

Understanding, Modality, Logical Operators

Christopher Peacocke

Columbia University

Timothy Williamson's *The Philosophy of Philosophy* stimulates on every page. I would like to discuss every chapter. To conform to house rules, I confine myself to two.

Knowledge of Metaphysical Modality

Tim proves these equivalences:

$$(17) \quad \Box A \Leftrightarrow (\sim A \Box \rightarrow \mathbf{F})$$

and

$$(18) \quad \Diamond A \Leftrightarrow \sim(A \Box \rightarrow \mathbf{F}),$$

where \mathbf{F} is a contradiction (I use his numbering).

Tim says about these equivalences: (1) “modulo the implicit recognition of this equivalence, the epistemology of metaphysically modal thinking is a special case of the epistemology of counterfactual thinking. Whoever has what it takes to understand the counterfactual conditional and the elementary logical auxiliaries \sim and \mathbf{F} has what it takes to understand possibility and necessity operators.” (158) (2) The idea that we evaluate metaphysically modal claims “by some quite different means [from those we use to evaluate counterfactuals – CP] is highly fanciful, since it indicates a bizarre lack of cognitive economy and has no plausible explanation of where the alternative cognitive resources might come from” (162) (3) “the capacity to handle metaphysical modality is an “accidental” byproduct of the cognitive mechanisms that provide our capacity to handle counterfactual conditionals” (162).

The biconditional corresponding to a proposed definition of a concept C may be necessary and a priori. But that leaves open these questions:

- (1) Does the definition contribute to an explanation of why what are in fact truths containing C are true? (The Explanation Question)
- (2) Does the definition contribute to an explanation of our knowledge of certain contents containing C? (Epistemic Question)
- (3) Does the definition contribute to an explanation of our understanding or grasp of contents containing C? (Understanding Question)

I argue that applied to (17) and (18), considered as definitions, the answers to these three questions are negative. It should not be surprising that there is sometimes a gulf between a true, necessary, a priori biconditional and the issues of explanation, knowledge and understanding. There are multiple definitions of natural numbers and their properties in set theory, second-order logic, and much else. It is widely agreed that these definitions do not answer the corresponding three questions about natural numbers. The various definitions of natural numbers do not answer these questions in the numerical case because each of the various definitions identifies numbers in some particular background domain that has structure that goes beyond what is required for the explanation, epistemology and understanding of primitive arithmetic.

Consider first the Explanation Question for definitions (17) and (18). Suppose we have some contradiction A, and we ask why A is metaphysically impossible. Suppose too that the F in the definitions (17) and (18) is understood as some particular contradiction $p \& \sim p$. Under definitions (17) and (18) so understood, for A to be metaphysically impossible is for it to be that

$$\sim(A \Box \rightarrow (p \& \sim p)),$$

that is,

$$A \Box \rightarrow (p \& \sim p).$$

In the special case in which A is $p \& \sim p$ itself, this becomes

$$(p \& \sim p) \Box \rightarrow (p \& \sim p).$$

The truth of this, as of any instance of the form $B \Box \rightarrow B$, is not in doubt. But does it give us any insight into why contradictions are metaphysically impossible? It tells us that if a contradiction were to hold, a contradiction would hold. Very well; but could a contradiction hold? No doubt it could not; but, as Frege would say, that is no thanks to the definitions (17) and (18).

Maybe we should not take **F** as some particular contradiction. This alternative course accords better with Tim's description of **F** as a logical auxiliary. If we are operating with a conception under which sentences determine functions from possible worlds to truth-values, **F** can simply be that function that maps every possible world to the False. (Merely mapping the actual world to the False would not be enough for modal purposes.) Under the definition (18), for $q \& \sim q$ to be metaphysically impossible is, once we eliminate the double negation, for $(q \& \sim q) \Box \rightarrow \mathbf{F}$ to hold: that is, for either (a) some possible world in which $q \& \sim q$ holds and which is mapped by **F** to the True to be closer than any.....and we need not continue this condition, since no world is mapped by **F** to the True; or (b) for there to be no world in which $q \& \sim q$ holds. This alternation of (a) with (b) is extensionally correct in classifying cases of metaphysical impossibility. But when a complex proposition is classified under this treatment as metaphysically impossible, its status as such is not exhibited as consequential in part upon its the content or logical form. An explanatory definition should contribute to an explanation of why the metaphysical impossibility of $q \& \sim q$ results from the semantic contributions of negation and conjunction, and their mode of combination in $q \& \sim q$.

I am inclined to draw the conclusion that the equivalences (17) and (18) can be rationally accepted as extensionally correct only by someone who has some other means, independent of these equivalences themselves, of coming to know that certain propositions are metaphysically impossible.

A simple explanation of the metaphysical impossibility of $p \& \sim p$ draws on the fact

that that necessity is truth in all possible worlds (taken as sets of propositions), where genuine possibility is constrained, inter alia, by the requirement that in any genuinely possible world, the same rules for evaluating logically complex expressions must hold as actually hold for outright truth. For outright truth, $\sim p$ is true iff p is not true; and, also for outright truth, $p \& q$ is true iff p is true and q is true. For any world w for which these rules also hold for truth-in- w , nothing of the form $p \& \sim p$ is true in w . This is an explanation of why something of the form $p \& \sim p$ does not hold in any possible world, an explanation that is explanatorily sensitive to the logical form of the proposition.

On the Epistemic Question, number (2) above, Tim argues we have a (fallible) capacity for perceptual imagination, which, properly operated, can give us knowledge of counterfactuals such as “If the bush had not been in the way, the rock would have rolled all the way down the slope into the lake”. I agree. Does this capacity for perceptual imagination mean that, in combination with definitions (17) and (18), we thereby have an epistemology of metaphysical modality? Our knowledge of metaphysical modalities goes beyond what can be explained by a faculty of merely perceptual imagination. Some counterfactuals concern the imperceptible (say, the unobservable consequences for the behaviour of electrons of a change in the earth's magnetic field). Even when the subject-matter is perceptible states of affairs, perceptual imagination is not going to account for such matters as our knowledge of the Banach-Tarski Theorem, that a sphere can be decomposed into finitely many sets of points that can be reassembled to form two spheres equal in size to the original. It seems to me that we hold on to certain set-theoretic, mathematical and logical principles in any genuinely possible world in a way whose legitimacy is not fully explained by perceptual imagination.

Tim is of course well aware, and explicitly notes, that not all imagination is perceptual imagination. But when imagination is not perceptual, it can yield knowledge of possibilities only if it is constrained by the requirement that logical, mathematical and various other principles be respected in the state of affairs that is imagined. Again, one plausible explanation in the logical case of why they must be respected in any genuine possibility is that some specification of a state of affairs is genuinely possible only if it respects the same rules for assigning semantic values to concepts and expressions as hold in the actual world. For the mathematical and set-theoretic cases, a plausible explanation

in a similar spirit is available if genuinely possible states of affairs respect what individuates the mathematical and set-theoretic objects, and if the mathematical and set-theoretic principles are consequences of a correct articulation of this individuation.

On the Understanding Question, number (3) above, these points suggest an alternative model to Tim's account involving the definitions (17) and (18). Understanding metaphysical necessity involves tacit knowledge that it is truth in all possible worlds, combined with tacit knowledge of the constraints that must be met for a description of a putative world to be genuinely possible. Tacit knowledge that in any specification of a genuine possibility, the same rules must be used in determining semantic value as are used in determining actual semantic value is one such piece of tacit knowledge involved in understanding. This approach has its own challenges to address, including the nature of the relation between the constitutive and the modal, and the epistemology of the constitutive. But it has certain attractions, a salient one being its ability to account for the fallibility of imaginative simulation even in perceptual cases. Some people imagine that if a ball were to be thrown by a rider from a moving carousel, at a tangent to the circle that is the rider's locus, the ball will follow a curved path when released. This need not be a defect in these people's understanding of counterfactuals or of metaphysical modalities. It is an error of their intuitive mechanics, the error consisting in a mistake about what the laws of motion are that will be preserved in nearby genuinely possible worlds. Whenever a method is widely reliable but fallible in respect of our reaching knowledge of a particular kind of proposition, we must have some deeper understanding of what it is for such propositions to be true that is distinct from what is reachable by that fallible method. An account of understanding of counterfactuals and modals that draws on tacit knowledge of constraints that must be met for genuine possibility can contribute to an account of this more fundamental level of understanding.

Logical Principles, Understanding and Assent

Tim argues that there are no necessary connections between understanding certain expressions and assent to certain sentences or principles containing those expressions. His view is that any sentence or principle may be intelligibly rejected by a thinker,

consistently with understanding, provided there is enough of the right sort of overlap between the thinker's practice with the expressions elsewhere and that of the rest of the community using the expressions in question. More generally, the meaning of an expression is socially determined: "when an individual does use a shared language as such, individual meaning is parasitic on social meaning" (125). There is such a thing as identity of meaning of expressions in causally unrelated linguistic communities, and this is best explained in terms of truth-conditional semantics. The character of an expression in David Kaplan's sense, that is the function from context to referential content associated with the expression, is a "better approximation" to linguistic meaning (128).¹

Not all of Tim's examples convinced me. In arguing for the intelligible rejection of conjunction-elimination, Tim gives the example of Simon, who "regards both truth and indefiniteness as designated (acceptable) semantic values for an assertion: what matters is to avoid falsity" (96). If A is simply false and B is indefinite, and A&B is counted as indefinite, then the transition A&B/A can lead from indefiniteness to falsity. "On these grounds Simon rejects the conclusion of that instance while accepting its premise", even though Simon may agree that asserting the premise may be misleading (96). It seems to me that although Simon may not have misunderstood any individual expression or mode of composition, he certainly does not properly appreciate the practice of assertion in which he is engaging. By Tim's own view – which I accept - of asserting something as representing oneself as knowing it, no one who thinks A is indefinite should be sincerely asserting A. If the believer in understanding/assent links restricts himself to those who have an adequate grasp of assertion (and judgement), the case of Simon need not trouble him.

But would the existence of some necessary understanding/assent links really suffice to establish the epistemological conception of analyticity that Tim attacks? Early in his chapter, Tim distinguishes, but does not return to, the distinction between constitutive and necessary formulations of the target doctrine. Even if someone who understands conjunction and appreciates assertion must be willing to accept conjunction-elimination in some form, it by no means follows that this is what *makes* it the case that

¹ D. Kaplan, 'Demonstratives', in *Themes from Kaplan* (New York: Oxford University Press, 1989), ed. J. Almog, J. Perry and H. Wettstein.

he understands conjunction. (Compare: even if it is necessary of four that it is the minimum number of colours required to colour an arbitrary map without adjacent colours on a boundary, it does not follow that this is what *makes* something the number four.) The intuitive, and it seems to me correct, view is that a thinker's understanding of conjunction explains his acceptance of instances of conjunction-elimination. Nothing can consist in what it explains. What explains must be distinct from what it explains.

What could understanding be, if it explains such acceptance? I would argue that understanding an expression, or grasping a concept, consists in tacit knowledge of the fundamental reference rule for that expression (or concept). The fundamental reference rule specifies what makes something the reference of the expression (concept) on a given occasion of its use. Tacit knowledge of the fundamental reference rule is drawn upon in subpersonal computations that explain why it strikes a thinker as reasonable to make certain judgements involving the expression or concept, usually in the presence of various additional pieces of information. A thinker's willingness to make a judgement in given circumstances is very rarely if ever explained only by his tacit knowledge of the fundamental references rules for the expressions or concepts involved. Much empirical information, and some subpersonal computational capacities, will be involved in the explanation of the judgement. Concepts, even families of concepts, have no conceptual role independently of such additional information and states. This is a further challenge to the conceptual-role theories of the sort Tim attacks.

If we use the resource of an understanding-condition, we can distinguish between radical and moderate versions of the doctrine that, when a shared language exists, "individual meaning is parasitic on social meaning". The radical version states that there is no understanding-condition, at the level of sense, that does not explicitly mention society and other language-users. The qualification "at the level of sense" matters. No one, however radical, need be in disagreement with the thesis that, for an expression which has only a reference and not a sense in the public language, there is some entity of the appropriate category such that: to understand the expression is to know of that entity that the expression refers to it. The radical doctrine is saying rather that there is nothing that plausibly determines a sense (not just a reference) and which is both an understanding-condition for an expression, and does not mention society and other

language-users. By contrast, the moderate version of the parasitism doctrine states that in some cases – the cases where there is sense as well as reference – there does exist an understanding-condition at the level of sense that does not mention society and other language-users. This moderate version still recognizes parasitism because it will agree that the individual aims to use an expression in accord with its public meaning. It will also agree that we attribute propositional attitudes to a thinker using the public-language expression in a ‘that...’ clause even when a thinker partially misunderstands, or fails to understand the expression.² On the moderate version, however, what correct understanding consists in, at the level of sense, has a specification not involving society. (Of course, what makes it the correct specification has something to do with society.) Many very different substantive theories of sense might occupy this moderate position, be they assertibility theories, pragmatist theories, conceptual-role theories, or, as I favour, theories that identify understanding at the level of sense with tacit knowledge of a fundamental reference rule. On this last approach, different rules correspond to different senses, even if two concepts have, on a given occasion of use, the same reference.

The moderate and the radical views have different consequences for several philosophical issues. I give two examples. The first concerns the philosophical explanation of the rationality of judgement and assertion, and the source of a thinker’s rationality. Rational application founded in the nature of the understanding-condition, at the level of sense, seems to have nothing to do with fitting in with other language-users. Indeed a thinker may be able to lead other language users if the thinker has a deeper sense-based understanding. An example would be appreciation of the correctness of a new axiom, one that does not follow from those already accepted, for a domain described in terms that are already understood, and understood in the same way before the new axiom was introduced. Such phenomena have nothing to do with deference and the parasitism of individual meaning on social meaning. They have rather to do with the understanding of a domain, thought about in a certain way. New axioms provide a sharp and striking type of case. But on the moderate view they are only vivid instances of a

² T. Burge, ‘Individualism and the Mental’, in his *Foundations of Mind* (Oxford: Oxford University Press, 2007).

more general phenomenon found also in the humblest cases, that of appreciating the rationality of accepting something on the basis of a certain kind of understanding.

The second example concerns the existence of sameness of meaning of a word across different languages and even causally isolated traditions. There is no apparent difficulty in the idea that an ancient Greek word and an Aztec word both had the same meaning of sentence-negation, or expressed the same shape-concept. On the moderate view, this can consist in the Greek and Aztec words having the same understanding-condition at the level of sense. The radical view cannot take that line, but must nevertheless acknowledge the possibility of some kind of sameness of meaning between words of causally unconnected languages. As we noted, Tim, in whom there are hints of the more radical view, suggests for this role sameness of structured character. But there are two sources of instability in combining the radical view with structured character as the elucidation of cross-language meaning.

The first source of instability is that we have a legitimate conception of meaning that in some cases cuts more finely than that. It is possible that there is a language with a semantically and syntactically unstructured word *W* for the shape perceived as a regular-diamond (same length sides, right angles). This shape is, at the level of reference, the same shape as a square. But it can be informative that squares are *Ws*. The story can be developed in such a way as to make it compelling that the understanding-condition for *W* mentions shapes perceived as diamonds. Yet *W* and 'square' have the same character. If this case is allowed, we must slice more finely than Kaplanian character to capture word-meaning. Tim says, very reasonably, that we should not circumscribe the resources of semantics in advance. But it is hard to see how the right way of accommodating the difference in this sort of example won't equally determine society-independent understanding-conditions of the sort the moderate, but not the radical, version of parasitism will recognize.

The second source of instability arises from the fact that a good formulation of a character-rule, especially at the level of concepts, will plausibly coincide with a fundamental reference-rule. But, provided that we are not concerned with the special subcases in which the fundamental reference or character rule actually mentions other language-users, this implies that character-rules can fix a society-independent

understanding-condition at the level of sense. The fundamental reference or character rule for a concept contributes to the explanation of what are good reasons for making judgements containing that concept.³ Effectively, the rule individuates a sense (or sense-type). Using character as the elucidation of society-independent meaning thus imports a resource that permits the elucidation of meaning and understanding in precisely the way the radical view rules out.

The understanding-conditions recognized by the moderate view will not usually have the direct connections with assent that Tim rejects. Understanding-conditions, like virtually everything else, contribute to explanations of particular cognitive performances only in combination with a raft of other conditions. But they do articulate conditions for understanding particular expressions much more specific than an undifferentiated holism. There is middle ground between the epistemological conception of analyticity and a constitutive holism about the ascription of specific meanings and contents. It is in this middle ground that much of reason operates.

³ For arguments for this view, see chapters 2 and 3 of my *Truly Understood* (Oxford: Oxford University Press, 2008).