MODAL LOGIC

IN THE MODAL SENSE OF MODALITY

for Jouko Väänänen

ABSTRACT. One reason is explained why the pursuit of modal logic's original aim of evaluating the formal validity of arguments turning on necessity and possibility in a 'metaphysical' sense must cross borders and draw the logician into 'metaphilosophy'.

John P. Burgess
Department of Philosophy
Princeton University
Princeton, NJ 08544-1006, USA
jburgess@princeton.edu

1

MODAL LOGIC IN THE MODAL SENSE OF MODALITY

1.   MODAL LOGIC & BORDER-CROSSING

Originally 'geometry' named an enterprise aiming to describe the forms of physical space. Today geometry is the study of a vast range of mathematical structures, closely enough related to structures once thought to model physical space that they can usefully be studied by the same methods or adaptations thereof. Some of these structures have important applications, as with the use of phase space in statistical mechanics; but as this example shows, the applications may have little to do with the original aims of the subject. The part of geometry devoted to the original aim is now but a small corner of a large field, and it is the more marginal because it involves a crossing of disciplinary borders. It does so because it is no longer tenable to view geometry as determining in advance the form of the arena in which physical forces act, before electromagnetic and gravitational theory, say, get to work describing those forces. On the contrary, the pursuit of what were the original aims of geometry is now inseparable from electromagnetic and gravitational theory themselves (which to begin with have taught us that we must not speak of physical space, but rather of physical spacetime).

Originally 'logic' named an enterprise aiming to prescribe the norms

of deductive argumentation in philosophy and elsewhere. Today logic is the study of a vast range of formal systems, closely enough related to systems once thought to model deductive argumentation that they can usefully be studied by the same methods or adaptations thereof. Some of theses systems have important applications, as with the use of boolean logic in circuit design; but as this example shows, the applications may have little to do with the original aims of the subject. The part of logic devoted to the original aim is now but a small corner of a large field, and it is the more marginal because it involves a crossing of disciplinary borders. It must do so because it is no longer uncontroversial to assume that logic can serve as an organon, determining in advance the norms according to which substantive philosophical or other issues should be debated, before philosophers or others get to work debating. On the contrary, the pursuit of what were the original aims of logic is now for many inseparable from philosophy generally, and philosophy of language especially.

As it is with logic in general, so it is with modal logic in particular. But what is meant by 'modal' here? Mood-and-modality is the linguistic category pertaining to the expression of irreality of one kind or another, insofar as this is done grammatically rather than lexically, which is to say, insofar as it is expressed by inflections of the verb, or by use of special

modal auxiliary verbs. Linguists distinguish deontic, epistemic, and dynamic modality. All are of interest in philosophy, but most of all the third, which more or less amounts to what philosophers call 'metaphysical' modality. Its key notions are expressible in various ways, but chiefly by use of the modal auxiliaries 'could' and 'would' and 'might'. For those key notions are the necessary in the sense of *what is and would have been (could not have failed to be) no matter what*, and the possible in the sense of *what is or isn't but potentially might (or could) have been*. Henceforth 'necessary' and 'possible', and with them 'modality', when used without distinguishing epithet, should be understood as pertaining to this kind rather than any other. It will here be our default sense of modality. Something very like this sense — or anyhow, something much more like it than like any other sense in play today — seems to have been assumed in the first work on modal logic, Aristotle's modal syllogistic. And therefore without much exaggeration or distortion we can say that to articulate the logic of this notion of modality was modal logic's original aim.

Modal logic today is an enormously broader enterprise than it was in Aristotle's time. Much of modal logic seems to have been developed for its purely mathematical interest, since the number of formal systems far exceeds the number of conceivable intuitive interpretations. But there are

some significant external applications, as with the use of dynamic logic in theoretical computer science. The model-theoretic machinery developed to bring order out of the chaos of competing systems is indeed found invoked today in many areas of logic and across the border from pure logic in theoretical computer science and linguistics. Very little of modal logic is devoted to the attempt to determine what are the valid forms of argument involving necessity and possibility in *any* sense, and vanishingly little is devoted to the original aim of modal logic, that of dealing with the logic of modality in our default sense.

In fact, apart from some scolding by critics of modern modal logic about the dangers of 'Aristotelian essentialism' — which were met by firm denials from modal logicians that their logic harbored any 'essentialist commitments' — until 1970 the very topic of modality in our default sense was lost sight of, not only in logic, but in philosophy as well. Today what were the original aims of modal logic have begun to be pursued again, though the pursuit may not as yet have gotten very far. As with geometry and with logic in general, pursuit of the original aims will require crossing borders into another domain.

In the remarks that follow I will indicate how, in particular, one must enter into the realm of metaphysics, and beyond that into the realm of

'metaphilosophy', where the issues are about *what the issues are about* when debating metaphysics. One must confront a notorious 'modal paradox', and consider not only various views about how to deal with the paradox, but also various meta-views about what 'dealing with' a paradox can mean.

2. MODALITY LOST

But how did the neglect, until a few decades ago, of modality, in what I have claimed is the default sense of the term, come about? Well, not all at once, but in stages. The 'metaphysical' notion of *what would have been no matter what* (the necessary) was conflated with the epistemological notion of *what independently of sense-experience can be known to be* (the *a priori*), which in turn was identified with the semantical notion of *what is true by virtue of meaning* (the analytic), which in turn was reduced to a mere product of human convention. And what motivated these reductions? Mainly, I would suggest, efforts at 'demystification'. Let me outline the story.

The mystery of modality, for early modern philosophers, was how we can have any knowledge of it. Here is how the question arises. We think that when things are some way, in some cases they could have been otherwise,

and in other cases they couldn't. That is the modal distinction between the contingent and the necessary. And we not only think that there is such a distinction, we also think that we know some examples. For instance, though there has been no female U. S. president, there could have been, whereas the number 29 not only has no nontrivial divisors but couldn't have had any. Thus we may speculate counterfactually about 'What if Hilary Clinton had been president in September, 2001?' but won't speculate counterfactually about 'What if 5 had been a divisor of 29?' And what is called the question of modal epistemology is just this: How do we know that the examples in question *are* examples of that of which they are supposed to be examples?

And why should this question be considered a difficult problem, a kind of mystery? Well, that is because, on the one hand, when we ask about most other items of purported knowledge how it is we can know them, sense-experience seems to be the source, or anyhow the chief source of our knowledge, but, on the other hand, sense-experience seems able only to provide knowledge about what is or isn't, not what could have been or couldn't have been. How do we bridge the gap between 'is' and 'could'? The classic statement of the problem was given by Immanuel Kant, early (p.3) in the introduction to the second or B edition of his first critique, *The Critique of Pure Reason*: 'Experience teaches us that a thing is so, but not that it

cannot be otherwise.'

Note that this formulation allows that experience *can* teach us that a necessary truth is true; what it is not supposed to be able to teach is that it is necessary. For example, experience with what they call 'manipulatives' in pre-school, trying to arrange 29 blocks in various patterns, may convince one that 29 is a prime number, one that has no nontrivial factorization; but granted that it doesn't *actually* have one, how do we know that it *couldn't possibly* have had one, and is necessarily a prime number? How can experience teach us that *extra* fact?

The problem becomes more vivid if one adopts the language that was once used by G. W. Leibniz, and much later repopularized by Saul Kripke in his famous work on model theory for formal modal systems, the usage according to which the necessary is that which is 'true in all possible worlds'. In these terms the problem is that the senses only show us *this* world, the world we live in, the *actual* world as it is called, whereas when we claim to know about what could or couldn't have been, we are claiming knowledge of what is going on in some or all *other* worlds. For *that* kind of knowledge, it seems, we would need a kind of sixth sense, or extrasensory perception, or nonperceptual mode of apprehension, to see beyond the world in which we live to these various other worlds.

That is the problem. Let me now rapidly review some of the highlights of the history of attempted solution, beginning with Kant himself. Kant concludes, in the very next sentence after the one I quoted earlier, that our knowledge of necessity must be what he calls *a priori* knowledge or knowledge that is 'prior to' or before or *independent of* experience, rather than what he calls *a posteriori* knowledge or knowledge that is 'posterior to' or after or *dependant on* experience. And so the problem of the origin of our knowledge of necessity becomes for Kant the problem of the origin of our *a priori* knowledge.

Well, that is not *quite* the right way to describe Kant's position, since there is one special class of cases where Kant thinks it isn't really so hard to understand how we can have *a priori* knowledge. He doesn't think *all* of our *a priori* knowledge is mysterious, but only most of it. He distinguishes what he calls *analytic* from what he calls *synthetic* judgments, and holds that *a priori* knowledge of the former is unproblematic, since it is not really knowledge of external objects, but only knowledge of the content of our own concepts, a form of self-knowledge.

We can generate any number of examples of analytic truths by the following three-step process. First, take a simple logical truth of the form 'Anything that is both an A and a B is a B', for instance, 'Anyone who is

both a man and unmarried is unmarried'. Second, find a synonym C for the phrase 'thing that is both an A and a B', for instance, 'bachelor' for 'one who is both a man and unmarried'. Third, substitute the shorter synonym for the longer phrase in the original logical truth to get the truth 'Any C is a B', or in our example, the truth 'Any bachelor is unmarried'. Our knowledge of such a truth seems unproblematic because it seems to reduce to our knowledge of the meanings of our own words.

So the problem for Kant is not exactly how knowledge *a priori* is possible, but more precisely how *synthetic* knowledge *a priori* is possible. Kant thought we do have examples of such knowledge, and indeed the example I have been using, that 29 is a prime number, is one of them. Arithmetic, according to Kant, was supposed to be synthetic *a priori*, and geometry, too — all of pure mathematics. In his *Prolegomena to Any Future Metaphysics*, Kant listed 'How is pure mathematics possible?' as the first question for metaphysics, for the branch of philosophy concerned with space, time, substance, cause, and other grand general concepts — including modality.

Kant offered an elaborate explanation of how synthetic *a priori* knowledge is supposed to be possible, an explanation reducing it to a form of self-knowledge, but later philosophers questioned whether there really

were any examples of the synthetic *a priori*. Geometry, so far as it is about

the physical space in which we live and move — and that was the original

conception, and the one still prevailing in Kant's day — came to be seen as,

not synthetic *a priori*, but rather *a posteriori*. (This development was one of

the main causes of the extended usage, alluded to earlier, of 'geometry' to

cover a vast range of mathematical investigations.) The great mathematician

Carl Friedrich Gauß had already come to suspect that geometry is *a

posteriori*, like the rest of physics, by some time in the first half of the

nineteenth century. Since the time of Albert Einstein in the early twentieth

century the *a posteriori* character of physical geometry has been the received

view (whence the need for border-crossing from mathematics into physics if

one is to pursue the original aim of geometry).

As for arithmetic, the great logician Gottlob Frege in the late

nineteenth century claimed that it was not synthetic *a priori*, but analytic —

of the same status as 'Any bachelor is unmarried', except that to obtain

something like '29 is a prime number' one needs to substitute synonyms in a

logical truth of a form much more complicated than 'Anything that is both

an A and a B is a B'. This view was subsequently adopted by many

philosophers in the analytic tradition of which Frege was a forerunner,

whether or not they immersed themselves in the details of Frege's program

for the reduction of arithmetic to logic.

Once Kant's synthetic *a priori* has been rejected, the question of how we have knowledge of necessity reduces to the question of how we have knowledge of analyticity, which in turn resolves into a pair of questions: On the one hand, how do we have knowledge of synonymy, which is to say, how do we have knowledge of meaning? On the other hand how do we have knowledge of logical truths? As to the first question, presumably we acquire knowledge, explicit or implicit, conscious or unconscious, of meaning as we learn to speak, by the time we are able to ask the question whether *this* is a synonym of *that*, we have the answer. But what about knowledge of logic? That question didn't loom large in Kant's day, when only a very rudimentary logic existed, but after Frege vastly expanded the realm of logic — only by doing so could he find any prospect of reducing arithmetic to logic — the question loomed larger.

Many philosophers, however, convinced themselves that knowledge of logic also reduces to knowledge of meaning, namely, of the meanings of logical particles, words like 'not' and 'and' and 'or' and 'all' and 'some'. To be sure, there are infinitely many logical truths, in Frege's expanded logic. But they all follow from or are generated by a finite list of logical rules, and philosophers were tempted to identify knowledge of the meanings of logical

particles with knowledge of rules for using them: Knowing the meaning of 'or', for instance, would be knowing that 'A or B' follows from A and follows from B, and that anything that follows both from A and from B follows from 'A or B'. So in the end, knowledge of necessity reduces to conscious or unconscious knowledge of explicit or implicit semantical rules or linguistics conventions or whatever.

On this picture we know that five isn't a divisor of 29 in any possible world, not because we have some kind of ESP that permits us to engage in telepathic remote viewing of other possible worlds while remaining in this one, but rather because, just as we have implicitly or unconsciously learned conventions or rules that tell us no one who is married may properly be described as a 'bachelor', so also we have implicitly or unconsciously learned conventions or rules that tell us, or from which it follows, that one may not describe any world as a world in which 5 is a 'divisor' of 29.

For instance, a world in which whenever one lays out five rows of six blocks and then counts the total, one finds 29, is not properly described as a world in which five times six is 29. Rather it is a world where, by some curious physical principle, whenever 30 blocks are brought together in five rows of six, one of them disappears; or where, by some curious psychological principle, whenever we try to count 30 blocks arranged in five

rows of six, we always miss one; or something of the sort.

Such is the sort of picture that had become the received wisdom in philosophy departments in the English speaking world by the middle decades of the last century. For instance, A. J. Ayer, the notorious logical positivist, and P. F. Strawson, the notorious ordinary-language philosopher, disagreed with each other across a whole range of issues, and for many mid-century analytic philosophers such disagreements were considered the main issues in philosophy (though some observers would speak of the 'narcissism of small differences' here). But both sides put forward a picture quite like the one I have just drawn, Ayer in his *Language Truth and Logic* (Golancz) of 1936 and Strawson in his *Introduction to Logical Theory* (Methuen) of 1952.

And people like Ayer and Strawson in the 1920s through 1960s would sometimes go on to speak as if linguistic convention were the source not only of our knowledge of modality, but of modality itself, and go on further to speak of the source of language lying in *ourselves*. Individually, as children growing up in a linguistic community, or foreigners seeking to enter one, we must consciously or unconsciously learn the explicit or implicit rules of the communal language as something with a source outside us to which we must conform. But by contrast, collectively, *as* a speech

community, we do not so much *learn* as *create* the language with its rules. And so if the origin of modality, of necessity and its distinction from contingency, lies in language, it therefore lies in a creation of ours, and so in us. 'We, the makers and users of language' are the ground and source and origin of necessity. Well, this is not a literal quotation from any one philosophical writer of the last century, but a pastiche of paraphrases of several.

3.   MODALITY REGAINED

Such was the climate of opinion before January, 1970, when Kripke delivered his Princeton lectures on 'Naming and Necessity'. Since then it has become clear that there are several things wrong with the conclusion that we, the makers and users of language, are the originators of necessity, and the line of thought from Kant onwards that led to this sort of conclusion, and that incidentally led to necessity, as something distinct from aprioriness or analyticity, being submerged and lost sight of. I won't try to say in advance just how *many* main things are wrong with the line of thought I will be criticizing, lest I end up like Cardinal Ximenes in the Monty Python sketch about the Spanish Inquisition, and find myself repeatedly having to revise my estimate upwards. I'll just mention — following Kripke, though not very

closely or strictly — several difficulties one by one; we can count them up afterwards if we wish.

For one thing, there is a difficulty to be found right at the beginning of the whole line of argument, with Kant, and indeed at the very first step, Kant's inference from the premise that experience does not teach us that a thing cannot be otherwise to the conclusion that our knowledge of necessity is *a priori*. Strawson was less dismissive of Kant than was Ayer, but he famously described one argument of Kant's as a '*non sequitur* of numbing grossness'. I don't know whether the argument to which this phrase was first applied really deserved words so harsh, but Kant's argument from 'experience isn't sufficient to teach us that a thing couldn't have been otherwise' to 'experience isn't needed to learn that a thing couldn't have been otherwise' does seem to call for a pretty severe rebuke. It's like arguing from 'sunshine isn't sufficient to grow corn (rainfall is also needed)' to 'sunshine isn't needed to grow corn (rainfall is sufficient)'. It really is a *non sequitur* of numbing grossness.

Kripke's achievement was not just to point out this logical lapse on the part of a famous philosopher — finding logical lapses in the works of famous philosophers is something it is all too easy to do, I'm afraid, and one doesn't become famous just for doing *that* — but rather, his achievement

was to offer plausible counterexamples to Kant's claim, plausible cases of *a posteriori* knowledge of necessity. One famous example, perhaps the most celebrated, is that the Queen — or to call her by name rather than title, Elizabeth Windsor — could not have been the daughter of Harry Truman. Suppose a tabloid publishes a story according to which she *is* the daughter of Harry Truman, smuggled into the royal family as an infant. The story could be disproved by a chain of documents and testimony leading back to persons who moved in royal circles in the 1920s. Or it could be directly disproved by DNA testing. But all of these methods of disproof require sense-experience. The tabloid's sensational conspiracy theory can't be disproved by so-called 'armchair methods', just sitting and thinking. The story is false, but its falsehood is an item of *a posteriori*, not *a priori* knowledge.

Nonetheless, according to Kripke, the fact that Elizabeth Windsor is not the daughter of Harry Truman is a necessary fact. She not only *actually isn't* but *couldn't possibly have been* the daughter of anyone other than her actual father George Windsor, and in particular, not of Harry Truman. A 'possible world' or counterfactual situation in which a daughter of Harry Truman was smuggled into the royal family as an infant and later became Queen, would be a situation in which a person other than Elizabeth Windsor became Queen (and presumably along the way came to be *called* 'Elizabeth

17

Windsor'). It is not a situation in which Elizabeth Windsor (the person we actually call 'Elizabeth Windsor') had a different father.

Such examples of *a posteriori* necessities show that what philosophers of the period of Ayer and Strawson tended to claim about our knowledge of necessary truths is wrong: it does *not* just reduce to knowledge of linguistic rules. In fact, one must keep carefully distinct the age-old notion of *necessity* from the modern notions of *aprioriness* and *analyticity*.

4.   NECESSITY AND CONVENTION RECONSIDERED

The purely negative conclusion that the primæval notion necessity is *not* to be equated with eighteenth-through-twenty-first-century notions of aprioriness or analyticity is what Kripke left us with at the end of his 1970 Princeton lectures. By two years later, however, when the transcript of the audiotape of those lectures was being published, he had something a bit more positive to say, and he said it in some addenda to the lectures proper. There he suggests that in the background, behind each of his examples of *a posteriori* necessity, there is some *a priori* principle according to which whatever is so about certain matters is so of necessity.

On this picture, we know *a priori* that whatever the facts are about whether one number divides another, they couldn't have been otherwise.

Experience with manipulatives may teach us that five *doesn't* divide 29, and combining what experience thus teaches us with our background *a priori* knowledge, we can conclude that five *couldn't have* divided 29. Similarly, we know *a priori* that, whoever a given person's parents were, someone with different parents wouldn't count as *that* person. We know *a posteriori* that the Queen *isn't* a daughter of Harry Truman, and combining this *a posteriori* knowledge with our *a priori* knowledge, we can conclude that the Queen *couldn't have been* a daughter of Harry Truman.

Kripke describes this *a priori* knowledge in the background as 'discoverable by philosophical analysis', which is near enough for present purposes to calling it 'analytic'. Thus the view appears to be that our knowledge of necessity reduces to something like a combination of *a posteriori* knowledge gained through sense-experience with analytic knowledge of semantic rules or linguistic conventions or whatever. In that case, the formulation that our knowledge of necessity derives from our knowledge of linguistic rules would be defensible after all, provided it is understood as meaning only this, that though our knowledge of the *truth* of a necessary truth may depend on sense-experience, our knowledge of the *necessity* of that truth, given that it is true, derives from knowledge of linguistic rules. The main point is that the epistemology of modality is still

demystified: there is no need for ESP. And perhaps that is what the logical

positivists and ordinary language philosophers really cared about.

(Unfortunately for them, Kripke has found *another* problem with the line of

thought leading to the conclusion that our knowledge of necessity reduces to

knowledge of linguistic rules, a 'skeptical paradox' about the notion of 'rule'

itself; but this is a topic I must leave aside here for lack of space among

other reasons.)

The point we have reached is this: We have learned that we must be

very careful about how we understand such a formulation as 'Our

knowledge of necessity derives from knowledge of linguistic rules'; but it

has not really been shown that such formulations are wholly off the mark.

There remains, however, to be evaluated the step from the intermediate

conclusion 'knowledge of modal distinctions derives knowledge of linguistic

rules' — presumably rules about using modal auxiliary verbs, such as

'would' and 'might' — to the further conclusion that modal distinctions

themselves derive from linguistic rules, and the final conclusion that the

origin of necessity lies in *us*, the makers and users of language.

One immediate difficulty with these further steps is that the

conclusion just enunciated, along with all talk about the 'origin' of necessity,

is simply nonsense: that the problem of the 'origin' of necessity is not a

mystery but a muddle. For if one takes the word 'origin' at all literally, then to speak of the origin of necessity is to speak of a time when necessity came into being, and before which there was none. And *that* certainly seems to be nonsense. The number 29 not only is today but *always has been* necessarily prime. Even if we substitute 'ground' for 'origin', to get rid of the temporal connotations, in asking after the ground of something we still seem to be asking what is it without which that thing would not have been, whereas the necessary as opposed to the contingent is precisely that which would have been *no matter what* as opposed to that which might not have been *if only*. So we seem to be asking, 'What is, without which that which could not have failed to be would have failed to be?' And surely such a question is nonsensical.

As for *our* role in the matter, as makers and users of language, certainly any suggestion that if only *we* had acted differently, if only we had adopted different semantical rules or linguistic conventions, different things would have been necessary, involves a kind of confusion. Suppose, for instance, we had adopted the conventions of a base-thirteen numeration system rather than our actual decimal or base-ten system. Then the number that we wrote as a two followed by a nine, and perhaps even pronounced 'twenty-nine', would indeed have been a number divisible by five, since it

would have been two times thirteen plus nine or 35 that we would have been calling 'twenty-nine', and that number is five times seven. But our adopting a different numeration system would not have changed any facts about the number 29, except the fact that we *call* it 'twenty-nine' and not 'twenty-three' as we might if we used base thirteen.

So in speaking of ourselves as the originators of modality, the philosophers of the mid-twentieth-century will have been speaking nonsense if they were speaking literally. But surely they, or the best of them, had the wit to see this, and so we must take them to be speaking figuratively, even if they make no particularly strenuous efforts to warn against literalism or to spell out in terms that *can* be taken literally what they express figuratively by speaking of ourselves, the makers and users of language, as being the creators of necessity. Let us grant this, and let us allow ourselves to go on speaking figuratively for a bit. Even so, the formulation according to which we are the creators of necessity is open to at least one further challenge.

Even granting that what is needed to get from the kind of knowledge sense-experience can give us to knowledge of the necessity of necessary truths is knowledge of linguistic rules, and even granting also (despite 'skeptical paradoxes' about 'rules') that we do have the requisite kind of knowledge of linguistic rules, still it is premature to conclude, even speaking

figuratively, that necessity itself derives from linguistic rules. Another

interpretation of the situation remains possible, and logical positivists and

ordinary language philosophers have given no real argument against this

rival interpretation, but merely assumed that the rival view is too

'metaphysical', in a pejorative sense, to be taken seriously by analytic

philosophers. The real problem or mystery of the so-called 'origin of

necessity' lies in the clash between these two interpretations.

5.   REALISM AND PRAGMATISM ABOUT MODALITY

Let me now try to describe what the two conflicting views about the

nature of modality amount to, before turning to consider the bearing

of their conflict on other issues concerning modality, and ultimately on

modal *logic*. I will give first the view that people like Ayer and Strawson

want to *reject*, and then the view they want to *accept*.

Here is the first view. Though the fact is not mentioned in *Genesis*,

the first thing God said on the first day of creation was 'Let there be

necessity'. And there was necessity. And God saw necessity, that it was

good. And God divided necessity from contingency. And only then did He

say 'Let there be light'. Several days later, Adam and Eve were introducing

names for the animals into their language, and during a break between the

23

fish and the birds, introduced also into their language modal auxiliary verbs, or devices that would be translated into English using modal auxiliary verbs, and rules for their use, rules according to which it can be said of some things that they 'could' have been otherwise, and of other things that they 'could not'. In so doing they were merely putting labels on a distinction that was no more their creation than were the fishes of the sea or the beasts of the field or the birds of the air.

And here is the rival view. The failure of *Genesis* to mention any command 'Let there be necessity' is to be explained simply by the fact that no such command was issued. We have no reason to suppose that the language in which God speaks to the angels contains modal auxiliary verbs or any equivalent device. (I am speaking of 'metaphysical' modality here as always, not of deontic modality of the kind involved in 'Thou shalt not…') Sometime after the Tower of Babel some tribes found that their purposes would be better served by introducing into their language certain modal auxiliary verbs, and fixing certain rules for their use. When we say that *this* is necessary while *that* is contingent, we are applying such rules, rules that are products of human, not divine intelligence.

I have been allowing myself here to use theological language in describing the distinction I want to draw. That would have been the natural

way for seventeenth or eighteenth century philosophers, who nearly all were

or professed to be theists or deists, to discuss the matter. For many today,

such language cannot be literally accepted, and if it is only taken

metaphorically, then I have not really done better than those who speak

figuratively and frame the question as that of whether the 'origin' of

necessity lies outside us or within us. So let me drop the theological

language, and try again, without it, to describe the division between two

views that I have in mind.

Well, here the first view: Ultimately reality as it is in itself,

independently of our attempts to conceptualize and comprehend it, contains

both facts about what is, and superfacts about what not only is but had to

have been. Our modal usages, for instance, the distinction between the

simple indicative 'is' and the construction 'had to have been', simply *reflect*

this fundamental distinction in the world, a distinction that is and from the

beginning always was there, independently of us and our concerns.

And here is the second view: We have reasons, connected with our

various purposes in life, to use certain words, including 'would' and 'might',

in certain ways, and thereby to make certain distinctions. The distinction

between those things in the world that *would have been* no matter what and

those that *might have failed to be* if only is a *projection* of the distinctions

made in our language. Our saying there were necessities there before us is a

retroactive application to the pre-human world of a way of speaking

invented and created by human beings in order to solve human problems.

Well, that's my second try. With it I have gotten rid of theology, but

unfortunately I have *not* gotten rid of all metaphors. The key remaining

metaphor is the *optical* one: reflection *vs* projection. Perhaps I should give

up the attempt to get rid of all metaphors, and admit that the two views I am

discussing are not so much philosophical *theses* or *doctrines* as

'metaphilosophical' *attitudes* or *orientations*: a stance that finds the

'reflection' metaphor congenial, and the stance that finds the 'projection'

metaphor congenial. But let me try a third time to describe the distinction

between the two outlooks in literal terms, avoiding optics as well as

theology.

To begin with, both sides grant that there is a correspondence or

parallelism between two items. On the one hand, there are facts about the

contrast between what is necessary and what is contingent, for instance,

between 29 being a prime number and 29 being the number of years it takes

for Saturn to orbit the sun. On the other hand, there are facts about our usage

of modal auxiliary verbs such as 'would' and 'might', and these include, for

instance, the fact that we have no use for questions of the form 'Would 29

still have been a prime number if such-and-such?' but may have use for

questions of the form 'Would 29 still have been the number of years it takes

for Saturn to orbit the sun if such-and-such?' The difference between the

two sides concerns the order of explanation of the relation between the two

parallel ranges of facts.

And what do I mean by *that*? Well, both sides grant that '29 is

necessarily prime', for instance, is a proper thing to say, but they differ in

the explanation *why* it is a proper thing to say. Asked why, the first side will

say that ultimately it is simply because 29 *is* necessarily prime. That makes

the proposition that 29 is necessarily prime true, and since the sentence '29

is necessarily prime' expresses that proposition, it is true also, and a proper

thing to say. The second side will say instead that '29 is necessarily prime' is

a proper thing to say because there is a rule of our language according to

which it is a proper thing to say. This formulation of the difference between

the two sides gets rid of metaphor, though it does put an awful lot of weight

on the perhaps fragile 'why' and 'because'.

Note that the adherents of the second view need not *deny* that 29 is

necessarily prime. On the contrary, having said that the sentence '29 is

necessarily prime' is, per rules of our language, a proper thing to say, they

will go on to say it. Nor need the adherents of the first view deny that

recognition of the propriety of saying '29 is necessarily prime' is enshrined in a rule of our language. The adherents of the first view need not even deny that proximately, as individuals, we learn that '29 is necessarily prime' is a proper thing to say by picking up the pertinent rule in the course of learning our language. But the adherents of the first view will maintain that the rule itself is only proper because collectively, as the creators of the language, we or our remote answers have, in setting up the rule, managed to achieve correspondence with a pre-existing fact, or rather, a pre-existing superfact, the superfact that 29 is necessarily prime. The difference between the two views is, as I said, in the order of explanation.

I will want labels for the two sides, or 'metaphilosophical' stances, and rather than invent new ones, I will simply take two of the most overworked terms in the philosophical lexicon and give them one more job to do. I will call the reflection view 'realism' about modality, and the projection view 'pragmatism'. That at least will be easy to remember, since 'realism' and 'reflection' begin with the same first two letters, as do 'pragmatism' and 'projection'. The realist/pragmatist distinction has bearing across a range of issues and problems, and above all it has bearing on the meta-issue of which issues are significant. For the two sides will, or ought to, recognize quite different questions as the central unsolved problems in

the theory of modality.

For those on the realist side, the old problem of the *ultimate* source of our knowledge of modality remains, even if it is granted that the *proximate* source lies in knowledge of linguistic conventions. For knowledge of linguistic conventions constitutes knowledge of a reality independent of us only insofar as our linguistic conventions reflect, at least to some degree, such an ultimate reality. So for the realist the problem remains of explaining how such degree of correspondence as there is between distinctions in language and distinctions in the world comes about. If the distinction in the world is something primary and independent, and not a mere projection of the distinction in language, then how the distinction in language comes to be even imperfectly aligned with the distinction in the world remains to be explained. For it cannot be said that we have faculties responsive to modal facts independent of us — not in any sense of 'responsive' implying that if the facts had been different, then our language would have been different, since *modal* facts *couldn't have been* different. What then *is* the explanation? This is the problem of the epistemology of modality as it confronts the realist, and addressing it is or ought to be at the top of the realist agenda.

As for the pragmatist side, a chief argument of thinkers from Kant to Ayer and Strawson and beyond for their anti-realist stance has been

precisely that if the distinction we perceive in reality is taken to be merely a projection of a distinction created by ourselves, then the epistemological problem dissolves. That seems more like a reason for *hoping* the Kantian or Ayerite or Strawsonian view is the right one, than for *believing* that it is; but in any case, even supposing the pragmatist view is the right one, and the problems of the epistemology of modality are dissolved, still the pragmatist side has an important unanswered question of its own to address. The pragmatist account, as I formulated it earlier, begins by saying that we have certain reasons, connected with our various purposes in life, to use certain words, including 'would' and 'might', in certain ways, and thereby to make certain distinctions. What the pragmatist owes us is an account of what these purposes are, and how the rules of our language help us to achieve them. Addressing *that* issue is or ought to be at the top of the pragmatists' to-do list.

While the positivist Ayer dismisses all metaphysics, the ordinary-language philosopher Strawson distinguishes *good* metaphysics, which he calls 'descriptive', from *bad* metaphysics, which he calls 'revisionary', but which I will call 'transcendental' (without intending any specifically Kantian connotations). Descriptive metaphysics aims to provide an explicit account of our 'conceptual scheme', of the most general categories of commonsense

thought, as embodied in ordinary language. Transcendental metaphysics aims to get beyond or behind all merely human conceptual schemes and represntations to ultimate reality as it is in itself, an aim that Ayer and Strawson agree is infeasible and probably unintelligible. The descriptive/transcendental divide in metaphysics is a paradigmatically 'metaphilosophical' issue, one about *what philosophy is about*. Realists about modality are paradigmatic transcendental metaphysicians. Pragmatists must in the first instance be descriptive metaphysicians, since we must to begin with understand much better than we currently do how our modal distinctions work and what work they do for us, before proposing any revisions or reforms. And so the difference between realists and pragmatists goes beyond the question of what issue should come first on the philosopher's agenda, being as it is an issue about *what philosophical agendas are about*.

6.   THE VARIED SORROWS OF MODAL LOGIC

So much for generalities about the realist/pragmatist distinction. Before taking up the issue of the bearing of the division on modal *logic*, I must return to my earlier historical sketch, and consider what was going on, in the period of Ayer and Strawson and Carnap, and more generally in the

pre-1970 period, in the domain of modal logic.

The subject began, as I have already had occasion to mention, with logic's founder, Aristotle. It was actively pursued through antiquity and the middle ages and in a more desultory fashion on into the early modern period as far as the nineteenth century. But Frege, the founder of truly modern logic, swept away such fragments of modal logic as had still been left in Kant and still remained in his own day. As he explains early (§4) in his *Begriffsschrift*, the founding work of truly modern logic, his concept-writing is to contain no symbols for modalities, on the grounds that modal distinctions only indicate something about how an asserted statement is *known* and contribute nothing to the content of the statement asserted.

Apparently Frege is here falling in with the tendency to identify necessity and aprioriness, which we have seen to be a mistake. A better reason for his leaving modal distinctions aside would have been that the task of developing a logic capable, as the traditional syllogsistic was not, of analyzing mathematical arguments was enough to keep him fully occupied. Distinctions of mood, like those of tense, play no role in mathematics, which deals with facts that are necessary, as well as timeless. Developing a logic for such notions, which do figure in philosophical as opposed to mathematical argumentation, could well be left for the future, after

mathematical reasoning has been taken care of.

Modality continued to be ignored by Frege's heirs: Russell, for one, and certainly Hilbert and his school. Eventually the eminent twentieth-century logicians Kurt Gödel and Alfred Tarski were to take an interest in modal systems, but in both cases involvement with modal logic was a sideline in a career devoted mainly to other matters, and the systems considered were ones that had already been developed by other, lesser twentieth-century figures, above all C. I. Lewis.

As is well-known, Lewis's main aim was to develop in the object language of a formal system a theory of matters that orthodox logicians express only in the metalanguage. His primary focus was on what he called 'strict implication'. Close reading shows that by this he did not mean formal logical deducibility or consequence in the strictest sense, but something more like analytic deducibility or consequence. (Thus 'Jones is unmarried' would be for Lewis 'strictly implied' not only by 'Jones is a man and Jones is unmarried', but also by 'Jones is a bachelor'.) It is perhaps not going too far to suggest that Lewis's enterprise, for which the label 'endometalogic' might be appropriate, came to be called 'modal' logic only because the modal notion of necessity had by Lewis's day become conflated and confounded with the notion of analyticity.

Now the chief difference between treating implication or entailment in the metalanguage and representing it by an operator in the object language is that such operators can be iterated and nested and embedded, whereas, if 'implies' or 'entails' appear only in metalinguistic formulations and only in application to object-language formulas, then there can be no question or such iteration or nesting or embedding. Lewis thought he had 'intuitions' directly about 'strict implication', and in devising a formal system he would consult these to decide which axioms involving iteration and nesting and embedding were acceptable and which not.

Sometimes he changed his mind. Does p's strictly implying q strictly imply that q's strictly implying r strictly implies p's strictly implying r? Or can we only say that p's strictly implying q and q's strictly implying r together strictly imply p's strictly implying r? Lewis leaned first towards the first position, then towards the second, and so were born the systems known as **S3** and **S2**. These were soon joined by others, even before the mathematically-minded, who claimed no intuitions about strict implication, took over and produced dozens and scores of variants.

Lewis rather belatedly introduced the diamond symbol for possibility, and allowed that 'p strictly implies q' could be analyzed as 'it is not the case that it is possible that p and not q'. But he never used the box symbol for

necessity, and never took the step of analyzing 'possibly' as 'not necessarily not' or '*p* strictly implies *q*' as 'necessarily, if *p* then *q*' (with the 'material' sense of 'if': it is not the case that *p* and not *q*). Formal systems with necessity as primitive derive from Gödel rather than Lewis.

When bringing in 'possibly', Lewis insists that it must be understood as what he calls 'logical' possibility, as not involving a contradiction (really meaning by this *analytical* possibility or not involving a contradiction-in-terms). Had he considered 'necessarily', he inevitably would have insisted on a 'logical', which is to say *analytical*, interpretation of it as well. His position was fully in accord with the tendency in his day to identify the necessary and the analytic.

Carnap, a pioneer of quantified modal logic, and one of the few modal logicians after Lewis to say much about the meaning of modal formulas, explicitly presented his key notion as an explication alike of 'Leibniz' notion of necessity and Kant's notion of analyticity'. And this provoked objections from W. V. Quine, who saw that quantification into the context 'the statement "----------" is analytic' is nonsense. Protesting against defenders of modal logic, who in his view were not answering his objections but changing the subject, he insisted that he was concerned about the status of quantification into modal contexts 'in the quantificational sense of

quantification and the modal sense of modality'; and by the latter he meant reading the box symbol in formal systems as 'it is analytic that'. For that is what the box was supposed to mean according to the writers Quine was most concerned to criticize, above all Carnap. But for us the 'modal sense of modality' should rather mean our default 'metaphysical' sense.

If Quine failed to convince, it was perhaps because many of his opponents, though officially committed to explaining necessity as analyticity, were unconsciously influenced by thoughts about necessity in the original, default, 'metaphysical' sense. The general pre-1970 failure to recognize both notions explicitly and to distinguish between them naturally made it impossible to appreciate that different logics may be appropriate to the different notions. That is part of what I meant by saying that the original aim of modal logic, to deal with 'metaphysical' modality, became submerged and lost sight of.

As for the dispute between Quine and the defenders of quantified modal logic, by hindsight, belatedly distinguishing notions formerly conflated, Quine seems to be right about the meaninglessness of 'There is someone x such that the statement "x did not defeat Nixon and win the 1968 U. S. presidential election" is not analytic', while o*f course* it makes sense to say that there is someone such *it might have happened* that he defeated

Nixon and won the 1968 U. S. presidential election. Indeed this not only makes sense to say, but is true: For *Hubert Humphrey* is someone such that it might have happened that he defeated Nixon and won the 1968 U. S. presidential election.

But we had to wait for Kripke (from whose 1970 lectures the foregoing example derives) before it was made plain that reading the box as 'it is analytic that…' (the official Ludovician explanation, to the extent that there was one) and reading the box as 'it could not have failed to be the case that…' (a reading which doubtless influenced the 'intuitions' of some) are two different and incompatible readings. Before then there was much arguing at cross-purposes, and much confused groping guided now by one reading of the box and now by another, without recognizing the switch.

It is said that when Cauchy lecture at the Académie des Sciences on the distinction between convergent and divergent series, Laplace rushed home to check the series in his *Mécanique Céleste*. But when Kripke lectured in Princeton on the distinction between 'logical' and 'metaphysical' necessity — or in his preferred terminology, simply between analyticity and necessity *tout court et sans phrase* — modal logicians did not rush home to check which formal system was the right one for which notion of necessity. For by 1970 the idea that there might be a 'right' modal logic — which had

been a serious issue between Aristotle and his student Theophrastus, and a serious issue for Lewis as he compared **S2** and **S3** — had come to seem quaint, so far had the subject drifted from its origins.

7.   R EALISM, P RAGMATISM, AND THE M ODAL P ARADOX

To this day the literature on the question which is the right modal logic for 'metaphysical' modality remains thin. The candidate systems most often considered have been **S5** and **S4**, though I have seen (as referee) a soon-to-be-published (I hope) paper making an ingenious case for McKinsey's **S4.1**. And there is a well-known argument, developed by Nathan Salmon and others, for the system **T**, or rather *against* the axiom that distinguishes **S4** from **T**: the putative law that if p is necessary, it is necessarily necessary (and dually, if p is possibly possible, it is possible).

The objection goes by various names, but is perhaps most often called simply 'the modal paradox'. To illustrate how issues of modal logic, in the original or old-fashioned sense of the search for the right system of logic for modality properly so called, are inseparable from issues of philosophy and even 'metaphilosophy', of metaphysics and even 'metametaphysics', I would like to describe the example, and indicate how the realist/pragmatist division bears on it.

The paradox goes back to some of the discussion of *a posteriori* necessities in 'Naming and Necessity', and in particular to some examples pertaining to the necessity of origins. I have in mind here under that head now not the necessity, spoken of earlier, of the genealogy of human beings, whether of royal or noble or commoner rank. Rather, the issue I wish to consider is over the original composition of inanimate artifacts.

In a much-discussed passage in his third 'Naming and Necessity' lecture and attendant footnotes, Kripke suggests that the wooden table before him could not have originally been made of wholly different pieces of wood — really one should say wood and hardware, I suppose. In so saying he is apparently tacitly granting that the table could have been made of *slightly* different pieces. Perhaps a different screw could have been used at a certain joint, for instance.

The same principle would seem to apply to, say, a ship made of a thousand planks. There is an intuition X to the effect that the same ship could have been made of 999 of the same planks, arranged the same way, plus a replacement for plank #473; but at the same time there is an intuition Y to the effect that a ship made of a thousand different planks would have been a different ship. The puzzle is that these two intuitions, X and Y, seem to conflict. Let me describe the conflict in Leibniziano-Kripkean 'possible

worlds' language.

Here in the actual world A we have a ship, let us name it the good ship *Theseus*, made of 1000 planks. Our first intuition X is that the same ship could have been made of 999 of these planks plus a replacement for plank #473. In possible-worlds terms that means there is another world B where the same good ship *Theseus* exists with all but one plank the same as in our world A, and only plank #473 different. But then in world B one has a good ship *Theseus* made of 1000 planks, and by the same sort of intuition, there must another world C where the same good ship *Theseus* exists with all but one plank the same as in the world B, but with plank #692 different. That means for us back in world A there is another world C where the good ship *Theseus* exists with all but *two* planks the same as in our world A, but with planks #473 and #692 different, so one could have two planks different and still have the same ship. The same sort of considerations can then be used to argue that one could have three planks different, or four, or five, or all 1000. But that is contrary to our other intuition Y.

The modal paradox resembles well-known paradoxes of vagueness, such as the *heap* and the *bald one*, for which proposed solutions are a dime a dozen — except that here what seems to be vague is the relation of *identity*. And the idea that 'is the very same thing as' could be vague is for many a far

more troubling idea than the idea that 'heap' or "bald" is vague. Indeed, according to many, it is an outright incoherent idea. Salmon has suggested that the solution here, the way to avoid treating identity as vague, is to recognize that though world B is a possibility for us in world A, and world C is a possibility for those in world B, still C is not a possibility for us in world A, but only a 'possible possibility'.

There are reasons, however, to doubt the cogency of this suggestion. For the modal puzzle seems entirely parallel to the famous temporal puzzle discussed by Plutarch and Thomas Hobbes. According to this puzzle, in the year A, the hero Theseus returned to Athens after slaying the Minotaur, and his ship was left as a monument in the harbor at Piraeus. In the year B, one plank, beginning to rot, was replaced, and in the year C another plank, and so on, until in the end none of the original planks were left. The puzzle is that we have an intuition that changing one plank will leave us with the same ship but that if we change them all we have only a replica, and not the original ship.

No one thinks the solution to this classical puzzle is to maintain that though time B may be later than time A, and time C later than time B, still time C is not later than year A, that it is at time A a time that 'is going to be going to come' but not a time that 'is going to come'. Parity of reasoning

41

suggests that if we are suspicious of the idea of a future future that is not a

present future, we should be equally suspicious of the idea of a possible

possibility that is not an actual possibility.

The two puzzles or paradoxes can be developed in a way that makes

the parallelism between the modal and temporal cases more strikingly

apparent. We have a thousand planks laid out in a warehouse, and

instructions for assembling them into a ship, indicating which plank is to go

where, and the following standing instruction. Each day we are to look out

for a signal at noon, and if it comes, we are to assemble the ship. If the

signal does not come, we are to inspect the planks, remove whichever one

seems least sound and burn it, replacing it with the top plank in a reserve

pile stored in a secure place. Once the signal does come and a ship is

assembled, we are to inspect it each day, and replace whichever plank seems

least sound, in the same manner.

In both versions of the puzzle, the signal comes the first day, and the

ship is assembled. In the temporal version, we compare the ship constructed

on the first day with the ship in the harbor on the second day, the ship in the

harbor on the third day, and so on through a thousand daily changes of a

plank, until the last day. In the modal version we consider ship constructed

on the first day with the ship that would have been constructed if the signal

had not arrived until the second day, with the ship that would have been constructed if the signal had not arrived until the third day, and so on. I leave it to the reader to think through the details.

The parallelism between the temporal and modal cases, and the clear untenability in the temporal case of the suggestion that 'later than' is not transitive, to me make Salmon's solution to the modal paradox is uninviting. Now Salmon's is not the only purported solution on offer, but I don't think any of the others on offer is any better, nor have I a better one of my own to suggest. My reason for mentioning this problem about the necessity of origins is not to showcase some new purported solution, but rather to underscore the significance of the difference between the realist and pragmatist views, thus indicating how that division may ultimately be relevant to the choice of the right modal logic, and specifically to the decision whether it includes the law that *possibly possible* implies *possible*. The difference between the realist and pragmatist views is not over which solution to the modal paradox is best, but over how confident one can be that there *is* any solution, and ultimately over how significant the puzzle is.

From a the standpoint of a genuine realist, presumably there just *has to be* some solution, even if it is one requiring us to adopt a radically revisionist metaphysics. Since the problem is that the vagueness of *identity*

seems to be intolerable, in a way other kinds of vagueness generally are not, the most obvious solution would be to get rid of the object, the ship, whose identity seems to be vague. This might be done by adopting a revisionist metaphysics in which such extended, composite objects as ships don't really exist at all, but only atoms in the void, or in more modern terms, leptons and quarks, or strings or something, some of which sometimes swarm in ship-like formation.

For the pragmatist, however, it is all only a matter of whether a certain object in a certain contemplated counterfactual situation would count as the same object as a certain other object in a certain other contemplated counterfactual situation. And that, in turn, is only really a question of whether the two *ought to* count as one according to some rules ultimately created by *us*, as makers and users of language. I produced early on a famous quotation from Kant. Let me now produce an even more famous one, from his 'Idea for a Universal History': 'Out of the crooked timber of humanity no straight thing was ever made.' Given the imperfection, the 'crookedness' of every creation of ourselves, it may well be that our rules and conventions yield in certain cases conflicting answers, or no answers; and in such a case the puzzle can have no solution. On the pragmatist view it is always well to be aware our imperfections, and the puzzle of the origin of

necessity is of significance insofar as it helps highlight one of those imperfections — but not because the puzzle demands a solution. For the pragmatist, until we are shown that some practical decision of genuine importance would hinge crucially on how we answer the question 'Would it be the same ship or not?' we may be content to leave the question unanswered. Moreover, if the question ever did prove to be, in some specific case, of practical importance, then according to the pragmatist view we could then settle it by *fiat*, making a new rule for ourselves, since it is rules we have made to which the answers to all questions about modality are ultimately to be traced.

In sum, for those who take a realist attitude towards the question of the origin of modality, there will be a real problem of principle about the necessity of origin. By contrast, to those who take a pragmatist attitude there will only be the potential for a practical problem, a potential that is thus far apparently unrealized, and that if it ever were realized could always be solved by fiat. On the realist view it seems that the question of which modal logic is the right one must remain in suspense until the right solution to the modal paradox has been found. On the pragmatist view, there simply may *be* no answer to the question of which modal logic is right, since there may be none to the question of which solution to the paradox is right. Clearly, then,

pursuit the right modal logic for 'metaphysical' modality, which was arguably modal logic's original aim, will take the logician deep into philosophy and beyond it into 'metametaphysics'.

NOTE

Portions of this paper derive from the second annual Kripke Lecture, which I presented at the CUNY graduate center in fall, 2012, and reprised a week later at a conference on modality at Princeton hosted by Boris Kment, and at the University of Connecticut in spring, 2013, using the title 'The Origin of Necessity and the Necessity of Origin' for the first two presentations, and 'Realism and Pragmatism about Modality' for the third, somewhat revised version. I am grateful to comments from the audience on all three occasions, which have led to revisions that I hope are improvements. The lecture material in turn was partly a summary, but partly an elaboration going beyond, my recent book *Saul Kripke: Puzzles and Mysteries* (Polity Press, 2013). In particular, I refer the reader to that book (especially the first sections of the Introduction and of chapter 2) for fuller treatment of the material in §§ 3, 4, and 6 above (and for discussion in chapter 5 of the 'skeptical paradox' about 'rules', which I have barely alluded to here), and especially for bibliographical references connected with this material. Portions of §6 trace back to an old unpublished paper of mine, 'The Varied Sorrows of Modality' (whose title is a phrase of Quine's), and I remain grateful to my colleagues Gilbert Harman and the late David Lewis for comments on that draft. Finally, I am indebted to Teresa Robertson for discussions of 'the modal paradox', and refer the reader to her *Stanford Encyclopedia of Philosophy* entry 'Essential vs. Accidental Properties' and its linked supplement 'Arguments for Origin Essentialism' for references to and discussion of the work of Salmon mentioned in §7 above, and the large body of related work I have not had space to discuss.