Chapter 2: Varieties of Scrutability

David J. Chalmers

1 Scrutability Theses

In this chapter I formulate and discuss a number of specific scrutability theses in more detail than I have so far. Scrutability theses can take a number of different forms. Some are restricted scrutability theses: one such thesis holds that all mental truths are scrutable from physical truths, for example. Most of the scrutability theses I am interested in will be unrestricted theses, however, saying that all truths are scrutable in some way.

The general form of an unrestricted scrutability thesis is: All truths are scrutable from base truths. This leaves three things to be clarified further: “truths”, “scrutable from”, and “base”. All three elements correspond to important dimensions of variation.

What are “truths”: true propositions, true sentences, true beliefs, or something else? In the introduction, scrutability theses were formulated in terms of propositions. In chapter 1, they were formulated in terms of sentences. I discuss this issue in section 2 and in much more detail in the third excursus, following this chapter.

What are “base truths”? These might be a class of truths, such as the class of physical truths, phenomenal truths, or fundamental truths. If so, a scrutability thesis will say that all truths are scrutable from that class. Alternatively, “base truths” might stand for a type of class of truths, such as the type: compact classes of truths. If so, a scrutability thesis will say that all truths are scrutable from some class of truths of that type. One important constraint on base truths is compactness (as characterized in section 4 of chapter 1), but numerous other candidates will be discussed. I characterize an initial class of base truths in the first section of chapter 3, and focus on minimal classes of base truths in chapters 7 and 8.

What is “scrutable from”? We have already encountered numerous scrutability relations: inferential scrutability, conditional scrutability, a priori scrutability, definitional scrutability, and so on. I will adopt the convention of using lower case for these scrutability relations (“inferential...
scrutability”), while using upper case for corresponding scrutability theses (“Inferential Scrutability”).

In principle one can mix and match different values for on each of these three dimensions (scrutability relations, base elements, sentences or propositions), yielding many different scrutability theses such as A Priori Fundamental Sentential Scrutability, Conditional Physical Propositional Scrutability, and so on. My convention will be that the default values are A Priori, Compact, and Sentential. So Scrutability (simpliciter) says that all true sentences are a priori scrutable from a compact class of true sentences. Fundamental Scrutability says that all true sentences are a priori scrutable from the class of fundamental true sentences. Conditional Propositional Scrutability says that all true propositions are conditionally scrutable from a compact class of true propositions. And so on.

Where scrutability relations are concerned, the three most important are those outlined in the introduction: inferential, conditional, and a priori scrutability. There I laid out corresponding theses in propositional form, but for present purposes it is useful to define the relations themselves and to do so in sentential form. Put this way, they will be relations between a sentence \( S \) and a class of sentences \( C \) for a subject \( s \).\(^1\) The basic characterization of these three notions is as follows.

\( S \) is inferentially scrutable from \( C \) for \( s \) iff \( s \) came to know \( C \), \( s \) would be in a position to know \( S \).

\( S \) is conditionally scrutable from \( C \) for \( s \) iff \( s \) is in a position to know that if \( C \), then \( S \).

\( S \) is a priori scrutable from \( C \) for \( s \) iff \( s \) is in a position to know a priori that if \( C \), then \( S \).

Here, “know \( C \)” is shorthand for “know all the sentences in \( C \)”, and “if \( C \)” is shorthand for “if the sentences in \( C \) are true”, where what it is to know a sentence (or to know a sentence to be true) is discussed in the next section. In all three cases, “in a position to know” should be understood as involving an idealization, with the rough idea being that the subject is in a position to know these

\(^1\)Strictly speaking, scrutability is always relative to a time and a world as well as to a subject, for example because of differences in what subjects are in a position to know at different times and in different worlds. So we should say that \( S \) is scrutable from \( C \) for a subject \( s \) at time \( t \) in world \( w \). To avoid clutter, I will usually mention only \( s \) and leave the relativization to time and world implicit. Where context-dependent sentences are concerned, we can also relativize scrutability to a context in order to reflect differences in which the sentences express in different contexts, as discussed later.
things through sufficiently ideal reasoning, if they were capable of that reasoning. I flesh out these three notions in the sections 3, 4, and 5, and the nature of the idealization in section 7.

Of the three relations, a priori scrutability theses is ultimately the most important for our purposes. Conditional scrutability plays an important role in arguing for a priori scrutability theses, however, and also provides a useful fallback thesis that is available even to those who are skeptical about the a priori. Inferential scrutability plays a less essential role, and is also more problematic in some respects, but I have included discussion of it both because it can help to motivate the other theses and because its problems are interesting in their own right.

For each of these scrutability relations, there is an analog conclusive scrutability relation concerning not knowledge but certainty. The conclusive relations can be defined by replacing “know” in the definitions above by “know with certainty”. For example, conclusive inferential scrutability requires that if \( s \) knew \( C \) with certainty, \( s \) would know \( S \) with certainty, while conclusive conditional scrutability requires that \( s \) is in a position to know with certainty that if \( C \), then \( S \).

Knowledge with certainty is a particularly strong sort of knowledge (I discuss the notion further at the start of chapter 4): intuitively, knowing \( S \) with certainty requires absolutely eliminating all hypotheses on which \( S \) is false. It is plausible that we cannot be certain of ordinary empirical claims, such as ‘There is a table in this room’, but it is at least arguable that we can be certain of some claims: perhaps ‘\( 2+2=4 \)’ and ‘I am conscious’.

There will then be conclusive variants of the scrutability theses we are concerned with: Conclusive A Priori Scrutability, for example. In some cases (certainly for Conditional Scrutability, and arguably for the others) the conclusive variant will be stronger than the original thesis. But many of the reasons for believing the original theses are also reasons for believing the conclusive theses (it is notable that Laplace himself talked of certainty), and the conclusive theses are useful for a number of explanatory purposes. So I will keep an eye on both the original theses and the conclusive theses in what follows.

In what follows, I discuss the objects of scrutability in section 2, the three main scrutability relations in 3 through 5, generalized scrutability in section 6, and issues about idealization in section 7. It is quite possible to skip the remainder of this chapter on a first reading. It goes over a number of separate foundational issues regarding the formulation of scrutability theses that need to be addressed and that are interesting in their own right but that are not essential to following the later discussion. For many readers, it may work better to look briefly at the following section and then move directly to Chapter 3, coming back to this chapter when necessary.
2 Sentences or Propositions?

Many different sorts of things are sometimes said to be truths: true propositions, true beliefs, true sentences, true utterances. Which of these is most relevant to scrutability theses?

Truths are most commonly understood as true propositions, where propositions are entities that are independent of any particular language, and that are the things that we assert and that we believe. My own view is that if propositions and scrutability are understood correctly, then all true propositions are indeed scrutable from a compact class of propositions. However, the nature of propositions is strongly contested, and different theories of propositions will have quite different results for scrutability.

On a possible-worlds view of propositions, the proposition expressed by a sentence is the set of possible worlds where the sentence is true. On this view, all necessary truths express the same proposition (the set of all worlds), a proposition that is itself knowable a priori. If so, then if it is necessary that water is H$_2$O, it follows that the proposition that water is H$_2$O is itself knowable a priori.

On a Russellian view of propositions, the proposition expressed by a sentence is a structure involving those objects and properties that are the extensions of parts of the sentence. On this view, not all necessary truths express the same proposition, but the proposition that Hesperus is a planet and the proposition that Phosphorus is a planet are identical.

On a Fregean view of propositions, the proposition expressed by a sentence is a structure of senses expressed by parts of a sentence, where senses are fine-grained entities reflecting the epistemic and cognitive significance of various expressions. On this view, not all necessary truths express the same proposition, and the proposition that Hesperus is a planet and the proposition that Phosphorus is a planet are distinct.

On an eliminative view of propositions, there are no propositions at all. There are only sentences and utterances, and perhaps acts of thinking and states of believing. But sentences do not express propositions, and thinking and believing do not involve relationships to propositions.

It is clear that if we antecedently assume one of these views of propositions, there will be very different results for a scrutability thesis cast in terms of propositions. On the possible-worlds view, all necessary truths will automatically be a priori scrutable from any basis. On the Russellian view, necessary truths will not usually be a priori scrutable from arbitrary bases. But some, such as the proposition that Hesperus is Phosphorus (if they exist), will arguably be scrutable from any basis, and there will be no epistemological differences between propositions expressed by pairs
of sentences involving ‘Hesperus’ and ‘Phosphorus’ respectively. On a Fregean view, necessary truths will not automatically be scrutable, and epistemological differences between sentences involving ‘Hesperus’ and ‘Phosphorus’ will be preserved. On an eliminative view, the thesis that all propositions are scrutable will be vacuously true.

My own purposes include the analysis of fine-grained epistemological properties of sentences and thoughts. For this purpose, a Fregean view of propositions is the most promising. But I cannot simply assume such a view at the outset. It is controversial whether there are Fregean propositions, and even among sympathizers, it is controversial just what sort of thing they might be and how they behave. Further, one of my purposes is to use scrutability theses to provide support for a Fregean view of propositions. If I were to assume such a view at the outset, there would be some circularity here. Perhaps some support for the view would accrue from demonstrating coherence and power in the resulting picture, but a flavor of assuming the conclusion and of begging the question against opponents would remain.

At the same time, it will not do to assume one of the other views of propositions. And it does not make sense to cast things in terms of propositions but stay neutral between these views, as the views yield very different results in evaluating scrutability theses. One can certainly cast scrutability theses in terms of all three notions of propositions, but scrutability bases for each of the three might look very different.

So I will set aside propositions for now. That being said: if one accepts Fregean propositions, or if one is at least prepared to allow that propositions are sufficiently fine-grained to reflect differences in cognitive significance as well as differences in reference, then one might well construe scrutability theses in terms of propositions. That way, many of the added complexities that come from appealing to sentences can be dispensed with, and much of what I say later will still apply. So those who accept fine-grained Fregean propositions should feel free to translate what I say into propositional terms (though a couple of obstacles to a perfect translation are noted in the third excursus).

It is not out of the question to cast a scrutability thesis in terms of mental states such as thoughts

\[\text{My own view is that the arguments in the next two chapters in principle can make the case for scrutability of all truths (or at least all truths expressible by the subject) from the relevant base even if Russellian or possible-worlds propositions are involved. But I do not expect manyRussellians to agree (see the tenth excursus for more on this); and while possible-worlds theorists might agree, they might also take the conclusion to be a weak one. And in practice, casting scrutability theses in terms of propositions of this sort, or in a neutral way, would lead to too many complications arising from disagreements among theorists about how to treat the epistemology of these propositions.}\]
or beliefs, holding for example that for any belief one might entertain, the truth or falsity of the thought could be inferred from a certain class of (potential) basic beliefs. But the individuation of thoughts and beliefs is also nontrivial, and it is more awkward to speak of mental items than of linguistic items, so I will set these aside too, although later I will give some role to thoughts in interpreting certain scrutability claims.

Instead, I will take the truths at issue to be linguistic items such as sentences: both sentence types (abstract sentences such as ‘The cat sat on the mat’) and sentence tokens (sentences as uttered on a given occasion). It is somewhat awkward to speak of knowing sentences, and this manner of speech is certain less familiar than talk of knowing propositions or facts. But it is not hard to motivate belief-like and knowledge-like relations between speakers and linguistic items, not least through considerations about sincere assertion, knowledgeable assertion, and so on. Doing so brings out certain dialectical advantages of proceeding this way, compared to proceeding via propositions.

To see this, consider the Russellian about propositions who holds that the proposition that Hesperus is Hesperus and the proposition that Hesperus is Phosphorus are identical, so that there can be no psychological or epistemological differences between these propositions. Even if the Russellian is right, the sentences ‘Hesperus is Hesperus’ and ‘Hesperus is Phosphorus’ are not identical, and likewise utterances of these sentences are quite distinct. So it is at least open to the Russellian to associate different psychological and epistemological properties with these sentences. Furthermore, there are very good reasons for the Russellian to do so.

One way to bring this out is through what we might call the argument from assertion against Russellian views of propositions. Suppose that Sue knows that the morning star is a planet but believes that the evening star is not a planet. Like others in her community, she associates ‘Hesperus’ with the evening star and ‘Phosphorus’ with the morning star. Intending to deceive John, she says ‘Hesperus is a planet’. We can then argue as follows:

1. Sue’s assertion is not sincere.
2. An assertion is sincere iff the speaker believes the proposition asserted.
3. Sue asserts the proposition that Hesperus is a planet.
4. If the Russellian view is correct, the proposition that Hesperus is a planet is identical to the proposition that Phosphorus is a planet.
5. Sue believes the proposition that Phosphorus is a planet.
6. The Russellian view is false.

This argument has some force, but it is not a knockdown argument against the Russellian view. Premise 1 cannot be plausibly rejected: it is clear that insofar that there is a sincerity norm on assertion, for example, Sue’s utterance violates it. A standard Russellian view is committed to premises 3-5. A typical Russellian will reject premise 2, giving an alternative account of sincere assertion. They might hold, for example, that an assertion is sincere if the speaker believes the proposition asserted under the guise under which it was asserted, or if the speaker believes that the assertion expresses a true proposition. But this is enough to make the point. A reconstrued version of the argument, with premise 2 omitted, leads to the conditional conclusion that if the Russellian view is correct, sincere assertion comes apart from belief in the proposition asserted. That is, to save the data about sincere assertion, the Russellian needs a way to associate cognitive properties (such as belief) with utterances so that utterances that express the same proposition can nevertheless have different cognitive properties.

One can make a similar point using notions such as knowledgeable assertion and justified assertion. It is clear that Sue’s assertion is not a knowledgeable assertion. It is also clear that Sue’s assertion is not a justified assertion. But on a standard Russellian view, Sue knows the proposition she asserts, and she is justified in believing the proposition that she asserts. So standard Russelians need to distinguish a knowledgeable assertion from assertion of a proposition that the speaker knows. They also need to distinguish making a justified assertion from assertion of a proposition which the speaker is justified in believing.

We need some language to distinguish these notions. I will say that when an assertion of a sentence $S$ is a sincere assertion (or a believed assertion), the speaker believes $S$. When an assertion of a sentence $S$ is a knowledgeable assertion (or a believed assertion, as in the previous

---

3 A nonstandard Russellian view suggested by Soames (2002) holds that in cases like these, speakers assert propositions that are not semantically expressed by the asserted sentence. This view will deny premise 3, but the upshot will be the same as in the text. Making a sincere assertion will come apart from believing the proposition semantically expressed, and utterances that semantically express the same proposition can nevertheless have different epistemic properties.

4 The case of sincerity is perhaps more complicated than the case of knowledgeable and justified assertion, as there are arguably other cases in which one might tease sincerity apart from believing the proposition asserted. For example, if Sue says something that she believes to be true but that she knows will mislead John, then arguably her assertion is not sincere. To avoid these complications one could use the somewhat more constrained (if somewhat less familiar) notion of a believed assertion instead.
footnote), the speaker knows $S$. When an assertion of a sentence $S$ is a justified assertion, the speaker is justified in believing $S$. The same goes for other properties such as being a justifiable assertion, being an a priori justifiable assertion, and so on. When $S$ is a context-dependent sentence type (‘He is ready’), it is best to relativize knowledge to context: so the speaker knows or believes $S$ relative to the current context, or knows or believes the current token of $S$.

This analysis is developed further in the third excursus. The key idea there is to analyze the epistemic status of a sentence $S$ in terms of epistemic properties of mental states that $S$ expresses or is apt to express. For example, one knows $S$ when one has knowledge that is apt to be expressed by $S$. Likewise, one knows $S$ a priori when one has a priori knowledge that is apt to be expressed by $S$, and one believes $S$ when one has a belief that is apt to be expressed by $S$. This allows the notion of knowing a sentence to be extended to the case in which the sentence is not uttered. The language of believing and knowing sentences is somewhat nonstandard, but it provides an efficient way to capture the association of doxastic and epistemic properties with sentences and assertions, and not just with propositions.

This account of knowing a sentence makes no appeal to the notion of knowing a proposition. Given a positive theory of propositions, there will presumably be some connection between knowing a sentence and knowing a proposition, as any such theory needs some way to capture the data about sincerity, knowledgeable assertion, and so on. But different theorists of propositions will make the connection in different ways. It is open to a Fregean to simply hold that a speaker knows $S$ iff they know the proposition expressed by $S$.Russellians (such as Salmon 1986) who allow that propositions are presented under guises might say that the speaker knows a sentence $S$ when they know the proposition expressed by $S$ under the guise associated with $S$. Other Russellians might say that the speaker knows $S$ when they know some ancillary proposition that is not semantically expressed by $S$ but is otherwise associated with $S$.

It might be suggested that one knows $S$ iff one knows that $S$ is true. On a literal reading of “knows that $S$ is true”, this involves a certain sort of metalinguistic knowledge about $S$, in which case the equation between the two is not plausible. It seems clear that Sue might knowledgeably assert ‘Phosphorus is a planet’ in a case such as the above without having any metalinguistic beliefs.

\[5\] If associated propositions or guises can vary between utterances of a sentence, then one can relativize these connections to contexts: one knows $S$ in a context if one knows the proposition associated with $S$ in that context, perhaps under the guise associated with $S$ in that context. Alternatively, one can make the connection at the level of sentence tokens: one knows a sentence token $S$ if one knows the proposition expressed by $S$, perhaps under the guise associated with $S$. 

8
about her assertion: all that is required is the right sort of astronomical beliefs. Still, there is a weaker reading of ‘knows that $S$ is true’ that does not require any metalinguistic knowledge, and that comes to a simple rephrasal of the claim that one knows $S$ in the sense from two paragraphs ago. I will sometimes use this locution, as it is more idiomatic and convenient for some purposes than the alternative. It should be understood, though, that to know that $S$ is true is not to have metalinguistic knowledge: it is just to know $S$ in the sense above. On this usage, to know that ‘bachelors are unmarried’ is true is just to know ‘bachelors are unmarried’, which as characterized above is more akin to knowing that bachelors are unmarried than to knowing something about the word ‘bachelor’.

In the third excursus, following this chapter, I discuss issues about sentential and propositional scrutability in much more detail. I spell out an account of knowledge of sentences in terms of thoughts; I discuss complications arising from context-dependence; I go into the precise formulation of sentential scrutability theses; and I analyze the relationship between sentential scrutability theses and propositional scrutability theses. Those who are interested in those issues should feel free to skip to the excursus now. Those details become complex, however, and they are not crucial to following the remainder of this chapter, so other readers should feel free to proceed to further issues.

In the remainder of this chapter, I discuss the formulation and interpretation of empirical, conditional, and a priori scrutability theses. Most of these issues apply to both propositions and sentences, but I will typically assume a formulation in terms of sentences.

3 Inferential Scrutability

As defined in section 1, $S$ is inferentially scrutable from $C$ for $s$ iff, if $s$ came to know $C$, $s$ would be in a position to know $S$. Inferential scrutability differs structurally from the other two sorts of scrutability. Conditional and a priori scrutability involve knowledge of conditionals, while inferential scrutability involves knowledge of unconditional sentences, given knowledge of other sentences. As we saw in the introduction, this distinctive form raises distinctive problems for

Note to the reader: One way to simplify the path through this book is to ignore all material on inferential scrutability (mainly in this chapter and the next). This material is interesting in its own right, especially if one is interested in Fitch-style issues, and it plays a useful minor role in motivating Conditional and A Priori Scrutability, but it does not play an essential role. Conditional scrutability plays a more essential role in supporting A Priori Scrutability in the arguments of the following chapters.
inferential scrutability theses that do not apply to the other two. These problems will be the main subject matter of this section.

Typically, inferential scrutability will involve knowledge by inference: if \( s \) came to know \( C \), \( s \) would be in a position to know \( S \) by inference from \( C \). The definition does not require knowledge by inference, however, so it can be satisfied in cases where \( S \) is in a position to know \( C \) some other way: perhaps \( s \) knew \( S \) already, or perhaps \( S \) is a theorem that \( s \) can prove independently of \( C \). It is arguable that in any such case, \( s \) is also in a position to infer \( S \) from \( C \) (perhaps \( s \) could start with \( C \), conjoin \( S \) after recalling or proving it, and then drop \( C \)). But even if this does not count as inference, we can simply note that inferential scrutability does not require inferrability. The label “inferential” is imperfect, but it captures the idea that transitions from knowledge states to knowledge states (rather than knowledge of conditionals) is what is central here.

The Inferential Scrutability thesis says: there is a compact class of truths \( C \) such that for all subjects, all truths are inferentially scrutable from truths in \( C \). That is: there is a compact class \( C \) of sentences such that for all subjects and all truths \( S \), if the subject were to come to know the truths in \( C \), they would be in a position to know \( S \).

Here the stipulation that the subject knows the truths in \( C \) is best understood as a stipulation that the subject knows the truths in \( C \) and continues to know them, insofar as this is possible. Without this stipulation, some subjects who are on the point of inferring \( S \) from \( C \) might respond by rejecting \( C \) instead, perhaps because they disbelieve \( S \). But for our purposes we are really interested in the consequences of knowledge of \( C \) where this knowledge is held fixed. Alternatively, we can require in the antecedent that the subject knows \( C \) with certainty. In this case, it is arguable that the subject should never respond by rejecting \( C \) (given that they are not certain of \( \neg S \), which they will not be if they are rational and \( S \) is true).

A subject \( s \) at time \( t \) is in a position to know \( S \) iff it is possible that \( s \) comes to know \( S \) at some later time \( t' \), starting from \( s \)'s position at \( t \) and without acquiring any further empirical information.\(^8\) The process of coming to know \( S \) need not involve a priori reasoning alone, as it can use any empirical beliefs that the subject has at the time of utterance, but it cannot involve new empirical discovery (based on perception, testimony, and so on) after this point. There are

\(^7\)That is, the first two can be represented as \( PK(C \rightarrow S) \) and \( PKA(C \rightarrow S) \), while inferential scrutability can be represented as \( K(C) \Rightarrow PK(S) \). Here \( K, PK, \) and \( PKA \) involve “knows”, “is in a position to know” and “is in a position to know a priori” respectively. “\( \rightarrow \)” is an indicative (or perhaps material) conditional connective, and \( \Rightarrow \) is a counterfactual connective.

\(^8\)In previous versions of this work I used “empirical scrutability”, because of the empiricist flavor discussed below, but that name misleadingly suggests a parallel to a priori scrutability with a defining role for a posteriori knowledge.
various notions of possibility, but for our purposes it is most natural to appeal to metaphysical possibility here. This makes for a significant idealization, allowing possible futures in which the subject manifests reasoning capacities that are not present at \( t \), but such an idealization is useful for our purposes in any case.

One can illustrate the thesis by first selecting a scrutability base. To anticipate the next chapter (where more details can be found in section 1), let us say that \( PQTI \) includes all truths in the following vocabulary: (i) microphysical expressions (in the language of a final physics); (ii) macrophysical expressions (in the language of classical physics), (iii) expressions for arbitrary types of conscious experience; (iv) expressions for lawhood, causation, and counterfactual dependence, (v) logical and mathematical expressions, and (vi) the indexicals ‘I’ and ‘now’. \( PQTI \) will also contain a that’s-all truth saying that the world is a minimal world in which the previous truths obtain. Not much will depend on this choice of basis (and one could add expressions for secondary qualities and normative expressions if one likes, along with others). But it is supposed to be a reasonably generous basis, so that there is at least a chance that many or all truths will be scrutable from it.

It seems clear that if a subject knew all truths in \( PQTI \), then they would thereby be in a position to know many more truths about the world. And it does not seem immediately out of the question (at least modulo the problems discussed below) that for any truth \( T \), they would be in a position to know \( T \). If so, all truths would be inferentially scrutable from \( PQTI \).

The inferential scrutability thesis is useful for our purposes in part because it avoids technical notions such as apriority and analyticity, and in part because it lacks the rationalist flavor of some other scrutability theses. If anything, this thesis has an empiricist flavor, saying that one could come to know a great many things by knowing certain (largely empirical) truths. So it provides a good entry point to scrutability theses for one who is skeptical of rationalism and the a priori.

As stated, the thesis requires considering a scenario in which the subject comes to know all truths in a compact class of truths \( C \) (I will henceforth put this by saying that the subject knows \( C \), which may be a huge class of truths about all of spacetime. Of course to know all truths in such a class, the subject would need to have cognitive capacities greater than any actual human...

---

8Bringing in worlds: A subject \( s \) at time \( t \) in world \( w \) is in a position to know \( S \) if there is a world \( w' \) (possible relative to \( w \)) and a time \( t' \) (later than \( t \)) such that \( s \) is in the same position at \( t \) in both \( w \) and \( w' \) and such that \( s \) comes to know \( S \) at \( t' \) in \( w' \). I will typically leave the world-relativization tacit, however, with the note that “starting from their position at \( r' \) always invokes the position in the original world. I will not try to analyze the notion of position, but it will probably suffice for \( w \) and \( w' \) to be duplicates at all times up to \( t \).
subject. So as before, we need to idealize to make sense of this scenario.

There is a worse problem than idealization, however. For the inferential scrutability thesis to be plausible, C will have to be sufficiently encompassing that its truths are jointly true of this world and this world alone. But then, assuming that no-one actually knows C, it will be impossible to know C. Any world in which someone knows C would have to differ from the actual world, and therefore must be a world in which C is false. But there are no worlds in which someone knows C and C is false. So no-one can know C. If so, then the Inferential Scrutability thesis is vacuously true (if counterfactuals with impossible antecedents are vacuously true), or at least hard to assess (if not).

This problem is closely related to Fitch’s paradox of knowability, discussed in the first excursus: if there is a truth S that no-one ever knows, there is a truth S₁ = ‘S and no-one knows S’ that no-one can possibly know (if someone knew S₁, they would know S, so S₁ would be false). Fitch’s paradox itself poses a closely related problem for the Inferential Scrutability thesis. On the face of it, Inferential Scrutability entails that all truths can be known, at least if C can be known. But we know that not all truths can be known. So either Inferential Scrutability is false or C cannot be known. Of course the paragraph above suggests that C cannot be known. But then we are still left with the dilemma at the end.

The remainder of this section is devoted to a number of possible responses to this Fitch-style problem. First, one might try construing “being in a position to know” in terms of warrant rather than in terms of possible knowledge, as in the fourth excursus. On such a construal, inferential scrutability of S from C requires that knowing C would provide a warrant for believing S. This might arguably help with the case of the Fitch sentence S₁: it is argued that if one knew all the truths in C, one would have warrant for each conjunct of S₁ and therefore for S₁ itself, even though one could not use this warrant to know S₁. But it does not obviously help with the worry that it is impossible to know all the truths in C. At this point one could try allowing that counterfactuals with impossible antecedents are not always vacuously true. Then one might allow (nontrivially) that if, per impossibile, one came to know all truths in C, then one would have warrant for sentences such as S₁ above. This strategy is not uninteresting, but I will set it aside in what follows.

Second, one could address the problem concerning the knowability of C by weakening the notion of inferential scrutability. We can say that a truth S is inferentially scrutable from a class C of truths if there is a knowable subclass of C such that if the subject were to come to know the truths in that subclass, they would be in a position to know S. In most cases, the required subclass will be much smaller than C itself, thus both avoiding the need for an overly strong idealization
and avoiding the problem arising from the unknowability of $C$.

I will adopt this useful amendment in what follows, as it at least allows that some truths are clearly inferentially scrutable. But there remain obvious problems with the corresponding Inferential Scrutability thesis. One worry is that it weakens the thesis considerably. For many truths $S$, there will be some $C$-truths such that knowing these truths justifies belief in $S$, but such that knowing further $C$-truths would justify rejecting $S$. For this reason one will have to be careful in making inferences from these weakened inferential scrutability claims to other scrutability claims.

Another problem for the amended thesis is that the Fitch sentence $S_1$ yields a counterexample to the thesis, at least on a modal understanding of being in a position to know. Because $S_1$ is unknowable, no knowable subclass of $C$ can be such that knowing $C$ would put a subject in a position to know $S$. There are related problem cases: say that $S_2$ is ‘I know no sentences in $C$’, uttered by a subject who has beliefs about some sentences in $C$ but does not know them, and is not now in a position to know that she does not know them. Then $S_2$ is true, but the subject is not in a position to know it by coming to know any subclass of sentences of $C$.

A third response stems from the observation that in all these cases, it is not ruled out that one can come to know whether the sentence $S$ is true (where knowledge of whether $S$ is true is the natural generalization of knowledge that $S$ is true as explained above). It is just that the very process of coming to know whether $S$ is true (by the procedure of coming to know $C$-truths) will render $S$ false.

We might call truths like this Fitchian truths, because Fitch’s unknowable truth ‘$P$ and I don’t know $P$’ is a paradigm. We might define a Fitchian truth as an alethically fragile truth: a truth $S$ such that properly investigating the truth-value of $S$ will change the truth-value of $S$. Here to investigate the truth-value of $S$ is to investigate whether $S$ is true, and to do so properly is, roughly, to do so as well as could be done. Of course there are different methods by which one might come to know the truth-value of $S$, so one might also say that $S$ is Fitchian with respect to a method iff properly investigating the truth-value of $S$ by that method will render $S$ false. Then $S_1$ above is Fitchian with respect to any method, as is the conjunction of all truths in $C$, while $S_2$ is Fitchian.

Another worry concerns the scrutability of negative truths, which as discussed in the next chapter may require the use of a ‘that’s-all’ sentence. On some formulations of the that’s-all sentence, knowing it will entail knowing all the other sentences in the class. Then the considerations here suggest that the that’s-all sentence will be unknowable, but then remaining subclasses will not allow scrutability of negative truths. To avoid this problem one would have to use other formulations of the that’s-all sentence or restrict scrutability theses to positive truths, as in the next chapter.

Thanks to Wolfgang Schwarz for discussion here.
with respect to the method of determining truth-value via knowledge of sentences in $C$.

One might then suggest a modified thesis, saying that there is some subclass of $C$ such that if one were to come to know this subclass, one would be in a position to know whether $S$ is true. Or better, one can suggest that there is some subclass such that if one were to come to know whether the sentences in this subclass are true, one would be in a position to know whether $S$ is true. This latter formulation allows for the possibility that the process of coming to know the sentences may also change the truth-value of sentences in $C$, as well as the truth-value of $S$.

This modified thesis is not threatened by Fitchian truths or by any of the cases above. There are some smaller residual worries. One worry is that in cases where one cannot know that $S$ is true without knowing whether a very large subclass of $C$-sentences is true, worlds where one knows that subclass may be so different from our world (in the cognitive capacity they allow, for example) that they are nomologically impossible. In such a world, knowing whether $S$ is true might require knowing about alien features of that world not described by $C$-sentences. If so, then no knowledge of $C$-sentences will put one in a position to know whether $S$ is true. It is unclear whether this scenario can arise. But a milder version of the worry applies more generally. For any truths $S$ such that one can know the truth-value of $S$ only by knowing some (actual) truths in $C$ to be false, then the inferential scrutability relation between $S$ and $C$ does not tell us directly about the status and grounds of $S$ in the actual world. Correspondingly, one will not be able to use inferential scrutability to argue directly for the scrutability of the truth of $S$ from the truth of $C$ in the actual world. So the force of the scrutability thesis is weakened somewhat. Still, this modified inferential scrutability thesis remains interesting and important.

A fourth strategy is simply to exclude Fitchian cases, and require only that all non-Fitchian truths (with respect to the method of inferential scrutability from $C$-truths) are inferentially scrutable from $C$-truths. (For some purposes one will need to modify this thesis somewhat further, for reasons I will discuss in the next chapter.) This strategy loses the universal scope of the thesis, of course, and it may appear somewhat ad hoc. Nevertheless, this strategy can help in supporting other scrutability theses. In the next chapter, I will argue that all non-Fitchian truths (of a certain sort) are inferentially scrutable. If inferential scrutability entails conditional scrutability, and if conditional scrutability is not subject to worries about Fitchian cases, then there is at least a reasonable prima facie case that all truths (of the relevant sort) are conditionally scrutable.

A fifth fix, and perhaps the most natural, is to move to conditional scrutability, below.

These imperfections in the Inferential Scrutability thesis will not matter too much for my purposes. For these purposes, the thesis is mainly of instrumental value. It is mainly valuable
for its role in helping to argue for other scrutability theses, such as Conditional and A Priori Scrutability, and in providing initial motivation for these scrutability theses for those who may be skeptical about them. If it turns out that the only problem for Inferential Scrutability is the Fitch-style problem, and that this problem does not affect the other theses, then the Inferential Scrutability thesis can still play these roles reasonably well.

4 Conditional Scrutability

A truth $S$ is conditionally scrutable from a class of truths $C$, for a subject, iff the subject is in a position to know that if the members of $C$ are true, then $S$ is true. The conditional scrutability thesis says that there is a compact class of sentences $C$ such that for all subjects, all truths $S$ are conditionally scrutable from the $C$-truths for that subject. As we saw in the first excursus, the conditional formulation avoids the Fitchian problems above. Even if $S$ is unknowable, there is usually no problem knowing that if some other sentence $T$ is true, then $S$ is true. And in the cases above, there seems to be no problem with the idea that the relevant subjects are in a position to know (on idealized reflection) that if the sentences in $PQT1$ are true, then $S_1$ and $S_2$ are true.

This thesis uses the notion of conditional knowledge: that is, knowledge of conditionals such as ‘If $P$, then $Q$’. Such claims are common in English: it would be natural to say that I know that if it rains today, then my car will get wet. Such claims are about as common as claims about conditional belief, as when I say that I believe that if Australia bats first, Ponting will score a century. It is natural to hold that conditional knowledge stands to conditional belief much as knowledge stands to belief.

The correct analysis of conditional belief and conditional knowledge is nontrivial. It is implausible that conditional belief and conditional knowledge, at least as ordinarily understood in English, simply involve belief in or knowledge of a material conditional. For example, I might know that it is not raining, and thereby know (and believe) the material conditional ‘If it is raining, then my car is dry’. But if my car is out in the open, it is intuitively incorrect to say that I know (or believe) that if it is raining, then my car is dry.

It is somewhat more plausible to say that that conditional belief and knowledge involve belief in and knowledge of an indicative conditional. For reasons analogous to those above, most theorists deny that indicative conditionals are equivalent to material conditionals: intuitively, for an indicative conditional such as ‘If it is raining, then my car is dry’ to be acceptable, then there must be a stronger connection (perhaps an epistemological connection) between the antecedent and the
consequent than the material conditional requires. But it is still not entirely clear what belief in or knowledge of an indicative conditional involves. For example, Lewis (1976) gives good reason to think that conditional belief cannot simply be a matter of believing or knowing a proposition, at least while preserving one’s epistemological intuitions. If so, then if conditional belief is belief in an indicative conditional, indicative conditionals cannot be understood as propositions.

The most common view of conditional belief (associated especially with Ramsey 19xx) holds that a subject believes that if \( p \), then \( q \) (for propositions \( p \) and \( q \)) if the subject’s conditional credence in \( q \) given \( p \), \( cr(q|p) \), is sufficiently high. Here we adopt a view on which subjects have credence between 0 and 1 in various propositions: \( cr(p) = 1 \) when the subject is certain of \( p \), \( cr(p) = 0 \) when the subject rejects \( p \) with certainty, \( cr(p) = 0.5 \) when the subject is entirely agnostic between the two, and so on. If \( p \) is the proposition that the dice will come up double-six, then my credence \( cr(p) \) might be 1/36. In cases where the subject believes that \( p \), then \( cr(p) \) will be well over 1/2. To a first approximation, we can say that a subject believes that \( p \) iff \( cr(p) \) is sufficiently high. It is plausible that the threshold for belief is context-dependent, vague, and differs between different propositions: for example, a credence of 0.999 may suffice for belief in some cases (belief that it will rain today) but not in others (belief that one will lose the lottery). But we can understand “sufficiently high” to be context-dependent, vague, and variable between propositions in a similar way.

Subjects can also have conditional credences in one proposition given another. For example, if \( q \) is the proposition that the red die will come up six, then my conditional credence \( cr(q|p) \) might be 1/6. In cases where \( cr(q) \) is greater than zero and where the subject is fully rational, \( cr(q|p) \) will be equal to \( cr(p&q)/cr(q) \). But for familiar reasons (Hajek 2003), it is reasonable to hold that subjects can have a conditional credence \( cr(q|p) \) even in some cases where \( cr(q)=0 \). For example, if \( p \) is as above, and \( q \) is the proposition that a randomly thrown dart lands exactly at position \( \pi \) on an interval, then \( cr(q) \) might reasonably be 0, while \( cr(q|p) \) might nevertheless reasonably be 1/6. So a subject’s conditional credence in \( p \) given \( q \) should not in general be understood as deriving wholly from the subject’s credences in \( p, q, \) and \( p&q \). Rather, it should be understood as capturing some more complex cognitive dependence between the subject’s attitudes to \( p \) and to \( q \).

Just as we can say that a subject believes that \( p \) when their credence \( cr(p) \) is sufficiently high,
we can likewise say that a subject believes that if $p$ then $q$ when their conditional credence $cr(q|p)$ is sufficiently high. Of course, much more needs to be said about just what “sufficiently high” involves. As before, we should expect that the threshold for conditional belief will be context-dependent, vague, and will differ for different pairs of propositions. But it is not implausible that what goes for unconditional belief also goes for conditional belief.

What about knowledge? In the case of unconditional knowledge that $p$, the justification requirement on knowledge plausibly corresponds to a claim that the subject is justified in having a sufficiently high credence $cr(p)$. This requires the idea that a subject’s credences can be justified (that is, that a subject can be justified in having a certain credence), and are subject to normative assessment. Some radical subjectivists reject this claim, holding that all credences in nonlogical propositions are equally reasonable; but this path leads easily to skepticism. If one holds that beliefs can be justified, it seems reasonable to hold that credences can be justified too. Of course unconditional knowledge also requires that the proposition be true, and that some sort of anti-Gettier condition be satisfied, requiring for example that one’s justification for the proposition is appropriately connected to the truth of the proposition.

Conditional knowledge that if $p$, then $q$ also plausibly requires that the subject is justified in having a sufficiently high credence $cr(q|p)$. Insofar as unconditional credences can be justified, it is also reasonable to hold that conditional credences can be justified. A difficult question concerns whether there is a truth requirement on conditional knowledge. It is not at all clear what it means to say that the conditional ‘if $p$, then $q$’ is true, as opposed to acceptable for a subject. Still, there are plausibly cases in which the subject has a high justified credence $cr(q|p)$, but does not know that if $p$, then $q$. This can happen if $p$ is true and $q$ is false, or if the subject infers $q$ from $p$ only with the aid of a false but justified belief in $R$, for example. So the question of just what needs to be added to justified conditional belief to obtain conditional knowledge remains an open question.

For present purposes, it might suffice to rely on our intuitive understanding of conditional knowledge, just as philosophers often rely on their intuitive understanding of knowledge even without an analysis of what an anti-Gettier condition involves. But an alternative way to proceed is to stipulate that for the purposes of the Conditional Scrutability thesis, what matters is justified conditional belief, not conditional knowledge. That is, we can modify the definition of scrutability so that conditional scrutability of $S$ from $C$ requires only that the subject is in a position to have a justified conditional belief that if the sentences in $C$ are true, then $S$ is true. And we could understand this notion in turn by saying that the subject is in a position to be justified in having a sufficiently high conditional credence in $S$ given $C$. (Here the notion of being in a position to be
justified can be understood in terms of the existence of a justification, as in the fourth excursus.)

In what follows, I will move back and forth between these related conceptions of conditional scrutability. The official thesis will be cast in terms of conditional knowledge, but I will often analyze things in terms of conditional credence. This is justified in part by the plausible thesis that conditional knowledge requires a sufficiently high justified conditional credence. When moving in the reverse direction, we can explicitly attend to the possibility of high conditional credence without conditional knowledge when it is relevant.

I will also adopt the idea that for at least some subjects and some propositions, there is a rational credence for the subject to have in the proposition, the credence that the subject ideally ought to have in the proposition. Or better, I will assume that for some subjects and propositions, there is a rational range of credences, in that one or more credences are rational, and some credences are irrational. Of course a subject’s rational credence in a proposition may differ from the subject’s actual credence in that proposition, if the subject is not ideally rational. We can say that the rational credence for a subject in a proposition is high when only high credences in the proposition are rational for the subject. As before, if we deny that subjects have high rational credences in some propositions, it is not easy to avoid skepticism.

We can then say that \( p \) is conditionally scrutable from a class of propositions \( c \), for a subject, if the subject’s rational credence \( cr'(p|c) \) is high, where \( cr'(p|c) \) is stipulated to be \( cr'(p|cc) \), where \( cc \) is a conjunction of all the propositions in \( c \). Here the notion of rational conditional credence in a pair of propositions is understood in a way parallel to the understanding of rational unconditional credence in a single proposition, above.

We can also define notions of conclusive knowledge and scrutability, involving knowledge with certainty. Intuitively, knowledge with certainty of \( p \) requires that one is justified in having credence 1 in \( p \), so that one’s rational credence in \( p \) is 1. Likewise, conditional knowledge of \( p \) given \( q \) requires that one’s rational credence in \( p \) given \( q \) is 1. We can then say that \( p \) is conclusively conditionally scrutable from \( c \) for \( s \) if \( cr'(p|c)=1 \). This relation is stronger than the nonconclusive conditional scrutability relation, but it is also better defined and better behaved, and it will be useful for some purposes.

\[ \text{It is arguable that credence 1 does not suffice for certainty. I might have credence 1 that the value of a random real number between 0 and 10 is not } \pi, \text{ without being certain of it. Some will say that the credence here is infinitesimally less than 1, or at least that it is not “true 1”. But in any case, rational credence 1 is plausibly a necessary condition for certainty, and the notion of conclusive scrutability defined this way is at least a good approximation to conclusive conditional scrutability. As before we can attend the the possibility of exceptions as they come up.} \]
What about credences and conditional credences in sentences? We might define these directly, for example in terms of the rational betting odds associated with sentences and with pairs of sentences for a subject. They can also be defined in terms of the thoughts apt to be expressed by the relevant sentences. I spell this sort of analysis out in the third excursus. For now, for ease of discussion, we can take it that a subject’s credence $cr(S)$ in sentence $S$ (in a context) is the subject’s credence in the proposition that $S$ expresses (relative to that context), or perhaps the subject’s credence in the proposition that $S$ expresses under the guise associated with $S$ (in that context). A subject’s rational credence $cr'(S)$ in $S$ is the subject’s rational credence in the proposition expressed by $S$ (under the guise, in the context). With appropriate modifications, the same goes for a subject’s conditional credence $cr(S_1|S_2)$ and the subject’s rational conditional credence $cr'(S_1|S_2)$ in a sentence $S_1$ conditional on another sentence $S_2$.

We can then say that $S$ is conditionally scrutable from $C$ (for a subject $s$ at time $t$) iff $cr'(S|C)$ is high for $s$ at $t$, where $cr'(S|C)$ is stipulated to be $cr'(S|CC)$ and $CC$ is a conjunction of all sentences in $C$. The Conditional Scrutability thesis says that there is a compact class $C$ of truths such that for all subjects $s$ and times $t$, all truths are conditionally scrutable from $C$ for $s$ at $t$. (A slightly modified version to accommodate context-dependence is given in the third excursus.) We can likewise define conclusive versions of the conditional scrutability relation and the corresponding thesis, by replacing “high” by “1” in the definition.

It has taken a while to unpack the conditional scrutability thesis, but it remains plausible. In particular, it remains plausible that given such a class as $PQT1$ above (or some disciplined subclass of it), then for at least many truths $S$, ideal reasoning would support a high conditional credence in $S$ given the hypothesis that all the sentences in $PQT1$ are true.

In many cases, when $S$ is inferentially scrutable from $C$ for a subject, it will be plausible that $S$ is conditionally scrutable from $C$ for that subject. This follows from a version of the Bayesian principle of conditionalization (discussed in chapter 4), at least if we assume that $C$ is the total relevant evidence that the subject acquires. This assumption may be false in some cases: for

---

12Of course this requires that the class $C$ is sufficiently disciplined, and/or that the language is sufficiently infinitary, so that a conjunctive sentence $C'$ exists. If $C$ itself contains arbitrary conjunctions, we can avoid regress by stipulating that $C'$ is a conjunction of all nonconjunctive sentences that are either members of $C$ or conjuncts of conjunctive sentences in $C$.

13Whereas other scrutability thesis trivially rule out the possibility that false sentences are scrutable from truths, this is not quite so trivial for conditional scrutability. So one might consider explicitly making the thesis a biconditional, saying that $S$ is true iff it is conditionally scrutable from $C$. Still, this claim plausibly follows from the original version, along with the claim that when $S_1$ and $S_2$ express incompatible thoughts, $cr'(S_1|C)$ and $cr'(S_2|C)$ cannot both be high.
example, in coming to know $p$ one may come to know that one believes $p$. But one might instead appeal to a mildly strengthened inferential scrutability thesis: if one were to come to know the truths in $C$ and use no additional evidence other than the truths in $C$, one could thereby come to know $S$. This strengthened thesis remains plausible in the core cases, and leads more directly to conditional scrutability.

So there is good reason to accept that in the core cases in which inferential scrutability holds, conditional scrutability holds. Furthermore, even in the Fitchian cases in which inferential scrutability fails, there is no corresponding reason to think that conditional scrutability fails. So if inferential scrutability is true of all non-Fitchian cases, then there is a good case for thinking that conditional scrutability holds in general. The relationship between empirical and conditional scrutability is discussed further in chapter 3.

One concern about the conditional scrutability thesis is that the idealization involved in it is not as easy to attenuate as with other scrutability theses. In the case of inferential scrutability, we attenuated the idealization by moving to a notion on which $S$ is inferentially scrutable from $C$ if there is some subclass $C'$ of $C$ such that knowing $C'$ would put one in a position to know $S$. One might correspondingly move to a notion on which $S$ is conditionally scrutable from $C$ if there is some subclass $C'$ of $C$ such that $cr'(S|C')$ is high. The trouble with doing this is that for almost any $S$, including false $S$, this definition will be satisfied. For most false $S$, there are some truths that taken collectively would constitute strong misleading evidence for $S$. If $C$ contains such sentences, constituting a subclass $C'$, then $cr'(S|C')$ will be high, and $S$ will be conditionally scrutable from $C$ even though $S$ is false.

A version of this problem arose for inferential scrutability, but it was not nearly as acute: it is impossible to know a false sentence, so no false sentence is inferentially scrutable. To avoid this worry in the case of conditional scrutability, it is best not to employ the weaker subclass-involving notion of conditional scrutability. Of course this means that to make an inference from inferential scrutability to conditional scrutability, one will also have to employ the stronger notion of inferential scrutability that avoids subclasses (although one might still use the weaker sort of inferential scrutability to help make the case for the stronger sort).

Where conditional scrutability is concerned, there is no getting around the need for a strong idealization at some point. In fact, the idealization needs to be subtly modified to handle problems tied to self-doubt, in which subjects are not confident of their own reliability. I discuss this modification in the sixth excursus.
5 A Priori Scrutability

The A Priori Scrutability thesis says that all truths are a priori scrutable from a compact class of base truths. The informal characterization in section 3 says that \( S \) is a priori scrutable from \( C \) for \( s \) if \( s \) is in a position to know a priori that if \( C \), then \( S \) (where the antecedent should be read as involving a conjunction of all the sentences in \( C \)).

The parallel with conditional scrutability might suggest that a priori scrutability is defined in terms of conditional knowledge, as in the last section, with the added requirement that the knowledge is a priori. But in the case of a priori scrutability we can avoid the complex issues about conditional knowledge entirely. Instead, we need require only that a material conditional ‘if \( C \) then \( S \)’ can be known a priori. Using material conditionals would trivialize conditional scrutability, but it does not trivialize a priori scrutability. While we can know the material conditional ‘if \( C \) then \( S \)’ just by knowing \( S \), we cannot know the material conditional a priori in any such simple way. (One can know it a priori just by knowing \( S \) is a priori, but this is the right result for a priori scrutability.)

This leads to the definition of a priori scrutability in the last chapter: \( S \) is a priori scrutable from \( C \) (for \( s \)) iff a material conditional from \( C \) to \( S \) is a priori. We can also say that \( S \) is a priori scrutable from \( C \) when there is some conjunction \( D \) of sentences in \( C \) such that the material conditional \( D \rightarrow S \) is a priori. Near-equivalently, \( S \) is a priori scrutable from \( C \) iff \( S \) can be logically derived from some sentences in \( C \) and some a priori truths.

Apriority of a sentence is discussed in the third excursus and in chapter 4. To a first approximation, we can say that \( S \) is a priori when it is a priori knowable: that is, when it is metaphysically possible for someone to know \( S \) a priori. On the analysis in the third excursus, one knows \( S \) a priori if one has a priori knowledge that is apt to be expressed by \( S \). Alternatively, one can say that \( S \) is a priori when there is an a priori warrant for \( S \), as discussed in the fourth excursus. The definition in terms of warrant has the advantage that it is not hostage to claims about what is metaphysically possible, and that it extends more naturally to the case of propositions while avoiding the difficulties about semantic fragility discussed at the end of the third excursus.

We can relativize apriority of a sentence to a subject if we want to: \( S \) is a priori for a subject when it is metaphysically possible for the subject to know \( S \), or when there is an a priori warrant for the subject to believe \( S \). This yields a notion of a priori scrutability that is itself subject-relative. We could then state an A Priori Scrutability thesis that quantifies over subjects (as Inferential and Conditional Scrutability do): there is a compact class \( C \) such that for all subjects, \( S \) is a priori scrutable from \( C \). But if the relevant sentences are restricted to context-independent sentence (see
the third excursus), then a sentence is plausibly a priori for one subject iff it is a priori for all subjects, so this relativization is not really necessary. When $S$ is a priori scrutable from $C$ in the non-subject-relative sense, it follows that $S$ is a priori scrutable from $C$ for some subject, so (given the previous sentence) $S$ is a priori scrutable from $C$ for all subjects. If so, the A Priori Scrutability thesis that does not mention subjects is equivalent to the one that quantifies over all subjects. I discuss the extension to context-dependent sentences, where some relativization to subjects or contexts is needed, in the third excursus.

Note that a priori scrutability, unlike conditional scrutability, requires that $S$ be epistemically related to some conjunction of members of $C$, rather than to the conjunction of all members. We could define it the latter way, but the two definitions are near-equivalent. If $S$ is a priori entailed by a conjunction of all members of $C$, it is trivially a priori entailed by some conjunction of members of $C$. If $S$ is a priori entailed by some conjunction of members of $C$, then $S$ is a priori entailed by a conjunction of all members of $C$, at least if such a conjunction exists. This parallels the observation that if a material conditional $A \rightarrow S$ is a priori, then so is any material conditional $A \& B \rightarrow S$, because the former conditional entails the latter. As before, this attenuates the required idealization somewhat: to establish a priori scrutability of $S$ from $C$, we need not always consider the conjunction of all members of $C$. A priori entailment by some proper subclass of $C$ will suffice.

A priori scrutability is clearly not subject to the Fitchian problems that arose for inferential scrutability. It also involves much less subject-relativity than conditional scrutability. In a number of respects, a priori scrutability is better behaved than the other sorts of scrutability, and it is the notion that I will concentrate on the most. The main downside of the notion compared to the others is that it invokes the more theoretical and controversial notion of the a priori. So it is useful to have the other notions too, to help motivate and argue for scrutability claims. Still, a priori scrutability will be the central focus.

6 Generalized Scrutability

Scrutability theses need not be restricted to the actual world. If the a priori scrutability thesis is true, then it is plausible that it still would have been true if the world had turned out differently. To see this, note that we can evaluate the truth of various sentences even given hypothetical information about ways the world might be. For example, in the Gettier case, it is irrelevant whether Smith’s case is actual: a subject can know that if Smith’s case as described is actual, then Smith does not know that someone owns a Ford. Or in the case of water, given an appropriate speci-
fication of the distribution, behavior, and appearance of clusters of XYZ molecules (information analogous to the information we have about H₂O in the actual world), a subject is in a position to conclude that if the specification is correct, then water is XYZ.

One might formulate a stronger scrutability thesis by requiring that the A Priori Scrutability thesis be necessary, or better, by requiring that it be a priori. If A Priori Scrutability is itself a priori (and conclusive a priori, in the sense outlined in chapter 4), then it will be true however the world turns out. That is, for every epistemically possible scenario, it will be a priori that if the scenario obtains, then A Priori Scrutability is true. So for every scenario, it will be a priori that truths (with respect to that scenario) are scrutable from a compact class of truths (with respect to that scenario). This would require that there is a compact scrutability base for each scenario, although there might be entirely different scrutability bases for different scenarios.

For my purposes, it is useful to formulate a somewhat stronger thesis still, according to which there is a single scrutability base that applies to all scenarios. We can formulate such a thesis as follows. Let us say that a sentence S is epistemically possible iff S is truth-apt and ¬S is not a priori, and that a class of truths is epistemically possible if every conjunction (finite or infinite) of truths in that class is epistemically possible.

14 Generalized Scrutability: There is a compact class C of sentences such that for all sentences S, if S is epistemically possible, then there is an epistemically possible subclass C′ of C such that S is scrutable from C′.

In effect, the generalized scrutability thesis says that there is a compact class of sentences that will serve as a scrutability base however the world turns out. The relevant compact class might be thought of as yielding a sort of scrutability base for the entire space of epistemically possible scenarios. Of course there are epistemically possible scenarios that involve all sorts of alien properties that are not present in the actual world. As a result, a scrutability base for the entire space of such scenarios may need to involve many more families of expressions than a scrutability base for the actual world. So for the purposes of generalized scrutability, the understanding of “compactness” might need to be weakened significantly, compared to the understanding that is required for actual-world scrutability. Perhaps we might even need to allow an infinite number of

---

14 As always in this book, this use of “epistemically possible” is stipulative and does not reflect the ordinary usage of the expression, according to which S is epistemically possible roughly when one does not know ¬S. See “The Nature of Epistemic Space” for more on the relations between the notions.
families in the base, for example if there are infinitely many sorts of fundamental alien properties. But the hope is that some interestingly limited class will suffice all the same.

In principle, one can formulate generalized versions of A Priori Scrutability, Conditional Scrutability, Inferential Scrutability, and various other scrutability theses. But as one might expect, it will be Generalized A Priori Scrutability that is most central for our purposes.

7 Idealization

Laplace’s scrutability thesis, in the introduction, appealed to an intellect “vast enough” to know all the physical truths and to “submit these data for analysis”. In effect, Laplace is invoking an idealization of cognitive capacities, allowing capacities far greater than normal human capacities.

All of our central notions of scrutability involve such an idealization. Inferential scrutability appeals to what it is metaphysically possible for a subject to know. Conditional scrutability appeals to what a subject ideally ought to believe. A priori scrutability appeals to what it is metaphysically possible for a subject to know a priori.

These idealizations, like Laplace’s, have at least two dimensions. First, they idealize thoughts: they allow subjects to entertain thoughts whose complexity is far beyond normal human capacity (the conjunction of all base truths, for example), or that involve concepts that humans do not possess (truths about what it is like to be a bat, for example). Second, they idealize reasoning: they allow subjects to engage in forms of reasoning are far beyond normal human capacity (proofs of enormously complex mathematical theorems, for example). The idealization of thought is required to allow subjects to entertain base truths and dependent truths, both of which may involve new concepts and great complexity. The idealization of reasoning is required to allow the subjects to make connections between the two.

The idealization of thought involves two fairly clear dimensions in turn: any concept that it is possible to possess is allowed, and arbitrary complexity (finite or infinite) is allowed. The idealization of reasoning has less clear dimensions, but any possible pattern of reasoning is allowed. We need not dictate what counts as reasoning. Any process by which someone comes to know something (or by which they come to know something a priori) is good enough for our purposes. As far as I can tell, however, the specific idealized forms of reasoning that I invoke in this book all involve natural extensions of familiar sorts of human reasoning (deduction, induction, inference, intuition, and various others), perhaps extended to arbitrary complexity or to infinitary processing. Given some epistemological views on which knowledge is cheap, one may have to constrain the
allowed forms of reasoning to avoid trivializing scrutability theses (as discussed briefly in the next chapter), but on my own views this is not necessary.

One could in principle define nonidealized analogs of the idealized notions of scrutability, perhaps in terms of what it is practically possible for subjects to know, or in terms of what a subject ought to believe in an ordinary nonidealized sense. But scrutability theses corresponding to the theses we have considered would then be much less plausible. For example, it is not practically possible for most subjects to entertain enormous world-descriptions; likewise, it is not clear that subjects ought to believe complex mathematical theorems in any ordinary nonidealized sense. It is not out of the question that by building more into the base (all mathematical truths, for example), the required idealization on reasoning might be diminished; and it is not out of the question that by reformulating scrutability theses in an appropriate way, the required idealization on entertaining thoughts might be diminished. (One such less idealized thesis is considered in chapter 9.) But at least for a version turning on a priori entailment of all truths from base truths, an idealization is hard to avoid.

Idealizations of cognitive capacity can be made in at least three different ways for our purposes. First, there are modal idealizations, cast in terms of what it is metaphysically possible for a subject to know or believe. Second, there are normative idealizations, cast in terms of what subjects ideally ought to believe. Third, there are warrant idealizations (discussed in the fourth excursus), cast in terms of what there is an (ideal) warrant for subjects to believe.

In principle, a given scrutability thesis can be cast in any of these three ways. We can say that a sentence is a priori, for example, when it is possible for someone to know it a priori, or when someone who cares about its truth and is restricted to a priori reasoning ideally ought to believe it, or when there exists an ideal warrant for someone to believe it.

For example, given a complex mathematical theorem $M$ and a nonideal subject Fred, it is metaphysically possible that Fred come to know $M$, although it is not practically possible. Fred ideally ought to believe $M$, at least if he were to consider it and to care about it, although it is not the case that he ought to believe $M$ in a less idealized sense. There is an ideal warrant for Fred to believe $M$, deriving from the proof of $M$, but Fred does not have this warrant to believe $M$.

The relation between these three sorts of idealization is complex. One might think that normative idealizations are answerable to modal idealizations (because ought implies can) or that warrant idealizations derive from normative idealizations. For my part, I think that warrant idealizations are the most fundamental. When someone ideally ought to believe $p$, this is because there is a warrant for them to believe $p$. When someone is in a position to know $p$, again this is because
there is a warrant for them to believe $p$.

Modal idealizations are perhaps the most familiar of the three, but they have some disadvantages. One is the problem of semantic fragility, discussed at the end of the third excursus. Another is that if we define apriority and the like using a modal idealization, claims about apriority are then hostage to certain theoretical views about modality. For example, some hold that there are brute constraints on metaphysical modality. If cognitive capacities beyond a certain size are metaphysically impossible, for example, then certain apparent apriorities, such as mathematical theorems that require enormous proofs, will not count as a priori. But there may still be an ideal a priori warrant for such a theorem, in virtue of there being a proof for the theorem. Related constraints may arise from views on which anything nomologically impossible is metaphysically impossible. I do not think that metaphysical modality is constrained in this sort of way, but someone who holds that there are such constraints should appeal to warrant idealizations (or perhaps normative idealizations, invoking ideal norms such that ought does not imply can) instead of modal idealizations.

In this book I appeal more often to modal idealizations than to normative and warrant idealizations, mainly because the notions involved (possibility and knowledge) are more familiar than those of ideal norms and ideal warrants. But granted an understanding of normative idealizations or of warrant idealizations, it is certainly possible to interpret most of what I say in this book using those idealizations alone. I discuss warrant idealizations at more length in the fourth excursus.

The use of large idealizations such as these can provoke a number of objections. These include objections from incoherence, from knowledge, from triviality, and from applicability. I discuss these objections in section 7 of chapter 3, after the role of idealization in my arguments has become somewhat clearer. One problem in particular, the problem of self-doubt, requires some modifications of the idealizations I have discussed so far, invoking a special notion of insulated idealization. I discuss those issues in the sixth excursus, following the next chapter.
Third Excursus: Sentential and Propositional Scrutability

In chapter 2, I introduced the notion of knowing a sentence, and suggested that scrutability theses might be formulated in terms of it. In what follows, I analyze this notion in more detail. I also formulate the sentential scrutability theses that result, and discuss their relations to propositional scrutability theses, along with associated issues about context-dependence, semantic fragility, propositional warrant, and the persistence of thoughts.

Sentences and thoughts

Earlier in the chapter, I suggested various ways of understanding the notion of knowing a sentence that are available to theorists of various different stripes. Different theorists can say that to know $S$ is to know the Fregean proposition that it expresses, or to know the Russellian proposition that it expresses under the guise of expression, or to know that $S$ is true. Here I will develop another way to understand the notion that is available to theorists of many different stripes. This understanding is not obligatory, but it helps to give an idea of just how the notion works.

The approach taken here relies on connections between utterances and mental states of the speaker. Intuitively, when Sue utters ‘Hesperus is a planet’, her utterance is associated with a mental state—a specific state of entertaining the proposition that Hesperus is a planet—that is itself a state of belief and a state of knowledge. When Sue utters ‘Phosphorus is a planet’, her utterance is associated with a very different sort of mental state—a specific state of entertaining the proposition that Phosphorus is a planet—that is neither a state of belief nor a state of knowledge.

To make this more precise, let us say that entertaining is the maximally general propositional attitude (occurrent or non-occurrent) with a mind-to-world direction of fit. So when one believes that $p$, knows that $p$, expects that $p$, hypothesizes that $p$, or supposes that $p$, one entertains $p$. Now let us say that a thought is a specific state of entertaining. This notion is parallel to that of a belief (a specific state of believing), a supposition (a specific state of supposing), a knowing (a specific state of knowing), and so on.

Wherever there is a belief, there is a thought. Likewise, wherever there is a knowing, there is a thought. The thought is intimately related to the belief and the knowing. On some views, the thought, the belief, and the knowing are all identical to each other. On another view, they are distinct states, but they stand in some other strong relation to each other: a relation of coincidence or realization, for example. I will not adjudicate this question here, but I will say that when this relation holds, the thought constitutes the belief and constitutes the knowing. When a thought con-
stitutes a belief or constitutes a knowing, we can also say more simply that the thought constitutes belief or that the thought constitutes knowledge (as with my current thought that 2+2=4, for example). We can also say that a thought constitutes justified belief iff it constitutes a justified belief, and that a thought constitutes a priori knowledge if it constitutes an item of a priori knowledge: that is, a knowing whose justification is independent of experience.

I will be especially (although not only) concerned with *occurrent* thoughts. In this special case, thoughts are not mere dispositional states, as with beliefs that are currently dormant. Instead, we can take occurrent thoughts to be specific acts of entertaining occurring in a subject’s stream of thought. These acts will themselves constitute acts of judging, of supposing, and so on. I take it that acts of judging can at least sometimes constitute states of believing and knowing, so that occurrent thoughts can also constitutes states of believing and knowing. I will not restrict the notion of thoughts to occurrent thoughts, and non-occurrent thoughts will sometimes be relevant, but occurrent thoughts will play the central role.

We can then appeal to the idea that utterances of truth-apt sentences typically *express* thoughts (and indeed occurrent thoughts). Sincere utterances typically express beliefs, and correlatively express thoughts. Insincere utterances do not express beliefs, but they nevertheless typically express thoughts. Note that expression is here construed as a relation between utterances and mental states, and should not be confused with the different notion of expression construed as a relation between utterances and propositions. However, there is a close relationship between the notions. Intuitively, an utterance and the thought it expresses have the same propositional content. I will not build this in as a definitional constraint, as there may be views on which thoughts and utterances have contents of different sorts. But it is at least a constraint that when an utterance expresses a thought, the utterance and the thought have the same truth-value, and it is natural to hold that they must have the same truth-conditions as well.15 These are not the only constraints on the notion: there must also be a causal link between the thought and the utterance, and an appropriate psychological relation. I will not try to define these things here, and will take the notion as an intuitive primitive (unanalyzed at least for now) instead.

There are weaker notions on expression that do not have this constraint. For example, according to a weaker notion, one could express a false thought with a true utterance by misusing a word, or merely by conveying the content of the thought while asserting something else. According to the stronger notion I am using here, these do not qualify as cases of expression. In fact it is best to think of expression as a relation that comes with an a priori guarantee of truth-preservation (roughly as deductive inference might), so that the utterance is guaranteed to have the
same truth-value as the thought.

It may be that there are some utterances that do not express thoughts. One might absent-mindedly utter a sentence by rote, without entertaining its content at all. If one is grasping for words, one might use a word whose content is not that of a thought. But at least for typical utterances, it is plausible that they express thoughts. So we can utterances that express thoughts to be our paradigm cases for initial analysis.\footnote{This restriction will affect only the analysis of context-dependent sentences below, not the analysis of context-independent sentences. In particular it will affect the analysis of what it is to be in a position to know a sentence in a context. In the first instance we can restrict attention to contexts in which a sentence is uttered expressing a thought. To handle other contexts, we can appeal to the notion of having warrant to accept a sentence in a context, along the lines of the following excursus, or we can appeal to possible thoughts that endorse the utterance.} Insofar as there are utterances that do not express thoughts, we can handle these derivatively.

In the reverse direction, we can say that when an utterance expresses a thought, the thought *endorses* the utterance. Such a thought will not be metalinguistic thought about the utterance: rather, if the utterance is about bachelors (say), the thought will be about bachelors. It is possible in principle even for a thought that follows an utterance to endorse that utterance, although understanding this idea properly requires appeal to the notion of persistence discussed below.

For present purposes, we can start with a subject who is making a fully competent utterance of a sentence $S$ (one in which the expressions in $S$ are used correctly and nondeferentially). We can say that if the utterance expresses a thought that constitutes belief, the subject believes $S$. If the utterance expresses a thought that constitutes knowledge, the subject knows $S$. This fits naturally with the accounts of these notions in terms of sincere and knowledgeable assertion: it is plausible that an assertion is sincere precisely when it expresses a thought that constitutes belief, and that it is knowledgeable precisely when it expresses a thought that constitutes knowledge. Then in the case of Sue, it is plausible that Sue knows ‘Phosphorus is a planet’ but not ‘Hesperus is a planet’: the thought expressed by her utterance of the former constitutes knowledge (and belief), but the thought expressed by her utterance of the latter does not.

This analysis is available to many different theorists of propositions. The argument from assertion earlier strongly suggests that there are different mental states associated with Sue’s utterances of ‘Hesperus is a planet’ and ‘Phosphorus is a planet’. Whatever one’s account of propositions, it is hard to deny that there are distinct thoughts (acts of entertaining) with Sue’s utterances of ‘Hesperus is a planet’ and ‘Phosphorus is a planet’, and that these thoughts have relevant psychological and epistemological differences (one constitutes an act of belief and an act of knowledge,
the other does not). Rather than denying these claims, it is more plausible for a Russellian to deny that a thought constitutes knowledge if the subject knows the proposition that is the content of the thought. In cases such as the above, the subject may have two thoughts with the same content, one of which constitutes knowledge and the other one does not. The same goes for belief.

Here it might help to briefly adopt a common model according to which thoughts correspond to sentence tokens in the language of thought. Each token has some content. When such a sentence token is in the “belief box”, it corresponds to a belief in the content. When the token is in the desire box, it corresponds to a desire in the content. When a token in the belief box was brought about by the right sort of process, it constitutes a justified belief. When further conditions are met, it constitutes knowledge. In a case such as the above, Sue’s utterance $S_1$ might be triggered by the sentence ‘$H = H$’ in her language of thought, while $S_2$ might be triggered by the sentence ‘$H = P$’ in her language of thought. The first sentence will be in the belief box, will meet the relevant further conditions, and will constitute knowledge. The second sentence will not be in the belief box, so it will not constitute belief or knowledge.

Of course the model involving a language of thought and belief boxes may be a fiction. But it remains plausible that thoughts correspond to specific states of a cognitive system, playing specific functional roles. If the state plays the right sort of role, the corresponding thought will be a belief, and so on. Even without a language of thought, Sue’s utterances of $S_1$ and $S_2$ are plausibly brought about by quite different states, one of which plays the functional role of a belief and the other one of which does not. So one state corresponds to a belief, and the other to a thought that is not a belief. All of this is so far quite compatible with a view on which these two thoughts have the same content.

We can extend the thought-based analysis to sentences that the speaker is not currently uttering by by saying: a subject knows (or knows a priori, or believes) a sentence $S$ if the subject has a thought apt to be expressed by $S$ and that constitutes knowledge (or a priori knowledge, or belief). Here, a thought is apt to be expressed by a sentence type $S$ if it could be expressed by a fully competent utterance of $S$. In this case, unlike the case where the thought is actually expressed, the thought may well be non-occurrent. Furthermore, knowing and believing a sentence type $S$ does not require that the subject is competent with the expressions in $S$: a French speaker who believes $2 + 2 = 4$ believes ‘Two plus two equals four’, in virtue of having a belief apt to be expressed by the sentence, if the subject were competent in English.

One define credences in sentences in a similar way. We can assume that thoughts are at least sometimes associated with credences $cr(T)$ and conditional credences $cr(T_1|T_2)$. We can then say
that a value $x$ is in the credence range $cr(S)$ for a subject at a time iff $x$ is in $cr(T)$, where $T$ is a thought by $s$ at $t$ that is apt to be expressed by $S$. Likewise, we can analyze $cr(S_1|S_2)$ in terms of $cr(T_1|T_2)$, where $T_1$ and $T_2$ are apt to be expressed by $S_1$ and $S_2$ respectively. The same goes for rational credences $cr'(S)$ and $cr'(S_1|S_2)$ and insulated rational credences $cr^*(S)$ and $cr^*(S_1|S_2)$.

Some philosophers are doubtful about the very idea of token thoughts. For example, Lewis (1994) suggests that “beliefs” is a “bogus plural”. On this view, subjects can certainly believe that $p$ for various $p$, but there are no token entities called beliefs to undergird this believing, except perhaps for trivial derivative entities such as the instantiation of the property of believing that $p$. Presumably Lewis would take a similar view of thoughts. For someone with this view, the current definitions of key notions such as knowing a sentence will be problematic.

It is worth noting that some of the doubts about token beliefs apply less clearly to occurrent thoughts. I think that it is an introspective datum that there are acts of judging and acts of entertaining a hypothesis. Given that there are such acts, it is hard to deny that there are at least occurrent thoughts.

If someone rejects the very idea of token thoughts, the key notions such as that of knowing a sentence $S$ (and the apriority of $S$, discussed below) will have to be understood differently. If the theorist takes a fine-grained view of propositions, so that ‘Hesperus is Hesperus’ and ‘Hesperus is Phosphorus’ express distinct propositions, it will suffice for present purposes for them to understanding knowledge of $S$ as knowledge of the proposition it expresses. If the theorist takes a coarse-grained view of propositions but allows that there are guises, it will suffice for them to understand knowledge of a $S$ as knowledge of a proposition under the guise associated with $S$, as above. If the theorist accepts none of these notions, an alternative approach will be needed. But in any case, the intuitive distinction in the status of various utterances is clear, and any theory that cannot explain it is an incomplete theory. So I will take it that even for theorists who reject all of the notions just mentioned, some way of understanding the key notion of knowing or believing a sentence can be found.

It is also not out of the question that one could take the notion of a thought as a useful fiction.

---

17This definition requires that there is a thought corresponding to any sentence with a credence, so unless we are very liberal about thoughts, the model will lead to undefined credences in many sentences. An alternative model says that $cr(S)$ is the credence one would have if one harmlessly entertained $S$—that is, entertained $S$ without disturbing any existing thoughts. This model is subject to certain counterexamples, for example involving sentences such as ‘I am entertaining a sentence about entertaining’. As I discuss in the third excursus, these counterexamples do not arise for insulated rational credences, and can be avoided in the case of non-insulated rational credences by giving a warrant-based analysis.
for motivating an intuitive idea of the apriority or scrutability of an utterance, and leave that notion unanalyzed in the short term. Once one has motivated scrutability theses and the like in this way, one can use them to motivate a Fregean notion of proposition, and of the propositional content of a sentence. With this done, one could return to a characterization of apriority in terms of Fregean propositions.

**Sentential scrutability and context-dependence**

How do we formulate scrutability theses in terms of sentences? We need to say when it is for a sentence to be (inferentially, conditionally, a priori) scrutable from a class of sentences for a subject, and then plug these notions into the general form of a scrutability thesis. I will start with context-independent sentences, which are fairly straightforward, and then move to complications raised by context-dependence.

Inferential scrutability for sentences was defined in section 4, in terms of possible knowledge of sentences. Given the notion of knowledge of a sentence $S$ from the last section, this definition is straightforward. Conditional scrutability of sentences was defined in section 7 in terms of a subject’s rational credences $cr^*(S_1|S_2)$ in certain sentences $S_1$ given sentences $S_2$. This notion was also defined in the last section (subject to a modification later in this excursus). A priori scrutability for sentences was defined in section 5, in terms of the apriority of sentences. Using the framework in the last section, one can say that $S$ is a priori when it is possible that someone comes to know $S$ a priori, and that $S$ is a priori for a subject when it is possible that that subject comes to know $S$ a priori.

In the first instance, we can take the sentences here to be context-independent sentence types in any possible language. Such a thesis has roughly the scope of a propositional scrutability thesis, at least if we assume that for every proposition, there is some possible sentence that expresses $p$ context-independently. For this assumption (discussed in the next section) to have a chance of being true, we must quantify over more than just English sentences or sentences in languages that are actually spoken. The restriction to context-independent sentences simplifies things while still allowing considerable power.

---

18I will be neutral on just what sort of abstract object a sentence type is, but I will take it that they are structures composed from simpler expression types such as words. I will take it that expression types belong to their languages essentially, so that the same expression type cannot recur in English and French. I will also take it that where there are ambiguous strings such as ‘bank’ there is more than one corresponding expression type in English. It follows that expression types are more than uninterpreted strings.
For some purposes, we need to apply scrutability theses to context-dependent sentences. After all, many sentences in natural language are context-dependent, and we would like to be able to apply scrutability theses to them. This is particularly important if we want to use scrutability to define a sort of content for utterances of such sentences, as I do. In general, the content of a sentence is defined in terms of what that sentence is scrutable from, and when a sentence is context-dependent, we should expect both the content and the scrutability of the sentence to vary with context.

Here, sentences in the dependent class and in the base class need to be treated somewhat differently. Sentences in the base class need not include natural-language sentences: the project of the *Aufbau* does not require that there is a natural-language expression for recollected similarity, for example. So we can allow that the sentences in the base class be sentences in a merely possible language, if appropriate natural-language sentences are not available. We can also largely allow that base sentences to be restricted to context-independent sentences, apart from (as we will see) a couple of exceptions such as ‘I’ and ‘now’.

The sentences in the dependent class must include sentences of natural language, however, at least for purposes such as defining the content of those sentences. When these sentences are context-dependent, the truth of these sentences will be context-dependent, as will their scrutability. For example, there may be no context-independent fact of the matter about whether ‘87 is a large number’ is true, or whether it is scrutable from a certain base.\(^{19}\)

Furthermore, there are cases in which scrutability is context-dependent even if truth is not. Ordinary proper names are often used with different modes of presentation in different contexts. For example, Leverrier might have used ‘Neptune’ solely under a mode of presentation characterizing it as whatever perturbs the orbit of Uranus, while later speakers might not. Then the truth of ‘Neptune is a planet’ is not context-dependent. But nevertheless, it seems reasonable to say that ‘Neptune is a planet’ is scrutable from ‘A planet perturbs the orbit of Uranus’ on Leverrier’s usage, but not on later usage. One can get the same behavior with any ordinary proper name. Any such name can be associated with different modes of presentation for different speakers and occasions: for example, the name ‘Adolf Hitler’ will involve a different mode of presentation for his mother

---

\(^{19}\)Knowledge of and belief in context-dependent sentence types can still be defined as in the previous section, but it will not yield useful results, because the definition will put only weak constraints on what is required to know or believe \(S\). For example, any belief about the location of an object is apt to be expressed by ‘It is there’, so almost any subject will believe this sentence. Likewise, almost any subject will believe ‘It is not there’. For useful results, we need to relativize to context somehow.
and for a student today. We can expect that this difference in mode of presentation will give rise to differences in scrutability.

We might put this by saying that these expressions are **epistemically context-dependent**, even if they are not **extensionally context-dependent**. 20 Officially, an expression \( E \) is epistemically context-dependent when there exists a sentence \( S \) containing \( E \) such that \( S \) is a priori scrutable from some base sentences in one context and not in another, wholly in virtue of a difference in the use of \( E \) in those contexts. 21 \( E \) is extensionally context-dependent when the extension of \( E \) can vary between contexts.

By contrast, an expression such as ‘I’ is extensionally context-dependent, but it is plausibly not epistemically context-dependent. That is, ‘I’ is associated with a canonical first-person mode of presentation, and there is no associated context-dependence in whether it is scrutable from various bases. The same goes for ‘now’, at least if we ignore any context-dependence associated with whether it refers to an instant or a longer span of time. ‘I’ and ‘now’ are both extensionally context-dependent, though, in that they can refer to different referents on different occasions of use. Expressions such as ‘tall’, ‘ready’, and ‘that’ are plausibly both epistemically and extensionally context-dependent.

When an expression is extensionally context-dependent, the truth of sentences containing it will depend on context. When an expression is epistemically context-dependent, the scrutability of sentences containing it will depend on context. So if scrutability theses are to accommodate such expressions, we need them to accommodate this sort of context-dependence.

To do this, we can apply scrutability not just to sentence types but to sentence tokens. The sentences discussed so far are sentence types: abstract expressions such as ‘The cat sat on the mat’, not anchored to any specific occasion of utterance. By contrast, sentence tokens are anchored to

---

20 On a standard Russellian view of content, names will have the same content relative to all contexts. If so, epistemic context-dependence will not be a form of semantic context-dependence, in the sense that an expression has different contents relative to different contexts. But not all forms of context-dependence are semantic, and even a Russellian should acknowledge the epistemological variations between different utterances of names.

21 The relevant contexts here should be restricted to those in which \( S \) is uttered with full competence. For the definition of scrutability in a context, see below. An awkward consequence of the definition of epistemic context-dependence is that someone who rejects apriority (or rejects a priori scrutability for names from any base) will then be committed to the claim that no expressions (or no names) are epistemically context-dependent. I argue against these views later, but given a view on which they are correct, it is probably best to replace the appeal to a priori scrutability in the definition by an appeal to conditional scrutability, or to understand epistemic context-dependence in some other way.
occasions of utterance: different tokens of ‘The cat sat on the mat’ will be uttered different speakers on different occasions. Sentence tokens are sometimes understood as inscriptions (concrete objects such as ink patterns or sound waves), but this model has the disadvantage that there will often be multiple tokens per utterance. For our purposes, it is best to understand sentence tokens as ordered pairs of consisting of a sentence $S$ and an utterance of $S$. Or better, we can understand a sentence token as an ordered pair consisting of a sentence $S$ and a context of utterance of $S$, where such a context of utterance is a centered world centered on an utterance of $S$.

Equivalently, we can relativize scrutability of a sentence type $S$ to a context (again, a centered world centered on an utterance of $S$). For example, ‘Neptune is a planet’ might be scrutable from a given base relative to Leverrier’s context but not relative to a later context. In practice I will move back and forth between talk of sentence tokens, utterances, and sentences in contexts. Sentence tokens and sentences in contexts are identical as defined, and they correspond one-to-one to utterances (in worlds). For ease of discussion, I will sometimes speak of a sentence token $S$ rather than of a token of a sentence $S$, and will speak of a token expressing a thought when the corresponding utterance expresses a thought.

We can first define knowledge of sentence tokens or sentences in contexts. A subject who utters a sentence $S$ knows the relevant token of $S$ (or knows $S$ relative to the context of utterance) if the utterance of $S$ expresses a thought that constitutes knowledge. More generally, a subject who utters a sentence $S$ at time $t_1$ knows the relevant token of $S$ (or knows $S$ relative to the context of utterance) at time $t_2$ if the utterance of $S$ at $t_1$ expresses a thought that constitutes knowledge at $t_2$. (Persistence of thoughts over time is discussed later in this excursus.) We will not need to characterize knowledge of a sentence token by a subject other than the original speaker.

Different models of contexts and context-dependence are available. On a standard view of extensional context-dependence, contexts involve sets of parameters (speakers, times, salient individuals, standards of strictness, and so on), and the extension of an expression depends on some formal way on those parameters. It is not obvious how to generalize this model to epistemic context-dependence, however. We could suppose that there are some parameters such that the epistemic role of an expression depend on those parameters. But it is not obvious what the parameters are and how to represent epistemic roles. One might take Fregean senses themselves to be parameters (and/or to be epistemic roles), but then we have gained little over working directly with Fregean propositions. The centered-world understanding of context has the advantage that the world itself fixes all the features of context. Restricting these centered worlds to worlds in which the speaker at the center is uttering the sentence in question allows the association between utterance and thought to fix an epistemic role and consequently to fix facts about scrutability. I do not rule out other models of context for analyzing scrutability, but the centered-world model will suffice for our purposes here. (It should be noted that the use of centered worlds to represent contexts is quite distinct from their use to represent epistemically possible scenarios.)
Corresponding definitions can be given for a priori knowledge and for belief. For our purposes, these definitions need not be restricted to fully competent utterances; they also apply to deferential and incompetent utterances.

We can then define scrutability for sentence tokens or sentences in contexts. For example: a token of $S$ is inferentially scrutable from a class $C$ of (context-independent) sentence types (or: $S$ is inferentially scrutable from $C$ relative to the context $c$ of utterance) iff, were the speaker to know all the true sentences in $C$, they would be in a position to know the token of $S$ (or: to know $S$ relative to $c$). The corresponding inferential scrutability thesis will then say: there is a class of (context-independent) sentences $C$ such that all sentence tokens are scrutable from $C$. I discuss the (empirical, conditional, a priori) scrutability of sentence tokens in more detail later in this excursus.

This scrutability thesis for sentence tokens is much weaker in some respects than a scrutability thesis for sentence types or for propositions. The thesis is restricted to tokens of sentences that are uttered in the actual world. Because of this, the less that is said in the actual world, the easier that it is for the thesis to be true! Still, this thesis is strong enough to be applied to all actual utterances, say for the purposes of defining content. For a stronger thesis, one can require that the thesis holds not just in our world, but in all nearby worlds, and perhaps in all nomologically possible worlds (all worlds with the same fundamental properties and laws as ours). We could in principle extend it to all metaphysically possible worlds (including those with alien properties and laws), but such a strengthening would be stronger than is needed for initial purposes.

In any case, we still have scrutability theses for sentence types to capture the central intended force. Theses for sentence tokens can be thought of as auxiliary devices to handle the application to utterances of context-dependent sentences.

A final issue arises from extensional context-dependence. Where this arises in the dependent class, we can handle it as above. However, a residual issue concerns extensional context-dependence. Someone might object that whether an expression is context-dependent will depend on what counts as the content of an utterance of an expression, which depends on one’s theory of propositions. A Fregean theory may count ‘Neptune’ as context-dependent where a Russellian theory does not, for example. However, the sorts of context-dependence that matter here are extensional and epistemic context-dependence. These notions are not defined in terms of propositional content, and are available irrespective of whether one is Fregean or Russellian. Both views can utilize the notion of an expression’s extension in a context, and both views can utilize the notion of an expression’s epistemic role in a context. It is true that given a Fregean view, epistemic context-dependence will correspond at least roughly to context-dependence of Fregean content (likewise for extensional context-dependence and Russellian content). But the notion of epistemic context-dependence does not presuppose a Fregean theory of content.
dependence in the base. Base sentences typically need to contain the indexicals ‘I’ and ‘now’. These sentences are not true and false absolutely (in the actual world): ‘I am hungry now’ can be true relative to one subject at one time, and false relative to another subject at another time. For this reason, we cannot require that base sentences be true absolutely. Rather, they must be true or false relative to a subject at a time. Given that our key theses concern the scrutability of truths from truths, we need to allow the relevant class of base truths to be different for different subjects and different times. On a moment’s reflection, this is just what one should expect, if the base truths include a specification of the subject’s place in the world.  

We might try formulating the thesis by saying: for all subjects s at all times t, there is a compact class of truths (true relative to s and t) such that all truths (relative to s and t) are scrutable (for s at t) from that class. This works when the dependent truths are sentence types, but if the dependent truths are restricted to sentence tokens, this formulation threatens to trivialize the thesis. It may be that every true sentence token is scrutable from a single truth (namely itself, or its counterpart in the base language). If so, then as long as subjects can produce only a finite number of tokens at one time, then for any s and t, there will be a finite class of truths from which the true tokens (relative to s and t) are scrutable.

In light of all this, one can formulate the thesis more generally (for both sentence types and tokens) by saying that there is a compact class C of sentences such that for all subjects s and all times t, all sentences that are true relative to s and t are scrutable (for s at t) from the subclass of sentences in C that are true relative to s and t. If dependent truths are restricted to sentence tokens, the thesis can be put more simply by saying that there is a compact class C of sentences such that every true sentence token S (uttered by subject s at time t) is scrutable (for s at t) from the subclass of sentences in C that are true relative to s and t. We can also extend this thesis to all nomologically possible sentence tokens, as suggested above, if we also relativize truth to the world of utterance.

In what follows, for simplicity, I will typically suppose that we have chosen an (arbitrary) subject and time, and will allow arbitrary sentence types or sentence tokens to be in the dependent class. (When those tokens have not actually been uttered, we can adopt the fiction that they have been uttered.) Then we can consider the question of whether there are sets of base truths from

---

24 Note that this relativization of truth of sentences to subjects should be distinguished from relativization of scrutability of sentences to subjects. The discussion in chapter 2 suggests that empirical and conditional scrutability of a sentence from a class of sentences is always relativized to a subject and a time, and that a priori scrutability can be so relativized but need not be so relativized.
which all truths are scrutable; the relativization to the subject will often be left explicit.²⁵

Scrutability of sentence tokens and the persistence of thoughts.

A residual issue concerns the scrutability of sentence tokens from sentence types. In defining scrutability of a sentence type $S$ from a class of sentence types $C$, we have used notions such as:

- if the subject came to know the sentences in $C$, they would be in a position to know $S$ (inferential scrutability);
- the subject is in a position to know that if $D$, then $S$ (conditional scrutability); and
- $D \rightarrow S$ is a priori (a priori scrutability). However, it is not obvious how to apply these definitions to sentence tokens, or to sentences in contexts.

One option here involves understanding these notions in terms of warrant, as discussed below: for example, being in a position to know a sentence token $S$ would involve there being a warrant for the thought expressed by $S$. This option will not raise the issues that follow. However, we may also want to understand these notions in broadly modal terms. A modal account needs further analysis.

In the case of inferential scrutability, we want to analyze what it is for a speaker to be in a position to know a sentence token $S$, both at the time of uttering $S$ and at certain later times (such as after coming to know truths in $C$). The natural suggestion, following the modal account for sentence types, is that at time $t$, the speaker is in a position to know $S$ iff it is possible that they later come to know $S$, starting from their position at $t$ and without acquiring further empirical information.

This analysis invokes the notion of knowing a sentence token $S$ at a time after the time of

²⁵Another worry about extensional context-dependence is that knowledge takes time, and that an indexical sentence type such as $PQT_{I}'$ (the conjunction of sentences in $PQT_{I}$) or ‘It is now $t$’ that is true at $t$ will not be true after $t$, so cannot be known then. This worry especially affects Inferential Scrutability: here we might modify our understanding of “in a position to know $S$ at $t'$ to require not possible knowledge at a later time $t'$ of thoughts apt to be expressed by $S$ at $t'$, but instead possible knowledge at $t'$ of thoughts apt to be expressed by $S$ at $t$. This requires an appeal to the persistence of thoughts, as discussed in the final section. The worry is not so bad for A Priori Scrutability, as long as we recognize that conditional sentences such as ‘$PQT_{I}' \rightarrow \text{It is now } t'$ can come to be known a priori at times $t'$ after $t$. It is a delicate question whether this should require a thought at $t'$ (constituting a priori knowledge) that is apt to be expressed by the conditional at $t'$ or at $t$; I think there is a case for the latter (so that the relevant thought in effect involves \text{It was then} ... rather than \text{It is now} ... on both sides), but either claim is plausible. The issue does not arise for Conditional Scrutability, which is understood normatively rather than modally. Perhaps this is another point in favor of the view that normative or warrant-based understandings of apriority and the like are more basic than modal understandings.
utterance. To define that notion, we can say that the speaker of a sentence token \( S \), uttered at time \( t_1 \), knows \( S \) at some later time \( t_2 \), if the thought expressed by \( S \) at \( t_1 \) constitutes knowledge at \( t_2 \). As discussed shortly, this requires reidentification of thoughts over time. Here the idea is that there is some possible process of reasoning that starts from the subject’s thought at \( t \) and justifies that thought, ending with the thought constituting knowledge. For the analysis of being in a position to know, the reasoning process need not involve a priori reasoning alone, as it can use any empirical beliefs that the subject has at the time of utterance, but it cannot involve new empirical discovery.

This analysis yields a straightforward analysis of inferential scrutability: a sentence token \( S \) is inferentially scrutable from \( C \) iff there is some subclass of \( C \) such that if the subject came to know the sentences in that subclass, they would be in a position to know \( S \). What about a priori and conditional scrutability?

We can say that a thought is a priori iff it is possible that on (perhaps idealized) reflection, the thought comes to constitute a priori knowledge. A sentence token \( S \) is a priori (equivalently: one is in a position to know \( S \) a priori) iff it expresses an a priori thought. We can then say that a sentence token \( S \) is a priori scrutable from a class \( C \) when there is a possible thought \( T' \), apt to be expressed by some conjunction of sentences in \( C \), such that \( T' \) implies \( T \), where \( T \) is the thought expressed by \( S \). For thoughts \( T_1 \) and \( T_2 \) had by the same subject, \( T_1 \) implies \( T_2 \) when a disjunction of \( T_2 \) with a negation of \( T_1 \) is a priori.

Likewise, \( S \) is conditionally scrutable from \( C \) when \( T' \) rationalizes \( T \), where \( T \) is the thought expressed by \( S \) and \( T' \) is the thought one would have if one harmlessly entertained the conjunction of all sentences in \( C \). For thoughts \( T_1 \) and \( T_2 \) had by the same subject, \( T_1 \) rationalizes \( T_2 \) when a conditional thought in \( T_2 \) given \( T_1 \) is rational, according to the standards of ideal insulated rationality.

These definitions invoke certain relations between thoughts. One thought can be a negation of another, intuitively when it is formed by negating the former thought. One thought can be the conjunction of some others, or the disjunction of some others, intuitively when it is formed by

\[ cr'(S_1 | S_2) = cr'(T_1) \]

\[ cr'(S_1 | S_2) = cr'(T_1 | T_2) \]

These claim about non-insulated rationality is not quite right: for example, if \( S_1 \) is ‘I am entertaining \( S_2 \)’, then \( cr'(S_1 | S_2) \) may intuitively be low while \( cr'(T_1 | T_2) \) may be high. (When thinking \( T_2 \), one will entertain \( S_2 \) and will plausibly thereby be in a position to know that one is entertaining \( S_2 \).) Fortunately this problem does not arise for the insulated rational credence \( cr'(S_1 | S_2) \), which is what is relevant to rationalization and conditional scrutability, because here \( cr'(T_1 | T_2) \) will be low due to the barring of introspection. To handle non-insulated rational credences, one can define \( cr'(S_1 | S_2) \) in terms of there being warrant for having a certain credence \( cr(S_1 | S_2) \), as discussed at the end of the fourth excursus.
conjoining or disjoining those thoughts. A conditional thought can be the conditionalization of one thought on another, intuitively when it is formed by accepting the latter conditional on the former. I think we have a clear intuitive grasp of these notions. Given the notion of persistence in the next paragraph, all we really need here is a synchronic relation: the idea that a thought can be a negation of a simultaneous thought, or a disjunction, a conjunction, or a conditionalization of two simultaneous thoughts.

These definitions presupposes that thoughts can be reidentified over time, or that thoughts can persist over time. The notion of persistence enters into the analysis of being in a position to know when we say “the thought expressed by $S$ at $t_1$ constitutes knowledge at $t_2$”. It enters into the analysis of apriority of a thought when we say that it is possible that a thought “comes to constitute a priori knowledge”. It also enters tacitly into the analysis of a priori and conditional scrutability in that the relevant conditional or disjunctive thoughts involving $T$, the thought expressed by $S$, may arise later than the original time of utterance of $S$.

This notion of persistence over time might give rise to objections. What is it for a thought at $t_1$ to persist as a thought at $t_2$? One might suggest that this is simply for the thought to have the same content. But then one’s conclusions about potential knowledge and apriority will be hostage to one’s theory of content, and it is not clear that they can then be used to ground a theory of content.

However, persistence should not be understood as sameness of content. First: On a fine-grained Fregean view of content, persistence may not require sameness of content. For example, my thought that I am hungry now might persist as my thought that I was hungry then, which arguably has a different fine-grained Fregean content.

Second: Even on a Russellian view of content, cases of semantic fragility suggest a possible gap between persistence and sameness of content. Suppose that I express a thought $T$ with ‘Snow is white iff actually snow is white’, but I do not reflect on $T$ or attempt to justify it. Then $T$ does not express a priori knowledge, but it is natural to say that $T$ is justifiable and expresses potential a priori knowledge. But if I had followed through, the content of the ensuing a priori knowledge would have been $q i f f q – i n – w$, whereas the actual content of $T$ is $q i f f q – i n – @$. So it appears that $T$ is justifiable in virtue of a later possible thought with a different content being justified. One way to put this is to say that $T$ could have persisted as a later thought with a different content. If we say this, we also need to say that $T$ itself could have had the different content $q – i f f – q – i n – w$. This leads to a picture on which the relevant thoughts are themselves semantically fragile: if one had investigated them, they would have had different contents. So one might take these cases to motivate a view on which thoughts do not have their Russellian contents essentially.
The alternative for a Russellian is to say that if one had engaged in the relevant investigation, one would not have had \( T \) at all, but instead would have had a different thought with a different content. This preserves the individuation of thoughts by contents, but makes it harder to associate properties such as justifiability with the relevant thoughts. On this view, \( T \) is not justifiable at all by the process in question: only a counterpart thought will be justified. To capture the phenomenon, one will end up saying something like: a thought is quasi-justifiable iff it is possible that a counterpart thought is justified. The upshot for present purposes will be much the same as on the previous view: to understand the phenomena of justifiability and potential a priori knowledge, one needs to appeal to a relation between a thought and possible future thoughts that does not require sameness of content.

Third: on almost any view of content, sameness of content does not entail persistence: if I think that \( p \) at \( t_1 \), and at \( t_2 \) I have a causally independent thought that \( p \), then the latter does not persist as the former. This applies even on a Fregean view, but is particularly clear on a Russellian view of content: when Sue has thought \( T_2 \) (expressed by ‘Phosphorus is a planet’) shortly after \( T_1 \) (expressed by ‘Hesperus is a planet’), these thoughts may have the same content, but \( T_1 \) certainly does not persist as \( T_2 \).

Persistence requires an appropriate continuity between thoughts over time. It is plausible that this continuity requires some sort of relatedness of content, perhaps involving common or related guises, but this relatedness of content does not suffice for persistence. Causal and psychological continuity is also required. I will not attempt to define the notion of persistence here, but will leave the notion unanalyzed. For present purposes, the most essential applications of the notions of persistence (in the case of a priori scrutability) can be restricted to occurrent thoughts within a single brief stream of thought: we can imagine an idealized thinker entertaining the thought, and coming moments later to justify it.

Most importantly, persistence is an intuitive notion that everyone needs, whatever they think about the theory of content. Anyone who believes in thoughts should allow that a thought can come to be justified, or that it can come to be confirmed by evidence. Making sense of these notions requires the notion of persistence. The notion, along with the related notions of negation, disjunction, and conjunction in thought, are also crucial to understanding the notion of inference in thought. Suppose one reasons: \( A, B \), therefore \( A \& B \). For this to be a valid inference in thought, conferring immediate justification of the conclusion, one’s initial thought that \( A \) must be appropriately related to one’s later thought \( A \& B \), intuitively acting as a conjunct of that thought. If one formed an independent thought with the same content, then this thought would not acquire the
same sort of immediate justification. So an appeal to the notion of persistence does not presuppose commitment to any theoretical account of content, or to any technical notions such as apriority. The notion is already manifest in our ordinary notions of justification, confirmation, and inference in thought.

**Sentential vs. propositional scrutability**

There are at least three potential differences in strength between sentential and propositional scrutability theses. One potential differences arises from a course we have considered already: differences between theories of propositions. We have seen that on Fregean accounts, there will plausibly be a close link between sentential scrutability as defined here and propositional scrutability: roughly, a sentence $S$ will be scrutable from a class $C$ of sentences iff the proposition expressed by $S$ is scrutable from the class of propositions expressed by sentences in $C$ (though see below for limitations of scope and exceptions). On a Russelian or possible-worlds theory of propositions, however, there will not be such a close parallel. For example, ‘Hesperus is a planet’ is not a priori scrutable from ‘Phosphorus is a planet’, but the associated Russelian propositions are identical, so the first is a priori scrutable from the second.

On a Russelian or possible-worlds account, one might bring propositional scrutability theses into a closer alignment with sentential scrutability theses by replacing propositions in the scrutability thesis by proposition/guise pairs. Then it is attractive to hold that a sentence (type or token) is scrutable from a class of sentences if the corresponding proposition is scrutable from the corresponding class of propositions, under the guises associated with the sentences.

A second difference arises from the phenomenon of *inexpressibility*: propositions that cannot be expressed by sentences. If there are inexpressible propositions, then it may happen that all expressible propositions are scrutable from a certain base but some inexpressible propositions are not. If so, an unrestricted sentential scrutability thesis may be true: sentences express only expressible propositions, so all sentences are scrutable. But an unrestricted propositional scrutability thesis is false: some inexpressible propositions are inscrutable. If we grant that sentences are scrutable from each other iff the propositions they express are scrutable from each other (as perhaps on a Fregean view of propositions), sentential scrutability theses are akin to propositional scrutability theses restricted to expressible (or entertainable) propositions. A thesis cast in terms of guise/proposition pairs may in effect have the same restriction, at least if we assume that there
are no guises associated with unentertainable propositions.\textsuperscript{27}

For example: on some views, there are certain properties that cannot be grasped or expressed: the intrinsic quiddities of matter discussed in chapter 7, for example. If these properties are involved in propositions, then on this view there may well be propositions that are not expressible or thinkable. These propositions may then yield counterexamples to propositional scrutability without corresponding counterexamples to sentential scrutability.

Another inexpressibility issue will arise if there are propositions expressible by sentence tokens but not expressible by any context-independent sentence types. For example, suppose that object-involving propositions can only be expressed using ordinary proper names and that ordinary proper names are always epistemically context-dependent. Then object-involving propositions fall under the scope of a propositional scrutability thesis, but no corresponding sentence falls under the scope of a sentential scrutability thesis for context-independent sentence types. On my own view, this case cannot arise: for any sentence token there is some possible epistemically invariant sentence type all of whose tokens behave like the original token (perhaps by stipulation). Even on other views, a token expressing the proposition in question will at least fall into the scope of scrutability theses cast in terms of sentence tokens. But the case is worth keeping in mind.

A third difference arises from the phenomenon of \textit{semantic fragility}. A sentence $S$ is semantically fragile when investigating whether $S$ is true can change the proposition that $S$ expresses.\textsuperscript{28} One example is ‘Snow is white iff actually snow is white’. On one common view, this sentence $S$ expresses a proposition $p$ involving the actual world @. Now, suppose that no one actually ever investigates $S$. Then it is still possible that someone come to know that $S$ is true. But if they did so, $S$ would express a proposition different from the one that it actually expresses. It would express a proposition $p'$ involving a different world $w$: the world in which $S$ is investigated and known. For example, $p$ may be the proposition $q$ iff $q$-in-@ while $p'$ is the proposition $q$ iff $q$-in-$w$, where $q$ is the proposition that snow is white.

It is natural to say that in this process, the subject would come to know $S$. But they would not

\textsuperscript{27}If graspability of Fregean propositions varies between subjects, as on Frege’s own view of first-person propositions, one might also want to relative to subjects and times, saying: for all subjects $s$ and times $t$, all true propositions graspable by $s$ at $t$ are scrutable from a compact class of propositions graspable by $s$ and $t$. On a Russellian view, one could do much the same for guise-proposition pairs.

\textsuperscript{28}Semantic fragility is roughly analogous to alethic fragility, discussed in section 3. In the former, investigating $S$ changes the proposition it expresses, while in the latter, investigating $S$ changes its truth-value. The Fitch sentence is alethically fragile without being semantically fragile. ‘Snow is white iff actually snow is white’ is semantically fragile without being alethically fragile.
come to know $p$, the proposition $S$ actually expresses; instead they would come to know $p'$. In these cases, we need to distinguish knowledge of a sentence from knowledge of the proposition that it actually expresses. Rather, knowledge of a sentence in a world should be knowledge of the proposition that the sentence expresses relative to that world (perhaps under the guise associated with that sentence in that world).

In section 2, I discussed two reasons for distinguishing sentential from propositional knowledge: dialectical reasons and reasons based on the argument from assertion. The phenomenon of semantic fragility constitutes a third reason. Its force is slightly different. The argument from assertion suggested that knowledge of $p$ does not suffice for knowledge of $S$, at least on coarse-grained views of propositional content. The argument from semantic fragility does the reverse: it suggests that knowledge of $S$ (in counterfactual worlds) does not suffice for knowledge of $p$, at least on object-involving views of propositional content.

Semantic fragility draws a wedge between sentential and propositional scrutability. In cases such as the above, it can happen that $S$ is scrutable from some base when $p$ is not. In fact, in the case above, it is plausible that $S$ is a priori knowable, so that it is a priori scrutable from any base. But $p$ may not be a priori knowable: if one does not actually know $S$, then any world in which one knows $S$ will be a world in which $S$ does not express $p$. So even if one comes to know $S$ a priori, one will not thereby come to know $p$ a priori. One can show (Chalmers 2011) that in some such cases, $p$ is not knowable at all. If so, then $p$ is not a priori scrutable from any base. So a priori scrutability of a sentence can come apart from a priori scrutability of the proposition it expresses.

Another example: say I introduce ‘Bigthink’ as a name for the most powerful reasoner in the world. Let us suppose that the most powerful reasoner is in fact Einstein, and let $S'$ be ‘Bigthink is German’. Assume a view on which $S'$ expresses a proposition $p'$ involving Einstein. Then my utterance of $S'$ may well be a priori scrutable from a giant world-sentence $G$, in that ‘If $G$, then $S'$’ is knowable a priori. But if I were to come to know this (extraordinarily complex) conditional a priori, I would be the most powerful reasoner in the world. If so, ‘Bigthink’ would refer to me. So in deriving $S'$ from $G$, I would not derive the proposition $p'$ (involving Einstein) that $S'$ actually expresses, but a different proposition involving me. So $S'$ is scrutable from $G$ by this method, but $p'$ is not. On a Russellian view where $p'$ is the same as the proposition that Einstein is German, $p'$ may be scrutable under a different guise, but on a Fregean view where both guises and objects are built into propositions, $p'$ may not be scrutable at all. Once again, the semantic fragility of $S'$ leads to a difference between sentential and propositional scrutability.

Semantic fragility tends to suggest that many propositional scrutability theses will be false
where corresponding sentential scrutability theses are true. For the overarching Scrutability thesis, for example, there may be no compact base of true propositions from which all true propositions are a priori scrutable: if a proposition such as \( p \) is not actually derived a priori from base sentences, it may be impossible to derive it. One might take this as a further reason to formulate scrutability theses in sentential rather than propositional terms.

Alternatively, to avoid the problem posed by semantic fragility, and to hold onto propositional scrutability theses in the spirit of the introduction one can try to understand propositional scrutability and propositional apriority in nonmodal terms: that is, not in terms of what it is possible to know, or to know a priori. An alternative nonmodal understanding in terms of warrants is developed in the next excursus.
Fourth Excursus: Warrants and Support Structures

So far, I have often cast notions such as apriority, scrutability, and knowability in modal terms: in terms of what it is possible to know, or to know a priori. We have seen that modal idealizations can lead to difficulties in some circumstances: they cause problems for propositional apriority and scrutability in cases of semantic fragility (see the end of the third excursus), and they cause other problems if there are brute modal constraints on possible reasoners (see section 7 of chapter 2). It is also arguable that modal idealizations are not explanatorily fundamental: even when a scrutability thesis involving a modal idealizations is true, it derives from more fundamental epistemological facts. So it is worth exploring nonmodal ways of understanding these notions.29

In the introduction, scrutability theses are cast in terms of what one is in a position to know. This notion can be cashed out in modal terms, but it can also be cashed out in other ways. In particular, a relevant notion of one’s being in a position to know \( p \) can be cashed out in terms of there being an \textit{warrant} for one to believe \( p \). A warrant is a conclusive justification, or a justification suitable for knowledge. There can be a warrant for one to believe \( p \) even if one does not in fact know or believe \( p \). For example, when there exists a proof of \( p \), this yields a warrant for believing \( p \) regardless of whether anyone proves \( p \). These warrants are a form of propositional justification: a justification that supports belief in \( p \) for a subject, whether or not the subject believes \( p \). This notion is standardly distinguished from doxastic justification: justification on which someone’s belief in \( p \) is based.

On one notion of propositional warrant, one says that a subject has a propositional warrant to believe \( p \) when the warrant is (in some sense) within the subject’s grasp. In this sense, the mere existence of a complex proof for \( p \) does not entail that a mathematically ignorant subject has a warrant for believing \( p \). The more relevant notion here is that of there being a propositional warrant for a subject to believe \( p \) (or more briefly, there being a propositional warrant for \( p \)), whether or not the subject has that warrant. This notion does not come with the requirement that the warrant is within the subject’s grasp. There can be a warrant for one to believe \( p \) even when knowing or believing \( p \) is beyond one’s cognitive capacities. For example, even when a proof of \( p \) is enormously complex, it yields a warrant for a mathematically ignorant subject to believe \( p \). We might call the first sort of warrant a \textit{nonideal warrant}, and the second sort an \textit{ideal warrant}.30

---

29The first few paragraphs of this excursus will make most sense in light of the last few paragraphs of the previous excursus, on semantic fragility (or alternatively, in light of “Actuality and Knowability”). Readers can either read that material first, or alternatively skip directly to “What is a warrant, exactly?” below.

30One might require that even an ideal warrant is within a subject’s idealized grasp (in some nonmodal sense), where
The framework of warrants makes a difference in cases of semantic fragility. When \( p \) is the proposition expressed by the semantically fragile sentence \( S \) discussed at the end of the third excursus (‘Snow is white iff actually snow is white’), one can argue that there exists a proof of \( p \) even though it is impossible to use it to prove \( p \). In particular, there exists an abstract proof of \( S \) using the logic of ‘actually’. \( S \) expresses \( p \) in the actual world, so this abstract proof of \( S \) is also an abstract proof of \( p \). But if one were to use the proof to prove \( S \), \( S \) would express \( p' \) rather than \( p \), so one would not prove \( p \).

What goes for proof goes also for warrant. In this case, there exists an (ideal a priori) warrant for believing \( p \) even though the warrant cannot be used to know \( p \). This warrant is a propositional warrant that cannot be used as a doxastic warrant. If the subject is sophisticated and the proof is easy, the subject may even have a (nonideal a priori) warrant for believing \( p \), even though the subject cannot use that warrant to believe \( p \). This last issue depends on precisely one understands the conditions for having a (nonideal) warrant, but either way, there will certainly be a priori warrant for believing \( p \). So there can be an a priori warrant for believing \( p \) even when it is not possible to know \( p \) a priori.

This suggests a nonmodal conception of propositional apriority: \( p \) is a priori in the nonmodal sense when there is an a priori warrant for some subject to believe \( p \). Propositions expressed by semantically fragile sentences such as ‘\( S \) iff actually \( S \)’ above may be a priori in this sense even if they are not a priori in the modal sense. Likewise, this suggests a nonmodal conception of knowability: \( p \) is knowable by a subject in the nonmodal sense if that subject has warrant for believing \( p \).

One can define nonmodal apriority and knowability for sentences in a similar way: \( S \) is nonmodally a priori if there is an a priori warrant for some subject to believe \( S \) (in the sense familiar from the third excursus), while \( S \) is nonmodally knowable for a subject if there is a warrant for that subject to believe \( S \). Unlike the propositional versions, these nonmodal notions of sentential apriority and knowability will not come apart from modal apriority and knowability in cases of semantic fragility. They may still come apart if there are brute modal constraints on possible thinkers, however. For example, if no possible thinker can carry out a proof that involves more than a million steps, then certain true mathematical sentences (and propositions) will be a priori in the modal sense but not in the nonmodal sense.

the ideaization involves idealization of reasoning. So nonideal and ideal warrants could both be seen as providing “subjective” reasons (reasons for a subject that are in some sense available to the subject), although the latter is a highly idealized variety. Thanks to John Bengson, Jon Kvanvig, Nico Silins, and Chris Tucker for discussion of warrant.
One can also appeal to warrant to define nonmodal notions of scrutability. For example: $q$ is inferentially scrutable from $p$ when knowing $p$ would provide a warrant for $q$; $q$ is conditionally scrutable from $p$ when one has a warrant for accepting if $p,$ then $q;$ and $q$ is a priori scrutable from $p$ when there is an a priori warrant for accepting if $p,$ then $q.$ These definitions are unaffected by semantic fragility. In the cases from the third excursus on which a semantically fragile sentence $P$ is scrutable from $Q$ on a modal definition of scrutability, the corresponding proposition $p$ is not scrutable from $q$ on a modal definition of propositional scrutability, but $p$ will plausibly be scrutable from $q$ on the nonmodal definition of propositional scrutability.

I often cast notions such as apriority and scrutability in modal terms elsewhere in this book, in part because modal analyses of notions such as apriority are more familiar than those in terms of warrant, and partly because the problem of semantic fragility does not affect sentential scrutability. But it is reasonable to hold that the warrant-based notions are more fundamental than the corresponding modal notions. When it is possible to know $p$ a priori, this is typically possible because there is an a priori warrant for $p.$ Likewise, when $S$ is scrutable from $C$ in a modal sense, this is typically because $S$ is scrutable from $C$ in a warrant-based sense.

What is a warrant, exactly? A warrant is plausibly a sort of justification: a justification suitable for knowledge. But what is a justification? Various understandings are possible, but here I will develop an understanding of warrants and justifications as support structures. This understanding is inspired by the special case of proof. When belief in $p$ is warranted by a proof, the warrant for $p$ will be the structure of propositions in the proof, with support relations corresponding to logical steps. We can count this structure as a support structure. More generally, a support structure will involve directed graphs of labeled propositions with support relations between them, capturing the justification for a proposition.

It is more common in the epistemological literature to view warrants and justifications as propositions or perhaps as sets of propositions. But these propositions only play their role in virtue of their position in a support structure, and for various purposes it is useful to make this structure explicit. Furthermore, the case of a proof suggests that there can sometimes be a warrant (in an intuitive sense) for someone to believe a proposition even though there is no clear proposition or set of propositions that constitute the warrant (the proposition expressed by the first step of the proof? the penultimate step? all thrown together?). Viewing warrants as support structures avoids these problems. This need not involve a substantive disagreement with those who view warrants as propositions: to be fully explicit we could call these structures “warrant structures” instead of warrants, and likewise for justifications. The details of this framework are not crucial.
for addressing the problem of semantic fragility, but they will play a role for other epistemological purposes in chapters 3 and 4. So I will develop the framework in some detail.

We can start with doxastic justification. Take any justified belief that \( p \). Something (a belief or an experience, for example) \textit{justifies} belief in \( p \) when it supports \( p \) or provides evidence for \( p \) in a way that yields prima facie justification for the belief that \( p \). It \textit{directly justifies} belief in \( p \) when it justifies belief in \( p \) and does not justify belief in \( p \) wholly in virtue of justifying something else that justifies belief in \( p \).\(^3\)

A \textit{direct justification} for \( p \) can be represented as a graph consisting of a node for \( p \) along with nodes for elements (if any) that directly justify belief in \( p \), with arrows from those nodes to \( p \). For example, when a belief in \( p \) is justified inferentially, it will be directly justified by inference from one or more other justified beliefs: a belief that \( q \), a belief that \( r \), and so on. Here we can say that these other beliefs, collectively, provide a direct justification for the belief that \( p \). A direct justification can here be represented as a node for \( p \) with nodes for \( q \), \( r \), and so on, with arrows from them to \( p \). If the justification is redundant, so that the belief that \( q \) suffices on its own to inferentially justify the belief that \( p \), there will also be a direct justification consisting of a node for \( q \) with an arrow to a node for \( p \).

When a belief in \( p \) is justified noninferentially, it will either be directly justified by some evidence distinct from the belief (e.g. perceptual evidence) or it will be justified by no such evidence (as on some views of basic belief). In the first case, we can say that \( p \) is justified noninferentially and evidentially. Here a direct justification for the belief that \( p \) can be represented as a node for \( p \) with arrows from nodes for any directly justifying evidence. In the second case, we can say that the belief that \( p \) is justified noninferentially and nonevidentially. Here a direct justification for \( p \) can be represented by a node for \( p \) alone, with \( p \) labeled as a basic belief. If the belief that \( p \) is self-justifying (if this is possible), the graph will include an arrow from \( p \) to itself. If there are beliefs that are justified both inferentially and noninferentially, then both sorts of support can be included in the structure; though as above, in cases where the justification is redundant, there will

\(^3\)The justification relation is an epistemic grounding relation, and should be distinguished from the metaphysical and conceptual grounding relations discussed in the excursus on grounding and elsewhere. For something to stand in this relation to \( p \) is not for it to be the metaphysical grounds for the belief that \( p \) or the metaphysical grounds for the justification of the belief that \( p \). For example, on a reliabilist view, a reliable process might serve as metaphysical grounds for the justification of a basic belief that \( p \), but this process will not itself stand in the epistemic grounding relation to the belief that \( p \). When a belief that \( p \) directly justifies a belief that \( q \), however, it is plausible that their standing in this support relation (or the metaphysical grounds of their standing in this relation) will serve as part of the metaphysical grounds for the justification of the belief that \( p \).
also be direct justifications that exclude redundant elements.

An *indirect* justification for *p* will include a direct justification for *p* and also justifications for one or more elements in the structure that supports *p*. A *justification* for *p* is a direct or indirect justification for *p*. A *full* justification is a justification that includes a justification for every belief in the structure. A *partial* justification is a justification that is not a full justification.\(^{32}\)

For example, when belief in *p* is inferentially justified by belief in *q*, and belief in *q* is inferentially justified by belief in *r*, and belief in *r* is noninferentially justified by evidence *e*, a direct justification will include only a link from *q* to *p*. A full justification *p* will include links from *e* to *r* to *q* to *p*. There will also be an indirect partial justification with links from *r* to *q* to *p*.

If there can be circles of justificational support (e.g. from *p* to *q* to *r* to *p*), then justifications can include these circles. When there can be infinite chains of support, then justifications can include these chains. When there are no such circles or infinite chains, we can say that a justification is *classical*. The *grounds* of a classical justification are its initial elements. When a belief that *p* has a classical justification (whether partial or full), we can say that the grounds of that justification *ground* belief in *p*. When a belief that *p* has a full classical justification, we can say that the grounds for the justification *fully ground* belief in *p*. The grounds of a full classical justification will be *basic evidence*: these may include basic beliefs and/or basic nondoxastic evidence.\(^{33}\)

I will adopt a model on which evidence always involves propositions. So introspective evidence might involve the proposition that one is in a given mental state: for example, the introspective evidence for the belief that one is in pain may be the proposition that one is in pain. Likewise, perceptual evidence may involve the propositional contents of perceptual experience: for example, the perceptual evidence for the perceptual belief that there is something red in front of one may be the (perceptually represented) proposition that there is something red in front of one. Something similar may go for evidence provided by (for example) intuition, memory, and testimony, if one thinks that these involve sources of noninferential justification distinct from perceptual or

\(^{32}\)Note that a partial justification is partial in the sense that it omits some elements that play a justifying role, but not in the sense that it yields support that merely weighs in favor of *p* without justifying belief in *p*. I am taking it that all of the justifications I consider here are strong enough to justify belief in *p*. The elements omitted in a partial justification will either be redundant elements (whose contribution is not required for justification) or indirectly justifying elements (whose contribution is mediated by another element).

\(^{33}\)Even when there is no full classical justification for *p*, there may still be a partial classical justification for *p*. For example, when full justification involves an infinite chain, there will always be a partial justification without such a chain. Even when full justification involves a circle, it may be that there is a partial justification without a circle.
introspective evidence.

To distinguish the different roles for various propositions here, the structure will label a proposition \( p \) as a belief proposition (\( p \) is believed), perceptual evidence (\( p \) is the the content of a perceptual states), introspective evidence (\( p \) is introspectively experienced, or perhaps need only be true), and so on. So in the case above, the justification might look like: \( e \) (experienced) \( \rightarrow p \) (believed) \( \rightarrow q \) (believed) \( \rightarrow r \) (believed).\(^{34}\) This can be read as saying that experience as of \( e \) justifies belief in \( p \), which justifies belief in \( q \), which justifies belief in \( r \). One could also represent this by adopting a model on which the nodes are all mental states, such as beliefs that \( p \) or experiences as of \( e \), but the propositional model provides a more natural way of representing propositional justification, where the justification may exist even though no-one believes that \( p \). The model can also be easily adapted to views on which some of the relata are facts (or perhaps property instances or other entities) rather than propositions.

For completeness, this structure should be elaborated in various ways. One could add more structure to support relations, indicating for example that \( p \) and \( q \) conjointly support \( p \& q \lor r \) while \( r \) supports it separately. One can also allow support relations themselves to be supported by evidence. As I discuss in chapter 4, for example, an inferential relation between \( s \) and \( t \) can itself be grounded in prior perceptual evidence \( e \). This can be represented by an arrow from \( e \) to the arrow between \( s \) and \( t \). It is even possible to allow that experiences can sometimes be supported, as I discuss briefly in chapter 3.

A propositional justification is the same sort of item as a doxastic justification, with the difference that there can be a propositional justification for a subject to believe \( p \) even without the subject believing the belief propositions in the justification. It is still plausibly required that for a support structure to be a propositional justification to believe \( p \), the subject must have (or at least have had) perceptual states corresponding to perceptual evidence and mental states corresponding to introspective evidence.\(^{35}\) A propositional justification yields a doxastic justification when the subject has all the relevant beliefs, and when the beliefs are properly based on each other and on the evidence in a way that reflects the support relations in the structure. A (propositional or doxastic) justification is also a warrant when it meets certain further conditions to make it suitable for knowledge for the subject. These arguably include the conditions that all the included propositions

\(^{34}\)I am using “evidence” in a broad sense on which all epistemic grounds count as evidence, as opposed to a sense on which only justifiers that are true or that are known count as evidence. Whatever one says about “evidence”, I think it is plausible that false propositions can serve as direct justifiers: for example, a belief that \( p \) or an experience as of \( p \) can justify a belief that \( q \) whether \( p \) is true or false.
be true, that any initial belief propositions (especially in a doxastic warrant) be known, and that there are no defeaters and no Gettier circumstances. These conditions might be varied or extended, but I will stay neutral on the precise conditions required.36

The model has a foundationalist flavor to it, but it does not presuppose foundationalism. As long as a view acknowledges the distinction between inferential and noninferential justification (even if it holds that all justification falls on one side), the model will be coherent. Coherentist and infinitist views will allow nonclassical justifications, and may or may not give a role to basic evidence. Reliabilist and other externalist views may sometimes find noninferential justification where other views find inferential justification, and may or may not give a role to noninferential evidential justification. (If a view does not recognize the notion of noninferential evidential justification, we can count it as classifying all noninferential justification as nonevidential.) Speaking for myself, I think it is enormously plausible that there is much inferential justification and much noninferential evidential justification, so I think that full justifications will often be quite complex.

We can use this framework to help analyze the distinction between a priori and a posteriori justification. At least among full classical justifications, an a posteriori justification will be one with some empirical grounds, while an a priori justification will be one with no empirical grounds. Empirical grounds will include all perceptual and introspective evidence, and perhaps other basic evidence depend on one’s views. If one accepts (as I do) that all basic empirical evidence is perceptual or introspective evidence, we can say more simply that a justification is a posteriori if its grounds include perceptual or introspective evidence. If one holds that there are other sorts of

36 What is the relation between these notions and the standard notions of subjective and objective reasons for belief? Let us say that a support structure is a justification freed of the requirement that the subject have states corresponding to the relevant perceptual and introspective evidence. Then a support structure provides subjective reason to believe \( p \) roughly when the subject either has mental states corresponding to the initial elements in the structure or perhaps when the subject is in a position to have them (at least where these elements correspond to a priori beliefs). A support structure provides objective reason to believe \( p \) when all the initial elements are true. All of the justifications I have talked about yield at least subjective reasons (although these may be idealized subjective reasons in some cases). Warrants yield both subjective and objective reasons. Here support structures are roughly analogous to valid arguments, subjective reasons are roughly analogous to valid arguments where the premises are justifiably believed, objective reasons are roughly analogous to sound arguments (valid arguments where the premises are true), and warrants are roughly analogous to sound arguments where the premises are known. (The fact that experiences are not objects of justification complicates the analogy, however, as does the fact that not every argument transmits justification.) One might also develop a more general notion of a “basing structure”, analogous to an argument (whether valid or invalid): such a structure might reflect only the basing (or potential basing) of certain beliefs on others, whether or not this basing goes along with justification.
basic empirical evidence, one will need further criteria for classifying basic evidence as empirical or nonempirical. If there are full nonclassical justifications, one will need further criteria to classify these as a priori or a posteriori; here the existence of an empirical ground will serve at least as a sufficient condition for such a justification to be a posteriori.

To analyze a priori doxastic justification, we can then say that a subject’s belief is justified a priori if it has an a priori doxastic justification. To analyze a priori propositional justification, we need only invoke the idea that there is an a priori justification for a subject to believe a proposition. We can make parallel claims about a posteriori justification, and about a priori and a posteriori warrant.

Other analyses of warrant are possible. There may be analyses of propositional warrant on which there is no such thing as a warrant per se, but simply a relation between a subject $s$ and a proposition $p$ that is misleading labeled “there is a warrant for $s$ to believe $p$. We can straightforwardly extend the current analysis to warrant for sentences and for thoughts. One could extend the current model to analyze cases in which prima facie justification is defeated or in which support is only partial. Finally, one might extend it to analyze justification for having certain credences in propositions (or sentences or thoughts), based on credences in other propositions (or sentences or thoughts) and evidence. The last analysis can yield a warrant-based analysis of rational credences, helping to avoid the problems for modal analysis of credences discussed in the previous excursus.