

OBLIGATION AND MODAL LOGIC

H. N. CASTANEDA

My purpose is to argue that the logic of obligation (also called deontic ⁽¹⁾ and normative ⁽²⁾ logic) cannot be fully formalized as a mere part of, or reduced to a system of ordinary logic, i.e., a system of symbolic logic which is interpreted as a formalization of logical necessity, like Lewis' systems S4 and S5. Such a 'reduction' of obligation to logical necessity commits what G. E. Moore called the 'naturalistic fallacy'.

Since the most complete and forceful argument for such a reduction has been given by Professor Alan Ross Anderson ⁽³⁾, the following discussion concentrates almost exclusively on his work.

1. *Anderson's reduction.* Anderson starts with a system X of modal logic, i.e., a system whose formulas are interpreted as describing states of affairs or as saying that states of affairs are necessary or possible; he then adjoins to it a propositional constant S, to be interpreted as the description of a sanction or penalty (FANS, 33f; RDL./103; LN, 86). To X he adjoins just one axiom and four definitions. Axiom A15: S is possibly false (symbolized as $M\sim S$).

Def. P. 'A is permitted' is short for 'it is possible that A is true and the sanction is false' (in symbols: 'PA' is short for ' $M(A.\sim S)$ ', where 'M' means it is possible that').

Def. O. 'A is obligatory' is short for 'not-A is not permitted' (in symbols: 'OA' is short for ' $\sim P\sim A$ ').

Def. F. 'A is forbidden' is short for 'A is not permitted' (in symbols: 'FA' is short for ' $\sim PA$ ').

⁽¹⁾ G. H. VON WRIGHT, *Deontic Logic*, in *Mind*, n.s. lx (1951), 1-15; A. N. PRIOR, *Formal Logic*, Oxford, Clarendon Press, 1955, 220-229.

⁽²⁾ H. N. CASTANEDA, *La lógica de las normas y la ética*, in *Universidad de San Carlos* (Guatemala), XXX (1954), 129-196; A. R. ANDERSON, *The Formal Analysis of Normative Systems*, New Haven, Yale Sociology Department, 1956, to be cited as FANS.

⁽³⁾ A. R. ANDERSON, FANS, sections II, III and X; *A Reduction of Deontic Logic to Alethic Modal Logic*, in *Mind*, n.s. lxvii (1958), 100-103, to be cited as RDL; and *The Logic of Norms*, in *Logique et Analyse* I (1958), 94-91, to be cited as LN.

The result is a system OX, which Anderson believes to be an adequate formalization of the logic of obligation. Searching for a more elegant and economical reduction of deontic logic to modal alethic logic, Anderson achieves the elimination of A15 by proving the proposition below, which I will call *Anderson's contention*: 'Addition of a propositional constant B and definitions of 'S', 'P', 'O' (and 'F'), to any normal alethic modal logic X yields a normal deontic logic OX.' (FANS, 45; only the definition of 'Pp' as an abbreviation for $M[p.\sim(M\sim B.B)]$ is mentioned in RDL, 101).

Def. S. 'S' is short for 'possibly not-B, but B' or 'B is possibly false, but it is true' (i.e., $(M\sim B.B)$) (FANS, 43).

Anderson's conditions for the normality of both modal alethic and deontic systems are certainly necessary conditions. But I shall argue that his conditions for the latter are not sufficient (in the normal sense of the word 'normal').

2. *Generality of OX.* The objection that Anderson's system OX has the same sanction for all obligatory acts can be easily met. As Anderson points out (FANS, 85), we can take S to be one of a class of sanctions, or the disjunction of all sanctions. Some of the definitions and theorems will have to be modified, but only slightly.

Anderson argues that OX is, as it must be, neutral with respect to the controversy between deontological and teleological theories of norms. "Nothing whatever in the formalism dictates the particular choice of an interpretation for 'S'. In particular, the 'sanction' might be the disapproval of God, or pain (on a hedonistic theory of ethics), or simply the statement that something wrong has been done... Presumably on *any* account of obligation, if an obligation is not executed, there is some way in which things are worse off than if the obligation had been executed — and this 'way in which things are worse off', no matter how vague, can function as an interpretation of 'S'... (The aim of this remark is not to recommend that 'S' be given a vague interpretation, but rather to dispel the notion that we are committed by the formalism to some *particular* choice of an interpretation for 'S'.)" (FANS, 88; Cf NL, 85f).

3. *Penalties.* Anderson's contention looks somewhat too generous. It does not seem proper to say that the above definitions of 'S', 'P', 'O', and 'F' formulate the meanings of the ordinary deontic terms 'permissible', 'obligatory', 'forbidden', unless some restriction is put

on the interpretation of the propositional constant 'B'. If 'B' is interpreted as 'There is a blue book', which is certainly a contingent proposition, we are forced by Def. F to say it is forbidden that there is a blue book which may have failed to be blue. This, of course, can certainly be a fanciful way of saying that some person ought not to have produced such a blue book; but neither of the two ways of talking could be a mere shorthand for the tautology 'If there is a blue book which may have failed to be blue, then there is a blue book which may have failed to be blue.'

Anderson's remark to the effect that 'S' is to be interpreted as a sanction or penalty or wrong is a necessary (semantical) condition for his formal system to be interpretable as a moral deontic logic. But he cannot leave this requirement in such a vague state. It might be urged that a correct analysis of 'punishment' or 'penalty' has to be made in terms of 'obligation' or 'duty' or 'ought'. It may be adduced that obligations are required by penalties, but not vice versa. An obligation or duty without punishments attached to it may be usually unfulfilled, but nevertheless it is an obligation or duty. Moral duties are not sanctioned by penalties. On the other hand, we can institute penalties only if we have previously decided on a system of duties. A dictator who imprisons and kills at the change of his whim creates no obligations or duties just because of the harm he causes; for that reason, his acts are not properly penalties *in the absence of previously prescribed duties*. Indeed, it may be argued that the people have a duty to stage a revolution even in the face of threat to their lives.

It is clear in any case that while the conditions are not rigorously formulated for the proper *deontic* interpretation of 'B' or 'S', we cannot be sure how much philosophical clarification is produced by system OX. Let us assume, however, that S is interpreted as either

- (i) 'Penalty... will be inflicted' or
- (ii) 'Penalty... ought to be inflicted' or
- (iii) 'Something bad or wrong will happen'.

Any other consequence, as Anderson well knows, cannot do.

It is easy to see that (i) is wholly inadequate. Failure to do one's duty never entails or strictly implies that punishment will be inflicted. It is commonplace that many criminals die unprosecuted, not to mention those acquitted by a miscarriage of justice. In fact, punishments (in the ordinary meanings of this word) are not even the effects of performing or omitting certain acts. No doubt, we use

the word 'ought' to reason about actions to be performed by us or by somebody else in such a way that if the acts have certain consequences we conclude that we ought (or ought not) to do them. But (1) such consequences are related to the acts causally, not by entailment; (2) they are not punishments or penalties, but (3) reasons which may lead (or may actually have led) us both to make the actions duties and to create punishments for failure to fulfill such duties.

Interpretation (ii) for 'S' is mentioned by A. N. Prior⁽⁴⁾ in a discussion of Anderson's system. It must be acknowledged that 'Act A is obligatory' does seem to entail 'If A is not performed punishment ought to ensue' or 'He who fails to do A ought to be punished'. However, it may be argued that this entailment holds if and only if it is presupposed that duties, obligations, and *ought's* are related to a system of punishments.

At any rate, Prior's proposed interpretation (ii) suffers from the difficulty that in order to accomplish the reduction of deontic to modal alethic logic we must introduce an un-reduced deontic interpretation. Obviously, any attempt to reduce the 'ought' contained in 'S' by means of another sanction S interpreted in accordance with (ii) will bring in another hidden 'ought'.

An immediate theorem in OX is OM48.
S is forbidden (FANS, 62).

This is, according to Def. F and Def. P, the tautology 'S entails S' or 'It is necessary that: if S, then S'. However, on interpretation (i), it means that the sanction or the punishment is forbidden. Suppose that the punishment is 15 years in jail. Thus, OM48 says 'it is forbidden to go to jail for 15 years'. Now (1) this is not a tautology or a truth of logic, but a very definite rule. Hence, it cannot be an instance of OM48, which is an abbreviation of, and is itself, a truth of logic. (2) It certainly seems implausible to say that the established punishments are forbidden. This implies that a wrong is done whenever a lawbreaker is properly punished or that by being punished he is made to do something forbidden, namely, go through the punishment. Thus a necessarily effective procedure to diminish unlawfulness or lawbreaking would simply be not to punish the criminals. (3) If the state-of-affairs of being

⁽⁴⁾ A. N. PRIOR, *Time and Modality*, Oxford, Clarendon Press, 1957, Appendix Dip. 140.

punished is forbidden, a person ought to try to avoid being punished. But then how could it be wrong for a criminal to try to escape? And how could it be right to punish him for escaping? To do so is to bring about *another* state-of-affairs which is forbidden.

If we follow interpretation (ii), OM48 does not fare much better. It now means that it is forbidden that punishment ought to be inflicted. This is nonsensical, unless it just means that punishment is forbidden; but then we are back to (i).

Indeed, it seems that a necessary condition for a formal system to be (interpreted as) a normal deontic logic is that in it punishments must not be forbidden, *provided*, of course, that in such a system punishments are attached to obligations.

4. *Obligation and entailment.* Let us turn to interpretation (iii), *viz.*, that the sanction entailed by failure to do one's duty is that something bad or wrong happens, or that the world will be worse off. This interpretation would seem to have an advantage, for (a) below does entail (b) in practically any sense of 'wrong' and many senses of 'bad':

(a) 'action A is obligatory or a duty' [in symbols 'Oa'],

(b) 'if A is not performed, something wrong (or bad) has been done'.

But, just as in the case of (i), Anderson has moved from the fact that (a) entails (b) to the view that (b) is itself an entailment-statement (Cf. FANS, 89). This is exactly what the definitions of 'O' and 'P' establish. It should be noted that, as in the case of interpretation (ii), the words 'wrong', and 'bad' are of deontic not modal alethic logic, so that once again the former is apparently reduced to the latter only because an un-reduced deontic interpretation is smuggled in.

In general, if Anderson's system OX were an adequate formalization of deontic logic, to assert (a) would be to assert the statement about entailment (c) below:

(c) 'failure to do A entails, or strictly implies, that the sanction will be actualized' [in symbols, 'Nec ($\sim p \supset S$)'].

That is, on Anderson's formalization, to say A is a duty is to say that the statement (d) below is logically necessarily true;

(d) 'If A is not performed the sanction will be actualized' [in symbols; ' $p \supset S$ '].

It seems that a true statement asserting that other statement is

logically true must itself be logically true⁽⁵⁾. Thus, if (c) is true, i.e., if (d) is logically true, one would expect (c) itself to be logically true. Since (a) is just shorthand for (c), that would make it logically true, too. Hence, all true statements formulating duties or obligations would be logically true. But this is absurd. A moral principle, or in general a principle of action, must be synthetic. A logical truth is not affected by what we do or by what happens in the world; in the case of deontic assertions, a logically true or analytic deontic assertion is one which cannot be disobeyed, because it prescribes no specific course of action, like 'Either A is permitted or not-A is permitted'. On the other hand, a principle of action *must* be disobeyable. Therefore, the important statements about obligations cannot be logically true; they must be enacted, and voidable in principle.

Now it is not clear that in Anderson's systems statements like 'A is obligatory' are logically true, if A is not formulated in a logically true description. If they are not, he allows of true statements about entailment which are not logically true, viz., those for which his definitions of 'P', 'O', and 'F' provide abbreviations. Thus, here his modal logic might be charged with incompleteness.

In any case, even if his deontic analogues are not logically true and his modal logic is completely satisfactory, it does not seem correct to identify an ordinary deontic assertion like 'You (Peter) ought to pay Smith before the end of the month' with a statement to the effect that a certain entailment holds. It may very well be true that making the deontic statement involves presupposing, or conveying that certain entailments hold, just as it may express a willingness to make other utterances. The deontic assertion is such because it prescribes or enjoins a course of action; the entailment statement reports that a certain logical relationship exists, and it exists regardless of whether Peter or anybody else has obligations or not. When Anderson tells Peter 'Your not paying Smith soon entails that S will be true', Peter can press him quite correctly 'Fine, but *ought* I to pay him soon or not?' This question will be out of place only if Peter knew that he *ought* to avoid making S true. But, once more, this later knowledge cannot be that the tautology 'If S, then S' is true; nor is it that the falsity of S entails that S' is true, unless he knows that he ought to avoid S'. But again, this knowledge... *ad infinitum*.

(5) For a thorough argument for this cf. a pap. *Semantics and Necessary Truth* (New Haven, Yale Univ. Press, 1958), pp. 119-127.

Hence, I conclude that Anderson's theorems below (FANS, 49) are false in the deontic interpretation he proposes^(5a):

OM7. $Op \equiv Nec (\sim p \supset S)$ 'p is obligatory and only if not-p entails the sanction'.

OM8. $Fp \equiv Nec (p \supset S)$ 'p is forbidden if and only if p entails the sanction'.

OM9. $Pp \equiv M (p, \sim S)$ 'p is permitted if and only if it is (logically) possible that p is true and the sanction is false'.

Let us take an example for OM9. Let 'S', i.e. the sanction mentioned in the Penal Code, be 'you will be put in jail for 10 years'; and 'p' be 'you will be put in jail for 9 years'. Clearly, it is logically possible to put you in jail for 9, but not 10 years. Thus, it follows logically that it is permitted to put you in jail for 9 years — without ado!

In general, according to OM9 one can find what one is permitted to do by drawing a list of all punishments: *Whatever is consistent with their denial is permitted*. Similarly, the task of legislators could be reduced simply to making lists of punishments. At this point one wonders why it should be said that driving on the left is not permitted, for it is compatible with all sanctions listed in the laws: jail terms, fines, gas chamber, electric chair, etc.

5. *The 'naturalistic' fallacy*. Moore⁽⁶⁾ says that the 'naturalistic fallacy' consists in confusing goodness or *being-good* with any non-ethical object or property whatever (PE, 13, 38f, 114). In particular this fallacy is committed in any attempt at defining 'good', for this word designates at least part of what is 'common and peculiar' to all ethical concepts (PE, cfr transition from end of § 1 to end of § 2). Since 'ought (to exist)' and 'good' are absolutely synonymous (PE, viii, 115 17), every attempted definition of the former commits the fallacy⁽⁷⁾. However, he emphasized that 'ought (to do)' is de-

^(5a) They are proved in von Wright's system M, which is contained in Lewis' S4 and S5. Hence deontic logic cannot be reduced to S4 or S5.

⁽⁶⁾ G. E. MOORE, *Principia Ethica*, Cambridge University Press, 1903, to be referred to as PE.

⁽⁷⁾ In connection with FRANKENA's *The Naturalistic Fallacy*, in *Mind*, XLVIII (1939) and reprinted in SELLARS and HOSPERS, *Reading in Ethical Theory*, New York, Appleton-Century-Crofts, 1952, it should be remarked that, according to Moore, if the "fallacy" is committed even when we identify goodness with an indefinable property (PE, 13, 16 f), and (ii) it completely covers the ought-is bifurcation.

finable (PE, 24f, *et al.*). It might be thought that Anderson defines 'ought (to do)' and so avoids Moore's charge. But since he does not make Moore's distinction, his definition of 'O' will have to cover both 'ought's, so that he is after all defining 'ought (to exist)'. In fact his symbol 'O' can be prefixed to any statement whatsoever, regardless of whether it is about acts or not. Anderson quite consistently often reads his 'O' as 'ought to be true' (FANS, 62, 64) and as 'ought to be the case' (FANS, 61).

One of Moore's arguments against the so-called fallacy is roughly, that an ethical principle (or, more generally, a principle of action) has to be a synthetic proposition and not one about entailments, whereas every proposed definition of 'ought (to exist)' would transform some principles of action into analytic or logically true propositions (PE, 7, 12, etc.). In § 4 above we have seen how on Anderson's definitions every deontic statement becomes a mere abbreviation of a statement about entailment, or lack of entailment.

6. *Acts.* My arguments have shown that Anderson's proposed definitions of deontic concepts do not provide adequate analyses because he conceives the relation between duty and punishment as one of entailment, whereas in actual deontic language it is much weaker. But there is another important point overlooked in those definitions. 'Ought', 'obligation', 'right', 'wrong', etc. are terms which apply primarily to actions, not *qua* happenings or states of affairs. To say that something ought to exist, as Prichard remarked a long time ago⁽⁸⁾, is to say in an abbreviated form that somebody did not do what he ought to have done, or is expected to do what he ought to do. Of course, some persons may have used the word in a different sense; but the deontic sense is that which it has in the context of action, and that is obviously the one Anderson is interested in: only in the context of actions and agents makes sense to speak of punishments.

My point is that aside from the question of entailment, Anderson errs in prefixing the deontic operators to descriptions of states of affairs. They must be prefixed to names or descriptions of acts in a sense in which these are not viewed as, or merely as, possible events or states of affairs. This is a very important respect in which Von Wright's system of propositional normative logic was more insightful: he simply assumed that his deontic operators were pre-

⁽⁸⁾ H. PRICHARD, *Does Moral Philosophy Rest on a Mistake?*, in *Mind*, n.s. XXI (1912), reprinted in SELLARS and HOSPERS, *op.cit.*, 152.

fixed to names of acts, even though by not discussing the difference between an act in that sense and an event he left the door open to Prior's and Anderson's identification of acts with possible events.

I cannot develop this here, but it seems to me that the sense of act needed is that in which an act is 'described' or 'named' by an imperative sentence, or by a sentence used to formulate decisions to perform (which I have called *resolutives*). That is, deontic or normative logic seems to me to be a modal logic of imperatives and resolutives, but not fully analogous to ordinary alethic logic, because (as we have seen) the 'necessity' of a norm is far from being a necessity of meaning — let alone the necessity of entailment.

(⁹) A detailed formalization of such a deontic logic with imperatives and statements of fact, which is both sentential and quantificational, can be found in a system N* presented in the work mentioned in footnote 2 and *Un sistema general de lógica normativa*, in *Dianoia* (Mexico), III (1957), 303-333; a system M* with improvements is outlined in *The Logic of Obligation*, in *Philosophical Studies*, X (1959), 17-23. For philosophical points connected with the logic of norms, cf. *Imperative and Deontic Logic*, in *Analysis*, XIX (1958), *Imperative Inference*, in *Philosophy and Phenomenological Research* (forthcoming) and 'Ought' and Assumption in Moral Philosophy to appear in *The Journal of Philosophy*.

Wayne State University, Detroit