# COUNTERFACTUALS AND THE ANALYSIS OF NECESSITY*

Boris Kment
University of Michigan, Ann Arbor

This paper is, in part, a straightforward exercise in philosophical analysis: I will try to define metaphysical necessity. But I will combine this aim with another: I want to know which cognitive practices of ordinary life gave rise to the concept of necessity, and what role the notion plays in these practices. Let me describe this goal in more detail, before giving an overview of my strategy.

## 1. Goals and overview

### 1.1. Internal and External Perspectives

Although the expression 'metaphysical necessity' is a technical term of analytic philosophy, I think that non-philosophers usually either have an inchoate, implicit grasp of the concept expressed by it, or can at least very easily be gotten to cotton on when the concept is explained to them, even if they are not given an explicit definition. When presenting a highly intelligent person without philosophical background with a carefully designed series of thought experiments, we might get her to say, "I could not have been born of different parents. A person with different parents would not have been me." (This is essentially what happened to me a couple of years ago as I conversed with a very intelligent non-philosopher about some moral issues concerning human cloning.) I also remember that Kripke's examples in *Naming and Necessity* and his conclusions about the necessity of origin immediately struck a chord in me when I first read the book as an undergraduate, long before I had a developed philosophical notion of metaphysical necessity or had learned to carefully distinguish epistemic from metaphysical necessity. I think that this suggests that the concept of metaphysical necessity in some sense arises out of certain ordinary-life practices of thinking and talking.

If this is correct, then a comprehensive philosophical account of modality should not only tell us what metaphysical necessity is, but should also tell us

which ordinary-life practices give rise to modal notions, and what role modal concepts play in them. It should thereby elucidate what the purpose of these notions is, why creatures with our interests and concerns have developed them. In this way, a comprehensive theory of modality should combine a theory of necessity with a theory of the practice of modalizing. The point can be explained by the metaphor of the difference between an internal and an external viewpoint. Before we start to philosophize about necessity, we have an implicit theory about it. The philosopher provides this pre-philosophical system of beliefs about modality with a foundation, and refines, extends and corrects it from within. He acts as a participant in our practice of modalizing; his standpoint is internal to this practice. But the philosopher should also make the practice of modalizing itself an object of study. He should, as it were, take a standpoint external to the practice, in order to describe the practice, and explain what its function is, why it exists.

It is a familiar fact that these two kinds of interest can pull in opposite directions. If we concentrate exclusively on the task of giving a metaphysical account of what necessity is, we might end up with a theory that makes it hard to explain why we are interested in modality. It is a common charge—whether it is justified or not I shall not endeavor to decide—against Lewis's account of necessity that it fails in just this way.[1] Philosophers have objected that, even if there were other worlds in Lewis's sense, we would have no apparent reason to be interested in what goes on in them. Hence, Lewis's account, so the objection continues, makes it a mystery why we should bother to think about modal facts. At the other end of the spectrum there are theories that do well at explaining the purpose of our modal notions, but do so at the expense of giving implausible accounts of what necessity is. Consider the conventionalist theory that for a proposition to be necessary is for it to owe its truth to a convention, perhaps the convention that we ought to regard the proposition as true come what may. If some propositions are indeed conventionally exempted from empirical testing, then it benefits our epistemic practices if we possess a concept that singles them out. The conventionalist therefore has no great difficulty with explaining the point of our modal notions. But this advantage is purchased at the price of several well-known drawbacks in her theory of necessity.

The task is thus to develop an approach to modality that permits us to achieve both of our goals, a credible metaphysical theory of necessity and a plausible account of the practice of modalizing. I think that the best way of doing this is not to neglect either objective while pursuing the other, but to integrate the two goals in a single enterprise: an account of the nature of metaphysical necessity can be guided by a hypothesis about the ordinary-life practice in which the concept of necessity originated, while assumptions about what necessity is can in turn suggest ways of developing one's ideas about the practice of modalizing. This is the methodology I will employ.

## 1.2. Overview of the Project

In trying to find a plausible metaphysical account of necessity, we may naturally start from one or the other of two intuitions about what necessity is. On the one hand, there is the intuition that modal discourse, talk about which things could or could not have been different, is concerned with situations other than the one that actually obtains, with alternatives to the way things are. Talk about whether something could have been different essentially concerns the question whether there is an alternative to the way things are in which the thing *is* different. Call this the *otherworldliness intuition*. On the other hand, it seems appealing to say that the core difference between a truth that could have been false and one that could not have been false is that the truth of the latter is somehow more *metaphysically secure* than that of the former, more *unshakable*, or *inexorable* (to borrow a term that Edward Craig used to articulate the intuition). The literature on modality abounds with passages that give expression to this intuition.[2] According to this idea, necessity is at bottom a special *mode of truth*, a special way of being true, namely that of being true in an especially secure and inexorable way. I will call this the *modalist intuition*.

As Peter Railton very helpfully suggested to me, the images underlying the otherworldliness and modalist intuitions are, respectively, that of *freedom* and that of *force*. According to the *otherworldliness* intuition, the range of possibilities circumscribes what we may call the *degrees of freedom* of the world (in the same sense in which we speak of the degrees of freedom of a physical system). The necessary truths impose restrictions on the degrees of freedom; they restrict the range of alternatives to the way the world actually is. According to the *modalist* intuition, there is a certain special *force* that attaches to those propositions that are necessary, and which secures their truth.

Different philosophers take different modal notions as basic. One group defines the modal operators—'necessarily,' 'possibly,' and so forth—by quantifying over possible situations. To say that it is possible that Fred has black hair is to say that there exists a possible situation in which Fred has black hair. Others take the modal operators to be basic and prefer to define the notion of a possible situation by using these operators, e.g. by identifying possible situations with sets of propositions that could have been jointly true. I suspect that each of these approaches is guided by one of the intuitions described in the previous paragraphs. If one assumes that modal discourse is ultimately about alternatives to the way things are, then it becomes natural to analyze the modal operators by quantifying over these alternatives, i.e. over possible situations. On the other hand, if one starts from the modalist intuition, then it is tempting to regard the concept of the special mode of truth characterized above as the most basic notion of modal discourse. Now, if one regards the necessity operator as expressing the property of having this special mode of truth (so that the sentence 'necessarily, *P*' is an ascription of this special mode of truth to the claim *P*), then it appears

natural to regard the necessity operator as expressing the most basic concept of modal discourse, and to define other modal notions (like that of a possible world) in terms of modal operators.[3,4]

I think that both intuitions, the otherworldliness intuition and the modalist intuition, are very forceful, and I believe that a plausible metaphysical account of modality needs to be capable of accommodating both intuitions. It is important to be clear about what is required in order to accommodate them. I take both the otherworldliness intuition and the modalist intuition to be intuitions about what it *is* for a proposition to be necessary. According to the first intuition, to be necessary just is to be true in all scenarios that are alternatives to the way things actually are; according to the modalist intuition, it is to be true in some particularly secure and inexorable way. Now suppose that we propose, as our account of necessity, that to be necessary is to be $F$ (for some predicate '$F$'). In order for this view to accommodate the two intuitions it is *not* sufficient that it can be shown that the propositions that have the property of $F$-ness are just those that are true in all alternatives to the way things are, or that the propositions that are $F$ are just those whose truth is particularly secure in the sense underlying the modalist intuition. Rather, the property of $F$-ness must *be* (identical with) the property of being true in all alternatives to the way things are, and it must *be* (identical with) the property of being true in a particularly secure way. In order to accommodate both intuitions, we need to find some way of interpreting the phrases "true in all alternatives to the way things are" and "true in a particularly inexorable way" on which the two phrases single out the same property, and then identify necessity with that property.

My attempt to do this starts from the otherworldliness intuition. I argue for a certain way of interpreting that intuition and of transforming it into an approach to the question what necessity is. After that, I begin the quest for an account of necessity anew, this time starting from the modalist intuition. I argue that this intuition, too, can be transformed by natural steps into an account of necessity. As it turns out, both maneuvers lead us to the same theory of necessity. I take this to support my contention that this view can be regarded as capturing both the otherworldliness and the modalist intuition.

Consider first the otherworldliness intuition, i.e. the intuition that necessary propositions are those that are true in all scenarios that are alternatives to the way things actually are. This intuition suggests that the ordinary-life practice that gives rise to the concept of necessity is one in which we consider situations that we do not believe to obtain, and ask ourselves what is true in them. There is more than one common cognitive procedure that fits this description. In section 2, I argue that the one in which modal concepts originate is our routine of representing to ourselves certain scenarios and considering what *would have been true if they had obtained*. We commonly express the outcome of such a thought experiment by a counterfactual conditional. The core idea of my theory is that, roughly speaking, for a proposition to be necessary is for it to play a certain distinctive role in the truth-conditions of counterfactuals.

This view reverses the customary order of explanation. For the standard view of counterfactuals, as propounded by Stalnaker and Lewis,[5] analyzes them in terms of the modal notion of a (metaphysically) possible world: 'if it had been the case that *p*, then it would have been the case that *q*' is (to simplify a bit) true just in case *q* is true in all those possible worlds in which *p* is true and which are otherwise as similar ('close') to the actual world as is compatible with the truth of *p*. I argue against this view in section 2.2. My discussion centers on a well-known problem for the standard view: if the antecedent of a counterfactual is metaphysically impossible, then there are no possible antecedent-worlds, so that the standard account entails that the counterfactual is vacuously true. But it seems very implausible to me that all counterfactuals with metaphysically impossible antecedents are true. (It is metaphysically impossible for water to be an element, but it does not seem true to say that if water were an element, everything would be the case.) I adopt one of the obvious candidate solutions to this problem (which has been developed in more detail by Daniel Nolan[6]): I reformulate the standard account by simply replacing the concept of a possible world with the wider, non-modal notion of a world. Worlds, which I think of as abstract entities of some sort (possibly sets of propositions), comprise both possible and impossible worlds. Impossible worlds are ordered by their closeness to the actual world, just as possible worlds are. A counterfactual is true just in case the consequent is true in the closest (possible or impossible) antecedent-worlds. The resulting theory of counterfactuals does not use the concept of a possible world, and therefore leaves us free to use counterfactuals to analyze necessity.

Such an analysis gains support from the observation that it seems natural to paraphrase

*P* could not have failed to be true,

as

*P* would have been true no matter what (else had been the case),

or, equivalently, as: *P* is true and, for any situation whatsoever, if that situation had obtained, *P* would still have been true. In section 2, I argue that, by qualifying and refining this idea, we can develop an account of necessity. I also show that the central idea of this account can be reformulated in terms of the familiar closeness (similarity) account of counterfactuals: for a proposition to be necessary is for it to be true in all worlds that have at least a certain degree of closeness to the actual world.

Consider next the modalist intuition that the necessary truths are those propositions whose truth is particularly *secure* or *inexorable*. This intuition appears to rest on the idea that there is a dimension of inexorability on which we can locate different truths. The necessary truths are those truths whose value on that dimension is above a certain point. In section 3, I will argue that it is this dimension of security or inexorability that we are talking about when we

ask ourselves, concerning some fact about our world, how easily it could have failed to obtain. To say that a certain proposition has a high degree of security or inexorability is simply to say that it could not easily have failed to be true. On the account I will propound, how easily a state of affairs could have obtained depends on the range of counterfactual situations in which it does obtain. If it obtains in situations that depart only minimally from the actual world, then we are inclined to say that it could easily have obtained. (Suppose that your favorite soccer team would have won their last game if the goalkeeper had stood just an inch further to the right in the fiftieth minute of the game. This gives you reasons for saying that your team could easily have won.) If a state of affairs obtains only in situations that depart very much from actuality, then we will instead say that it could *not* easily have obtained. How easily a state of affairs could have obtained is thus determined by the degree of closeness between the actual world and the closest world in which the state of affairs does obtain. The degree of inexorability of a true proposition is accordingly measured by the distance from the actual world to the closest worlds in which the proposition is false. If we combine this with the thought that metaphysical necessity is simply a high degree of inexorability, we arrive, once again, at the conclusion that for a proposition to be metaphysically necessary is for it to be true in every world that has at least a certain degree of closeness to the actual world.

In section 3.3 I argue that this approach to metaphysical necessity can be generalized to other kinds of necessity, such as nomic and conceptual necessity. I think that these other kinds of necessity mark off degrees on the same dimension of inexorability as metaphysical necessity, although they mark off different degrees. The degree of inexorability that a proposition needs to have in order to be conceptually necessary is higher, that which it needs to have in order to be nomically necessary is lower, than that which is required for metaphysical necessity. In section 4, I use these ideas to formulate formal definitions of conceptual, metaphysical and nomic necessity.

My discussion of the modalist intuition yields the result that modal properties—necessity and possibility—come in degrees. I argue that, when we talk about how easily a certain state of affairs could have obtained, we are talking about its degree of possibility. The relation of comparative closeness of worlds, too, can be interpreted as a relation of comparative possibility: to say that a certain world is close to the actual world is to say that it could easily have been actualized, i.e. that it has a high degree of possibility.[7] The closeness relation is therefore itself a modal relation. The definition of necessity in terms of closeness that I give in section 4 reduces one modal concept to another. In order to reduce the modal to the non-modal, it is still necessary to give a non-modal account of the closeness relation. I tackle this task in section 6.

I find it plausible that modal facts must be grounded in non-modal facts. If a proposition is necessary, then this modal status must ultimately be grounded in its non-modal properties. We can sharpen this idea by drawing on the modalist intuition that to be necessary is to be true in an especially secure

way. This intuition is naturally accompanied by another: if a certain proposition is impossible, then this is so because that there is some kind of particularly formidable obstacle that prevents it from being true. What secures the truth of a necessary proposition is the fact that its negation runs up against especially inexorable obstacles, against particularly deep features of the world order. This makes it natural to say that the non-modal feature of a necessary truth that grounds its special modal status is that it is underwritten by certain very deep features of the world order (e.g. by the metaphysical or mathematical order of the world). In sections 4 and 5, I try to explain these ideas in less metaphorical terms and, drawing on the account of counterfactuals that I expounded in my (2006), I argue that my theory can capture them.

In section 7, I attempt to show that, by directing our attention to the cognitive practice in which the concept of metaphysical necessity originates, the account of this paper permits us to see why this concept is useful to us: it facilitates counterfactual reasoning, by allowing us to single out those propositions that play a certain special role in that practice.

A deeper account of the utility of modal concepts would also require us to say what the use of counterfactual reasoning itself is. The answer is, I think, that counterfactual reasoning is a component of a reasoning strategy that is useful for a variety of purposes, such as predicting the likely outcomes of possible future actions, and evaluating claims about the causal and explanatory interrelations between different facts. Unfortunately, I cannot expound that part of my view in this paper. That is a task for another occasion.

The core idea of the account I will propound has been foreshadowed (though not developed) by Davis Lewis in the 1970's in his book on counterfactuals, and in a posthumously published paper by Ian McFetridge.[8] In the last couple of years, Marc Lange, Timothy Williamson, Christopher Hill, and I have worked out similar ideas independently of each other.[9] In this paper, I intend to make further progress on motivating and developing the approach.

## 1.3. Resources

As a preliminary to carrying out the project I have outlined, I will give a brief overview of the concepts and presuppositions on which my account rests.

*Firstly*, I will use the concepts of a thing's essence or nature, and of its essential properties. The essence or nature of an object is what it is to be that object. The essence of propane, e.g., is to be $C_3H_8$, since to be propane just *is* to be $C_3H_8$. The essential properties of an entity are those that are part of what it is to be that entity, i.e. those that make it the entity it is. It is, e.g., an essential feature of propane that it is a compound of hydrogen and carbon; being composed of these elements is part of *what it is* to be propane. By contrast, it is merely accidental to propane that it is used for cooking.

It has long been common to define the notions of essence and of an essential property in modal terms.[10] (On this account, a property of a thing is essential to

it just in case the thing cannot exist and fail to have the property. The essence of an object is a property that it cannot fail to have (if it exists) and that no other thing could have had.) But recently Kit Fine[11] argued that this characterization of essence cannot adequately capture the underlying intuitive idea. As Fine points out, it is a necessary feature of the number 2 to be a member of the set {2}, and a necessary property of {2} to have 2 as a member. But while having 2 as an element is part of what it is to be {2}, being a member of {2} is not part of what it is to be 2. If Fine is right, as I think that he is, then there is no apparent reason for thinking that the concepts of essence and essentiality need to be explained in modal terms. The two notions can therefore be used in an account of necessity without obvious threat of circularity. That is what I will do.

The way in which I conceive of my project is intimately bound up with the concept of essence. I take the otherworldliness and modalist intuitions to be intuitions about the *essence* of necessity. They amount, respectively, to the intuition that truth in all alternatives to the way things are is the essence of necessity, and that inexorability is the essence of necessity. And I understand my present task as that of specifying the essence of metaphysical necessity, i.e. of giving what is sometimes called a *real definition* of the property of metaphysical necessity.

*Secondly*, I will avail myself of a conception of propositions as entities with syntactic (sentence-like) structure. This conception allows us to understand the notion of (narrowly) logical truth for propositions in a very intuitive way: a logical truth is a proposition that is true in virtue of its logical structure alone. This concept of logical truth seems to me to be non-modal, and the standard (model-theoretic) way of spelling out its details makes no use of any modal locutions.[12]

*Thirdly*, I will use the notion of a conceptual or analytic truth. I think that this concept, too, can be understood non-modally: a proposition is an analytic or conceptual truth just in case it is true in virtue of the fact that it is built up from certain concepts in a certain way, and in virtue of the natures of these concepts.

Given that the notions of conceptual and narrowly logical truth (and the concomitant concepts of logical consequence, logical consistency, and so forth) are non-modal, they can be used in an account of necessity without circularity.

We can use the notions of narrowly logical consistency and conceptual truth to define the concepts of *analytic consistency* and *analytic consequence*: a proposition is analytically consistent just in case it is narrowly logically consistent with the set of all conceptual truths. *P* is an analytic consequence of a set *S* of propositions just in case *P* is a narrowly logical consequence of the union of *S* with the set of all conceptual truths. The concepts of analytic consistency and analytic consequence are non-modal. I will make extensive use of them throughout this paper.

## 2. The Otherworldliness Intuition

The attempt to develop an account of necessity on the basis of the otherworldliness intuition is beset with difficulty. Consider the biconditional

(1) A proposition is necessary if and only if it is true, not only as things actually are, but in all possible situations.

(1) as it stands does not look like a promising starting point for an informative theory of modality. The notion of a possible situation seems *prima facie* to be a modal concept, and (1), far from being an informative account of what necessity is, merely seems to articulate a fairly trivial interconnection between two modal concepts.

One way out of this difficulty is Lewis's: Regard possible situations, or at least possible worlds (*maximal*[13] possible situations), as things just like the actual world, understood as the mereological sum of everything spatio-temporally connected to you and me. A possible world is simply (to simplify somewhat) a big spatio-temporal object.[14] On this view, the notion of a possible world can be explained in non-modal terms, and necessity can therefore be defined in terms of possible worlds without circularity. Most people, myself included, are not willing to accept the ontological commitments of this position. If we set aside Lewis's view of possible worlds, we are left with what is usually called an 'ersatzist' conception. Ersatzists identify possible situations with entities of some more palatable sort than Lewis's worlds, such as sets of propositions. Of course, not every set of propositions can count as a *possible* situation. In defining the notion of a possible situation, the ersatzist therefore needs to formulate some criterion for distinguishing those sets of propositions that are possible situations from those that are not. But the obvious candidate criteria (e.g., consistency with all necessary truths) are modal ones. And if we use one of these criteria, we remain trapped in the narrow and uninformative circle of modal concepts. Suppose, for the purpose of illustration, that the ersatzist defines a possible world as a set of propositions that is maximal (i.e., which, for every proposition $P$, contain either $P$ or $\sim P$) and narrowly logically consistent with all necessary truths. Assume further that he says that a proposition is true in a possible world just in case it is an element of this world, and that he interprets (1) as the claim that a proposition is necessary just in case it is true in all possible worlds. On this account, all that (1) tells us is that the necessary truths are related in a certain way to themselves: the necessary truths are just those propositions that are elements of all the maximal sets of propositions that are consistent with all necessary truths. But it is true of *every* deductively closed set $S$ of truths that it contains all and only the truths that are elements of all maximal sets of propositions that are consistent with all members of $S$. Hence, someone who does not yet know anything about necessity will, by reading (1), learn know no more about necessity than that the class of

necessary propositions is a deductively closed set of truths. There is therefore a sense in which (1) is almost completely uninformative as an attempt to explain what necessity is. But this conflicts with the intuition that we started from, namely the idea that truth in all possible situations is simply *what necessity is*, so that (1) encapsulates the essence of necessity.

Is there a way of understanding (1) that permits us to break out of the narrow circle of modal concepts and to develop an informative account of necessity on the basis of (1)? I will present what appears to me to be a suitable construal of (1) and sketch a strategy for formulating an account of necessity on the basis of it.

## 2.1. The Counterfactual Concept of Truth in a Situation

I suggested in section 1.1 that the notion of metaphysical necessity arises from certain ordinary-life practices of thinking and talking. Since our starting point is the thought that the necessity of a proposition consists in its truth, not only in the world as it actually is, but in every possible situation, we might suspect that the ordinary-life practice that gives rise to the concept of necessity is one in which we consider situations that we believe not to obtain, and ask ourselves what is true in them.[15] We can readily distinguish between two common cognitive procedures of this kind. In the one we consider the situation as *actual*, i.e. we consider it in order to determine what *is* true if the situation *actually obtains*. In the other, we consider it as *counterfactual*, i.e. with the intention of finding out what *would* have been true if the situation *had obtained*. Both of these familiar practices are often conducted by what is called 'hypothetical reasoning': we hypothetically entertain the thought that a certain situation obtains, and reason from this hypothesis and certain background knowledge to other propositions.

Consider two examples:

Bob: I hope that Fred didn't get to read Susie's letter to Bugsy.

Mary: If he did, he got over it very quickly. For I saw him earlier and he was jesting merrily.

Bob: I'm glad that Fred didn't get to read Susie's letter to Bugsy. Just imagine how he would have reacted!

Mary: He would have been writhing in agony!

In the first case, Mary considers the situation as actual, in the second case, as counterfactual. As the two examples indicate, we have special linguistic tools for expressing the outcomes of these cognitive processes: the conditional connectives. The antecedent describes the situation being supposed, and the consequent is the proposition that we believe to be true in that situation. If we consider a situation as *actual* and conclude that a certain proposition is true in it, we report this by

using the indicative conditional. When considering the situation as *counterfactual*, we use the subjunctive conditional. Philosophical research on the two kinds of cognitive procedure has often been conducted under the heading 'the semantics of conditionals.'

How do these reflections help us in interpreting (1)? Let us center on the two crucial terms on the right-hand side of the biconditional: 'possible situation' and 'true in'. Some interpretations of (1) restrict the extension of 'possible situation' to possible *worlds*, i.e. to *maximal* possible situations (i.e., possible situations in which every proposition is either true or false). The foregoing considerations, however, seem to show that there are ordinary-life practices of considering what is true in non-actual situations in which our interest is not restricted to situations of the maximal sort. I therefore suggest that we explore a broader construal of 'situation,' on which its extension includes all situations, maximal and non-maximal. On this interpretation, we can think of possible situations as sets of mutually compossible propositions; or (replacing sets of propositions by the conjunctions of their members) we can identify them simply with possible propositions.

Next, what about the notion of truth in a situation? If we construe 'true in' in (1) as expressing the concept of truth-in-the-situation-considered-as-*actual*, we obtain the following version of (1):

A proposition $Q$ is necessary if and only if $Q$ is true and, for every possible proposition $P$, if $P$ is actually true, then $Q$ is true.

But this version of (1) is false: consider the possible proposition that the yellow stuff that fills our cavities, that the rich use to decorate their ears and fingers, etc., is not a metal. We would say that, if this situation actually obtains, then gold is not a metal. Considered as actual, this situation is therefore *not* one in which gold is a metal. Hence, although the proposition that gold is a metal is necessary, it is not true in every possible situation considered as actual. This seems to show that, if we are to choose one of the aforementioned two ordinary-life notions of truth-in-a-situation in interpreting (1), it ought to be that of truth-in-a-situation-considered-as-counterfactual.

Suppose, then, that we allow the extension of 'possible situation' in (1) to include non-maximal situations and that we understand 'true in' in (1) in counterfactual terms. This yields the following version of (1): a proposition $Q$ is necessary if and only if $Q$ is true and, for every possible situation $S$, $Q$ would still have been true if $S$ had obtained. If we simply identify possible situations with possible propositions in the way suggested above, we can interpret (1) as amounting to the following principle:

(2) A proposition $Q$ is necessary if and only if $Q$ is true and, for every possible proposition $P$, $Q$ would still have been true if $P$ had been true.

This interpretation of (1) is supported, I think, by the fact that there is strong independent evidence for the claim that we are intuitively inclined to accept something like (2) as an explanation of what necessity is. For, when asked what it is for a proposition $Q$ to be necessary, it is surely natural to say something like this:

(3) To say that $Q$ could not have failed to be true is to say that $Q$ not only *is* true, but that it would have been true no matter what (else had been the case);

which seems to mean this: a proposition $Q$ is necessary just in case $Q$ is true and, for every proposition $P$, if $P$ had been true, $Q$ would still have been true.

But how are we to interpret the quantifier in 'for every proposition $P$'? Is it an unrestricted quantifier that ranges over all propositions whatsoever, or is there some restriction on its range (as there is on most quantifiers we use in ordinary discourse)? I think that there are very strong reasons for thinking that we do not pre-theoretically take every counterfactual with necessary consequent to be true. In particular, I find it fairly clear that we do not accept all counterfactuals with necessary consequents and impossible antecedents. For instance, Thatcher is not my mother, and I think that it is necessary that she is not. But I do not think that it is true to say that, if Thatcher were my mother, she would (still) not be my mother. There is therefore presumably some restriction on the quantifier over propositions that is implicit in (3), and the most obvious candidate is a restriction to possible propositions. This suggests that we read (3) as saying that

(2) A proposition $Q$ is necessary if and only if $Q$ is true and, for every *possible* proposition $P$, $Q$ would still have been true if $P$ had been true.

On this reading, (3) simply amounts to (2).

## 2.2. The Analysis of Counterfactuals and the Concept of a Possible World

So far, it will no doubt seem that the special interpretation that I have put on (1) does nothing to make (1) look like a promising starting point for an informative account of modality. There is still the modal term 'possible proposition' on the right-hand side of (2). And my counterfactual interpretation of truth-in-a-situation might appear only to make matters worse. For, on the standard account, counterfactuals are to be analyzed in terms of the notion of a possible world. If they are, then on my reading of (1), the right-hand side contains

(implicitly and explicitly respectively) both the notion of a possible world and the concept of a possible proposition.

Let me consider these worries one by one. I will first turn to the question whether counterfactuals are to be analyzed in terms of the notion of a possible world. The problematic occurrence of the term 'possible proposition' on the right-hand side of (2) will be the subject of the next section.

The basic idea underlying the standard account of counterfactuals has been nicely stated by David Lewis in the opening sentence of his book on the matter:

> "'*If kangaroos had no tails, they would topple over*' seems to me to mean something like this: in any possible state of affairs in which kangaroos have no tails, and which resembles our actual state of affairs as much as kangaroos having no tails permits it to, the kangaroos topple over."[16]

More formally, the counterfactual connective (for which I will use the symbol '□→') is defined along the following lines:

> (4) '$P \,\square\!\!\rightarrow Q$' is true just in case $Q$ is true in all the closest (i.e. most similar) metaphysically possible $P$-worlds.[17]

The possible-worlds account faces a well-known problem: If $P$ is a metaphysically impossible proposition, then there are no possible worlds in which $P$ is true. In that case, it is vacuously true that $Q$ holds in all the closest possible $P$-worlds. The possible-worlds account therefore entails that, for any metaphysically impossible proposition $P$, all counterfactuals '$P \,\square\!\!\rightarrow Q$' are true. But we have already seen in section 1.2 and in the last section that this is counterintuitive. One strategy for avoiding this problem is to formulate the closeness account, not in terms of the concept of a *possible* world, but in terms of the wider (and non-modal) notion of a world, which covers impossible worlds—worlds in which impossible propositions are true—as well as possible worlds. Daniel Nolan, among others, has argued in detail for this solution to the problem.[18] The basic idea is simple: *all* worlds, not just the possible ones, are ordered with respect to their closeness to the actual world. We can restate the closeness account as follows:

> (5) '$P \,\square\!\!\rightarrow Q$' is true just in case $Q$ is true in all the closest $P$-worlds.

(5) is just like (4), except that the modal notion of a possible world has been replaced by the wider concept of a world. For any impossible proposition $P$, there are impossible worlds in which $P$ is true. '$P \,\square\!\!\rightarrow Q$' is true if $Q$ is true in all the closest impossible $P$-worlds; it is false otherwise.

I believe that the foregoing considerations show that there are reasons quite independent of the present project for replacing the standard account (4) by (5), which makes no use of the notion of a possible world.

(5) presupposes that there are impossible worlds. That is problematic on a Lewisian realist conception of worlds, but unproblematic for an ersatzist who regards worlds as abstract entities, such as sets of propositions that meet certain conditions. Such an ersatzist can think of impossible worlds simply as sets of propositions that meet these conditions and which contain some impossible propositions. I suggest that we adopt this way of thinking of worlds for the purposes of this paper.[19] (I will say a little more about the ersatzist conception of worlds in section 3.3.)

The concept of truth-in-a-world that is used in (5) is distinct from the two notions of truth-in-a-situation that I distinguished in section 2.1 (viz. the concepts of truth-in-a-situation-considered-as-actual and truth-in-a-situation-considered-as-counterfactual). If we are thinking of worlds as sets of propositions, then we can say that a proposition is true in a world in the sense relevant to (5) just in case it is a member of the world.

## 2.3.  Ramsifying out of the Circle

Even if counterfactuals are not to be analyzed in terms of the notion of a possible world, the right-hand side of (2) still contains the modal term 'possible proposition.' We are therefore trapped in the narrow and uninformative circle of modal concepts as long as we center our attention on the property specified on one side of the biconditional in order to give an account of the modal property mentioned on the other side. I suggest that we instead focus on the property that is ascribed to the necessary truths by the biconditional *as a whole*, thereby in effect using (2) as (a partial) implicit definition of the concept of necessity. Let me explain what I mean by this.

A proposition is metaphysically possible just in case it is analytically consistent with the set of all and only the metaphysical necessities. We can use this principle to reformulate (2) as a statement, not about individual necessary propositions, but about the set containing all and only the necessary propositions:

> A proposition $Q$ is a member of the set of all and only the metaphysical necessities if and only if $Q$ is true and, for every proposition $P$ that is analytically consistent with the set of all and only the metaphysical necessities, $Q$ would still have been true if $P$ had been true.

If we uniformly replace both occurrences of expressions for the set of metaphysical necessities by instances of a variable, we obtain the following open sentence:

(2′) A proposition $Q$ is a member of $x$ if and only if $Q$ is true and, for every proposition $P$ that is analytically consistent with $x$, $Q$ would still have been true if $P$ had been true.

(2′) does not contain the notion of a possible situation or of a possible proposition. I will therefore try to break out of the narrow circle of modal notions by giving an account of necessity in terms of the property expressed by (2′).[20]

Let us call a world $w$ *analytically consistent* with a set of propositions $P$ just in case the propositions that are true in $w$ are jointly analytically consistent with $P$. And let us call a set $S$ of worlds a 'sphere' around world $w$ just in case every world in $S$ is closer to $w$ than any world not in $S$. Given certain plausible assumptions it can be proven that

(6) A deductively closed set $S$ of true propositions has the property expressed by (2′) just in case the worlds that are analytically consistent with $S$ form a sphere around the actual world.

The term 'the actual world' in (6) is to be understood in a non-rigid way, as synonymous with 'whatever world is actualized.' When (6) is evaluated with respect to another possible world $w$, the term 'the actual world' denotes $w$. It can be shown that, when so interpreted, (6) is metaphysically necessary: it is true in any metaphysically possible world $w$ that a deductively closed set $S$ of true propositions has the property expressed by (2′) if and only if the worlds that are analytically consistent with $S$ form a sphere around $w$ (For the proof, see the appendix.)

To say that the set of necessary truths has the property expressed by (2′) is thus provably equivalent to saying that the worlds that are analytically consistent with the necessary truths—the possible worlds—are closer than all the other worlds. In other words, the possible worlds are all and only those worlds that are no more than a certain distance away from the actual world. A proposition is necessary just in case it holds in all worlds that are no more than a certain distance away. On the present interpretation of (1), this is essentially what (1) tells us. But we started from the idea that (1) tells us what it is for a proposition to be necessary. We therefore naturally arrive at the idea that for a proposition to be necessary is for it to be true in every world that is no more than a certain distance away from the actual world.

Needless to say, this is not as yet a full account of necessity. It is merely a rough idea that can form the starting point for developing a theory of necessity. I will try to develop such an account in section 4. But before that, I will try to provide further motivation for going down this road. I have so far tried to motivate the idea as one way of explaining the otherworldliness intuition. I think that the same idea can also be presented as a way of developing the modalist intuition. This will be the task of section 3.

### 2.4. The Context-Dependence of Counterfactuals and the Standard Interpretation

The truth-conditions of counterfactuals are widely held to vary across contexts of use. To borrow an example from Jackson,[21] suppose Frank lives on the tenth floor of a building, and that there is nothing that could break the fall of someone jumping out of his window onto the street below. We can safely say that Frank would get hurt if he were to jump. But assume that Frank says: 'I'm a sensible chap. I would never jump from a tenth-floor window, unless I had made sure that there was a safety net. So, if I were to jump, a net would be in place, and I would be fine.' Frank's reasoning might convince us of the truth of his counterfactual. It seems that the context has shifted. In the context as it was before Frank's utterance, worlds in which Frank jumps despite the absence of a net count as closer than worlds in which he first places a net below the window and then jumps. After Frank's utterance, it is the other way around.

Such considerations convinced me that the truth-conditions of counterfactuals are context-dependent. However, like David Lewis, I believe that there is such a thing as a *default* or *standard* assignment of truth-conditions to counterfactuals, an assignment we choose when interpreting the utterance of a counterfactual unless our presumption in favor of it is cancelled by distinctive features of the context.[22] This seems plausible enough in the example of the last paragraph: If presented with the case out of the blue, we would say that Frank would get hurt if he were to jump. It requires some stage-setting (like that provided by Frank's utterance) to create a context in which we are willing to say that he would be fine.

(2′) can express different properties, depending on which reading we give to the counterfactual in (2′). I do not claim that the set of metaphysical necessities has each of the properties that can be expressed by (2′) on some reading of the counterfactual. (That would be a bold claim, given the wide variation in the truth-conditions of counterfactuals across contexts.) Instead, I stipulate that the counterfactual in (2′) is to be given the standard interpretation. I maintain that the set of metaphysical necessities has the property that (2′) expresses when it is so interpreted, and this is the claim on which I will build my theory of necessity. I do not claim that the set has the properties that are expressed by (2′) on other readings of the counterfactual.

### 3. The Modalist Intuition

### 3.1. The Inexorability Scale

The modalist intuition is the thought that necessity is a special way of being true. The necessary truths are those propositions whose truth is secure and

inexorable in a way in which the truth of contingent propositions is not. This implies that there is some dimension of inexorability on which different truths occupy different positions, and that to be metaphysically necessary is simply to have a value above a specific point on that scale.
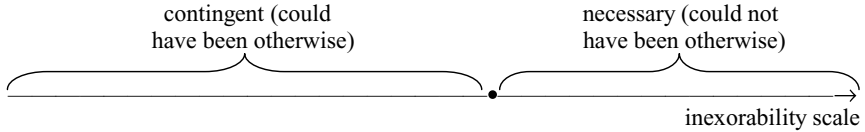


contingent (could have been otherwise)      necessary (could not have been otherwise)

inexorability scale

I suggested in section 1.1 that the notion of metaphysical necessity arises from certain ordinary-life practices of thinking and talking. We should therefore expect the inexorability dimension to be one that figures in such everyday practices. It would help in developing the modalist intuition if we could find out which practices these are. One approach to this task is to ask whether there are ordinary-language expressions other than the plain 'could' (and the concomitant expressions 'must,' 'necessary,' 'possible,' etc.) that express degrees on the inexorability scale. Now, usually we talk about the values of different objects on a single scale by using expressions to which intensifiers ('very,' 'exceedingly,' etc.) can be attached and which allow for the formation of comparatives ('more . . . than,' 'less . . . than'). When talking about the heights of different people, e.g., we use expressions (such as "tall," and "short") that allow for both kinds of transformation. We may therefore wonder whether there are English expressions that are just like "could," except that they allow for the addition of intensifiers and the formation of comparatives.

Now, we cannot simply add intensifiers to, or form comparatives of, 'could have been the case.' We cannot say, 'this very could have been the case,' or 'this could more have been the case than that.' But we can obtain an expression that allows for the addition of intensifiers and the formation of comparatives by adding an adverbial modifier to 'could have been the case':

   could easily have been the case.

We can say that such-and-such could very easily have been the case, or all too easily, or not very easily. And we can say that *x* could more easily have been the case than *y*.

But how are we to analyze the claim that a certain state of affairs could or could not easily have been the case? One approach to this question is to consider how we ordinarily evaluate the claim that a certain non-actual state of affairs could or could not easily have obtained. The usual way of doing so is to consider the range of counterfactual circumstances in which the state of affairs does obtain. We may say, e.g.,

(7) They could so easily have won the game. If the goalkeeper had stood just an inch closer to the goalpost, the other team would not have scored their second goal.

In general, if the state of affairs obtains in some worlds that depart only minimally from the actual situation (as in the case of (7)), then we want to say that it could easily have obtained. If it obtains only in worlds that depart very much from actuality, then we say that it could *not* easily have obtained:

A: They could easily have won the game.
B: I don't think so. I think that they would have won only if Fred and Susie had signed up for the team, Martha hadn't had a sore foot, Bugsy had been sober, the weather had been nice, . . . . "

This suggests that how easily a proposition $P$ could have been true is a matter of how much the worlds in which $P$ is true depart from actuality. The closer the closest $P$-worlds are, the more easily $P$ could have been true.[23]

It seems plausible to me that 'could easily have been the case' is only one among a whole range of ordinary-language expressions that can be used to express degrees on this dimension. Other expressions of this kind arguably include:

It almost happened.
It was within a hair's breadth.
This was a close call.
This was a narrow escape.[24]

And so forth. For example:

He was almost killed. If he had stood five inches further to the right, the brick would have smashed his head.

There is a counterfactual situation that departs only minimally from ours, such that he would have been killed if that situation had obtained. This supports the claim that he was almost killed.

I suggest that 'could have been otherwise' and 'could not have been otherwise' express degrees on the very same dimension that we are talking about when we ask how easily things could have been otherwise. This idea seems very plausible. Consider:

Fred's house could easily have been destroyed by yesterday's earthquake. (If the earthquake had been just a little bit stronger, the house would have been destroyed.)

Susie's house could not as easily have been destroyed. (The earthquake would have had to be much stronger to destroy Susie's house.)

Bugsy's house could not possibly have been destroyed. (Earthquakes of sufficient strength never occur in this area.)

This sequence of sentences feels intuitively like a progression along a single dimension.

This suggests that, when we are speaking about a proposition's degree of inexorability, we are talking about how easily it could have been false. To say that it has a high degree of inexorability is to say that it could not easily have been false, i.e. that the closest worlds where it is false are far away from the actual world. In the example of the last paragraph, e.g., the fact that Fred's house survived the earthquake has a low degree of inexorability, the fact that Susie's house survived has a somewhat higher degree, the fact that Bugsy's house survived has the highest degree of the three facts.

Which degree on the inexorability scale is expressed by 'could not have been otherwise' depends on the context. In philosophical discussions, the phrase often expresses metaphysical necessity. In ordinary life, it often expresses a lower degree of inexorability. (Seeing that Susie was head and shoulders above the competition, I can truly say that she could not have failed to win the tournament. This does not mean that her victory was metaphysically necessary.) It is an interesting question how it is determined which degree of inexorability is expressed by the phrase in any given context.

The key to the answer may lie in a conspicuous feature of our use of 'could': When we say that neither *A* nor *B* could possibly have failed to be true, we cannot add, in the same breath, that *A* could more easily have failed to be true than *B*. We cannot say, 'Neither Smith nor Jones could possibly have won the tournament, and Smith could have done so more easily than Jones.' By saying that *A* and *B* could not have failed to be true, we take them out of the range of propositions for which we can raise the question of how easily they could have failed to be true. In a different context, it may be true to say that *A* and *B* *could* have failed to be true, and in *that* context, it makes perfect sense to say that *A* could more easily have failed to be true than *B*. For instance, when we move into a context in which 'could' expresses metaphysical possibility, it becomes true to say that both Smith and Jones could have won the tournament. And if Smith is much better than Jones, then we can say in this new context that Smith could more easily have won than Jones.

The following analogy illuminates my attempt to make sense of these data. Suppose that I want to buy a car. I am interested in the differences between the prices of different models. But only within certain limits. I do not much care how the price of a Rolls Royce compares to that of a Mercedes, since I know that I cannot afford either car. There is a certain maximum amount of money I can spend, and differences between prices that are higher than that are of no interest

to me. That is how it often is when we are comparing the values of different objects along a certain dimension, and I claim that it is also how it is when we are comparing the values of different truths on the inexorability scale. For the purposes at hand, we may be interested in the differences between degrees of inexorability below a certain point on the scale, but we may not care about the differences between the different values that are above that point. I would like to suggest that 'could not have been otherwise' marks what we may call the "indifference point" on the inexorability scale, i.e. that point on the scale above which we do not wish to distinguish degrees of inexorability in the given context: when it is true to say, in a given context, that neither $A$ nor $B$ could possibly have been the case, then we cannot, in the same context, ask which of $A$ and $B$ could more easily have been the case. The location of the indifference point on the inexorability scale can vary from context to context.

I suggest that the use of 'could not have been otherwise' to express meta-physical necessity is simply a special case of this. The contexts in which the phrase expresses the notion of metaphysical necessity are simply those in which the indifference point has a specific (very high) value on the inexorability scale.[25] On this view, for a proposition to be metaphysically necessary is for it to be true in every world that is at most a certain distance away from the actual world. This is just the idea that we arrived at in section 2, as a result of trying to develop the otherworldliness intuition. We have thus gained independent support for the view via two different routes.
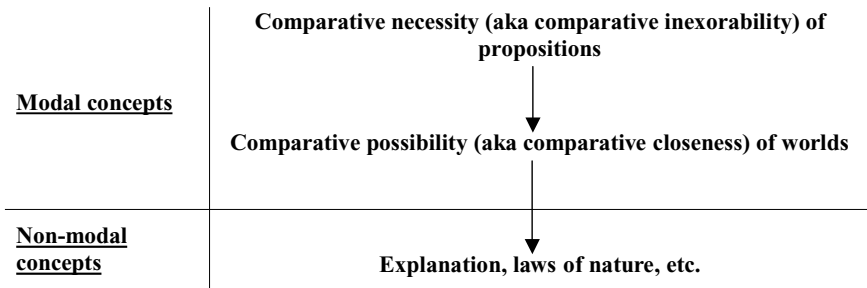
### 3.2. Modal Properties Come in Degrees

According to the previous section, for a state of affairs to be metaphysically possible is simply for it to occupy a certain position on a continuous dimension, the dimension that measures how easily different states of affairs could have obtained. Now, 'could easily have been true' is an overtly modal idiom; it is more or less just like 'could,' except that it allows for the addition of intensifiers and the formation of comparatives. It is therefore natural to speak of the dimension as measuring a proposition's *degree of possibility*. Corresponding to the degrees of possibility, there are *degrees of necessity*. The lower the degree of possibility of $\sim P$, the higher $P$'s degree of necessity. (A proposition's degree of necessity is just its degree of inexorability.) We thus arrive at the view that modal properties of propositions come in degrees. (The idea that possibility of propositions is a matter of degree, and that degrees of possibility can be defined in terms of closeness in the way described, was suggested, but not motivated or developed, by David Lewis in his book on counterfactuals[26].)

I argued in the last section that how easily a state of affairs could have obtained is measured by the closeness of the closest worlds where it obtains. When we apply this principle to maximal states of affairs or worlds, we obtain the result that a world $w$'s degree of possibility equals the degree of closeness of

the closest world where $w$ obtains. Since the closest world in which $w$ obtains is $w$ itself, this amounts to saying that a world's degree of possibility is simply its degree of closeness. The relation of comparative closeness can thus naturally be thought of as a relation of comparative possibility of worlds. To ask how close a world is to the actual world is to ask how easily the world could have been actualized, what degree of possibility it possesses.

In the last section, I analyzed inexorability in terms of closeness. This account amounts to a definition of comparative necessity for propositions in terms of comparative possibility for worlds. It reduces one modal notion to a more basic modal concept. In order to reduce the modal to the non-modal, we still need to give a non-modal account of the closeness relation. This will be the task of section 6. Drawing on the results of the previous sections and on those of my (2006), I will define closeness in terms of concepts which, in my opinion, are not themselves to be analyzed in modal terms, such as the concepts of explanation and of a law of nature.

The picture I sketched is summarized in the table below (an arrow means "is definable in terms of"):

| | **Comparative necessity (aka comparative inexorability) of propositions** |
|---|---|
| **Modal concepts** | ↓ |
| | **Comparative possibility (aka comparative closeness) of worlds** |
| | ↓ |
| **Non-modal concepts** | **Explanation, laws of nature, etc.** |

In the next section, I will attempt to provide further evidence for the claim that necessity comes in degrees, by arguing that this view allows us to give a unified account of different kinds of necessity.

### 3.3. Conceptual and Nomic Necessity

The concept of metaphysical necessity is only one of several notions of necessity that are of interest to philosophers. The most prominent other examples are the notions of conceptual and nomic (natural, nomological, physical) necessity. Like McFetridge and Lange,[27] I find it plausible that there should be something that metaphysical necessity and these other kinds of necessity have in common (and which is not possessed by any other properties), some common feature *in virtue of which* they are kinds of necessity. That is, we should expect there to be some property that is the defining feature of the higher-order property of being a kind of necessity, and which is therefore possessed by all and only the

properties that are kinds of necessity. Moreover, it seems plausible that it is an *essential* property of metaphysical necessity that it is a kind of necessity. Hence, whatever the defining feature might be of the higher-order property of being a kind of necessity, the fact that metaphysical necessity has this feature should be reflected in a correct specification of its essence. We should therefore expect a good account of the essence of metaphysical necessity to give us some idea of what this feature might be. I will argue that my account meets this condition.

The modalist intuition helps us to determine the common feature of the different kinds of necessity that makes them species of a common genus. According to the modalist intuition, for a proposition to be metaphysically necessary is for it to be true in an especially secure or inexorable way. This thought is intimately connected to the idea that, if a certain proposition is metaphysically impossible, then this is so because that there is some particularly formidable obstacle that prevents the proposition from being true. The truth of a necessary proposition is thus secured by the fact that its negation runs up against especially inexorable obstacles—perhaps that it runs up against the very metaphysical order of things.

Now, I believe that the modalist intuition also applies to nomic and conceptual necessity. Consider nomic necessity first, a kind of modal status associated in some way with the laws of nature (as well as with the metaphysical necessities).[28] It seems plausible to say that the nomic necessities are true propositions whose truth is particularly secure and inexorable, inasmuch as it is underwritten, not necessarily by the *metaphysical* order of things, but either by the metaphysical or by the *natural* (*nomological*) order: their negations run up against the metaphysical or the natural order of the world.

The modalist intuition also generalizes to conceptual necessity, but before we can fully appreciate this fact, we need to be sure to carefully distinguish between the properties of conceptual *necessity* and conceptual *truth*. The two properties are coextensive: all and only the conceptual truths are conceptually necessary. But (taking up an idea by Kit Fine[29]) I believe that they are nonetheless distinct. The property of conceptual necessity is a *modal* property; it relates to a proposition's modal force or modal status. The property of conceptual or analytic truth, by contrast, is non-modal. It is a proposition's property of being true for a certain reason, namely its being true in virtue of the fact that it is built up from certain concepts in a certain way, and in virtue of the natures of these concepts. Now, something akin to the modalist intuition about metaphysical necessity seems intuitively plausible for conceptual necessity as well: for a proposition to be conceptually necessary is for it to be true in an especially secure or inexorable way.

Conceptual, metaphysical and nomic necessity are usually thought to be ordered by their strictness, by the hardness of the 'must' associated with them (to borrow a phrase from Wittgenstein[30]): conceptual necessity is the strictest of the three kinds of necessity, nomic necessity is the least strict. This suggests that the degree of inexorability or necessity that a proposition needs to possess in order to be conceptually necessary is higher than that which it needs to possess in order

to be metaphysically necessary, and the latter degree is, in turn, higher than that which is required in order for a proposition to be nomically necessary.

On this view, there are points on the inexorability scale, $v_n$, $v_m$, $v_c$, such that to be nomically necessary is to have some value above $v_n$ on the inexorability scale, to be metaphysically necessary is to have some value above $v_m$, and to be conceptually necessary is to have some value above the still higher point $v_c$.
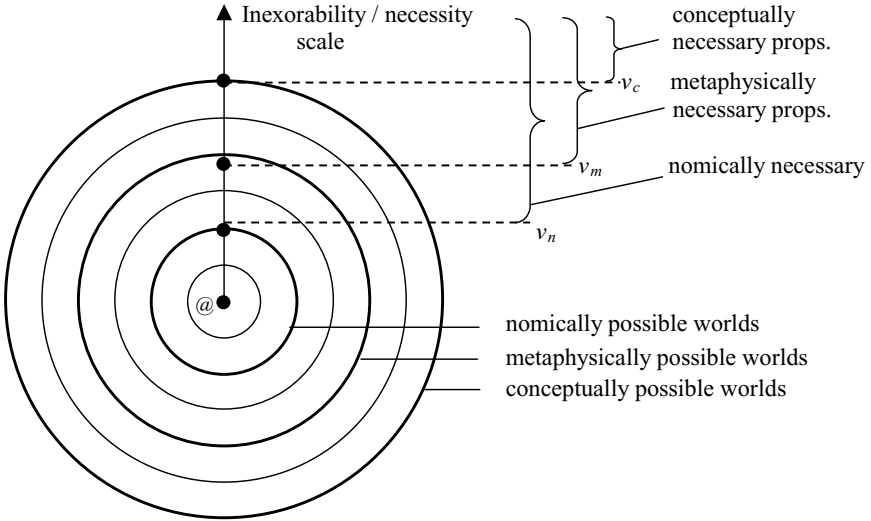


According to my account of section 3.1, a true proposition's degree of inexorability equals the distance from the actual world to the closest worlds in which the proposition is false (or, equivalently, the radius of the largest sphere throughout which the proposition is true). The individual degrees of inexorability therefore correspond directly to the individual spheres around the actual world (as represented in the foregoing diagram). According to the account of conceptual and nomic necessity suggested above, the conceptually possible worlds must therefore form a sphere around the actual world, and the same must be true of the nomically possible worlds. I think that these predictions of my account are borne out. Let me discuss the two cases one by one, starting with conceptual necessity.

*Conceptual necessity.* Before we can discuss this case, some preparatory considerations will be necessary. I argued in section 2.2 that counterpossibles, i.e. counterfactuals with metaphysically impossible antecedents, need not be vacuously true. I tried to support this claim by considering examples of counter-possibles and appealing to the intuition that their truth-conditions are not trivial. Most people I asked agree with my judgment about these examples. However, in the examples I used, the antecedents (such as 'Thatcher is my mother') were never analytically false. The responses of my subjects became more varied

when they were presented with counterfactuals whose antecedents are analytic falsehoods, such as 'If there were triangles that did not have three angles, . . . .' Such counterfactuals are quite different from the likes of 'If Thatcher were my mother . . . ,' inasmuch as their antecedents contradict themselves. Some of my subjects felt that this made it hard for them to understand what scenario they were supposed to envisage and reason about hypothetically, and therefore made it difficult for them to assign truth-values to the counterfactuals in a non-trivial way. Other subjects saw no such difficulty. I am undecided.

I said in section 2.2 that we can permit counterfactuals with metaphysically impossible antecedents to have non-trivial truth-conditions if we let impossible worlds figure in the theory of counterfactuals alongside possible worlds. A metaphysically impossible world can be analytically consistent (i.e., the propositions true at the world may be jointly narrowly logically consistent with all conceptual truths). Whether or not we want to take the further step and allow analytically inconsistent worlds to figure in our theory depends on whether we want counterfactuals with analytically false antecedents to have non-trivial truth-conditions. If we do not, then we will have no need for analytically inconsistent worlds. We can then define worlds as maximal analytically consistent sets of propositions. If we want counterfactuals with analytically false antecedents to have non-trivial truth-conditions, then we need to make room for analytically inconsistent worlds. We can do so by defining worlds, not as maximal analytically consistent sets of propositions, but simply as maximal sets of propositions, without stipulating that they need to be analytically consistent.

For our present purposes, we need not decide between the two views. We merely need to argue that on either view, the analytically consistent worlds form a sphere around the actual world, i.e.

> (8) Analytically consistent worlds are closer to the actual world than analytically inconsistent worlds.

If there are no analytically inconsistent worlds, then (8) is vacuously true. But I think that (8) is also true if there *are* analytically inconsistent worlds. In order to understand what motivates this assumption, note that, as long as an assumption we hypothetically entertain is analytically consistent, no analytic inconsistency should emerge in the course of our reflection on what would have been the case if the assumption had been true. More precisely, where $A$ is any analytically consistent proposition, every proposition $C$ for which '$A \;\square\rightarrow\; C$' is true is analytically consistent with $A$. This finding can be explained by (8), as is easy to see: Suppose that $A$ is an analytically consistent proposition, so that there are analytically consistent worlds in which $A$ is true.[31] If (8) is true, then the analytically consistent $A$-worlds must be the closest of all $A$-worlds. And since the propositions true at these worlds must be mutually analytically consistent, there can be no proposition $C$ that is analytically inconsistent with $A$ such that $A \;\square\rightarrow\; C$. (8) seems to me to provide the best explanation for the fact that

no analytic inconsistency can emerge in hypothetical reasoning from an analytically consistent assumption, and it is strongly supported by this explanatory power.

Since the analytically consistent worlds are all and only the conceptually possible worlds, we can infer from (8) that the conceptually possible worlds form a sphere around the actual world.

*Nomic necessity*. I characterized nomic necessity as a kind of necessity that is, in some way, associated with the laws of nature (as well as with the metaphysical necessities). But what exactly is the association between nomic necessity and the laws? There is no consensus on the matter. According to some philosophers, nomic necessity is a kind of necessity that attaches to the laws of nature, and to all the propositions that are metaphysically necessitated by the laws. On another view, nomic necessity also attaches to the truths about which principles are laws, e.g. to true propositions of the form *P is a law* and *P is not a law*.[32] On yet another account, truths about which principles are laws are nomically necessary, whereas the laws themselves are not.

My sympathies are with the third of these views. Let me explain why. Analytic philosophers often (though by no means universally) assume that a universal generalization cannot be a law of nature unless it is true without exception. But it seems to me that, inasmuch as we are in the business of trying to capture the pre-theoretical notion of lawhood, this assumption ought to be controversial. That is, it does not seem obvious to me that there is anything in the folk concept of a law that precludes the existence of exceptions to a law. (For many centuries belief in miracles was very common as a central component of popular religious faith, and on one natural and common way of understanding the notion of a miracle, it involves a violation of natural law.) Moreover, I think that there are theoretical pressures that make the view that laws can have small exceptions attractive. (I say more about this issue in my (2006). See Lange (2000) for a considerably more detailed discussion.) On this view, a universally quantified proposition can be a law, even though it is not, strictly speaking, true. But a false proposition cannot have any kind of necessity. I therefore do not subscribe to the idea that there is any kind of necessity that attaches to all laws of nature. Instead, I think that nomic necessity attaches to the truths about which principles are laws of nature, as well as to all propositions that are metaphysically necessitated by the truths about which principle are laws.

On this view, the nomically possible worlds are just those metaphysically possible worlds that have the same laws of nature as the actual world. It is controversial whether these worlds form a sphere around the actual world, but I have argued in detail in my (2006) that this is true.[33]

The discussion of this section suggests an answer to the question of which commonality between conceptual, metaphysical and nomic necessity makes them species of a common genus: for each of the three properties, there is a point on the inexorability scale such that for a proposition to have the property is for it to have a degree of inexorability above that point. There are many other species

of the same genus. For every sphere, there is a kind of necessity that attaches to all and only those propositions that are true in every world in that sphere. In addition to the kinds of necessity that have familiar names, such as conceptual, metaphysical and nomic necessity, there are innumerable other, nameless kinds of necessity. For any world $w$, there is a sphere that contains all and only those worlds that are at least as close to the actual world as $w$, and there is a kind of necessity that attaches to all and only those propositions that are true in every world in that sphere. This kind of necessity may not have a familiar name in English, although we can describe it. We can use the phrase "the negation of $P$ could less easily have obtained than $w$."

## 3.4. Other Modal Notions

I argued in the last section that my account of metaphysical necessity can be generalized to conceptual and nomic necessity. Now, the concepts of conceptual, metaphysical and nomic necessity are only three of the many modal notions we possess, and we may naturally wonder whether my account also generalizes to any of the others. What about the notions of

(9) Mathematical necessity
Biological necessity
Technological necessity
Practical necessity
Epistemic necessity?

I do not think that the properties listed under (9) belong to the same genus as the properties of conceptual, metaphysical and nomic necessity. I do not think that they mark degrees on the same scale of inexorability. Intuitively, it seems that the dimension of security or inexorability on which the latter three notions mark degrees is a dimension of metaphysical, mind-independent, objective security or inexorability. The metaphysical necessities are especially secure, inasmuch as their truth is secured by the metaphysical order of the world, by certain deep features of the world order (e.g., the natures of things, or the mathematical facts) that are objectively there, independently of the activities of any minds. Their negations run up against obstacles of objective, mind-independent weight and inexorability. Now, it seems that not all modal notions can plausibly be said to ascribe degrees on such a scale of objective, metaphysical security. The notion of practical necessity, e.g., seems to be a poor candidate. The practical necessities might have a special kind of inexorability *for us*, but I do not think that there is a high degree of objective, metaphysical security or inexorability that is distinctive of all and only the practical necessities. Similarly for technological necessity. Or consider the various notions of epistemic necessity. For a proposition to be

epistemically necessary is for its truth to be particularly secure in *some* sense; but the security is *epistemic*, not *ontic* or *metaphysical*. The security of the proposition consists in the fact that its truth is guaranteed by something we know, or that we have an a priori guarantee that the proposition is true, or something of that kind. There is no objective *metaphysical* security that is distinctive of the epistemic necessities.

I am equally doubtful that any of the other properties listed under (9) capture degrees of metaphysical inexorability (in the sense that having that property consists simply in being located above a certain point on the inexorability scale). Consider the case of biological necessity. If biological necessity captures a degree of metaphysical inexorability, then the truth of every biological necessity must be more secure and inexorable than that of any proposition that is not biologically necessary. But this seems intuitively implausible. There are laws of physics and chemistry that are not biological necessities (since they are not laws of *biology*), but it seems implausible that their truth is less inexorable than that of biological necessities. Similarly, I am not convinced that mathematical necessity captures a degree of inexorability. It is mathematically necessary that 2 exists; it is not mathematically necessary that Aristotle is not a musical performance (since this is not a truth of *mathematics*).[34] But the truth of the latter proposition does not seem less inexorable than that of the former. (The obstacles that prevent Aristotle from being a musical performance do not seem less weighty than those that prevent 2 from failing to exist.)

Although the notions of biological and mathematical necessity do not mark off degrees of inexorability, they may be definable in terms of notions that do. Two ways of defining new modal concepts from notions of necessity deserve attention in this context. Following Kit Fine, I will call them "restriction" and "relativization."[35] To say that a notion $N$ is defined from a kind of necessity $N^*$ by *restriction* is to say that a proposition's falling under $N$ consists in the combination of two things: (i) The fact that it possesses the kind of necessity $N^*$; and (ii) the fact that it meets certain additional conditions. To use an example of Fine's,[36] there is some plausibility to the idea that the concept of mathematical necessity is defined from the notion of metaphysical necessity by restriction. I find it attractive to say that a proposition's being mathematically necessary consists in its being metaphysically necessary for a certain reason, viz. the reason that it is a mathematical truth. To say that a notion $N$ is defined from a kind of necessity $N^*$ by *relativization* is to say that a proposition's falling under concept $N$ consists in its being necessitated, relative to the kind of necessity $N^*$, by certain propositions $P$. For example, it has some plausibility to say that the concept of biological necessity is defined from the concept of metaphysical or conceptual necessity by relativization to the laws of biology: a proposition's being biologically necessary consists in its being metaphysically or conceptually necessitated by the laws of biology.[37]

In the remaining sections I will be concerned exclusively with those modal notions that ascribe degrees of metaphysical inexorability, as conceptual,

metaphysical and nomic necessity do, whereas the concepts listed in (9) do not. Whenever I speak of 'kinds of necessity' in what follows, I will have only those properties in mind.

## 3.5. A Defense of my Direction of Explanation

On the account propounded in this paper, for a world to be metaphysically possible is simply for it to have at least a certain degree of closeness to the actual world. When first presented with this account, many philosophers grant its extensional correctness. They do not doubt that the metaphysically possible worlds are closer than the impossible ones. But they deny that this is what their metaphysical possibility *consists in*. Instead, they say, the modal notions are more basic and, in some sense, prior to the notion of closeness. They believe that it is the modal status of a world that *explains* its position in the closeness ordering: the metaphysically possible worlds are closer than other worlds *because* they are metaphysically possible.

Since this is a fairly common response to my view, I will address it in this section. My discussion will take the form of a fictional dialogue with a proponent of the response described in the last paragraph, whom I shall call "Albert."

*I*: In section 3.1, I presented evidence which, I think, supports the claim that necessity and possibility come in degrees: just as we can say that such-and-such could have been the case, we can ask how easily a certain state of affairs could have obtained. I noted that we evaluate claims about how easily a certain state of affairs could have obtained by considering the worlds in which the state of affairs obtains. The closer those worlds are to actuality, the greater the degree of possibility of the state of affairs. When applied to worlds, this reduces to the principle that the closer a world is, the greater its degree of possibility. I would like a theory of necessity to be able to accommodate these findings. Can your account do that?

*Albert*: Sure, it can. I agree that the evidence considered in section 3.1 supports the view that possibility of worlds comes in degrees. And the data about the way we evaluate claims about degrees of possibility suggests that we take the degree of possibility of a world to be higher the closer the world is. I therefore agree that

(10) World $w$ has a higher degree of possibility than world $w^*$ if and only if $w$ is closer than $w^*$.

(10) tells us that the relations of comparative possibility and of comparative closeness induce the same ordering of worlds, that they are, as we may put it, perfectly correlated. But it does not say that the two relations are *identical*, and the assumption of identity is not needed to make sense of the way we evaluate claims about comparative possibility. I, for one, resist such an identification, since

I believe that modal classifications are prior to, and more basic than, the concept of closeness.

*I*: What, then, do you make of the perfect correlation between the two relations? How do you explain it? If we identify the two relations, as I propose we do, then this question does not arise. But if we resist the identification, then we need some explanation.

*Albert*: Consider again the fact that the metaphysically possible worlds are closer than the metaphysically impossible worlds. I think that it is the special modal status of the metaphysically possible worlds that explains their position in the closeness ordering: the metaphysically possible worlds are closer than other worlds *because* they are metaphysically possible. Now, the fact that the metaphysically possible worlds are closer than the metaphysically impossible worlds is just a special case of the general correlation between comparative possibility and comparative closeness. We should expect to find a unified explanation of this correlation. Hence, if we say that the special position of the metaphysically possible worlds in the closeness ordering is due to the fact that they are metaphysically possible, then we should say that it is true quite generally that a world's degree of possibility determines and explains its position in the closeness ordering. The closeness ordering, in turn, determines the truth-values of counterfactuals in the familiar way: a counterfactual is true just in case its consequent is true in the closest antecedent-worlds.

*I*: This account strikes me as unnecessarily complicated. If the relations of comparative closeness and comparative possibility are perfectly correlated, then we do not need both of them in the theory of counterfactuals, and it seems to be a demand of good methodology to get rid of one of them. Instead of explaining the closeness ordering by appeal to comparative possibility and then explaining the truth-conditions of counterfactuals in terms of closeness, you could as well directly define the counterfactual connective in terms of comparative possibility. You could simply say that '$P \:\square\!\!\rightarrow Q$' is true just in case $Q$ is true in those $P$-worlds that have the highest degree of possibility.

*Albert*: So much the better. This move only makes my theory simpler and more attractive. I can now do without the relation of closeness altogether. Instead, I am using the notion of comparative possibility, which plays exactly the same role in my theory as the concept of closeness does in the standard account of counterfactuals.

*I*: I deny that you have really gotten rid of the closeness relation. The meaning of the term 'closeness' as used in the literature on counterfactuals (or at least as I am using it) is determined by its theoretical role: 'closeness' refers to whatever relation induces the ordering of worlds that figures in the correct theory of counterfactuals. If it is the relation of comparative possibility that induces this ordering (as on your account), then the term 'closeness' refers to the relation of comparative possibility. Hence, the relations of comparative closeness and comparative possibility are not distinct after all. They are one and the same relation.

Now, it seems plausible that to say of a world that it is metaphysically possible is to say something about its degree of possibility: it is to say that it has at least such-and-such a degree of possibility. Given the identity of the relations of comparative closeness and comparative possibility, this amounts to the view that for a world to be metaphysically possible is for it to have at least a certain degree of closeness. But this, of course, is just the account I am proposing in this paper.

## 4. Defining Necessity

According to the results of the previous section, conceptual, metaphysical and nomic necessity are different species of a common genus: each of them is a proposition's property of having at least a certain degree of inexorability. Each of the three properties should therefore have a definition of the form:

> (11) To be metaphysically (conceptually, nomically) necessary is to be true throughout that sphere of worlds around the actual world that meets condition $C_M$ ($C_C$, $C_N$),[38]

where the term "actual" is, once again, to be understood in a non-rigid sense. In order to complete our quest for a definition of metaphysical (conceptual, nomic) necessity, we still need to find a suitable replacement for '$C_M$' ('$C_C$,' '$C_N$'). We must replace this dummy by the statement of some condition which, in every possible world, singles out the sphere of all and only the metaphysically (conceptually, nomically) possible worlds. This section will be devoted to finding a suitable condition.

My strategy for achieving this goal will be somewhat indirect. I mentioned in the introduction that I find it plausible that the modal properties of a proposition need to be grounded in its non-modal features. In section 4.1 I will propose a tentative answer to the question of which non-modal features can ground the conceptual, metaphysical and nomic necessity of a proposition. I think that a correct definition of necessity needs to be compatible with that answer. In section 4.2, I will motivate a specific way of completing the definitional schema (11) by arguing that it yields a definition of necessity that meets that constraint.

### 4.1. What Grounds Modal Force?

It seems plausible to me that the special modal force of conceptually necessary propositions is grounded in their non-modal feature of being conceptual truths. The cases of metaphysical and nomic necessity are a little more difficult.

In trying to develop a working account for these cases, we may start from work by Kit Fine. As was described in section 1.3, Fine argued in detail, and

quite convincingly in my opinion, against modal definitions of the concepts of essence and essentiality. However, even if the concepts of essence and essential properties are not to be explained in modal terms, there is no denying that there is a close connection between modality and the realm of essence. The obvious suggestion, eloquently put forward by Fine,[39] is that necessity in some way flows from the essences or natures of things. The necessary truths are those that are true in virtue of the natures of things. On one first-shot way of explaining this idea, the metaphysically necessary truths are all and only those propositions that are analytic consequences of true ascriptions to entities of their essential properties, i.e. of truths of the form *a is P*, where *P* is an essential property of *a*.

Let me develop and revise this view a little bit. The true ascriptions of essential properties include, it would seem, the proposition that Elvis is a human being. But I do not think that this claim is itself metaphysically necessary. I think that there are possible worlds in which Elvis does not exist. Moreover, I believe that Elvis cannot have any properties in a world in which he does not exist,[40] and therefore cannot have the property of humanity in such a world. This point, of course, is a familiar one. And the obvious remedy is to say that, in order to obtain a necessary truth, we need to conditionalize the proposition that Elvis is human on the existence of Elvis. No comparable problem arises for the claim that this conditionalized ascription of humanity to Elvis is necessary, provided we understand the conditional connective used in it as a material conditional: the proposition *Elvis exists ⊃ Elvis is human* is equivalent to *Either Elvis does not exist or Elvis is human*, and this proposition is true even at possible worlds in which Elvis fails to exist.[41]

The conditionalization on existence yields a problem, however. It seems plausible to me that the entities of pure mathematics exist necessarily. But the existence claims of mathematics are not analytic consequences of the conditionalized statements of essential properties. (Suppose that *a* is some mathematical entity and that *P* is one of *a*'s essential properties. The conditionalized ascriptions of essential properties include the proposition *a is P if a exists*. But this proposition does not entail that *a* exists.[42]) I therefore think that not all necessary truths are analytic consequences of conditionalized essential truths. Some mathematical truths are not.

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

about the world. I will use this view as my working account of the features that ground the modal status of metaphysically necessary propositions.

I do not claim that this working account is right in all its details. It is enough for my purposes if it is approximately correct. It might be that the mathematical truths do not owe their special modal status to the fact that they are truths of mathematics, but to some more general feature, e.g. to the fact that they are truths that are purely about abstract entities, and not even in part about the concrete world. Such a view would merit exploration. For the purposes of this paper, however, I will make do with the working account stated in the last paragraph.

Given our view about what grounds metaphysical necessity, it is not hard to come up with an account of what grounds nomic necessity. On my account, the extension of nomic necessity contains all and only those propositions that are metaphysically necessitated by the truths about which principles are laws of nature. It is therefore natural to think that every nomic necessity owes its special modal status to the fact that it flows from the mathematical and (conditionalized) essential truths and the truths about which principles are laws.

The reader may worry that the mathematical truths, the conditionalized essential truths and the truths about which principles are laws do not form an interesting and unified class. Why should it be precisely these three categories of truths that are capable of grounding modal force? In section 5, I will try to assuage such worries by arguing that, if we endorse the view I stated in this section, we have an attractive way of cashing out the idea that the necessary propositions are those propositions whose truth is underwritten by particularly deep features of the world order. The picture I will sketch in section 5 will make it understandable why it is precisely the relevant three categories of truth that ground the special modal status of the metaphysical and nomic necessities.

## 4.2. Rules of Weighting, and the Definition of Necessity

A good account of necessity should be able to accommodate the finding that the modal status of metaphysical and nomic necessities is grounded in the non-modal properties mentioned in the last section. I will propose a simple-minded way of meeting this condition.

In section 3.3, I left it open whether we should allow analytically inconsistent worlds to figure in our account of counterfactuals, or whether we should define the notion of a world in such a way as to make it true by definition that worlds are analytically consistent. But I think that, *if* we allow for analytically inconsistent worlds, then we should make the following principle part of our definition of the concept of closeness:

(8) Analytically consistent worlds are closer to the actual world than analytically inconsistent worlds.

I also claim that the following two principles are part of the definition of the concept of closeness:

($R_M$) Analytically consistent worlds in which all actual mathematical and conditionalized essential truths hold are closer to the actual world than all other worlds.

($R_N$) Analytically consistent worlds in which all actual mathematical and conditionalized essential truths hold, and which have the same laws of nature as the actual world, are closer to the actual world than all other worlds.

(If we do not allow for analytically inconsistent worlds, then the words "analytically consistent" can be deleted from ($R_M$) and ($R_N$).) The expression 'actual' in (8), ($R_M$) and ($R_N$) is to be understood in the non-rigid sense. As I said in section 1.3, I take the notion of analytic consistency to be non-modal, and therefore believe that we can make (8), ($R_M$) and ($R_N$) parts of our definition of closeness without thereby importing any modal content into the notion of closeness.

The degree of closeness between two worlds is the resultant of weighting different similarities and dissimilarities between them. A theory of closeness that incorporates the rules ($R_M$) and ($R_N$) suggests the following picture of the principles for weighting similarities. These principles single out three disjoint sets of truths and order them by the weight that attaches to the truth of them at an analytically consistent antecedent-world: *firstly*, there are mathematical and conditionalized essential truths. If all actual mathematical and conditionalized essential truths hold at an analytically consistent antecedent-world, then this makes the world closer than any world that does not meet this condition. *Secondly*, there are truths about which principles are laws. If all actual mathematical and conditionalized essential truths and all actual truths about the laws hold at an analytically consistent world, then this world is closer than any world that does not meet this condition. *Thirdly*, there are truths about the individual events that take place in the natural world. These truths carry less weight than truths of the two other kinds.

On the basis of the foregoing considerations, I suggest the following definitions of conceptual, metaphysical and nomic necessity:

To be conceptually necessary is to be true throughout the sphere around the actual world that contains all and only the analytically consistent worlds.

To be metaphysically necessary is to be true throughout the sphere around the actual world that contains all and only those analytically consistent worlds in which all actual mathematical and conditionalized essential truths hold.

> To be nomically necessary is to be true throughout the sphere around the actual world that contains all and only those analytically consistent worlds in which all actual mathematical and conditionalized essential truths hold and which have the same laws of nature as the actual world.

Once again, the term 'actual' is to be understood in the non-rigid way.

This account can accommodate the views discussed in section 4.1 about which non-modal features ground the modal force of conceptual, metaphysical and nomic necessities. Consider the case of metaphysical necessity as an example. Suppose that $P$ is an analytic consequence of the mathematical and (conditionalized) essential truths. According to $(R_M)$, the analytically consistent worlds in which all actual mathematical and conditionalized essential truths hold form a sphere around the actual world. $P$ is true throughout this sphere, and in virtue of this fact $P$ is metaphysically necessary. The *reason* why $P$ is true throughout the sphere is that it is an analytic consequence of the mathematical and conditionalized essential truths. There is therefore a clear sense in which the latter feature is what *grounds* $P$'s metaphysical necessity, what P *owes* its metaphysical necessity to. This illustrates a general model for how we can explain the fact that the modal status of a certain proposition $P$ is grounded in one of its non-modal features.

## 4.3. A Simpler Account?

I suggested that the metaphysical necessities are all and only the analytic consequences of mathematical and conditionalized essential truths. If this is true, an objector might say, why bring closeness and spheres into the account of necessity at all? It seems much simpler to define metaphysical necessity as follows:

> (12) For a proposition to be metaphysically necessary is for it to be an analytic consequence of the mathematical and conditionalized essential truths.

I will argue that, despite its apparent simplicity, this account is not a viable alternative.

In section 3.3 I said that a plausible account of metaphysical necessity should permit us to identify the defining feature of the higher-order property of being a kind of necessity. I argued that my account does that. Account (12), by contrast, does not seem to offer a similar prospect of a unified account of the different kinds of necessity. Note that an account of, e.g., nomic necessity in the same spirit as (12) would probably run as follows:

> To be nomically necessary is to be an analytic consequence of the mathematical and conditionalized essential truths and the truths about which principles are laws of nature.

But what, on such a view, is the commonality between nomic necessity and metaphysical necessity that makes them both kinds of necessity? Each of the two kinds of necessity is the property of being an analytic consequence of a certain set of propositions. But not every property of being a logical consequence of a certain set of propositions is a kind of necessity. What distinguishes those properties of this kind that are kinds of necessity from those that are not? The account envisaged leaves us completely in the dark about the answer.

It is one of the most fundamental facts about metaphysical necessity that it is a kind of necessity. Hence, whatever else is true of metaphysical necessity, it must have the defining feature of the higher-order property of being a kind of necessity. An account of metaphysical necessity that makes it a mystery what this feature might be thereby lays itself open to the suspicion that it has failed to capture that feature of metaphysical necessity which makes it a kind of necessity; that it has, in fact, failed to capture that which makes it a *modal* property. But a theory of metaphysical necessity that does not capture the fact that it is a modal property must be fundamentally flawed.

I suspect that the account (12) will also be incapable of accommodating the modalist intuition, and will thereby fail to achieve one of the central goals of a metaphysical account of modality. According to the modalist intuition, it is the distinctive feature of the metaphysical necessities that their truth is particularly secure and inexorable. There must be some dimension of security or inexorability, such that for a proposition to be necessary is simply for it to have a particularly high value on that dimension. A good account of metaphysical necessity should tell us what this dimension is. And it should offer as the defining feature of metaphysical necessity the property of having a high value on that dimension. The account I offered in this paper meets this condition: The inexorability of a proposition is measured by the distance from the actual world to the closest worlds in which it fails to be true, and necessity is defined as the property of having a high value on that scale. But what about (12)? There is no mention of the inexorability scale in this proposed definition of necessity.

The failure to accommodate the modalist intuition is ultimately the same as the failure to explain what makes metaphysical necessity a kind of necessity. For I argued in section 3.3 that what makes a property a kind of necessity is that it is the property of having at least a certain degree of inexorability. The modalist intuition therefore in effect amounts to the intuition that metaphysical necessity is a kind of necessity.

I expect the proponent of (12) to reply as follows: "I can see no reason why I should not endorse the view outlined in section 3.1 about the dimension of inexorability that is relevant: the inexorability of a truth is measured by the distance from the actual world to the closest worlds in which the proposition is false. It is true that this dimension of inexorability is not mentioned in my *definition* of the concept of metaphysical necessity. But there is nonetheless a clear sense in which, even on my account, the concept of metaphysical necessity singles out a segment of the inexorability scale; for it singles out just those propositions

whose position on the inexorability scale is above a certain point. By the same token, I can give an account of what makes a property a kind of necessity. What the different kinds of necessity have in common is that each singles out the set of just those propositions that are true in all worlds that are at most a certain distance away from the actual world. And I can say that that is what makes those different properties kinds of necessity."

Such an approach to modality would be rather close to my own. In particular, its proponent would have foregone the one advantage that account (12) seemed to have over my theory, namely its greater simplicity. It is true that the definition of metaphysical necessity proffered in section 4.2 mentions spheres, whereas there is no mention of them in (12). But if we combine (12) with the account mentioned in the last paragraph of what makes a property a kind of necessity, then we have not, after all, gotten rid of the story about closeness and spheres, with all its complexity. We have just relocated it: instead of making it a part of the definition of metaphysical necessity, we have only made it part of our account of what makes metaphysical necessity a kind of necessity. Our *overall* theory of modality has not become simpler.

Simplicity considerations are therefore not relevant to a decision between the two accounts. Instead, we need to ask whether it is better to write the connection between necessity and closeness into the *definition* of metaphysical necessity, or to mention it only in the account of what makes metaphysical necessity a kind of necessity. I think that the former option is better. Let me briefly say why. The proponent of the rival account believes that the defining feature of the property of being a kind of necessity is the property of singling out those propositions that have a specific location on the inexorability scale. But this location on the inexorability scale is not mentioned in her *definition* of metaphysical necessity. This means that the fact that metaphysical necessity is a kind of necessity is not reflected in her specification of what metaphysical necessity *is*, in her proposed statement of the *essence* of metaphysical necessity. On the rival account, it is therefore not an essential, but merely an accidental, feature of metaphysical necessity that it is a kind of necessity. On the rival account, one could know what metaphysical necessity is, and also know what it is for a property to be a kind of necessity, and yet be in no position to know that metaphysical necessity is a kind of necessity. But this, I think, is rather implausible.

By the same token, I also believe that the rival account, even when developed in the way suggested in the fictional reply of its proponent, still fails to accommodate the modalist intuition. As I said in sections 1.2 and 1.3, the modalist intuition is an intuition about what it *is* for a proposition to be necessary, about the *essence* of necessity. According to the modalist intuition, it is the essence of necessity to be the property of having a high degree of inexorability. Whatever property of the necessary propositions constitutes their high degree of inexorability must be the defining feature of necessity, and must figure as such in a specification of the essence of necessity. But when we now consider (12), it is obvious that the feature

that constitutes the high degree of inexorability of the necessary truths—that of being true in all worlds that are at most a certain distance away from the actual world—is not mentioned in (12) at all; for (12) says nothing about closeness. (12) therefore does not accommodate the modalist intuition.

## 5. The Grounding of the Modal in the Non-Modal

As I mentioned in sections 1.2 and 3.3, it seems intuitive that a necessary proposition owes its modal status to the fact that its truth flows from particularly deep features of the world order, so that its truth is secured by the fact that its negation runs up against especially formidable obstacles. The deeper the features of the world order from which the truth of the proposition flows, the higher its degree of modal force. So far, I have only expressed this idea in purely metaphorical terms. In this section I will try to make the idea a little bit clearer, and try to show that my account can capture it.

More concretely, I will work on three tasks: I need to explain the non-metaphorical content of talk about *deep* features of the world order. Moreover, I will argue that on my account, degrees of depth are correlated with degrees of necessity in the way described in the last paragraph. In section 4.1 I argued that metaphysical necessities owe their modal status to the fact that they flow from the mathematical and conditionalized essential truths, and that nomic necessities owe their special modal force to the fact that they flow from the mathematical and conditionalized essential truths, and truths about the laws. This view requires us to be able to explain how these features of propositions are capable of grounding modal force. Given the correlation between depth and modal force, this requires showing that, in the relevant sense, there is something particularly deep and fundamental about the mathematical and essential truths, and also (although to a lesser degree) about the truths about the laws.

My discussion in this part of the paper will not be very rigorous. I will merely attempt to give a rough and preliminary sketch of the intuitive picture underlying my theory.

## 5.1. Counterfactuals and Explanation[43]

First, we need to consider some background facts about the relation of closeness or overall similarity between worlds that enters into the standard truth-conditions of counterfactuals. The degree of closeness between two worlds is the resultant of weighting different similarities and dissimilarities between them.[44] Not every respect of similarity carries non-zero weight. Some similarities between worlds are simply irrelevant to closeness. I will propound a principle about which similarities are relevant and which are not. I argue for this principle in detail in my (2006). In this section, I will need to be brief.

Before stating my principle, I need to discuss a couple of examples that motivate it. Let us begin by considering a kind of example that has been discussed extensively in the literature on counterfactuals. Fred is about to toss a coin. The coin-tossing setup is genuinely indeterministic and fair. You are looking on. Immediately before the coin toss, someone offers you a bet on heads. You decline. The coin comes up heads. It seems true to say, 'If you had betted on heads, you would have won.' But this presupposes that:

If you had betted on heads, the coin would still have come up heads.

Now consider:

If the coin had been tossed by Susie rather than by Fred, the coin would still have come up heads.

Almost no one believes that this counterfactual is true. If Susie had tossed the coin, then it *might* have come up heads, or it *might* have come up tails. It is not true that it *would* have come up heads.

As our intuitive judgments about the counterfactuals show, in the first example antecedent-worlds in which the coin toss has the same outcome as in our world are (all other things being equal) closer than antecedent-worlds in which it has a different outcome. By contrast, in the second example, some antecedent-worlds with a *different* outcome are among the closest antecedent-worlds. In other words, in the one case, the similarity in outcome contributes to the closeness between the worlds, while in the other case it does not.

It is not difficult to come up with an intuitively plausible explanation for this difference: Your decision whether or not to accept the bet does not make a difference to the outcome; that is, it does not *causally affect* the outcome. This is why we think that the outcome would have been just the same if you had made a different decision. The second example is different. If Susie tosses the coin, then the causal history of the outcome of the coin toss is different. Several authors who discuss pairs of examples of this kind provide diagnoses that are at least roughly along these lines.[45]

The above examples suggest, then, that

(c) Similarities between two worlds $w$ and $w^*$ with respect to individual events contribute to the closeness between the worlds if and only if these events have the same causal histories in the two worlds.[46]

I think that this point needs to be generalized and revised. My argument for this generalization will rest on the following assumption:

(L) Where $L$ is a law of nature and $E$ is an event that instantiates $L$, the fact that $L$ is a law contributes to explaining $E$. Moreover, for any law $L$, the

fact that $L$ is a law explains the general fact that events conform to $L$ (i.e. the fact that $L$ is a law explains why $L$ is true).[47]

To take an example, consider

(Law of Gravitation) Any two bodies of masses $m_1$ and $m_2$ that are at distance $d$ of each other attract one another with a force of strength $Gm_1m_2/d^2$,

where $G$ is the gravitational constant. Assume that (Law of Gravitation) is a law of nature. Last week, Mars took a certain path through space in accordance with (Law of Gravitation). I believe that the fact that (Law of Gravitation) is a law is one of the factors that together explain why Mars took the path it did. And I also believe that the fact that (Law of Gravitation) is a law explains the general fact that events conform to (Law of Gravitation); that is, it explains why bodies of masses $m_1$ and $m_2$ that are at distance $d$ of each other attract one another with a force of strength $Gm_1m_2 \ / \ d^2$.

(L) seems very plausible to me. Most other people I have asked find the principle plausible, too, and this makes me hope that the reader will find it reasonable as well. Unfortunately, any serious discussion of (L) would have to take up a lot of space, and is therefore beyond the scope of this paper.[48]

Now suppose that (Law of Gravitation) is a fundamental law of our world, i.e. that it is a law, and that this fact cannot be explained by appeal to other, more fundamental laws. Consider,

(13) If (Law of Gravitation) had not been a law, then events would still have at least approximately conformed to it.

No one I asked believed that this counterfactual was true.

How can we explain our unwillingness to accept (13)? I will focus on an explanation with considerable intuitive appeal. The only reason why events conform to (Law of Gravitation) in our world is that (Law of Gravitation) is a law. But that reason is absent in an antecedent-world. Hence, even if the events of an antecedent-world (approximately or even perfectly) conform to (Law of Gravitation), their conformity to the law does not have the same explanation as in our world. That is why it contributes nothing to closeness. Antecedent-worlds that conform to (Law of Gravitation) are no closer than those that do not. This is why there is no reason for accepting (13).

The foregoing considerations suggest the following generalization of (c):

(C) If some fact $f$ obtains in both of two worlds, then this similarity contributes to the closeness between the two worlds if and only if $f$ has the same explanation in the two worlds. (In the special case in which

> *f* has no explanation in either world, this condition counts as vacuously satisfied.)[49, 50]

I argue in my (2006) (sct. 8) that this formulation is a little bit too simple. But I think that it can serve as a working account for our present purposes.

I am using the term 'explanation' in (C) in what is sometimes called the 'ontic' sense (Salmon 1984). On this reading, the term 'explanation' expresses an objective metaphysical, non-epistemic relation: to say that fact *f* is one of the facts that jointly explain fact *g* is to say that *f* is part of the reason why *g* obtains, that *f* is one of the factors that jointly give rise to, or are responsible for, *g*. I take causation to provide the paradigmatic examples of the relation I have in mind: if *X* is a cause of *Y*, then *X* is one of the factors that are jointly responsible for, or explain, *Y*. (The reason why the ball started to move is that the player kicked it. The kick is one of the factors that explain the ball's beginning to move.) However, I think that the relation of one thing's explaining another is more general than that of causation, that it can hold between things that cannot cause each other. We have already considered an example: I believe that the fact that (Law of Gravitation) is a law explains why events conform to this law (but it would be odd to say that the lawhood of the law *causes* events to conform to the law). Similarly, I think that the lawhood of one law can explain the lawhood of another; for example, the fact that (Law of Gravitation) is a law might explain why Kepler's Laws are laws. Mathematics supplies additional examples. I think that mathematical truths often explain other mathematical truths: one mathematical fact can be the reason why another mathematical fact obtains. Moreover, truths of mathematics often contribute to explaining why certain laws of nature are laws. For example, the lawhood of derived quantitative laws can be explained by that of more fundamental quantitative laws and the mathematical truths that are used in the derivation.

So much for examples of the relation of explanation. I will assume that the relation is transitive. If *f* is one of the factors that explain *g* and *g* is one of the factors that explain *h*, then *f* is one of the factors that explain *h*.

## 5.2. Explanatory Depth and Modal Force

Let the nodes in the diagram below represent the individual facts that obtain in our world. An arrow leading from one node to another symbolizes that the fact represented by the first node contributes to explaining the fact represented by the second node. Given the assumption that explanation is transitive, if a chain of arrows leads from one node to another, then the fact represented by the first node contributes to explaining the fact represented by the second. *f*, for example, contributes to explaining *g*.
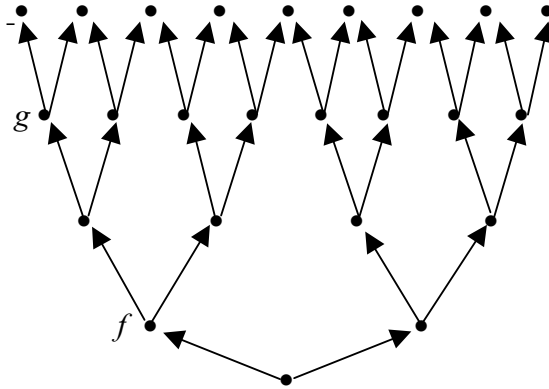
*Diagram* 1

According to principle (C) of section 5.1, the closeness of an analytically consistent world is determined by weighting its different similarities to our world in facts that have the same explanations as in our world. We can imagine that every fact $F$ represented in Diagram 1 has a certain number $x$ attached to it. If an analytically consistent world $w$ shares $F$ with our world and $F$ has the same explanation in $w$ and in our world, then the sharing of $F$ contributes with weight $x$ to the closeness of $w$ to our world. If $w$ shares $F$ with our world but $F$ does not have the same explanation in $w$ and in our world, then the sharing of $F$ has zero weight.

Now, consider the questions: How close are the closest worlds in which $f$ fails to obtain, and how close are the closest worlds in which $g$ fails to obtain? Suppose that it is analytically consistent to assume that $f$ fails to obtain and also analytically consistent to assume that $g$ fails to obtain, so that (according to (8)) the closest no-$f$ worlds and the closest no-$g$ worlds are analytically consistent
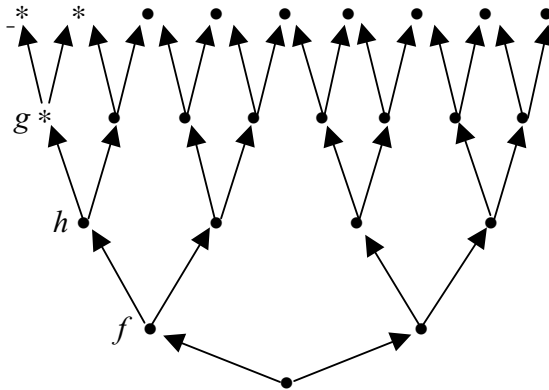


*Diagram* 2

worlds. According to the results of section 5.1, the closest no-$g$ worlds are those analytically consistent no-$g$ worlds that maximize match in facts with the same explanations as in our world. We can represent the situation diagrammatically: $g$ and facts that are explained (at least in part) by $g$ are represented by asterisks in Diagram 2, while all other facts are represented by dots. $g$ cannot obtain in any analytically consistent no-$g$ world. And if one of the other facts represented by asterisks obtains in a no-$g$ world $w$, it cannot have the same explanations as in our world, and the sharing of the fact is therefore irrelevant to the closeness between $w$ and our world. Only match in facts represented by dots ('dot facts,' for short) can contribute to closeness. The closest no-$g$ worlds will be those analytically consistent worlds that maximize match in dot facts that have the same explanations as in our world. If there are analytically consistent no-$g$ worlds that share *all* the dot facts with our world, then these will be the closest no-$g$ worlds.

Of course, there may be no such worlds. It could be that the dot facts of Diagram 2—the facts that are not $g$ and are not explained by $g$—together entail $g$, so that no analytically consistent no-$g$ world can contain all the dot facts. There are many cases of this kind. Suppose, e.g. that determinism is true,[51] and that causation always proceeds forward in time. Let $x$ be the fact that a specific kind of event occurred at time $t$. Given the temporal asymmetry of causation, $x$ does not contribute to explaining any events before $t$. $x$ also does not contribute to explaining any laws of nature (for laws of nature are not explained by individual events).[52] But under determinism the facts about the world's history before $t$, together with the laws of nature, entail that $x$ obtains.[53] Hence, either the history before $t$ is different in the closest no-$x$ worlds, or some actual law is violated in these worlds (and thus fails to be true), or both.[54] To take another example, suppose that $L_0$ is a fundamental law of nature, and that $L_1$ is a derived law whose lawhood is explained by the lawhood of $L_0$ and by certain mathematical truths that are used in the derivation. The fact that $L_0$ is a law and the relevant mathematical truths jointly entail that $L_1$ is a law. Thus, there are facts that are distinct from the fact that $L_1$ is a law and are not explained by the lawhood of $L_1$, and which jointly entail that $L_1$ is a law, and some of these facts must be absent in the closest worlds in which $L_1$ fails to be a law.

If the dot facts of Diagram 2 entail $g$, then in every analytically consistent no-$g$ world *some* of the dot facts are absent. But the closest no-$g$ worlds are those that *maximize* match in dot facts with the same explanations. Hence, they contain *most* of the dot facts of Diagram 2. Let us say, for the sake of definiteness, that $h$ is absent in the closest no-$g$ worlds (which means that match with respect to facts that are actually explained by $h$ cannot contribute to the closeness of these worlds, so that the facts that are actually explained by $h$ might also be absent in the closest no-$g$ worlds), but that all dot facts except $h$ and the facts actually explained by $h$ are present in the closest no-$g$ worlds. Thus, the facts represented by dots in Diagram 3 are all and only those actual facts that obtain in the closest no-$g$ worlds and have the same explanations as in our world:
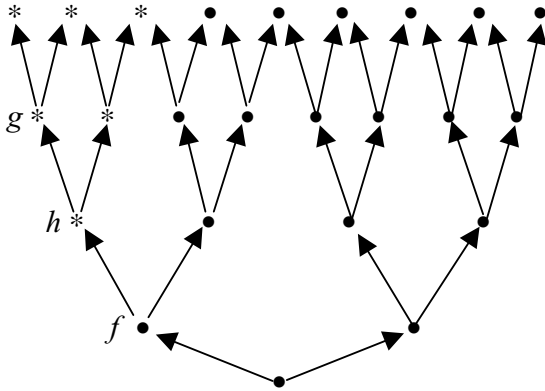
*Diagram* 3

Now ask yourself what the closest worlds in which *f* fails to obtain look like. Once again, the closest no-*f* worlds will be those analytically consistent no-*f* worlds that maximize match in facts with the same explanations as in our world. The asterisks in Diagram 4 represent *f* and the facts explained (at least in part) by *f*. All other facts are represented by dots.
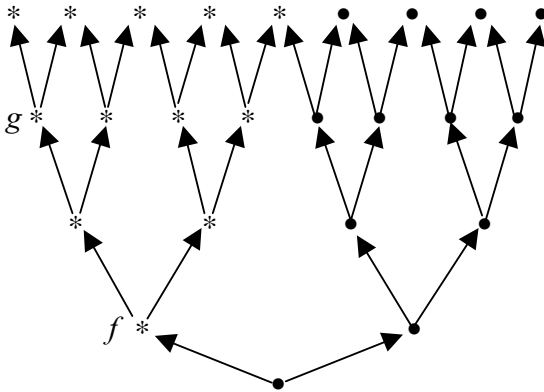


*Diagram* 4

*f* cannot obtain in any analytically consistent no-*f* world, and if other asterisk facts of Diagram 4 obtain in an analytically consistent no-*f* world *w*, they cannot have the same explanations as in our world, and the match in those facts is therefore irrelevant to the closeness between *w* and our world. The only similarities between an analytically consistent no-*f* world and the actual world that can be relevant to closeness are similarities in the dot facts of Diagram 4.

But even if the dot facts of Diagram 4 are jointly analytically consistent with the non-obtaining of $f$, so that there is an analytically consistent no-$f$ world that shares all the dot facts with our world, the combined weight of these similarities is smaller than the combined weight of match in the facts represented by dots in Diagram 3, simply because the dot facts of Diagram 4 are a proper subset of those of Diagram 3. Hence, the closest no-$f$ worlds are less close than the closest no-$g$ worlds. The difference in closeness is even more striking if the dot facts of Diagram 4 are jointly analytically inconsistent with the non-obtaining of $f$, so that there are no analytically consistent no-$f$ worlds that share all of these facts with our world.

The upshot of the foregoing considerations (which were not rigorous, and were merely intended to convey the intuitive picture) is the following rule of thumb:

> (D) By and large, if a fact $f$ is explanatorily prior to a fact $g$, in the sense that $f$ contributes to explaining $g$ but not vice versa,[55] then the closest no-$f$ worlds are farther away than the closest no-$g$ worlds; that is, $f$ could less easily have failed to obtain than $g$.[56]

For the most part, if $f$ is represented by a node closer to the bottom of Diagram 1 than $g$, then the truth that $f$ obtains is more secure and inexorable than the truth that $g$ obtains. By and large, the more explanatorily basic and fundamental a fact is (the greater the range of facts which it underlies and explains), the greater its degree of necessity.[57]

## 5.3. Mathematical Truths and Truths about the Laws

With principle (D) in mind, let us return to the principles of weighting that we considered in section 4.2. According to these principles, it is of the first importance to hold fixed the mathematical and conditionalized essential truths, and of the second importance to hold fixed the truths about what the laws are; it is less important to hold fixed truths about individual events. Let us leave aside the conditionalized essential truths for the moment (I will consider them in detail in section 5.4), and center on the other three classes of truths. It seems to me that there are what we might call "explanatory asymmetries" between these classes of truths. It seems plausible to me that truths of pure mathematics often contribute to explaining facts about lawhood. For example, the lawhood of derived quantitative laws is explained by that of more fundamental quantitative laws and the mathematical truths that are used in the derivation. Moreover, by explaining facts about lawhood, truths of mathematics can figure in the explanations of events. The converse, however, does not hold. No truth of pure mathematics is explained, in whole or in part, by facts about which principles are laws of nature or by the events in the natural world. The relation between the class of facts about lawhood and the class of facts about individual

events is exactly parallel: while facts about which principles are laws contribute to explaining facts about individual events (in accordance with principle (L) of section 5.1), facts about individual events do not explain facts about lawhood. (It it important not to misunderstand these theses of explanatory asymmetry. I do not, of course, claim that every mathematical fact figures in the explanation of every truth about what the laws are, or even that every mathematical fact figures in the explanation of *some* members of the category of truths about the laws. My thesis is that there is an explanatory asymmetry between the two *categories* or *classes* of propositions, not that there is an explanatory asymmetry between every *individual* mathematical truth and every *individual* truth about lawhood. The explanatory asymmetry between the categories consists in the fact that *some* members of the category of mathematical truths figure in the explanation of *some* members of the category of truths about the laws, whereas *no* truths about the laws figure in the explanation of *any* mathematical truths. Similarly for the relationship between the class of truths about lawhood and the class of truths about individual events.)

The foregoing considerations suggest the following picture: There is a classification of facts that carves nature at its joints, that corresponds to an important metaphysical distinction. On the one side of it, there are the facts that are, at least in part, about the realm of concrete entities, on the other side there lie the facts of pure mathematics (and perhaps other facts as well—perhaps all facts that are purely about abstract entities). Facts of the first kind are further subdivided into those about which laws govern the realm of concrete objects, and those about the individual events that occur in the concrete realm. These three major metaphysical categories of facts are ordered with respect to how explanatorily fundamental they are: the class of mathematical facts (or: of all facts about pure abstract entities?) is explanatorily prior to, or more fundamental than, the class of facts about concrete entities. And among the facts about concrete entities, those about the laws that govern the concrete realm are explanatorily more fundamental than those about individual events. This ordering of the three classes of facts by their degree of explanatory fundamentality is represented in
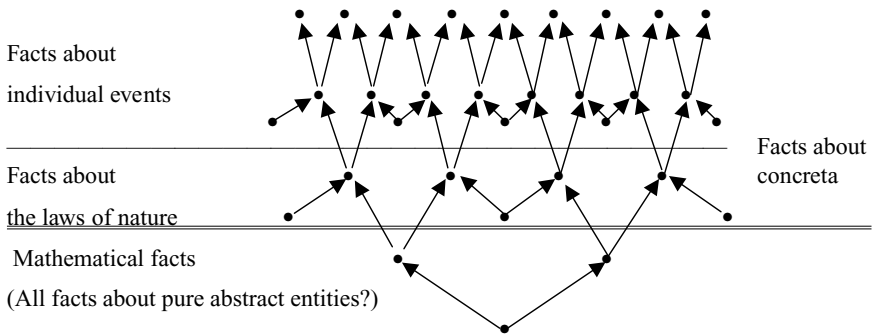


Facts about
individual events

Facts about
concreta

Facts about
the laws of nature

Mathematical facts
(All facts about pure abstract entities?)

*Diagram* 5

the diagram. The more fundamental facts are lower in the diagram, and arrows indicate explanatory relations in the usual way.

Our rules for determining the closeness ordering of worlds follow the contours of these three fundamental metaphysical categories of facts. In accordance with principle (D), they determine that the closest worlds in which the actual truths of mathematics fail to hold are less close than the closest worlds with different laws of nature, and that the latter are, in turn, less close than worlds that differ from ours only with respect to the individual events that occur in the natural world.

Let us look at the same issue from a slightly different angle. Suppose that you want to know what would have been the case if a certain fact $x$ had not obtained. (I will once again assume that it is analytically consistent to say that $x$ fails to obtain, so that (according to (8)) the closest no-$x$ worlds are analytically consistent worlds.) You need to find out what the closest no-$x$ worlds are like. Now, given that $x$ actually obtains, the totality of all facts is obviously analytically inconsistent with the non-obtaining of $x$. Hence, not all actual facts can obtain in the closest no-$x$ worlds. The question is which of the actual facts obtain in these worlds.

How can we find the answer to this question? Think of it as a game. You are given Diagram 1, and you need to remove enough nodes from the diagram to ensure that the facts represented by the remaining nodes are analytically consistent with the assumption that $x$ does not obtain. Whenever you remove one of the nodes, $n$, from the diagram, every node to which a chain of arrows leads from $n$ also disappears. Each of the nodes is assigned a certain number, which measures the weight with which the sharing of this fact contributes to the closeness between two worlds if the fact has the same explanation in the two worlds. You remove nodes from the diagram until you reach a point at which the remaining nodes represent facts that are jointly analytically consistent with the non-obtaining of $x$. At this point, your score will equal the sum of the numbers assigned to the nodes that remain in the diagram. The goal of the game is to maximize this final score. Each optimal way of playing the game (each way of playing the game that maximizes the final score) corresponds to a set of closest no-$x$ worlds, namely those analytically consistent no-$x$ worlds that share with our world all the facts represented by nodes that remain in the diagram at the end of the game.

It is important for our purposes to take into account one important rule of strategy of the game: consider again the facts $f$ and $g$ represented in Diagram 1. $f$ is explanatorily prior to $g$, i.e. $f$ figures in the explanation of $g$ but not vice versa. (Thus, there is a chain of arrows that lead from the node representing $f$ to the one representing $g$, but there is no chain of arrows leading from the $g$-node to the $f$-node.) Now, suppose that in your attempt to achieve analytic consistency with the non-obtaining of $x$, you are wondering whether to remove the dot representing $f$ or the one representing $g$. If you remove $f$, then you will lose the points attached to $f$, as well as those attached to the facts that are actually explained by $f$. If you

remove $g$, you will lose the points attached to $g$, as well as those attached to the facts that are actually explained by $g$. The latter set of facts is a proper subset of the former. Hence, by removing $f$ you would lose more points than by removing $g$. In consequence, it seems that, *ceteris paribus*, removing $g$ is the better option. More generally, when, in reasoning hypothetically from a certain assumption, we need to choose between holding fixed a fact $g$ and holding fixed another fact $f$ that is explanatorily prior to $g$, we should, all other things being equal, hold $f$ fixed.

Given that the aforementioned three classes of truths—the mathematical truths, the truths about the laws, and those about the individual events in the natural world—are ordered by explanatory priority, we can now see that the strategic principle described in the previous paragraph is enshrined in principles $(R_M)$ and $(R_N)$ of section 4.2. For what these rules tell us is that we should choose to hold fixed the mathematical truths over holding fixed the truths about the laws, and the truths about the laws over the truths about individual events. In other words, we should choose to hold fixed truths from the explanatorily more fundamental metaphysical categories of truths over holding fixed truths from the explanatorily less fundamental categories.

## 5.4. Essential Truths

So far, I have discussed only two of the three classes of propositions mentioned in $(R_M)$ and $(R_N)$. It remains to deal with the third class: the conditionalized essential truths. It seems plausible to me that the same kind of explanatory asymmetry that holds between the classes of propositions discussed in this section so far also holds between the essential truths about a thing and the accidental truths about the same thing: the essential properties of the thing usually explain many of the accidental properties, but not vice versa. For example, it seems plausible to me that it is an essential property of water to be made up molecules that consist of two hydrogen atoms and one oxygen atom. The fact that water has this molecular structure contributes to explaining many of its other (observable and unobservable) properties, but it cannot itself be explained by appeal to other properties of water. Water's having the microstructure it does is its *deepest* explanatory feature. When we are in the business of explaining some of the properties of water by others of its properties, we can trace back many of its accidental properties to its property of having the molecular structure $H_2O$, but we cannot trace the latter property to anything that is explanatorily still deeper. Once we have explained the other features by appeal to the microstructure of water, we have reached explanatory bedrock.

Consider another example: it seems plausible to me that organisms have their origins essentially. For instance, I believe that it is essential to Fred to have originated from the sperm and egg he actually did. Now, Fred's property of originating from this sperm and egg contributes to explaining many of the other

properties Fred has. (For instance, the fact that Fred originated from sperm Bob and egg cell Susie, together with certain properties of Bob and Susie (e.g. the property of containing certain genes), might contribute to explaining why Fred has a certain height.) But it seems to me that we cannot, in turn, explain *why* Fred originated from Bob and Susie by appealing to any other properties of Fred. We can causally explain the properties that Fred had at various times by appealing to properties he had at earlier times, and we can explain those properties by other properties he had at yet earlier times, and so forth. But when, in tracing back the causal chain in this way, we have arrived at Fred's property of originating from Bob and Susie, we cannot go back any further to properties Fred had at yet earlier times, since Fred did not exist at any earlier times (before he originated from the two gametes). Fred's property of originating from Bob and Susie is the *terminus* of the causal chain that links Fred's different properties, just as the microstructure of water is the terminus of the explanatory chain that links the different properties of water.[58]

Like the distinction between mathematical facts and facts about the concrete realm, and like the distinction among the latter facts between those about what the laws are and those about individual events, the distinction between the essential and the accidental facts about an entity can plausibly be regarded as a metaphysically very important distinction, a distinction that carves nature at its joints. Our principles of weighting similarities between worlds follow the contours of these metaphysical categories of facts about an entity. Under the influence of (D), the weight ordering once again follows the explanatory ordering: the essential truths about an entity are given greater weight than the accidental truths. When deciding whether to hold fixed a certain essential truth about an entity or a certain accidental truth, we should not choose to hold fixed the accidental truth over holding fixed the essential truth. We should hold fixed the essential truths if we hold fixed any truths about the entity.

$(R_M)$ enshrines this principle, as is easy to see. $(R_M)$ lays down that the analytically consistent worlds in which all actual mathematical and conditionalized essential truths hold are closer than all other worlds. Now, as I said in section 2.3 (and as is proven in the appendix), this implies that, for any proposition $P$ that is analytically consistent with all mathematical and conditionalized essential truths, all mathematical and conditionalized essential truths would still have been true if $P$ had been true. Hence, $(R_M)$ lays down that (provided the antecedent of a counterfactual is analytically consistent with the mathematical truths and the conditionalized essential truths) the conditionalized essential truths are to be held fixed. This rule in effect guarantees that the essential truths about an entity will be held fixed as long as any true ascription of a property to the entity is held fixed. For let $P$ be any property of $a$, and let $E$ be some essential property of the entity $a$. Suppose that we are wondering what would have happened if a certain antecedent $A$ had been true (where $A$ is analytically consistent with all truths of mathematics and conditionalized essential truths). Suppose that we hold fixed the truth that $a$ has property $P$. Since (in my view) this truth entails that $a$ exists,[59]

we also need to hold fixed the existence of *a*. But if we hold fixed the existence of *a*, and also hold fixed all the conditionalized ascriptions of essential properties to *a*, then we need to hold fixed all unconditionalized essential truths about *a* as well. So, ($R_M$) guarantees that, if we hold fixed any true ascriptions of properties to *a*, we need also to hold fixed all true ascriptions of *essential* properties to *a*.

As I mentioned in sections 1.2 and 3.3, it seems intuitively plausible that a proposition that has a high degree of necessity owes this special modal status to the fact that it is a consequence of propositions that state particularly deep and fundamental features of the world order. The deeper these features, the higher the degree of necessity that attaches to the proposition. The results of this section allow us to make the ideas of depth and fundamentality a little more precise: According to (D), what matters is the *explanatory* fundamentality of the facts that make a proposition true. By and large, if the truth of a proposition flows from the explanatorily most fundamental and basic facts of the world, then the proposition has a high degree of necessity. I suggested above that the difference in degree of necessity between mathematical truths, truths about the laws and truths about individual events is grounded in the differences in explanatory depth between these three ontological categories of propositions. The special modal force of conditionalized essential truths derives from the relative explanatory depth of the essential truths about the entity, compared to the accidental truths about the same entity.

The picture I sketched in the foregoing sections is connected to a certain strategy I endorse for explaining why our counterfactual connective is governed by principles ($R_M$) and ($R_N$). I cannot state this explanation in this paper. I will merely sketch the explanatory strategy and leave it to another occasion to fill in the details.

Counterfactuals serve a number of practical purposes. They figure, e.g., in decision-making ('If I were to do such-and-such, then such-and-such would happen'). Moreover, we use them to establish claims about the explanatory interrelations between different facts. (Suppose we know that *q* would not have been the case if *p* had not been the case. Given certain background assumptions, we can conclude that the fact that *p* plays a role in explaining the fact that *q*.) In order for counterfactuals to serve their purposes adequately, the rules that determine the closeness ordering need to have certain features, and we can appeal to the purposes of counterfactuals to give a functional explanation of these features. I think that it can be shown, for example, that the rules defining the closeness ordering need to incorporate principle (C) in order for counterfactuals to adequately serve their purpose in establishing claims about the explanatory relations between different facts. As we saw above, this means that these rules need to conform to the constraint stated by (D). Moreover, if counterfactual reasoning is to be a useful reasoning strategy, then the rules that determine the closeness ordering ought to be easy for us to handle. In order for this to be so, they should rest on a classification of truths that is easy to apply. This will be so if they rest on a classification that follows the contours of a conspicuous division

of truths into fundamental ontological categories. As we saw above, $(R_M)$ and $(R_N)$ rest on such a classification of truths. These two rules therefore satisfy the desiderata of conforming to (D) and being straightforward to apply.[60]

## 6. A Non-Modal Account of the Closeness Relation

In the foregoing sections, I have offered an account of modal properties of propositions in terms of the modal notion of comparative closeness. In order to complete the reduction of the modal to the non-modal, we still need an account of the closeness relation in non-modal terms. In this section, I will draw on the results of the previous discussion and on those of my (2006) to give a partial account of closeness.

The closeness ordering is the resultant of weighting different similarities between worlds, and we can define the closeness relation by specifying what weights attach to the different respects of similarity. In section 4.2, I stated some principles about the way similarities between analytically consistent worlds ought to be weighed against each other: match in mathematical and conditionalized essential truths carries the most weight, match in facts about what the laws are is next in the ordering, match in facts about individual events carries less weight than similarities of the first two kinds. It remains to consider how similarities with respect to individual events are to be weighted against each other.

Truths about such events come in two forms: on the one hand, there are what we may call *particular truths* about events, or *matters of particular fact*: truths about which events occur in the individual regions of space-time. The truth that I had coffee for breakfast this morning is an example. On the other hand, there are *general truths* about events: generalizations about individual events, i.e. truths about the general patterns that can be found in the events that occur in the world. Among these generalizations, we can distinguish between laws of nature, such as (Law of Gravitation) of section 5.1, and accidental generalizations (generalizations that do not have the status of laws). We know from Nelson Goodman's work[61] that laws can support counterfactuals, while accidental regularities cannot. In other words, the degree to which other worlds conform to the actual laws often matters to their closeness, while the degree to which they conform to the accidental regularities of the actual world is never relevant. We can therefore ignore accidental regularities in what follows, and concentrate on matters of particular fact and laws.

A law of nature must be distinguished from the fact that it is a law. For instance, (Law of Gravitation) is distinct from the fact that (Law of Gravitation) is a law. On my view, the two truths are two-way independent: (Law of Gravitation) is true in some metaphysically possible worlds where it is not a law. Hence, (Law of Gravitation) does not entail that (Law of Gravitation) is a law. Conversely, since I believe that laws of nature can have exceptions, and therefore do not think that a law of nature must be, strictly speaking, true, I do not think that

the proposition that (Law of Gravitation) is a law entails (Law of Gravitation). There are metaphysically possible worlds in which (Law of Gravitation) is a law, but is not true without exception.

Principle (C) tells us that not all similarities in truths about individual events are relevant to the closeness between two worlds. Such similarities matter only if they concern truths that have the same explanations in the two worlds. It remains to consider how we ought to weight different kinds of similarity in truths about individual events that meet this condition. For reasons discussed in detail in my (2006),[62] I accept a modified version of David Lewis's answer to this question. Let us call a violation of the actual laws that occurs in some other world but does not occur in the actual world an "alien" violation. I think that large and conspicuous alien violations of the actual laws detract more from the closeness of another world than small and inconspicuous alien violations. While small alien violations can be outweighed by increases in match in matters of particular fact, large alien violations cannot be outweighed in this way. All other things being equal, if $w$ has a greater amount of match in matters of particular fact than $w^*$ but $w$ contains more large and conspicuous alien violations of the actual laws, then $w^*$ is closer than $w$. All other things being equal, if $w$ contains more small and inconspicuous alien violations of the actual laws than $w^*$ but $w$ also has a greater amount of match in matters of particular fact, then $w$ is closer than $w^*$.

We have arrived at the following principles for weighting the similarities of an analytically consistent world $w$ to the actual world:

1. It is most important to ensure that all actual mathematical and conditionalized essential truths holds in $w$.
2. It is of the second importance to ensure sameness of the laws.
3. It is of the third importance to avoid large and conspicuous alien violations of the actual laws, provided the facts that explain the conformity of events to these laws in the actual world also obtain in $w$.
4. It is of the fourth importance to maximize match in matters of particular fact with the same explanations.
5. It is of the fifth importance to avoid small and inconspicuous alien violations of the actual laws, provided the facts that explain the conformity of events to these laws in the actual world also obtain in $w$.[63]

As before, the term "actual" is to be understood in a non-rigid sense.

So far, we have considered how similarities between analytically consistent worlds are to be weighed against each other. If we allow for analytically inconsistent worlds, we only need to make one small modification to our system of weights, in order to accommodate (8): we need to add that the single most important factor in determining the degree of closeness of another world $w$ to the actual world is whether or not $w$ is analytically consistent. Analytically inconsistent worlds are less close than analytically consistent worlds. The relative weighting of all other closeness-relevant factors remains the same.

## 7. The Utility of the Concept of Necessity

I mentioned in the introduction that it is one of the advantages of the present account of modality that it permits us to see why the notion of metaphysical necessity is useful to us: it facilitates counterfactual reasoning. Let me expand on this idea.

Counterfactual thoughts have a number of important functions in our thinking. We use them, for instance, when trying to explain things ('It is only because of you that he is not here. He would have come if you had not been so nasty to him.'), and when thinking about the likely outcomes of a contemplated course of action ('If I were to call him now, he would be able to come tomorrow'). It is therefore important for a variety of ordinary-life purposes to evaluate counterfactuals.

As we noted in section 2.1, one common way of establishing a counterfactual '$P \,\square\!\!\rightarrow Q$' is by a technique of hypothetical reasoning. To borrow an example from Jackson,[64] suppose that we want to know whether Frank would get hurt if he were to jump from the window of his tenth-floor apartment. We might start by hypothetically entertaining the hypothesis that Frank jumps, and then adding further truths as supplementary premises according to certain rules: the propositions that there is no net that could break Frank's fall, that the street is a hundred feet below the window, that bodies gain a high speed when falling a hundred feet, that human beings get hurt when hitting an asphalted surface at that speed, and so forth. Finally, we can reason from these propositions and our hypothetical assumption that Frank jumps from the window to the conclusion that Frank gets hurt. This shows that the counterfactual is true.[65, 66]

Of course, we cannot add just any old truth as supplementary premise to the antecedent. Even though Frank never in fact jumps out of his window, we cannot simply add the assumption that he does not. We can hold fixed a truth $P$ in reasoning hypothetically from an antecedent $A$ if and only if $P$ is cotenable with $A$, i.e. if and only if $P$ would still have been true if $A$ had been true.[67] Hence, before we can use a proposition as supplementary premise, we need to make sure that it is cotenable with the antecedent. This means that hypothetical reasoning needs to be based on rules that permit us to determine which propositions are cotenable with a given antecedent.

$(R_M)$ is such a rule. For $(R_M)$ entails that all analytic consequences of the set $S$ of all mathematical and conditionalized essential truths are cotenable with every antecedent that is analytically consistent with $S$. This is a corollary of the result proven in the appendix, and it is easy to see that it holds: $(R_M)$ lays down that the analytically consistent worlds in which all propositions in $S$ hold form a sphere around the actual world. Suppose that $P$ is a proposition that is analytically consistent with $S$, and is therefore true in some world in the sphere specified in $(R_M)$. Since the worlds in that sphere are closer than all other worlds, the $P$-worlds in the sphere are the closest $P$-worlds. This means that all propositions that are true throughout this sphere must be true in the closest $P$-worlds, and must

therefore be cotenable with $P$. But the propositions that are true throughout the sphere are all and only the analytic consequences of $S$. The analytic consequences of $S$ are therefore available as ancillary premises in reasoning from any antecedent that is analytically consistent with $S$.

Applications of $(R_M)$ in hypothetical reasoning are very common. Assume that we are reasoning from an antecedent that is analytically consistent with the set $S$ of all mathematical and conditionalized essential truths. Let $P$ and $Q$ be two truths, and suppose that $P$ is an analytic consequence of $S$, but $Q$ is not. If $P$ and $Q$ are individually but not jointly analytically consistent with the antecedent, then at most one of them can be used as ancillary premise. We can apply $(R_M)$ to conclude that it is $P$, and not $Q$, that can be used. It is easy to think of ordinary-life examples in which we are using $(R_M)$ in this way. Suppose that Susie, a shop keeper, receives \$10 from one of her customers and \$20 from another. What if the two customers together had given her, not \$30, but \$50? When reasoning hypothetically from this antecedent, we cannot use *both* the mathematical truth that $10 + 20 \neq 50$, *and* the truth that the two customers gave Susie \$10 and \$20 respectively. We can regard only one of these propositions as cotenable with the antecedent, and there is no doubt that it is the mathematical truth that we would regard as cotenable: We would say that if Susie had received \$50 from the two customers, then at least one of them would have given her more money than he actually did. We would *not* say that the sum of 10 and 20 would have been 50.

As for mathematical truths, so for true ascriptions of essential properties. We can use a Kripkean example to illustrate the point. Imagine a situation in which people usually use a certain odorless, colorless, tasteless liquid to quench their thirst, brush their teeth, brew their coffee, etc. More generally, they use that liquid in just the way in which we use water in our world. However, the liquid they use is not composed of hydrogen and oxygen, but of some other elements. What if a situation of this kind had obtained? We would say that, if that had been the case, people would have used some stuff other than water to quench their thirst, brush their teeth, and so forth. I think that we would *not* say that the stuff used for these purposes would have been water, but that water would not have been composed of hydrogen and oxygen. In this example, the antecedent ('People usually use a certain colorless, odorless, tasteless liquid that is not composed of hydrogen and oxygen to quench their thirst, etc.') is analytically consistent with each of the following two truths, but not with their conjunction:

> Water (if it exists) is composed of hydrogen and oxygen.

> Water is the colorless, odorless, tasteless liquid that people usually use to quench their thirst, brush their teeth, etc.

At most one of these propositions can be used as ancillary premise in hypothetical reasoning from the antecedent. As our reactions show, it is the first proposition that we regard as permissible auxiliary premise. This is simply an application of $(R_M)$: it is essential to water to be a compound of hydrogen and oxygen. Therefore,

according to ($R_M$), the proposition that water (if it exists) is composed of these elements is cotenable with the antecedent.

Now, when a speaker wants to use ($R_M$) in hypothetical reasoning from a given antecedent, she needs to proceed in two steps:

(1) First, she needs to find out whether ($R_M$) applies to the antecedent at issue, and in order to do that, she must determine whether the antecedent has the property of being true in some world in the sphere of worlds specified in ($R_M$). Determining whether a given antecedent has this property is thus a routine task that speakers face when applying ($R_M$). It would therefore be useful for the speakers to have the concept of that property. On my account, they do. It is the concept of metaphysical possibility.

(2) Suppose that a speaker has ascertained that the antecedent $A$ is true in some world in the sphere specified by ($R_M$). She can conclude that all the propositions that are true throughout this sphere are cotenable with $A$, and are therefore available as ancillary premises in reasoning hypothetically from $A$. In order to make use of this knowledge, she needs now to determine which propositions have the property of being true throughout the sphere. Determining whether a proposition has this property is therefore a routine task in evaluating counterfactuals, and it would surely be very useful for the users of counterfactuals to have the concept of that property. On my account, they do. It is the concept of metaphysical necessity.

## 8. Conclusion

In section 1 of this paper, I argued that a good theory of modality should combine an attractive metaphysical account of necessity with a plausible theory of our practice of modalizing. I also described a number of ideas about necessity that I think a good metaphysical account of modality should be able to capture: the otherworldliness intuition, the modalist intuition, and the thought that the modal status of a necessary proposition is grounded in its non-modal feature of flowing from particularly deep and fundamental features of the world order. I formulated an account of necessity according to which modal properties of propositions come in degrees, and a proposition's degree of necessity is determined by the distance from the actual world to the closest worlds in which the proposition is false. I argued that this account can accommodate the otherworldliness and modalist intuitions. Drawing on the account of counterfactuals that I developed in my (2006), I tried to show that my theory of modality also allows us to give a plausible account of how the modal properties of propositions are grounded in their non-modal features. Finally, I argued that my account permits us to explain why the concept of necessity is of use to us: it facilitates counterfactual

reasoning, by singling out those propositions that play a certain special role in that practice.

## Appendix

I will prove that

(6) A deductively closed set $S$ of true propositions has the property expressed by the open sentence (2') if and only if the worlds that are analytically consistent with $S$ form a sphere around the actual world.

As I said in section 2.3 the term "actual world" in (6) is to be understood in the non-rigid sense. After proving that (6) is true, I will argue that (6) is metaphysically necessary.

First, some terminological clarifications. Let us say that two propositions $P$ and $Q$ are *mutually analytically consistent* just in case the set containing $P$, $Q$ and all conceptual truths is narrowly logically consistent. Let us say that a world $w$ is *analytically consistent* just in case the propositions that are true in $w$ are mutually analytically consistent. (As I say in section 3.3, I wish to remain neutral on the question whether every world is analytically consistent.) A world $w$ is *analytically consistent with a set $S$ of propositions* just in case the propositions that are true in $w$ are jointly analytically consistent with $S$. A proposition $Q$ is an *analytic consequence* of $P$ just in case $Q$ is a narrowly logical consequence of $P$ and the set of all conceptual truths. I call a set *deductively closed* just in case it is closed under analytic consequence. A proposition $Q$ is *cotenable with* a proposition $P$ just in case $Q$ is true and $P \;\Box\!\!\rightarrow Q$.

Suppose that $S$ is a deductively closed set of true propositions. Using the terminology just introduced, the claim that $S$ has the property expressed by (2') can be reformulated by saying that $S$ contains all and only those propositions that are cotenable with all propositions that are analytically consistent with $S$. (6) therefore amounts to the following claim: for any deductively closed set $D$ of true propositions,

(14) $D$ contains all and only the propositions that are cotenable with all propositions that are analytically consistent with $D$

if and only if

the worlds that are analytically consistent with $D$ form a sphere around the actual world.

My proof will make use of a number of assumptions that I find plausible. Firstly, I will assume that

(Lindenbaum)  For every analytically consistent set of propositions, there is
some analytically consistent world in which all the proposi-
tions in the set are true.

In other words, every analytically consistent set of propositions can be embedded
in a *maximal* analytically consistent set. This assumption seems very plausible to
me. (Thanks to Harold Hodes for making me aware of the fact that I need this
assumption.) Secondly, I will assume that

(8)  Analytically consistent worlds are closer to the actual world than ana-
lytically inconsistent worlds.

I argue for principle (8) in section 3.3. I will also use a certain principle of
plenitude for propositions:

(15)  For any finite set $C$ of analytically consistent worlds, there is some
proposition that is true in all the worlds in $C$ and in no other analytically
consistent world.

Where $P$ is any proposition or set of propositions, let us use the expression
'$R_P$' for the class containing all and only the worlds that are analytically consistent
with $P$. If a world $w$ is analytically consistent with a set $D$ of propositions and
$P \in D$, then $\sim P$ cannot be true in $w$, and so (by the maximality of worlds) $P$
must be true in $w$. Hence, if $w$ is analytically consistent with $D$ (i.e. if $w$ is in $R_D$),
then every proposition in $D$ is true in $w$.

I will prove the two directions of my *demonstrandum* separately:

*For any deductively closed set D of true propositions, if $R_D$ is a sphere around
the actual world, then (14) is true of D.*

*Proof*: Suppose that $D$ is a deductively closed set of true propositions and that $R_D$
is a sphere around the actual world. Let $P$ be any proposition that is analytically
consistent with $D$. We can use (Lindenbaum) to infer that there are analytically
consistent worlds in which $P$ and all propositions in $D$ are true. That is to say,
there are $P$-worlds in $R_D$. Since $R_D$ is a sphere around the actual world, the
$P$-worlds in $R_D$ must be closer than all other $P$-worlds. Hence, for any $S$ in $D$,
there are $P$ & $S$-worlds (viz. those $P$-worlds that are in $R_D$) that are closer to
the actual world than any $P$ & $\sim S$-world. Hence, '$P \;\square\!\!\rightarrow S$' is true. Moreover,
since $S$ is in $D$, $S$ is true. Hence, $S$ is cotenable with $P$. We can conclude that
every proposition in $D$ is cotenable with every proposition that is analytically
consistent with $D$.

Now, let $S$ be some proposition that is not $D$. Since $D$ is deductively closed
and $S$ is not in $D$, $S$ cannot be an analytic consequence of $D$. Hence, $\sim S$
is analytically consistent with $D$. Since $\sim S$ is analytically consistent with $D$,

$\sim S$ must be analytically (self-)consistent. Hence, by (Lindenbaum), there must be analytically consistent worlds in which $\sim S$ is true. By (8), the analytically consistent $\sim S$-worlds are closer than all other $\sim S$-worlds. Hence, all the closest $\sim S$-worlds are analytically consistent worlds. No proposition that is analytically inconsistent with $\sim S$ is true in any analytically consistent $\sim S$-world. Hence, no proposition that is analytically inconsistent with $\sim S$ is true in the closest $\sim S$-worlds. Hence, $S$ is not true in the closest $\sim S$-worlds. Hence, $\sim S \, \square\!\!\rightarrow S$ is not true. So, $S$ is not cotenable with $\sim S$. Therefore, $S$ is not cotenable with all propositions that are analytically consistent with $D$. We can conclude that no proposition that is not in $D$ is cotenable with all propositions that are analytically consistent with $D$. Hence, $D$ contains every proposition that is cotenable with all propositions that are analytically consistent with $D$.

This shows that $D$ contains all and only the propositions that are contenable with all propositions that are analytically consistent with $D$. Hence, (14) is true of $D$.

*For any deductively closed set D of true propositions, if (14) is true of D, then $R_D$ is a sphere around the actual world.*

*Proof*: I will prove the contrapositive. Let $D$ be a deductively closed set of true propositions, and suppose that $R_D$ is *not* a sphere around the actual world. Since $R_D$ is not a sphere, there must be some world $w$ that is not in $R_D$ and some world $w_D$ in $R_D$, such that $w$ is no less close to the actual world than $w_D$. Since $w_D$ is an analytically consistent world and $w$ is no less close to the actual world than $w_D$, it follows by (8) that $w$, too, is an analytically consistent world. Since $w$ is both analytically consistent and analytically inconsistent with $D$, not all the propositions in $D$ can be true in $w$. That is, there must be some proposition $S$ that is in $D$ but is not true in $w$. By (15), there is a proposition $P_{w,wD}$ that is true in $w$ and $w_D$ and in no other analytically consistent world. The only worlds in which this $P_{w,wD}$ is true are $w$, $w_D$, and possibly some analytically inconsistent worlds. By (8), $w$ and $w_D$ (being analytically consistent worlds) are closer to the actual world than any analytically inconsistent $P_{w,wD}$-world. Hence, $w$ and $w_D$ are closer to the actual world than any other $P_{w,wD}$-worlds. Therefore, since $w$ is no less close to the actual world than $w_D$, $w$ is at least as close to the actual world as any other $P_{w,wD}$-world. But $S$ is not true in $w$. Therefore, $S$ is not true in all the closest $P_{w,wD}$-worlds. Hence, '$P_{w,wD} \, \square\!\!\rightarrow S$' is false; i.e. $S$ is not cotenable with $P_{w,wD}$. Thus, it is not the case that every proposition in $D$ is cotenable with $P_{w,wD}$. But there is an analytically consistent world, viz. $w_D$, in which both $P_{w,wD}$ and the propositions in $D$ are true. Hence, $P_{w,wD}$ is analytically consistent with $D$. It is therefore not the case that all propositions in $D$ are cotenable with every proposition that is analytically consistent with $D$. Thus, (14) is not true of $D$.

Principle (8), with the term 'actual' understood non-rigidly, (Lindenbaum), and principle (15) can plausibly be regarded as metaphysically necessary. (What

contingent feature of the world could their truth rest on?) Using necessitated versions of these three assumptions, we can prove the necessitated version of (6), i.e. the claim that

> ($6_N$) It is true in any metaphysically possible world $w$ that a deductive closed set $S$ of true propositions has the property expressed by (2′) if and only if the worlds that are analytically consistent with $S$ from a sphere around $w$.

## Notes

1. For discussions of the point, see Kripke (1980), p. 45, Hazen (1979), pp. 320ff., Blackburn (1984), pp. 214f., Hale (1986), p. 77, McFetridge (1990), sct. 2, and Rosen (1990), pp. 349ff.

2. Here is a typical example, taken from a paper by Edward Craig, in which the author comments on the fact that conventionalist accounts of necessary truth once enjoyed widespread popularity:

   This popularity can hardly be due to any immediate plausibility or intuitive appeal. Far from it: there is *prima facie* an element of paradox about Conventionalism, since [necessary] truths ... seem to be anything but conventional, and attract attention precisely by their *air of inexorability*. (1975, p.1, my emphasis)

3. Because of its association with the conception of necessity as a mode of truth, the view that takes modal operators as basic is often called 'modalism.'

4. My discussion of the otherworldliness and modalist intuitions, and of the way they are connected to different choices of modal primitives, owes much to Fine (1977), pp. 116–8.

5. Stalnaker (1968), Lewis (1973), (1986b).
6. (1997).
7. This idea was proposed by Lewis in his (1973), sct. 2.5.
8. (1973), sct. 2.5, McFetridge (1990), pp. 150ff.
9. Lange (1999), (2004), (2005), Williamson (2005), Kment (dissertation), Hill (forthcoming).
10. Fine (1994), p. 3, mentions Mill ((1986), bk. 1, ch. vi, §2) and Moore ((1922), pp. 293, 302) as proponents of this conception. More recent examples include Kripke (1980), and Forbes (1985), pp. 96-100, among others.
11. (1994).
12. For the purposes of this paper, I need a notion of logical truth according to which all logically true propositions are metaphysically necessary. I believe that the proposition that I exist is not metaphysically necessary. But this proposition is a truth of standard first-order logic, since in that logic we can derive it by applying the rule of universal instantiation to the logical truth that everything exists. In this paper, I will therefore be operating with a notion of logical truth that is based on free logic, rather than standard logic.
13. A situation $s$ counts as maximal if and only if, for every proposition $P$, either $P$ or $\sim P$ is true in $s$.
14. Lewis (1986a), ch. 1.
15. When I say that in the two cognitive practices I will describe we consider situations that we believe not to obtain, I merely mean that among the situations that we consider in these practices there are *some* that we believe not to obtain. There are also cases in which we use these practices to consider situations that we do believe to obtain.
16. Lewis (1973), p. 1.
17. The theoretical framework described is due to Stalnaker (1968) and Lewis (1973). Other significant work done in that framework includes Jackson (1977), Lewis (1986b), and Bennett (1984).
18. Nolan (1997).
19. While the conception of worlds as sets of propositions is a convenient working account for our purposes, it ultimately needs to be modified to get around the objection that maximal collections of propositions are too large to form sets.
20. Marc Lange ((1999), (2004), (2005)) has developed an account of logical and nomic necessity in terms of a property very similar to the one expressed by (2'). Lange arrived at his views by a route rather different from mine.
21. (1977), p.9.
22. (1986b), p. 34f.
23. I am not denying that there are also uses of the expression 'could easily have been the case' that are different from the one considered in this section. I think that in some contexts, to say that a certain state of affairs could easily have obtained is simply to say that there was a high chance that it would obtain. But I think that this cannot be the only meaning of the phrase. Even under determinism, there are situations in which it is in some sense true to say, of some event that did not actually happen, that it could easily have happened. If I know that our team would have won the match if the goalkeeper had stood an inch closer to the goalpost, then this entitles me to say that we could easily have won the game, even if I have no idea whether determinism is true. My analysis allows us to make sense of this fact.

24. I do not claim that it is the *sole* function of any of the expressions listed here to express degrees on the inexorability dimension. I am inclined to think that the relevant phrases can be used in different ways. But I believe that each of the phrases has one reading on which it ascribes degrees on the relevant dimension.

25. Needless to say, the word 'could' has a number of functions in English, including some that are quite different from the one described in this section (some of those will be considered in section 3.4). But I claim that it is *one* of the functions of 'could' to allow us to talk about degrees of inexorability.

26. Lewis (1973), sct. 2.5.

27. McFetridge (1990), Lange (2005).

28. It is commonly assumed that the metaphysical necessities are also nomically necessary.

29. Fine distinguishes between logical truth and logical necessity in his (2002), pp. 264f.

30. Wittgenstein (1953), §437.

31. I am here using an assumption that I call '(Lindenbaum)' in the appendix: for every analytically consistent set of propositions, there is some analytically consistent world in which all the propositions in the set are true. The assumption seems very plausible to me. (Thanks to Harold Hodes for making me aware of the fact that I need this assumption.)

32. See, e.g., Fine (2002), sct. 2, in particular p. 260.

33. David Lewis (1986b) defends the opposite view. Also see Bennett (1984), (2001), (2003), section 80.

34. The sentence

    (A) Aristotle is not a musical performance,

has two readings, depending on whether we give narrow or wide scope to the negation operator. On the narrow-scope reading, it ascribes to Aristotle the property of not being a musical performance. On the wide-scope reading, it says that Aristotle does not have the property of being a musical performance. Now, I believe that an entity cannot have a property at a world in which it does not exist (see footnote 39). There are many possible worlds in which Aristotle does not exist, and it is not true at such worlds that Aristotle has the property of not being a musical performance. It is, however, true at such worlds that it is not the case that Aristotle has the property of being a musical performance. The proposition expressed by (A) on the wide-scope reading is true, that expressed on the narrow-scope reading is false in the worlds under consideration. (See Adams (1981).)
I intend the sentence (A) to be given the wide-scope reading. My point is that the truth of the proposition expressed on this reading does not seem to have a lower degree of inexorability than that of the proposition that the number 2 exists. I do not think that it could more easily have been true that Aristotle is a musical performance than that 2 fails to exist.

35. Fine (2002), pp.254f..

36. Fine (2002), p. 255.

37. Note that the term 'consists in' that I am using here is a hyperintensional, constitutive idiom: It is necessarily true that a triangle is equilateral just in case it is equiangular. But its equilaterality does not *consist in* its equiangularity. Similarly, to claim that being necessary in one sense *consists in* the holding of some other

kind of necessity is to make a stronger claim than that of necessary equivalence. I believe, e.g., that it is a necessary truth that a proposition is a nomic necessity just in case it is metaphysically necessitated by the totality of truths about which propositions are laws of nature. But this is not to say that its being nomically necessary *consists in* its being metaphysically necessitated by the truths about which propositions are laws. I do not believe the latter claim. Instead, I think that being nomically necessary consists in being true in all worlds that are no more than a certain distance away from the actual world.

38. (1994), pp. 9f.

39. Thus, I subscribe to the view that an entity can have a property in a world only if it exists in that world.

40. I think that it would not have solved the problem if, instead of conditionalizing the ascriptions of essential properties, we would have maintained that the essential properties themselves are conditional, e.g. if we had said that it is not the property of humanity that is essential to Elvis, but the property of being-human-if-existing. I maintain that Elvis can have *no* property in a possible world where he does not exist, not even the property of being-human-if-existing. So, the proposition that Elvis has the property of being-human-if-he-exists is not true in all possible worlds.

41. At least this is true if our logic is a free logic, as I suggested in footnote 12 that it must be.

42. What about propositions that affirm or deny the identity between things, such as the propositions expressed by the sentences

> (1) Hesperus is Phosphorus,
> (2) Hesperus is not Alpha Centauri,

where "Hesperus," "Phosphorus" and "Alpha Centauri" are understood as directly referential expressions?

I am thinking of propositions as structured entities. The proposition expressed by a sentence containing a directly referential expression for the entity $e$ is an object-dependent proposition that contains $e$ as one of its constituents. The object-dependent proposition expressed by (1) contains Hesperus (i.e., Phosphorus) as constituent twice over. It is therefore of the logical form $a = a$. It seems plausible to say that the proposition is true in virtue of having this logical form, and is in this sense a logical truth. As a logical truth, it is a logical consequence, and therefore an analytic consequence, of every set of propositions.

Consider next the proposition expressed by (2). This proposition is of the logical form $\sim(a = b)$, but not of the form $\sim(a = a)$ (since there is no constituent that occurs twice in it). Every proposition that instantiates the first form but not the second is a true statement of non-identity, and it seems plausible to me to say that the proposition under consideration is true in virtue of the relevant two facts about its logical form. The proposition is a logical truth in the sense of being true in virtue of its logical form. As a logical truth, it is a logical consequence, and therefore an analytic consequence, of every set of propositions.

43. This section consists mainly of material from my (2006).

44. For reason which I cannot expand on here (but which I discuss in Kment (MS)), I think that this account of the closeness relation is a little bit too simple. But I think that it is a good approximation to the truth, and it can serve as a perfectly good working account for my purposes.

45. A causal diagnosis of this kind of our intuitions about relevant examples was already given in Adams (1975), ch. IV, sct. 8 (in particular pp. 132f.), though it was not formulated in the closeness framework. Causal diagnoses formulated on the basis of the closeness account can be found, e.g., in Mårtensson (1999); Edgington (2003); Bennett (2003), ch.15, Schaffer (2004), Hiddleston (2005), Kment (2006), and Wasserman (forthcoming). The different causal diagnoses differ in matters of detail.

46. The principle is called '(c)' for '*c*ausal history.'

47. In section 3.3, I said that I believed that laws can have exceptions, in the sense that a principle $L$ can be a law in a possible world $w$ even if there are exceptions to $L$ in $w$. Once we accept such a view, we do not want to rule out the possibility that there are exceptions to the actual laws in our world. That is, we want to leave open the possibility that our world does not perfectly conform to all the actual laws. But even if our world only approximately conforms to a given actual law $L$, I think that this approximate conformity can still be explained by the fact that $L$ is a law.

48. Thanks to Harold Hodes, Marc Lange, Geoffrey Sayre-McCord and Thomas Hofweber for useful discussion of the point.

49. It may be objected that there are similarities in facts with the same explanation that do not contribute to closeness, for example similarities with respect to disjunctive facts. (Perhaps it contributes nothing to the closeness between two worlds that emeralds are grue (i.e. either green and first observed before 2000, or blue and first observed after 2000) in both of them, even if the explanation of this fact is the same in the two worlds.) In order to accommodate this point, one might have to impose some restriction on the range of facts that (C) quantifies over. There are several ways in which this might be done. Unfortunately, I have no space to discuss them.

50. I called the principle '(C),' since it is a generalization of (c) but has a larger scope (hence the capital letter).

51. By "determinism" I understand the thesis that any two possible worlds that perfectly conform to the laws of our world and are alike throughout some extended initial segment of their histories are alike throughout their histories. (Note that, on the present assumption that laws can have exceptions, the actual world might not be among the worlds that perfectly conform to the actual laws.)

     If we allow for the possibility that laws have exceptions, then we cannot instead formulate determinism as the thesis that any two possible worlds that are perfectly alike throughout some extended initial segment of their histories *and which have the same laws as our world* must be alike throughout their histories. If laws can have exceptions, then this thesis can never be true. For a possible world can have the same laws as our world and yet not perfectly conform to these laws. Hence, no matter what the actual laws are like, there will be two possible worlds that have the same laws as our world and which are alike throughout some extended initial segment of history that ends at time $t$, and one of which features a violation of law after $t$ while the other one does not, so that the two worlds are not entirely alike throughout their histories.

52. In accordance with (L), I believe that, where $L$ is any law, the fact that $L$ is true is explained by the fact that $L$ is a law. In other words, on my view the laws are explained by the fact that they are laws, not by facts about individual events.

53. As I said in section 3.3, I believe that laws of nature can have exceptions. In the light of this assumption, I need to add some qualification to the claim that the facts about the history before $x$, together with the laws of nature, entail that $x$ obtains. The claim is false in one special case: it could be that the facts about the history of the world up to the time of $x$, together with the laws, entail that $x$ *fails* to obtain, but that at the time of $x$ some law is violated, so that $x$ obtains anyway.

54. I agree with Lewis that the closest no-$x$ worlds are worlds that contain tiny violations of the actual laws, and differ from our world with respect to some events immediately before the time of $x$ (see Lewis 1986 and my 2006).

55. There are possible cases in which a fact $f$ contributes to explaining another fact $g$, and $g$ also contributes to explaining $f$, e.g. in the case of causal loops. In such cases I call neither of the two facts explanatorily prior to the other.

56. The principle is called '(D)' for '(explanatory) *d*epth,' since it links a fact's degree of explanatory depth to its degree of inexorability.

57. (D) is only a rule of thumb. It is not true without exception that, if $f$ is explanatorily prior to $g$, then the closest no-$f$ worlds are farther away from our world than the closest no-$g$ worlds. Here are two examples of cases that are exceptions to the principle:

    (i) The factors that together explain a certain fact often entail the fact. Let $y$ be some fact of which this is true, and assume that it is analytically consistent to suppose that $y$ fails to obtain, so that (given principle (8)) the closest no-$y$ worlds are analytically consistent worlds. The closest no-$y$ worlds must be worlds in which some of the facts that actually explain $y$ fail to obtain. Let $x$ be one of these facts. Given that $x$ contributes to explaining $y$ and assuming that $y$ does not also contribute to explaining $x$, $x$ is explanatorily prior to $y$. But since $x$ fails to obtain in the closest no-$y$ worlds, it cannot be true that the closest no-$x$ worlds are farther away than the closest no-$y$ worlds. (The case just considered is represented in Diagram 3: fact $h$ is explanatorily prior to fact $g$; but $h$ fails to obtain in the closest no-$g$ worlds, so that the closest no-$h$ worlds can be no farther away than the closest no-$g$ worlds.)

    (ii) Consider a case of causal over-determination in which $A$ and $B$ are counterfactually independent of each other and are individually sufficient causes of $C$. In the closest no-$A$ worlds, $B$ is still present and causes $C$. Likewise, in the closest no-$B$ worlds, $A$ is still present and causes $C$. The closest worlds in which $C$ is absent are worlds in which neither $A$ nor $B$ is present, and those worlds are farther away than the closest no-$A$ worlds and the closest no-$B$ worlds. Even though $A$ and $B$ are explanatorily prior to $C$, $C$ could less easily have failed to be the case than either $A$ or $B$.

58. To be more precise, I do not want to rule out completely the possibility that an essential property of a thing can be explained by other properties of the thing. It might be so explicable, but if it is, then the other properties by appeal to which we can explain it must themselves be essential. (For instance, suppose that it is essential to Fred that he is a hippopotamus. This property might be explicable by the fact that Fred originated from sperm Bob and egg Susie, together with the fact that Bob and Susie are hippopotamus gametes. But note that the property of Fred's that we appeal to in order to explain the fact that he is a hippopotamus

(viz. that of originating from Fred and Susie) is itself an essential property of Fred.) My thesis is merely that no essential property of an entity can be explained by an *accidental* property of the same entity, whereas accidental properties can often be explained by essential properties. There is a clear explanatory asymmetry between the essential and the accidental facts about an entity.

59. As I mentioned in footnote 39, I believe that no entity has properties in any metaphysically possible world where it does not exist.

60. Principle (D) constrains what the principles of weighting can be like. But it does not entail ($R_M$) or ($R_N$). There are other sets of principles that we could have used to weight similarities, and which also conform to the constraint imposed by (D). Nonetheless, I think that there is something special about a system of rules that includes ($R_M$) and ($R_N$). Every possible system of weights divides up the facts about our world into different classes and orders these by how weightily it contributes to the closeness of another world if it shares with the actual world facts in the relevant classes. Some possible sets of rules rely on classifications of facts that carve nature at its joints, others rely on more gerrymandered classifications. As I said, ($R_M$) and ($R_N$) are rules of the former kind. They rely on metaphysically very significant and central classifications of facts, perhaps the most significant metaphysical distinctions between facts which there are. I conjecture that, of all the possible principles of weighting that conform to the constraints imposed by (D), there are none that rest on a metaphysically deeper and more important classification of facts. In consequence, I surmise that ($R_M$) and ($R_N$) optimally satisfy the desideratum of being easy to apply while conforming to the constraint imposed by (D).

61. (1947).

62. Scts. 5–7.

63. This specification of the closeness relation is cast in the same form as Lewis's well-known formulation in his (1986b), pp. 47f.

64. Jackson (1977), p. 9.

65. If the antecedent is analytically inconsistent and we allow for analytically inconsistent worlds, then the closest antecedent-worlds are analytically inconsistent. In that case, there may be some restriction on the principles of inference that can be used in the hypothetical reasoning, since some of the inference rules that are actually valid might be invalid in the closest antecedent-worlds. But if the antecedent is analytically consistent, then all actually valid principles of inference are applicable without restriction.

66. From the perspective of the closeness account, the process of hypothetical reasoning can be described as follows: Suppose that we want to find out whether '$P \,\Box\!\!\rightarrow Q$' is true. This requires us to determine whether the $P$-worlds closest to the actual world are $Q$-worlds. We can do so by applying both our background knowledge about the actual world and the rules defining the closeness relation to show that certain true propositions are also true in the closest $P$-worlds, and by then reasoning from $P$ and the relevant truths in order to determine whether they entail $Q$. If they do, then we can conclude that $Q$ is true in the closest $P$-worlds and that '$P \,\Box\!\!\rightarrow Q$' is therefore true.

67. This definition is slightly different from the one Goodman originally proposed ((1955), p. 15). On his definition, a proposition $P$ is cotenable with an antecedent

*A* just in case *P* is true and it is not the case that *P* would not have been true if *A* had been true.

# References

Adams, E. (1975), *The Logic of Conditionals*, Dordrecht (Reidel)

Adams, R. M. (1981), "Actualism and Thisness," *Synthese* 57, pp. 3–42

Bennett, J. (1984), "Counterfactuals and temporal direction," *Philosophical Review* 93, pp. 57–91

Bennett, J. (2001), "On Forward and Backward Counterfactual Conditionals," in: *Reality and Humean Supervenience*, ed. G. Preyer and F. Siebelt, Lanham (Rowman and Littlefield), pp. 177–202

Bennett, J. (2003), *A Philosophical Guide to Conditionals*, Oxford (Clarendon)

Blackburn, S. (1984), *Spreading the Word. Groundings in the Philosophy of Language*, Oxford (Clarendon)

Craig, E. (1975), "The Problem of Necessary Truth," in: *Meaning, Reference and Necessity*, ed. S. Blackburn, Cambridge (Cambridge UP), 1–31.

Edgington, D. (2003), "Counterfactuals and the Benefit of Hindsight," in: *Causation and Counterfactuals*, ed. P. Dowe and P. Noordh of, London (Routledge)

Fine, K. (1977), postscript to A. N. Prior's *Worlds, Times, and Selves*, London (Duckworth), 116–8

Fine, K. (1994), "Essence and Modality," *Philosophical Perspectives* 8, pp. 1–16

Fine, K. (2002), "The Varieties of Necessity," in: *Imagination, Conceivability and Possibility*, ed. J. Hawthorne and T. Gendler, Oxford (Clarendon), pp. 253–81

Forbes, G. (1985), *The Metaphysics of Modality*, Oxford (Clarendon)

Goodman, N. (1947), "The problem of counterfactual conditionals," *Journal of Philosophy* 44, pp. 113–28

Goodman, N. (1955), *Fact, Fiction and Forecast*, Indianapolis (Bobbs-Merrill).

Hale, B. (1986), "The Compleat Projectivist," *Philosophical Quarterly* 36, pp. 65–84

Hazen, A. (1979), "Counterpart-Theoretic Semantics for Modal Logic," *Journal of Philosophy* 76, pp. 319–38

Hiddleston, E. (2005), "A Causal Theory of Counterfactuals," *Noûs* 39(4), pp. 632–657

Hill, C. (forthcoming), "Modality, Modal Epistemology, and the Metaphysics of Consciousness," in S. Nichols (ed.), *The Architecture of the Imagination: New Essays on Pretense, Possibility, and Fiction*, Oxford (UP)

Jackson, F. (1977), "A Causal Theory of Counterfactuals," *Australasian Journal of Philosophy* 55, pp. 3–21

Kment, B. (2006), "Counterfactuals and Explanation," *Mind* 115, pp. 261–310

Kment, B., MS, "The Similarity Account of Counterfactuals"

Kment, B. (dissertation), *A Theory of Counterfactuals and a Counterfactual Theory of Necessity*, Princeton 2006

Kripke, S. (1980), *Naming and Necessity*, Cambridge, Mass. (Harvard UP)

Lange, M. (1999), "Laws, Counterfactuals, Stability, and Degrees of Lawhood,"*Philosophy of Science* 66, 243–67

Lange, M. (2000): Natural Laws in Scientific Practice. New York: Oxford University Press.

Lange, M. (2004), "A Note on Scientific Essentialism, Laws of Nature, and Counterfactual Conditionals," *Australasian Journal of Philosophy* 82, 227–241

Lange, M. (2005), "A Counterfactual Analysis of the Concepts of Logical Truth and Necessity," *Philosophical Studies* 125, pp. 277–303

Lewis, D. K. (1973), *Counterfactuals*, Oxford (Blackwell)

Lewis, D. K. (1986a), *On the plurality of worlds*, Oxford (Blackwell)

Lewis, D. K. (1986b), "Counterfactual Dependence and Time's Arrow," in Lewis 1986c, pp. 32–66.

Lewis, D. K. (1986c), *Philosophical Papers*. New York Oxford: Oxford University Press, vol. ii.

Mårtensson, J. (1999), *Subjunctive conditionals and Time*, Göteborg (Kompendiet), online at http://www.phil.gu.se/johan/ThesisGM.pdf

McFetridge, I. (1990), "Logical Necessity: Some Issues," in: *Logical necessity, and other essays*, ed. by J. Haldane and R. Scruton, London (Aristotelian Society), pp.135–54

Mill, J. S. (1986), *A System of Logic, Ratiocinative and Inductive*, Charlottesville, Va. (Lincoln-Rembrandt)

Moore, G. E. (1922), "External and Internal Relations," in: *Philosophical Studies*, London (Routledge and Kegan), pp. 276–309

Nolan, D. (1997), "Impossible Worlds: A modest approach," *Notre Dame Journal of Formal Logic* 38, pp.535–73

Rosen, G. (1990), "Modal Fictionalism," *Mind* 99, pp. 327–54

Salmon, W. C. (1984), "Scientific Explanation: Three General Conceptions," in: *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1984*, ed. P. D. Asquith and P. Kitcher, East Lansing (MI: Philosophy of Science Association), pp. 293–305

Schaffer, J. (2004), "Counterfactuals, Causal Independence and Conceptual Circularity," *Analysis* 64, pp. 299–309

Slote, M.A. (1978), "Time in Counterfactuals," *Philosophical Review* 87, pp. 3–27

Stalnaker, R. (1968), "A Theory of Conditionals," in *Studies in Logical Theory, American Philosophical Quarterly Monograph Series* 2, Oxford (Blackwell), pp. 98–112

Tichý, P. (1976), "A counterexample to the Stalnaker-Lewis analysis of counterfactuals," *Philosophical Studies* 29, pp. 271–3

Wasserman, Ryan. Forthcoming: 'The Future Similarity Objection Revisited'. Forthcoming in *Synthese*.

Williamson, T. (2005), 'Armchair Philosophy, Metaphysical Modality and Counterfactual Thinking,' Proceedings of the Aristotelian Society 105(1).

Wittgenstein, L. (1953), *Philosophical Investigations*, New York (Macmillan)