DEONTIC LOGIC AND THE LOGIC OF IMPERATIVES

E. J. LEMMON

In [8], R. M. Hare advanced the thesis that there is both a similarity and a connexion between ethical judgments and imperatives: both are forms of prescriptive language (1), and ethical judgments, at least if they are action-guiding, entail imperatives (2). If the second of these propositions is to be tested, or even fully understood, then the notion of entailment must be suitably extended, from its normal application to statements or sentences, to cover imperatives as well; and this requires a logic of imperatives in parallel with or as a supplement to the usual logic(s) of indicatives. There have been sporadic attempts to provide this over the last 20 or so years (3). There have also been advanced many systems of deontic logic, which purport to explore the logic of such notions as 'ought', 'obligatory', and 'permitted' (3). The purpose of the present paper is to survey somewhat critically a few of these attempts and to make an assessment of Hare's theses in the light of this survey; its generally negative tone is alleviated, it is hoped, by some constructive suggestions towards the end concerning a logic of commands.

1. *The structure of deontic logics.* There are three main brands of deontic logic currently on the market, which we may call the naive absolute, the alethic absolute, and the relativistic variety. Let us state briefly the formal outline of each of these.

Naive absolute systems of deontic logic are obtained in a fairly definite way from so-called alethic modal systems, such as Lewis's S1-S5. I shall illustrate this procedure with respect to a fairly weak propositional modal logic, which I labelled E2 in [12] (4). The formation-rules of E2, as of all such systems, can be taken as follows: (i) propositional variables 'p', 'q', 'r', ... are wffs; (ii) if A and B are wffs then so are $\sim A$, $\Box A$, and $(A \supset B)$. Axiom-schemata for E2 are:

$$\text{(A1)} \quad A \supset (B \supset A);$$

(1) Hare [8], p. 3.

(2) Hare [8], p. 163ff.

(3) See the bibliography for some main references.

(4) E2 is contained in and closely related to S2; it has no theorem of the form $\Box A$.

(A2) $(A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))$;
(A3) $(\sim A \supset \sim B) \supset (B \supset A)$;
(A4) $\Box(A \supset B) \supset (\Box A \supset \Box B)$;
(A5) $\Box A \supset A$;

and its two rules of inference are:

(R1) from A and $A \supset B$, to derive B;
(R2) From $A \supset B$, to derive $\Box A \supset \Box B$.

Then the corresponding deontic logic D2 is obtained by dropping (A5) in favour of:

(D5) $\Box A \supset \sim \Box \sim A$.

Most axiomatizations of alethic modal logics contain either (A5) or some deductively equivalent schema, and a corresponding deontic system can be obtained by replacing either (A5) or its equivalent by either (D5) or some equivalent, together perhaps with other slight modifications (5).

In alethic modal logics, '$\Box$' is to be read 'it is necessary that...'; the operator '$\sim \Box \sim$', often abbreviated to '$\Diamond$', can be read 'it is possible that...'. There is some doubt concerning the correct interpretation of '$\Box$' in deontic systems; thus Prior [17] initially suggests 'it is (morally) obligatory that ...', but later uses what he seems to regard as mere stylistic variants: 'we are obliged', 'we have an obligation', 'we ought'. It is also expected that '$\Diamond$', as defined, will deontically bear the interpretation 'it is permitted that...'. Under this sort of interpretation (A4) may be taken as the assertion that what is obligatory, like what is necessarily the case, is closed under detachment, and the rule (R2) corresponds to the principle that the *logical* consequences of what is obligatory are obligatory. The dropping of (A5), however, in favour of (D5) reflects the conviction on the part of deontic logicians that, whilst what is necessarily the case is the case, it is not always true that what is obligatory is actually performed or that what we ought to do we in fact do; on the other hand, it is always false that an action and its 'negation' are both obligatory, and this is what (D5) on the present interpretation affirms.

The following can readily be proved as theorem-schemata of D2,

(5) D2 is equivalent to a deontic system discussed in Prior [17], pp. 220-229, which is closely akin to von Wright's original system in [24] and [25].

if we adopt the obvious definitions of propositional calculus connectives:

(1) $\Box(A \& B) \equiv \Box A \& \Box B$;
(2) $\Diamond(A \vee B) \equiv \Diamond A \vee \Diamond B$;
(3) $\Box A \vee \Box B \supset \Box(A \vee B)$;
(4) $\Diamond(A \& B) \supset \Diamond A \& \Diamond B$;
(5) $\Box A \supset \Box(B \supset A)$;
(6) $\Box \sim A \supset \Box(A \supset B)$;
(7) $\Box(A \& \sim A) \supset \Box B$.

It is important for what follows that none of the proofs of (1)-(7) uses (D5). As Prior has shown in [17], the rule

(R3) from A, to derive $\Diamond A$

can be established for D2; but its establishment rests essentially on (D5). Prior suggests (p. 224) that from (R3) follows the Kantian principle 'What I ought, I can'. Richer naive absolute systems of deontic logic may be presumed to contain as much and more besides; for example, the schema

(8) $\Box(\Box A \supset A)$

has been suggested by Prior as a deontico-logical truth ('it ought to be the case that what ought to be the case is the case'), yet this is not provable in D2.

In [1], A. R. Anderson suggested in effect a 'reduction' of deontic logic to alethic modal logic. To such a logic let there be added a propositional constant 'σ', which is to be understood as a moral sanction (it may be read, according to Prior [18], p. 140, as 'the world will be worse off' or 'punishment ought to follow' or 'something of that sort'). Then we may define:

$$OA =_{df} \Box(\sim A \supset \sigma);$$

A is obligatory if and only if the non-fulfilment of A strictly implies the sanction. Thus if 'σ' is added to E2, (A 4) with 'O' replacing '$\Box$' is provable in it and (R2) with similar replacement derivable for it. Further, (D5) under the same replacement is provable if we add to E2 the axiom:

(A6) $\sim \Box \sigma$.

Let us label that extension of a modal logic L which results from adding the constant 'σ' and the axiom (A6) L(d); and by A′ let us understand the wff which results from A by replacing each occurrence of '$\Box$' by 'O'. Then it is immediate that if A is a theorem of D2, A′ is a theorem of E2 (d).

The 'Andersonian simplification' of deontic logic, as the above definitions of other deontic concepts have sometimes been called, is of great importance because it enables us to study the properties of deontic logics by reference to known results concerning alethic modal systems. As a trivial example, it can readily be shown that (A5) is *not* a theorem-schema of D2 on the grounds that counterexamples to (A5)′ can be found for E2(d) from known properties of E2. We can also see that, although (8)′ is not a theorem-schema of E2(d), it is a theorem-schema of S2(d), and indeed of the deontic extension of any modal system which contains $\Box(\Box A \supset A)$ (and substitutivity of material equivalents). Indeed, L(d) enables us to study the interrelation between deontic operators and ordinary alethic ones for which naive absolute systems do not have the expressive power. For example, in E2(d) we have the implicational chain:

$$\begin{array}{ccc} \Box A & \searrow & \\ \downarrow & & OA \\ A & & \downarrow \\ \downarrow & & PA \\ \Diamond A & \swarrow & \end{array}$$

in which PA represents $\sim O \sim A$ (with the interpretation 'A is permitted') and '$\rightarrow$' represents a material implication provable in the system. '"Ought" implies "can"' receives thus a formal fiat. Deontic logics obtained in this manner from alethic modal logics I am terming systems of the alethic absolute variety.

At about the same time that Anderson propounded this reduction of deontic to ordinary modal logic, von Wright published a paper [26] proposing a system of *conditional permission*, for which the primitive symbol was 'P(p/c)' with the interpretation 'p is permitted under condition c', which was mainly intended to avoid the "paradoxes" of deontic logic (5) and (6) to which Prior [16] had drawn attention. The system was enriched by Rescher [19], attacked by Anderson [2], modified by Rescher [20], and finally reattacked by Anderson [3]. It is perhaps fair to suggest that the meat of this controversy was the question whether 'doing A commits us to doing B' can be satisfactorily formalized within the framework of *absolute* deontic logics of the kind already discussed or whether a system of *relativized* permission and obligation should be designed to this end. If 'A commits us to B' is represented as '$O(A \supset B)$', then by (5) the doing of any act commits us to an obligatory one and by (6) the doing of a forbidden act commits us to any act. Even if we represent 'A commits us to B' by '$A \supset OB$', the first paradox, though not the second, survives.

It is hard to discuss the merits of this relativistic brand of deontic logic, since at present there is neither an agreed axiomatization nor any clearly defined interpretation. But, if we write '$O(p/c)$' for 'p is obligatory in circumstances c', then the relativized version of (D5), namely:

$$(9)\ O(p/c) \supset \sim O(\sim p/c)$$

is accepted as a theorem by both von Wright, Rescher, and Anderson. To this extent at least some of the strictures of the next section apply with equal force to the third variety of deontic logic.

2. *A critique of deontic logics.* It will be seen from the preceding section that there is a wide diversity of views on what is and what is not a logical truth involving deontic concepts, which is reflected in the multiplicity of approaches to deontic logic in the literature. Yet there remains a measure of agreement: in particular, it is widely agreed that (A5) has to be rejected on deontic interpretation and (D5) accepted in its place. This is one fundamental feature of deontic logics which I wish to question.

In the first place, however, it seems obvious that much more needs to be supplied for the interpretation (in ordinary or semi-technical language) of deontic logics before a rational choice between them can be made. The same, I think, applies to a lesser degree to ordinary modal systems: for a precise interpretation of 'it is necessary that...' some modal logics are correct, others incorrect, as I argued in [13]. Here interpretational problems beset us from the beginning. Over what kind of assertions are our variables to be allowed to range? Prior [17], p. 221, following in principle von Wright [25], suggests that they range only over assertions of the form 'an act of such-and-such a sort is done'. With this in mind we can perhaps construe '$\Box A$' as 'it is obligatory that the (perhaps complex) act A be done'. But on this view '$\Box A$' will presumably not itself be an assertion to the effect that an act of some kind is done, so that we are left with no interpretation of iterated modalities, or indeed of any formula (such as (8)) in which one occurrence of '$\Box$' occurs within the scope of another. Any discussion of the interpretational validity of (8) is thus doomed to futility so long as the range of propositional variables is restricted in the proposed manner. (In fact such a restriction should be reflected either in the formation-rules or in the rule of variable substitution for the proposed system; von Wright originally followed the former course, but later writers, myself included, have not been so cautious.)

It may be that a sense exists for, or can be given to, 'it is obligatory that...' in connexion with a wider class of assertions than those asserting that an action is done, so that no such restriction is needed. But if so, the wider class should be specified and the sense explained. If no restriction at all is given, then we are compelled to do our best to construe such sentences as:

(10) it is obligatory that Brutus killed Caesar;
(11) it is obligatory that there is life on Mars;
(12) it is obligatory that 7 is prime;

and this seems not easy. My case is (a) that we are not told enough about the range of propositional variables to give an exact interpretation to deontic logics and (b) that it is far from clear what if any interpretational ruling would be satisfactory.

Next, as I remarked in the last section, the deontic interpretation of '$\Box$' is itself unclear. Though von Wright sticks carefully to 'obligatory', Prior and others have freely used 'ought' or 'obliged' in their analyses. To see that this makes a great difference, let us begin with 'x is obliged to...'. It seems clear, first, that the most normal employment of this idiom is as a synonym for 'x is *forced* to...' in quite non-moral contexts. 'Because the road was blocked, I was obliged to take an alternative route'. But second, inasmuch as 'obliged to' is usually synonymous with 'forced to', it satisfies the stronger principle (A5) as well as (D5): what we are obliged to do or forced to do, we do, in the sense that if we do not it it cannot have been true that we were obliged to; just as what we know to be the case, is the case, in the sense that if it turns out not to be the case then it cannot have been true that we knew it. Thus it is implicitly contradictory to say that we were obliged to take another route but did not do so; if we did not take another route, then we might have thought that we would be obliged to do so but cannot actually have been so obliged.

Thus 'obliged', like 'forced', 'had to', and 'must', obeys it seems the usual alethic principle (A5) and is not characterized by its failure, as some deontic logicians would have us believe. However, it may be argued that, even if 'obliged' is usually non-moral and satisfies (A5), the genuinely moral 'obligatory' does not do so, any more than 'x is under an obligation to...'. Now the situation seems here to me unclear. I think there *may* be a sense of 'obligatory' which is like 'obliged' in that what is accepted by someone as obligatory *for him* must be done *by him*. Suppose for example that we consider it obligatory to stand up for a lady on a bus: then if Jones does not do so, this seems to show that he does not accept this as obligatory even if we do. If I do not do something, it seems to follow that I do not accept it as obligatory to do that thing.

This is not the whole story, however. For there seems also to be a sense of 'obligatory' in which that which is obligatory is just that which I am under an obligation (*not* obliged) to do. Yet I can certainly agree that I am under an obligation to do A and yet do not do A, if only because I am prevented by circumstances (e.g, illness) from doing A. And the very notion of my failing (for whatever reason) to fulfil my obligations must surely have as part of its meaning that there was something which I was under an obligation to do yet did not do. So that 'obligatory', in the sense of that which I am under an obligation to do, does not obey (A5). However, *this* sense of 'obligatory' does not obey (D5) either.

We need only, I think, consider the case of conflicting obligations: here I am both under an obligation to do A and under an obligation to do B, where the doing of B entails the non-doing of A. It is surely natural do describe this as a situation where I am both under an obligation to do A and under an obligation to do not-A, since not-A follows from B, so that it provides a counterexample to (D5). It may be argued that these are merely *prima facie* obligations, one of which will «disappear» when our true moral situation, what we «really» ought to do, has been revealed to us. This view seems to me to make the moral life too easy. Perhaps Ross's term *prima facie* is here ill-judged; rather it is essential to our perplexity when faced with conflicting obligations that we *really are* under an obligation to do A and *also* under one to do not-A (e.g. we *really did* give conflicting promises). The successful moral resolution of this jam, if it comes at all, will not, I think, take the form of the discovery of what we really ought to do; in one sense, we already know that and precisely cannot do it; it will rather take the form that, given that we cannot do what we ought to do, the *best* thing to do is ....

Nor will it do, I suspect, to say that, although (D5) may have to be dropped, nonetheless (9) may be retained: it might be argued that the circumstances (reasons or whatever) that yield the obligation to do A (my promising to do *that*) are different from the circumstances that yield the obligation to do not-A (my promising to do B which is incompatible with A); thus the conflicting obligations are relative to different circumstances, and (9), though not (D5), is preserved. But, insofar as I understand the notion of conditional obligation, I would contend that, if A is obligatory in circumstances c, then it is *a fortiori* obligatory in circumstances c & d; hence we can construct from our present case a counterexample to (9) by forming the conjunction of the circumstances that yield the conflicting obligations. In any case, in the chosen example the different circumstances have been realised (we have given the conflicting promises), and it appears to me inescapable that in these composite circumstances we have both an obligation to do A and one to do not-A.

It may further be argued that our being faced with this moral situation merely reflects an implicit inconsistency in our existing moral code; we are forced, if we are to remain both moral and logical, by the situation to restore consistency to our code by adding exception-clauses to our present principles, or by giving priority to one principle over another, or by some such device. The situation is as it is in mathematics: there, if an inconsistency is revealed by derivation, we are compelled to modify our axioms; here, if an inconsistency is revealed in application, we are forced to revise our principles. This argument is hard to rebut, partly because the notion of consistency for a code, containing as it does principles, maxims, and the like rather than axioms and rules of inference, is by no means as clear as the corresponding notion for a formal system (in this respect we have a problem analogous to one awaiting us in the logic of imperatives: what is it for a set of rules or a set of commands to be consistent ?). But two lines of attack seem possible. Firstly, in the sense required, it is unlikely that any moral code containing at least two principles of conduct is consistent, i.e. that all *possibilities* of situations of conflict have been catered for by exception-clauses and the like. However, this reply suffers from the defect of being as obscure as the argument it is meant to rebut. Secondly and more usefully, even if the case of conflicting obligations does show that our code is *in some sense* inconsistent, it does not follow that it is the part of a moral and logical man to modify it. For 'I am under an obligation to do A' and 'I am under an obligation to do not-A' are not *formally* inconsistent; nor, as we shall see later, can a contradiction be deduced from them without begging the very question at issue; so it is not necessarily *illogical* to retain both, together with the principle(s) from which they come. And I cannot see that it is *immoral* to retain one's principles even though they have landed one in a fix: indeed this may be what we expect from a *very* moral man. There is nothing either morally or logically weak about a man who says that one ought always to keep one's promises, though sometimes this may become impossible to do.

It will no doubt be seen that I am in effect denying the Kantian '"ought", implies "can"', so to this I must turn. First, it seems to me that the above considerations concerning 'I am under an obligation' apply with equal force to 'ought': namely that 'I ought to do A' and 'I ought to do not-A' are not contrary, and therefore 'ought' does not obey (D5) either; for, I would claim, it follows from my being under an obligation to do A that I ought to do A. But I have argued this in [14], and will not repeat the case here. It was also pointed out in [14], note 8, p. 150, that abandoning (D5) for 'ought' entails abandoning '"ought" implies "can"' ('logically can'), at least if 'ought' obeys (1) as I think it does. Put contrapositively, if 'I ought to do A' implies 'I can do A', then 'I ought to do A' implies 'It is not the case that I ought to do not-A'; for suppose I both ought to do A and ought to do not-A; then I ought to do both A and not-A by (1), whence I can do both A and not-A by the Kantian principle, which is modally absurd. The converse also holds, as the derivation of the rule (R3) shows. Let us informally use 'OA' for 'I ought to do A', '$\Box$A' for 'I have (logically) to do A', and '$\Diamond$A' for 'I can (logically) do A', and suppose that whenever OA then $\sim O \sim A$. Suppose now OA and yet $\sim \Diamond A$. Then $\Box \sim A$ (by obvious modal principles), whence $\sim A$ is a logical truth, so that so is $A \supset \sim A$ by propositional calculus. Hence, by (R2) applied to 'ought', $OA \supset O \sim A$ is a logical truth, so that $O \sim A$ holds, contravening $OA \supset \sim O \sim A$. Thus for *any* deontic operator, the Kantian principle and (D5) in application to it stand or fall together; the principles are deductively equivalent in any natural modal logic equipped to express both, and we need only (1) and (R2), together with simple alethic principles, to establish this.

The fact that these two principles are equivalent means, of course, that I cannot *show* that 'ought' does not imply 'can' except by appealing to what I consider to be counterexamples to (D5); for counterexamples to one, if there are any, will also be counterexamples to the other. But similarly anyone wishing to *uphold* (D5) must equally be prepared to uphold the Kantian principle. So that the form my opponent's argument should now take will be: since the Kantian principle is intuitively *so* obvious, (D5) must thereby be accepted. My reply is equally clear: since there are counterexamples to (D5), our intuitions concerning the Kantian principle must be at fault.

To loosen up our intuitions a bit, if I may use the phrase, let us consider some clear counterexamples to the Kantian doctrine, in which however the 'can' is not the 'can' of logical possibility. I ought (for the sake of my health) to touch my toes each morning, but simply cannot (I no longer have the physical power). This is perhaps not a moral 'ought'. I ought (I think *morally*) to give up smoking but simply cannot (I no longer have, perhaps never have had, enough will power). But this is mere physical or psychological impossibility. To get done everything I have to do, I ought to be in two places at once (this could be a moral 'ought'), yet I evidently cannot without breaking the laws of nature (or logic ?) — often said by way of tacit *modus tollens* to show that I cannot get everything done which I have to do. I ought to burn these manuscripts (the poet made me promise to do so on his deathbed), but I simply cannot (of course I have the physical and psychological power—it is just that my aesthetic sense would be outraged by the act). I ought to be practising my recorder, but I cannot (I haven't the opportunity, I forgot to bring it with me). If the logical 'can' obeys the Kantian principle, it appears to be the only 'can' that does.

It will be as well in this connexion to consider Hare's recently expressed views in [9], ch. 4, since they differ somewhat from his original position in [8], which was the starting-point of this paper. He agrees that 'ought' does not universally imply 'can' (p. 51), and that exceptions occur not only when 'ought' has its "inverted commas" sense, invoking a moral convention to which the speaker does not necessarily himself adhere (p. 52), but also when 'ought' is prescriptive; this use, however, 'falls short of ... a *universal* prescription which would apply to [the speaker's] own case' (p. 53). Such a usage is an example of a 'declension which, to match the human weakness of their users, moral judgments commonly undergo'. To say 'I ought but I can't' is as if to say 'If I were able, it would be the case that I ought (full force); but since I am not able, that lets me out'. (In the following chapter, Hare uses this kind of analysis to explain *acrasia* or backsliding.) Now first, the employment of 'ought but can't' by an agent is by no means always a sign of moral weakness; the weak man, finding that he cannot do what he believes he ought to do, is more likely to *deny* that he ought, and it may require considerable moral courage to admit the moral truth. Thus 'ought but can't' is often far from being a 'let-out', as Hare's analysis suggests. Second, Hare has, in a way, conceded that 'ought' does not imply 'can', although he maintains that often, perhaps nearly always, what ought to be done can be done. If the Kantian principle is to be maintained despite the counterexample, it must be shown that 'ought' here has a different *sense* from the one in the principle. And Hare does think that its *force* is different. But this difference of force is simply made to exist by Hare's account of universal prescription (since the man who says 'ought but can't' cannot be prescibing for his own case, the prescription fails of full universality: why ? — because prescription implies possibility of action: at this point the question is begged). But surely we want to say that the 'ought' in 'I ought because I promised; so will' has exactly the same sense as the 'ought' in 'I ought because I promised; but can't'.

Hare's second line of defence (pp. 53-56) is that 'implies' in ' "ought" implies "can" ' is not entailment but a relation analogous to Strawsonian presupposition. Just as 'the present king of France is wise' implies or presupposes that there is a unique king of France in the sense that unless the latter statement is true the question of the truth or falsity of the former does not arise (6), so, he claims, 'I ought to do A' (with the full force of 'ought') implies or presupposes that the *prescriptive question* (p. 56). «What ought I to do ?' to which this is a possible answer arises, 'which it would not, unless the person in question were able to do the acts referred to' (p. 54). I ignore the question-begging of this argument for the more serious consideration as to whether anything can be made of the Strawsonian analogy. Can we, more strictly, propose that 'I ought to do A' is *neither true nor false* in case A is impossible for me ? I can see no justification for this move, and only the usual complications for those who adopt it.

Hare's real, and rather vague, reason for believing that 'ought'

(6) See STRAWSON [23]. Hare, incidentally, seems completely to misunderstand Strawson's position on p. 59, by proposing that the question of truth or falsity arises if and only if the question, if asked, would be *comprehensible*; this suggests that the truthvalueless 'The present king of France is wise' is meaningless, which Strawson was at pains *not* to say.

implies 'can' is that a prescriptive question can only arise if a practical question 'What shall I do?' arises, and the latter cannot arise if a man cannot but do what he is going to do (pp. 55-56). But this by no means shows that 'ought' implies 'can', and is quite consistent with my earlier counterexample. No doubt there too the question 'What shall I do?' has arisen (is comprehensible) and with it the question 'What ought I to do?' The trouble is that the answer to the latter question, 'I ought to do both A and not-A', does not provide an executable answer to the former: hence the dilemma.

Hare further believes that the prescriptivity of 'ought' explains the commonly accepted notion that 'ought' implies 'can'. Since I am here denying that 'ought' implies 'can', it is reasonable for me also to try and explain its common acceptance, which I do not dispute. I think there are two possible, and not incompatible, explanations. Firstly, a common employment to which the idiom 'You ought to have done A' is put is to *blame* someone for not doing A by accusing him of failing to do A which was his duty or which he was under an obligation to do; and it is commonly held that blaming someone for not doing A is only in place if he could have done A, just as praising him for doing A is only in place if he might have done not-A. But this common view is wrong: it is sometimes proper and even useful to blame someone for failing to do something which he could not help failing to do, as a means of getting him to avoid in future the circumstances in virtue of which he was unable not to do it. Secondly, we commonly use the idiom 'You ought to do A' as a way of *advising* people what to do by pointing out to them where their duty or obligation lies; and it is certainly pointless to advise people to do what we know they cannot do (though, I think possible, and even sometimes done). However, this employment, far from suggesting that 'ought' implies 'can', suggests only the weaker thesis that 'x ought to do A' implies that the speaker *thinks* that x can do A. In any case, this is not the only use of 'you ought to do A': in the same *sense* of the words, I suggest, we can employ the idiom merely to point out to someone what his moral position is, leaving it to him to draw any conclusions relevant to action.

I have been contending, in effect, that, while some deontic modes, such as 'obligatory' in one sense and the ethical 'must' and 'have to', obey the the principle (A5) as well as (D5), and so are pretty much like the usual alethic modality '$\Box$', others, such as 'obligatory' in another sense, and 'ought', obey neither (A5) nor (D5). Hence the system (A1)-(A4) and (D5) catches the logical properties of no deontic concept with which I am familiar. But this is related to a different, and more important, point, as I hope is now clear: that deontic logic, as usually formulated, is beset by the two unduly optimistic assumptions concerning morality (a) that cases of real moral conflict do not — indeed cannot — occur and, equivalently, (b) that what we ought (are under an obligation, etc.) to do we always can do.

Although I have been chiefly concerned to attack naive absolute systems, the burden of the attack carries over to the other varieties since they accept, in one form or another, the principle (D5). There are particular interpretational difficulties concerning the sanction 'σ' in the Andersonian variety, as was argued in [15], to which I do not think the reply in Smiley [22] is sufficient answer. But rather than discuss these here, it is perhaps worth asking what remains to naive deontic logic if we drop (D5).

As was observed earlier, surprisingly much survives: at least the laws (1)-(7) of the previous section. Still, it is hard to feel they have much 'bite'; they consist either of truisms ((1)-(4)) or paradoxes ((5)-(7)). In particular, it may be felt that (7) is a good deal more paradoxical if the standpoint of the present section is adopted. For I have argued that $\Box(A \mathbin{\&} \sim A)$ is sometimes true under deontic interpretation, in which circumstance anything at all is obligatory. This might be felt to cast doubt on the principle (R2) by which (7) is directly obtained from the corresponding paradox of material implication.

Perhaps my greatest unease concerning deontic logic is as yet unexpressed, and not easy to express. It is that it is hard to imagine an agent *using* deontic logic in other than quite trivial ways; since (D5) in effect afffirms the consistency of our moral commitment, it at least has some content, though I think it is false. There is nothing particularly 'deontic' about the remaining axiom-schemata. It remains true that we argue, often with some subtlety but more often with a feeling of despair, in and about moral situations; but none of these arguments, or only their very dullest features, find codification in deontic logic. Perhaps the unease can be crystallized in this form: so far there have not been enough parameters to deontic operators, and the symbol '$\Box$A' leaves too many questions unanswered — 'it is obligatory that A', but on whom, to whom is this obligation due and when, how was it incurred and when ? For a useful deontic logic, we shall need, I suspect, at least variables for people, variables for individual acts, and variables for times, together with appropriate quantifiers (7).

(7) One excellent move in this direction is in HINTIKKA [10], but there only

3. *The logic of imperatives.* It would not be proper, as in the case of deontic logics, to speak of different brands of imperative logic, for as yet there is no consensus of opinion in this field and only one formal system (as distinct from sketches) available — that of Hofstadter and McKinsey [11]. These authors begin by contrasting *directives* (imperatives together with an agent's name, such as 'Henry, close the door') with *fiats* (such as 'let there be light', where no agent is mentioned), and decide to concentrate on the logic of the latter. I shall outline the formal system which they present, ignoring for the sake of simplicity the handling of quantifiers, though they do include these. To some formulation of the classical propositional calculus, with 'p', 'q', 'r', ... as sentential variables and truth-functional operators '$\sim$', '$\supset$', '&', 'v', '$\equiv$', we add the symbols '!', '—', '+', '$\times$', '$\rightarrow$', and '='. To the obvious formation-rules we add the following: (i) if S is a sentence, then !S is an imperative; (ii) if C_1 and C_2 are imperatives, then $-C_1$, $(C_1 + C_2)$, $(C_1 \times C_2)$ are imperatives; (iii) if S is a sentence and C and imperative, then $(S \rightarrow C)$ is an imperative; (iv) if C_1 and C_2 are imperatives, then $(C_1 = C_2)$ is a sentence. Thus the clauses (i)-(iv) together with the usual clauses provide a simultaneous definition by recursion of *sentence* and *imperative*.

To motivate this procedure interpretationally, it should be recalled that it has been the hope of several authors (e.g. Hare [7]) that the logic of imperatives could be developed in fairly complete parallel with the familiar logic of indicative sentences. Thus the fact that the imperative mood is, in most languages, sadly deficient in tenses and persons is regarded as a remediable and contingent feature by Hare [7], pp. 25-27, who concludes 'I shall therefore assume that a logician is entitled to construct imperatives in all persons and in all tenses'. This hope is daunted at the outset by the uncertainty as to what are the proper analogues of *truth-values*. There are at first sight two possibilities: (a) an order may be either *obeyed* or *disobeyed* in something of the way in which an assertion may be either true or false; (b) an order may be either *in force* or *not in force*, as an assertion may be either true or not true. Now it is fairly clear that suggestion (b) will not in fact lead to a development at all parallel with assertions. When an assertion is not true, then there is an assertion, namely its negation, which is true. But it may be that an order is not

variables for individual acts are introduced and quantified over, not for people or times. For a suggestion concerning the analysis of obligation, but no logic to go with it, see Section 7.

in force without there being any 'counter-order' which *is* in force: there may be no orders at all in force in a given context.

Suggestion (a) is immediately more fruitful. We say that an order to do p is *obeyed* if and only if p becomes the case (Hofstadter and Mckinsey use 'satisfied' for 'obeyed' and 'is' in place of 'becomes'; Geach [6] speaks of 'fulfilment' and 'coming true'). Interpretationally, we may think of imperatives as being constructed out of future-tense sentences of a certain kind ('it will become light'): this operation is '!' in [11]. Then we may think of the *complement* of an imperative !S, symbolically —!S, as that imperative which is obeyed if and only if S comes out false, i.e. as identical with the imperative $!\sim S$. 'Don't shut the door', it may be hoped, is the complement in this sense of 'shut the door', and obeyed if and only if the door is not shut. Similarly, we may define the *sum* of imperatives !S and !T, !S + !T, as that imperative which is obeyed if and only if either S or T comes out true, i.e. as identical with $!(S \vee T)$: 'either shut the door or open the window' is obeyed if and only if either the door is shut or the window is opened. The *product* of two imperatives, $!S \times !T$, has a similar relation to '&', and is identical with !(S & T). If we write '$(C_1 = C_2)$' to mean 'C_1 is obeyed if and only if C_2 is obeyed', then the following, which Hofstadter and McKinsey offer as axiom-schemata, are evident:

(13) $-!S = !\sim S$;
(14) $!S + !T = !(S \vee T)$;
(15) $!S \times !T = !(S \,\&\, T)$.

In English, apart from compounding complex imperatives out of simpler ones by means of 'not', 'and', and 'or', we also frequently use conditionals whose antecedents are indicative but whose consequents are imperative. Thus 'if it is fine, open the window; but if it is raining, close the door'. It is not clear, however, what sense if any attaches to 'conditionals' whose antecedents and consequents are imperatives, or whose antecedents are imperatives and consequents indicatives, though we have a use for 'only if it is raining shut the door'. The 'only if' form may in any case be reducible to the other: the example given seems equivalent to 'if it is not raining do not shut the door'. Hence Hofstadter and McKinsey introduce $(S \rightarrow C)$ as an operator forming an imperative out of a sentence S as antecedent and imperative C as consequent. Its obedience conditions are formed by analogy with material implication: $(S \rightarrow C)$ is obeyed if and only if either S is false or C is obeyed. This gives the schema:

(16) $(S \rightarrow C) = -!S + C$.

The remainder of their axioms concern quantifiers and substitutivity and other properties of '='. For example, and most obviously:

(17) $(!S = !T) \equiv (S \equiv T)$.

On the basis of these axioms, they establish one metatheorem to the effect, roughly, that imperative logic is trivial. For the special case of the propositional calculus (PC), letting PC(i) be the extended system which results from PC by adding the Hofstadter-McKinsey symbols and axioms, we may state the result thus: if S is a sentence of PC(i) then there is a sentence T of PC such that $\vdash_{PC(i)} S \equiv T$; if C is an imperative of PC(i) then there is a sentence T of PC such that $\vdash_{PC(i)} C = !T$. Thus all imperative connectives can be eliminated from a sentence and all imperative connectives save an initial occurrence of '!' can be eliminated from an imperative. Further, the sentence T correlated with an imperative C in this result is unique to within deriviability, since if $\vdash_{PC(i)} C = !T_1$ and $\vdash_{PC(i)} C = !T_2$ then $\vdash_{PC} T_1 \equiv T_2$. Thus we may call imperatives provable, refutable, etc., according as T is provable, refutable, etc. Notions such as consistency, defined for sets of imperatives, can similarly be taken over from the corresponding sentential notions. Roughly, a set of imperatives will be consistent if and only if they can be simultaneously obeyed.

In fact, as Fisher [5] remarks or implies ([8]), a simple decision procedure for the calculus is forthcoming by a combination of 'obedience values' O and D ('obeyed' and 'disobeyed') and truth-values T and F. E.g., the tables for '!', '→', and '=' are as follows:

([8]) Fisher, I think, is wrong (p. 233) to use the same symbols for both sentence-forming operators and imperative-forming operators; in this respect Hofstadter and McKinsey are models of clarity. In particular, it is impossible properly to extend the idea of a truth-table test to cover obedience-values if this is done.

S	!S
T	O
F	D

$S \rightarrow C$		C: O	C: D
S	T	O	D
	F	O	O

$C_1 = C_2$		C_2: O	C_2: D
C_1	O	T	F
	D	F	T

Then if S is a tautological sentence of PC, $\vdash_{PC(i)}$ C = !S if and only if C takes value O for all possible assignments of truth-values to its propositional variables.

4. *A critique of the logic of imperatives.* How trivial is the Hofstadter-McKinsey imperative logic, and is this a good or bad thing? Geach [6] writes that the triviality 'is the source of everything that can be said about the inferability, incompatibility, etc. of imperatives; their being imperatives does not affect these logical interrelations'. Hare [7], after drawing his well-known contrast between phrastic and neustic but in other terminology, contends that most inferences are from phrastic to phrastic and that we may add whichever set of neustics (whether 'yes' or 'please') we wish *salva validitate*, though he expresses reservations (which take shape in [8]) concerning inferential mixtures of indicatives and imperatives. A similar, though rather more cautious, view is expressed by Castaneda [4], who proposes the principle: 'an imperative inference is valid if it is the result of replacing one and the same imperative for each occurrence of a given indicative throughout a valid inference'. We embarked on a search for some exciting symmetry between indicative and imperative logic, but our navigation ends in a kind of dull isomorphism between the two.

To see some of the snags, let us begin with the problem of defining entailment for a 'mixed' set of indicatives and imperatives. It seems natural to say that such a set is *consistent* if it can be simultaneously the case that all the indicatives are true and all the imperatives obeyed, and otherwise *inconsistent*. Then, by analogy with indicative logic, we may say that an imperative C is *entailed* by indicatives S_1, ..., S_m and imperatives C_1, ..., C_n if the set consisting of S_1, ..., S_m, C_1, ..., C_n, —C is inconsistent. Now the set S, —!S is obviously inconsistent, so that S entails !S. ($S \rightarrow$!S is an 'analytic' imperative in the Hofstadter-McKinsey system.) But if there is this definable sense in which 'there will be a sea-battle tomorrow' entails 'let there be a sea-battle tomorrow', it is certainly not the sense of entailment which we hope to capture for imperative logic ([9]).

Our worry about the above purported entailment *may* be connected with Hare's principle in [8], p. 28: 'No imperative can be validly drawn from a set of premisses which does not contain at least one imperative'; for if it were correct it would be a counterexample to this principle. Geach attacks this principle in [7], but his own counterexample raises new difficulties: it has the form $a = b$. $\therefore - !Fb + !Fa$ ('Grimbly Hughes is the largest grocer in Oxford, therefore either do not go to the largest grocer in Oxford or go to Grimbly Hughes'), and therefore tacitly assumes the extensionality of imperative contexts. It seems, however, in fact likely that the imperative mood generates referential opacity: Krushchev may well have given to Gagarin the order 'be the first man in space', but, given that Gagarin was the first man in space, we cannot conclude that Krushchev's order to him was to be Gagarin; insofar as this latter order is comprehensible, it presumably has distinct obedience-conditions from the former ([10]).

Of course, we can perhaps discount this apparent exception to the extensionality of imperatives by pointing out that at the time when the order was given the identity was not true; indeed it was *made* true by Gagarin's very obedience to the order. And this at least reveals that in the testing for entailment or consistency of imperatives by or with indicatives due attention must be paid to *the time at which* indicatives are supposed to be true or false; for the obeying of an order normally *changes* the truth-value of a sentence — this we admitted when we defined obedience in terms of a future-tense sentence's *becoming* true — so that it matters enormously whether we are defining entailment, for example, in terms of *truth when the order is put in force* or rather in terms of *truth after obedience to the order*.

Let us remind ourselves of some truisms which may lie behind Hare's principle: no description of the present state of affairs alone can *entail* that a future event will take place (the truism behind the induction problem), but if a future event does not take place this, together with a description of the present state of affairs, may entail

([9]) It is fair to point out that this is *not* a valid inference by Casteneda's criterion, cited in the previous paragraph; but we shall find reason to reject his principle also later on.

([10]) That Geach's inference is unsound will, I hope, appear on other grounds from the discussion in Section 5.

that a universal law has been broken (a truism based on *modus tollens*). Hence no set of sentences true of the present can entail that an order *will* be obeyed, but if an order is not obeyed this fact together with these sentences may entail that some law has been broken. The rules of chess may be such that, your present position being what it is, unless you move your queen you are bound to lose; thus it may seem that your acceptance of these rules, together with your desire not to lose, 'entails' the imperative 'move your queen' — an entailment that does not hold between the mere *description* of your position and the order. We have only to add the idea that rules are somehow 'presciptive' (misleading because so many are permissive) to obtain what, I think, is more than a caricature of Hare's thesis: an order (prescription) will not follow from a mere description, but will follow from a description together with some other prescription. However, Hare's principle, insofar as it applies to imperatives and not to prescriptions in general, seems correct; the most plausible potential counterexamples which I can think of are 'tautologous' orders like 'either close the door or do not close the door', obedience to which, it may be claimed, is logically certain and so is entailed trivially by any proposition; and to these I will return. Equally, of course, it is clear that future-tenses sentences *can* entail that orders will be obeyed: if there will be a sea-battle tomorrow, it follows that the order 'let there be a sea-battle tomorrow' will be obeyed.

I conclude that, although the obedience of imperatives is in some ways analogous to the truth of indicatives, no very helpful account of *entailment* of imperatives is forthcoming by following out this analogy. It further appears that a proper logic of imperatives must wait upon a proper logic of tenses, as Geach [7] p. 51 hints, precisely because the obeying of an order usually involves changing the truth-value of a sentence. In this connexion, it will be worth our while to take a second look at 'negative' imperatives, orders such as 'do not close the door'. The command 'close the door' presupposes that the door is not *already* closed (a squad given the order 'present arms' when they already have their arms presented would be hard pressed to understand, let alone obey, the order given), and is in effect an instruction to turn the presently false sentence 'the door is closed' into a true one. Hence its obedience conditions are ordinarily undefined for the situation where the door is closed, in something of the way in which the truth-conditions of an ordinary conditional are undefined when its antecedent is false. Following Hofstadter and McKinsey, we cheerfully explained the complement of an imperative

C as that imperative which is obeyed if and only if C is not obeyed; we can now see how much this definition leaves open. In fact, the obedience conditions of 'do not close the door' are as indeterminate as those of 'close the door' — normally the presupposition is the same for both imperatives, that the door is not closed. (The order 'don't present arms' is just as puzzling to our imagined squad as the order 'present arms'.) Thus the order 'don't close the door' is normally equivalent to 'leave the door open'; if 'close the door' has the crude form 'change not-p to p', where p is the proposition that the door is closed, 'do not close the door has the crude form 'change p to p', i.e. make no change with respect to p. This is by no means the end of the complexity of 'negative' imperatives, however. For example 'don't smoke', said to a man *not* smoking, will tend to have the force 'continue not smoking', but, said to a man who *is* smoking, it will tend to have the force 'stop smoking'. It becomes increasingly clear, as we study the examples, that a proper study of imperatives cannot ignore the factor of *change from one state of affairs to another.*

5. *Imperatives and the logic of change* (11). To fix ideas, let us consider a simple change model which takes into account two successive state descriptions; we may write '(A/B)' to mean 'state description A changes to state description B', where A and B are understood to be propositional calculus expressions. For the one-variable case, there are 4 possible 'changes', two of which are intuitively not changes at all: (i) $(\sim p/p)$ (p becomes true); (ii) (p/p) (p stays true); (iii) $(p/\sim p)$ (p becomes false); (iv) $(\sim p/\sim p)$ (p stays false). For n variables there are correspondingly 2^{2n} such changes. There are some obvious laws concerning complex changes, in virtue of which a simple decision-procedure for truth-functional compounds of change-assertions is available. Here are some examples:

(18) $(A/B \,\&\, C) \equiv (A/B) \,\&\, (A/C)$;
(19) $(A \,\&\, B/C) \equiv (A/C) \,\&\, (B/C)$;
(20) $(A/B \vee C) \equiv (A/B) \vee (A/C)$;
(21) $\sim((A/B) \,\&\, (A/\sim B))$;
(22) $\sim((A/B) \,\&\, (\sim A/B))$

I shall not here develop an axiomatic structure for a propositional change logic, but pass directly to an application of this logic to com-

(11) The basic ideas of the logic of change, and its possible application to ordering, were suggested by G. H. von Wright [27], especially chapters II-IV.

mands. The simplest commands concerning a proposition p are the four commands related to (i)-(iv) above. Thus:

(i)	$!(\sim p/p)$	('close the door')
(ii)	$!(p/p)$	('leave the door closed')
(iii)	$!(p/\sim p)$	('open the door')
(iv)	$!(\sim p/\sim p)$	('leave the door open').

If we say now that the command ! (A/B) is obeyed if and only if (A/B), i.e. if and only if the change from A to B takes place, the obedience conditions of these four commands is fully determinate. Thus the command 'close the door' is obeyed if the door *becomes* closed but otherwise disobeyed; in particular, it is automatically disobeyed if the door is *already* closed, whether it becomes open or not — we are here in effect filling 'obedience gaps' which exist in ordinary discourse, much as we fill truth-value gaps in logic where e.g. definite descriptions fail of reference.

The complements of these four commands are now determined by the Hofstadter-McKinsey convention, namely as the commands which are obeyed if and only if the original commands are not obeyed:

(v)	$!\sim(\sim p/p)$	('refrain from closing the door')
(vi)	$!\sim(p/p)$	('refrain from leaving the door open')
(vii)	$!\sim(p/\sim p)$	('refrain from opening the door')
(viii)	$!\sim(\sim p/\sim p)$	('refrain from leaving the door open').

Something of the ambiguity of the negative imperative in ordinary speech (to which Ross [21] drew attention) is now clear. 'Do not smoke', said to a non-smoker, may have the form either of (iv) ('continue not smoking') or of (v) ('refrain from starting to smoke'); these two orders have differents obedience values in case the man is already in fact smoking. 'Do not smoke', said to a smoker, may have the form either of (iii) ('stop smoking') or of (vi) ('refrain from continuing to smoke'), which two again have different obedience values in case the man already does not smoke. It seems to me that full clarity concerning negative imperatives can only be obtained by means of a change-analysis.

Let us define a *simple change-expression* as an expression of the form (A/B) where A and B are truth-functional expressions, and a *change-expression* generally as any truth-functional compound of change-expressions. Then, to develop the logic of imperatives, I propose that we modify conditions (i) and (iii) of the Hofstadter-McKinsey formation-rules as follows: (i)′ if S is a change-expression then !S is an imperative; (iii)′ if S is a change-expression and C an imperative then $(S \rightarrow C)$ is an imperative. Clauses (ii) and (iv) are to remain the same, but we are no longer to imagine that these rules are adjoined to a set of propositional calculus formation-rules; rather the *sentences* of the language are truth-functional compounds of change-expressions or sentences as provided by (iv). Thus '$(p \supset p)$' is not a sentence, though '$((p/{\sim}p) \supset ({\sim}p/p))$' is. Intuitively any sentence refers to two distinct states of affairs, one conceived as following the other, and imperatives concern changes from the one to the other ([12]). We cannot construe '$(p \mathbin{\&} (q/{\sim}q))$' because it is unclear whether p is here asserted simultaneously with q or with ${\sim}q$. In this language, the conditional imperative 'if A is the case now then bring B about' will most naturally be represented as '$((A/A \vee {\sim}A) \rightarrow !({\sim}B/B))$', where the antecedent asserts that A is the case when the order is given whatever may happen at the later time.

If we now adopt the Hofstadter-McKinsey axioms (e.g. (13)-(17)) ([13]) together with suitable axioms for a change logic, we may again establish the triviality metatheorem in the form: if C is an imperative then there is a change-expression S such that $\vdash C = !S$. Further, for any set of imperatives $C_1, \ldots, C_m$ we can form the 'total' imperative C', namely $C_1 \times \ldots \times C_m$, and for this total imperative we have S' such that $\vdash C' = !S'$. Now, bearing in mind that '(A/B)' is equivalent to the conjunction 'A at the first time and B at the later time', we may develop principles of change logic in virtue of which any change-expression S' can be reduced to a 'normal form' of a disjunction of n simple change-expressions $(A_1/B_1) \vee \ldots \vee (A_n/B_n)$, in which each A_i and B_i are either conjunctions containing all propositional variables in S' either negated or non-negated or else contradictions. The procedure is a natural extension of the usual one for obtaining canonical disjunctive normal forms. For S', the corresponding normal form represents in an obvious way the complex change required by the

([12]) Despite this, I hope it is clear that the symbolism is neutral with respect to the Geach-Hare controversy as to whether there could logically be past-tense imperatives; for if there were such their force would presumably be to require a *change in the past*.

([13]) Of course, in them, and from now on, S, S_1', ... are understood to be restricted to change-expressions.

set of imperatives $C_1, \ldots, C_m$. If B_i is a contradiction for all i ($1 \leqslant i \leqslant n$), then we may define the imperatives as *inconsistent,* since they are collectively incapable of completion. If A_i is a contradiction for all i, we may define the imperatives as *incoherent,* since they are collectively incapable of initiation. A set of imperatives cannot be obeyed if it is either incoherent or inconsistent, but it may be neither incoherent nor inconsistent and still not obeyable; thus it may be possible to initiate the set and also possible to complete it, but not to do both. If S′ contains j variables, then there are 2^{2j} distinct simple change-expressions appropriate to its normal form, and if all of them appear we may call the set of imperatives *tautologous* ([14]), since they logically must be obeyed. In general, we can read off the obedience conditions directly from the normal form.

So far so good. We seem to have a clear account at last of the possible meanings of negative imperatives, together with an adequate criterion of consistency, which we can readily extend to cover the case of a mixed set of imperatives and change-expressions. Entailment, however, is as far from our grasp as before. Thus the order $!(\sim p/\sim p)$ ('let there not be a seabattle tomorrow') is inconsistent with $(\sim p/p)$ ('there will be a sea-battle tomorrow'), yet we do not wish to say that $(\sim p/p)$ entails $-!(\sim p/\sim p)$ ('let it not be the case that there continue to be no sea-battle'). I think we need to go back to the contrast made earlier between an imperative's being in force or not in force and an imperative's being obeyed or disobeyed, and recognize that, whilst consistency of imperatives can naturally be defined in terms of obedience, *entailment needs to be defined in terms of what imperatives are in force at a given time.* These two criteria are quite independent: an order may be in force but not obeyed, or be obeyed when it is not in force. A useful start may be made by reconsidering the soundness of our earlier deontic principles when interpreted as concerning orders. Let us provisionally write 'OC' for 'the imperative C is in force'. Then 'O!(A/B)' will mean 'the order to effect change (A/B) is in force'. (A4) in the form:

$$\text{(O4)} \quad O(S \rightarrow C) \supset (O!S \supset OC)$$

seems correct. If one is ordered to do C in case of change S, then if one is ordered to effect S one is implicitly under orders to do C. This in fact is the principle that orders in force at a given time are

([14]) This notion is intended to be a clearer version of 'objective validity' in Ross [21] — a notion of which that author despairs.

closed with respect to detachment. The analogue to (A5) clearly fails: ordered changes do not always take place. We may, however, accept (R2) in the form:

(RO2) from $S \supset T$ to derive $O!S \supset O!T$.

If change S logically involves change T, then if the order to effect S is in force then so implicitly is the order to effect T. With this material we can derive the analogues of (1)-(7) for orders-in-force, as well as more special principles such as:

(23) $O!(A/B \supset C) \supset (O!(A/B) \supset O!(A/C))$;
(24) $O!(A/B \,\&\, C) \equiv O!(A/B) \,\&\, O!(A/C)$;
(25) $O!(A \,\&\, B/C) \equiv O!(A/C) \,\&\, O!(B/C)$;
(26) $O!(A/B \,\&\, {\sim}B) \supset O!(A/C)$.

(26) says roughly that if the orders in force are inconsistent then every order is in force. We have indeed the general principle:

(RO3) from $C_1 = C_2$ to derive $OC_1 \equiv OC_2$;

if orders C_1 and C_2 have demonstrably the same obedience conditions, then C_1 is in force if and only if C_2 is in force.

The analogue to (D5) clearly fails: it may well be the case that there are inconsistent orders in force — this is merely hard luck on those who attempt to obey them: very hard luck in view of (26), since *every* order is then in force. If it be felt that we should add something like:

(O5) $O!(A/B) \supset {\sim}O!(A/{\sim}B)$,

in order to study the logic of consistent orders-in-force, I think an appropriate reply is that there is no need since the consistency of orders is, as we have seen, a matter of their simultaneous obedience and can be studied independently of the O-operator; and, as we shall see, the addition of (O5) sanctifies entailments which I think are invalid.

Thus the logic of orders-in-force is closely analogous to the modified deontic logic we obtained at the end of Section 2. In particular, there are no *theorems* of the form OC: though the order either to smoke or not to smoke cannot but be obeyed, it is not as a matter of logic in force ([15]). And there may be orders in force, i.e. inconsistent ones, which as a matter of logic cannot be obeyed. Hence a sharp contrast is afforded by our system between whether an order is obeyed or not and whether it is in force or not.

It is now easy to define entailment for this simple calculus. Change-epxressions $S_1, \ldots, S_m$ and imperatives $C_1, \ldots, C_n$ entail the imperative C if and only if it is a logical truth (theorem of the system) that

$$S_1 \,\&\, \ldots \,\&\, S_m \,\&\, OC_1 \,\&\, \ldots\, OC_n \supset OC.$$

Roughly, imperative C is entailed if its being in force is entailed by the indicatives together with the being in force of the imperatives.

With qualifications to follow, this account validates many of the inferences canvassed in the literature. Let us consider two examples. The argument:

It is raining;
if it is raining, close the door;
therefore close the door

seems to be valid. And of course it *is* valid if we represent the second premiss as $(p/p) \supset O!(q/{\sim}q)$ and the conclusion as $O!(q/{\sim}q)$. It will *not* be valid, however if we represent the second premiss as $O((p/p) \rightarrow !(q/{\sim}q)$. Now this order is equivalent to $!((p/{\sim}p) \vee ({\sim}p/p) \vee ({\sim}p/{\sim}p) \vee (q/{\sim}q))$ and is obeyed by e.g. stopping the rain; what is true is that, given that it is raining now and the rain goes on ((p/p)), then if the second order is obeyed the last order $!(q/{\sim}q)$ must have been obeyed. But the conditional order's being in force by no means entails that the conclusion is in force — from 'either stop the rain or ... or close the door' 'close the door' does not follow, even if it does go on raining. This perhaps suggests that the Hofstadter-McKinsey analysis of conditional imperatives is unhelpful for showing their correct entailment relations. It also, I think, shows that the schema:

$$(27) \quad S \,\&\, O(S \rightarrow C) \supset OC$$

(compare (O4)) should not be accepted, despite its *prima facie* plau-

([15]) Though if *any* order is in force them this is, by the theorem $OC \supset O\,!(A \vee {\sim}A/B \vee {\sim}B)$.

sibility [16].

Consider also the argument:

> If he comes, leave the files open;
> do not leave the files open;
> therefore he does not come,

which Castaneda [4] accepts as sound and as a counterexample to Hare's thesis in [8], p. 28, that no indicative conclusion can be validly drawn from a set of premisses which cannot be validly drawn from the indicatives among them alone. It is hard to be sure here what is meant by the premisses, but if the argument has the form:

$$(\sim p/p) \supset O!(q/q);$$
$$O!(q/\sim q)$$
$$\text{therefore } \sim(\sim p/p),$$

then the argument is not sound unless something like (O5) is added. It seems to me that the argument is *not* sound: if he does come, then, given the premisses, inconsistent orders are in force, but this is not a contradiction from which we can deduce that he does not come. Castenada argues that, if the boss gives the two orders to his secretary, she may *infer* that the other person will not come; rather surely she may *hope* that he will not come (because if he does then she certainly cannot obey her boss's orders) — or she may conclude from her knowledge of her boss that he does not *think* the other person will come. The same objection applies to Geach's two examples in [6], p. 57, which have the same form. In view of this, Castaneda's principle, quoted at the beginning of Section 4, has to be abandoned, and I do not think that a convincing counterexample to Hare's thesis has been produced.

6. *Some further doubts.* I think the above account of both consistency and entailment is on the right lines, and that the distinction between obedience conditions on the one hand, in terms of which consistency can be defined, and being in force on the other, in terms of which entailment can be defined, helps to clarify confusions in the logic

(16) Actually, *via* the paradox $O!\sim S \supset O(S \rightarrow C)$, we could prove from (27) the very implausible $S \,\&\, O!\sim S \supset OC$ (whenever an order is in force but unobeyed then any order is in force). The analogous $O!S \,\&\, (S' \supset OC) \supset OC$ is also unsound for similar reasons: the antecedent is true if S is false though the order to achieve S is in force, whether C is in force or not.

of imperatives. But when we attempts to extend the account to more complex imperative inferences containing quantifiers and identity, grave problems arise which are related directly or indirectly to the possible *referential opacity* of imperatives, as was suggested earlier.

Consider first one of Hare's examples of a valid inference:

> Take all the boxes to the station;
> this is one of the boxes;
> therefore take this to the station.

If we let 'Bx' mean 'x is a box', 'Sx' mean 'x is taken to the station', and 'a' 'this', then we may represent, by an obvious quantificational extension of our earlier notation, the argument thus:

$$(x)\ ((Bx/Bx)\ \supset\ O!(\sim Sx/Sx))$$
$$(Ba/Ba)$$
$$\text{therefore } O!(\sim Sa/Sa),$$

which is sound in virtue of quantifier rules. If however we represent the first premiss as $O!(x)\ ((Bx/Bx)\ \supset\ (\sim Sx/Sx))$, validity can no longer be shown. Nor is it in fact clear that we really wish to regard the argument as sound on reflexion: suppose I order all the boxes to be taken to the station, thinking that this is not one of the boxes but, say, my piano, when in fact it is one of the boxes. I did not intend this to be included in the order: have I in fact ordered this to be taken ? Of course my order will only be fully obeyed if this is taken to the station, so that obedience to the conclusion follows from obedience to the premisses given that this is one of the boxes; but it remains unclear whether, thinking what I did, that was what I ordered.

There are similar problems concerning identity. If George IV ordered Scott to come to dinner, not knowing that Scott was the author of *Waverley*, did he in fact order the author of *Waverley* to come to dinner ? If I am the orderly officer, I may, on relinquishing my duties, order you to be orderly officer, but I do not thereby order you to be me, thouh I do perhaps order you to become what I was. And there is the Krushchev-Gagarin case mentioned earlier. Some, but not all, of these problems can be resolved by using change-expressions and so paying due attention to the times at which different identities are true or false. But the problem of quantifying into 0-

contexts and identity-substitution in them is beyond the scope of this paper ([17]).

Another respect in which our present notation and ideas are deficient may be brought out as follows. Suppose I order Jones to close the door and simultaneously order Smith not to close it, i.e. to leave it open. Then my orders to Jones are consistent and so are my orders to Smith; but it is not clear whether they are jointly consistent or not. Certainly if Jones closes the door while Smith goes on reading, then both have obeyed the orders I gave them. Yet my two orders suggest that I want the door both closed and not closed, a logically impossible requirement. We surely have to say that the orders are consistent and that the appearance of inconsistency is generated, partly by the fact that I seem to want impossible things (though I may well in fact precisely want the door closed by Jones but not by Smith), but mainly by the fact that if the orders were given to the same person then they would be inconsistent. If we are to study the consistency of *directives* as opposed to fiats, then we need to be able to mark in our symbolism the person to whom the order is given; and clearly we should at the same time mark the orderer. We need at least 'Oxy!(A/B)' with the sense 'x orders y to effect the change (A/B)', in which characteristically 'y' will be a free variable in A and B. Then principles such as (O4), (RO2) and (RO3) will carry over by obvious extension to the richer symbolism, and we shall be able to study more clearly the movement of quantifiers across O-contexts.

Again, the restriction to changes between only two state-descriptions becomes intolerable if we are to study the consistency even of, let alone entailment relations between, sets of orders which refer to a succession of different times: 'do A; then if B do C; then if D do E', etc. We shall also wish to allow orders to appear within orders; thus a principle which we so far have not the means to state is that if x_1 orders x_2 to order x_3 to ... order x_n to effect (A/B) then x_1 has implicitly ordered x_n to effect (A/B) — what might be called the chain of command principle. A related inference is mentioned by Ross [21]: from your parents' order to do as the teacher tells you and the teacher's order to prepare your lessons there follows, as an order from your *parents*, that the lessons are to be prepared. This however has the form:

([17]) The critical case is this: suppose a = b but not necessarily so, and I order a to *stay* b (O!(a = b/a =b)); then by identity substitution O!(a = a/a = a) — yet this was not what I ordered, or even part of what I ordered.

$$Oxy(OzyC \rightarrow C),\ OzyC \therefore OxyC$$

and will be justified by a suitable extension of (O4), provided we allow the first premiss to be well-formed. Ideally, to handle successive states of affairs as well as nesting of O-operators, we need three time parameters on each simple order—the time at which the order comes into force or is given, the time of the initial state of affairs for the ordered change, and the time at which the change is supposed to be completed. Thus we might write

$$O_{t1}\,xy!\ (A_{t_2}/B_{t_3})$$

to mean 'x orders y at t_1 to effect the change from A at t_2 to B at t_3'. I suggest that ordering, fully studied in the framework of a logic of tenses, is at least a heptadic relation, between two people, two states of affairs, and three times. Another heptadic relation, with, so far as I can see, identical formal properties, is requesting.

7. *Conclusion.* Are ethical judgments like imperatives ? There can be no doubt, I think, that the formal properties of 'ought' and the formal properties of 'it is ordered that' are very similar, as the preceding discussions have separately shown. We have only to compare (A4) and (O4), (R2) and (RO2), and the failure of both (D5) and (O5) for which I have argued, to see this. It is even, I think, arguable that deontic logic, insofar as it purports to relate to human action, would be richer and better if based on the calculus of change as I have attempted to base the logic of ordering; then the formal similarities would be even more striking. Further, there is a case for supposing that being under an obligation is also at least heptadic — 'x is under an obligation to y at time t_1 to effect the change from A at t_2 to B at t_3'. [18]

I think it would be wrong to base upon any such formal similarities, either detected or detectable, Hare-like conclusions of the form that both kinds of language are *prescriptive,* or even to explain the similarities in this way. For example, the relation of being taller than and that of being older than have identical formal properties: both are transitive, irreflexive, but not connected in the class of human beings, so that they impose a strict partial ordering on this class, but there is no deep affinity that I can see between height and age. The formal similarities that we are here noting are based on (A4) and (R2); and these similarities will be shared by any operator which selects from a class of propositions a subclass which is closed with

[18] It is perhaps worth remarking that the chain of command principle for obligation is counterintuitive.

respect to detachment and logical consequence (in the sense that the logical consequences of members of the subclass are in the subclass unless it is empty) but which is not assumed to be consistent (failure of (D5)) or true (failure of (A5)). Hence they are shared by the operator 'it is a theorem of T that...' for any logical theory T, for example, yet there is no clear sense in which this operator is prescriptive. They are also perhaps shared by 'I assert that ...' and 'I promise that ...', to mention two more 'performative' verbs. If anything, these similarities merely reflect the dull fact that a system of orders-in-force or a system of 'ought'-stamements resembles structurally a logical theory. Hence I find no support for Hare's first thesis in these formal considerations.

Do ethical judgments ever entail imperatives ? Again, it cannot be a consequence of our work that they do, simply because we have studied the logic of each separately. We can certainly *make* 'ought'-statements entail imperatives by adding to a composite logic of both the schema:

(H) $\Box S \supset O!S$

(where '$\Box$' is to be read 'it ought to be the case that ... come about'), whereby whenever it ought to be the case that S the order to effect the change S is in force. Now I have already indicated earlier what I think to be part of the source of the philosophical opinion that 'ought'-statements entail imperatives: just as the rules of chess may be such that in your present position, unless you move your queen you will lose, so your moral principles may be such that in your present position, unless you do A, some moral catastrophe may ensue. It is natural in the first case to say that you *have* to move your queen, at least if you want not to lose, and in the second case that you have to or ought to do A, at least if you want to stick with morality. And these steps, though informal, may be permitted. But that *then* follows the command 'move your queen' or 'do A' seems to me highly suspect. It is the case, the situations being what they are, that unless these commands are obeyed you will lose or be engulfed in disaster. But I cannot see that these results, even coupled with a healthy desire to avoid defeat or calamity, in any recognizable sense *entail* the respective orders, though they do make obedience to them a rational course. I find it only too easy to imagine that a man may recognize in the fullest sense as true all that we tell him concerning his position, that he admits that he ought to move his queen, that he dearly wishes to win, and yet that he does not move his queen and so does not 'accept' the order to do so. If any moral philosopher is inclined to doubt this, I can only invite him to take another, and closer, look at his fellow-men. Though such a man is absurd, I think he cannot be legislated out of existence by a logical trick such as the adoption of (H).

I suspect that my main claim here, as in the case of 'ought' implies 'can', is that any deontic logic or logic of imperatives should be neutral with respect to irrational behaviour. But I hope also that enough has been said about the entailment of imperatives to enable the various theses of Hare and others concerning this matter to be evaluated a little more clearly. In particular, I should like to conclude by saying a further word concerning the two special principles of imperative reasoning enunciated by Hare [8], p. 28, attacked by Geach and Castaneda, and discussed *passim* in this paper. For I think that an informal justification can be given for both of them in terms of the earlier discussion, provided that it be conceded that *in any possible state of affairs any orders may be in force or none*. Granted this, it follows that

$$S_1 \mathbin{\&} \ldots \mathbin{\&} S_m \supset OC$$

cannot be an logical truth if $S_1, \ldots, S_m$ are jointly consistent. For then $S_1 \mathbin{\&} \ldots \mathbin{\&} S_m$ describes a possible state of affairs which must remain consistent upon the addition of $\sim OC$. Thus (a) no imperative is entailed by a set of indicatives (provided they are consistent). Secondly, suppose that

$$S_1 \mathbin{\&} \ldots \mathbin{\&} S_m \mathbin{\&} OC_1 \mathbin{\&} \ldots OC_n \supset S$$

is a logical truth. Then so is

$$S_1 \mathbin{\&} \ldots \mathbin{\&} S_m \mathbin{\&} \sim S \supset \sim (OC_1 \mathbin{\&} \ldots \mathbin{\&} OC_n).$$

Therefore $S_1, \ldots, S_m, \sim S$ must be inconsistent: for if they were consistent it would follow that at least one order was not in force in the state of affairs they jointly describe. Therefore

$$S_1 \mathbin{\&} \ldots \mathbin{\&} S_m \supset S$$

is a logical truth. Thus (b) no indicative can be validly drawn from a set of premisses which cannot be validly drawn from the indicatives among them alone.

I find it pleasant to end on a note of such accord with Mr Hare, though I fear he may have qualms about its foundation.

Claremont, California. E. J. LEMMON

BIBLIOGRAPHY

(1) A. R. ANDERSON, *The formal analysis of normative concepts,* Technical Report No. 2, U.S.O.N.R., 1956.

(2) A. R. ANDERSON, *On the logic of «commitment», Phil. Studies,* vol. 10 (1959), pp. 23-27.

(3) A. R. ANSERSON, *Reply to Mr. Rescher, Phil. Studies,* vol. 13 (1962), pp. 6-8.

(4) H. N. CASTANEDA, *Imperative reasonings, Philosophy and Phenomenological Research,* vol. 21 (1960), pp. 21-49.

(5) M. FISHER, *A system of deontic-alethic modal logic, Mind,* vol. 71 (1962), pp. 231-236.

(6) P. T. GEACH, *Imperatives and deontic logic, Analysis,* vol. 18 (1958), pp. 49-56.

(7) R. M. HARE, *Imperative sentences, Mind,* vol. 58 (1949), pp. 21-39.

(8) R. M. HARE, *The language of morals,* Oxford 1952.

(9) R. M. HARE, *Freedom and reason,* Oxford 1963.

(10) K. J. J. HINTIKKA, *Quantifiers in deontic logic,* Helsinki 1957.

(11) A. HOFSTADTER and J. C. C. MCKINSEY, *On the logic of imperatives, Philosophy of Science,* Vol. 6 (1939), pp. 446-457.

(12) E. J. LEMMON, *New foundations for Lewis modal systems, Journ. Symb. Log.,* vol. 22 (1957), pp. 176-186.

(13) E. J. LEMMON, *Is there only one correct system of modal logic ?, Proc. Ar. Soc.* Vol. 33 (1959), pp. 23-40.

(14) E. J. LEMMON, *Moral dilemmas, Phil. Review,* vol. 71 (1962), pp. 139-158.

(15) P. H. NOWELL-SMITH and E. J. LEMMON, *Escapism: the logical basis of ethics, Mind,* vol. 69 (1960), pp. 289-300.

(16) A. N. PRIOR, *The paradoxes of derived obligation, Mind,* vol. 63 (1954), pp. 64-65.

(17) A. N. PRIOR, *Formal Logic,* 2nd ed., Oxford 1962.

(18) A. N. PRIOR, *Time and modality,* Oxford 1957.

(19) N. RESCHER, *An axiom system for deontic logic, Phil. Studies,* Vol. 9 (1958), pp. 24-30.

(20) N. RESCHER,*Conditional permission in deontic logic, Phil. Studies,* Vol. 13 (1962), pp. 1-6.

(21) A. Ross, *Imperatives and logic*, *Philosophy of Science*, Vol. 11 (1944), pp. 30-46.
(22) T. J. Smiley, *The logical basis of ethics*, *Acta Phil. Fennica*, Fasc. 16 (1963), pp. 237-246.
(23) P. F. Strawson, *On referring*, *Mind*, vol. 59 (1950), pp. 320-344.
(24) G. H. von Wright, *Deontic logic*, *Mind*, vol. 60 (1951), pp. 1-15.
(25) G. H. von Wright, *An essay in modal logic*, Amsterdam 1951.
(26) G. H. von Wright, *A note on deontic logic and derived obligation*, *Mind*, vol. 65 (1956), pp. 507-509.
(27) G. H. von Wright, *Norm and action*, New York, 1963.