are not in general supposed to be maximal, like the linguistic
ersatzer’s “possible worlds”. They can often be less idealized, more down-to-earth things.
For example, if I say ‘I could have had toast for breakfast this morning’ (where this isn’t
meant epistemically, but as a description of an alternative way that things might have gone),
I have given a CSD, albeit a brief one. So we should instead understand this not varying as follows: the proposition’s negation does not appear in any CSD.

This yields: a proposition is ICI iff, when we hold it true, then no matter what we do, the proposition’s negation does not appear in any CSD that we produce.

We are making progress, but this formulation still leaves us with some pressing questions. For the remainder of this section, I will consider such questions one at a time, refining and explaining things in response.

Question 1: Is the reference to us (‘we’) essential? (It may seem a bit out of place. or at least like something that could be abstracted away from - especially given the fact that we have explicitly distanced ourselves, with the addition of ‘no matter what we do’, from the idea that our behaviour can influence whether a proposition is ICI.)

My first suggestion is that maybe this isn’t a big problem. As long as we have an explanation we can understand, which seems reasonable, and gives the right answers, perhaps we shouldn’t get too hung up about this. (Of course, another worry might be that it hasn’t been made clear who or what is included as being one of us here. A good answer to that seems to be: any language user.)

That said, if the reference to language users can be avoided, the essentials of our account may become clearer if we do avoid it. Its main function in the above formulation seems to be to link propositions held true to CSDs; we can’t just say a proposition is ICI iff its negation doesn’t appear in any CSD. This would rule out most if not all of our examples, propositions which are paradigm cases of ICI; as long as you can hold their negations true, you can then produce CSDs containing those negations, which by the ‘only if’ part of this faulty account would yield the undesired result that they aren’t ICI.

---

8 It is obvious that this example counts against the ‘appears in’ version of the proposal. It is perhaps less obvious that it also counts against the ‘is true according to’ version, but I think that on reflection, and if we stick with a common sense understanding it ‘is true according to’, it does. Just think how wrong it sounds to say that, according to the description embedded in ‘I could have had toast for breakfast this morning’, Canberra is in the Australian Capital Territory.

9 You may worry here that if the proposition in question is already a negation, then its negation may fail to appear in any CSD, not because the proposition in question is ICI in the sense we are reaching for, but because its negation would exceed the limits of what a CSD we produce could contain, i.e. by being too long or having too many negations in it. I have two responses to this. Firstly: as will become clear, I am talking not just about CSDs that actually get produced, but possible CSDs, and it seems plausible that, on the understanding of possible CSDs’ we want here (more on which in the answer to Question 2 below), there is no good reason to think that that this problem could arise. We are talking, not about what is actually feasible for us given our resources, but about what descriptions are producible in principle, and how could a proposition producible in principle not have a negation which is producible in principle? Secondly: if you aren’t convinced by the first response, amend the present proposal so that it runs: a proposition is ICI iff, when we hold it true, then no matter what we do, the proposition’s negation - or, if the proposition is a negation, the proposition it is a negation of - does not appear in any CSD that we produce.
The idea is clear enough: when we produce CSDs, we are holding certain things true. So we can abstract from the producers of the CSDs and form a concept linking CSDs directly to propositions held true: we may speak of the propositions held true for a CSD.

It is worth emphasizing that this holding a proposition true for a CSD is not the same thing holding it fixed across counterfactual scenarios. For example, when Kripke says\(^\text{10}\) things like ‘Given that Hesperus is Phosphorus, it couldn’t have been otherwise’, the first part is indicating that it is being held true - true in the actual world, so to speak - that Hesperus is Phosphorus. The second part is making the distinct point that, if we hold this true, then we hold it fixed when describing counterfactual scenarios. Thus, when it comes to non-ICI propositions, holding them true when producing a CSD - holding them true for that CSD - does not mean they won’t be contradicted in that CSD. For instance, if I say ‘I’ll grant that it was a bad performance, but it would have been good if I had only omitted the opening lines’, it is being held true for the CSD occurring in the second part of that sentence that it was actually a bad performance, while it is true according to the CSD that the performance was not bad, but good - i.e. in the scenario described by the CSD, the performance was good.

So, we now have:

A proposition is ICI iff its negation does not appear in any CSD for which it is held true.

Thus we have eliminated any explicit mention of speakers. Perhaps the way we should think about this issue is as follows: this last formulation is cleaner and sticks more to the essentials, but it is more technical and a bit harder to understand. So, the earlier formulation involving ‘we’ still has some heuristic value as an expeditious way of getting the basic idea of ICI across. (That is why I began with it and then introduced this refinement.)

**Question 2:** Is there a modal element here, and if so, how should we understand this? (Surely it wouldn’t do for ‘ICI’ to be defined in terms of just the CSDs which actually get produced.)

Yes, I think we should admit that there is a modal element here. We are appealing to a space of possible CSDs.\(^\text{11}\)

Before we continue: it may be that there is an alternative here. You could have a more abstract, Platonic conception of CSDs and think of them as all existing whether or not they are ever produced by language users. On the other hand, it could be argued that such a Platonic conception is tacitly modal. Be that as it may, I’m not afraid of a modal element in the account, and I think it’s politic to just allow that there is such an element at this point, and be upfront about it.

---

\(^{\text{10}}\) E.g. in Kripke (1980, p. 47, p. 103, p. 123, p. 125, and twice on p. 126).

\(^{\text{11}}\) I do not mean to say that this is an irreducibly modal element - only that no reduction of it comes with my account. I happen to suspect that it *is* in some sense an irreducibly modal element, but that is no part of the account.
Now, happily going along with the idea that here we are really talking about possible CSDs, the question arises: how should we understand this talk? Hopefully, it makes some intuitive sense as it stands. One reason that this question arises, I think, is because there is more than one option here for cashing this talk out a bit further. I want to distinguish two.

The difference between these two ways of cashing out talk of possible CSDs concerns what we think of as being, so to speak, the host of the possibilities. For instance, if I speak of all the possible configurations of a Rubik’s Cube, we may call the Rubik’s Cube the host of these possibilities. Now, in the case of possible CSDs, two quite differently flavoured options suggest themselves. (I do not mean to say there are only two options here - these are just the ones I find most salient.) We may think of the host as being the world at large, or something more circumscribed, like a system of language (or thought, or representation).

On the first option, we explain ‘possible CSDs' by saying something like: here we mean all the CSDs that there can be. Here we aren’t appealing to any particular host of these possibilities, other than - if you like - the world itself. On the second option, we specify something. We may relativize the host to the proposition whose ICI status is in question - regarding it, for instance, as being the system of language (or thought, or representation) to which that proposition belongs. So we might say that a proposition is ICI iff its negation does not appear in any CSD producible by the system that the proposition belongs to for which it is held true.

I think it is fairly unimportant which conception we let dominate. They seem to come to the same thing. From a rhetorical point of view, each has an advantage and a disadvantage. Thinking of the host as, if anything, the world, feels more metaphysical, so to speak - more like a bold and potentially dizzying abstraction. Thinking of the host as a system of language (or thought, or representation) on the other hand avoids some of this feeling, but at the cost of appealing to something which may raise difficult questions of its own. The second option, it is worth noting, will be seen to chime with the conception of propositions I will offer in Chapter 6 - on that conception, one aspect of a proposition’s meaning is reckoned as the role it plays in the system of language to which it belongs.

To be clear: I am not here trying to give a theoretical account of talk of possible CSDs. Rather, I am just appealing to it for the purposes of a theoretical account of necessity de dicto. I just need it to be intelligible and legitimate. If we want some sort of theoretical account of it as well, then that is a task for another day.

I hope this position - that that is a task for another day - seems legitimate already. But I also have a couple of suggestions and reminders in case it seems dubious.

One reason this may seem dubious is historical (though the history is mostly quite recent). Since Kripke isolated the notion of subjunctive necessity de dicto and distinguished it from the notion of apriority, philosophers pursuing big, grand projects directed at explaining all modality, or at least all non-epistemic modality, have often begun with this notion. For instance, Sider’s quasi-conventionalism, which is meant to show how to reduce all such modality, begins there. Lewis’s account begins with notions of necessity and possibility de
dicto as well.\textsuperscript{12} For this reason, it would come very naturally to a contemporary philosopher assessing my account here, if they do not pay sufficient attention to my explicit statement of my task (given in Section 1.2.), to treat it as if it is another one of these big, grand projects. It requires active effort not to think of it that way. One of my main themes in this thesis is that this effort is well worth it: by not asking so much of an account of subjunctive necessity \textit{de dicto} - I am tempted to say `by not asking the impossible' - we open the way to some genuine, solid theoretical progress, deepening our understanding of the notion and how it relates to other notions, and vindicating the semantic hunch. My account may not be quite as exciting and provocative as, say, modal realism, but it more than makes up for that by being \textit{true} - or more diplomatically, being plausible.

Relatedly: my talk here of possible CSDs, like all modal talk, \textit{can} give rise to metaphysical puzzlement. However, and this is important, it can also just be taken in its stride - used rather than problematized. It is a question of the attitude we take, and of our expectations. And on the matter of expectations: my account makes no claim of holding the key for clearing up all metaphysical puzzlement surrounding modality.

This brings up another general theme of this thesis, which is that we need to be wary of a tendency to not do enough distinguishing of the varieties of modality (although awareness of the richness of the terrain here seems to be growing, as indicated by the appearance of Kment’s (2012) encyclopedia entry on the topic). Kripke’s distinction between epistemic and non-epistemic modality isn’t the only one we should be making. There are different kinds of non-epistemic (objective, metaphysical) modality as well.

Here is how I would put it. You can, by appealing to a bit of what we might call indicative \textit{de re} modality (`possible CSDs'), plus some other notions such as that of a CSD, the notion of truth, and the notion of implication, give an accurate and illuminating account of subjunctive necessity \textit{de dicto}, doing justice to the semantic hunch expressed in Section 1.3., in the form of a statement of the conditions under which a proposition is subjunctively necessary.

Whether or not you would also put it this way, I want to emphasize how natural it is to think that the modality being appealed to is not, at least on the face of it, subjunctive \textit{de dicto} modality. We are talking about a space of possible things there might be, in a way that feels somewhat analogous to talk of a space of possible configurations of some device. (Compare: this Rubik’s Cube can be put into approximately 43 quintillion configurations.) Now, someone might maintain that, on analysis, we are really talking about \textit{de dicto} modality here, i.e. about the status of some proposition, and about subjunctive modality, i.e. not what may be, but what might have been. But I see no reason to think that is the case.

Finally, I want to register a fallback position here. You might think this account is circular on this score, despite these arguments - that there is some kind of appeal here to the notion

\textsuperscript{12} As far as I know, Lewis was never very clear or explicit about the notions in question being subjunctively as opposed to indicatively understood, but it is natural to think that his target analysandum ‘Necessarily, \textit{P}’ can be understood as trafficking in basically the same notion of necessity \textit{de dicto} that Kripke isolated. (Perhaps it is flexible and can be understood a number of ways, this being one of them.)
being analyzed. If I could not convince you otherwise, I would still urge that this doesn’t make the account trivial or theoretically worthless.

Question 3: Regarding ‘its negation does not appear’: what about the fact that, when we describe counterfactual scenarios, we often use subjunctive language and sub-propositional phrases? (It seems like ‘I might not have gone to the movies’ should perhaps count as a CSD in which the negation of ‘I went to the movies’ appears, and yet in a narrow sense it seems not quite right to say that the negation of that proposition does appear there.)

One response to this worry would be to have a suitably wide conception of a proposition’s negation. (It may be a bit tricky to define this explicitly and exactly, but that some such thing is available seems plausible. Note that the definition need not be purely syntactic.)

Another response would be to stick to a narrower conception of a proposition’s negation, and introduce a relation linking negations to the appropriate phrases or components of CSDs. We might for instance call these phrases or components cognates of the negations in question, and then amend the relevant part of account to read something like ‘its negation (or a cognate thereof) does not appear’. (Again, an exact definition of this relation may be a bit tricky, but that some such thing is available seems plausible.)

A third response is as follows. We can appeal to the idea of a “normal form” for CSDs. It might as well just take the form of a list of propositions, themselves free of any modifications signalling that this is a subjunctive or counterfactual scenario description. So, if we want to use this normal form, we can say, instead of ‘I might not have gone to the movies’, ‘The following might have been the case: ~(I went to the movies)’. Now, since there will always be a possible normal form CSD corresponding to any other, we can just rely on them.

So we have at least three possible responses here. Which one we go for seems to me like a matter of unimportant detail. I tentatively prefer the third one, because it doesn’t require us to modify the account itself in any way or give promissory notes - it is just a consideration which shows that the account as it stands won’t give any wrong answers because of this wrinkle.

What exactly is a CSD, and what makes a CSD a CSD?

The first thing to clarify here is that, despite what the term ‘CSD’ may suggest, it is not the case that a CSD is just a description of a scenario which doesn’t describe things as they actually are. There are descriptions which describe things as being other than they actually are which aren’t CSDs, and there are CSDs which do describe how things actually are.

For example, if someone asks ‘What happened at the party?’ and I say ‘John turned up with a cake and then went home early’, then whether or not this is true, it isn’t a CSD. Likewise if I say ‘Maybe John turned up with a cake and then went home early’ or ‘Let’s assume that John actually turned up with a cake and then went home early’. On the other hand, if I happen to think (or simply be supposing for the sake of argument) that John went to the party when he actually didn’t, but rather sent someone else to deliver his cake, then if I say ‘If John hadn’t gone to the party, it wouldn’t have gone as well’ or ‘Imagine if John hadn’t
gone to the party but sent someone else to deliver his cake’, then these expressions do contain CSDs, even though the CSDs, not that this was the point of them, do happen to describe how things actually are.

A further subtlety is that, despite what the above may have suggested, it seems like we don’t have to think or even suppose anything contrary to a given CSD for us to produce one and treat it as such. This is brought out by the famous example Anderson (1951) used to show that it’s not always the case that counterfactual (i.e. subjunctive) conditionals imply or suggest that their antecedents are false. Some doctors are standing over a patient and one says ‘If he had taken arsenic, he would have developed these symptoms’, suggesting that perhaps the patient has taken arsenic. The antecedent here should I think be counted as a CSD, even though the doctor saying it doesn’t think, and isn’t even supposing, anything to the contrary of it.

For these reasons ‘subjunctive’ might be a better word than ‘counterfactual’. On the other hand, that choice has its own drawbacks: it is less vivid (‘counterfactual’ isn’t wholly misleading), feels a bit arcane, and it is connected with subtle grammatical disputes in linguistics which I suspect aren’t relevant to us here. I have decided to opt for ‘counterfactual’ despite its drawbacks.

So we’ve made two negative points about CSDs. A further negative point to make is that we are not trying to characterize CSD-hood syntactically: while there are characteristic markers, at least in some contexts (e.g. ‘had’), I am not proposing that any particular word or structural feature of expressions is such that its presence ensures CSD-hood, or its absence ensures non-CSD-hood. Rather, CSD-hood is a broadly semantic property, and whether a piece of language has it depends on what it going on around it as well as how it is constituted.

Now for a more positive point: the distinction made by two-dimensional semanticists between considering a scenario as actual and considering it as counterfactual seems to be essentially what we are after here. (The terminology is due to Davies and Humberstone (1981) and has been taken up by prominent two-dimensionalists such as Stalnaker, e.g. in his (2001), and Chalmers, e.g. in his (2006).) CSDs, we may say, are descriptions of scenarios of a sort where we are considering these scenarios as counterfactual. And again, as the Anderson example shows, this “considering as counterfactual” is not to be confused with presupposing, suggesting, or implying that the scenarios don’t obtain (despite what the use of ‘counterfactual’ may suggest).

So, that is the notion. The terminology here is a bit tricky, and zooming in on the notion as we have done here can make it seem quite delicate, but it is nevertheless based directly on a real notion already in use in analytic philosophy - that of considering a scenario as counterfactual.

My purpose here was just to avoid some misunderstandings and convey the notion. There may well be more to say about what the considering-as-counterfactual vs. considering-as-actual distinction amounts to, what grounds it, etc., but these are questions
for another day. (Compare: you don’t have to answer the philosophical question ‘What makes a token a name?’ to be able to appeal intelligibly to the idea of a name. Or in place of ‘name’ you could put ‘predicate’, or ‘subjunctive conditional’, or one of many other categories found in analytic philosophy.)

There is a further issue, still undealt with, which may be behind an instance of the above question, having to do with the bounds or extent of the notion of a CSD. This is brought out by the following question.

_Sometimes we talk, so to speak, counterpossibly - counterpossibly by our own lights, even. We may talk, for instance, about what would have been the case had something impossible been the case, e.g. ‘If a logician had succeeded in giving an effective decision procedure for first-order logic, they would have become famous for it’. Doesn’t this make your account give wrong answers?_

For all that has been said so far, there is a real issue here. What this issue reveals is the need for a notion marking off the CSDs, or the ones we want to include in our consideration, from CSD-like things we do not want to include. Whether we call all these ‘CSDs’, or just the desired ones, seems an unimportant terminological matter.

I call the descriptions we want to exclude ‘non-genuine CSDs’, and reserve ‘CSD’ or if I want to be more explicit ‘genuine CSD’, for the ones we do not want to exclude. (I do not mean, with the terminology ‘non-genuine’, to denigrate non-genuine CSDs, or talk involving them. These may often have value.) I do not have a definition for this notion of genuineness - as my account currently stands at least, it is a primitive notion. Let us try to get a feel for it.

Perhaps the first order of business here is to make it clear that we are not here simply smuggling in the notion of subjunctive possibility _de dicto_ - which of course would constitute a serious threat to the significance of our account.\(^{13}\) Plenty of putative CSDs containing necessarily false propositions count as genuine. For example, if I believe that Hesperus is not Phosphorus, I might talk of a counterfactual situation where Hesperus comes closer to Phosphorus and then moves away again. Or I may simply throw in ‘Hesperus is not Phosphorus’ (which I in this example think also holds of things as they actually are, but that doesn’t matter) as part of the description. This could be a perfectly genuine CSD. (This, by the way, highlights the point - already made in Section 5.1. above - that CSDs in my account do not play the role of possible worlds, ersatz or otherwise.) On the other hand, if I am holding it true for a putative CSD that Hesperus _is_ Phosphorus, these inclusions _would_ make it a non-genuine CSD. Conversely, if I am holding it true that Hesperus _isn’t_ Phosphorus, then a putative CSD I produce according to which Hesperus _is_ Phosphorus will be non-genuine.

A more general point which may help us grasp this notion: when everything being held true for a CSD is in fact true, then that CSD will be genuine iff what it says is subjunctively

---

\(^{13}\) This may not automatically be a dealbreaker, but it would raise the suspicion that the appeal here is doing all the work and other elements in the account are idle cogs.
possible. It is when things are being held true for a CSD which are not in fact true that you get genuine CSDs containing subjunctive impossibilities.\footnote{It seems to me that the notion of a genuine CSD may be useful for a variant definition of ‘rigid designator’ more fundamental in some respects than Kripke’s. Let me explain. Suppose you have read \textit{Naming and Necessity} and as a result acquired a notion which you express with the words ‘rigid designator’, ‘rigid designation’, ‘rigid’, ‘rigidity’, etc. But suppose this notion has become quite intuitive for you, so that you don’t translate it in thought using Kripke’s characterization involving the notion of a possible world.

Now, suppose you come across an expression such that you cannot tell whether it is a real definite description or a name which has the form of one (e.g. ‘the Jonesenator’ used as a frivolous name for a man called ‘Jones’). Let’s say it’s ‘the opener’ (or ‘the Opener’, but let’s suppose you encounter it in speech so that you can’t tell whether ‘opener’ should be capitalized in writing). Now suppose you hear Bob using the expression in the following context: ‘Imagine if things had instead gone as follows: Julia is running away from Mary and calls the opener, who in this scenario is Harry - not Tom like in real-life …’. On the basis of this, you might say ‘OK, “the opener” isn’t rigid’.

But suppose that Julia and Mary are actually the same person, and you know this, but Bob thinks otherwise, and you know that he does. (Also suppose for the sake of argument that it’s impossible to run away from yourself.) In that case, the scenario being described is not subjunctively possible: Julia couldn’t have run away from Mary, because Julia is Mary.

So, how did you arrive at the conclusion that ‘the opener’ isn’t rigid? It could be that your use of ‘rigid’ conforms exactly to the Kripkean definition, and that you inferred from the appearance of ‘the opener’ in the above counterfactual scenario description that, while that description doesn’t hold in any possible world, there are descriptions which are relevantly the same but which are possible, and so ‘the opener’ must designate different objects at different possible worlds. But it seems to me that your procedure could easily have \textit{not} been like that. It could have been a more direct inference than that. You could have reasoned: ‘Look, there’s “the opener” in a genuine counterfactual scenario description being used to designate a different object from the one it is actually supposed to designate - so it’s not rigid’. If this is what you did, then I would suggest that your working notion of rigidity is actually not exactly the one Kripke defined, but another one - perhaps a more fundamental one - where genuine counterfactual scenario descriptions are what really matters, not possible worlds. The CSDs don’t actually have to be \textit{possible} for them to be relevant, just genuine. In other words, what matters is how a designator is systematically used with respect to CSDs which are, so to speak, possible \textit{from the point of view of what is being held true for the purposes of them}.

This suggests to me that the notion of a genuine CSD, which plays a crucial role in my account of necessity \textit{de dicto}, could also be of use in defining a notion of rigidity that is more fundamental in some respects than Kripke’s official notion. At least, it seems to me that it might be more fundamental from the point of view of categorizing different kinds of designators based on how they work; in other connections, such as arguments from rigidity to conclusions about what is and is not necessary, the official Kripkean definition might be more to the point. More tentatively, putting aside questions of what is more fundamental, I want to suggest that the notion of a genuine CSD might be of use in the project of distinguishing and clarifying various notions in the ballpark of Kripke’s official notion of rigidity. (Existing contributions to this project include discussions of \textit{de facto} vs. \textit{de jure} rigidity, discussions of obstinate vs. non-obstinate rigidity, and discussions dealing with the extension of the notion of rigidity to non-singular terms. For an overview, see LaPorte (2016).)
though I am in fact a human. I don’t think there’s any real problem here, but this might seem a bit odd. I have two suggestions. Firstly, we have to look at the role being played by this notion in the overall analysis, and be wary of thinking of ‘genuine CSD’ along the lines of ‘subjunctively possible CSD’ - that is not a good way to think about it in general. (Perhaps my terminology here is partly to blame, but I haven’t been able to come up with anything better. I suspect that the main thing needed might just be more familiarity with these notions.) Secondly, we have to emphasize that what is (apparently) being held true in some of these worrying cases may be pretty far out, and thus difficult or even impossible to really hold true.

This notion of a (genuine) CSD clearly has some of the character of the notion of subjunctive necessity _de dicto_. Like with that notion, there will be plenty of unclear cases, but that should not be allowed to prevent us from grasping it. Unlike the notion of necessity _de dicto_ this one, as well as the notion of ICI of which it forms a part, seems _a priori_ tractable; the element that makes questions of necessity _de dicto_ sometimes _a posteriori_ has been factored out, so to speak, in the form of the truth requirement. The implication clause has also been distinguished as a separate element.

Some may be skeptical that there is a distinction here at all between genuine and non-genuine CSDs, or that there are cases on the non-genuine side. But such people will probably also be skeptical that anything is really absolutely necessary _de dicto_ - or if they aren’t, I suggest that their overall best and most consistent option, if they really want to deny any distinction (where both sides are instantiated) between genuine and non-genuine CSDs, is to deny also that anything is really necessary _de dicto_, or reject the notion of necessity _de dicto_ altogether. Taking the first option, they can still accept my account\(^\text{15}\) as an account of what it would take for a proposition to be necessary _de dicto_. I think they’d be missing something - would have a blind spot, so to speak - but at least they could agree that my account is true.

### 5.4. Implication and its Role in the Account

The account I am proposing says that a proposition is necessary if it is, _or is implied by_, a proposition which is both true and ICI. What is the purpose of the italicized clause, and what exactly does it mean?

Something of its purpose should already be clear from some of the examples worked through early in this chapter. In general terms, the clause is required for propositions which are necessary but not ICI: propositions which may be held true in such a way that their negations _can_ appear in counterfactual scenario descriptions for which they are held true. The principal sort of case is that of a disjunction like ‘Either Hesperus is Phosphorus, or my hat is on the table’.

This sort of case may suggest, instead of an appeal to implication, some sort of mereological clause; it strikes us in this case that the proposition has a _component_ which is ICI. But this

---

\(^{15}\) Minus the bit about genuine as opposed to non-genuine CSDs, perhaps - or they could just say that there are no non-genuine CSDs
will not always be so, as the following example makes clear: ‘Everything is either such that it is either not a cat or is an animal, or such that it is either less than 100 kilograms in weight or not in my room’. This proposition doesn’t have any components which are themselves propositions. But it is necessarily true, and not ICI.

These problems for an analysis like ours but lacking the implication clause arise, we might say, because necessity is closed under implication; the logical consequences of necessary truths are themselves necessary. Bringing in the implication clause reflects this fact and gives us what we need. Our first case is implied by ‘Hesperus is Phosphorus’, which is ICI and true, and our second case is implied by ‘Everything is either such that it is either not a cat or is an animal’, which is ICI and true.

Are there other sorts of cases for which the implication clause is needed? I do not know, but I conjecture that all such cases will have a disjunctive character. However, they need not all themselves contain a disjunctive word like ‘or’. I may define a predicate as follows:

\[ x \text{ is P} = df x \text{ is either identical to Phosphorus or was talked about by me yesterday.} \]

Now the proposition ‘Hesperus is P’ is necessary but not ICI, but does not itself contain ‘or’ or the like. Still, our account gives the right answer here, since ‘Hesperus is P’ follows from ‘Hesperus is Phosphorus’ which is ICI and true.

Another question which arises in connection with the implication clause is: could there be necessary truths which are implied by two (or more) true ICI propositions jointly, but not by any single true ICI proposition by itself?\(^{16}\) I do not think so, because if a proposition P is implied by A and B jointly, then it will be implied by the conjunction A \& B, and it seems to me that if two propositions A and B are both true and ICI, then their conjunction will be as well.\(^{17}\) If any doubt remains, my account could be modified to run: a proposition is necessary iff it is in the closure (under implication) of the set of true ICI propositions. (I regard this to be equivalent to my main formulation.)

So much for the function of the implication clause. Now for the question of what exactly it means. So far, I have been using familiar language like ‘implies’, ‘logical consequence’ and ‘follows from’ as though it is all about a single relation which holds between propositions and which behaves in the required way, and as though we already understand this language and have a grasp of this relation.

\(^{16}\) Thanks to N.J.J. Smith for raising and discussing this issue with me.

\(^{17}\) Suppose A is true and ICI and B is true and ICI. Then of course A \& B is true. Now suppose A \& B isn’t ICI: that means \(\neg(A \& B)\) appears in some CSD for which A \& B is held true. Now this may not lead to a contradiction by logic alone, but if we assume (i) that all genuine CSDs for which A \& B is held true are also, or (depending on how CSDs are individuated) have identical twins which are, CSDs for which A is held true and B is held true, and (ii) that if \(\neg(A \& B)\) appears in some CSD for which some set of things S is held true, then either \(\neg A\) or \(\neg B\) will appear in some CSD for which the same set of things S is held true, then we get a contradiction. And assumptions (i) and (ii) seem right to me.
That would be the simplest situation for the present account. And the idea that this is pretty much how things are has a good deal of intuitive support. Note that this being the situation doesn’t rule out that there may be unclear or indeterminate cases when it comes to the question of what implies what.

Of course, many philosophers dispute that this is the situation (for starters, see Carnap (1937), Beall & Restall (2000, 2006)). I do not think my account, in any of its essentials, is threatened by this possibility. If the simple monist picture is in order, things are completely straightforward. If not, we still have plenty of options. If there are multiple notions, or corresponding relations, of logical consequence, perhaps one of them will, when plugged into my account, yield results which accord better with our intuitive judgements of necessity de dicto than those yielded when any other candidate is plugged in. If that is the case, our account should be understood as appealing to that one. If that isn’t the case, then we may regard the notion of necessity de dicto as being indeterminate between the different options. We may also, or instead, distinguish different notions of necessity de dicto. Some may even want to take a line akin to Williamson’s (1994) epistemicism about vagueness, and hold that one of the candidates is the right one for my account but that we cannot know which. There are are probably still other options I haven’t thought of. None of these would prevent my account, in all its essentials, from being a good one.

So, it seems like there would only be a problem here if there were actually no legitimate notion or relation of logical consequence at all, or none which behaves in the way required for our account to give the intuitively right answers.\(^\text{18}\) I see no reason to think either of these things is the case. The first seems far-fetched indeed, so I will say no more about it. In support of the second not being the case either, note that the cases which have so far presented themselves as requiring the implication clause involve what seem like very straightforward cases of implication.

Another question raised by the appeal to implication is whether implication should be understood or explained along modal lines, and if so, what that would mean for our account. This is not the place to pursue the analysis of implication, but a few remarks are in order here.

\(^{18}\) One constraint worth mentioning is highlighted by an interesting result obtained by Leslie Tharp, reported in his (1974) and finally demonstrated in his posthumously published (1989). Tharp showed that every necessary truth is a priori materially equivalent to a contingent one. The general recipe for getting a contingent truth given a necessary one is to take any contingent a priori truth and tack it onto the necessary truth with a material biconditional. E.g. \(2 + 2 = 4\), which is necessary, is a priori equivalent to \(2 + 2 = 4 \equiv \text{Julius invented the zip, if anyone did}\), since both are true a priori, and the material biconditional as a whole is of course contingent. Thus, however we understand ‘implies’, it had better not come out true that ‘\(2 + 2 = 4\)’ implies this material biconditional, for then the latter would wrongly be deemed necessary by my account. (Tharp also showed that every truth is a priori materially equivalent to a necessary truth. Take any truth \(p\), and introduce a name ‘A’ with the stipulation that ‘A’ is to (rigidly) designate the truth-value of ‘p’, conceived as a number for simplicity’s sake. Now ‘A = 1’ is a necessary truth, and ‘\(p \equiv A = 1\)’ is a priori. This would also falsify my account of necessity given a wrong understanding of ‘implies’ on which ‘A = 1’ implies ‘p’.) I am not very worried by this, for it seems intuitively right to deny that ‘\(2 + 2 = 4\)’ logically implies, all by itself, the material biconditional, even though the two are a priori materially equivalent. After all, all mathematical truths - or at least all knowable ones - are a priori materially equivalent to each other, but we don’t ordinarily think that (knowable) mathematical truths all logically imply one another.
The simplest sort of modal analysis of implication, according to which A implies B iff it is impossible for A to be true and B false, is widely known to have unwelcome consequences (see Beall & Restall (2014, §1)). Of course, this shouldn’t be taken to show that there is no hope for a more sophisticated modality-involving analysis. For instance, perhaps the stated condition is necessary but not sufficient, and could be supplemented. So, for all that is being said here at least, a modal element could be implicit in the notion of implication.

A further question is whether the analysis of implication should involve modal notions of the subjunctive de dicto sort we are trying to analyze in this thesis. If so, then the account will be circular (or recursive). If not, all that follows is that the account cannot be said to reduce necessity de dicto to non-modal notions - something we have already accepted in our discussion of ICI above.

The following consideration suggests that the modal element in the analysis of implication, if there is to be one, will not be of the subjunctive de dicto sort. Consider for concreteness the simple proposal above: for it to be impossible (in some sense) for A to be true and B to be false is for the negation of the conjunction of A with B’s negation to be necessary (in the same sort of sense). Now, whether a proposition is subjunctively necessary de dicto is, at least orthodoxy and according to a working assumption of this thesis, sometimes an empirical matter. But surely it is never an empirical matter whether a given proposition A implies another proposition B - logical consequence seems like a paradigm case of an a priori matter. Therefore it seems natural to expect that the modal element in the analysis of implication, if such there be, will be of the subjunctive de dicto kind. We would rather expect it to be a notion whose application is an a priori matter (for instance, the notion of indicative necessity discussed briefly in Section 7.2.).

Granted, this is not a knockdown argument - it could be argued that other parts of a good analysis of implication ensure overall apriority despite the application of the modal notion used in the analysis not being an a priori matter in general. For instance, if implication is regarded as necessary truth-preservation in virtue of form. For my part, I find it natural to suppose that the modal element in the analysis of implication, if there is to be one, will not be of the subjunctive de dicto sort. However, my being wrong about that wouldn’t, as far as I can tell, render my account false or unilluminating.

5.5. The Account Reviewed

I have now given a new account of necessity de dicto, shown how it works with a number of examples, developed the key notion of inherent counterfactual invariance, and discussed the role of implication in the account. To review:

Propositions which are true and could not have been otherwise, no matter how things had turned out, we call ‘necessary’.
A proposition is necessary iff it is, or is implied by, a proposition which is both true and inherently counterfactually invariant.\(^{19}\)

A proposition is inherently counterfactually invariant iff its negation does not appear in any (genuine) CSD for which it is held true.

Before considering some objections and replies, I want to outline a fallback position.

5.6. A Fallback Position

This would be especially relevant to a reader who is not happy with the notion of inherent counterfactual invariance being in the account - such a reader may (misguidedly in my view, of course) think it isn’t a legitimate or sufficiently clear notion, or that it does not behave as required. Also, the fallback position is instructive in its own right, even if you accept the main account.

Consider the property some propositions have of being such that it is a priori that they are necessary if true. Expressed semi-formally: the property a proposition \(p\) has iff Apri-(\(p \rightarrow \text{Nec}-p\)). We might refer to these propositions as those with a priori necessary character. The fallback position I want to offer results from substituting, for inherent counterfactual invariance, this notion of a priori necessary character. Thus:

A proposition is necessary iff it is, or is implied by, a proposition which is both true and of a priori necessary character.\(^{20}\)

Of course, this is blatantly circular (or recursive), and may not deserve to be called an account, or an analysis, of necessity de dicto. Still, it seems like a true and instructive proposition.

Note that this notion of a priori necessary character does not line up exactly with the notion of ICI. Consider for example the proposition:

‘First-order logic is undecidable or my hat is on the table’.

This is clearly not ICI; you can hold it true by holding it true that my hat is on the table but false that first-order logic is undecidable, and in that case you would be in a position to produce genuine counterfactual scenarios for which it is held true. However, it seems that it

\(^{19}\) Also recall the alternative touched on in Section 5.4. which, as I argued in f.n. 17 above, delivers the same results: a proposition is necessary iff it is in the closure (under implication) of the set of true ICI propositions.

\(^{20}\) If, as I believe, there is a notion of indicative necessity which can be cashed out non-epistemically - without using any notion of knowledge or a knowing subject - and which lines up, exactly or nearly, with the notion of apriority, we can also speak of the propositions which are such that it is indicatively necessary that they are subjunctively necessary if true, and use that notion instead. This too seems to have some interest in its own right. For simplicity’s sake I will for the rest of this section proceed as if only the apriority-involving version has been put forward, but really two things have been put forward and the following remarks apply to both.
does have a priori necessary character. It seems like it is a priori that this proposition is necessary if true - after all, it seems to be a priori that it is true, because it is a priori that its first disjunct is true, and that the first disjunct is necessary if true seems clearly to be an a priori matter.

So it seems that the two biconditionals - the main account and the present proposal - deliver the same results, but that the implication clause comes into effect in fewer cases here than in the main account.

This is no substitute for the main account on offer. The main account, with its notion of inherent counterfactual invariance, goes deeper and reveals things which this account does not, such as the importance of counterfactual scenario descriptions, the subjunctive-modality-like but a priori tractable notion of ICI (and the underlying notion of a genuine CSD), and the importance of the notion of holding something true for a counterfactual scenario description. Still, it is not without interest, and seems likely to be more readily accepted by more philosophers, as it does not involve any unfamiliar notions.

Let us now return to the main account and consider some objections which may be raised against it.

5.7. Objections and Replies

Objection 1: This account fails to answer the question: <insert some deeply puzzling question about modality which is distinct from the question ‘Under what conditions is a proposition necessary?’>.

Reply: Firstly: So what? Granted, in recent history, some philosophers (e.g. Lewis and Sider, whose accounts we considered in the last two chapters) have made influential attempts to answer some deeply puzzling questions other than ‘Under what conditions is a proposition necessary?’ by beginning with an account of necessity de dicto. But that doesn’t mean there can’t be a good account of necessity de dicto which leaves these other questions unanswered. There is, as far as I know, no reason to think there can’t be. And now that my account is on the table, I think we should conclude that there can be.

One possibility is that by applying my account in a larger story, or by pushing the analysis further (analyzing things that I leave as primitive, such as the appeal to possible CSDs and notion of a genuine CSD), answers to some of these remaining questions could be arrived at. Or if not, perhaps the present account could still help to guide us in searching for answers to some of these questions - by analogy, or even negatively, by leading us to avoid certain avenues of research which would be unfruitful or needlessly circuitous. Another possibility is that an account of necessity de dicto is of no help when it comes to some, or even all, of these other questions. All of these possibilities seem compatible with my account being a good one.

---

21 Some examples: ‘Can modal notions be reduced to non-modal notions, and if so, how?’, ‘What is the metaphysical ground of modal truths?’, ‘What is it for an individual to have a property necessarily?’, ‘How do we come by modal knowledge?’.
Objection 2: Doesn't an account such as this, which explains what makes a proposition necessary rather than contingent in broadly semantic terms, trivialize hard modal questions?

Reply: No, I don't think it does. Take, for instance, what is perhaps the most visible hard, philosophically loaded case here: the question of p-zombies (i.e. beings physically like normal human beings, but lacking consciousness, as discussed by Campbell (1970), Kirk (1974a, 1974b) and most famously Chalmers (1996)). Are p-zombies metaphysically possible? Put into de dicto form, and put directly in terms of necessity, the question is: Is the proposition ‘There are no p-zombies’ necessary? Supposing that we have already answered the question ‘Is “There are no p-zombies” true?’ in the affirmative (and it is obvious that my account doesn’t trivialize that question), the remaining necessity question, on my account, boils down to the question of whether ‘There are no p-zombies’ is ICI, or is implied by a true ICI proposition. Assuming (plausibly I think) that this is not one of the cases where the implication clause comes into effect, the question boils down to: is ‘There are no p-zombies’ ICI?

Now, of course we can construct CSD-like things according to which there are p-zombies. We can say, for instance, ‘Suppose there had been beings physically just like us but without phenomenal consciousness’. Now, the question is: is the description in this sentence a genuine CSD?

My response to this question, at present, is: I don’t know what to say about this. It’s puzzling. And this is just the same as my take on the canonical question ‘Are p-zombies metaphysically possible?’ It may be that putting things in terms of my account leads to some traction here - that would be welcome. It may well be that it doesn’t lead to any traction - that would be perfectly OK too. The point is: my account is not, it seems, guilty of illicitly transforming a difficult, puzzling philosophical question into a too-easily answered one.

Objection 3: You have given a characterization of inherent counterfactual invariance, and this may be intelligible to some degree, but is the notion of a proposition - the notion of the things which are meant to bear ICI - legitimate, and if so, can a proposition really have such a property?

Reply: In a word: wait until the next chapter. To explain a bit further: this sort of objection, as I understand it, comes from the following sort of place. A philosopher has, in response to arguments and puzzles in the philosophy of language (e.g. Kripke’s arguments against descriptivism, Quinean arguments against analyticity and intensional notions, intuitions and considerations which seem to support Millianism and Russelian propositions), been driven to skepticism about propositions, or to a conception of propositions which makes it seem like they couldn’t have such a property as ICI (or which makes it seem unclear whether they could or not). Thus, to respond to this, one has to show that the position the philosopher has been driven to is not mandatory - that there is another way of going which does justice to the arguments and puzzles, on which propositions can legitimately be appealed to, and on which they are clearly the sort of thing which could be ICI. I attempt just this in the next chapter.
5.8. Conclusion

We have developed a new, substantive answer to the question ‘Under what conditions is a proposition necessarily true?’ It shares a structure with Sider’s account, and employs a new notion - that of inherent counterfactual invariance - in place of Sider’s arbitrary ‘being of a certain sort’. It could thus be said that the task of analyzing necessity de dicto involved synthesizing a notion to play a key role in the analysis.

The account is simple and elegant - at least, it is once you’ve mastered the new notion of inherent counterfactual invariance. But it’s just complicated enough, and the new notion required was far enough off the radar, that it’s understandable that the account wasn’t discovered immediately after Kripke isolated its target.

Regarding the notion of inherent counterfactual invariance as a broadly semantic notion, the account satisfies the semantic hunch expressed in Section 1.3. while avoiding the dubiousness of truth by convention, as well as any dubiousness about analyticity.\(^{22}\) Perhaps more to the point, the account avoids the dubiousness of the idea that either of these notions has a role to play in an account of necessity de dicto. Indeed, a highlight of the account is that it finally gives semantic considerations their proper place in accounting for necessity de dicto: they come into the picture in explaining what makes necessary truths necessary, rather than what makes them true. And they come in in a somewhat subtle but compelling way.

\(^{22}\) An avenue for further investigation is what might be called the sub-propositional bases of ICI, and in turn modal, status. ICI propositions are not made that way one at a time, so to speak; it seems plausible, for example, to say that ‘If there is a cat there, then there is an animal there’ and ‘All cats are animals’ owe their ICI status to the very same features of the system of language to which they belong. To capture this, we might speak of a counterfactually invariant connection between the words ‘cat’ and ‘animal’ such that, when we describe counterfactual scenarios, everything in the extension of ‘cat’ in that scenario will also be in the extension of ‘animal’.

We may also distinguish between empirically defeasible and empirically indefeasible connections. Talking this way, we might say that there is:

- An empirically defeasible, counterfactually variable connection between ‘John’ and ‘standing over there’.
- An empirically indefeasible, counterfactually variable connection between ‘Julius’ and ‘invented the zip’.
- An empirically defeasible, counterfactually invariant connection between ‘cat’ and ‘animal’.
- An empirically indefeasible, counterfactually invariant connection between ‘bachelor’ and ‘unmarried’.

Or we may speak of concepts here, and connections between them. There may be room for considerable refinement and elaboration of this sort of classification. We may not want to talk about only connections, but other formations as well. Also, this sort of classification may be useful for modelling technically certain aspects of our language and thought. Perhaps a new kind of model theory for modal predicate logic could be constructed along these lines, for example. (A strategy which might be interesting would be to develop a formal treatment of the notion of inherent counterfactual invariance (treated as an operator or a predicate) along the lines of the above suggestions together with the explanation of ICI from this chapter, and combine this with a formal treatment of the notion of logical implication to give a definition, based on the account given in this thesis, of an operator or predicate representing necessity de dicto.)
way; it is not always the ICI nature of the proposition itself which explains why it is necessary rather than contingent. After all, some necessary propositions aren’t ICI. But of such cases you can say it is the ICI nature of the truths that imply them that makes them necessary rather than contingent. (You can also still maintain that with all necessary propositions, something about the nature of those very propositions explains why they are necessary rather than contingent; if a proposition P is necessary but not ICI, it will be implied by a true proposition Q that is ICI, and you can regard the property of being implied by Q as part of P’s nature.)

The account allows straightforwardly for the necessary a posteriori and the contingent a priori; it is easy to see that a proposition can be inherently counterfactually invariant while being a posteriori, and that a proposition can be counterfactually variable while being a priori.23

The account leaves to one side questions about de re modality and quantification into modal contexts. For a brief discussion of these, see Appendix 2 below.

In the next chapter, we will turn to the task of sketching a philosophical approach to propositions and their meanings which dovetails with this account of necessity de dicto, in particular with the notion of inherent counterfactual invariance. The approach is by no means the only way of going, but it shows by example that there is at least one way of thinking of propositions which makes good sense of the idea that some of them are ICI.

Appendix 1: Arguments for the Coextensiveness of Necessity and Apriority

In this appendix I will indicate briefly how two arguments for the coextensiveness of the subjunctively necessary and the a priori, adumbrated by Kripke, go wrong. First, let us see these arguments as rehearsed by Kripke, and what he says by way of preliminary criticism of them:

I think people have thought that these two things [‘necessary’ and ‘a priori’] must mean the same for these reasons:

First, if something not only happens to be true in the actual world but is also true in all possible worlds, then, of course, just by running through all the possible worlds in our heads, we ought to be able with enough effort to see, if a statement is necessary, that it is necessary, and thus know it a priori. But really this is not so obviously feasible at all.

Second, I guess it’s thought that, conversely, if something is known a priori it must be necessary, because it was known without looking at the world. If it depended on some contingent feature of the actual world, how could you know it without looking? Maybe the actual world is one of the possible worlds in which it would have been

---

23 The account I have offered allows straightforwardly for the existence of the necessary a posteriori and the contingent a priori, in the sense that nothing about the account precludes them or could easily be thought to do so. But it remains that there are arguments which may seem to show that the subjunctively necessary and the a priori coincide. For a brief treatment of these, see Appendix 1 at the end of this chapter.
false. This depends on the thesis that there can't be a way of knowing about the actual world without looking that wouldn't be a way of knowing the same thing about every possible world. (Kripke (1980), p. 38.)

As you can see, at the end of each of the last two paragraphs Kripke has a sentence offering something like a preliminary diagnosis of what went wrong with the preceding argument. (I use the hedge ‘something like’ because calling these remarks of Kripke’s diagnoses, even preliminary ones, might be too strong. Certainly Kripke makes no claim in Naming and Necessity to have fully diffused these arguments, or to have solved the puzzles raised by them. These remarks are just things he said which presumably seemed to him to be relevant and potentially helpful.) I think these suggestions leave something to be desired.

What Kripke says about the argument from necessity to apriority, that ‘running through all the possible worlds in our heads’ is ‘not so obviously feasible at all’ may make it seem like the main problem is that there are too many possible worlds to get through, or that the worlds are too large. I think it is more to the point to observe that we cannot make sure a priori that we are only considering possible worlds. We would need to run through all and only the possible worlds, and the ‘only’ part is the fundamental problem. To illustrate: suppose we know that we don’t know whether a and b, which we have observed empirically in different situations, are one and the same. In that case, we can imagine all sorts of scenarios in which a is distinct from b and various other things are the case, and all sorts of scenarios in which a is identical to b and various other things are the case. But we don’t know which ones correspond to “possible worlds”. That is, we don’t know whether our descriptions of these scenarios are subjunctively possible or not. Put in terms of necessity, we don’t know whether the negations of these descriptions are non-necessary or necessary. And it’s clear why we don’t know that: to know that they are necessary, we would have to know whether they were actually true, and we don’t know that. My account, which explicitly requires putative necessary propositions to be true, or to follow from propositions which are true (which would of course make them true in turn), and which straightforwardly allows that whether they are true or not might be an empirical matter, comports well with this diagnosis.

Now let us consider what Kripke says about the argument in the other direction, from apriority to necessity. The argument suggests that if something is a priori, then it can’t be contingent, so it must be necessary. Why couldn’t it be contingent? Because, the argument runs, if it’s a priori then we can know it without looking at the actual world, and if we haven’t looked at the actual world, then “[m]aybe the actual world is one of the possible worlds in which it would have been false”, as Kripke puts it. Put in the form of a puzzle: how could we know a priori that our world isn’t one of the possible worlds where the contingent proposition is false?

Kripke’s initial suggestion here, that this line of thought ‘depends on the thesis that there can’t be a way of knowing about the actual world without looking that wouldn’t be a way of knowing the same thing about every possible world’ doesn’t really solve the puzzle so much as give rise to it in a more complicated form; how could there be a way of knowing about the actual world without looking that wouldn’t be a way of knowing the same thing about every possible world?

I suggest that the solution lies in considering what propositions might exist, or be available, at what worlds, while keeping in mind that propositions have two aspects of meaning, internal and external. (The distinction between internal and external meaning is developed in Chapter 6.) In worlds where a contingent a priori proposition is false, that proposition cannot
occur. Take the proposition ‘Julius invented the zip, if anyone did’, which is contingent a priori given that we have stipulated that ‘Julius’ is to refer to whoever invented the zip, if anyone did. In a world where someone else invented the zip, there could be a proposition which is alike syntactically and in internal meaning: someone in that world could make the same stipulation about ‘Julius’ and then formulate a contingent a priori proposition running ‘Julius invented the zip, if anyone did’. But this proposition would not be alike in external meaning: ‘Julius’ would refer to someone else, namely the person who invented the zip in that world. There could also be a proposition in a world where someone else invented the zip which is like ours syntactically and in external meaning: there could be a proposition running ‘Julius invented the zip, if anyone did’ where ‘Julius’ refers to Julius, the person who actually (in our world) invented the zip. And if ‘invented the zip, if anyone did’ meant the same as it does in our world, this proposition would of course be false. But this proposition could not have the same internal meaning as ours; the name ‘Julius’ in this proposition couldn’t be governed by the stipulation that it refers to whoever invented the zip, if anyone did, since then it would refer, not to Julius, but whoever actually invented the zip in that world. In short, in the worlds where someone else invented the zip, you can have a proposition which is like ours in internal meaning, and you can have a proposition which is like ours in external meaning, but you can’t have a single proposition which is both.

Appendix 2: De Re Modality and Quantifying In

Are there legitimate constructions like ‘Aristotle was necessarily human’ which are not about propositions, but in which ‘necessarily’ has the same broad sort of subjunctive meaning isolated by Kripke? And if there are such constructions, how should they be understood, and what is their connection with subjunctive necessity de dicto? And finally, is it legitimate, and if so what does it mean, to quantify into modal contexts (as in ‘There is something such that it is necessary that that thing is human’ or ‘There is something which is necessarily human’)? (The coherence of de re modality and quantifying in was famously argued against by Quine in his (1953). He was building on criticisms in his (1943), which according to Føllesdal (2004, p. 24) ‘was, as far as I know, the first objection ever raised against quantification into modal contexts’.) Kripke (1980) defends de re modality from Quine’s skepticism, mainly on intuitive grounds. Other influential defences, from different philosophical points of view, are the counterpart theory of Lewis (1968) mentioned in Chapter 3, and the essentialism of Fine (1994), mentioned in Chapter 2.

Supposing we answer the first question, as to the legitimacy of subjunctive de re modal talk, in the affirmative, the question which then arises as to this talk’s connection with necessity de dicto can be broken down into several more specific questions.

Does ‘This proposition is necessary’ - a proposition attributing necessity de dicto to a proposition - amount to the same as ‘This proposition is necessarily true’ construed as a de re modal attribution? That is, can de dicto modal talk be seen as a special case of de re modal talk, where the res is a proposition and the property at issue is truth? If so, then my account of attributions of necessity de dicto is also an account of this class of de re modal attributions.

We may also ask whether de re modal talk more generally, or a lot of it at least, can be explained in terms of de dicto. One strategy for doing that is offered by Sider (2011, pp. 287 - 288): use the notion of necessity de dicto to construct a space of ersatz possible worlds, and then give a counterpart-theoretic account of de re modal discourse à la Lewis (1968). (Actually, Sider offers two strategies for dealing with de re modality in his (2011), but the
other one is specific to his account of necessity de dicto, which we discussed in Chapter 4. It consists in adding further “modal axioms”, and for quantifying in, relativizing modal axiomhood to variable assignments. See Sider (2011, p. 287.) But this might seem artificial. Or, while it may be viable as far as it goes, maybe it doesn’t give us the whole story. There seem to be more direct ways of linking de re modal attributions, and cases of quantifying into modal contexts, to de dicto claims, but in such a way that in some cases the strategy breaks down.

For instance, you might think that ‘Aristotle is necessarily human’ comes to the same thing as “Aristotle is human” is necessary’. One question is whether this is correct at all - maybe the two could come apart. This raises in turn a perplexing question: could two propositions which name the same object and attribute the same property to it ever differ in ICI, and in turn modal, status? I am inclined to answer in the negative, but the very question makes me uneasy, in that I don’t see any clear way of deciding it, although it arises naturally. Thus it seems like there might be an opportunity here to get clearer about something. Another question which arises, assuming that ‘Aristotle is necessarily human’ and “Aristotle is human” is necessary’ and the like cannot come apart, is what their relationship is exactly. Does the former just mean the latter? This may seem objectionable - it may seem to “make everything a matter of language” in some unwholesome sense.

Another issue is that, even if we can in some sense explicate ‘Aristotle is necessarily human’ by means of “Aristotle is human” is necessary’, it is hard to see how this could yield a general account de re modality; if we can say ‘Aristotle is necessarily human’ then of course we have a name for Aristotle and we have the proposition ‘Aristotle is human’. But what about objects that we don’t have names for? Can we explicate ‘There is an unnamed thing which is necessarily human’ in de dicto terms? Perhaps we could say that this comes to the same thing as 'There is something such that, if you named it and attributed humanity to it, you would get a necessary proposition’. But then what about ‘There is something which is contingently unnamed”? A parallel treatment would seem to yield a wrong answer: if you named an unnamed object and predicated unnameness of it, you would get something, not contingently true, but false, since by naming it you would have stopped it from being unnamed.

Another thing which might also bother us here is the question of the point, or meaning, of all this. We might worry, with Divers (2007), about the utility of (broad, unrestricted) de re subjunctive modal talk. The difficulties we find here have sometimes reminded me of Wittgenstein’s remark that

[p]hilosophers often behave like little children who scribble some marks on a piece of paper at random and then ask the grown-up "What’s that?” — It happened like this: the grown-up had drawn pictures for the child several times and said "this is a man," "this is a house," etc. And then the child makes some marks too and asks: what’s this then? (Wittgenstein (1980, p. 17e).)

But is that a fair comparison in this case? Perhaps that thought can seem appealing here just because the questions are difficult and tiring. Furthermore, even if that is a fair comparison, does that absolve us from wrestling with these perplexing questions, now that they have arisen? Perhaps it offers, rather, a clue about how best to approach them.

**Chapter 5 References**


6. Propositions and Meaning

My task here is to sketch an account of propositions which is independently attractive, and which makes sense of the idea that propositions can have a property like that of inherent counterfactual invariance, which I appealed to in my account of necessity de dicto in the previous chapter. In the course of doing this I will develop some ideas about linguistic meaning more generally.

The plan for this chapter is as follows. In Section 6.1., I will explain in a preliminary way what I mean by ‘proposition’ and situate my project here with respect to other projects in philosophy which give a central place to the word ‘proposition’. In Section 6.2., I will be a bit more specific about what I mean by ‘proposition’, occasioning some methodological remarks as well. In Section 6.3., I will separate two aspects of linguistic meaning, internal and external. In this I am guided by Putnam-inspired, Twin-Earth type considerations. In Section 6.4., I will further develop the notion of internal meaning, using the notion of a role in a language system. In this I am guided by Wittgenstein (his middle period especially). By way of further explanation, I will compare and contrast my notion of internal meaning with certain ideas associated with the phrases ‘conceptual role semantics’ and ‘narrow content’. In Section 6.5., I will say a bit more about external meaning. In Section 6.6., I will apply my account of internal meaning to names, and show how it enables us to solve some puzzles in the philosophy of language.

In Section 6.7., I will sketch a further aspect of my account of internal meaning, having to do with the granularity or individuation of internal meanings (and in turn, propositions). I will suggest that internal meanings can, in a certain sense, be carved up at different granularities - that there is a special kind of flexibility inherent in our talk of meaning, which should be reflected in our philosophical accounts of language. This, I will argue, will help us to assuage general worries about meanings and their “criteria of identity” (famously pressed by Quine), and can also be applied to a number of philosophical puzzles and issues.

Finally, in Section 6.8., I will tie the account of propositions just sketched back to the account of necessity de dicto given in the previous chapter, showing how the former supports the latter.

6.1. Preliminary Explanation

By ‘proposition’ I mean, roughly, a declarative sentence in meaningful use.\(^1\) Thus I consider propositions to be linguistic items which have semantic properties. They can be true, false, be about particular things, and - the idea is - be subjunctively necessary.

\(^1\) Thus my usage differs from another kind of usage, perhaps more common within present day analytic philosophy, according to which sentences express propositions and sentences in different languages can express the same proposition. (I of course can say that two propositions mean the same.) While my usage is perhaps not the most common within analytic philosophy nowadays, I think it has both ample historical precedent and a good deal else to recommend it. However, I am not arguing for or relying on that here - rather, I am simply exercising the author’s freedom to use technical words in any reasonable way as long as the usage is made clear.
It is a commonplace that ‘sentence’ is type/token ambiguous; using it the type way, we say that this:

Snow is white.

And this:

Snow is white.

Are instances of a sentence - the same sentence. Using ‘sentence’ the token way, we say that each of the things I drew your attention to is a sentence. I think it is useful to use ‘proposition’ ambiguously in this way also (disambiguating when necessary, of course).

Using ‘proposition’ the token way, we can say quite literally that propositions come out of people’s mouths, get written in ink, are heard and seen, etc.

Now for some points, mostly negative, about what sort of account of propositions I am going to sketch here.

No stance taken on the “primary truth-bearer” issue. Some authors have discussed, and taken positions on, an issue about what the primary bearers of properties like truth and falsity! Some, for instance, may allow that the things I call propositions can be said to possess properties like truth, falsity, necessity, contingency, and the like, but maintain that they possess these properties derivatively. Others may maintain that the things I call propositions are the primary bearers, and that other things, for instance beliefs, possess these properties derivatively. I am not taking a stance on any such issue. I am not even taking a stance on whether it is an intelligible issue, or whether it is an important issue, or any other question having to do with this issue. All I require is to be able to literally and correctly apply predicates like ‘is true’, ‘is false’, ‘is necessary’, ‘is ICI’, and ‘implies’ to the things I call propositions.

My use of ‘proposition’ not functional or role-based. Some authors begin with the roles that propositions are to play, and then propose theories about what sort of thing propositions should be taken to be. This is particularly clear when the word ‘proposition’ is introduced by means of the stipulation that it is to apply to the things, if there are any, which play, or best play, the nominated role or roles. I have no objection to such a practise, but it is not what I am doing here. At least, this doesn’t seem to be a fitting way of thinking about what I am doing here. Rather, I am, so to speak, specifying directly what sort of thing I mean by ‘proposition’. Having specified what sort of thing I mean by ‘proposition’, I claim - not

---

2 For a thorough, book-length study of this issue see Rojszczak (2005). For a more recent treatment of the issue see Hanks (2014). Mosteller (2014, p. 110) shows how Russell changed his views back and forth on this issue. Haack (1978, p. 73) argues that this debate has been ‘neither very conclusive nor very fruitful’.

3 Two clear examples of this are McGrath (2014, Intro.) and Smith (2016, pp. 84 - 85).
stipulate - that propositions can bear the property of inherent counterfactual invariance and, in turn, the property of subjunctive necessity.4

No theoretical stance on belief-reports or "propositional attitudes". Philosophers sometimes want the things they call 'propositions' to play a role in explaining the meaning of belief-reports like 'Smith believes that Jones is in town' - and perhaps in explaining, in turn, the things we explain with such reports, such as the behaviour of people. For instance, some defend the idea that belief-reports, at least in some uses, ascribe a relation, that of believing, which holds between agents and (the things they call) propositions. Belief is sometimes called a 'propositional attitude', and sometimes other such attitudes are proposed, such as desire and fear (and there is debate about whether this is appropriate). I am not here taking a stance on any of these topics.

However things are, I think it would be hard to deny that connections exist between what I call propositions and the topics of belief-reports, "propositional attitudes" and the like. Indeed, some of the interest of what I have to say may depend on that being the case. But I am not trying to elucidate those connections here.

I will allow myself to talk of accepting propositions, holding them true, and even believing them, but none of this is to be taken to reflect any serious theoretical commitment in this "propositional attitudes" part of philosophy. On the other hand, talk of holding propositions true is a key part of the account of necessity in Chapter 5. Now, it could in principle turn out that it is a bad idea, at least in philosophy, to talk of 'believing' or 'holding true' the things I call propositions. But I think it's pretty clear that on any reasonable view there will be something near to hand and acceptable to say which does the job I wanted to do by talking of believing, or holding true, propositions.

6.2. Being More Specific

In the initial gloss above, I said that by 'proposition' I mean, roughly, a declarative sentence in meaningful use. But this is not quite right, since declarative sentences can be given non-propositional meaningful uses. For example, they may be used performatively (e.g. 'I pronounce you man and wife'). Another problem with this gloss might be that signs which we are not inclined to call 'declarative sentences' may be used propositionally (more on what that comes to below).

One way of improving on the initial gloss might be to swap in something more tailor-made for the notion of a declarative sentence. For example, Wittgenstein in the Tractatus (3.12) wrote:

---

4 This raises the question of the dialectical relationship between my approach and these role-based approaches. First and foremost, I think they are just somewhat different sorts of projects. But of course, they are far from unrelated. There may be some important philosophical differences in the background here. Also, perhaps the two sorts of projects could complement each other. In any case, this will not be the focus of the present chapter. I want to say: surely there exist the things I want to call 'propositions', and surely they are of philosophical interest. Let me try to sketch an account of them.
The sign through which we express the thought I call the propositional sign. And the proposition is the propositional sign in its projective relation to the world.

Or, we could supplement the ‘in meaningful use’ part of the initial gloss, and say that a proposition is any sign whatever in this kind of use. Here I will pursue this latter course, since it will be more straightforward.

So, we want to supplement ‘in meaningful use’ so that it applies in all and only the cases we want to call cases of propositions. ‘Propositional’ is a good word for this - we can say that a proposition is a sign ‘in meaningful propositional use’. We can now drop ‘meaningful’ too, since all propositional uses will be meaningful. But the question now is: what does ‘propositional’ mean here?

We can say that propositional uses of signs are characterized by the fact that, given such a use, we can ask: is that true or false?

We can say of signs used propositionally that they express claims about how things are. And of course we can give many clear examples of propositional uses of signs, such as this very sentence.

Can we say much more than that? For my part, I am inclined to think not. For one thing, I am sympathetic to the idea that the concept I am trying to convey here is open-textured, i.e. that it is not everywhere definite. Furthermore, I think this is essential to this sort of concept, in the sense that, if you “made it definite”, you would in fact be replacing it with something else. To be sure, there is a place in philosophy for constructing definite concepts “corresponding” to indefinite ones, but I think there is also a place for working with and theorizing about indefinite concepts themselves, perhaps refining and elaborating them, but retaining much of their original character, indefiniteness included. We should allow both methods, and here I am mostly using the latter. For another thing, whether definite or not, I am doubtful that there is any further spelling out of ‘propositional’ to be done here. (This is a separate point, because you can sometimes give insightful analyses or definitions of notions that are indefinite. In fact, I hope my account in the previous chapter is an example of that.)

Having said that, it is worth noting that the attitude just suggested can be detached from much or all of the rest of what I am saying here. You might accept that there are signs in propositional uses, and that these are interesting objects of philosophical study, and you might agree with much of what I am going to say below about the meanings of what I call ‘propositions’, but differ from me in thinking that, in the present philosophical context, a precise definition of ‘proposition’, or a characterization more substantial than any I have just proposed, can and should be given. In that case you might regard what I am saying here as lacking in that respect but still find some of it helpful.

6.3. Internal and External Meaning

5 Here I am influenced by Wittgenstein. See for example his (1953), §§59, §§65, §§68, §§71 especially, and §§92 - 137 (§135 especially).
The *locus classicus* for the distinction I am going to make here is Putnam’s famous Twin Earth thought experiment, as given in his (1973) and (1975). At least, the Twin Earth thought experiment is the first port of call for the distinction I want to make here; it may be that the distinction itself cannot be attributed to Putnam, since it seems he was mainly interested in another lesson which can be drawn from the thought experiment.6

Here is the setting up of the thought experiment as originally given by Putnam:

For the purpose of the following science-fiction examples, we shall suppose that somewhere there is a planet we shall call Twin Earth. Twin Earth is very much like Earth: in fact, people on Twin Earth even speak *English*. In fact, apart from the differences we shall specify in our science-fiction examples, the reader may suppose that Twin Earth is *exactly* like Earth. He may even suppose that he has a *Doppelganger* - an identical copy - on Twin Earth, if he wishes, although my stories will not depend on this.

Although some of the people on Twin Earth (say, those who call themselves “Americans” and those who call themselves “Canadians” and those who call themselves “Englishmen,” etc.) speak English, there are, not surprisingly, a few tiny differences between the dialects of English spoken on Twin Earth and standard English.

One of the peculiarities of Twin Earth is that the liquid called “water” is not H2O but a different liquid whose chemical formula is very long and complicated. I shall abbreviate this chemical formula simply as XYZ. I shall suppose that XYZ is indistinguishable from water at normal temperatures and pressures. Also, I shall suppose that the oceans and lakes and seas of Twin Earth contain XYZ and not water, that it rains XYZ on Twin Earth and not water, etc. (Putnam (1973), pp. 700 - 701.)

Of course, since the advent of modern science, people have become aware that water is H2O, and this could be expected to make things here on Earth different in further ways from how things are on Twin Earth. To get around this, Putnam has us consider things as they were for ordinary people in 1750, and has us imagine this having been paralleled on Twin Earth (pp. 701 - 702). He has us imagine a particular ordinary speaker from 1750, Oscar, along with his Twin Earth counterpart, Twin Oscar.7

That is the setup. I will now take over and use this to make the distinction I want to make between internal and external meaning. I will then consider some worries about the case. I think these worries can be adequately responded to, and will say how. Also, I will offer a different Twin-Earth-like case, involving individuals rather than stuffs, for which the same

---

6 Putnam’s lesson was, in brief, that mental states do not determine extension. I am using the Twin Earth thought experiment to help separate out an aspect of meaning, internal meaning, which also does not determine extension (in general), but which is not to be thought of in terms of mental states.

7 Putnam used ‘Oscar,’ and ‘Oscar,’ to name them, but it has since become customary to call these characters ‘Oscar’ and ‘Twin Oscar’.
worries don’t arise. I will also suggest that the case of indexicals further illustrates the distinction between internal and external meaning.

The meaning of ‘water’ as Oscar uses it seems to differ in an important respect from the meaning of ‘water’ as Twin Oscar uses it; the word as Oscar uses it applies to a different stuff from what it applies to as used by Twin Oscar. Similarly, the meaning of a proposition uttered by Oscar containing the word ‘water’ will differ in an important respect from the meaning of the corresponding proposition uttered by Twin Oscar.

But obviously, there is something in common between Oscar’s uses of ‘water’ and propositions containing it, and Twin Oscar’s corresponding uses. And it is not just that they are a bit similar but also a bit different - rather, it seems that, in virtue of how the case has been set up, they are in an important respect exactly the same, but also different in another important respect. The case reveals two different aspects, or factors, in what we pre-theoretically call the ‘meaning’ of a word or a proposition. I call the factor which is the same for Oscar and Twin Oscar internal meaning, and I call the factor which differs external meaning.

I will now consider two worries about this case of Oscar and Twin Oscar and their use of ‘water’.

The first worry is that maybe ‘water’ as Oscar uses it, given that he is a typical speaker from 1750, actually applies to the same stuff (or stuffs) as ‘water’ as Twin Oscar uses it, so that both $\text{H}_2\text{O}$ and XYZ fall under both uses. In philosophical jargon, maybe the word should be understood phenomenologically, or functionally, rather than as a natural kind term - i.e. maybe it applies to anything that seems watery, or plays a certain role (such as sustaining life, etc.). Or maybe some combination of those sorts of things.

My response to this is as follows. Maybe that’s right, or closer to the truth, when it comes to ‘water’ as typically meant in 1750. That is at least partly an empirical matter, and it need not concern us here. It is enough that we can see a possible meaning for ‘water’ which behaves in the imagined way.

The second worry is that XYZ - a liquid different from $\text{H}_2\text{O}$ whose chemical formula is long and complicated - couldn’t possibly behave in the same way as $\text{H}_2\text{O}$. There may be good scientific reasons for thinking this.

My response to this is that, while this may be physically impossible, or impossible in any other number of senses, there is a more unrestricted sense in which it is possible, and that’s all that matters for the thought experiment.  

---

8 A complementary point: for a Twin Earth case - a case where you have two different expressions occurring in different places which have the same internal meaning but differ in external meaning - not everything other than the salient difference needs to be the same. There just needs to be enough of the right kinds of similarities that we may count the expressions as having the same internal meaning. Differences which don’t lead to differences in internal meaning don’t matter.
Happily, to make the distinction I want to make, I need not rely entirely on these responses being adequate. We can imagine a different sort of case which does the job, but for which these worries do not arise.

Let us remove the stuff about XYZ from Putnam’s setup, and just imagine another planet exactly like Earth. We still have Oscar and Twin Oscar, but the watery stuff on this Twin Earth is exactly the same as here. Suppose Oscar comes home one day, and someone he lives with says ‘Oscar’s home!’, announcing his arrival to the rest of the household. And suppose that exactly the same thing happens with Twin Oscar (remember, ‘Twin Oscar’ is our name for Twin Oscar - we may suppose that in his own life he is called simply ‘Oscar’). Now, the two propositions - ‘Oscar’s home!’ here and on Twin Earth - clearly differ in meaning in an important respect. One is about Oscar, and one is about Twin Oscar. The first would have been false if Oscar hadn’t come home then, but that isn’t true of the second, and the second would have been false if Twin Oscar hadn’t come home then, but that isn’t true of the first. But again, it seems like in another respect, these two propositions are exactly alike in meaning. And not just because they both say, of someone, that they are home. There’s more to the similarity than that. The word ‘Oscar’ for both speakers plays an exactly similar role in their speaking and thinking. It is elicited by exactly the same stimuli. Its pattern of use is the same in both cases. And again, I call the factor which is the same internal meaning, and the factor which differs external meaning.

Clearly, the two worries that arose in the previous case do not come up here. It cannot reasonably be maintained that in these cases ‘Oscar’ is used as a kind of general term applying to anyone sufficiently Oscar-like. That’s obviously not how names work. The second worry doesn’t arise, because no different chemical is being posited.

However, this case has its own downsides, not affecting the ‘water’ case. (It is good to have both kinds of case to offer, since they have different pros and cons.) One downside is that it may set off alarm-bells for philosophers sympathetic to, or even just aware of, the Millian view of proper names according to which they have no meaning, or no meaning except for their referent. That is not an internal problem for my account here, since in Section 6.6. I will be proposing a contrary view of names using the very notion of internal meaning I am beginning to talk about here, and I will try to explain what is going wrong when the Millian view gets adopted. Furthermore, what I have said about there being a factor of meaning in common seems, from an intuitive point of view, hard to deny. Still, this is a downside. Another worry which might arise about this case, even for someone not sympathetic to the Millian view of names, concerns the individuation of internal meanings; it may be clear enough in this case, where everything is the same except for which planet and which individuals are involved, that there is a meaning-factor in common, but what about cases involving further differences? What does it take for a proposition to have this internal meaning? You might have the hunch that there isn’t going to be a single, definitive answer here, and this may be worrying. (The fact that the propositions in question here are about particular individuals in a very specific sort of situation rather than common stuffs seems to make this worry more pressing than in the ‘water’ case.) To this worry, my reply is: wait for Section 6.7., but for now I will just say that, in my view, it’s flexible - the hunch that there isn’t going to be a single, definitive answer is correct, but this isn’t a bad thing.
Finally, another sort of case which helps to illustrate the distinction between internal and external meaning is that of indexicals: you and I may use ‘you’ and ‘I’ with the same internal meaning, but different external meanings. ‘Today’ has the same internal meaning from day to day, but its external meaning changes from day to day.

So, that is the distinction between internal and external meaning, at least in a preliminary form. The distinction as I have made it here, based as it is on an intuitive notion of meaning, may seem a bit fragile. It will be strengthened in the following sections when I sketch a philosophical account of internal meaning and say a bit more about external meaning.

**6.4. Internal Meanings as Roles in Language Systems**

So far we have got a preliminary, intuitive characterization of internal meaning as distinct from external meaning. But it is based on intuitive, unrefined talk of meaning, and our ordinary linguistic practice with ‘meaning’ and similar words is very multi-faceted. Here I will try to refine the notion of internal meaning, abstracting away from some aspects of ordinary usage of ‘meaning’ and accentuating what is important for our present philosophical purposes (namely, motivating an account of propositions by using it to solve some puzzles in the philosophy of language, and then using that account to underpin an account of subjunctive necessity *de dicto*).

Thus we are arriving at an account of internal meaning in two steps: starting with a loose, intuitive idea of meaning, we made the division between internal and external meaning. Now, having put external meaning to one side (to return to later), we will try to refine the idea of internal meaning by abstracting away from certain aspects of the ordinary use of ‘meaning’, aspects which remain after we put external meaning to one side.

This second step can be broken down into two parts: first, a rough preliminary characterization of what we are abstracting away from, and secondly, a more elaborated positive characterisation of the aspect we are interested in. (I want to do the abstracting in large part by specifying what internal meaning *is* to be thought of as, and letting the rest fall away, rather than describing in detail what is to be ignored.)

So now for the first part: what, roughly, are we abstracting away from here? We want to abstract away from, so to speak, incidental mentalistic accompaniments of language. (Incidental for us, that is. In other connections these might be just what we are interested in.) The word ‘meaning’ is sometimes used to talk about associations and feelings associated with bits of language (and other things). One particularly clear example of this aspect of the use of ‘meaning’ is the way we say that if you repeat a word over and over it ‘loses its meaning’.

Wittgenstein expressed versions of this type of attitude in various places. For instance, in a lecture in the early 1930’s (Wittgenstein (1979), §2):
The meaning of a word is to be defined by the rules for its use, not by the feeling that attaches to the words.

And in Part I of *Philosophical Grammar*, ‘The Proposition and its Sense’:

44. What interests *us* in the sign, the meaning which matters for us is what is embodied in the grammar of the sign.

We ask “How do you use the word, what do you do with it” - that will tell us how you understand it.

Grammar is the account books of language. They must show the actual transactions of language, everything that is not a matter of accompanying sensations.

In a certain sense one might say that we are not concerned with nuances.

These quotes may help you to understand the account being sketched here, because they express, if not the same attitude, then an attitude which is similar with respect to the sort of thing being abstracted away from.

There are important differences too. For instance, in the first quote, the mention of ‘rules’ does not accurately represent anything about my account. I don’t want to say that internal meanings can in general be captured by rules. But the negative part of this quote is a fitting expression of the kind of refinement of the ordinary use of ‘meaning’ I want to make here.

The second quote, considered as a description of the account I want to uphold, may be criticized for making the attitude seem more austere and fixedly coarse-grained than it is. There *is* room for inclusion, under the heading of internal meaning, of what might be called ‘nuances’, particularly given the ideas about granularity to be explained in Section 6.7. (But as Wittgenstein said, it is in a ‘certain sense’ - not every sense - that one might say that we are not concerned with nuances.)

So much for our preliminary characterization of what aspects of the ordinary use of ‘meaning’ we will be abstracting away from. We will now proceed to a positive specification of what it is we are interested in.

The leading idea of this section, expressed in a formula, is this: the internal meaning of an expression is to be regarded as the role it plays in the system of language to which it belongs.10 11

---

9 You might ask ‘How then can they be captured’? The questions of how, and how fully, internal meanings in general may be described are interesting questions, but I cannot pursue them here. (Recall that this is only a sketch of an account of propositions and meaning - I am trying to show the philosophical usefulness of a notion of internal meaning, and offering a suggestion about how we should think of internal meanings. I am not pretending to have a complete account of them which answers all important questions about them.)

10 Compare:

A name has meaning, a proposition has sense in the calculus to which it belongs. (...)
One problem you might have with this formulation is that it mentions this thing, ‘the system of language’ to which an expression belongs. I will take it as beyond dispute that, in some sense, language is systematic - that the notion of a system can be applied to it.\(^1\) But what is the system to which an expression belongs? Is it something which a speaker can carry around with them? Is it something which can differ from speaker to speaker? Or is it something more universal and public?

One thing to keep in mind here is that expressions, regarded as mere shapes or sounds or whatever, may have more than one meaning, and may belong to more than one system (however we think of these systems). In the slogan above about ‘the internal meaning of an expression’, we are really talking about meaningful occurrences of expressions, or types thereof. But still, you might think there are multiple candidates things here which could deserve to be called ‘the system of language’ to which the occurrence or occurrence-type belongs.

There are a couple of ways to respond to this worry. Firstly, having more than one candidate is surely better than not having any. Perhaps further investigation will narrow things down. Secondly, perhaps we should think of the notion of internal meaning as being flexible on this point, or coming in different versions corresponding to different ways of thinking about ‘the system’. When dealing with some expressions, we may be able to think of ‘the system’ as something understood and shared by all speakers conversant with the language being used. At the other extreme, when dealing with other expressions, such as proper names or very specialized terms, we may, for some purposes at least, prefer to think of ‘the system’ as something belonging to a small group of people, or a pair of people in a conversation, or even just one person.

Further light can be shed on the notion of ‘the system’ and the work it is doing here by considering the following question: why not simply say that the internal meaning of an

\[\text{I might say: the only thing that is of interest to me is the content of a proposition and the content of a proposition is something internal to it. A proposition has its content as part of a calculus.}\]

The meaning is the role of the word in the calculus. (Wittgenstein (1974), §27, p. 63.)

\(^1\) This is, by the way, perfectly consistent with the point made above in Section 6.1. about my use of ‘proposition’ not being functional or role-based. The point there was that I am specifying directly what sort of things I take propositions to be, rather than outlining a theoretical role and stipulating that ‘proposition’ is to be used for whatever best satisfies that role. The suggestion here is that the internal meanings of bits of language, including the things I call propositions, be thought of as roles in language systems.

\(^2\) Johnson (2004) argues that language is not systematic, given an understanding of systematicity based on the idea of being able to substitute expressions of the same category while preserving grammaticality. The data he presents are interesting and instructive, but needless to say, they do not establish that language is not systematic in the ordinary sense. After all, he proposes counterexamples to systematicity understood in a particular way and expects us to agree, even if we haven’t encountered the counterexamples before. It can be said that here he is relying on the systematicity of language in order to prove that language does not have what he calls ‘systematicity’.  

93
expression is the role it plays simpliciter? (Or the role it plays in our lives, or in the world?)
To see why not, consider the following example of Wittgenstein’s:

498. When I say that the orders "Bring me sugar" and "Bring me milk" make sense, but not the combination "Milk me sugar", that does not mean that the utterance of this combination of words has no effect. And if its effect is that the other person stares at me and gapes, I don't on that account call it the order to stare and gape, even if that was precisely the effect that I wanted to produce. (Wittgenstein (1953), §496. Also appears in Wittgenstein (1974), end of §136, p. 188.)

What makes this example so instructive is that it shows us a use of words which is meaningless but not pointless. It has a function, or a role in life, but the string of words in question has no internal meaning because it is not used significantly, or as part of a system of language. Its use is, so to speak, isolated. I hasten to add that the distinction here, between being meaningful and being meaningless yet not pointless, is probably vague, but it is no less real for that.

Another potential worry about the idea of this section concerns propositions (or other units of language) which are likely to be used only once, or some very small number of times. It is one thing to think of a single word (’of’, ’cat’, ’description’), or a phrase (’the point is’, ’no pun intended’, ’plumb the depths’), or a short and often-repeated proposition (’The sky is blue’, ’3 x 4 = 12’) as playing a role in a system, but what about a proposition like the one at the beginning of this paragraph (’Another potential worry […]’)? It can seem a bit odd to speak of such a proposition as playing a role in a system. I accept that there is a slight oddity here, but I don’t think it seriously affects the account. We can maintain that the proposition at the start of this paragraph plays a role in a language system, even though this role only ever opens up and gets played in the context of this part of this thesis, as long as we are ready to allow that this is an extended, theoretical use of the notion of ’playing a role’ which may accordingly seem a bit odd.

I have now said something about internal meaning - that an expression’s internal meaning may be regarded as the role it plays in the system of language to which it belongs - and given an indication of what work is being done by the reference to the system of language. (Perhaps more should ultimately be said here, but recall that the task of this chapter is to sketch an attractive account of propositions which is capable of underpinning our analysis of necessity de dicto.)

I will now try to clarify my notion of internal meaning further by saying more about what it is not. First I will get some easy clarifications out of the way, distinguishing my approach from inferentialism, rule-based approaches and the like. I will then suggest a rough “insensitivity” heuristic for thinking about what internal meaning doesn’t involve. Finally, I will examine a couple of notions visible in the literature under the headings ‘conceptual role semantics’ and ‘narrow content’ which need to be quite carefully distinguished from my notion of internal meaning.
To begin with: it is no part of my account that the roles I speak of can properly be thought of
as being constituted just by inferential relationships between bits of language. So this is not
a form of inferentialism. Also, it is no part of my account that the roles I speak of can be
captured in explicit rules of usage. (In general, I am wary of what I would call overly
reductive accounts of the role an expression plays in a language system, or of the language
system itself.)

Language systems are essentially open systems - i.e. it is essential that they have
connections to things which aren't symbols, things which are not part of the apparatus of
language. That said, we do not want to count having $H_2O$ in its extension and having Oscar
as a referent as part of the internal meanings of 'water' and 'Oscar' as used in the thought
experiments of Section 6.3.

We are steering between two extremes in our conception of (a role in) a system of language:
thinking of it as, so to speak, a disconnected structure, vs. thinking of it as external meaning
involving.

Guided by Twin Earth considerations, we can perhaps give a good idea of the extent of
internal meaning, so to speak, with the following "insensitivity" heuristic: internal meaning is
insensitive, at least for a large class of cases, to the numerical identities of the objects being
talked about, and to the underlying natures of the objects or stuffs being talked about.

Before moving on to further explain external meaning, let us briefly compare and contrast
this notion of internal meaning with two proposals belonging to a cluster of theories in the
philosophy of language and mind. Theories in this cluster are versions of what is called
'conceptual role semantics'. (For a recent bibliography on this see Båve (2015).) Seeing how
the present notion of internal meaning differs from these proposals, which are strikingly
similar in some respects but also importantly different, will help us get a more precise idea of
the former. We can find them both in a paper by Ned Block (1998). We will call the first one
'Simple CRS', and the second one (following Block) 'Two-Factor CRS'.

First, Simple CRS. Here is Block’s characterization:

According to Conceptual Role Semantics ("CRS"), the meaning of a representation is
the role of that representation in the cognitive life of the agent, e.g. in perception,
thought and decision-making. (Block (1998), Abstract.)

What is being said here about 'the meaning of a representation' is strikingly similar to what I
say about the internal meaning of an expression. One difference is that the 'role' being
referred to is characterized more psychologically and less linguistically than in my account of
internal meaning. This difference is minor, however, compared to the one which emerges
when Block writes that

CRS supplements external use by including the role of a symbol inside a computer or
a brain. (Block (1998), Abstract.)
So, the notion of 'conceptual role' Block is working with here differs from mine not just in being cashed out in more psychological terms, but also in being exclusively about what is internal to an agent. (For me the 'internal' in 'internal meaning' means 'internal to language', not 'internal to a person'.)

Another difference is that CRS is here put forward as an account of 'the meaning' of a representation - not just one aspect of meaning, as it is in my account. But Block goes on to consider a challenge to Simple CRS - Putnam's Twin Earth thought experiment, which we used above to distinguish internal from external meaning - and describes Two-Factor CRS, a view designed to overcome it:

Putnam (1975) raised what might seem to be a powerful objection to any CRS. He pointed out that many "natural kind concepts," such as water and gold, depend in part for their meaning upon something other than the role of a representation in a person's head, namely upon what happens to be in their external environment. (...) Some proponents of CRS have responded by favoring a "two-factor" version of CRS. On this view, meaning consists of an internal, "narrow" aspect of meaning - which is handled by functional roles that are within the body - and an external referential/truth-theoretic aspect of meaning, which might handled by some of the other metaphysical theories of meaning (e.g. a causal one) that we mentioned earlier. According to the external factor, 'Superman flies' and 'Clark Kent flies' are semantically the same since Superman = Clark Kent; the internal factor is what distinguishes them. But the internal factor counts 'Water is more greenish than bluish' as semantically the same in my mouth as in the mouth of my twin on Twin Earth. In this case, it is the external factor that distinguishes them. (Block (1998), §3.)

This version of CRS shares a further similarity with my notion of internal meaning: the desire to separate two factors of meaning, such that one factor - the internal factor - is the same for expressions here and their Twin Earth counterparts. But the bit about 'within the body' is no part of my account. The internal factor Block has in mind here, or a mental-state based analogue of it, has been called 'narrow content' and discussed in the philosophy of mind. As the introduction to Brown's (2011) encyclopedia article on 'Narrow Mental Content' says:

A narrow content of a particular belief is a content of that belief that is completely determined by the individual's intrinsic properties. An intrinsic property of an individual is a property that does not depend at all on the individual's environment.

Again, my notion of internal meaning, unlike 'narrow content' and Block's notion of conceptual role, is not about what is intrinsic to an agent. It goes beyond that, and what goes on inside the brain is not emphasized. Indeed, it is tempting to say that that is in a certain sense irrelevant from our point of view. In this, I think my notion of internal meaning is closer to Wittgenstein than the notion of conceptual role used in Two-Factor CRS.

6.5. External Meaning Further Explained
Based on the intuitive introduction of the distinction between internal and external meaning, given in Section 6.3., it would be quite natural to think of external meaning as being something like extension. After all, that is the most obvious thing which differs in the cases used: in the first case ‘water’ has a different extension here from what it has on Twin Earth, and similarly, in the second case ‘Oscar’ refers to one man here and a different man on Twin Earth.

In most cases, you won’t go far wrong thinking of it this way. But it’s not quite right. This is shown by cases where a term happens to lack extension or reference. For instance, Oscar might be out walking on a misty night and fancy he has seen a shrouded figure, when in fact no one was there. He might introduce the name ‘Enigma’, intending for it to refer to the figure he believes he has seen. Suppose that exactly the same thing happens with Twin Oscar. Now, we want to distinguish semantically between ‘Enigma’ here and on Twin Earth, and between propositions here containing ‘Enigma’ and the corresponding propositions on Twin Earth. But we can’t say that the difference here lies in the extension of ‘Enigma’, because both names are alike as regards extension - neither refers to anything. The difference is, so to speak, that they aim at different places, although neither hits anything.

So, our concept of the external aspect of meaning must include this. ‘Enigma’ here has a different external meaning from ‘Enigma’ on Twin Earth. If we want a pithy phrase, we might say that the external meaning of an expression involves, at least in a large range of cases, the external projective relations it bears to the world.¹³

This raises some delicate questions about how exactly our notion of external meaning is to work. What about expressions whose internal meaning determines their extension? For example, expressions like ‘4’, ‘is an even number’, perhaps ‘the property of happiness’. (This talk of ‘determining extension’ may be equivocal, not least of all because ‘extension’ may be equivocal, and perhaps we should carefully distinguish a number of things this might mean, but we won’t go into that here.) Should we say they have external meaning, but don’t bear any external projective relations to the world? Or should we say they have no external meaning? This may be an unimportant terminological matter, but I tend to favour the former. So, I propose that we think of what I call ‘external meaning’ as involving any extensions of the expression and subexpressions involved (so that John himself is involved in the external meaning of, not just ‘John’, but ‘John runs’), together with any external projective relations the expression and subexpressions involved bear to the world. Thus there will be cases where an expression’s external meaning involves just extension (‘4’, ‘2 + 2 = 4’), cases where it involves just external projective relations (‘Enigma’), and cases where it involves both (‘Oscar’, ‘water’, ‘Oscar is home’). (There also seem to be cases where we have neither. Perhaps the right thing to say about ‘yes’ or ‘hello’ for example, is that these words, in their primary use at least, have no external meaning.)

¹³ This bit of jargon is adapted from the Tractatus 3.12, where Wittgenstein says that a proposition is ‘a propositional sign in its projective relation to the world’. Since in the Tractatus the ‘projective relation’ is the only component besides the sign, whereas I have the sign, internal meaning, and external projective relations, it seems reasonable to suppose that the Tractarian notion of projective relations may have covered more, so to speak, than my notion of external projective relations.
6.6. Names

A major question about proper names in the philosophy of language is:

(N) Do proper names have meaning, and if so, what does the meaning of a proper name consist in?

The account of internal and external meaning being sketched here, in the first instance with respect to propositions, can also be applied to proper names. This, I will now argue, yields an answer to (N) which offers satisfying and elegant solutions to problems which other answers struggle with.

There are two classic types of answer to (N): Frege-Russell style descriptivism and Millianism. I will now briefly rehearse these, highlighting their strengths and weaknesses. Then I will put forward my answer, and show how it combines the strengths of both while avoiding the weaknesses.

A word on the dialectical situation here is in order. There are sophisticated descendants of Frege-Russell descriptivism and sophisticated versions of Millianism. I make no claim to have demonstrated in this section that my answer to (N) outperforms any such sophisticated views, let alone all of them. Having said that, sophisticated Millians have in my view made a good case against the views of sophisticated descriptivists, and vice versa. So - going with the thrust of both of sets of criticisms, without trying to add substantially to either - my starting point here is that it is time to consider another approach. Given that starting point, I use classic Frege-Russell descriptivism, Millianism unsupplemented with sophisticated auxiliary considerations, and the problems with each, to motivate and explain my approach. Let us begin with classic Frege-Russell descriptivism.

Frege held that names can have two aspects of meaning: sense and reference. The referent, if the name has one, is the object to which it refers. Its sense, Frege held, was a *mode of presentation* (Frege (1892), p. 26). Frege seems to have held that senses are, or can (at least sometimes) be given by, definite descriptions. He writes things like this, for example:

> In the case of an actual proper name such as ‘Aristotle’ opinions as to the sense may differ. It might, for instance, be taken to be the following: the pupil of Plato and teacher of Alexander the Great. (Frege (1892), p. 27, f.n. 2).

Russell held that names - at least ordinary proper names, which are our topic here - are disguised definite descriptions:

---

16 Compare Fine (2007, p. 37): ‘Current philosophical thinking on Frege’s puzzles has reached an impasse’.
17 Regarding my use of ‘aspects of meaning’: here I am describing Frege’s view in my terms.
18 The page reference is to the original German publication. The translation is Max Black’s.
We may even go so far as to say that, in all such knowledge as can be expressed in words, with the exception of “this” and “that” and a few other words of which the meaning varies on different occasions - no names in the strict sense occur, but what seem like names are really descriptions (Russell (1919), p. 178, ‘Descriptions’.)

These two answers to (N) share some great strengths. They offer solutions to the following puzzles:

(1) How can a true identity statement of the form ‘a is b’ differ in meaning from the corresponding ‘a is a’?

(2) How can a true negative existential statement ‘a does not exist’ have a meaning at all?

(3) How can different true negative existential statements involving names differ in meaning from one another?

Frege-style solution to (1): a true identity statement of the form ‘a is b’ can differ in meaning from the corresponding ‘a is a’, because while ‘a’ and ‘b’ will have the same referent, they can differ in sense, leading to a difference in meaning between the statements.

Russell-style solution to (1): a true identity statement of the form ‘a is b’ can differ in meaning from the corresponding ‘a is a’, because ‘a’ and ‘b’ may be different disguised descriptions.

Frege-style solution to (2): ‘a does not exist’ can be true and have a meaning, because while the name ‘a’ will lack reference in that case, it can still have sense.

Russell-style solution to (2): ‘a does not exist’ can be true and have a meaning, since in that case this will be a disguised statement of the form ‘The so-and-so does not exist’. (If puzzlement persists, i.e. if it is made to seem problematic that a true statement of the form ‘The so-and-so does not exist’ can have a meaning while ‘The so-and-so’ lacks reference, then we can go on to apply Russell’s theory of descriptions and analyze this in turn as saying that it is not the case that there is exactly one so-and-so.19)

Frege-style solution to (3): True negative existentials involving names can differ in meaning from one another because the names involved can have different senses.

Russell-style solution to (3): True negative existentials involving names can differ in meaning from one another because the names involved can be different disguised descriptions.

However, these two answers to (N) also share some serious weaknesses. This was forcefully argued by Kripke in Naming and Necessity. In rehearsing Kripke’s arguments, I will for ease of exposition stick to Russell’s version of descriptivism - in particular, its clear

---

19 This is, of course, just a handy way of summing up the way Russell would analyze such a statement, not the actual full analysis itself, which does not contain a word for ‘exactly’ and which makes use of scope distinctions.
implication that ordinary proper names are synonymous with definite descriptions (nothing extra, about disguise or abbreviation, is relevant to the weaknesses). Frege’s version is a little more open to interpretation, and whether the arguments also apply equally to Frege’s view of names as he understood it may perhaps be doubted, but it is hard to escape the conclusion that the view of sense and reference suggested by his writings, particularly (1892), is vulnerable here.20

Modal.21 Names, Kripke pointed out, are rigid designators. When we describe counterfactual scenarios using names which have a reference in the actual world, those names always designate, with respect to those counterfactual scenarios, the same object: the object that they actually designate. Definite descriptions, on the other hand, can designate different objects with respect to different counterfactual scenarios. ‘The teacher of Alexander’, for instance, which actually designates Aristotle, would designate, with respect to counterfactual scenarios in which someone else taught Aristotle, whoever that was, rather than Aristotle. This difference is reflected in the fact that ‘Had things gone differently, Aristotle might not have been the teacher of Alexander’ has a true reading, while ‘Had things gone differently, Aristotle might not have been Aristotle’ does not (at least, none corresponding to what the first sentence means on its nearest-to-hand true reading).22

Epistemic.23 If a name like ‘Aristotle’ were synonymous with some definite description, like ‘the teacher of Alexander’, then the proposition ‘If Aristotle exists, Aristotle is the teacher of Alexander’ would be analytic and therefore knowable a priori. But it is not, so the name ‘Aristotle’ cannot be synonymous with ‘the teacher of Alexander’.

Semantic.24 If the description we associated with a proper name turned out not to apply to the object we thought it applied to, we would not ordinarily conclude that the name designates, not the object we thought it did, but the object which actually satisfies the description. Kripke’s main example is known as the Gödel-Schmidt case: suppose we

20 Jeff Speaks navigates this matter well in a course handout: Russell explicitly claimed that the meanings of proper names were equivalent to the meanings of descriptions associated with those names by speakers, and Frege consistently uses definite descriptions in explaining the sense of proper names, which indicates that he thought that there was some very close relationship between the sense of names and the sense of descriptions. (http://www3.nd.edu/~jspeaks/courses/mcgill/415/kripke-descriptivism.html, last accessed: 12 Oct 2016.)

21 See Kripke (1980, pp. 48 - 49, 71 - 77).

22 One neo-descriptivist response to this argument, found in Dummett (1973, pp. 110 - 151), is to hold that a name’s associated description must always take wide scope, so that ‘Necessarily, a is F’, if we associate the description ‘the G’ with ‘a’, gets analyzed as ‘The G is such that, necessarily, it is F’ rather than as ‘Necessarily, the G is F’. Kripke argues against this in the preface to the second edition of Naming and Necessity (Kripke 1980, pp. 13 - 14), complaining that this privileging of a wide scope reading is ‘unaccountable’, and adding (in effect) that, in any case, we could always just utter ‘a is F’ and then say ‘What that sentence expresses could not have failed to be the case’. Another neo-descriptivist response to this argument is to ‘rigidify’ the descriptions they appeal to using an ‘actually’ operator (see Jackson (1998b)), or Kaplan’s ‘dthat’ operator (see Kaplan (1978), Chalmers (1996)). The rigidification strategy is criticized extensively in Soames (2002, Chapter 2), along with further criticism of the wide scope strategy already discredited by Kripke.

23 See Kripke (1980, pp. 78, 86 - 87).

24 See Kripke (1980, pp. 78 - 85).
associate with the name ‘Kurt Gödel’ the description ‘the prover of the incompleteness of arithmetic’. Now, if it turned out that Gödel actually stole the proof from someone else, Schmidt, and passed it off as his own, we would not conclude that ‘Kurt Gödel’ designates Schmidt.  

Frege’s and Russell’s descriptivisms face some more direct intuitive difficulties as well. You might say:

‘When you have named an object as a preliminary to saying something about it, you haven’t, just by naming it, already said anything about it - haven’t yet described it at all. You have just picked it out.’

Now, a Frege-Russell descriptivist can agree with this, by maintaining that it is only by using a complete sentence that you can say something. But it is hard to shake the feeling that this objection is still getting at something. Suppose you do utter a complete sentence beginning with a name, and thus end up saying something. Now, while the Frege-Russell descriptivist can maintain that, when you had just said the name, you hadn’t yet said anything, it would seem that they cannot agree that the name, in the context of the whole sentence, just picks an object out without describing it.  

‘It sounds funny to ask “What does ‘John’ mean?”’

---

25 Neo-descriptivists have generally focused more on Kripke’s modal argument than on the epistemic or semantic arguments, but they have responded to these latter by offering more sophisticated descriptive contents, or opting for a more attenuated connection between names and descriptions. Chalmers pursues this last strategy, suggesting that, while his view could be thought of as a ‘highly attenuated form of descriptivism’, he prefers calling it ‘two-dimensionalism’ and sees it as a distinct, but still Frege-inspired, view. (Chalmers (2006), p. 3.) Also, for a response to the semantic argument appealing to a minimal notion of linguistic competence see Stanley (1999).

26 Compare this passage from Kripke’s ‘Identity and Necessity’:

At least if one is not familiar with the philosophical literature about this matter, one naively feels something like the following about proper names. First, if someone says “Cicero was an orator,” then he uses the name ‘Cicero’ in that statement simply to pick out a certain object and then to ascribe a certain property to the object, namely, in this case, he ascribes to a certain man the property of having been an orator. If someone else uses another name, such as, say, ‘Tully’, he is still speaking about the same man. One ascribes the same property, if one says “Tully is an orator,” to the same man. So to speak, the fact, or state of affairs, represented by the statement is the same whether one says “Cicero is an orator” or one says “Tully is an orator.” It would, therefore, seem that the function of names is simply to refer, and not to describe the objects so named by such properties as “being the inventor of bifocals” or “being the first Postmaster General.” (Kripke (2011), p. 5.)

27 Compare:

One response is to insist that proper names do indeed have meaning (...). But this seems strange. One does not find them in the dictionary, and the question ‘What does “David” mean?’ sounds confused. (Whiting (2016), 4.b.)

(What the response is a response to doesn’t matter here.) Incidentally, the claim that names don’t appear in the dictionary appears to be false. For instance, the current (October 2016) Merriam-Webster dictionary has an entry for ‘France’, and even one for ‘Aristotle’. (The latter entry just gives his dates and says ‘Greek philos.’.) And that names don’t all appear in the dictionary is not a serious problem for the idea that they have meaning. After all, people certainly use expressions
This is puzzling for Fregean descriptivism, which would seem to suggest that this question can be answered by specifying the sense of the name (and thereby any reference as well) using a definite description. Likewise for Russellian descriptivism, which would seem to suggest that this question can be answered by giving the undisguised description that ‘John’ is a disguised version of. (You might think that there is a way to avoid this puzzle by appealing to the fact, certainly admitted by Frege, that in ordinary language people’s senses don’t always match up - that they use expressions with different senses, something which wouldn’t be allowed in a proper scientific language, but manage to get on anyway. But then the following puzzle arises: if anything, ‘What does “John” mean in your usage?’ sounds even stranger than ‘What does “John” mean?’.)

Now, let us compare Millianism. Millianism may be regarded as the thesis that proper names have no meaning, or it may be regarded as the thesis that the meaning of a proper name is just its referent, if it has one. Both versions include the thesis that a proper name has no meaning beyond its referent. (According to the “no meaning” version this will be true a fortiori.)

One great strength of Millianism is its compatibility with, indeed its implication or near-implication of, Kripke’s insight that proper names are rigid designators; if there’s nothing to the meaning of a proper name beyond its referent, then of course it will designate that same referent with respect to any counterfactual scenario, as long as that referent exists in that scenario.

Millianism, at least given the plausible assumption that ‘the teacher of Alexander’ does have a meaning beyond its referent, correctly predicts that ‘Had things gone differently, Aristotle might not have been the teacher of Alexander’ differs in meaning from ‘Had things gone differently, Aristotle might not have been Aristotle’. So the fact that the first is naturally read as true and the second is naturally read as trivially false is no problem for Millianism.

Millianism is also free from descriptivism’s epistemic problem - it doesn’t wrongly imply that any proposition like ‘Aristotle is the teacher of Alexander’ is analytic and therefore a priori - and it avoids the semantic problem, suggesting the right answer in the Gödel-Schmidt case: finding out that Schmidt and not Gödel proved the incompleteness of arithmetic wouldn’t make us think that ‘Gödel’ designates Schmidt, and that is just what Millianism would suggest, since according to Millianism ‘Gödel’ has no meaning beyond Gödel himself.

which are not names, and which clearly have meanings, despite not appearing in dictionaries. Think of technical usages, and usages confined to close-knit groups, or pairs, of people.

28 J.S. Mill held two theses about names which for him were intimately related. (1) That they denote but do not connote, i.e. do not ‘indicate or imply any attributes as belonging to’ their referents. This is supported with the famous Dartmouth argument. (See Mill (1843, I, ii, §5, p. 20).) (2) That they have no ‘meaning’ or ‘signification’. (See Mill (1843, p. 21).) As we are understanding Millianism here, and as it is commonly understood in contemporary philosophy, it is (2) that is (the “no meaning” version of) Millianism. (1) all by itself does not imply Millianism as understood here; the view of names I propose, which contradicts Millianism, is compatible with (1). An expression may not indicate or imply any attributes, but still have a meaning (beyond any referent it may have). I am not taking a position on (1) in this thesis.
Finally, Millianism comports well with the intuitive thoughts about naming expressed above. To review them:

‘When you have named an object as a preliminary to saying something about it, you haven’t, just by naming it, already said anything about it - haven’t yet described it at all. You have just picked it out.’

That is exactly what Millianism predicts. How could you say something about an object just by using its name, if that name has no meaning beyond the object?²⁹

“It sounds funny to ask ‘What does ‘John’ mean?’"

This is no problem for Millianism. Obviously, the no-meaning version of Millianism is in an especially good position here: if we hold that thesis, we can reply ‘Of course it sounds funny - “John” doesn’t have a meaning.’ The meaning-is-the-referent version is, while perhaps less glaringly free from this problem, also without cause for embarrassment. After all, if we were asking for a specification of the referent of a man’s name, we wouldn’t say ‘What’, but ‘Who’.

So, Millianism is strong just where descriptivism is weak. But the reverse is also true. For Millianism, puzzles (1) - (3) above are serious problems.

(1) How can a true identity statement of the form ‘a is b’ differ in meaning from the corresponding ‘a is a’?

In responding to this question, a defender of Millianism could either deny the presupposition of the question, that such statements can differ in meaning, or accept that they can and try to reconcile it with Millianism. Both options are problematic. The first option is problematic chiefly because it just does intuitively seem like a true identity statement of the form ‘a is b’ can differ in meaning from the corresponding ‘a is a’. For instance, ‘Clark Kent is Superman’ seems to mean something other than ‘Clark Kent is Clark Kent’. This seems so basic, so undeniable, that any sophisticated theoretical attempts to deny it are bound to be suspicious. (It is one thing to develop a theoretical account of names which uses ‘meaning’ in such a way that ‘Clark Kent is Superman’ and ‘Clark Kent is Clark Kent’ are said to mean the same, and uses other language to describe the difference, but it seems reasonable to think that any such theory must be using ‘meaning’ and the like in a specialized sense, and not in any intuitive sense, in particular, not in the sense it is used in in the question (N).)

²⁹ You might feel that this raises another issue, on which Millianism fares worse than Frege-Russell descriptivism: the issue of how reference is determined. If so, see Objection 1 to my answer to (N) below, and the reply.
³⁰ See Salmon (1986), Soames (1987, 1989, 2002), and Báve (2008). These authors attempt to explain away the anti-Millian appearances by means of pragmatic considerations, making especial use of Grice’s notion of conversational implicature (see Grice (1989)).
The second option\(^{31}\) - accepting the meaning-difference and trying to reconcile it with Millianism - is problematic, because the salient difference between ‘Clark Kent is Clark Kent’ and ‘Clark Kent is Superman’ is that the second statement contains a different name. The natural thing to do is to lay the difference in meaning between the two statements at the door of the different names: ‘Superman’ differs in meaning from ‘Clark Kent’, leading to a difference of meaning in the two statements. The contrary idea, that the names have, if any meaning at all, exactly the same meaning, but that the two statements differ in meaning anyway, seems unnatural and mysterious.

(2) How can a true negative existential statement ‘a does not exist’ have a meaning at all?

Millianism has trouble here, since if ‘a does not exist’ is true, then ‘a’ has no referent, and according to Millianism it must therefore have no meaning. But it seems to be making a crucial contribution to the meaning of the statement. Denying that such statements have meanings is implausible. Accepting that they have meanings despite the names in them not having meanings makes it unclear how such statements are supposed to work, and what the names are doing in them. This may lead to attempts to give an analysis of such statements, unpacking them into a different construction. But such analyses face serious difficulties.\(^{32}\) Secondly, it may be questioned whether such analyses, even if they did work, actually put Millianism in the clear: if we have a kind of analysis which yields different meaningful statements for different inputs of the form ‘a does not exist’, depending on what goes in place of ‘a’, isn’t it for that very reason problematic to say that the names which go in the ‘a’ position lack meaning?\(^{33}\)

(3) How can different true negative existential statements involving names differ in meaning from one another?

For Millianism, this question just compounds the two sorts of difficulty we saw above.

Let us now see how the present account of internal and external meaning, when applied to names, yields a view which has all of the above strengths of the Frege-Russell descriptivisms and of Millianism, and none of the weaknesses. This view, I will argue, combines the meaning-conferring and difference-making power of the descriptivisms, i.e. the ability to solve puzzles (1) - (3), with Millianism’s compatibility with names being rigid designators and invulnerability to Kripke’s anti-descriptivist arguments. It also squares well with the intuitive ideas about naming we have been considering.

---

\(^{31}\) See Fine (2007). One of the drawbacks of Fine’s account, pointed out in a review by Ostertag (2009, p. 348), is that it does not offer an analogous solution to what Ostertag calls ‘the monadic form of Frege’s Puzzle’: ‘Tully was an orator’ and ‘Cicero was an orator’ are not, in Fine’s account, counted as saying different things.

\(^{32}\) See the last chapter of Kripke (2013).

\(^{33}\) I am not aware of an existing source for this (proto-)objection. This is not the place to develop it, but to do so might be worthwhile.
According to the present account, the answer to (N) is as follows: the meaning of a proper name consists in its external meaning (if it has one\textsuperscript{34}) together with its internal meaning, i.e. the role it plays in the language system to which it belongs.

Let us first see how this enables us to solve puzzles (1) - (3) just as satisfyingly as could Frege’s and Russell’s descriptivisms:

(1) How can a true identity statement of the form ‘a is b’ differ in meaning from the corresponding ‘a is a’?

Solution: the names ‘a’ and ‘b’ can have different internal meanings - play different roles in the language system to which they belong - leading to a difference in internal meaning between the statements. (For instance, ‘Clark Kent’ and ‘Superman’ play different roles in language. I think this is intuitively correct, and it explains how Lois Lane could be surprised by ‘Clark Kent is Superman’ but not by ‘Clark Kent is Clark Kent’.)

(2) How can a true negative existential statement ‘a does not exist’ have a meaning at all?

Solution: the statement can have a meaning because, while ‘a’ lacks a referent in this case, it will still have an internal meaning, i.e. it will still play a role in its language system.

(3) How can different true negative existential statements involving names differ in meaning from one another?

Solution: they can differ in meaning because the names involved may differ in internal meaning, i.e. play different roles in their language system, leading to a difference in internal meaning between the statements.

These solutions seem every bit as straightforward and plausible as those given by the classic descriptivisms. Let us now see how our proposal fares with Kripke’s anti-descriptivist arguments.

\textit{Modal}. Names being rigid designators is perfectly compatible with their playing roles in language systems, and with these roles being an aspect of their meaning. Indeed, the rigid designation thesis gives us information about what sort of role they play; it is part of the sort of role names play that, when we consider counterfactual scenarios, they designate, with respect to the scenarios, the same objects that they actually designate, at least provided those objects exist in the scenarios. And so there is no problem either in the fact that ‘Had things gone differently, Aristotle might not have been the teacher of Alexander’ differs in meaning from ‘Had things gone differently, Aristotle might not have been Aristotle’ - this is because ‘Aristotle’ and ‘the teacher of Alexander’ differ in meaning, a fact which is perfectly compatible with the present account.

\textsuperscript{34} Perhaps a name in mathematics whose putative reference is fixed using a definite description which in fact doesn’t and couldn’t refer should be thought of as having no external meaning.
Epistemic. For a name to play a role in a language system, it does not need to be synonymous with, or otherwise closely tied to, any particular definite description. Again, ‘Aristotle’ and ‘the teacher of Alexander’ differ in meaning, and there is no tendency for the present view to push us toward the implausible claim that a proposition like ‘If Aristotle existed, Aristotle was the teacher of Alexander’ is analytic, or knowable a priori.

Semantic. Again, no problem. If we discovered that Gödel didn’t prove incompleteness, but rather Schmidt did, we wouldn’t be pushed towards the view that ‘Gödel’ designates Schmidt, since the role played by ‘Gödel’ does not make it designate whoever proved completeness - ‘Gödel’ has a different meaning from ‘the prover of incompleteness’.

So, Kripke’s arguments against descriptivism leave the present account unscathed. What about the intuitive thoughts about names which we saw descriptivism struggle with, and which Millianism handled so well?:

‘When you have named an object as a preliminary to saying something about it, you haven’t, just by naming it, already said anything about it - haven’t yet described it at all. You have just picked it out.’

No problem for the present view here. Names having internal meanings, i.e. playing roles in language systems, is perfectly compatible with the idea that just by naming something, you haven’t yet described it, but just picked it out. We can agree with this and say: yes, the role a name plays is not that of a description.

“It sounds funny to ask “What does ‘John’ mean?”

Just because it sounds funny to ask for the meaning of some expressions, that doesn’t mean they don’t have meanings. It is natural to construe such a question as asking for a synonymous expression, or a definition - something else which has the same meaning, and which perhaps unpacks the meaning of, the expression being asked about. But a proper name can have an internal meaning - play a role in a language system - without it being the case that there is another expression, perhaps a more complex one, which has the same internal meaning. Names, we might say, are, in a sense, indefinable - and in that case, no wonder the question sounds funny.35

I have just indicated how this chapter’s account of internal and external meaning, when extended to names, yields an answer to (N) which combines the major strengths of the classic descriptivists with those of Millianism, while avoiding the major weaknesses of both. Before proceeding to the next section, in which I will elaborate a further part of my account of internal meanings, having to do with their granularity or individuation, I will consider some objections to my answer to (N). The last of these objections will lead us straight into the next section.

---

35 While I lean towards the idea that names are, at least typically, indefinable in this sense, there being a good reply here does not require this. If names were definable but not easily, or not usefully, this would lead us to expect questions as to their meaning to be unusual or even unheard of, at least outside of philosophy. This would suffice to account for our question sounding funny.
Objection 1: Your answer to (N) does not give us an account of the determination of reference. This is a major strength of the classic descriptivisms: according to them, the reference of a name is determined by means of the associated description. So perhaps your answer to (N) is not so good after all.

Reply: OK, so the classic descriptivisms give you something to say there, and my answer to (N) does not. But I think that, on reflection, we should not think that this gives the former any real edge over the latter. Here are three reasons for thinking it doesn’t.

Firstly, the account of the determination of reference provided by the classic descriptivisms seems to give wrong answers, in view of Kripke’s Gödel-Schmidt case and similar thought experiments. And surely it is better for an answer to (N) to be silent on some issue than to make mistakes about it.

Secondly, as far as I can see, there is no independent reason to think that any good answer to
(N) Do proper names have meaning, and if so, what does the meaning of a proper name consist in?

ought to be able to answer

(D) What determines the reference of a proper name?

They seem to be pretty different questions. Millianism, for all its faults, sets an instructive precedent here. It doesn’t, by itself, give an answer to (D) either.

Thirdly, while my answer to (N) does not answer (D), it presents no special difficulties either. Internal meaning may be held to constrain reference, with other things doing the rest. This seems plausible; a name with the internal meaning of ‘Julius Caesar’ could never refer to, say, a number. So we might say that internal meaning constrains reference, without (in general) determining it. This leaves at least two options for a straight answer to (D): appeal to internal meaning as part of the story, or “go deeper” and frame an answer in terms of stuff which determines both internal meaning and reference. Alternatively, we might avoid giving a straight answer and instead subject the question (D) to scrutiny.

Finally, it is worth noting in this connection that Kripke’s causal-historical ‘picture’ of how reference is determined seems quite compatible with the present view - though I hasten to add that the present view does not require that picture, or push us towards it in any special way of its own.

Objection 2: Your answer to (N) may be able to account for the fact that different true negative existentials can differ in meaning from one another, and it may also differ from

---

36 See Kripke (1980, pp. 88 - 97.)

107
Millianism in not creating a special obstacle to their having meanings at all - but still, it does not remove all puzzlement surrounding existence statements involving proper names.

Reply: I agree that puzzlement may remain, but would suggest that we look for a resolution of it some place other than in an answer to (N). That said, I think the notion of internal meaning being developed here may be of further use in this connection. Puzzlement about statements of this sort, in my view, stems from assimilating them with (other) subject-predicate statements, and failing to make sufficient room for them in our taxonomies of language. By attending to the special sort of role they play in our language systems - i.e. by attending to the sort of internal meaning they have - and letting them be themselves, I think we can resolve the puzzlement. (Kant’s famous contention that existence is not a predicate contained a lot of truth in this regard, but its significance needs to be made clearer.)

You might think that existence statements involving proper names require an analysis, in the sense of an account which spells out what they mean. I don’t think so. Analysis, in this sense, is one method in philosophy. Yes, classic descriptivism, at least in Russell’s version, employs that method with these statements, but as we have seen, it runs into problems elsewhere. Millianism, while it doesn’t provide an analysis, pushes one strongly toward thinking that singular existence statements, and especially their negations, need to be analyzed. They clearly have meaning and can be true and false, and to square this with Millianism, it would seem you need to give some account of what they really say. Thus you get into the tortuous territory Kripke tried his best to navigate in the last lecture of Reference and Existence. But just because these two approaches to names furnish (Russellian descriptivism), or make you need (Millianism), an analysis of existence statements involving proper names, that doesn’t mean we automatically need an analysis no matter what view we take of names. I think once we have a better view of names, we don’t need an analysis anymore. We can just attend carefully to how existence statements work and make room for their peculiarities.

Objection 3: OK, you have shown that your answer to (N) outperforms the classic descriptivisms and unsupplemented Millianism on some key issues. But what about <insert here the name of some sophisticated neo-descriptivist proposal, or a supplemented version of Millianism38>? 

Giving a substantive answer to a question of this form is beyond the scope of the present section. Recall, my starting point here is that such sophisticated versions of descriptivism and Millianism face serious objections, and so it is time to consider another approach. Still, for all I have said here, a sophisticated version of descriptivism or Millianism may yet win the day.

Also, it is worth mentioning here that what I have said about names here may be compatible with some sophisticated versions of descriptivism; my account specifies, at a certain level of

---

37 It is less clear exactly how, or whether, Frege dealt with problems pertaining to negative existentials and empty (ordinary proper) names.

38 For examples of each see f.n. 14 and f.n. 15 above.
abstraction, what the meaning of a proper name consists in, and thus solves the puzzles. So for those puzzles, you don’t really need a sophisticated descriptivism - or so I think. Still, that does not mean such a theory might not have something to offer. Certain forms of descriptivism may be regarded as attempting to model, spell out or describe, at a lower level of abstraction, the internal meanings of names. Having said that, I hasten to add, for those readers hostile to any such theory, that what I have said doesn’t require or push us toward this. It is neutral on the matter.

Objection 4: You have made your answer to (N) look rather good on paper, but you have been evading the chief difficulty for any approach like yours: how are these things (internal meanings in your case) to be counted? That is, how are they individuated? What is their granularity? When do two expression-occurrences have the same internal meaning? If you cannot give general, principled answers to these questions, your account falls apart and your internal meanings are just a will o’ the wisp.

Reply: This is a very serious objection, and I have not (with the exception of a couple of hints) said anything which deals with it. One way of dealing with it would be to go along with the last sentence of the objection (‘If you cannot...’) and try to provide general, principled answers to the questions, or at least good reason to think that such answers are out there to be had. Another way of dealing with it would be to motivate a rejection of the idea of the last sentence of the objection - a rejection of the idea that, if you cannot give general, principled answers to questions about the individuation of meanings, you are in philosophical trouble. I strongly prefer the latter strategy. I believe that a rejection of this idea can be powerfully motivated, and that following through with the rejection opens up exciting philosophical prospects. I will turn to this now. (Others may have some sympathy with my account of propositions and meaning as I have expounded it so far, or parts of it, but want to take the other road here. Accordingly, the following should be regarded as a detachable but important part of my account of propositions and meaning.)

6.7. Flexible Granularity

The leading idea of this section, expressed in a slogan, is that meanings can be carved up at different granularities. This way of expressing it may be confusing, however. For one thing, it may sound like it is saying that there is this particular class of things, the meanings, which can be carved up in different ways - and what could that mean? The idea is really about how the concept of meaning works. For another thing, the reference to ‘granularities’ may make it sound like there are a few different granularities which may be distinguished, at each of which there is a set of truths about which expression occurrences have the same meanings as which others. But that would be overly simplistic.

More carefully, the idea is this: that a sentence of the form ‘X means the same as Y’ may be true in one context while false in another, even though the concept of meaning invoked in each case is, in an important sense, one and the same. I think this is a good way of thinking about much of our ordinary talk of meaning, in particular talk of synonymy, of ‘meaning the same as’. I also think it is a good way of thinking about the notion of internal meaning I am trying to develop in this chapter. Here I will mostly concentrate on substantiating this latter
point, but the two are related - after all, I appealed to intuitive talk of meaning when introducing the distinction between internal and external meaning.

Now, how can one and the same concept be such that, in one context, a pair of things fall under it, while in another context that same pair of things do not fall under it? Let me try to explain using an analogy. You might have a simple tool for slicing pies, which always slices them into five pieces, and you might have a tool which is similar but slices the pies into eight pieces. Or, you might have a more sophisticated tool which can be adjusted, so that on one setting it slices a pie into five pieces and on another setting it slices it into eight. I think we can think of the concept of meaning, or at any rate the concept of internal meaning being developed here, as being like the more sophisticated tool.39

(The above analogy is helpful on this point, but may be misleading in other ways. For instance, the concept of internal meaning may not be well thought of as having a definite number of discrete settings, and the idea, which may be suggested by the above, that it sorts all possible occurrences of meaningful expressions into a definite number of non-fuzzy bundles, is no doubt an idealization.)

I will motivate this idea using Kripke’s puzzle about belief - or rather, a closely related linguistic puzzle arising from the story Kripke uses for his puzzle. I will then argue that, despite perhaps seeming radical from within analytic philosophy of language, the idea makes good intuitive sense, especially given the conception of internal meanings as roles in language systems. (While these main parts of my defence will be purely philosophical, at the end of this section, in f.n. 47, I will show that essentially this idea is in common use in AI research into natural language processing, a fact which I take to provide further support.) Following that, I will suggest some further applications of the idea to difficulties in philosophy. Following that, I will respond to some possible objections.

Here is the puzzle setup as given by Kripke:

Suppose Pierre is a normal French speaker who lives in France and speaks not a word of English or of any other language except French. Of course he has heard of that famous distant city, London (which he of course calls 'Londres') though he himself has never left France. On the basis of what he has heard of London, he is inclined to think that it is pretty. So he says, in French, "Londres est jolie."

(...)

Later, Pierre, through fortunate or unfortunate vicissitudes, moves to England, in fact to London itself, though to an unattractive part of the city with fairly uneducated inhabitants. He, like most of his neighbors, rarely ever leaves this part of the city.

---

39 Note that we need not always say, when ‘X means the same as Y’ is true in one context and false in another, that ‘means the same as’ invokes the same concept, or means the same, in each case. Applying the idea of the flexible granularity of synonymy statements to instances concerning synonymy statements themselves, we can say that “means the same as” in instance X means the same as “means the same as” in instance Y may be true in one context and false in another.
None of his neighbors know any French, so he must learn English by 'direct method', without using any translation of English into French: by talking and mixing with the people he eventually begins to pick up English. In particular, everyone speaks of the city, 'London', where they all live. Let us suppose for the moment - though we will see below that this is not crucial - that the local population are so uneducated that they know few of the facts that Pierre heard about London in France. Pierre learns from them everything they know about London, but there is little overlap with what he heard before. He learns, of course - speaking English - to call the city he lives in 'London'. Pierre's surroundings are, as I said, unattractive, and he is unimpressed with most of the rest of what he happens to see. So he is inclined to assent to the English sentence:

(5) London is not pretty.

He has no inclination to assent to:

(6) London is pretty.

Of course he does not for a moment withdraw his assent from the French sentence, "Londres est jolie"; he merely takes it for granted that the ugly city in which he is now stuck is distinct from the enchanting city he heard about in France. But he has no inclination to change his mind for a moment about the city he stills calls 'Londres'. (Kripke (1979), p. 392.)

Now, Kripke frames his puzzle in terms of belief, and employs a principle linking sincere assertion to belief. But much of what is puzzling about Kripke’s puzzle about belief arises already with Pierre’s linguistic behaviour, and questions about what his utterances mean. Since my topics here are propositions and internal meaning, rather than belief or belief-reports (which I do not want to take a position on), I will stick with linguistic items and their meanings and frame a puzzle analogous to Kripke’s which concerns them. Another reason why this is a good thing to try is that the principle linking sincere assertion to belief, and another principle Kripke invokes involving translation, might be problematic in independent ways, which could allow Kripke’s puzzle to be diffused “too easily”, by simply poking a hole in one of his principles. The analogous puzzle I will use here, on the other hand, does not require them, and therefore cannot be diffused in that way.

I will add one thing to the story above: Pierre is not only inclined to assent to ‘London is not pretty’. He actually says it. So in France he says ‘Londres est jolie’, and then later in England he says ‘London is not pretty’.

If we were to be told only the first half of the story about Pierre, featuring him in France, we would have little hesitation in agreeing that when he says ‘Londres est jolie’, this means the same as ‘London is pretty’ would mean if you or I were to say it. Likewise, if we were told only the second half of the story, we would have little hesitation in agreeing that when he says ‘London is not pretty’, this means the same as ‘London is not pretty’ would mean if you or I were to say it.
Now, with the whole story in mind, would we agree that if Pierre in London reverted to French for a moment, began to think of the city he had heard about which sounded so appealing, and said ‘Londres est jolie’, this would mean the same as what ‘London is pretty’ would mean if he were to say that? Here I think we would hesitate. Indeed, we might be inclined to say that, on the contrary, these would mean different things in Pierre’s mouth, and that ‘Londres’ as used by him doesn’t mean the same as ‘London’ as used by him.

But: given that, how is it that, when told only the first part of Pierre’s story, we would agree that ‘Londres est jolie’ when he says it means what ‘London is pretty’ would mean if we said it? And don’t I, with my account of propositions and internal meaning, want to say that the proposition ‘Londres est jolie’ that came out of Pierre’s mouth has the same internal meaning, plays the same role in its language system, as ‘London is pretty’ as uttered by you or I, and likewise for just the names ‘Londres’ and ‘London’? And yet, pondering the whole story, I wanted to say that ‘Londres’ and ‘London’ when used by Pierre differ in internal meaning. Putting these two things side by side, it looks like I want to say that ‘London’ when I say it means the same as ‘Londres’ when Pierre says it, and yet ‘London’ and ‘Londres’ don’t mean the same when Pierre says them, implying that ‘London’ when I say it doesn’t mean the same as ‘London’ when Pierre says it! But why would that be the case? Why would ‘Londres’ as said by Pierre have a greater claim to meaning the same as ‘London’ when I say it than ‘London’ as said by Pierre? That seems like arbitrary favouritism.

The solution is to see that two different granularities are clashing here. When we say that Pierre’s ‘Londres est jolie’ means the same as our ‘London is pretty’, we are operating at a granularity coarser, at least with respect to the utterances in question, than the one we are operating at when we say that Pierre’s ‘Londres’ and his ‘London’ differ in internal meaning. Without allowing for flexible granularity - without embracing the idea of this section - we can have one of these claims, but not the other (if we want to avoid the kind of arbitrariness which threatened above, at least). To my way of thinking, that just won’t do. Thus I think we should embrace the idea that meanings can be carved up at different granularities. This raises all sorts of questions, but I think doing it is worthwhile. In fact, I think the questions are potentially very fruitful, and that by developing this idea we stand to learn a lot.

I will now try to strengthen the intuitive case for this idea that meanings can be carved up at different granularities. Though it may seem radical within analytic philosophy, it is actually just good sense.

Compare talk about games: faced with two instances of games being played, where the players follow slightly different rules in each case, we might say that both groups are playing the same game, albeit slightly different versions of it, or we might say that they are playing different games. It seems like just good sense to allow that, in some cases of this sort at least, we could correctly say either of the two possible things - if we say the first, we are carving up games more coarsely than if we say the second. Which option we take might depend on our interests, or sometimes either option may be equally good. It’s not as though there’s always some unique correct answer in a case like this, an answer which we may not be able to find out, as to whether the two groups are playing different versions of one game.
or two different games. At least, that is what common sense suggests, I think. In short, games can be carved up at different granularities; ‘is playing the same game as’ has a kind of flexibility to it. I think this is part and parcel of its usefulness. And I want to say that ‘means the same as’, and my more technical ‘has the same internal meaning as’, have a similar kind of flexibility.

Another comparison: you go to a ceremony held by a group of people, and watch a ritual dance. You recognize the way the dancers are moving from another ceremony you attended five years earlier, and say to your companion ‘I’ve seen people dance that way before’. Meanwhile, someone else in the audience who was also at the other ceremony five years ago - someone who knows a lot more than you do about these dances, and has received training in them - thinks of the dance performed five years ago, and notices some systematic differences: when the dancers raise their arms, they always flick their wrists slightly in a characteristic way, which didn’t happen in the earlier dance, and one or two other things like that. Now they say to their companion, who is also thinking of the earlier dance, ‘They’re dancing in a different way this time’. Is one person right here and the other wrong? For instance, was what you said not strictly right? I don’t think we have to say that. In support of this: it could be that, if you had your attention drawn to the slight differences which the other, more in-the-know person was occupied with, you would say ‘OK, but it’s still the same general way of dancing’ or something like that. Your interest in talking to your companion centred on the many striking similarities between the two performances. The interest of the more in-the-know person centred on the differences. I think we can say that you both said something true, but that you were individuating ways of dancing at different granularities: ‘is dancing in the same way as’ is flexible in a similar way to ‘is playing the same game as’ and ‘means the same as’.

Now for something which, given my account, is a bit more than just a comparison. Mary works for an insurance company for a number of years and then moves to another one. At the new company she has a similar title, and mostly does the same sort of work for the new company as she did for the old: her main duties are to meet and negotiate with suppliers and to keep an inventory of the company’s supplies. But there are some differences: the new company is smaller, and where at the previous company she had a team below her who assisted with taking inventory, at the new one she does all that herself. At dinner one evening while she is still settling into the new job, she might say to her husband ‘I’m playing a much more hands-on role here than I played at my old job’. But years later, once she has maybe retired or moved into another field, she might, while recounting some of her life to a recently made acquaintance, say ‘I then worked for an insurance company, managing supplies. Actually, I had that role at a couple of places’. So, while in the thick of things, she distinguished two roles when talking to her husband. At a greater remove, when the differences between the two jobs were not particularly relevant, she said she played the same role at two companies. Was she wrong in one instance? Common sense suggests that no, she wasn’t. The lesson, I say, is that roles - like games and ways of dancing - can be carved up at different granularities.40 Thus, the idea that meanings can be carved up at

---

40 Do not be distracted by the fact that ‘role’ in this sort of context can also just mean a particular job at a company - the point is that, even if Mary is using ‘role’ such that, in principle, the same role can be played at different companies, she might still want, when talking to her husband shortly after starting
different granularities fits well with my proposal that we consider the internal meaning of an expression as the role it plays in its language system.

I think this idea - call it ‘the idea of flexible granularity’ - may be quite helpful in philosophy. Some possible examples:

- It may help us to explain, without acquiescing in, Quine’s negative attitude to concepts like those of meaning and synonymy; a major strand of Quine’s negativity was his charge that these concepts are in bad shape because we lack a stable criterion of identity for meanings. With the idea of flexible granularity in hand, we can diagnose the problem: Quine has mistaken a feature for a bug here, due to prejudices about how language ought to work.
- Confusing questions and disputes about whether certain expressions or types of constructions are ambiguous or not may perhaps be alleviated using the idea of flexible granularity, at least to some extent - perhaps there is an ambiguity at a relatively fine granularity, while at a coarser granularity, all relevant uses of the language in question may be regarded as meaning the same.
- Perhaps we could give a resolution of the paradox of analysis using the idea of flexible granularity. If the expression on the right-hand side of an analysis gives the meaning of the expression on the left-hand side, then, if the analysis is correct, both sides mean the same. But in that case, how can an analysis be informative? We might say: it can be informative to someone if it is understood in such a way that, at a fine granularity, the LHS and the RHS differ in meaning, but at a coarser granularity, they mean the same.
- The idea of flexible granularity may shed light on the topic of compositionality in relation to natural languages, compositionality being the principle that the meaning of a complex expression is determined by the meanings of its parts (perhaps together

---

the new job, to distinguish the role she played at the old company from the role she is playing at the new company.

41 Here is a characteristic passage, from the 1951 edition of ‘Two Dogmas of Empiricism’:

For the theory of meaning the most conspicuous question is as to the nature of its objects: what sort of things are meanings? They are evidently intended to be ideas, somehow – mental ideas for some semanticists, Platonic ideas for others. Objects of either sort are so elusive, not to say debatable, that there seems little hope of erecting a fruitful science about them. It is not even clear, granted meanings, when we have two and when we have one; it is not clear when linguistic forms should be regarded as synonymous, or alike in meaning, and when they should not. (Quine (1951), p. 22.)

And another, from a section of Philosophy of Logic entitled ‘Propositions Dismissed’:

The uncritical acceptance of propositions as meanings of sentences is one manifestation of a widespread myth of meaning. It is as if there were a gallery of ideas, and each idea were tagged with the expression that means it; each proposition, in particular, with an appropriate sentence. In criticism of this attitude I have been airing the problem of individuation of propositions. (Quine (1970), p. 8.)

42 One classic case is Quine’s question about whether ‘hard’ is ambiguous (think of ‘hard chair’ vs. ‘hard test’) - see Quine (1960, p. 130). Some more recent cases: Pietroski and Hornstein (2002) take a controversial stance on some purported cases of scope ambiguity, arguing that they are not ambiguous after all. Dayal (2004) gives the reader a sense of how controversial the question ‘Is “Any” ambiguous?’ is. Breckenridge (2007) argues that the verb ‘look’ is not ambiguous, despite influential arguments to the contrary made by philosophers of perception.

43 See Moore (1942) and Langford & Schilpp (1943).
with the way they are put together).\textsuperscript{44} For instance, confusion and debate over whether certain complex expressions constitute “counterexamples” to, or failures to live up to, the principle, may be clarified by observing that perhaps, if the meanings of the parts are individuated at a fine granularity, compositionality holds, while if they are individuated more coarsely, compositionality fails.

- The idea of flexible granularity may shed light on a certain strand of the motivation for relativism about truth, and on a certain kind of relativist-sounding talk. When people are inclined to say things like ‘That’s true for you, but not true for me’ - something which happens especially in the aesthetic, ethical and religious domains - what is this thing being designated with the ‘that’? We might think: they aren’t just talking about a bare expression here, but they aren’t talking about a single, full-fledged proposition either, since one and the same proposition can’t be true for one person and false for another. We might say that the thing being designated with the ‘that’ is a proposition or somewhat proposition-like thing, but that the granularity at which it is being individuated is so coarse that different instances of it (we are talking about a type here of course, not a token) have different truth-values. What we say about the propriety of such talk is, in a way, secondary - what’s important is that the idea of flexible granularity can give us a way of understanding what is going on when people say this sort of thing.

- Let us characterize metaphysical realism as the view that there is exactly one true and complete description of the world, and let us characterize conceptual relativity as the denial of that claim.\textsuperscript{45} The appearance of an intractable disagreement between these two views may perhaps be alleviated using the idea of flexible granularity in something like the following way: if a complete description of a domain D is a set of propositions such that every proposition which says something true about D means the same as one of the propositions in the set (or a conjunction of them), then whether we say a description of the world is complete depends on the granularity we are operating at. Conceptual relativity may hold at a certain granularity, but once you make the granularity finer, it may collapse into metaphysical realism.

(The above are just some applications that have occurred to me - they are not supposed to be the most important.)

I have now explained and defended the idea of flexible granularity, and indicated some philosophical applications. I will conclude this section by considering some objections.

Objection 1: Doesn’t the idea of flexible granularity lead to a perplexing or objectionable kind of anti-realism about these things called meanings?

Reply: I don’t think it does. The idea is about how the concepts of meaning and synonymy work, and not about the “metaphysical status” of meanings. It seems to me to be perfectly

\textsuperscript{44} For an overview of the topic see Szabó’s (2013) encyclopedia entry. Cohnitz (2005), for one, seems to agree that more light on this topic would be welcome; his paper opens with the remark that ‘[a] superficial look at the literature on the principle of compositionality [...] could suggest that the discussion is as confused as a discussion can be’.

\textsuperscript{45} This opposition is drawn from Putnam (1978, 1981, 1985).
compatible with the metaphysical views that meanings really exist, that there are objective facts about them, and the like. Meanings exist every bit as much as games, roles you can play at different companies, and ways of dancing, but you might say: there’s not one all-purpose lot of meanings - you latch onto different “things” at different granularities. And the idea is not that, if ‘X’ means the same as ‘Y’ is true at one granularity and false at another, then there is no fact of the matter about whether X means the same as Y. For my part, I would rather say this: when interpreted such that it comes out true, the sentence states a fact, and when it is interpreted such that it comes out false, its negation states a fact and does not contradict the original sentence interpreted such that it comes out true. (Compare: I can truly say ‘There’s no milk anywhere’, restricting my attention to just what is in the tea room, and thus state a fact, without contradicting the proposition ‘There’s milk at the shop across the road’.)

Objection 2: Doesn’t the idea of flexible granularity lead to the view that when it comes to synonymy you can just say whatever you like?

Reply: No, it doesn’t. One consideration showing that you can’t just say whatever you like is that, for some pairs of expressions or expression-occurrences, there won’t be any legitimate granularity, however coarse, on which they mean the same. There are limits. To take a really obvious case: it just isn’t the case that ‘if’ means the same at ‘tiger’ at any granularity. On the other side, there are pairs of expressions or expression occurrences such that you can’t, at any legitimate granularity, truly say that they don’t mean the same: for instance, an occurrence of a set theory symbol in a mathematical proof, and another occurrence of the same symbol later in the proof. Another consideration is that, even if a pair of expressions are synonymous at one granularity and not at another, sometimes it will just be a fact that you’re operating at one granularity and not the other, and trying to wriggle out of it in the face of countervailing evidence will usually be seen for what it is. (Compare: Someone says ‘Is there any milk?’, looking in the fridge, and you say ‘Yes, there’s milk’. If it turns out that there is no milk in the fridge, and you look a bit perturbed and then offer ‘Oh, but when I said “there’s milk” I meant that milk exists somewhere’, this probably won’t fly. It might be difficult to give a satisfying philosophical account of what makes it the case that you meant it one way rather than another, but that you can sometimes just be wrong in this way is not something that can reasonably be denied.) The fundamental fact of life here is that, while our talk and thought of meaning and synonymy is certainly flexible, we do revise it, admit mistakes, and the like.

Objection 3: When you use a piece of language that has this kind of flexibility, what makes it the case that you are using it one way rather than another? For instance, if you utter a true (or false) instance of ‘X means the same as Y’ where this same sentence is false (or true) at another granularity, what makes it the case that you meant it the way you did, such that it had the truth-value that it had?

Reply: I already touched on this worry in the previous reply, and hinted at what I think the proper response is: this sort of question is vexing and difficult enough in general that the very idea of flexible granularity should not be rejected in lieu of a fully satisfying philosophical
treatment of it. (Compare Kripke’s famous problem\textsuperscript{46}, inspired by Wittgenstein: What makes it the case that by ‘+’ I mean the function \textit{plus} rather than \textit{quus}? Or the kind of case used in the previous reply: When someone uses a quantifier with implicit contextual restrictions, what makes it the case that those restrictions, and not others (or none) are in force?) It’s also worth noting that the task of answering the question may be approached at different levels. We might talk about intentions. Or behavioural dispositions (including verbal behaviour). Or concrete facts about what has been going with us and around us - which in turn can be described at different levels.

There are plenty of questions still to be answered, but we can proceed happily with them. In fact, this seems to me like an exciting direction for further research.\textsuperscript{47} In arguing that the idea

\textsuperscript{46} See Kripke (1982).

\textsuperscript{47} For instance, the applications of the idea suggested above could be developed, and further applications found. Also, the idea itself may be developed in various ways.

We might ask about the limits of the flexibility (touched on above in the reply to Objection 2). We might try to make generalizations about different uses of these flexible notions - perhaps characteristic patterns of individuation tend to occur in certain circumstances. (How much can be said here, both in the way of exceptionless generalizations and more heuristic stuff, I do not know.) And we might try incorporating the idea into technical models of language, thought and communication.

This last suggestion is addressed to philosophers. Addressed to everyone it might strike some as silly and ill-informed, for in a practical sense this idea seems to be used all the time in technical work. AI researchers working on natural language processing, particularly in the area known as ‘word sense disambiguation’, talk about granularity a lot. (I discovered this well after starting to think and write about the idea, motivated primarily by Kripke’s puzzle about belief and the closely related puzzle about sentence meanings discussed in this section.) Consider for instance this passage from the abstract of McCarthy (2006):

The granularity of word senses in current general purpose sense inventories is often too fine-grained, with narrow sense distinctions that are irrelevant for many NLP [natural language processing] applications. (...) We propose relating senses as a matter of degree to permit a softer notion of relationships between senses compared to fixed groupings so that granularity can be varied according to the needs of the application.

This bears striking similarities to what I have been saying, quite independently, in this thesis. (Recent overviews of the field of word sense disambiguation include Navigli (2009) and Pal & Saha (2015).) What better evidence could there be that analytic philosophy should take the idea of flexible granularity seriously?

We may also want to extend the idea to other notions. For instance, the notion of a fact. (Here I am thinking not of special philosophical notions of facts used in metaphysics, but of our working intuitive notion of a fact - not to say that there is a sharp distinction here, or that the special philosophical notions aren’t related to the intuitive notion.) Perhaps this notion also possesses an analogous flexibility, but a flexibility which follows its own characteristic pattern. In favour of this: it is natural both to distinguish the meaning of Pierre’s ‘\textit{Londres est jolie}’ from the meaning of our ‘London is pretty’, and to say they mean the same, given certain descriptive circumstances, but it is less natural to say that there are two facts here. (I am supposing that London is in fact pretty. If you’re worried about London’s prettiness, or whether there are facts about what is pretty, examples avoiding these worries are obviously near to hand.) So considering just that case, you might think that the idea of flexible granularity doesn’t apply to the notion of a fact. On the other hand, when we consider facts about what is identical to what - in particular, the sort of facts we state with identity statements involving proper names - I think we do find considerations suggesting that facts can be carved up at different
of flexible granularity makes good sense, I have tried to alleviate the worry that there is something fundamentally wrong with our account of propositions and meaning insofar as we can’t give a single, unified account of how meanings are individuated - a single criterion of identity for meanings. That worry presupposes something false: that no legitimate notion can have the kind of flexibility that the notions of playing the same game as, dancing the same way as and occupying the same role as all seem to have.

6.8. How This Account Fits With That of the Previous Chapter

The account of subjunctive necessity de dicto given in the previous chapter appeals to the notion of inherent counterfactual invariance. According to the account, a proposition is inherently counterfactually invariant iff its negation does not appear in any (genuine) counterfactual scenario description for which it is held true. A main desideratum of the present account of propositions and meaning is that it be able to underpin the idea that there are these things called propositions, and that they can bear a property like inherent counterfactual invariance.

We can now see that this desideratum is fulfilled. I call a ‘proposition’ an expression in meaningful propositional use. It is pretty clear that there are such things. With regard to their being able to bear a property like inherent counterfactual invariance, my account of them makes good sense of this. I distinguish two aspects of a proposition’s meaning, internal and external, and construe internal meaning as the role an expression plays in the language system to which it belongs. The inherent counterfactual invariance of a proposition which possesses it can naturally be regarded as an aspect of its internal meaning - on my account, as an aspect of the role it plays in the language system to which it belongs. After all, the notion of ICI is cashed out in terms of the proposition’s behaviour (or at least the behaviour of its negation, which may by extension be regarded as an aspect of the original proposition’s behaviour) in a certain type of context - namely that of (genuine) counterfactual scenario descriptions. And this is the sort of thing we think of as part of something’s role. (Compare: it is part of a bureaucrat’s role that they behave in certain ways in certain situations. Furthermore, just as you might use the words of the negation of some ICI proposition with a different meaning, such that they do appear in a CSD, a bureaucrat might behave contrary to their office in some situation - but insofar as they do so, they are not playing their assigned role.)

I hope I have succeeded in sketching an independently attractive account of propositions and meaning to go with my account of subjunctive necessity de dicto. At the very least, I hope the account is worthy of consideration. You might prefer another approach, but still think my account of necessity de dicto is true or on to something. Or you might not care for the account of necessity de dicto, but like the account of propositions and meaning. In any case, they fit together nicely.

Chapter 6 References

granularities. Putting these two things together, I think we should consider the idea that the notion of a fact does exhibit a similar flexibility, but on a different pattern.


7. Conclusion

7.1. Stepping Back

In this thesis, I have set out a task (that of giving an illuminating account of the conditions under which a proposition is subjunctively necessary), criticized some existing accounts (with special focus on modal realism and Sider’s quasi-conventionalism), presented my own account, and then sketched an account of propositions and meaning to underpin that account.

Stepping back a bit, one of the aims of this thesis has been to clarify and build on the insights of Kripke’s Naming and Necessity. (A subsidiary aim, emerging in Chapter 6, has been to bring Wittgensteinian ideas about meaning - especially the middle-period idea of an expression’s meaning as the role it plays in language - to bear on puzzles and problems in analytic philosophy of language.)

As I see it, Naming and Necessity presents us with four major problems, or four major obstacles to a deeper understanding of the topics it deals with (there are no doubt others, but there are at least these four):

1. Kripke’s curious doubling-down on certain verdicts concerning what may seem like unclear or borderline cases. (Necessity of origin, necessity of constitution.¹)
2. The apparent push toward Millianism about proper names, with all its attendant problems, which comes from the work.
3. Uncertainty about whether, or how, semantic considerations come into the picture when we ask whether a given truth is necessary or contingent.
4. The worry that Kripke, in - as Chalmers (1998) puts it - ‘appealing to the subjunctive’, has overlooked a distinct notion which may sometimes be a better candidate for what a philosopher means by ‘necessary’ (particularly in pre-Kripkean analytic philosophy), and skewed our collective understanding of modality toward the subjunctive as opposed to the indicative.

Regarding (1), I have suggested that we just not worry much about these cases; we should simply put them to one side as unclear or borderline. This, I think, helps in a preliminary way to demystify the notion of subjunctive necessity de dicto and puts us in better stead for a correct, sober account of it. Regarding (2), I have tried to resolve the difficulty by outlining, in Chapter 6, a conception of meaning which is applicable to names, allowing us to preserve Kripke’s insight that names are rigid while also enjoying a straightforward solution to Frege’s puzzle.² Regarding (3), I have responded by developing the account of subjunctive necessity

---

¹ See Kripke (1980), pp. 110 - 113 and pp. 113 - 114.
² By ‘Frege’s puzzle’, in this context, I mean puzzle (1) from Section 6.6., ‘How can a true identity statement of the form “a is b” differ in meaning from the corresponding “a is a”’?’. (In the present philosophical context, this is regarded as a puzzle from the point of view of the question of the meaning of proper names, without any hypothesis about identity playing a role. In Frege (1892), on the other hand, the puzzle is framed from the point of view of the idea that identity is a relation between objects, holding between all object and themselves. I discuss the puzzle from that point of view in my (forthcoming).)
**de dicto** of Chapter 5, which makes it clear that semantic considerations do come into the picture, and in a quite particular way.

Regarding (4), I have tried to avoid contributing further towards marginalizing the indicative by being explicit and emphatic about the fact that I am focusing narrowly on *subjunctive* necessity *de dicto*. However, in focusing on subjunctive necessity I have not done anything to remedy the neglect which indicative necessity may have been suffering. In the remainder of this chapter, I want to offer some tentative suggestions about this topic, and its relation to the topics of apriority and analyticity. I think the notion of internal meaning sketched in Chapter 6 may have a key role to play here. This is, I think, one of the most exciting opportunities for further research connected with the topics of this thesis, along with the idea of flexible granularity (also introduced in Chapter 6).

### 7.2. Indicative Necessity, Apriority and Analyticity

Focusing as it does on the notion of *subjunctive* necessity *de dicto*, the body of this thesis leaves questions about *indicative* necessity open. Should we recognize a notion of indicative necessity *de dicto*? If so, how should we understand that notion?

I suspect that we can and should recognize a notion of indicative necessity *de dicto*, and that this may actually be a better candidate for what pre-Kripkean analytic philosophers such as Ayer and Carnap and their contemporaries, and perhaps a lot of their predecessors in the history of philosophy, meant by ‘necessary’ (at least some of the time). One of my inspirations in thinking this is Chalmers (1998) (a short unpublished paper used in giving a talk). Chalmers conducts his discussion in a two-dimensional, possible worlds framework. Abstracting away from this for present purposes, we may as a rough initial gloss say that an indicatively necessary truth *must actually be* the case, while a subjunctively necessary truth *could not have failed to be* the case had things gone differently.

When Chalmers considers ‘six reasons for favoring the subjunctive’, the first he considers is:

(a) Indicative necessity is "merely epistemic". (Chalmers (1998).)

The response he gives is as follows:

---

3 For example in Ayer (1936) and Carnap (1947), which we touched on in Chapter 2. In this connection, it is interesting to note that, early on in the first lecture of *Naming and Necessity*, as Kripke is about to clarify the notion of necessity, he says ‘The second concept which is in question is that of necessity. Sometimes this is used in an epistemological way and might then just mean *a priori*’ (Kripke (1980), p. 35). While this is encouraging in that it shows Kripke ready to allow that some philosophical uses of ‘necessary’ may be best understood as not trafficking in the notion of subjunctive necessity which Kripke is going to be reserving that word for, it is also somewhat puzzling if you stop and think about it. Why would the word ‘necessary’, with its close relationship to ‘must’, and talk of *all* possibilities, get used to mean the same as ‘*a priori*’, which is traditionally explained in terms of the possibility of being known without recourse to experience? On the suggestion of the present section, things become clearer: what Kripke is saying here is close to the truth but not strictly correct. The uses of ‘necessary’ he has in mind should be thought of as invoking *indicative* necessity, which property *underlies* apriority traditionally conceived, explaining why *a priori* propositions are knowable without recourse to experience.
Answer: So? Before 1970, almost everyone thought necessity was tied to the epistemic (cf. Pap's book). Kripke *argued* that necessity and epistemic notions came apart, by appeal to the subjunctive, but one can't simply presuppose it. (Chalmers (1998).)

I have long felt that this is a bullet that shouldn't be bitten, although I am not sure exactly what 'merely epistemic' and 'tied to the epistemic' mean here. Putting exegesis aside, I think that there are good prospects for an account of indicative necessity which is, in a certain sense, not epistemic. The type of account I have in mind is not epistemic in the sense that it does not appeal to any notion of knowledge or a knowing subject. One reason I think this may be a virtue is that it may allow for a certain kind of explanation of why a priori truths are knowable a priori, tying this epistemic property to an underlying semantic one.

I want to suggest that the notion of internal meaning sketched in Chapter 6 may offer us a promising way of thinking of indicative necessity de dicto. I suggest that we try thinking of it as consisting in, roughly, determination of truth-value by internal meaning. Or: the indicatively necessary propositions are those which are such that any proposition with the same internal meaning must have the same truth-value.5

This understanding of indicative necessity, which does not appeal to considerations of knowledge, knowability, or experience, can then be used to explain apriority conceived along traditional lines as knowability without recourse to experience. That some propositions can be known without recourse to experience (in a certain sense) is not just some brute fact, nor on this suggestion is it something which is to be explained by appealing to special facts about knowing subjects. Rather, what makes a proposition a priori has to do with the nature of the proposition itself, with the kind of meaning it has. I find this to be intuitively compelling; paradigmatically a priori propositions seem to be different in some fundamental way from empirical ones. And we can say that what distinguishes them is that their internal meanings determine their truth-values.

---

4 Chalmers is referring to Pap (1958).
5 One source of difficulty here, which I want to mention without trying to resolve, is a class of propositions which seem to satisfy the constraint that any proposition with the same internal meaning must have the same truth-value, but which we may not want to count as a priori. These are propositions whose concrete occurrence ensures their truth, such as 'I exist' and 'Language exists'. Unlike propositions such as '2 + 2 = 4' and 'All bachelors are unmarried' (assuming a simple definition of 'bachelor' as 'unmarried man'), propositions of this class seem to reach out into the world, so to speak, for their truth-values. (Note by the way that their being contingent doesn't seem to be the distinguishing feature here, contrary to what you may at first think: 'If Julius exists, Julius invented the zip', where 'Julius' is stipulated to refer to the inventor of the zip (if there is one), is contingent but seems to lack this character of having to reach out into the world for its truth-value.) Faced with these problem cases, we have various options, most saliently: (i) allow that they are both indicatively necessary and a priori, (ii) allow that they are indicatively necessary but deny that they are a priori (on this option you could still hold that all a priori propositions are indicatively necessary, and that this may help explain their apriority, but then you may be pressed to explain why not all indicative necessities count as a priori), (iii) try to define indicative necessity more carefully (perhaps using a notion of grounding or explanation) so as to rule out the problem cases. For discussion of these problem cases in connection with two-dimensional semantics, see Chalmers (2006, §2.4).
Does this mean I am saying that all \textit{a priori} propositions, such as those of mathematics, are analytic? It depends on what you mean by ‘analytic’.\textsuperscript{6} My terminological preference is to reserve ‘analytic’ for a notion narrower than that of indicative necessity, so that it comes out that ‘synthetic \textit{a priori}’ is a non-empty category. (More on this narrower notion in a moment.) Having said that, I do think we can use the notion of indicative necessity, understood along the lines of determination of truth-value by internal meaning, to get at something true and important which philosophers have sometimes captured by saying ‘mathematics is analytic’.

I do not think that the idea that \textit{a priori} propositions’ internal meanings determine their truth-values has ever been seriously threatened, apart from general attacks on the notion of meaning. Nevertheless, the idea that \textit{a priori} truths, including those of mathematics, are true in virtue of meaning in \textit{any sense} is nowadays curiously unfashionable. It is worth noting that this wasn’t the case for much of the twentieth century. And nor was the idea confined to any particular school of thought. You could be forgiven for thinking that any such idea is part and parcel of things like logical positivism, verificationist approaches to meaning, and crude critiques of metaphysics. But this is not the case - Gödel, for instance, who was against positivism, and who had Platonistic, metaphysical tendencies of thought, also held that mathematical truths were true in virtue of meaning. For instance, after arguing against conventionalism about mathematical truth in his Gibbs Lecture, Gödel says the following:

> However, it seems to me that nevertheless one ingredient of this wrong theory of mathematical truth [i.e. conventionalism] is perfectly correct and really discloses the true nature of mathematics. Namely, it is correct that a mathematical proposition says nothing about the physical or psychical reality existing in space and time, because it is true already owing to the meaning of the terms occurring in it, irrespectively of the world of real things. What is wrong, however, is that the meaning of the terms (that is, the concepts they denote) is asserted to be something man-made and consisting merely in semantical conventions. (Gödel (1951/1995), p. 320.)

Perhaps the diverse philosophers who thought things like this, which are currently so unfashionable, were on to something which we should try to recover. The ideas about propositions and meaning sketched in Chapter 6 of this thesis, augmented with the present suggestion about indicative necessity \textit{de dicto}, offer one avenue for doing this: use Putnam-inspired Twin-Earth-style considerations to distinguish internal and external meaning, taking care to distinguish the former from “narrow content” and the like, then apply middle-Wittgenstein-style ideas about meanings being roles in systems to the internal factor (this part is optional and could be replaced with something else), and then finally explicate the notion of indicative necessity along the lines of determination of truth-value by internal meaning and explain the \textit{a priori} in terms of that. This can then be shielded from misunderstandings by emphasizing things such as the linguistic division of labour\textsuperscript{7},

---

\textsuperscript{6} Relevant here is Gillian Russell’s (2008) book attempting to bring the philosophical discussion of analyticity up to date. One of the themes of her book, which agrees in a broad way with a lot of what I am saying, is that different notions of meaning should be distinguished, and these may give different results when plugged in to yield a notion of analyticity.

\textsuperscript{7} See Putnam (1975).
Wittgensteinian considerations about meaning not being a queer mental process, Kripkensteinian considerations about meaning outrunning actual occurrences, and the like.

Finally, what about the notion I alluded to above which is narrower than that of indicative necessity de dicto, and for which I favour reserving the term 'analytic'? I think there are prospects for an interesting notion of analyticity which is conceptually and extensionally distinct from the notion of indicative necessity de dicto, and which captures some of the spirit of Kant's account of the analytic-synthetic distinction in the Critique of Pure Reason.

It is well known that Kant's definition, or principal explication, of 'analytic' and 'synthetic' is given in terms of subject and predicate:

In all judgments wherein the relation of a subject to the predicate is cogitated (I mention affirmative judgments only here; the application to negative will be very easy), this relation is possible in two different ways. Either the predicate B belongs to the subject A, as somewhat which is contained (though covertly) in the conception A; or the predicate B lies completely out of the conception A, although it stands in connection with it. In the first instance, I term the judgment analytical, in the second, synthetical. (Kant (1781), Introduction, IV.)

Since modern logic and philosophy of language have taught us not to regard every proposition as being composed of a subject and a predicate, this definition can't be adequate for us. But it is suggestive, as are some of the other things Kant says about the analytic-synthetic distinction. He also says of analytic and synthetic propositions that 'the former may be called explicative, the latter augmentative' (Kant (1781), Introduction, IV). And consider this elaborated version he gives of his main question, of how synthetic a priori knowledge is possible:

If I go out of and beyond the conception A, in order to recognize another B as connected with it, what foundation have I to rest on, whereby to render the synthesis possible? (Kant (1781), Introduction, IV.)

The idea that synthetical judgments are 'augmentative', that they 'go out of and beyond' 'conceptions', can, I think, be generalized or abstracted from Kant's discussion in such a way that it does not depend on construing all propositions as being of the subject-predicate form. And we get a hint of how to do this from the following passage about the syntheticality of the proposition '7 + 5 = 12':

We might, indeed, at first suppose that the proposition 7 + 5 = 12 is a merely analytical proposition, following (according to the principle of contradiction) from the conception of a sum of seven and five. But if we regard it more narrowly [my emphasis], we find that our conception of the sum of seven and five contains nothing more than the uniting of both sums into one, whereby it cannot at all be cogitated what this single number is which embraces both. The conception of twelve is by no

---

8 See Wittgenstein (1953).
9 See Kripke (1982).
means obtained by merely cogitating the union of seven and five; and we may
analyse our conception of such a possible sum as long as we will, still we shall never
discover in it the notion of twelve. We must go beyond these conceptions, [...] (Kant
(1781), Introduction, V.)

This regarding-more-narrowly is key. I suggested above that a proposition is a priori iff it is
indicatively necessary de dicto - this being conceived as an explanation, rather than a
conceptual analysis, of apriority. And I suggested that a proposition is indicatively necessary
de dicto iff its internal meaning determines its truth-value - this being conceived as an
account of the very notion of indicative necessity de dicto. The idea now is that a proposition
is analytic iff its internal meaning regarded more narrowly in a certain way - or iff a certain
aspect or part of its internal meaning - determines its truth-value. And so the next task would
be to characterize the aspect or part of internal meaning we are restricting our attention to
here. Perhaps we should draw on a notion of linguistic competence for this.

We could make such an account handy and memorable by calling the relevant aspect or part
of an expression’s internal meaning its ‘meaning-radical’, and then saying:

A proposition is analytic iff its meaning-radical determines its truth-value.

Now, we might want to say, with Kant, that arithmetical propositions - at least once you get
beyond the very most basic and trivial ones - are synthetic a priori.

Consider in this connection the fact that we can come to believe false arithmetical
propositions - for example on the basis of miscalculation, or misremembering, or false
testimony - and that we can apply them. Contrast the case of a paradigmatic analytic
proposition such as 'All bachelors are unmarried'. (Let us just suppose that 'bachelor' simply
means 'unmarried man'.) To be sure, someone can assent to the sentence 'Not all bachelors
are unmarried', and dissent from 'All bachelors are unmarried', but in such a case we would
say that they don't understand this latter as we do - for them it is not an instance of our
proposition 'All bachelors are unmarried'.

As we saw above, Kant says that we can become 'more clearly convinced' of the synthetically
of arithmetical propositions 'by trying large numbers'. Let us now, therefore, try to illustrate
the notion of a meaning-radical by considering an example of a false arithmetical proposition
involving numbers larger than 7, 5 and 12. Say '25 x 25 = 600'.

Despite being false a priori, the proposition '25 x 25 = 600' is something we can mistakenly
believe and apply, while understanding it correctly (in some suitably minimal sense of
'understand'). We have - wrongly - made a connection between our conception of the
product of 25 and 25, and our concept of 600.

All a priori propositions, then, on the account I am suggesting, will be such that their internal
meanings determine their truth-values. Analytic propositions have the further property that
their meaning-radicals determine their truth-values. As I have indicated, how we should think
of meaning-radicals is an issue for further research.
Chapter 7 References


