Chapter 3

EMPIRICISM AND THE A PRIORI

My aim in this chapter is develop an empiricist account of *a priori* knowledge and to defend it against objections raised by rationalists and anti-rationalist critics such as W. V. O. Quine. Since empiricists famously regard *a priori* truth as analytic, I shall offer here a clarification and defense of analytic truth. The clarification is needed because the upshot of Quine's influential criticism was that, for all its apparent reasonableness, a distinction between analytic and synthetic statements has not yet been satisfactorily drawn. The idea that such a distinction can be satisfactorily drawn is, he once said, a "metaphysical article of faith." My aim here is not only to draw a satisfactory distinction but to show that the notion of analytic truth, suitably clarified, provides a reasonable explanation of how a priori truths can have the universality and necessity that they are traditionally supposed to have.

Quine's Critique of Analytic Truth

In "Two Dogmas of Empiricism" Quine criticized three post-Kantian definitions of analytic truth.[1] The first one he criticized was essentially Frege's, though he did not identify it as such. His criticism was focused on the class of supposed analytic truths that, like "No bachelor is married," are not logically true. According to Frege, statements of this kind are analytic just when they can be proved to be true by general logical laws and definitions. Quine described these statements a little differently, saying that they can be turned into logical truths by "putting synonyms for synonyms," the synonyms being expressions (words, phrases) appearing in the *definiens* and *definiendum* of the relevant definitions. If the definitions are acceptable, these expressions must be "cognitively synonymous": with the exception of poetic quality and psychological associations, their meaning must be the same. But how, Quine asked, can the synonymy of two words be known in a particular case? Can this be known if the word "analytic" is not understood already? He argues that the answer is no, and proceeds to look about for an alternative definition.

Why did Quine think that the notion of synonymy could be understood only if the word "analytic" is understood already? His reasoning was this. The definitions needed for the demonstrations Frege described served as principles of substitution. If the predicate "is a prince" is defined as "is a royal son," then we may substitute the latter for the former in the logical truth, "A prince is a prince," and obtain another truth, which can be considered analytic—namely, "A prince is a royal son." Since the words that good definitions allow us to substitute for one another must be cognitively synonymous, a promising way of defining cognitive synonymy is by means of substitutions that preserve truth: If substituting W1 for W2 in any true statement containing W1 always results in another true statement, the words W1 and W2 must be synonymous: they do what a good definition permits. This strategy seems promising until one realizes that the full range of statements containing a word W1 will include statements that also contain the word "analytic" (for instance, "It is analytic that princes are royal sons") or words that, if empiricists are right, can be understood only by means of "analytic"—for instance, "it is necessary that." If any of these statements were excluded from the substitution test, the test would not

[1] Quine (1953).

identify synonyms. If they are allowed, we can apply the test only if we already understand what we are trying to understand or make sense of.

Although the intimate connection between "being synonymous" and "being analytic" makes it inadvisable to try to define "analytic" by means of "synonymy," Quine's strategy in making sense of analyticity was nevertheless highly peculiar from the beginning. He initially noted that the statements held to be analytic "by general philosophical acclaim" fall into two classes, the first including logical truths such as "No unmarried man is married." He expressed no difficulty in understanding what a truth of this first kind is. "If," he said, "we suppose a prior inventory of *logical* particles, comprising 'no', 'un-', 'not,' 'if', 'then', 'and', etc., then in general a logical truth is a statement which is true and remains true under all reinterpretations of its components other than the logical particles."[2] But even this initial, limited clarification is peculiar in a discussion of what an analytic truth is. Kant's definition[3] was intended to show us why analytic judgments are true, but Quine's characterization of a logical truth assumes that we can recognize the truth and the resultant truth of statements that are true and remain true under all reinterpretation of their components other than the logical particles. This gives us no insight into how we know that the relevant statements are true.

The same holds for Quine's proffered account of the second kind of presumed analytic truths, the kind containing "All bachelors are unmarried," and his suggested strategy for defining synonymy. His suggestion was that analytic truths of the second kind are statements that can be turned into logical truths by putting synonyms for synonyms. This could work only if we had some independent means of recognizing logical truths. His strategy for identifying synonymous expressions had a similar limitation. We were supposed to consider whether the result of substituting one expression for the other in all true statements would be a true statement. But if we were wondering whether a candidate analytic statement "All princes are royal sons" is true, the question whether "prince" and "royal son" are synonymous would oblige us to consider whether the result of substituting "prince" for the first occurrence of "royal son" in "All royal sons are royal sons" is true—which is to say whether "All princes are royal sons" is true." The strategy would simply take us in a circle and get us nowhere.

A satisfactory definition of "analytic" should give us an understanding of why all analytic statements are true, the first kind as well as the second kind. Kant's definition did not apply to the class of logical truths, and it worked only for a small part of the other class.[4] The problem is to find a definition that works for the totality of both classes and also provides the understanding that an empiricist, an opponent of epistemological rationalism, desires. Quine considered two further definitions, or groups of them, but neither, as he understood them, appeared to work for all cases or provide the desired understanding. One definition (one member of the class he considered) was applicable primarily to artificial, formal languages, the idea being

---

[2] Ibid, p. 22.

[3]According to the definition Kant gave in his *Critique of Pure Reason* (A6, B10), an affirmative judgment is analytic just when its predicate is contained, perhaps only covertly, in its subject concept. To ascertain the truth of such a judgment, one has only to become conscious, he says, of what is contained in the subject concept. If the predicate concept is affirmatively contained in the subject concept, the judgment must be true, because anything to which the subject applies will satisfy or fall under the predicate: the predicate will apply to it, too.

[4]As early as 1884, Frege emphasized that Kant's definition does not include relational judgments such as "If the relation of every member of a series to its successor is one- or many-one, and if m and y follow in that series after x, then either y comes in that series before m, or it coincides with m, or it follows after m." See Frege (1950), p. 103.

that a statement of such a language is analytic if its truth is a consequence of the semantical rules laid down for that language. The other definition was based on the notion of empirical confirmation, although Quine relates it to the Verification Theory of Meaning: An analytic statement is one that is "confirmed no matter what."[5] This last definition is not credible in view of current conceptions of empirical confirmation,[6] but the "semantical rules" approach is far better than Quine supposed, and I will discuss it further in a later section of this chapter. Quine took a more moderate approach to analyticity in a later paper, and it will be instructive to consider his view in this paper next.

Quine's Later View of Analyticity

Forty years after he published "Two Dogmas..," Quine published **"**Two Dogmas in Retrospect."[7] In this later paper he summarized the more generous attitude toward analyticity that he had expressed in some of his later work. According to this more generous attitude, "analyticity undeniably has a place at a common-sense level… It is intelligible and often useful in discussions," he said, "to point out that some disagreement is purely a matter of words rather than of fact." A paraphrase that avoids a troublesome word can often resolve the disagreement. Also, in talking with a foreigner we can sometimes recognize "some *impasse* as due to his having mislearned an English word rather than to his having a bizarre view of the subject matter."[8] To deal with such cases, Quine offered what he called a "rough definition of analyticity." According to this rough definition, a sentence is analytic for a native speaker if he learned its truth by "learning the use of one or more of its words." He improved on this rough definition by "providing for deductive closure, so that truths deducible from analytic ones by analytic steps would count as analytic in turn."[9]

Quine claimed that the augmented definition accommodates such sentences as "No bachelor is married" and also the basic laws of logic. "Anyone who goes counter to modus ponens," he said, or anyone "who affirms a conjunction and denies one of its components, is simply flouting what he learned in learning to use 'if' and 'and.'" (He limits this to native speakers, he said, because a foreigner could have learned our words indirectly by translation.) Given the deductive closure qualification, he concluded that all logical truths in his sense—"that is, the logic of truth functions, quantification, and identity—would then perhaps qualify as analytic, in view of Gödel's completeness proof."[10]

In "Two Dogmas…" Quine had insisted that no statement is in principle immune to revision: revision even of the law of excluded middle had been proposed,

---

[5] Ibid, pp. 32-42.

[6] According to the conception I favor, E confirms H when E raises H's probability. Since an analytic truth has a maximal probability already, it could not be confirmed in the way Quine suggested. See chapter six, p.266. Devitt (2005), opposing the very idea of a priori knowledge on the "holist" ground that even purely logical statements must be confirmed together with other statements and "even whole theories" (p. 106) on the basis of experience, gives no hint of how the probability of "p ∨ ~p' might be raised by this process. Could it have a lower initial probability to begin with?

[7] Quine (1991).

[8] *Ibid*, p. 270. The other words quoted in this paragraph appear on the same page.

[9] The notion of closure is a mathematical one. As for analytic truth, saying that the set of analytic truths is closed under deduction is equivalent to saying that if **T** is deducible from members of this set, **T** belongs to the set as well.

[10] *Ibid.* By means of this proof Gödel showed that all the truths of first-order logic are derivable from a standard set of first-order axioms and rules.

he noted, as a means of simplifying quantum mechanics.[11] In the retrospective paper, he returns to this claim, asking "If the logical truths are analytic—hence true by meanings of words—then what are we to say of revisions, such as the imagined case of the law of excluded middle?" Echoing a question that his claim about the case often prompted in the past, he raises the additional question, "Do we thereby change our [logical] theory or just change the subject, change the meaning of our words?" He answers both questions by saying, "My answer is that in elementary logic a change of theory *is* a change of meaning. Repudiation of the law of excluded middle would be a change in meaning, and no less a change of theory for that."[12]

Although Quine proceeds to say that this "more generous" view of analyticity is not really as generous as it may appear, it is important not to move on too quickly, because his rough new definition is not easy to apply. According to the new definition, a sentence is analytic for a native speaker if he learned its truth by "learning the use of one or more of its words." Of course, by the word "sentence" here Quine obviously means "sentence with a fixed interpretation." But how could one possibly learn the truth of any sentence by learning the use of one or more of its words? Exactly how could this feat be accomplished? If we do not understand this, we will not really understand the import of his rough new definition.

Since Quine said the definition "obviously works" for "No bachelor is married," this example is a good one to start with. How could one learn the truth of this sentence by learning the use of some word in it? Here is one possibility. Suppose Tommy already understands the words "no," "is", and "married." And suppose he is familiar with the grammatical structure exemplified by the sentence in question. What he does not understand in the sentence is the word "bachelor." He therefore asks his mother, "What is a bachelor, Mom?" His mother answers, "A bachelor is a man who is unmarried." How can this answer teach him that "No bachelor is married" is true? This way, I should think. The mother's utterance tells him what the unknown word applies to: it applies to any man who is unmarried. Could a man who is unmarried be married? Obviously not: No man who is unmarried is married. Since "bachelor," according to his mother, applies to a man who is unmarried, Tommy knows that no bachelor is married. He puts two and two together.

Tommy learns the truth of "No bachelor is married" in a way that recalls Kant's definition of an analytic judgment. When Kant presented his definition, he observed in passing that it could easily be extended to negative judgments.[13] The idea would be that a universally negative judgment—one of the form "No S is P"—is analytic just when the predicate concept is excluded by what is contained in the subject concept. In what way excluded? The answer is "logically excluded":[14] the ideas involved in the subject concept are logically incompatible with the predicate concept just as the ideas included in the concept of a bachelor—the ideas of being a man and being unmarried—are logically incompatible with the idea of being married. One can know that a universally negative analytic judgment is true because, on ascertaining what is contained in the concept of the subject, one will be logically assured that nothing falling under the subject concept could possibly fall under the predicate concept: the application conditions of the two concepts are logically incompatible.

---

[11] Quine (1953), p. 43.

[12] Quine (1991), p. 270.

[13] Kant (1997), A7, B11.

[14] The idea of a logical relation is also implicit in Kant's original definition, for he said that in affirmative analytical judgments the connection of the predicate [to the subject] is thought through [the relation of] identity. *Ibid.*

I am not certain that my description of the way Tommy learns the truth of "No bachelor is married" conforms to what Quine had in mind when he spoke of learning a sentence by learning the use of some word in it. But I cannot think of another way that such a thing could plausibly be learned. Still, the pattern of this description does not apply to the way one might learn the truth of a basic law of logic. To learn the truth of "No bachelor is married" Tommy applies logic to what his mother tells him about the meaning of a word in a sentence he otherwise understands; he concludes that "No bachelor is married" is true because it is equivalent to "No man who is unmarried is married," and he knows that the latter is true. Evidently we do not conclude that a basic law of logic is true because something else is logically true. We do not reason in this way. How, then, are we to understand the kind of learning Quine has in mind when he speaks of learning the truth of a basic law of logic?

I really do not know the answer to this question, but a plausible candidate quickly comes to mind. When philosophers think of logic, they think of formal logic; they do so because *logical truth* is a formal notion, as is *the validity of an inference*. Today, formal logic is expounded by means of various symbols, some representing logical operations such as negation, conjunction, or universal quantification, and others representing statements and their parts—for instance, individual variables, individual constants, and relation symbols. When we learn a truth of formal logic, we learn the truth of a symbolic formula, and when we learn the validity of an argument form, we learn the validity of a symbolic pattern or sequence. Quine may suppose that we can learn the truth of certain formulas and the validity of certain symbolic patterns by learning the use of symbols contained in them.

It is convenient to begin with a valid form of inference. I have described such forms of inference as symbolic patterns or sequences; these patterns consist of statements, or premises, and a conclusion that is validly inferred from them. One of the simplest of logically valid argument forms involves conjunctions: all arguments conforming to this pattern are logically valid:

$$(p \wedge q) \;/\therefore\; p$$

To learn that this argument form is valid, we must first learn that a valid argument form is one whose proper instances have true conclusions whenever they have true premises: a valid argument form is truth-preserving. When this information is in hand, we then learn that the symbol "$\wedge$" is used to assert the truth of two statements, the two it conjoins. In learning this we learn that if a premise having the form of "$p \wedge q$" is true, both of its conjuncts are true, its first conjunct as well as its second. To learn this is to know that the form represented above is valid.

The other valid argument form that Quine mentioned is a form of modus ponens. This argument form is usually represented by a pattern containing two premises, one containing the symbol " $\supset$ " or an equivalent such as "$\rightarrow$":

$$(p \supset q),\ p \;/\; \therefore\; q.$$

To learn the validity of this form of inference we need to learn the meaning of the horseshoe symbol, " $\supset$ ". This symbol corresponds to the English "if…, then…," but its meaning is special. Its peculiarity is that it forms a conditional statement that is true whenever its antecedent is false or its consequent is true. If both premises in an argument having the form of modus ponens are true, the antecedent of the conditional premise must be true, because it is the same as the second premise. Since a horseshoe conditional is true whenever its antecedent is false or its consequent is true, the consequent of the second premise must then be true, because its anecedent is not false. But the conclusion of the argument is the same as the consequent of the second premise. Since this consequent is true, the conclusion is true. The argument form is therefore valid: when the premises are true, the conclusion is true as well. This is guaranteed by the meaning of the horseshoe symbol and the concept of a valid argument form.

A little later in "Two Dogmas in Retrospect," after expressing his generous attitude toward analyticity, Quine becomes more negative, saying "In fact my reservations over analyticity are the same as ever, and they concern the tracing of any demarcation, even a vague and approximate one, across the domain of sentences in general." By "sentences in general" he means all sentences, not just the ones expressing logical laws and truths such as "No bachelors are married." He supports this generally negative attitude with two reasons. The first is that "we don't in *general* know how we learned a word, nor what truths were learned in the process." The second is that we have no reason to expect uniformity in this matter of learning from speaker to speaker" (p. 271). Although Quine does not take these two reasons as undermining the analyticity of logical laws and examples such as the one about bachelors, we might ask why he does not. If we do not in general know how we learned a word, do we know how we learned logical words and words such as "bachelor"? And do we all learn these words in basically the same way?

The answers to these questions bring out something special about logical words (or logical symbols) and words such as "bachelor." They have, at least on particular readings, precise meanings, and they are learned in the same basic ways. Words like "bachelor" (on certain readings) are short for longer clusters of words, and when we learn their meaning—whether we are given their meaning by a teacher or parent or whether we look them up in a dictionary—we learn what groups of words they abbreviate. Like little Tommy, we learn to substitute them for their equivalents in statements that are logically true, and we thereby come to know truths that are analytic in Quine's sense. The precision of logical words has a similar result. When we learn the meaning of a logical symbol such as the horseshoe, we learn to compute the value of conditionals containing it by means of the values of the statements it connects. There is just one truth-function associated with this symbol, and when we learn what this is, we understand that symbol; we do so whether we initially encounter it in a definition relating it to negation and disjunction or in an equivalent definition that relates it to negation and conjunction. The same is true of other logical symbols. When we know what they mean, we can "by analysis" compute the truth-value of many statements in which they occur.

Analyticity, Logic, and Everyday Language

If the only truths we can reasonably claim to be analytic are those of elementary logic and trivialities such as "Bachelors are unmarried males," then the concept of analytic truth does not have the importance that empiricists take it to have. This is Quine's position, and I think he is right in holding it. I intend to provide a more satisfactory account of analytic truth in what follows, but before attempting to do so, I must first resolve some issues left over from the last chapter. Resolving these matters will bring me closer to the analysis I want to defend.

When I criticized the rationalist claim that basic logical truths can be seen to be true by a kind of direct intuition, I emphasized the extreme generality of these truths and went so far as to find instances that appeared to falsify them. I cited examples of statements that, asserting other statements to have a certain truth-value, could apparently be proved to be both true and false themselves, and I offered other examples that, owing to vague expressions contained in them, could reasonably be said to be neither true nor false and that, together with statements like them, provided apparent counter-instances to basic logical laws such as the principle of excluded middle. I even cited examples of arguments, formulated in everyday English, that some philosophers have taken to be counter-examples to modus ponens. Since these examples could not possibly be surveyed by the direct intuitions focused on general or schematic formulas that rationalists appealed to as sources for their a priori knowledge, I concluded that the rationalist's belief in the epistemic efficacy and authority of these alleged intuitions was simply and clearly unfounded.

However successful my examples may have been in refuting the basic rationalist claim about intuitive certainty, they also raise a problem for the empiricist alternative, for they raise (or should raise) serious doubts about the certain truth of the supposed logical laws that even Quine eventually described as analytic. How could we possibly know that the schematic formulas that are supposed to hold true for *all* statements corresponding to them do not, in fact, have a single falsifying instance? Do empiricists have an infallible means of surveying all instances that is not available to the rationalist? If so, what is it?

Not all empiricists would answer these questions in the same way, but one answer is this:[15] The instances to which a schematic formula is intended to apply are prescribed rather than simply surveyed. A system of logic is commonly introduced in connection with an artificial language, a system of formulas that are constructed and interpreted in specific ways. The statements of such a language system are "well-formed formulas," and rules are introduced that describe how they are properly constructed. Such formulas are interpreted by means of semantical rules, which assign semantic values to the formulas and their functionally significant parts. Possible values for the "closed" formulas of classic systems[16] are restricted to truth and falsity, and no formula can possess both these values. The kind of semantic vagueness that make it appropriate to assign an indeterminate value to particular formulas is therefore not allowed in a classical system, and one can know in advance that any legitimate formula of the system will satisfy the schemas expressing the laws of excluded middle and non-contradiction. Similarly, by placing restrictions on the kinds of predicate that can be acceptably attached to statements of certain classes, one can disallow statements such as "The sentence in the triangle is false" and make it impossible to derive in the system the sort of contradiction that I discussed in the last chapter. Thus, by playing it safe—by excluding from a logical language the sort of statement that can cause logical trouble—we can insure that classical laws are preserved there. To make this assurance maximal, to banish any possible doubt from the simplest and most trouble-free vocabulary, we can go so far as to declare that any formula leading to trouble will count as deviant all along. The system never involved an error, we may say; it was simply set up or described incorrectly.

The arguments and assertions that we evaluate in everyday life do not, of course, belong to artificial languages and they do not consist of technical symbols that need to be assigned semantical values by technical rules. How can we use logic to evaluate them? One strategy is to adopt translations for them in a symbolic language. Thus, we might translate "If the ladder slips, the man will fall" into "L ⊃ F," taking "L", " ⊃ ", and "F" as translations, respectively, of "The ladder slips," "if," and "The man will fall." Since the formula "L ⊃ F" is easily evaluated by means of

---

[15] I am following Carnap (1958) here; see his chapter B.

[16] A formula is *closed* when any variable it contains is bound by a quantifier; "∃x(x is a prime number)" is a such a formula. Formulas with free variables may be "satisfied" by an object but they are not true or false.

our rule for formulas containing the symbol "⊃", our evaluation for "L ⊃ F" will apply to its translation in the vernacular, "If the ladder slips, the man will fall." The acceptability of this evaluation will obviously depend on the acceptability of translating "if" in the vernacular sentence by the symbol " ⊃."[17] If the meaning of "if" in this sentence is considered acceptably close to that of "⊃," the translation will be acceptable; if not, it will not be.

Another strategy for evaluating everyday arguments and assertions is to select a part of everyday language, possibly regiment it in ways that eliminate ambiguity and vagueness, and then create a logical language that is a hybrid of vernacular forms and technical symbols. A sentence of this sort of language might be "The ladder slips ⊃ the man will fall." We might even use everyday words in place of logical symbols, using "and" with the meaning of "∧" and "if" with the meaning of " ⊃ ". In this last case it will appear that we are using the language of everyday life, but we will be using just a selected part of it (not every grammatical sentence of English will count as a proper formula) and some words will not have their usual senses. To avoid paradoxes and violations of standard logical laws, we must impose restrictions on our total logical vocabulary.[18]

When Quine, "In Two Dogmas in Retrospect," agreed that the laws of classical logic and statements like the one about bachelors can be considered analytic in the rough sense he described, he left no doubt that he was thinking of logical truths as expressed in everyday language, for that is the language in which people learn the word "bachelor" as well as "if" and "and," which are the logical words he mentioned.[19] It is also clear that Quine was not thinking of the restrictions on everyday language that must be accepted if the formulas for basic logical laws are not to be falsified. (Thus, he had nothing to say about vagueness and the so-called semantic paradoxes exemplified by the statement about the false sentence in the triangle). Also, he ignored the fact the vernacular "if" is not always used in such a way that those who have mastered its use invariably recognize the validity of modus ponens. I noted in the last chapter that the validity of modus ponens and modus tollens are, in fact, sometimes challenged by philosophers who support their case by presenting examples formulated in the language of everyday life. Now is a good time to return to the examples I presented, for they underline the importance of tying logical problems to logical systems.

The first example was this:

> If it rained yesterday, it did not rain hard (yesterday).
> It did rain hard (yesterday).
> Therefore, it did not rain yesterday.

This argument seems to have the form of modus tollens; yet the conclusion must be false if the second premise is true. It would appear that the first premise could be true. The second premise could also be true. Yet if the truth of the second premise guarantees the falsity of the conclusion, it would appear that the argument cannot be valid.

Do we have a genuine counter-instance to modus tollens? The answer is "No, particularly not if the first premise is understood as a material conditional, one that

---

[17] It will also depend on the acceptability of taking a certain vernacular sentence as the translation of a statement constant that must be either true or false. I comment on this below.
[18] I emphasize the importance of this claim for current arguments about the justification of basic logical principles in Appendix 2, where I criticize some recent contentions by Paul Boghossian and Hartry Field.
[19] See Quine (1991), p. 270. Quine's view of formal logic and its relation to vernacular discourse is expounded most fully in Quine (1981).

can be represented by 'It rained yesterday ⊃ it did not rain hard yesterday'." If it could be represented this way, the falsity of the conclusion would guarantee the falsity of one of the premises. It is true, as I mentioned, that both premises could be true, but reflection shows that on this interpretation they could not be true *at the same time*: they are inconsistent. This can be seen as follows. If the second premise is true at some time, the consequent of the first premise must then be false. But if this consequent is false at that time, the antecedent of the first premise must equally be false if that premise is true. The falsity of this antecedent is therefore inconsistent with the truth of the second premise.

The second example concerned the participants in the 1980 U. S. presidential election, which was eventually won by the Republican, Ronald Reagan. Jimmy Carter, a Democrat, was second and Anderson, a Republican running as an Independent, was third. The example was as follows:

> If a Republican wins, then if Reagan does not win, Anderson will win.
> A Republican wins (=does win).
> Therefore, if Reagan does not win, Anderson will win.

Vann McGee, who discovered the example, thought it is a counter-example to modus ponens because the first and second premises seem obviously true while the conclusion seems false. Reagan won, and since he and Anderson were the only Republicans running, if he did not win, Anderson would. The conclusion seems false because the real race was between Reagan and Carter; Anderson was far behind. At the time of the election it would therefore be false to say, "If Reagan does not win, Anderson will win."

When I originally presented the example, I expressed the opinion that it is impossible to say decisively whether it is or is not an acceptable counterexample without some clarification of the English in which it is expressed. The logical word "if" featured in it is clear in some respects, but it is not clear in others, for arguments containing it can be expressed in nonequivalent symbols. Suppose we read the argument as having the following logical form:

> A Republican wins ⊃ [~(Reagan wins) ⊃ Anderson wins].
> A Republican wins.
> Therefore, ~(Reagan wins) ⊃ Anderson wins.

Read this way, the argument is clearly not a counter-instance, for the conclusion is plainly true: it is logically equivalent to "Reagan wins ∨ Anderson wins," which is guaranteed to be true if it has a true disjunct--and it does so in this case.

There are, of course, other ways of construing the argument. When I presented it as an ostensible counterexample, I suggested that the conclusion is false because the real race was between Reagan and Carter, Anderson being so far behind as to be effectively out of it. If the conclusion is read with this firmly in mind, it will appear to have a subjunctive force not captured by the horseshoe symbol. Suppose, therefore, that we interpret the "if"s in the argument as representing the counterfactual conditionality expressed by David Lewis's symbol " □→".[20] Conditionals of this kind are evaluated by reference to possible worlds or "ways the world might be." A conditional of the form "P □→ Q" is considered true just when, of all possible worlds in which P is true, Q holds in the one or the ones most similar to the actual world.

---

[20]See Lewis (1973), passim.

On this interpretation the argument takes the following form:

1*. A Republican wins ❑→ [~(Reagan wins) ❑→ Anderson wins]
2*. A Republican wins.
3*. Therefore, ~(Reagan wins) ❑→ Anderson wins.

Understood this way, the conclusion is no doubt false, for in a world in which Reagan does not win but that is otherwise minimally different from the actual world (the "closest world" in which Reagan does not win), Carter presumably wins instead of Anderson. Yet the first premise is now false, and it must be true if the argument is to provide a counterexample. The closest world in which a Republican wins in 1980 is the actual world, and in this world it is not true that if Reagan were not to win that election, Anderson would. Thus, when the vernacular "if" is replaced by the technical symbol "❑→", the resulting argument also fails to provide an acceptable counter-instance to modus ponens.

Not all occurrences of "if" need be replaced by the same technical symbol, of course. Two further arguments could be obtained if one of the following formulas were put in place of 1:

4*. A Republican wins ❑→ [~(Reagan wins) ⊃ Anderson wins]
5*. A Republican wins ⊃ [~(Reagan wins) ❑→ Anderson wins]

If 1* were replaced by 4*, the result would not be an instance of modus ponens, however; for the consequent of 4* differs from 3*. If 1* were replaced by 5*, we would have an instance of modus ponens, but the first premise would not then be true. 5* is logically equivalent to the disjunction of "~(A republican wins)" and 3*, both of which are false. Thus, on these further readings we still do not have an acceptable counterexample.

Other, nonstandard readings of the vernacular "if" are possible, and it is on one such reading that Christopher Gauker defends a counterexample to modus tollens.[21] The multiplicity of possible readings of the vernacular argument raises an important question: "Just what is modus ponens?" If we do not have a particular system of logic in mind, we cannot answer this precisely. We can say that modus ponens is an argument form in which a conclusion q is inferred from a premise p and a conditional premise having p as antecedent and q as consequent; but because formulas of significantly different logical powers can be described as conditionals, argument forms of significantly different kinds can count as instances of modus ponens, some lacking counter-instances and some, for all I know, having them. The vernacular "if" is not so precise in meaning that only a single interpretation is possible for it even in a given context. If we want to single out a definite class of argument forms in speaking about modus ponens, we shall have to restrict our reference to the argument forms that can be constructed from the vocabulary of some formal system or group of systems. As I noted earlier, a "regimented" part of English may count as such a system, the precision (or logical determinacy) of its formulas depending on the way it is regimented.

It should be clear to the reader that the arguments I could confidently declare to be, or not be, counterexamples to modus ponens contain logical symbols with precise interpretations. The horseshoe symbol is not a common term whose meaning is determined by the linguistic behavior of ordinary speakers; it is a technical symbol whose logical properties are fixed by logical convention. This and

[21] See above, p. 36.

other conventions permit an exact assessment of formulas whose implications are sufficiently parallel to those of certain vernacular statements to be considered the latter's symbolic transcriptions, but the vernacular statements are far less determinate in what they assert.[22] For an additional example, consider "Either something is red or everything red is green." A natural assessment of this statement is that it is a contingent truth, supported by the fact that red things obviously exist. But if it is interpreted as adequately symbolized by the formula "$\exists xRx \vee \forall x(Rx \supset Gx)$," it is easily seen to be a tautology, because "$\forall x(Rx \supset Gx)$" is true if no x is R.

Analyticity Extended

If we return to Quine's rough definition of analyticity, we see that it is acceptable only on certain idealizing assumptions—that the language is appropriately "regimented," as Quine put it in *Word and Object*, that certain sentences containing "is true" and "is false" are ignored, that vagueness is disregarded or evaluated by special conventions,[23] and that logical words have the sense of certain technical counterparts. Even allowing these assumptions as trouble-free, Quine's rough definition is, as he emphasized, significantly limited and ostensibly not sufficient to accommodate the problem statements that rationalists regard as expressing synthetic a priori truths, the statements I claimed to be analytic in the last chapter.

To obtain a more encompassing definition of analyticity, it will be instructive to consider another of the definitions of analyticity that Quine criticized in "Two Dogmas…," the one focused on semantical rules. Quine actually criticized several definitions of this kind, claiming that the fundamental defect common to them all is the appeal to semantical rules: the idea of such rules is as much in need of clarification as analyticity itself.[24] Rudolf Carnap, Quine's close friend but his opponent regarding analyticity, had claimed that "the concept of analyticity has an exact definition only in the case of a language system, namely a system of semantical rules, not in the case of ordinary language…."[25] In "Two Dogmas…" Quine denied this, saying in effect that this claim puts the cart before the horse: "Semantical rules determining the analytic statements of an artificial language are of interest only in so far as we already understand the notion of analyticity; they are of no help in gaining this understanding" (p. 36).

This last remark by Quine is seriously exaggerated. As Carnap said in his *Introduction to Symbolic Logic and Its Applications*, semantical rules are rules of interpretation for what would otherwise be an uninterpreted language or formal calculus (p. 80). There is nothing obscure about the purpose of some of these rules. As I noted earlier, the horseshoe, one of the basic symbols of elementary logic, has a technical meaning that cannot be adequately explained simply by relating it to the vernacular "if" (or some counterpart in another language). To explain it adequately for the purpose of a logical system, one must specify rules of interpretation that allow us to calculate the truth-value of compound formulas containing it and other formulas. In this case the rules can be reduced to this one: A formula of the form "$p \supset q$" is true just when the formula corresponding to "p" is false or the formula

---

[22] Frege emphasized the difference between the material conditional, "$P \supset Q$," and the "if…then" of everyday language in Frege (1962); see pp. 550ff of the reprint in Klemke (1968).

[23] When vernacular discourse is regimented for logical purposes, the vagueness of everyday assertions is commonly ignored. When this sort of vagueness is explicitly recognized, a number of different logical strategies are available. One possibility is to assume a qualification that makes a vague statement sufficiently determinate to deserve a value of T or F; for example, "Tom is thin" may be read as meaning "Tom is on the thin side." For other strategies, see van Fraassen (1966), Lewis (1983), pp. 244-46, and Williamson (1994).

[24] Quine (1953), p. 36.

[25] See Carnap (1990), a short paper written in 1952 and never published by Carnap himself.

corresponding to "q" is true. This is a simple, well-known rule, and to the extent that one understands it and the point of having it, one understands something about the meaning of the words "semantical rules" and the purpose of the rules they denote.

In criticizing the "semantical rule" definition of analyticity, Quine compared the notion of a semantical rule with that of a postulate. Just as no true statement is inherently a postulate, so no string of words is inherently a rule, semantical or otherwise. But Carnap agreed with this. In his view a particular semantical rule represents an interpretive decision, a decision about how some symbol or aggregate of symbols is to be understood in relation to the intended domain of discourse. Different decisions are always possible, but if particular decisions are made in a given case, words of a familiar kind can be used to express those decisions in that case. The same words could be used to express different decisions in a different case. As Quine said, no sentence is inherently a postulate.

The semantical value of a statement in relation to a domain of discourse is usually truth or falsity; the value of a proper name is usually a particular member of that domain; the value of a two-place predicate is a set of ordered couples in that domain; and so on. But we can also interpret some symbols by relating them to others whose interpretation is already known. Some definitions have this purpose. If "adult male" and "unmarried" are understood as belonging to the vocabulary of a regimented language-system, the word "bachelor" can be given a precise interpretation in relation to this vocabulary by the formula, "∀x(x is a bachelor ≡ (x is an adult, male human being ∧ x is unmarried))."[26] The idea would be that regardless of the meaning that the word "bachelor" might have in everyday language, in the context of the regimented system it is to be understood as an abbreviation of the words appearing in the right-hand side of the defining formula. One may wish to introduce a strict sense of "bachelor" if a strict sense is needed for special purposes.

Carnap was convinced that a precisely specified language system is needed for the concept of analyticity because he thought words have no "clearly defined meaning" in ordinary language. It is easy to miss the reasonableness of his view here. Consider "bachelor," a word for which a strict sense might conceivably be needed. In everyday life the word is not only ambiguous, but it is often used quite loosely. As for ambiguity, the word is now occasionally applied to young women living alone or to people possessing a B.A. or B.S. degree (see the OED); as for looseness, people are actually apt to disagree (as Gilbert Harman observed) about whether the word is applicable to the pope, who is not married in any ordinary sense, or whether it should be applied to a man who has lived with a woman for several years without getting married.[27]

In view of the controversy about truths that epistemological rationalists claim to be synthetic a priori, it is worth considering an example that arose in a dispute between Carnap and Quine on analyticity. The example was "Everything green is extended," which Quine said he hesitated to classify as analytic because of an incomplete understanding not of "green" or "extended" but of "analyticity." Carnap said it seemed "completely clear" to him that the difficulty lies in the unclarity of "green," which betrays an indecision whether to apply the word to a single space-time point.[28] "Since one scarcely ever speaks of space-time points in everyday life," he said, "this unclarity about the meaning (or intended application) of 'green' plays

[26] Technical definitions have the form of a biconditional or an identity statement. See Suppes (1957), Ch. 8, "Theory of Definition."

[27] Harman (1996), p. 399.

[28] In his language form IIB described in Carnap (1958), Carnap defined space-time points as "the smallest non-empty spatial regions; see p. 160.

as small a role [in everyday life] as the unclarity about whether the term 'mouse' should also be applied to animals which, apart from their greenness, are completely similar to the mice we know, but are as large as cats."[29] This lack of clarity is unimportant for the practical purposes of everyday life, but it is vitally important for the philosophical question about the analyticity of "Everything green is extended." To settle the latter, Carnap thought, we must make our meaning of "green" or color words generally more precise in relation to our thought about points.

The idea of making one's meaning more precise in certain respects, or in some respects rather than others, was very important for Carnap and is, I believe, very important for the subject of analyticity. Carnap first called attention to the importance of a partial analysis in 1936, when he wished to introduce predicates for dispositions into the context of a technical language having the horseshoe as its sole symbol for conditionality. He could not define "x is water-soluble" by the conditional "x is immersed in water ⊃ x dissolves," because, owing to the truth of material conditionals with false antecedents, anything never immersed in water would then count as water-soluble. To avoid this difficulty, he introduced the idea of a "bilateral reduction sentence," a formula by which the meaning of a disposition predicate is specified incompletely, only for instances in which the relevant test condition is satisfied. The general form of such a reduction sentence is "$Q_1 \supset (Q_3 \equiv Q_2)$," where "$Q_1$" and "$Q_2$" represent preexisting predicates of the scientific language and "Q_3" represents the predicate whose meaning is being specified for cases in which the test condition "Q_1" is satisfied.[30] Applied to the predicate "water soluble," the reduction sentence lays down a necessary and sufficient condition for the application of this predicate to objects immersed in water. The predicate's application to objects not so immersed would remain undetermined in basically the way that the application of "is bald" is undetermined for cases in which a person showing a lot of scalp still has a significant amount of hair.

The practice of reconstructing the meaning of vernacular words, which I discussed in chapter one in connection with David Lewis's treatment of "S knows that P," Carnap called "explication." When the meaning of a word or formula is fully explicated, or completely reconstructed, it is introduced into technical language by explicit definitions whose *definiens* consist of words or symbols whose meaning is antecedently clear and unproblematic. For cases in which the meaning is explicated only incompletely, Carnap first used the label "meaning postulate" and later changed it to "A-postulate":[31] for him, A-postulates are the formulas providing the partial explications. These explications are not generally intended to specify some part or aspect of the meaning that a word or group of words already possesses; they are used to stipulate the meaning they have in a specified (or tacitly understood) context: either the context of a technical language or discourse, or that of some discussion.

Carnap illustrated the point of an incomplete stipulation in a paper called "Meaning Postulates."[32] Suppose a person constructing a certain system wishes to use the symbolic predicates "Bl" and "R" in a way corresponding to (but not necessarily the same as) the way "black" and "raven" are used in everyday life. Speaking of such a person, Carnap says:

> While the meaning of 'black' is fairly clear, that of 'raven' is rather vague in the everyday language. There is no point for him to make an elaborate

---

[29] Carnap, "Quine on Analyticity," p. 427.
[30] See Carnap (1936 and 1937).
[31] See Carnap (1966), p. 261.
[32] Carnap (1956) pp. 222-229.

> study, based either on introspection or on statistical investigation of common usage, in order to find out whether 'raven' always or mostly entails 'black.' It is rather his task to make up his mind whether he wishes the predicates 'R' and "Bl" of his system to be used in such a way that the first logically entails the second. If so, he has to add the postulate (P_2) '(x)(Rx ⊃ Bl x)' to the system, otherwise not" (p. 225).

If the postulate P2 is added to the system, the person constructing it has thereby stipulated how, in the context of the system, the predicate "R" is to be understood in relation to a symbolic predicate corresponding to "black." If "R" is applicable to a thing x, "Bl" must be applicable to it as well.

Gilbert Harman once said, "…stipulative definitions are assumptions. To give a definition is to say 'Let's assume for the time being that the following equivalence holds'."[33] This is wrong. Assumptions can be false; stipulative definitions cannot.[34] If I decide to use "raven" in accordance with the stipulation (holding for a certain context) that nothing non-black will count as a raven, I will not be proved wrong if something that might be called a raven in the ordinary sense--a bird indiscernible from a raven except for being white--should be observed. It would simply not be a raven in my stipulated sense. Using my special terminology, I might call it a "waven" and say that ravens and wavens in my sense of the words are pretty clearly subspecies of a distinct kind that might be called "dravens." Seeing such a bird might move me to bring my special terminology more into line with common usage and to use "raven" as people ordinarily do. But I would not have made an error in using "raven" as I formerly did.

A meaning postulate, as Carnap understood it, is very close to the sentences featured in the "modest" sort of analytical account that Williamson offered for the concept of knowing.[35] This kind of account discloses the conceptual connections between a target concept and certain others, and in doing so it provides a kind of non-reductive analysis of the target concept. In explaining how knowing can be understood as being the most general "factive, stative attitude," Williamson identified a number of analytic implications in which "knows" participates. Three obvious examples are the following:

If S knows that P, it is true that P.
If S remembers that P, S knows that P.
If S sees that P, S knows that P.

Carnap differs from Williamson in having serious reservations about the precision and determinacy of everyday language. As I have explained, his postulates are to be understood as stipulations rather than complete or partial analyses of existing usage. He generally expected them to reflect existing usage if there is no need, scientifically or philosophically, to diverge sharply from it; but he thought that we are bound to diverge in some degree if we wish to be clear and precise.

Although I am somewhere between Carnap and Williamson in my attitude toward everyday language, I have no doubt that Carnap's strategy of providing stipulative explications allows us to introduce a broader sense of analytic truth than the one given by Quine's "rough definition." Statements so explicated are analytic *for us* (not analytic generally) because they represent part or (conceivably) all of

[33] Harman (1996), p. 399.
[34] They can, of course, be revised, abandoned, and the like. But revision and so forth is not the same as falsification. In Appendix 3 I discuss some conditions that an acceptable stipulation must satisfy.
[35] See chapter one, p. 00.

what we mean in using the words they contain. People who speak "our language"—people who speak English, for example—need not mean what we mean by every word, and our explications need not be valid for what they say. This broader sense of analyticity does not therefore identify the analytic sentences of a whole natural language or dialect, though we may wish and even recommend that others adopt our usage in preference to theirs. Nevertheless, this limited and local conception of analyticity is sufficient for epistemology. It allows us to dispose of the issues rationalists raise by means of the problem examples I discussed in chapter two.

Consider again the statement, "Nothing can be both yellow and green all over at the same time." As my discussion in chapter two made clear, this statement need not even be true. "Yellow" and "green" are highly generic predicates that are not used in exactly the same way by all speakers of English. Although they are perhaps normally regarded as incompatible, they can be used, as Harry of my story did, in a way that makes them jointly applicable to the same part of a leaf or shrub. Something with the determinate shade Harry called "green-yellow" may be described as both green and yellow all over, for both colors are there, all over. What are clearly incompatible are determinate color shades: If something is green-yellow in Harry's sense or yellowish-green in Mary's sense, it cannot at that time also have any other determinate shade of color. This incompatibility is not a matter of ontological fact that is independent of classificatory conventions; it is a consequence of how we individuate a thing's specific color at a time. We *could* restrict ourselves to a purely generic means of attributing colors, calling things either yellow, green, red, or blue, and so on; and if we did so, there would be no definite error in our describing something with Harry's green-yellow shade (which we would not then distinguish as such) as both green and yellow at the same time.

In discussing color incompatibility in the last chapter, I said that we do in fact identify specific colors in a way that assumes indiscernibility as an identity condition for them. We consider a determinate color A to be the same as a determinate color B just when A and B are indistinguishable.[36] When we conceive of specific colors this way, we are tacitly accepting a convention that renders it analytic for us that nothing can have two different determinate colors at the same time.[37] The analyticity here is not peculiar to just a few of us; it holds for all who accept the convention—all who identify specific colors this way. Many of the tacit conventions that render statements analytic for members of a group govern aspects of the use of words or sentences that are as wholes vague or hard to define. It is not easy to say exactly what a fake object is, but there is no doubt that if something is a fake duck, it is not a real one, and there is no doubt that that if Nero fiddled while Rome burned, Rome was burning while Nero fiddled. Grammatical structures that do not appear in formal languages also warrant inferences that are valid for those who use them. If someone says of a friend, "Lacking an umbrella, she hit him with a shoe," we are normally entitled to infer that, if the speaker is right, the hitter lacked an umbrella, hit a person or animal with a shoe, and did the latter because of the former. The truths of these conditionals and the acceptability of this last inference are not ascertained in Quine or Frege's way, by making deductions from logical truths and accepted definitions; they immediately come to mind as the consequence of tacit conventions accepted by all who use the relevant language in a normal way and can think abstractly about truth and validity.

---

[36] A more satisfactory of expressing this is to say that x and y (or regions on their surfaces) have the same determinate color just when *they* are indistinguishable in color. The point of this observation will become evident in chapter 4.

[37] See the proof given in Appendix 1.

At the present time 350 million people speak English as their first language and around 450 million speak it as a second language.[38] These people live in different parts of the globe, have conflicting interests and customs, and vary greatly in education and general knowledge. Generalizing about the structure of English or the meaning of this or that English word is therefore inherently risky. The same is true, of course, for any other widely used language. Realizing this, I am tempted say that we can justifiably speak of analytic truths only when we can relate them to logical systems and explicit stipulations, the latter being either complete or partial. But this attitude is really too cautious. The examples I gave in the last paragraph make it obvious that words, phrases, clauses and constructions in existing dialects of natural languages have implications so vital to the meaning of what they are used to say that any alert and attentive speakers of a relevant dialect would find it odd, puzzling, or paradoxical to question them. When this condition is satisfied by a word or symbol, it seems to me that a sentence of the dialect clearly and unambiguously expressing an appropriate implication can reasonably be regarded as analytically true for those alert and attentive speakers.[39]

In making this last claim I am obviously adding to the conception of analyticity that Carnap offered. I am not limiting analytic truths to statements that are true for certain speakers by virtue of explicitly identified semantical rules and complete or partial stipulative explications; I am also including statements whose truth is ensured by the conventions that those speakers tacitly apply in making them—conventions whose implications are so vital to the meaning of the words and structures being used that the speakers would find it odd, puzzling or paradoxical to question them. These latter statements can, of course, be related to the sort of semantical rules and complete or partial explications that Carnap described. The procedure is this: If explicit semantical rules and complete or partial explications sufficient to demonstrate the truth of those statements *were* formulated, brought to the attention of the relevant speakers, and satisfactory explained to them, the speakers *would* then accept them as making explicit the meaning they attach, wholly or partly, to the words, phrases, and constructions involved in those statements. If the speakers would not do this, and if no alternative explanation of their negative attitude were available, the statements in question could not reasonably be regarded as analytic for them there and then.

The meaning speakers attach to the words they use in saying this or that need not be associated with a dialect in a narrow sense of the word. This is an important matter, because the speakers might comprise a very small group, even a singleton, adopting special conventions for a particular publication or a serious conversation. Just the other day, in a discussion with another philosopher, I temporarily adopted a special convention for the word "variable." Because adjustments and qualifications pertinent to a person's usage are often partial, temporary, and relevant to just this or that audience, a satisfactory account of analyticity should always be related to some reasonably determinate context. The explicatum should be "Φ is analytic for Σ in context X," where Σ is a class that includes the relevant persons (the speakers and hearers, or just the speaker or speakers) and X includes the parameters identifying the context. S may be analytic for Tom and Sally in the context of a particular discussion; S′ may be analytic for me

[38] Ferguson (2002), p. 304.

[39] The analytic character of the informal inferences normally involved in evaluating formulas and argument forms by reference to semantical rules is to be understood along these lines. The inferences could, of course, be formalized, in which case their validity could be assessed by higher-order rules. But the assessment would not *make* a formula tautologous or an inference valid. "P ∨ ~P" is a tautology if it is an instance of a schematic formula all of whose proper instances are true. There are different ways of discovering whether "P ∨ ~P" has this property; one is by using a truth table.

in the context of a book or chapter I have written; and so on. The relevant explicans (or analytical account) that provides the explication should *ideally* list the relevant semantical rules and the full or partial explications that characterize the conceptually determinate aspects of the language used by the persons Σ in the context X. In practice this is an excessively demanding requirement for speakers of natural dialects, because they are normally accustomed to relying on tacit conventions that only experts can be expected to identify and describe.[40] But special meanings should nevertheless be clarified in this way. If the meanings are special, they are usually not associated with tacit conventions.

Some Examples and Arguments by Kripke

Having explained Carnap's approach to analyticity and my extension of it, I can now attempt to come to terms with some important unfinished business-- specifically, the examples illustrating the alleged necessity of a thing's origins that Kripke mentioned in two footnotes of his *Naming and Necessity.* I discussed these examples in the last chapter. I noted that Kripke said one of the examples is "susceptible of something like a proof," and reflection convinces me that the argument he seemed to have in mind for this example can be converted into arguments that apply to the others. These arguments depend on an axiom of modal logic that, like any logical axiom, is arguably analytic in Carnap's sense of being true by virtue of semantical rules. If the arguments succeed, the examples can then be considered analytic in the sense I have explained; they will provide no support for epistemological rationalism.

The arguments I shall consider make use of a strategy Kripke included in a footnote to the second edition of *Naming and Necessity* to support the principle that if an object has its origin in a certain hunk of matter, it could not have had its origin in any other matter. After formulating this argument in way that makes its logic easy to follow, I will show how it can be revised to support the other examples.

The principle to be proved by the first argument can be stated as follows. If M1 had its origin in a hunk of matter H1, then M1 could not have originated from any hunk H2, where H1≠H2. This principle is intended to hold for all M1, H1, and H2; the argument, in showing that it holds for any arbitrarily chosen values of these variables, shows that it holds for them all. The argument proceeds by conditional proof. Assume that there is a possible world in which M1 had its origin in H1 (as in the actual world) and that an object very like M1 was made from a different hunk of matter H2. Since H1 and H2 are distinct hunks of matter, M1 is distinct from M2 in this world. But if two objects are distinct in any possible world, they are distinct in every possible world. This is a theorem of Kripke's modal system. Yet if M1 could not be identical with M2, which represents *any* relevantly similar object made from a different hunk of matter H2, M1 could not have originated from any such hunk. Since an origin would be impossible.

I said above that the argument just given could be adapted to provide arguments supporting the other principles about the necessities of a thing's origins that Kripke discussed. Take the principle about parents: If C's biological parents are P1 and P2, C could not have been born to anyone other than P1 and P2. To prove this, assume that there is a possible world in which C's biological parents are P1 and P2, as in the actual world, but that a person D, just like oneself otherwise, was born from other parents, P3 and P4, at the very same time. In this world, clearly, D ≠ C, since they have different biological parents. By the kind of modal reasoning given in the last paragraph, it follows that there is no possible world in which D = C. Since D

[40] Consider the conventions for irregular verbs described in Pinker (1999).

is representative of *any* possible person born of different parents when C was born from P1 and P2, it follows that C could not have been born from different parents.

Kripke says (p. 114n) that the two arguments I have just given ultimately rest on a modal principle that he calls "the necessity of distinctness."[41] The principle described by these words is usually expressed by saying that if **a** and **b** are distinct things—that is, if **a** ≠ **b**—then it is necessary that **a** ≠ **b,** but the argument I have given above requires a stronger principle—namely, that if it is possible that **a** and **b** are distinct, then it is necessary that they are distinct. Expressed differently, the principle is that if there is a possible situation in which **a** and **b** are distinct things**,** then they are not be the same thing in any possible situation. It seems to me that this principle accords with what we mean in speaking of the same and different things. If there is a possible situation in which I am distinct from some other person, how could I possibly be myself and also be that person in some other situation? The meaning of vernacular words is not decisive for logic, of course, but the operators of Kripke's modal system are sufficiently parallel in meaning to their vernacular counterparts to be used in their place in a complicated argument. Since the semantical rules of his system allow us to prove that the strong necessity-of-distinctness principle is logically true, we can justifiably regard it as analytic in Carnap's sense. And since the arguments I have reconstructed depend on that and other logical principles, the claims Kripke supported by means of those arguments—if the arguments are in fact satisfactory—deserve to be regarded as analytic as well. No patently synthetic a priori premises are needed in their defense.

I added the qualification, "if the arguments I have given are in fact satisfactory," because I do not believe that they actually prove what they are intended to prove. Take the second argument, which is intended to prove that if the biological parents of a person C are P1 and P2, C could not have been born to anyone other than P1 and P2. To prove this, the argument supports the principle that if it is possible for a person actually born of parents P1 and P2 to have those parents when a very similar person has other parents, then the first person could not be identical to (or one and the same as) the second person. But this last principle is not equivalent to the principle the argument purported to prove, nor does it entail that principle. This is evident from the fact that the argument relies on the possibility of two very similar persons, C and D, with different parents coexisting in a possible situation. The fact that C is not identical with D in this situation does not show that there is no other situation in which C has the parents D has in this situation. If C had those parents in some other situation, C would not have his (or her) actual parents there, but C would still be himself (or herself), not some other person.

Although I am not convinced by the arguments I have considered, I would not insist that the principles about the necessity of origins that Kripke discusses are in fact false. I offer no opinion on that subject. I will say, though, that if those principles can be proved by some argument,[42] the argument will be analytical and the principles will be shown to be analytic. There is no plausibility in the idea that they are intuitively obvious or deducible from premises that are not analytic in the extended sense I have introduced in this section.

Beliefs, Propositions, and Analyticity

What I have been saying about analyticity in the last two sections supports a language-centered account of the subject. Can it be extended to accommodate the

[41]Kripke (1980), p. 114.
[42] Nathan Salmon discusses other arguments for Kripke's conclusion in Salmon (2005), ch. 7.

apparent fact that judgments and beliefs may also be analytically true? These psychological states may be expressed linguistically, or put into words, but they are evidently not themselves linguistic entities and it would appear that they are as susceptible of a priori truth or falsity as any statement. Stipulations about the meaning of words can hardly be pertinent to their falsity or truth. Or so it would seem.

To evaluate this important objection, it is vital to have a defensible conception of a judgment or belief, to know just what they are and how they are put together. Someone new to epistemology might think that the nature of these states is obvious to any thinking person, but the reality is quite otherwise, at least if we go by what philosophers say about them. A very common claim is that judgments, beliefs, doubts, suppositions and a host of other propositional attitudes consist in some relation to a "proposition."[43] A judgment is always a judgment that P (for some P); a belief is always a belief that P; and analogous claims hold true for the other attitudes. What is common to them is some proposition or other; they differ in the way they are related to a proposition. Believing and doubting involve relations that are virtual opposites; believing and suspecting are similar in some respects but different in others; believing and opining are substantially the same.

If we are to take this view of so-called propositional attitudes seriously, we have to know what a proposition is. The classic view of such a thing, the one worked out in what David Kaplan called "the Golden Age of Pure Semantics," was introduced by Gottlob Frege and refined by Rudolf Carnap.[44] Frege viewed a proposition as the "sense" or meaning of a sentence. Since the words of a meaningful sentence are themselves meaningful units that contribute to the meaning of the whole, the sense of a sentence is a function of the meanings (or senses) of its words. According to Frege, the names and predicates of a sentence have "concepts" as their senses, [45] and these concepts may be singular as well as general. Consider the sentence "Socrates is wise." Corresponding to the descriptive words in this sentence are two concepts, the individual concept corresponding to "Socrates" and the general concept corresponding to "wise." Frege's account of the relation between these concepts is somewhat confusing; in one place he appeared to describe it as a relation of subordination (of the individual concept to the general one).[46] Carnap described it as attribution or predication: presumably the general concept is predicated of the individual to which the individual concept applies.[47]

How are propositions so understood related to believing? Frege and Carnap appear to differ on this matter. Judged by his essay, "The Thought: A Logical Inquiry," Frege seemed to believe that propositions can be directly apprehended and so accepted by the believer independently of any sentence. For him, the basic relation between person and proposition was one of "apprehension."[48] Carnap, by contrast, held that our access to propositions involves the use of sentences. In his view, the statement "John believes that P" has the sense, approximately, of "John is disposed to an affirmative response to some sentence that expresses the proposition

---

[43] I comment on Scott Soames's version of this view in 9.
[44] Kaplan (1991a), p. 214.
[45] This common interpretation does not accord with some of Frege's explicit claims. In "On Concept and Object" (Frege [1892]) he described a concept as the "nominatum" of a predicate and perhaps considered the corresponding sense as the "mode of presentation" of this concept. Carnap (1956) says that the interpretation I assume here, which he and Alonzo Church accept, "is in accordance with Frege's intentions when [as he occasionally does] he regards a class as the (ordinary) nominatum of a…common noun and a property as its (ordinary) sense" (p. 125).
[46] See Frege (1892), p. 48.
[47] Carnap (1956), p.
[48] Frege (1965), p.307.

that P."[49] Since Carnap himself applied the predicate "is L-true" (his equivalent for "is analytic") only to sentences for which semantical rules have been given, anyone who accepts his analysis can apply "is analytic" to propositions only in some extended sense. The only plausible way of doing this is to say that a proposition is analytic if it is expressed by (or is the intension of) a sentence whose truth, in a system S, can be calculated on the basis of the semantical rules of S alone. The treatment of analyticity I outlined in the last sections obviously accommodates this strategy very well. One's belief is analytic just when the sentence to which one is belief-related is analytic in the sense I specified.

Frege's view of our access to a proposition is obviously far more attractive to a rationalist than Carnap's, but the relative merits of these views are no longer very significant since the classic view of propositions involved in them has been seriously undermined by recent work in semantics. The fundamental defects of the classical view can be traced to proper names whose supposed correlates in a proposition were taken to be individual concepts. Frege and Carnap thought that these concepts were needed to connect names to objects in the world, but the required individual concepts do not, in general, exist: there is no generally shared conceptions that single out the referents of commonly used names, and historical individuals such as Socrates and Aristotle may fail to satisfy the descriptions that people commonly associate with them. The connection between proper names and their referents is now generally thought to be "direct" rather than mediated by some associated concept. A connection is set up in a community by various talk and behavior, sometimes by acts of naming or dubbing, and the name is then spread through the community of language-users by talk and actions, moving from "link to link as if by a chain."[50] No individual concept, no uniquely identifying description, is needed in this process.

Demonstrative expressions such as "I," "here," "now," "he" and "she" are also not connected to their referent by some individual concept; they too directly refer to their referent. They have, it is true, as David Kaplan has emphasized, a distinctive character by means of which speakers and hearers can identify their referent in this or that context of utterance, but there are no propositional components, no concepts, that single out those referents. As a matter of fact, auditors will commonly interpret an utterance containing demonstratives by different words, even when speaker and auditor share the same language. I say "The book is here on this desk," and my hearer interprets me as saying that the book is there on that desk. Mary tells me "I will meet you on that corner tomorrow," and the next day I, waiting on the right corner, think, "She said she would meet me on this corner today." Her assertion and my thought of what she said have no common, classically conceived propositional object.[51]

What conception has emerged from the breakdown of the classical conception of propositions? No single conception appears to be dominant.[52] Some philosophers who once accepted the classic conception have simply given up on propositions altogether.[53] Others have retained classical propositions for fully general sentences but developed new conceptions for sentences containing proper names and demonstratives. One thing common to leading conceptions of singular propositions—the propositions expressed by atomic sentences containing proper names—is that the

---

[49] Ibid., pp. 54-62.
[50] Kripke (1980), p. 91; see also pp. 92-164.
[51] See Perry (1979). Another serious problem with classically conceived propositions is presented in Kripke (1979).
[52] See King (2001) and Fitch (2002).
[53] This is Chisholm's response; see Chisholm (1997),

referents of the names are said to exist *within* those propositions. David Kaplan, who was the first to develop a view of this kind in the late twentieth century, cited the early Bertrand Russell as his precedent. Kaplan himself described the proposition expressed by "Socrates is wise" as the ordered pair consisting of the man Socrates and the property *Wise*—that is, as <Socrates, *Wise*>.[54] Other philosophers have described singular propositions in other ways, but they have retained the Russell-Kaplan strategy of "loading" referents into these propositions.[55]

The truly revolutionary features of propositions so understood is that they are not themselves objects that represent the world, as classical propositions were, but helpers or interpreters (it is hard to say which) of other objects—namely, sentences—that do represent it. Such propositions are often informally referred to as "what is said"[56] by utterances of sentences in various contexts, but this way of speaking is not really appropriate. If I say that Tom Smith has a silly smile, I say something *about* Tom Smith; the man himself is not happily described as part of what I say. In fact, if the "property" that Kaplan takes to be the second component of the singular proposition <Socrates, *Wise*> is the sort of thing that can exist in the world (as many philosophers suppose) neither component of this proposition is reasonably considered a part of what someone might say. Both parts are rather things one may refer to or talk about in saying this or that.

Another fashionable conception of propositions is commonly advocated by philosophers concerned with the semantics of counterfactual conditionals and statements of necessity and possibility.[57] According to this conception, propositions are either sets of possible worlds or functions from possible worlds to truth-values. But sets of possible worlds can hardly be grasped by the mind in the way Frege and others thought propositions could be grasped, and the same is true of functions from worlds to truth-values, which are commonly viewed as sets of ordered couples, each couple consisting of a possible world and an associated truth-value, specifically truth. This conception is obviously quite technical, but it is not really hard to understand, and it has the merit, from my point of view, of being entirely compatible with the view of analytic truth that I developed in the last section. I want therefore to say some more about it here.

Consider the sentence "Bachelors are unmarried." Understood in the usual strict or idealized way, this sentence is true in a wide range of possible worlds or "ways the world might be."[58] Conceived of as a function (or many-one relation)[59] from possible worlds to truth-values, the proposition expressed by the sentence "Bachelors are unmarried" is the function that assigns the value T (truth) to a world just in case the bachelors in that world are unmarried. Conceived of more simply as a set of possible worlds, the proposition is the set of worlds in which all bachelors are unmarried. But which worlds are in this set? Or, equivalently, which worlds are assigned the value T by the relevant function? The answer is "All possible worlds whatever." How do I know that this answer is true? Because the sentence "Bachelors are unmarried," understood in the usual strict way, is analytically true. Any person in any possible world that counts as a bachelor is guaranteed to be unmarried. This is owing to the meaning of the predicate "bachelor" or to any

---

[54] Kaplan (1991a), p. 221.

[55] For a helpful discussion of specimen examples of these alternatives see the article "Propositions" in the *Stanford Encyclopedia of Philosophy.*

[56] See Appendix 4.

[57] Se Lewis (1973), p. 46f.

[58] David Lewis describes possible worlds as "ways the world might be." See Lewis (1986), p. 2.

[59] A relation R is said to be many-one just when for every object x in its domain (the entities it relates to something) there is just one object y in its range (the entities it relates something to). *The biological father of* is thus a many-one relation, since everyone has just one such father.

predicate that properly translates it. We do not have to examine the contents of a possible world to know that it is assigned the value T by the bachelors-are-unmarried function. We know this by knowing what a bachelor (in the sense in question) is supposed to be.

It is worth observing here that the possible-worlds conception of a proposition moves the notion of a proposition away from the classical conception for a reason I have not yet mentioned. According to the classical conception, a proposition is the fundamental bearer of truth: it is what is true in a fundamental sense. As Frege put it in a famous passage: "What does one call a sentence? A series of sounds; but only when it has a sense…. And when we really call a sentence true, we really mean its sense is."[60] (Recall that a proposition, for Frege, is the sense [*Sinn*] of a sentence.) But a function from worlds to truth-values or a set of possible worlds is not really a bearer or possessor of truth; it is not itself true at all. This point, oddly enough, seems to be overlooked even by philosophers who actually make it. In his excellent encyclopedia article on propositions, Jeffrey King says this:

> Intuitively, it [the intension of a sentence, a proposition] maps a world to the value true if the sentence is true at that world. Thus the intension of a sentence can be seen as the primary bearer of truth and falsity at a world: the sentence has the truth value it has at the world in virtue of its intension mapping that world to that truth value.[61]

What King actually says to be true here is a sentence, or possibly a world; the proposition is a "bearer" of truth only in the metaphorical sense that it "carries" (maps) the world to a truth-value. Thus, propositions on this conception not only fail to be "what is said"; they are no longer even true or false.[62]

In view of the general failure of the classical conception of propositions, it is important to consider an alternative to the attitude-object view of propositional attitudes, the one that describes them as relations to a propositional object. The standard alternative, historically speaking, is known as conceptualism, the view held by such philosophers as Kant. According to this view, propositional attitudes—believing, judging, supposing, and so forth—have "contents" rather than "objects." The content of a thought that Socrates is wise has two principal constituents. The first constituent is a singular idea, one that represents Socrates in the way that the name "Socrates" represents him. Following Kaplan, we can say that the idea represents him *directly*. The other constituent is a general idea, a concept in Kant's sense, one by means of which the referent of the subject idea is characterized as wise. This is substantially Kant's account of the matter, though his logical apparatus is simpler than what we would use today. Since the predicate concept in this last case is not contained in the subject (it could not be, since the subject has the character of a name) Kant would declare it to be synthetic. If the matter were otherwise—if the predicate were so contained—it would be analytic.

How would this conceptualist account of thought relate to my extended account of analyticity, the one involving semantical rules and complete or partial explications? This way: Just as such rules and explications tell us what reality (or an item of reality) must be like if a certain word or formula is applicable to it, so analogous rules and explications tell us what reality must be like if the idea or thought expressed by a given formula is applicable to it. The process by which we

---

[60] Frege (1965), p. 290.
[61] King (2001).
[62] It is interesting to note that Carnap, in a book where he espoused a basically Fregean view of language, insisted that "truth in the semantical sense is a property of sentences." See Carnap (1956), p. 93.

fix the content of an idea or thought is fundamentally the same as the process by which we fix the meaning of a word or formula: we adopt appropriate principles of reference, equivalence, and inference. To make my idea of a minimal if-then relation clear, for instance, I can adopt the principle that a compound thought involving this relation will be true just when either its antecedent is false or its consequent is true. In adopting this principle I am clarifying my idea. I can also form ideas in a deliberate way. The fundamental fact is that thoughts, as conceptualists describe them, have the semantic properties of the words that "express" them, and their content is determined by corresponding principles.

Philosophers who believe in classically conceived propositions say that they can be expressed in different languages by sentences that are good translations of one another. Since a claim of this kind is not intended to be a tautology, it should be possible to say, at least in a general way, what counts for mutual translatability without introducing the idea of a proposition. If I were asked how we could identify two sentences that have this status on a standard reading, I would say that each sentence should be built up in equivalent ways by words that apply to the same objects and have corresponding implications. To have corresponding implications the words would have to be such that, if they are applicable to certain objects, additional words that are good translations of one another are also applicable to those objects. The syntactical and semantical similarities that must exist here are very complicated, but if we can ascertain that they do exist, we will know everything we have to know to decide if they are good translations or not. We will not have to appeal to anything abstract that they both "express."

If propositions, understood as classically conceived abstract objects, are not needed to account for the translatability of one sentence by another, and if thoughts and statements can be semantic counterparts without being related to a common object of this kind, then such propositions are not really needed for an acceptable semantic account of either thought or speech. Since I accept both antecedents of this last conditional, I accept its consequent: classically conceived propositions are not in fact needed for a semantical analysis of either thought or speech. This conclusion is supported, as Kaplan and others have emphasized, by the directly referential character of names and demonstratives, but it is also supported, as I shall argue in the following chapter, by the directly attributive character of predicates. As they are used in properly formed sentences, predicates can be used to characterize or describe objects without relating them to a further object, a property in one sense of the word, that somehow does the job for them.

Near the end of the "Afterthoughts" that he appended to his paper "Demonstratives," David Kaplan expressed the view that "our connection with a community in which names and other meaning-bearing elements are passed down to us enables us to entertain thoughts *through the language* that would not otherwise be accessible to us." We become capable of thinking about things in the world as the result of having experienced various things ourselves, but we also gain the capacity "vicariously," he said, "through the symbolic resources that come to us through language. It is the latter—vocabulary power—that gives us our apprehensive advantage over the non-linguistic animals."[63] I agree with these sentiments completely. We do gain the capacity to think about occurrences in ancient history, exotic forces and fields in subatomic physics, and even certain traits and quirks of everyday acquaintances only by means of words we learn from parents and teachers, textbooks and dictionaries, newspapers and television. We do not master perfectly the words we accumulate, and our sources are also imperfect transmitters of collective verbal wisdom, so there is usually some lack of fit between our speech and

[63] See Kaplan (1992b), pp. 603f.

thought and the speech and thought of others. For this reason, there are not many words in common use with the precise univocal meanings that could justify the definitions and analyses that many philosophers construct—if those definitions and analyses are not partial or explicative in the sense I have explained. An analytic/synthetic distinction is not really possible for the whole of our language as it actually is. But clarifications and reconstructions are always possible. These provide the basis for an acceptable, philosophically useful account of analytic truth.