A LOGICAL THEORY OF COMMANDING

MARK FISHER

University of Khartoum

1. In many well-known systems of modal logic there occur theorems whose interpretation is a sentence in which a modal expression is re-iterated. Some systems contain reduction theorems; that is, they have a certain expression with a number of iterations equivalent to another expression with a smaller number, or none. One wellknown system, the Lewis system S5, allows iterations to be eliminated. The problem of interperting iterated modal operators in a deontic manner has received little attention. Some writers (1) consider that a system, to receive a proper deontic interpretation, must contain no well-formed iterations. The problem of interpreting iterated modal expressions in an alethic manner has been discussed along with the wider problem of what precise sense to give to the alethic expressions «necessarily» and «possibly». For instance, what precise meaning can be attached to an interpretation of *LLp* as «It is necessary that it is necessary that *p*» can only be decided when it has first been decided what exact meaning can be attached to *Lp* interpreted as «It is necessary that *p*». One very common interpretation of iterated modal operators treats *LLp,* for instance, as a statement about *Lp*. It has been this interpretation that those writers have had in mind, I think, who have said that there is no deontic interpretation available for *LLp*. As far as I know, only one writer on modal logic has offered a deontic interpretation for iterated modal operators. O. Becker (2) offers the following two interpretations of *LLp* (and similar ones of other formulae containing iterations). Under alethic interpretation *Lp* means «p is true of every member of a certain class», and *LLp*

(1) G. H. Von Wright, *An Essay in Modal Logic,* North-Holland Publishing Co., New Amsterdam, 1951.

A. R. Anderson, *The Formal Analysis of Normative Systems,* Yale Interaction Laboratory, New Haven, Connecticut, 1956; in ch. 8 there is a discussion of various deontic analogues of the Lewis systems, and reduction theorems are proved.

E. E. Dawson, personal communication, takes the same view as Von Wright of the undesirability of iterated deontic operators.

(2) O. Becker, *Untersuchungen über den Modalkalkül.*

means «p is true of every member of every class of a certain class of classes». Under deontic interpretations these two formulae mean «It is commanded that *p*» and «It is commanded that it be commanded that *p*», i.e. a higher authority commands a lower authority to command that *p*. Von Wright, on whose review (3) of Becker's work I rely, has pointed out a number of difficulties in his methods. Here I am only concerned with his method for deontic interpretation. From this point of view the chief weakness in his system, under interpretation, is his axiom *CLpp*. Becker, aware of the difficulties, suggests thas substituents for variables be restricted to expressions denoting *legitimate* acts; but as Von Wright points out this is equivalent to the simpler procedure of replacing *CLpp* by *CLpNLNp*, i.e. by *CLpMp*. Further improvements could easily be suggested, for Becker's system is equivalent to the Lewis system S2, one of the less satisfactory of the well-known systems from a deontic point of view. The strength of Becker's work lies in his suggestion of a deontic interpretation of iterated modal operators. I shall adopt this suggestion, and hope to improve upon it by making the notation more explicit and basing it on a stronger modal system.

2. Another writer who has contributed to the subject I am discussing is J. Łos. I rely on reviews of his papers by R. Suszko and H. Hiz (4). His work has also been discussed, in a different context, by Prior (5). Łos presents a system which contains propositional calculus, quantifiers, individual variables with range restricted to persons or other language-users, and the primitive symbol *L*. *Lxp*, where *x* is an individual variable and *p* a propositional variable, is well-formed, and so is any addition of *Lx* to the left of a wff. *Lxp* has as intended interpretation «*x* asserts that *p*»; Prior suggests that this should be modified to «*x* asserts that *p*, or to be consistent ought to».

In section 7 below I develop a similar system with the same intended interpretation but different and I think more satisfactory axioms and rules. I now adopt the essential feature of Łos's sym-

(3) G. H. Von Wright, review of Becker, *op.cit.*, *Mind* vol. 60, 1951, pp. 557-561.

(4) R. Suszko, review of article in Polish by J. Łos, *Journal of Symbolic Logic* vol. 14, 1949, pp. 64-65.

A. N. Prior, *Formal Logic*, Clarendon Press, Oxford, 1955; Łos's axioms as reported by Suszko are given on p. 313.

(5) A. N. Prior, *Time and Modality*, Clarendon Press, Oxford, 1957; Los's method is discussed on pp. 121-122.

bolism in order to construct a system with a deontic intended interpretation. I write Oxp and interpret it as «x commands that p».

In what follows the propositional variables, under interpretation, have a range of substituents which are not, as is usual, either indicative or imperative sentences, but what may be called gerundives, or in the well-known terminology introduced by Hare [6], phrastics. Thus, if for x in Oxp is substituted the first person pronoun «I» we obtain, when a *phastric* is substituted for p, an imperative sentence. If any other individual name or uniquely referring expression (referring to a person) is substituted we obtain an indicative sentence reporting a command. To avoid misunderstanding I emphasise that the many differences between these two types are in this paper being deliberately ignored.

3. I now define the system Ox. Under its intended interpretation I shall refer to it as the logic of commanding.

Primitive symbols: C, N, O, Π

Variables: $p, q, r, p_1, q_1, r_1, p_2, \ldots$ called propositional variables
$x, y, z, x_1, y_1, z_2, x_2, \ldots$ called individual variables
$\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{A}_1, \mathbf{B}_1, \mathbf{C}_1, \mathbf{A}_2, \ldots$ called arbitrary expressions

Constants: $a, b, c, a_1, b_1, c_1, a_2, \ldots$ called individual constants
$\mathrm{F, G, H, F_1, G_1, H_1, F_2}, \ldots$ called predicate constants

Rules of formation:

F.1. $Ox\mathbf{A}$ is a wff

F.2. If $\mathbf{A}$ and $\mathbf{B}$ are wffs, then $N\mathbf{A}$, $C\mathbf{AB}$, and $\Pi x\mathbf{A}$ are wffs

F.3. $\mathbf{A}$ is wff if and only if its being so follows from a finite number of applications of F.1. and F.2.

Definitions:

D.1. $K\mathbf{AB}$ abbreviates $NC\mathbf{A}N\mathbf{B}$

D.2. $A\mathbf{AB}$ abbreviates $CN\mathbf{AB}$

D.3. $E\mathbf{AB}$ abbreviates $KC\mathbf{AB}C\mathbf{BA}$

D.4. $Px\mathbf{A}$ abbreviates $NOxN\mathbf{A}$

D.5. $\Sigma x\mathbf{A}$ abbreviates $N\Pi xN\mathbf{A}$

Rules of inference: In stating the rules of inference I write «$\vdash \mathbf{A}$» to mean «$\mathbf{A}$ is a theorem of Ox», and «$\mathbf{A} \rightarrow \mathbf{B}$» to mean «Infer $\mathbf{B}$ from $\mathbf{A}$». $\mathbf{A}_x$ is an arbitrary expression containing x as a free variable. A theorem is the last line of a proof, which is a finite se-

(6) R. M. Hare, *The Language of Morals*, Clarendon Press, Oxford, 1952,

quence of lines each of which either is an axiom or has been inferred from one or more previous lines using one or more of the rules of inference.

R.1. $\vdash$ **CAB** and $\vdash$ **A** $\rightarrow$ $\vdash$ **B**
R.2. $\vdash$ **A** $\rightarrow$ **B**, where **A** contains one or more occurrences of p, and **B** is the result of replacing each occurrence of p in **A** by **C**
R.3. $\vdash \mathbf{A}_x \rightarrow \vdash \mathbf{A}_y$, provided that there is no free occurrence of x in a wf part of **A** of the form Π*y***B**
R.4. $\vdash$ **CAB** $\rightarrow$ CΠ*x***AB**
R.5. $\vdash$ **CAB** $\rightarrow$ $\vdash$ **C**Π*x***B**, if there is no free occurrence of x in **A**
R.6. If **A** can be proved from R.1, R.2, and $CCCCCpqCNrNp_1rq_1CCq_1pCp_1p$, then $\vdash$ **A**
R.7. $\vdash$ **A** $\rightarrow$ $\vdash$ *Ox***A**

Axioms:
(100) C*Ox***A***NOxN***A**
(200) C*Ox*C**AB**C*Ox***A***Ox***B**
(300) CΠ*xOy***A***Oy*Π*x***A**

Readers familiar with modal logic will easily recognise Ox as a version of the system M of Feys-Von Wright. It is weaker than M in having (100) and R.4. in place of the usual alethic axiom CL**AA**. Some remarks will now be made about the reasons for adopting these axioms and rules.

R.6. allows all theorems of classical propositional calculus to be proved. The method of adopting a «propositional calculus base» is now familiar in modal logic.

R.7. is adopted because, as shown by Hintikka [7], it can be derived in any system which is as strong as Von Wright's original system and contains the theorem that not everything is permitted, i.e. rejection of *Px***A**. I have not found a rejection-method for the system Ox, but I conjecture that *Px***A** is not a theorem of it, as seems desirable, and therefore adopt R.7. provisionally, in the hope that if a rejection method is found it will be possible to show the non-independence of R.7.

The other rules require no comment.

(200) is common to a well-known modal systems.

[7] Von Wright, *op.cit.*, pp. 38-39; a proof is given in detail by Prior, *Formal Logic*, pp. 222-223.

(100) is the deontic form of the other common modal axiom.

A set of systems of increasing strength can be constructed using these two axioms, R.1., R.2., and various additional rules and axioms. Lemmon [8] gives a set of deontic systems using the same procedure and the deontic (100), along with a rule that ⊢ **CAB** → ⊢ C**O**A**O**B in place of R.7., and a special rule ⊢ CO**AA**, if **A** is fully modalised, i.e. either is O**A** or results from O**A** by substitution and prefixing of O. This rule is designed to preserve the «reduction theorems» which are available in the stronger systems, i.e. to allow the number of non-equivalent modalities to be finite in them. An alternative method of developing deontic analogues to these modal systems is developed by Anderson [9], in which also the reduction theorems are available. It is interesting to note that these reduction theorems are not available in Ox strengthened with a form of the typical S4 axiom, but are available in Ox if it is strengthened with a form of the typical S5 axiom. These assertions will be proved later.

(300) is a version of an axiom introduced by Mrs Barcan Marcus [10], who was unable to prove it in a system S2 with quantifiers. Prior [11] later showed that in quantified S5 it is provable. It is interesting to note that this theorem is not available in Ox strengthened with a form of the S5 axiom, although a theorem rather similar to it is available. It should be understood that an expression like $Ox\mathbf{A}_y$ is to be interpreted as either «x commands y to do **A**» or «*x* commands that $\mathbf{A}_y$ (be the case)». That is, we do not at present distinguish between variables whose range is the set of expressions referring to agents, commanders, in general persons and other language-users and those whose range is the set of expressions referring to occasions, situations, etc. It will be possible to do so, using a two-sorted functional calculus. This development will not be studied in this paper.

Theorems: The following diagrams summarise a number of theorems which can easily be proved. An arrow between two expres-

(8) E. J. Lemmon, «New Foundations for Lewis Modal systems», *Journal of Symbolic Logic*, vol. 22, 1957, 176-186; deontic systems discussed on pp. 184-186.

(9) A. R. Anderson, *op.cit.*, ch. 8.

(10) R. Barcan (now Marcus), «A functional calculus of first order based on strict implication», *Journal of Symbolic Logic* vol. 11, 1946, pp. 1-16.

(11) A. N. Prior, «Modality and Quantification in S5», *Journal of Symbolic Logic*, vol. 21, 1956, pp. 60-62.

sions **A** and **B**, thus $\mathbf{A} \rightarrow \mathbf{B}$, shows that $C\mathbf{AB}$ is a theorem. The sign $\equiv$ between two expressions, thus $\mathbf{A} \equiv \mathbf{B}$, shows that $E\mathbf{AB}$ is a theorem. The first diagram is for reference purposes, and shows familiar relations between doubly quantified expressions. The second shows that Ox behaves like Πx to a large extent; the third shows that it is not entirely analogous.

I

$$\begin{array}{ccc} \Pi x \Pi y \mathbf{A}_{x,y} & \equiv & \Pi y \Pi x \mathbf{A}_{x,y} \\ \downarrow & & \downarrow \\ \Sigma x \Pi y \mathbf{A}_{x,y} & & \Sigma y \Pi x \mathbf{A}_{x,y} \\ \downarrow & & \downarrow \\ \Pi y \Sigma x \mathbf{A}_{x,y} & & \Pi x \Sigma y \mathbf{A}_{x,y} \\ \downarrow & & \downarrow \\ \Sigma y \Sigma x \mathbf{A}_{x,y} & \equiv & \Sigma x \Sigma y \mathbf{A}_{x,y} \end{array}$$

II

$$\begin{array}{ccc} \Pi x O y \mathbf{A}_{x} & \equiv & O y \Pi x \mathbf{A}_{x} \\ \downarrow & & \downarrow \\ \Sigma x O y \mathbf{A}_{x} & & P y \Pi x A_{x} \\ \downarrow & & \downarrow \\ O y \Sigma x \mathbf{A}_{x} & & \Pi x P y \mathbf{A}_{x} \\ \downarrow & & \downarrow \\ P y \Sigma x \mathbf{A}_{x} & \equiv & \Sigma x P y \mathbf{A}_{x} \end{array}$$

III

$$\begin{array}{ccc} \Pi x O x \mathbf{A} & \rightarrow & O x \Pi x \mathbf{A} \\ \downarrow & & \downarrow \\ \Sigma x O x \mathbf{A} & & P x \Pi x \mathbf{A} \end{array} \qquad \begin{array}{ccc} O x \Sigma x \mathbf{A} & & \Pi x P x \mathbf{A} \\ \downarrow & & \downarrow \\ P x \Sigma x \mathbf{A} & \rightarrow & \Sigma x P x \mathbf{A} \end{array}$$

It is now possible to consider a number of special axioms which might be added to Ox, which will give it a distinctive character. The possibilities are listed here.

(400) $COxOy\mathbf{A}Ox\mathbf{A}$
(401) $COxOy\mathbf{A}Oy\mathbf{A}$
(402) $COx\mathbf{A}OxOy\mathbf{A}$
(403) $COx\mathbf{A}OyOx\mathbf{A}$
(500) $CNOx\mathbf{A}OxNOy\mathbf{A}$
(501) $CNOx\mathbf{A}OyNOx\mathbf{A}$

It can be seen that (402) and (403) are versions of the special S4 axiom *CLpLLp*, while (500) and (501) are versions of the special S5 axiom *CNLpLNLp*. Some formal consequences of adding various combinations of these will be investigated; then the question of which are logical truths under interpretation will be considered.

In alethic systems as strong as or stronger than M, *CLLpLp*, of which (400) and (401) are versions, is provable. These axioms are not provable in Ox, however.

If (400) is added to Ox, then the following theorem is provable:

(400.1) $COx_1Ox_2...Ox_n\mathbf{A}Ox_1Ox_2...Ox_{n-m}\mathbf{A}$, for $n>m\geqslant 1$.

The usual proof that M has infinitely many non-equivalent modalities can then be carried through. This proof will now be sketched.

Consider the matrix whose values are infinite sequences of the digits 1 and 0 in any order, a non-denumerably infinite number of non-identical sequences each containing a denumerably infinite sequence of digits. The designated sequence or value is $(\dot{1})$. The finite matrix for propositional calculus is adapted thus: if $x, y, ...$ are digits in a sequence, then $N(x, y, ...) = (Nx, Ny, ...)$; $C(x, y, ...) (x_1, y_1, ...) = (Cxx_1, Cyy_1)$; and so on for other connectives. Let a be a sequence having an initial part consisting of n l's; then Oxa is the sequence consisting of n—1 l's followed by O but otherwise identical with a. It is easy to prove that this matrix is satisfied by Ox; in fact it does not distinguish Ox from M. Now consider two separate modalities, $Ox...Ox_n$ and $Ox_1...Ox_m$, for $m>n\geqslant 1$. Let the value of $\mathbf{A}$ be the sequence consisting of m l's followed by O. Then the value of $Ox\mathbf{A}$ is the sequence consisting of m—1 l's followed by O; hence the value of $Ox_1...Ox_m\mathbf{A}$ is the sequence O. Now the value of $Ox_1...Ox_n\mathbf{A}$ is the sequence consisting of m—n l's followed by O. Hence, since $m>n$, the value of $Ox_1...Ox_n\mathbf{A}$ must contain l at its initial place at least. Hence the value of $COx_1...Ox_n\mathbf{A}Ox_1...Ox_m\mathbf{A}$ has C1O in its initial place at least and therefore is undesignated. Hence $Ox_1...Ox_n$ and $Ox_1...Ox_m$ are not equivalent.

A similar proof can be given if (401) is added to Ox in place of (400). It is interesting to note that the proof fails if either (402) is added as well as (400), or (403) is added as well as (401); in these cases the two modalities considered in the proof are equivalent.

In these two systems, Ox with (400) and (402) or with (401) and (403), the normal reduction theorems cannot be proved in the usual way, although I have not found any proof that they cannot be proved at all. Two different types of what I may call quasi-reduction theorems can however be proved: I now proceed to these proofs.

If we restrict the added axioms by using only one individual variable, then the system is equivalent to S4 with deontic modification. In it the reduction theorems are once more available in the usual way. Anderson ([12]) gives a proof of this fact, and it can also be proved in the present system without using his special definitions of operators. Thus

([12]) A. R. Anderson, *op.cit.*, ch. 8.

(400.2) $EOxPxOxPx\mathbf{A}OxPx\mathbf{A}$ and
(400.3) $EPxOxPxOx\mathbf{A}PxOx\mathbf{A}$

are theorems of Ox with (400) and (402), since the restricted axiom follow by substitution from (400) and (402). Similarly, if instead of (500) we take a restricted version of it with only a single individual variable, then the S5 reduction theorems can be proved; i.e. (400.2), (400.3) and

(500.1) $EPxOx\mathbf{A}Ox\mathbf{A}$ and
(500.2) $EOxPxOx\mathbf{A}PxOx\mathbf{A}$

are all theorems of Ox with (500). So incidentally is

(500.3) $EOxOx\mathbf{A}Ox\mathbf{A}$.

A more general type of reduction theorem can be proved in Ox with (400) and (402). If we try to carry through Parry's proofs [13] of the S4 reduction theorems we find that proof of each conditional can be obtained, but the two converses turn out not to yield the biconditional. We have

(400.4) $COxPy\mathbf{A}OxPzOx_1Py\mathbf{A}$

Proof	1. $COzPy\mathbf{A}OzPy\mathbf{A}$	(R.6)
	2. $COz\mathbf{A}OzOx_1\mathbf{A}$	((402))
	3. $COzPy\mathbf{A}OzOx_1Py\mathbf{A}$	(2.)
	4. $COzPy\mathbf{A}PzOx_1Py\mathbf{A}$	((100), R.6.)
	5. $COxOzPy\mathbf{A}OxPzOx_1Py\mathbf{A}$	((200))
	6. $COxPy\mathbf{A}OxPzOx_1Py\mathbf{A}$	((400), (402), R.6.)

To prove our version of the converse we need the following derived rule:

R.D.I. $\vdash \mathbf{CAB} \rightarrow \vdash CPx\mathbf{A}Px\mathbf{B}$

It is easy to prove this rule from R.7.; proof will be omitted.

(400.5) $COxPyOzPx_1\mathbf{A}OxPy\mathbf{A}$

Proof.	1. $COz\mathbf{A}OzOx_1\mathbf{A}$	((402)
	2. $CPzPx_1\mathbf{A}Pz\mathbf{A}$	(1., Df.*P*, R.6.)
	3. $COzPx_1\mathbf{A}PzPx_1\mathbf{A}$	((100), df.*P*)
	4. $COzPx_1\mathbf{A}Pz\mathbf{A}$	(2., 3., R.1.)
	5. $CPyOzPx_1\mathbf{A}PyPz\mathbf{A}$	(4., R.D.1.)
	6. $CPyOzPx_1\mathbf{A}Py\mathbf{A}$	((400), (402), 5.)
	7. $COxPyOzPx_1\mathbf{A}OxPy\mathbf{A}$	(6., (200))

(13) W. T. Parry, «Modalities in the *Survey* system of strict implication», *Journal of Symbolic Logic,* vol. 4, 1939, pp. 137-154.

Conjunction of these theorems does not give the equivalence which would be analogous to the S4 reduction theorem. Thus (400.4) and (400.5) are quasi-reduction theorems in this system. The corresponding theorems available in Ox with (401) and (403) are:

(401.1) $COxPy\mathbf{A}Ox_1PzOxPy\mathbf{A}$ and
(401.2) $COxPx_1OzPy\mathbf{A}OxPy\mathbf{A}$

I have not been able to prove the S4 reduction theorems in Ox with (400)and (402), and I leave open the question of whether they are provable or not.

If we add (500) alone to Ox reduction theorems are not (at least, not likely to be) provable, because a version of *CLpp* is now lacking. Curiously enough, if we add (500) and (401) to Ox all S5 reduction theorems are provable, but if we substitute (400) for (401) here they do not seem to be; at least, the usual proof fails.

(500.4) $EOxOy\mathbf{A}Ox\mathbf{A}$

Proof. 1.	$COx\mathbf{A}OxOy\mathbf{A}$	((401))
2.	$CPxPy\mathbf{A}Px\mathbf{A}$	(1., $N\mathbf{A}/\mathbf{A}$, df.*P*., R.6.)
3.	$COxPy\mathbf{A}Px\mathbf{A}$	((100), $Py/\mathbf{A}$, 2., R.6)
4.	$CPx\mathbf{A}OxPy\mathbf{A}$	((500))
5.	$EOxPy\mathbf{A}Px\mathbf{A}$	(3., 4., R.6.)
6.	$EOxNOy\mathbf{A}NOx\mathbf{A}$	(5., df.*P*)
7.	$ENOxNOy\mathbf{A}NNOx\mathbf{A}$	(6., R.6.)
8.	$EPxOy\mathbf{A}Ox\mathbf{A}$	(df.*P*., R.6.)
9.	$EOxPyOz\mathbf{A}PxOz\mathbf{A}$	(5., $Oz\mathbf{A}/\mathbf{A}$)
10.	$EOxOy\mathbf{A}PxOz\mathbf{A}$	(8., 9.)
11.	$EOxOy\mathbf{A}Ox\mathbf{A}$	(8., 10.)

Lines 8. and 9. of this proof are the S5 reduction theorems.
(500.4) itself is the S4 reduction axiom.

Taken together these results suggest that, if reduction theorems are desirable, the best special axioms to add to Ox will be (401) and (500). However, under interpretation these axioms are undesirable for other reasons. The question of plausibility under interpretation of various axioms and theorems is discussed later.

One further technical result will be of interest. Prior ([14]) has shown that Mrs Barcan Marcus's axiom, of which (300) is a version in Ox, is provable in S5 with quantifiers added. The proof cannot be reproduced in Ox with, say, (400) and (500), again because of the absence of any version of *CLpp*; but again a related result can be

([14]) Prior, *op.cit.* in note 11.

proved. Curiously, the proof of this result can only be proved if we add (401) and (500), since theorem (500.4) is necessary; at least, I have been unable to prove it otherwise.

(500.5)	$CPx\Pi zOy\mathbf{A}_z Ox\Pi z\mathbf{A}_z$	
Proof. 1.	$COxC\mathbf{AB}CPx\mathbf{A}Px\mathbf{B}$	(line 10. of Prior's proof; proof in Ox similar)
2.	$CPx\mathbf{A}PxPy\mathbf{A}$	(from (500.4))
3.	$CPxOy\mathbf{A}Ox\mathbf{A}$	((500), R.6., df.*P*)
4.	$COxPy\mathbf{A}_z Ox\Sigma zPy\mathbf{A}_z$	(R.6., R.4., R.5., R.7., (200))
5.	$C\Sigma zPx\mathbf{A}_z Ox\Sigma zPy\mathbf{A}_z$	((500), df.*P*, R.4-6.)
6.	$CPx_1\Sigma zPx\mathbf{A}_z Px_1Ox\Sigma zPy\mathbf{A}_z$	(5., R.7., 1.)
7.	$CPx_1\Sigma zPx\mathbf{A}_z Ox_1\Sigma zPy\mathbf{A}_z$	(3., 6., R.6.)
8.	$COx_1\Sigma zPy\mathbf{A}_z Ox_1Ox_2\Sigma zPy\mathbf{A}_z$	((500.4), R.2.)
9.	$CPx_1\Sigma zPx\mathbf{A}_z Ox_1Ox_2\Sigma zPy\mathbf{A}_z$	(7., 8., R.6.)
10.	$CPx_1\Sigma z\mathbf{A}_z Px_1\Sigma zPxPy\mathbf{A}_z$	(2., R.4-5., R.7, 1., R.2., R.1.)
11.	$CPx_1\Sigma z\mathbf{A}_z Ox_1Ox_2\Sigma zPy\mathbf{A}_z$	(9., 10., R.1.)
12.	$CPx_1\Sigma z\mathbf{A}_z Ox_1\Sigma zPy\mathbf{A}_z$	((500.4), R.6.)
13.	$CPx\Pi zOy\mathbf{A}_z Ox\Pi z\mathbf{A}_z$	(R.3., df.Σ, df.*P*., R.6.)

The question of plausibility under interpretation will now be discussed.

4. *Choice of special axioms.* Before we can decide which of the special axioms to adopt, an objection must be answered which applies to them all, namely the charge that none of them is true under interpretation. I shall answer this charge by showing that there is a legitimate sense in which (400) is a logical truth. I shall then discuss whether any of the others can also be accepted as logically true in the same sense.

The objection just mentioned might be specifically directed against (400) as follows. There are many occasions, it might be said, on which it is not true at all. For instance, suppose *a* wants *p* (to be the case), but for some reason does not want to give the command that *p* (be the case) himself — perhaps because he is embarrassed or afraid. Then he might well both command *b* to do so and refuse to do so himself. Here is a counter-example to (400), which therefore cannot be a logical truth. Although a strain of some kind arises in such situations, it is not the extreme strain of logical falsehood. *a*'s position would be weak, but not logically weak.

I think this charge can be rebutted. Consider a case in which we normally would say that *a* was contradicting himself, saying something logically false. For instance, suppose he asserts that *q* follows from p, but refuses to admit that the denial of p follows from the denial of q. Then a parallel objection could be raised to calling his position logically inconsistent, as follows. Clearly there is strain in what he says, it might be said, for his position could no doubt be shown, by means of logical truths he does accept, to lead to contradiction; but though this might cause him some embarrassment, the strain here is not the extreme strain of logical falsehood. This objection, we feel, is easily answered: if this, we reply, is not the extreme strain of logical falsehood, what *would* count as an example ? Why then is this reply not available to the objection to (400) ? It may said, because in that case the strain *is* extra-logical. When this is said it is surely clear that what is at issue is what is to count as logical and what to be excluded as extra-logical. This question is not so easy to answer as it is often thought to be.

Strawson ([15]), for example, attempts an informal explanation of logical truth in terms of contradiction. His proposal, without his qualifications (most of which I think are important) seems to be that someone asserts what is logically false if and only if he either explicitly contradicts himself, or commits himself to doing so. But this proposal is inadequate, since in order to understand what it is to be committed to self-contradiction I must already know what logical commitment is (it being logical commitment that is intended), and without this notion of commitment it will not be possible to specify a necessary condition of asserting a logical falsehood. And there are obviously other kinds of commitment besides logical commitment. So the explanation is so informal as to be, if we insist on any kind of formalities at all, circular.

A similar objection can be raised to Popper's method ([16]) of explaining the nature of logical truth. His primitive notion is that of one statement following from another. A statement is logically true if it follows from the null set of statements. Again, we have first to understand the distinction between what follows logically and what follows extra-logically. That a man decides to do some crucial deed, for example, may follow (inevitably) from his character's being what it is. It is often thought that if one *statement* follows from another

([15]) P. F. Strawson, *Introduction to Logical Theory*, Methuen, London, 1952; ch.s 1 and 2.

([16]) K. R. Popper, «New Foundations for Logic», *Mind* vol. 56, 1947, pp. 193-235.

K. R. Popper, «Logic without assumptions», *Proceedings of the Aristotelian Society*, 1946-47.

it must do so logically; but this is far from true. Thus Popper's account of logical truth, though not without merit, does not answer our question.

What do we have to understand in order to understand that something does follow ? In the case where we understand that the deed followed (inevitably) from the man's character, what is necessary is knowledge of a certain kind — crudely speaking, psychological or spiritual knowledge. To know that it follows we have to know what the man is. In the case where we understand that one statement follows logically from another, what is necessary again a certain kind of knowledge — but of *what* kind ? It is useless to say: logical knowledge.

The popular, and obviously right, answer is: knowledge of the meanings of certain words, namely what are commonly called the logical words (the logical constants, as they used to be called). This account need not be circular; for we can simply make a list of these words, label the list «logical words», and decide later what the words have in common. Then logical following, logical truth, and the other logical notions can be defined in terms of the meanings of words on this list. For instance, a statement is logically true if someone can understand that it is true whether he knows the meanings of the non-logical words used to make it or not, provided only that he does know the meanings of the logical words used to make it.

If this is accepted as being roughly right — I do not claim accuracy — I go on to ask: what words are on the list ? How is it decided whether a given word shall be on the list or not ? It should not be thought that no-one has ever disputed this. The quarrel between those who do and those who do not insist on a strictly extensional logic (excluding for instance modal logic *in toto*) is a quarrel about just this question. Is the word «necessarily» to go on the list, or not ? Similarly, it has been asserted, and denied, that the statement that *a* ought to do a certain thing logically implies (entails) that it is in *a*'s power both to do and not to do that thing. And it seems clear to me that, since the notion of the list of logical words is fundamental, all other logical notions being defined in terms of it, what is to be admitted to the list is, *to a large extent,* a matter for decision. There must be, I suppose, certain tests; but what they are I find it very difficult to say. There must of course be some possiblility of system: a disjointed collection of entailments, each discovered by a flash of insight, will not be admitted to be logical unless some order can be introduced. This is a vague stipulation, and besides, we shall never get as far as the insights unless we allow the logician freedom to investigate any words he chooses. For this reason I doubt if much good will come of trying to decide in advance what words or notions may be considered as potential or actual logical words or notions. It seems better to say: logic is concerned with what logicians are concerned with — and leave the consequences of this to logicians.

These considerations show, I hope, that (400) has at least some claim to the title of logical truth. Its claim stands or falls with that of the the notion of commanding, the word «commands», to being admitted to the list of logical words or notions. And I hope to have shown, developing a suitable formal system, that this word has at least as good a right to a place there as the ordinary modal words.

A second objection could be raised to (400). Whether it is logically or extra-logically false, it might be said, the fact remains that it is false, as the previous counter-example shows. To answer this objection it will be necessary to consider the example in more detail. Let us consider three persons *a, b,* and *c*. The command, issued to *c,* to do some act of type F I write «F*c*». Now suppose *a* says to *b*, «Tell *c* to do F», and *c* says to *a* «Do you tell me to do F ?» (not «Did you tell me...?» — *c* is asking *a* what command he is prepared to give him, not about what command he gave when speaking to *b*); and *a* then says «No, I do not». Then according to (400) *a* has said two inconsistent things. Suppose further that *a, b* and *c* are face to face, and that *a* cannot give any special reasons for refusing to give the command to *c* directly. Has *a* contradicted himself ? Plainly he has not said anything of the form *KpNp*; but in a slightly extended and quite obvious sense what he has said is inconsistent. He has said something, and then said something else the effect of which is to cancel or retract what he first said. Surely this is contradiction in a clear sense.

Suppose that *a* gives his indirect command to *b* in *c*'s absence, and his reply to *c* in *b*'s absence; does this make any difference ? If (400) is true it makes no difference. Consider a parallel case, in which statements are made. Suppose *a* says to *b* «S is P», and then, when *c* asks him «Do you say that S is P ?» he replies «No, I do not». Suppose further that *a* can give no special reasons, such as change of facts or of mind, for what he has said. If all three are face to face throughout the conversation then *a* has patently contradicted himself; but if he speaks to *b* and *c* separately the contradiction will be apparent only to him. Yet we do not hesitate to say that *a* has contradicted himself in either case.

A vague justification of this view is the thesis, which lies behind the formal logic of propositions, that the truth of a proposition is independent of the person to whom it is uttered. We are prepared to abstract propositions from their context to at least that extent. Some roughly similar thesis lies behind (400). It may be specified in this way: what would count as a command's being obeyed can be specified independently of the person to whom it is uttered. This is not entirely true of course; for most commands are obeyed by some action of the person to whom they are addressed, while only a small number of propositions mention the person to whom they are addressed. But the cases are similar enough for it to be possible, and I think desirable, to abstract from the context of commands also, at least from the identity of the addressee. To put this briefly: my assumption is simply that a command «*x*, do F» is obeyed if and only if the statement that *x* does F is true.

The objection to (400) might also take this form. In the case of propositions our objection is to logical inconsistency; but in the case of inconsistent commands, or an inconsistent refusal to command, our objection is to moral rather than logical inconsistency. When *a* refuses to give the command to *c* directly, his refusal is like that of man who commits himself to a course of action, and then refuses to do it — rather than that of a man who (logically) commits himself to a certain statement, and then refuses to make it; he is like a man who breaks a promise. My reply to this objection is that the laws of propositional logic also can be thought of as rules telling us what we must, may, and must not say — rules designed to prevent people from saying things which it is desirable to prevent, namely self-defeating utterances. The trouble with both logical and moral inconsistencies is that they defeat our purposes. The similarities between them are far-reaching. Of course, for most purposes it is desirable to distinguish between logical and moral rules. In proposing that (400) is a logical truth I am drawing the distinction in a slightly different place from that in which it is usual to draw it.

I assume, then, that (400) is acceptable, and go on to ask: are any of the other special axioms also plausible ? Let us first consider (401), perhaps the most interesting from a formal point of view. Under interpretation it asserts that if someone commands someone else to command that something be done, then that second person commands that it be done (or should, to avoid inconsistency). Thus all disobedience of indirect commands is inconsistent. The consequences of accepting this as a logical truth would indeed be alarming; (401) must certainly be rejected. The same is true of (403), which under interpretation asserts that if one person commands that something be done, then someone else commands that he command that it be done (or should, to avoid inconsistency). Thus all commands issued by subordinates are confirmed in advance — another alarming consequence, and (403) must be rejected. (500) is rather more difficult to decide about. A consideration which seems to me decisive, however, is this. (500) is equivalent to $CPxOy\mathbf{A}Ox\mathbf{A}$, and this, with (402), yields by syllogism $CPxOy\mathbf{A}OxOy\mathbf{A}$; while (500) with (400) yields the converse, $COxOy\mathbf{A}PxOy\mathbf{A}$. Hence $EPxOy\mathbf{A}OxOy\mathbf{A}$: there is no difference in logical commitment between commanding that a command be given and permitting that it be given. Since (402), as I shall show, is acceptable, (500) must therefore be rejected. (402) under interpretation asserts that if someone gives a command he must command another person to give it (to avoid inconsistency): I admit to a little discomfort in proposing this as a logical truth. The weaker $COx\mathbf{A}PxOy\mathbf{A}$ may be more plausible. But I think (402) can be accepted. With (400) it yields $EOxOy\mathbf{A}Ox\mathbf{A}$, which may be taken to mean that anyone at the head of a chain of command must be assumed to be directly responsible for the commands given on his instructions. Finally, can the quasi-reduction-theorems (400.4) and (400.5) be accepted? These under interpretation assert that (400.4) if x orders y to permit that something be done, then he must order z to permit x_1 to order y to permit that it be done; and (400.5) if x orders y to permit z to order x_1 to permit that something be done, then he orders y to permit that it be done. These principles are clearly much more complicated than any in ordinary use. However, I think it is possible to imagine a situation in which they, or some similar rules, would be needed. I mean a situation in which chains of commands and permissions commonly occur. Here (400.4) would allow an increase of indirectness; (400.5) would allow a decrease. In such circumstances I see no reasons why they should not be accepted.

There is a stronger axiom, containing a quantifier, from which (400) can be derived, namely

(600) $COx\Sigma yOy\mathbf{A}Ox\mathbf{A}$;

and a corresponding stronger axiom from which (402) can be derived:

(602) $COx\mathbf{A}Ox\Sigma yOy\mathbf{A}$.

Various other simple quantified axioms might be chosen on formal grounds, but none seem to have any plausibility, unless they are provable from these two.

Finally, we can combine these two axioms by conjunction:

(603) $EOx\Sigma yOy\mathbf{A}Ox\mathbf{A}$.

My final suggestion, then, for the logic of commanding is Ox with

(603); this is a weakened form of deontic S4 with quantifiers.

5. *The logic of asserting.* The axioms of the system given by Łos [17] are:
(1) *LxCC***ABCCBCCAC**
(2) *LxCCN***AAA**
(3) *LxC***A***CN***AB**
(4) *ELxN***A***NLx***A**
(5) *CLxC***AB***CLx***A***Lx***B**
(6) CΠ*xLx***AA**
(7) *ELxLx***A***Lx***A**

The rules are the same as R.1.-5. of Ox, with *L* for *O*. R.6. is a derived rule in Łos's system. (5) is the same as (200) in Ox, with *L* for *O*. I now discuss axioms (1)-(4), (6) and (7). (4) is equivalent to *ELx***A***NLxN***A**, and I think has just the same lack of plausibility as the corresponding version in Ox, which would assert the equivalence of commanding **A** and permitting it. The correct axiom here is surely the corresponding version of (100): *CLxN***A***NLx***A**.

(6), under interpretation, asserts that if everyone asserts that something is the case, then it is the case. This remarkable suggestion seems to have been adopted so that the theorems of propositional calculus can be proved in the system — as they can be from (1)-(3), quantifier rules and (6). But surely the weaker and altogether less incredible procedure is to adopt a version of R.6, and drop (1)-(3) and (6). (1)-(3) will then be provable as theorems, and (6), I conjecture, will not.

(7) is a weakened version of (603) of Ox. It asserts under interpretation that *x* asserts that he asserts something if and only if he does asserts it. This I think is plausible. But if we generalise (7) to the corresponding version of (603) in Ox, it loses all plausibility: for *x* can surely assert that *y* asserts something without being in the least committed to asserting it himself. This is, I think, one at least of the essential differences between asserting and commanding.

If my suggestions are adopted, the system appropriate to an interpretation as a logic of asserting is deontic S4. Once again, reduction theorems are not available in the general sense; but here, unlike in Ox, the quasi-reduction-theorems are not available, as there seems to be no plausible special axiom. The system does, however, provide an alternative, and I think in some ways more comprehensible, method of interpreting iterated alethic modalities.

University of Khartoum

([17]) See references in note 4.