The logic of obligation and the obligations of the logician[1]

by
A.N. Prior

§1. The unlawfulness of the impossible

Moral philosophers have sometimes claimed to deduce particular duties from the very nature of obligation. It would now be agreed by most that this cannot be done; and with this general view I have no serious quarrel. Yet there are forms of reasoning which in some sense owe their cogency to 'the very nature of obligation'[2]. For example, if it is not possible to do A without doing B, we may infer that if it is permissible to do A, then it is permissible to do B. 'Possible' is a notoriously ambiguous word; so let it be understood that bare logical possibility is intended. It is impossible, for example, that John Jones should at once be drinking tea in the kitchen and not drinking tea; ergo, if it is permissible that he be drinking tea in the kitchen, it is permissible that he be drinking tea, i.e. it is not obligatory that he be not drinking it. No one, I think, would quarrel with this form of inference. And it holds because obligation (in terms of which permissiblity may be defined) is what it is. Replace 'permissible' by, for example, 'not the case', and the inference loses its cogency. ('It is impossible to do A without doing B' does not imply 'If one is not doing A one is not doing B'). And it is with the principles of such inferences as these that the logic of obligation is concerned.
Let us look again at the principle just mentioned. If at-once-doing-A-and-notdoing-B is not possible, then if A is permissible so {2}[3] is B – whatever A and B might be. But – this is not the logic of obligation but logic pure and simple – if it is impossible to do A at all, then at-once-doing-A-and-not-doing-B is clearly impossible too (at-once-doing-A-and-anything-at-all is impossible, if doing A is impossible). Syllogize from these two, and the conclusion plainly is that if A is something it is impossible to do, then if A is permissible so is B. Or, if A is impossible and is also permissible, then B is permissible – B being anything at all. That is, if A is at once impossible and permissible, anything whatever is permissible. But some things are not permissible. Hence it cannot be that A is at once impossible and permissible.

This conclusion takes us somewhat further than Kant's "What I ought I can". That tells us only that what is not possible is not obligatory; but what we have now learnt is that it is not even permissible. However, if we were unaware of this prohibition before, we need not fear that we have inadvertently contravened it. What is forbidden is not attempting the impossible, but doing it, and anything that we have actually done, is not impossible [1][4]. And if this conclusion – which was first

[1] This paper has been edited by Anne Kathrine Kjær, Jörg Zeller, and Peter Øhrstrøm. The original is undated and kept in the Prior collection at Bodleian Library, Oxford. An earlier version of this text has been published in *Synthese* (2012) 188:423–448.
[2] Prior is somewhat inconsistent in his use of single and double quotation. We use the single quotation, when it refers to a name, a term or something figurative. In other cases, we use double quotation.
[3] By this, a number in curly braces, we mark the page numbers from the typewritten original.
[4] Prior's own notes are indicated with square brackets and have been places at the end of the text (i.e. as endnotes).

drawn by Hintikka as reported by von Wright[5] [2] – need not frighten us, neither, I am afraid, does it practically help us. It remains true that no specific precept which can actually guide our action is deducible from principles which express merely the logic of obligation. Hintikka's theorem only brings out an unsuspected abstract connection between obligation and necessity.

Of those who study the logic of obligation, the first interest with some is in logic and with others in obligation. And those who approach it from these two ends are apt to become involved in conflicts of attitude. Hintikka's theorem may serve to show up the difference. {3} The logician will be extremely interested in it, and will even find a use for it, in his own sense of 'use'; moralists regard it with suspicion, and wonder where the catch is. To the moralist, the logician – especially when he talks about obligation – is irresponsible; to the logician, the moralist is puritanical. I am frankly on the logician's party, and am anxious that moralists should understand a little better what our standards are. For we do have them; and the way we talk about obligation is part of a general way of talking about anything whatever. I propose, therefore, to go into some detail about logic generally, especially about logic of the modern 'formalised' sort, saying at first nothing at all directly about the logic of obligation, but laying a foundation for that, and also picking my illustrations with an eye to what I shall be saying about that afterwards – amassing a set of parallels, and creating a context. And then we shall return to Hintikka's theorem.

Formal logic, like other activities of men, has its well-tried ways, its almost unconscious traditions and criteria of excellence, and to evaluate adequately a 'logic of obligation' we must think not only of its immediate subject-matter (obligation), but must see it taking its place in the larger company of formalised systems and absorbing something of their style. It is to help moralists to look at current efforts at a 'deontic logic'[6] from this point of view (and also, on the side, to advance deontic logic itself a little) that this paper is written. I make no promises, then, to show that the logicians game can be of any tremendous service to the moralists purposes; I would like rather to interest the moralist in the logician's purposes – to have him joining in the game at the end. And that could perhaps be of value, not to moral philosophy, but to morality; for it is a narrow and bad {4} morality which can only take one sort of purpose into account.

2. Formalisation of general logic

Suppose we are playing a game in which we construct sentences containing, to begin with, only the symbols p, q, r, s, C, N, and K. Of the various possible sequences containing these symbols, some we shall call P-formulae[7] and some we shall not. The small letters are all P-formulae, and if α and β are any P-formulae, (i) Nα is a P-formula, (ii) Cαβ is a P-formula, and (iii) Kαβ is a P-formula; and these are all the P-formulae we can construct using these symbols only. The following are three examples of P-formulae [3]:

1. CCpqCCqrCpr
2. CCNppp
3. CpCNpq

[5] Prior refers to the Finnish philosopher and logician, Georg Henrik von Wright (1916–2003), and is clearly inspired by von Wright's works. Von Wright is the author of "An Essay in Modal Logic", to which Prior refers in his notes.

[6] Prior apparently got the term 'deontic' from von Wright, who first suggested it in his works "Deontic Logic" (*Mind*, New Series, Vol. 60, No. 237 (Jan., 1951), pp. 1–15) and *An Essay in Modal Logic* (North-Holland Publishing Company, 1951). In "Deontic Logic" von Wright remarks in a footnote: "For the term 'deontic' I am indebted to Professor C. D. Broad."

[7] P-formula is an abbreviation for a "propositionally well formed formula".

1 is a P-formula because (a) p, q, r are P-formulae, so that by (ii) above, Cpq,Cqr and Cpr are P-formulae; so that, again by (ii), CCqrCpr is a P-formula; so that, again by (ii), CCpqCCqrCpr is a P-formula. That 2 and 3 are P-formulae may be shown similarly.

Of P-formulae, some are T-formulae[8] and some are not. The formulae 1, 2, and 3 are T-formulae, and any formula obtained from a T-formula by certain rules is a T-formula. These are the rules: (i) In any T-formula, any small letter may be replaced by any P-formula, provided that the same replacement is made throughout. This may be called the Rule of Substitution. For example, since CpCNpq (i.e. formula 3) is a T-formula, so is CpCNpp, which is obtained from it by replacing the small letter q by the P-formula p. This derivation might be set out as follows:

3. CpCNpq
3 q/p = 4
{5}
4. CpCNpp

(The middle line asserts that 3 with q replaced by p yields 4. And we show that something is a T-formula by giving it a number). Here is a more complicated derivation of a new T-formula by the same rule:

1. CCpqCCqrCpr
1 q/CNpp, r/p = 5
5. CCpCNppCCCNpppCpp

Further, (ii) if α is any T-formula, and Cαβ is a T-formula, then β is also a T-formula.We call this the Rule of Detachment. For example, let α be CpCNpp and β be CCCNpppCpp. Then α is a T-formula, namely 4, and Cαβ is a T-formula, namely 5; hence this β is a T-formula. This reasoning may be set out thus:

5 = C4 – 6
6. CCCNpppCpp

Again, let α be CCNppp and β be Cpp. Then α is a T-formula, namely 2, and Cαβ is a T-formula, namely 6, so this β is a T-formula; or setting this out as before:

6 = C2 – 7
7. Cpp

This derivation of Cpp from our original three T-formulae may be set out more compendiously as follows:

1 q/CNpp, r/p = C3 q/p – C2 – 7
7. Cpp

[8] T-formula is an abbreviation of ‘theorem-formula’.

Finally, (iii) in any T-formula, any P-formula of the form NCαNβ may be replaced by the corresponding formula of the form Kαβ, or vice versa, and the result will also be a T-formula. This rule may be written:

Df. K: Kαβ = NCαNβ
{6}What it amounts to is that in this game Kαβ is simply an abbreviation for NCαNβ, and we may use the abbreviated or the unabbreviated form as we please. For example, the formula

8. CCpCqrCNCpNqr

may be shown to be a T-formula; hence, abbreviating the appropriate portion of it,

9. CCpCqrCKpqr

is a T-formula. This derivation may be written out thus:

8 x Df. K = 9.
9. CCpCqrCKpqr

These are our symbols and rules, and the game is simply to find T-formulae. All the information required for playing it has been given; in particular, in order to play the game successfully, we do not need to know what use or uses might be found for symbols manipulated according to these rules. To save tedium, however, two possible uses may be mentioned. The small letters could be understood as variables standing for numbers, like the x's, y's etc. of ordinary algebra; Nα for 1 + α and Cαβ for 1+α(1+β). Nα, it may be observed, is even if α is odd, and vice versa. Cαβ is odd if α and β are both odd, if α is even and β odd, and if α and β are both even, while it is even, if α is odd and β even. Kαβ, i.e.1 + (1 + α(1 + (1 + β))), works out as odd if and only if both α and β are odd. And all T-formulae will have the property of representing odd numbers regardless of whether their component variables represent odd numbers or even. (This can be shown by calculation in the case of 1, 2 and 3; and the rule of detachment on this interpretation asserts that if α is a formula that can only represent odd numbers, and so is Cαβ, then so is β; which can be shown to hold) [4]. {7}

Alternatively, the small letters may be taken as standing for statements: N for the phrase 'It is not the case that –'; C for 'If – then –' (it being understood that a statement of the form "If p then q" is false if p is true and q false, and true otherwise), and K for 'Both – and –'. All P-formulae will then represent statements, and all T-formulae will represent true statements regardless of whether their contained small letters represent true statements or false ones. T-formulae, we may say, will then express 'logical laws', or more specifically, laws of the propositional calculus or logic of truth-functions. For example, 1 (CCpqCCqrCpr) expresses the law "If p implies q, then if q implies r, p implies r "; 7(Cpp) the law "If p then p"; 9 (CCpCqrCKpqr) the law "If p implies that q implies r; then the combination both-p-and-q implies r "; and the rule of detachment becomes the principle that if α expresses a logical law and "If α then β" also does so, then β may be affirmed as a new logical law. (It was by assuming 9, thus interpreted, that we passed, in proving Hintikka's theorem, from "If A is impossible then if A is permissible B is permissible" to "If A is both impossible and permissible, B is permissible").

Let us now return to our game, and – again disregarding its possible uses–enlarge the notions of P-formula and T-formula a little. So far we have only been given means of constructing P-formulae out of other P-formulae. But let us now suppose we are going to be supplied with an indefinite number of other types of symbols out of which, in various specified ways, P-formulae may be constructed. And let us suppose that, as they come in, these symbols are divided into V-symbols (‘variables’) and K-symbols (‘constants’), and that of the ones we already have, the small letters are V-symbols and C, N and K are K-symbols. For each class of V-symbols there will be appropriate rules of substitution, and we may also form a P-formula by attaching either of the symbols Π and Σ to a V-symbol and prefixing the whole to a P-formula. That is, if x is {8} any V-symbol, and α any P-formula, Πxα and Σxα are P-formulae. For example, using the symbols we already have, ΠpCpq is a P-formula (since p is a V-symbol and Cpq is a P-formula). Π and Σ are said to ‘bind’ the V-symbol to which they are attached in the P-formula which they begin, and a V-symbol which is not thus bound by a Π or a Σ is said to be ‘free’. Associated with Π and Σ are certain new rules for deriving T-formulae from other T-formulae, namely:

Df.Σ: Σx = NΠxN

Π1: If Cαβ is a T-formula, so is CΠxαβ.

Π 2: If Cαβ is a T-formula, so is CαΠxβ, provided that the V-symbol x does not occur freely in α [5].

We also associate with Π and Σ certain restrictions on substitutability for V-symbols, namely that we cannot substitute for such a symbol at a point in a formula where it is bound by a Π or a Σ; nor can we substitute for a free V-symbol a V-symbol which is bound elsewhere in the formula.

If we interpret the P-formulae as statements, or more accurately as forms of statements, the prefixes Πx and Σx (where x may be any sort of V-symbol) may be interpreted as the quantifiers “For all x” and “For some x”. Thus ΠpCpq will mean “For all p, if p then q”, or “Everything implies q”); and with this interpretation in mind, we may now illustrate the use of the above rules.

Suppose we introduce a new set of V-symbols x, y and z, which we shall call N-symbols; and a pair of V-symbols ϕ and ψ such that ϕ or ψ followed by any N-symbol as a P-formula and, where α is an N-symbol ϕα or ψα may be replaced in a T-formula by any P-formula constructed out of α (provided that the P-formula thus substituted is one in which α occurs freely). Let us also add the K-symbol I, such that I followed by two N-symbols is understood to be a P-formula, and let {9} us associate with I the two new T-formulae (axiomatically laid down)

10. Ixx
11. CIxyCϕxϕy

If the N-symbols are interpreted as standing for names of individuals, and ϕ and ψ for verbs (forming statements out of names), I could be the predicate ‘- is identical with -’. 10 would then assert the law that anything is identical with itself, and 11 the law that if x and y are one and the same thing, then if xϕ’s, yϕ’s (whatever ϕ-ing is taken to be).With the basis which we now have, we may prove the following two theorems: (i) if x is identical with y, then whatever x does y does, and (ii) if whatever x does y does, then x is identical with y. (i) follows directly from 11 by Π2, thus:

11. CIxyCϕxϕy
    11 x Π2ϕ = 12
12. CIxyΠϕCϕxϕy

And this deduction may serve to illustrate the rationale of the use ofΠ2. In general it is a dangerous procedure to pass from “If so-and-so then if xϕ’s yϕ’s” to “If so-and-so, then y does everything that x does”, but if the former is laid down as true whatever we take ϕ–ing to be, and ϕ can vary independently of any variables occurring in the antecedent, the procedure is safe enough; and these conditions are here fulfilled. (It is in just the same way that, in proving Hintikka's theorem, we pass from "If A is both impossible and permissible, B is permissible", where B can be any act, and can vary independently of A, to "If A is both impossible and permissible, anything at all is permissible"). Our second theorem about identity, the converse of the other, is true because, by Π1 and the law of identity Cpp, if whatever x does y {10}does, then if xϕ's (say) yϕ's; so (substituting I xα for ϕα) if whatever x does y does, then if x has x identical with it, y has x identical with it; so (transposing antecedents in virtue of the law CCpCqrCqCpr, known to be provable from 1–3) if x has x identical with it; then if whatever x does y does, y has x identical with it; but x is identical with x (by 10); hence if whatever x does y does, x is identical with y. Or setting it out formally:

7. Cpp
14[9]. CCpCqrCqCpr
10. Ixx
  7 p/CϕxϕyX Π1ϕ = 15
15. CΠϕCϕxϕyCϕxϕy
  15 ϕ/Ix'[10][11] = 16
16. C ΠϕCϕxϕyC I xx I xy
  14 p/ΠϕCϕxϕy, q/I xx, r/Ixy
  = C16 –C10 - 17
17. CΠϕCϕxϕyI xy

(Note that in the substitution by which we pass from 15 to 16, the replacement for ϕ is made only at those points in 15 where its occurrence is 'free').

12 and 17 between them assert that Ixy and ΠϕCϕxϕy are logically equivalent, i.e. that it is a logical law that each implies the other; and with this result we obtain all T-formulae which we would have obtained by introducing I, not as a quite new symbol with special axioms, but as a simple abbreviation with the definition

Df.I: Ixy = ΠϕCϕxϕy.

Conversely, if we are given this definition, we may obtain all that we can obtain from the axioms 10 and 11, for with this definition 10 and 11 are provable by means of propositional calculus and rules {11}for quantifiers. Hence we obtain the same results in terms of T-formulae containing I whether we proceed in the first way or in the second; and as far as economy in initial data is concerned, the second way (using the definition) is clearly preferable. It could be argued, on the other hand, that "Whatever x does, y does" is not what we intuitively mean by "x is identical with y" (even when 'does' is so widely understood that having x identical with it counts as one of x's 'doings') and that consequently the deduction of 12 and 17 from 10 and 11 represents a process of discovery which we need to go through, and to record. To this, the manipulator of formal systems will reply that the question as to what is intuitively meant by this or that is too subjective a one to enter into his calculations. The formalist might add, too, that if the above definition of '- is identical with -' is

[9] Prior does not include 13.

[10] The upper case X in the typewritten original is apparently used erroneously instead of a lower case x.

[11] Prior uses the sign of single quote in Ix' as marker for an empty argument space of the 2-place predicate 'I'.

objected to on the ground that it does not give the intuitive meaning of the phrase, a similar objection would have been in order at a much earlier point. For even within the propositional calculus, can it be said that "It is not the case that if p then not q" is what we intuitively mean by "Both p and q"? Indeed, our intuition is violated in two ways here – both in giving this definition of 'and', and in not giving a definition of 'if'. For the 'if' which is symbolised by our C is notoriously not the 'if' of ordinary discourse, but a logician's invention; and the closest thing to Cpq in ordinary speech is a form in which 'and' already occurs, "Not both p and not q". But these considerations bear solely on the interpretation of a symbolic system, not on its elegance and ease of manipulation. As far as the system is concerned, symbols are divisible into two sorts: those which are introduced by a rule of abbreviation, and those which have no definition of this sort, but which are in a sense defined by the rules and axioms. And in particular, as far as the system is concerned, Cαβ does not mean {12} "If α then β", or "Not both α and not β", but means anything you like that will fit 1–3 and the rule of detachment (if 'if' in the ordinary sense will not do, that is no matter – try something else). The formalist's aim is to introduce as few symbols as possible in this way, and as many as possible by rules of abbreviation. Nowadays one carries this so far as to define not only 'and' but even 'not' in terms of an 'if' which is defined only by rules and axioms. Just as all the usual T-formulae containing I will be obtained if we use Ixy as an abbreviation for ΠϕCϕxϕy (whether I in this sense represents what we commonly understand by '- is identical with -' or not), so all the usual T-formulae containing N will be obtained if we introduce Np as an abbreviation for CpΠpp, 'If p is true, then everything is true'. (Remember that intuitively this amounts to "Not both p-is-true and not-everything-is-true", and since "not-everything-is-true" is true anyway, the whole complex will be false if p is true also, but true if p is false, and precisely this is, of course, the distinguishing property of "Not p") [6]. Given this definition (call it Df. N) we obtain the same T-formulae as with the other basis (1, 2, 3, etc.) if we have substitution, detachment, the rules for Π, and the single axiom

18. CCCpqrCCrpCsp [7]

Let us now make an addition to our symbolic material at a different point, by introducing the two K-symbols L and M, it being understood that if α is any P-symbol, Lα and Mα are P-symbols; and let us call any V-symbol 'modalised' if it is part or the whole of a P-formula with M or L prefixed. (Thus p is modalised in Mp and in LCpq, but not in CpLq.) And let us add the following new rules for deriving T-formulae:

Df.M: M = NLN
L1: If Cαβ is a T-formula, so is CLαβ.
L2: If Cαβ is a T-formula, so is CαLβ, {13} provided that no V-symbol occurs unmodalised in α.

From these it is possible to deduce two further rules:

M1: If Cαβ is a T-formula, so is CMαβ, provided that no V-symbol occurs unmodalised in β.
M2: If Cαβ is a T-formula, so is CαMβ.

The following is a typical piece of deduction on this basis:

19. CKpqp
20. CCpqCNqNp

19 x M2 x M1 = 21
21. CMKpqMp
20 p/MKpq, q/Mp = C 21 – 22
22. CNMpNMKpq.

Here 19 and 20 are known to be derivable from 1, 2 and 3 by substitution, detachment and Df. K (or from 18 by substitution, detachment, Π1,Π2, Df. K and Df. N). And on the propositional calculus interpretation, their truth is evident enough; they assert respectively "If both p and q then p" and "If p implies q, the falsity of q implies the falsity of p". If we read L as 'It is necessary that' and M as 'It is possible that', what our symbolic derivation amount to is, firstly, arguing that since the truth of the combination both-p-and-q entails the truth of the simple p, the possibility of the former entails the possibility of the latter (21), so that the impossibility of the latter entails the impossibility of the former (22) [8]. The concluding principle, it may be remembered, is used in the proof of Hintikka's theorem.

The interpretation just suggested for L and M is the usual one; {14} their introduction into the propositional calculus, with the rules Df. M, L1 and L2, can be regarded as providing the symbolic basis for a development of the logic of modality. It is true that if a man were asked to set down off-hand a few formulae expressing properties of possibility and necessity from which all other properties of these concepts might turn out to be deducible, he would be unlikely to hit upon L2, though Df. M and L1 would be possible. Df. M equates "Possibly p" with "Not necessarily not p", and according to L1, if the truth of α entails that of β, so does the necessity of α; and these are obvious features of possibility and necessity which might easily strike a man as fundamental. That L2 holds, however, is not at all clear intuitively, though of course neither is it clear intuitively that it does not. It is in fact the end-product of a long process of trial and error, in which one first sets out at random various propositions involving possibility and necessity which are obviously true in all cases, and others which are not, and tries one basis after another for obtaining everything one can think of under the first head and nothing under the second. Deduction using L2 will certainly not represent the means by which we come in fact to learn those modal principles of which we are most certain; but then that is not what is claimed for it. The manipulator of formal systems will be content if Df. M, L1 and L2 (subjoined to 1, 2 and 3 with substitution and detachment) will yield all modal principles to which on reflection we assert, and none which on reflection we are compelled to reject. If they also yield a few principles which no amount of reflection will enable us to decide upon intuitively, may we not count these as genuine logical discoveries, to be accepted no doubt with caution, but in their own way exhibiting the possible positive value of rigorous formalisation? They are by-products of the system which gives the most economical explanation of what is {15} definitively known; are not these to be welcomed in this science as in any other?

But is the modal logic which one would obtain by giving the modal interpretation to L and M in Df. M, L1 and L2 really in this position? This basis is known to yield all modal principles to which people can generally be counted on to assert; but does it yield nothing which ought definitely to be rejected? To illustrate the kind of data we have for making up our minds on this point, let us consider a few T-formulae which this basis definitely does yield. When our three rules are subjoined to propositional calculus, one T-formula which is easily obtainable is:

23. CMpLMp

asserting that whatever is possible is necessarily possible [9]. If the theory of quantifiers is added, one formula that is obtainable is:

24. CMΣxϕxΣxMϕx,[12]

"If it is possible that something ϕ's, then there is something that has the possibility of ϕ–ing" [10]. And if we add also the theory of identity we may obtain

25. CNIxyLNIxy

"If x is in fact other than y, then x could not but be other than y" [11]. 23 is the sort of principle about which most people seem incapable of deciding intuitively, and so could perhaps be accepted as a 'logical discovery', but there are people who do not like it. 24 might be queried on the ground that even in cases in which nothing that in fact exists is capable of ϕ–ing, there might have been something which both could and did ϕ. And 25 leaves one very uneasy indeed.

If the derivability of 23, 24 and 25 is put forward – as at least 24 and 25 could very well be – as making this particular modal system quite unacceptable, the formalist might simply say, "Oh well, let it {16} go then – my M and L cannot mean what you mean by 'Possibly' and 'Necessarily', so use them for something else. It's a pretty system anyway". Or he might, if in a more helpful mood, suggest alternative rules and axioms for M and L which would not yield 23, 24 and 25, but would still yield those modal principles of which most men are certain. This requirement could in fact be met by keeping Df.M and replacing L1 and L2 by the two axiomatically-laid-down T-formulae

26. CLpp
27. CLCpqCLpLq

(asserting that whatever is necessary is true, and that if p necessarily implies q, then if p is necessary q is necessary) and the rule

RL: If α is any T-formula, so is Lα [12].

All these are derivable from Df. M, L1 and L2, but the converse is not the case, and with this new basis we cannot derive 23, 24 or 25 as T-formulae, though we would have no difficulty in deriving, say, 21 and 22. It should be noted, however, that even with the weaker basis for modal logic, if we substitute Mp for p in 26 we will obtain, not 23, but its converse,

28. CLMpMp

"What is necessarily possible is actually possible" and if the objection to 23 is not that its truth is doubtful but that a modal qualification 'has no meaning' when it is attached to something that is already modally qualified (as in the form LMp) [13], this will apply to 28 also.

To construct, on the assumption that this odd objection is to be taken seriously, a formalised modal system which would not be open to it, would be extremely difficult. It would not, however, be impossible; the solution to this problem would be similar to one offered in the {17} next section for a similar problem arising in the logic of obligation.

[12] Prior wrote CMΣxϕxϕxMϕx, but he must have meant CMΣxϕxΣxMϕx.

§3. Formalisation of the logic of obligation

We are now in a position to see what it would be like to have a formalised logic of obligation. It is a matter of introducing, at least into the propositional calculus and perhaps into other already-formalised fields also, some appropriate K-symbol to express the notion of obligation, this being 'defined' by some adequate set of T-formulae and/or rules.

Suppose, then, we introduce the symbol O, it being understood that where α is a P-formula so is Oα [14]. The interpretation envisaged is more or less the following: Suppose p is a statement of the form "xϕ's"; the Op will mean "x ought to ϕ", or "It is obligatory that x should ϕ". It should be emphasised that the form Op is not thought of as making an assertion about a statement. It no more does this than, say, Np does. "It is not the case that xϕ's", or "x does not ϕ", is not a statement about the statement "xϕ's", rather, it is a new statement about x. Similarly "It is obligatory that x should ϕ", or "x ought to ϕ" is not a statement about the statement "xϕ's"; rather, the operator O forms from the statement "xϕ's", which is of course about x, another statement which is also about x.

In this respect the form Op resembles not only the form Np but also the modal forms Lp and Mp. And principles which would be generally accepted about obligation suggest that there are further parallels. For example, just as p's being necessary implies that its opposite is not necessary (i.e. that it is itself possible), so p's being obligatory implies that its opposite is not obligatory (i.e. that it is itself permissible). Or in symbols, just as our modal T-formulae ought to {18} include CLpNLNp, so the T-formulae of a logic of obligation ought to include

29. COpNONp

And just as we have found it convenient to introduce the symbol M as an abbreviation for NLN, so we would find it convenient to introduce, beside O, the further K-symbol p (to be read 'It is permissible that') by the definition

Df.P: P = NON

("It is permissible that p" = "It is not obligatory that not p"). On the other hand, although modal logic includes the law CLpp, asserting that whatever is necessary is true, no serious logic of obligation would contain as a law COpp, which would assert that whatever is obligatory is done.
At least part of the task of formalising the logic of obligation thus consists in constructing a system which is structurally similar to a modal system without CLpp – a task similar in its general character to that of constructing a modal system which does not contain CMpLMp. And this much at least is not difficult to do. We could do it, for example, by laying down Df.P, 29 (or its abbreviated variant COpPp), the further axiom

30. COCpqCOpOq

and the rule

RO: If Cαβ is a T-formula, so is COαOβ.

30 reads literally "If it is obligatory that if p then q, then if it is obligatory that p it is obligatory that q". "Obligatory that if p then q" has an odd sound, but if we remember that "If p then q" amounts to "Not both p and not q", we may see that 30 amounts to "If it is obligatory that not both p and not q, then if it is obligatory that p it is obligatory that q", i.e. "If the combination p-and-not-q is {19} not

permissible, then if p is obligatory so is q". This seems reasonable enough, and is, of course, the analogue of the modal law 27, CLCpqCLpLq. As to RO, an example of its use would be this:

19. CKpqp
 19 x RO = 31
31. COKpqOp

i.e. if the combination p-and-q is obligatory, then p is obligatory.

There are a number of quite clear principles in the logic of obligation, formulable in the symbolism now at our disposal, which 29, 30, Df. P and RO do not suffice to prove [15], but into this deficiency, and possible remedies for it, we need not here enter. The more important question, we shall find, is not whether in this system we can prove enough, but whether we cannot prove too much. But before turning to that, let us consider just what we have so far done. The logic of obligation – the study of forms of inference which owe their cogency to the nature of obligation – could no doubt be generally classified as 'applied' logic rather than 'pure'. But what does these terms mean? There is of course a clear distinction between the game of finding T-formulae, which can be played without considering what our symbols mean or even whether they mean anything, and the interpretation of the game as a means of discovering the laws of some meaningful discipline. But in terms of this distinction, the logic of obligation is as 'pure' as any other kind of logic which we choose to formalise. Indeed, the process of formalisation – meaning by that the representation of the laws of some discipline by symbols which can be drawn into our T-formula game and thereafter used as counters without regard to their origin – is directed towards precisely that 'levelling' result. A T-formula is any sequence of symbols which our rules will make a T-formula; whether all of its component symbols have interpretations {20} like 'If – then –' or some of them are given such interpretations as 'It is obligatory that–', has nothing to do with it. And whether the rules by which the T-formula is derived refer only to the symbols with such interpretations as 'If – then–', or include ones dealing specially with (say) O and P, has nothing to do with it either. Certainly in the exposition of the T-formula game given in the last section and this one, we introduced our symbols in a rather definite, not to say tendentious, order – C, N, K and Π first; then I; then L and M; and finally O and P. But there was no reason in principle why we should not have introduced them all in a lump at the beginning, and our rules and axioms likewise. Also, it happens that we have T-formulae containing no K-symbols but C, N, K and Π, but none containing L, M, O or P without any drawn from the other group. We have, however, Ixx without them, and even if we treat this as an abbreviation for ΠϕCϕxϕx, there is no reason in principle why there should not occur in the T-formula game similar cases in which there is no question of any such 'reduction'.

The distinction between 'pure' and 'applied' logic must therefore be found, if anywhere, within the interpretation of the T-formula game. And here I think it can be found; but we should not be worried if it proves to be a little hazy and a little arbitrary. If anything belongs to 'pure' logic, the propositional calculus and the theory of quantification belong to it; and there is much to be said for the view that nothing else belongs to it. These form the tree on to which we do in practice graft whatever else we may think worth formalising. Yet is not at least the theory of modality a part of the tree too, especially if it is 'logical' necessity and possibility to which it is taken to refer? I do not think it matters what we say about this or other cases that might be represented as borderline ones, so {21} long as we do not say that the laws of pure logic hold only in virtue of our conventions while the laws of other disciplines hold because that is the way things are. Where convention rules is in the T-formula game, but whether a T-formula game with such and such conventions may be interpreted in such and such a way depends on whether, when thus interpreted, it fits the facts with which the interpretation is concerned. “If p then p” is no more ‘true by convention’ than it is true by convention that what is obligatory is permissible. We cannot set up a convention to make what is obligatory not permissible; at most we can treat the word ‘obligatory’ as an uninterpreted symbol and then give it some meaning other than its usual one. And in the same way we cannot set up a convention to make implication non-reflexive; we can at most treat the word ‘If’ as an uninterpreted symbol and then give it a meaning other than its usual one. Cpp is no doubt ‘true by convention’ (if it is in order to talk about truth at all in this context) in the particular T-formula game that we have been playing; for it is derivable, be the rules we have chosen to adopt, from the T-formulae by which in this game C is in effect defined. But whether ‘If – then –’ (or ‘Not both – and not[13] –’) is an appropriate verbal translation of the symbol C depends in part on whether ‘If p then p’ is something which in fact always holds whether we like it or not.

It might be said, again, that ‘pure’ logic, by contrast with ‘applied’, is not only capable of being represented by a T-formula game, but is drawn upon in our actual playing of any T-formula game whatever. For example, in the course of a T-formula game we may argue that since any P-formula is substitutable for p in CLpp, and Mp is a P-formula, it follows that Mp is substitutable for p in CLpp; and here we are availing ourselves of a principle which could itself be {22} expressed by the T-formula CΠxCϕxψxCϕxψx (“If whatever ϕ’s ψ’s, then if x ϕ’s x ψ’s”), derivable by applying Π2 to Cpp with the substitution p/Cϕxψx. But if the propositional calculus and the theory of quantifiers are thus not only representable by T-formula games but presupposed in the playing of them, may not this also be said of, say, the logic of obligation? It is, in fact, one of the main contentions of the present paper that there is an ethics of T-formula game-construction which must be respected even when what we are formalising is the logic of obligation – this ‘station’ has its ‘duties’, and being a moralist is no excuse for not discharging them. Returning now to the details of our formalisation, let us consider what happens when we adjoin symbols, rules and axioms for the logic of obligation not merely to the propositional calculus as based on 1, 2 and 3 (with substitution and detachment), but to this supplemented by some basis for modal logic, say 26, 27, Df.M and RL (i.e. the weaker basis mentioned at the end of the last section). In a combined modal and ‘deontic’ logic of this sort, RO could be replaced by the axiom

32. CLCpqCOpOq

“If the performance of p necessarily implies the performance of q, then if p is obligatory so must q be”. From this basis it is not difficult to prove

33. CNMKpNqCPpPq

“If p cannot be done without doing q, then if p is permissible q is permissible”. This was the first principle appealed to in our proof of Hintikka’s theorem, and this proof can be easily formalised if to what we now have we add quantifiers with their rules and with the further axiom.

34. NΠqPq

“Not everything is permissible” (literally “It is not the case that {23} for all q it is permissible that q”). The proof is as follows:

1. CCpqCCqrCpr

[13] Prior has omitted the second ‘not’ in error.

9. CCpCqrCKpqr
20. CCpqCNqNp
22. CNMpNMKpq
33. CNMKpNqCPpPq
34. NΠqPq
1 p/NMp, q/NMKpNq, r/CPpPq = C22 q/Nq – C 33 – 35
35. CNMpCPpPq
9 p/NMp, q/Pp, r/Pq = C 35 – 36
36. CK NMpPpPq
36 x Π2q = 37
37. CK NMpPpΠqPq[14]
20 p/KNMpPp, q/ΠqPq = C 37 – C34 – 38
38. NKNMpPp

The conclusion reads "Not both it is not possible that p and it is permissible that p", i.e. nothing that is impossible is permissible.

So we are back where we began. What are we to make of Hintikka's conclusion now? There are some paradoxes arising in the logic of obligation which are a clear signal to revise our premises – we are to begin with inclined to lay down something as a general law, but its consequences make it clear that this inclination is mistaken [16]. But not many would wish to take this line here. A more common reaction to Hintikka's theorem I have found to be the following: The operators 'It is obligatory that -' and 'It is permissible that -' only have meaning when they take as arguments those statements which describe the performance of an action; but nothing is rightly called an action if it is something which is impossible to do. Hence if NMp is true, p is {24} not the sort of symbol to which O and P may be prefixed (assuming M, O and P to be interpreted in the usual ways). Is it possible to revise this part of our T-formula game in such a way as to meet this objection (in effect) to its suitableness for representing the logic of obligation? We can at all events try. Let us again introduce K-symbols O and P, but along with these let us introduce a new set of V-symbols a, b, and c, and let us introduce the general notion of an A-formula, by the following rules

(i) the V-symbols a, b and c are A-formulae, and if α and β are any A-formulae then Nα,Cαβ and Kαβ are A-formulae.
(ii) In any T-formula, the V-symbols a, b, and c may be replaced (systematically) by any A-formula.
(iii) All A-formulae are P-formulae, and so are substitutable for p, q etc. in T-formulae, but not all P-formulae are A-formulae (so that not all P-formulae are substitutable for a, b and c in T-formulae).
(iv) If α is any A-formula, then Oα, Pα, Lα and Mα are P-formulae (though they are not A-formulae).

A-formulae may now be interpreted as standing for statements which describe actions. The rule that if α is an A-formula so is Nα expresses the fact that an omission counts as an action for the purposes of this calculus (which is reasonable, since omissions as well as actions may be obligatory and permissible). Similarly if α and β are actions, their joint performance is also an action (if α and β are

[14] Prior has omitted the Π in error. The typewritten MS has a space for the Π, and all of the other Greek symbols have been entered in handwriting.

A-formulae, so is <u>K</u>αβ), and the omission of the combination of α with the omission of β is an action (if α and β are A-formulae, so is <u>C</u>αβ). But the assertions that an action is obligatory, permissible, necessary or possible, while being genuine statements, are not 'action-statements' (so that we cannot {25} prefix <u>O</u> and <u>P</u> to them, and say that it is obligatory or permissible that something be obligatory, permissible, necessary or possible) [17]. And let us again add Df. P, 29, 30, 32 and 34, but with <u>p</u> and <u>q</u> replaced throughout by <u>a</u> and <u>b</u>, and let us also add 39. Ma expressing the fact (if it be one) that any statement describing an action must describe something that can be done. 39 will not in fact prevent the appearance as a T-formula of <u>NKNMaPa</u> (the new version of 38); but – what will perhaps serve the purpose equally well – it will completely trivialise it. "a is not possible" is still never true together with "a is not permissible"; but then it is never true together with anything (e.g. with "a is not permissible"), for it is never true at all. And this way of solving this sort of problem has some interesting precedents. For example, it is notorious that if empty terms (like 'dragon', 'mermaid', etc.) are introduced into Aristotelian syllogistic logic, it is difficult to interpret the operators so that all the Aristotelian laws (e.g. the law "If every <u>a</u> is <u>b</u>, then some a is b") will hold. But we may introduce operators A, E, I and O (this is not, of course, the I of the theory of identity or the O of the theory of obligation) with the understanding that the term-variables <u>a</u>, <u>b</u>, <u>c</u>, etc. (not our A-symbols) to which they are attached may only stand for terms which are not empty, and this understanding is reflected within the system by the T-formula

40. <u>Iaa</u>

("Some a is an <u>a</u>"), which would not hold if the term a could be empty [18]. Again, the propositional calculus may be enriched by the introduction of a V-symbol ζ such that <u>ζp</u> may represent any truth-function into which p enters (e.g. <u>Np</u>, <u>Cpq</u>). And this restriction on the type of {26} statement forming operators on statements which can be substituted for ζ is reflected in the system by the presence of the T-formula

41. <u>Cζ pCζ Npζq</u>,

which asserts that if any ζ –function holds whether its argument be <u>p</u> or not-<u>p</u>, it will hold no matter what its argument may be. (This formula can in fact be used as a single axiom for the whole propositional calculus – a much simpler single axiom than any of those, such as 18, which are possible without using ζ [19].) If ζ were replaceable by, for example, <u>M</u>, 41 would not express a law; for we can have both 'Possibly p' and 'Possibly not <u>p</u>' without its being true that everything whatever is possible. Our proposed T-formula 39 (<u>Ma</u>) might therefore be regarded as performing the same function in the logic of obligation as 40 and 41 in the calculi to which they belong. Unfortunately, however, the system we have sketched is inconsistent. For since a is an A-formula, so is <u>Na</u>, and since a and <u>Na</u> are A-formulae, so is <u>KaNa</u>, hence by substitution for a in 39, we may obtain as a T-formula

42. <u>MKaNa</u>

asserting that it is possible for an action to be at once performed and omitted. Beside being obviously false, this is inconsistent with

43. <u>NMKaNa</u>

which is easily obtainable as a T-formula from the modal part of our system. It could still be argued, however, that the trouble lies not with 39 but with our rules for forming A-formulae. But if this is so, how should the rules be amended? We could say, no doubt, that if α and β are both A-formulae, so is Kαβ, provided that β is not Nα. But this proviso alone will not exclude all that we need to exclude. It would not do, for example, if for a in 39 we could substitute {27} KKabNa ("Both both-a-and-b and not-a") and assert that an action of this form is possible. In fact, if we are to retain 39 without inconsistency, we must allow no substitution α for a such that NMα is a T-formula. And in a way it is not difficult to give a summary description of the formulae which this consideration would include. For (i) M and L are so related that NM is equivalent to LN, and NMN to L, so that NMα is a T-formula if and only if LNα is one, and NMNα is a T-formula if and only if Lα is one; and (ii) by RL, if any formula α is a T-formula, so is Lα, and by substitution in 26 (CLpp) and detachment, if Lα is a T-formula so is α, or in sum, Lα is a T-formula if and only if α is, and so LNα is one if and only if Nα is. Syllogising from (i) and (ii), we conclude that NMα is a T-formula if and only if Nα is one, and NMNα if and only if α is one; or more compendiously, NMα is a T-formula if and only if either α is the negation of some T-formula or some T-formula is the negation of α. But we cannot make it one of the rules of a T-formula game that we must not substitute for a, b or c any formula which is the negation of some T-formula, or of which some T-formula is the negation; for this presupposes that all T-formulae are already determined independently of the rules, whereas in fact the rules help to determine what are T-formulae and what are not. So it is impossible to amend our definition of A-formula in the only way which would enable us to incorporate 39 into a consistent system. And I am not sure that this is not a perfect practical example of an impossible which is for that very reason not permissible. The system of 'formalised syllogistic' with 40 (Iaa) as a T-formula could easily be inserted into a context in which a very similar problem would arise. There is a discipline sometimes called 'ontology' {28} which is in effect the logic of statements constructed out of common nouns, regardless of whether these nouns have application or not [20]. Within this discipline, it is possible, given the term-symbols a and b, to construct the negative term-symbol na (for what is not an a), and the conjunctive term symbol kab (for what is at once an a and a b). We may also introduce into ontology, by a suitable abbreviative definition, the form exa, interpretable as 'An a exists', and of course

44. Nexkana

("An a-and-non-a does not exist") is a T-formula. If we introduce the syllogistic operator I into this discipline, with 40 still among its defining axioms, we shall be able to obtain by substitution (since the negation of any term is a term, and the conjunction of any two terms is a term)

45. Ikanakana

("Some a-and-non-a is an a-and-non-a"). And if I is interpreted in the ordinary way, this will be inconsistent with 44. However, we could overcome this by a device analogous to our introduction of A-symbols. For we could introduce into ontology a set of V-symbols called E-symbols, say x, y and z, these being substitutable for the ordinary term-variables of ontology, but without the converse substitution being permissible, and without any provision for the formation of negative or conjunctive E-symbols; and we could lay it down that the Aristotelian operators A, E, I and O form P-symbols when attached to pairs of E-symbols, but not otherwise. It is obviously the rule for introducing complex term-symbols that produces the trouble with Iaa, just as it is the rule for introducing complex A-symbols which gives rise to the inconsistency with Ma in the logic of obligation; and in the former case we can avoid the trouble by just not introducing complex term-symbols of the sort {29} to which the syllogistic operators attach. But without complex A-symbols the logic of obligation would be very barren indeed.

It is time we looked again at the assumption, which our 39 was supposed to reflect, that a statement expressing an impossibility cannot be a description of an 'action'. The view that 'It is obligatory that' and 'It is permissible that' only form statements when prefixed to statements which are descriptive of actions is a plausible one, though even this much should not pass without question. I am myself more inclined to say that "It is not permissible that two and two should be four" is a false statement" than that it is not a statement at all, and I certainly think that if in formalising the logic of obligation we are to confine the arguments of our O and p to A-formulae, the class of A-formulae ought to be not more narrowly but more broadly defined than it is in (i) of our present sketch. Consider, for example, the formula <u>COaa</u>, "If a is obligatory, then it is done". Our rule states that if α and β are both A-formulae, so is <u>Cαβ</u>; and here only β is an A-formula, so the whole is not an A-formula, and it is not possible to construct a P-formula by prefixing O to it. Yet it seems to me that

46. <u>OCOaa</u>

"It is obligatory that if a thing be obligatory, it be done" expresses about as obvious a law as this field contains, and ought to be not only a P-formula but a T-formula in any formalisation of the logic of obligation that claims to be complete. Our (i), therefore, might well be changed to read "If either α or β is an A-formula, so is <u>Cαβ</u>" [22][15]. But however it be worked out in detail, let it be granted that 'It is obligatory that' and 'It is permissible that' only form statements if prefixed to statements descriptions of actions. Even so, is not a statement still 'descriptive of an action' if it describes an action {30} which not only is not and has not been but could not be performed? "John at once drank tea and did not drink tea" is surely an action describing statement, even though it is just as surely, even necessarily, a false one. It is out of the same box as "John drank tea", not out of the same box as "It is obligatory that John should drink tea", or "Two and two are four". In brief, even if we yield the point that the obligatoriness of an action is not an action, and that the impossibility of an action is not an action, there seems no strong intuitive reason for denying that what is impossible is an action (though of course an unperformed one) [21]. And we have seen that there is the strongest possible formal reason against it, for if we take this course we can only say that if there is any logic of obligation it is unformalisable. At this point the logician is surely bound by the oath of his calling to dig his toes in, and to be ready to 'fight on the beaches'.

A less violent but less suicidal procedure – and incidentally a more successful one, as far as eliminating Hintikka's theorem is concerned – would be to admit that it makes sense and is sometimes even true to describe an action as impossible, but to weaken our axioms 29, 30 and 32 by making each of them conditional on the possibility of the actions to which they refer [23]. That is, we replace them respectively by

47. <u>CMaCOaPa</u>,
48. <u>CMaCMbCOCabCOaOb</u>, and
49. <u>CMaCLCabCOaOb</u>

[15] 15 Prior's note [22] precedes note [21] in the typewritten original because Prior apparently added it handwritten first after the typewritten manuscript was finished.

This change would not merely trivialise Hintikka's theorem but {31} definitely rid us of it, and replace it by a formula which is trivial even when 39 (Ma) is rejected, namely CMaNKNMaPa. This asserts not that nothing is both impossible and permissible, but only the truism that what is possible is not both impossible and permissible. And with this solution, I suspect very many will be entirely satisfied – even all, except anti-logical extremists who still want to say that the ascription of obligatoriness, permissibility and unlawfulness to what is impossible is not only untrue but 'does not make sense' (i.e. is not a genuine ascribing); and counter-extremists like myself who do not want to lose Hintikka's theorem anyway.

What weighs with this last minority is not merely that 47, 48 and 49 are longer and clumsier axioms than 29, 30 and 32. There is a further point which can only be made plain when Hintikka's theorem is stated in the form which he originally gave to it. We have presented it here as a T-formula in a mixed logic of obligation and modality: but Hintikka himself (as reported by von Wright) presented it as a derivative rule of inference in a logic of obligation without modality, namely the rule that if Nα is a T-formula, so is NPα. And his proof, instead of making use of our 32 (which contains the modal operator L), uses a rule similar to our RO. In our system, with RO ("If Cαβ is a T-formula, so is COαOβ") now stated with the proviso "if α and β are A-formulae", there is a similar proof. For in this system NPα can always be derived from Nα (provided that α is an A-formula) as follows:

1. Nα
50. CNpCpq
20. CCpqCNpNp
34. NΠbPb
    50 p/α, q/b = C1 – II

{32}

    II. Cαb
    20 p/α, q/b = CII – III
III. CNbNα
    III x RO = IV
IV. CONbONα
    20 p/ONb, q/ONα x Df.P = CIV –V
V. CPαPb
    V x Π2b = VI
VI. CPαΠbPb
    20 p/Pα, q/ΠbPb = C VI – C 34 –VII
VII. NPα

(Here 50 is a well-known law of the propositional calculus which makes sense if we remember that it amounts to "If not p then not both p and not q". And to ensure, the applicability of Π2 in obtaining VI, if b occurs freely in α replace it in II by some action-variable which does not). The net effect of this, as of the T-formula 38, is that all such formulae as "It is not permissible at once to do a and omit it" (NPKaNa), "It is not permissible at once to do both-a-and-b and to omit a" (NPKKabNa), etc. are T-formulae.

Hintikka's rule, as we may call it in this form, can be eliminated by introducing modal operators and modifying RO still further to "If Cαβ is a T-formula, so is CMαCOαOβ (provided that α and β are A-formulae)". But in the absence of any strong intuitive reason for rejecting Hintikka's result (and there is none, though of course there is none for accepting it either), the logician has a quite overwhelming motive for doing nothing of the kind, and for not merely putting up with Hintikka's result but positively clinging to it. For thus considered as a rule, it leads easily to the simpler rule that if α is a T-formula {33} so is Oα (for if α is a T-formula, NNα will follow from this by the law of double negation: from this, by Hintikka's rule, NPNα; and from this, Oα) [24]; and since we have this rule anyway, we can radically simplify our first formalisation of the logic of obligation by laying it down at the start instead of RO, and deriving from it both RO and the 'axiom' NΠqPq. For given the new rule ("If α is a T-formula so is Oα, provided that α is an A-formula"), if we have a T-formula Cαβ in which α and β are A-formulae, we may pass from it to OCαβ, and from this to COαOβ by 30 (COCabCOaOb). And the proof of NΠbPb would be as follows (calling the new rule RO', and drawing the law CpNNp from the propositional calculus):

7. Cpp
51. CpNNp
20. CCpqCNqNp
51 p/Caa = C 7 p/a – 52
52. NNCaa
52 x RO' = 53
53. ONNCaa
51 p/ONNCaa x Df. P = C 53 - 54
54. NPNCaa
7 p/Pb x Π1b = 55
55. CΠbPbPb
20 p/ΠbPb, q/PNCaa = C55 b/NCaa - C54 - 34
34. NΠbPb [25].

What we do here is in effect to argue that not everything is permissible because the failure to do-a-if-we-do-a is not permissible, this in turn being not permissible because it is not possible. {34} The proof we have just given does not, of course, represent how we normally learn that not everything is lawful – but we have gone into all that before, have we not? Psychologically, it is not 34 but Hintikka's rule that requires proof. In this, it is like the identity of indiscernibles (Proposition 17) and the equivalence of Np and CpΠpp; but, as in those cases, the proof can be given, and once we have it, we can place the rule wherever it will serve our systematising best, and that happens to be at the start. To ask a logician to do anything else with it – above all, to ask him to throw it away – is to ask him to turn aside from the one opportunity that has yet offered itself of effecting the same kind of spectacular simplification in this field that has marked his progress in others. It is asking him to do second-rate work; and to an honest craftsman that is the same as asking him to down tools and go home.

{35}

Notes and references

1. Dr. H.O. Pappe[16] has pointed out to me that if we assume the principle that what it is wrong to do it is wrong to attempt, it follows from the unlawfulness of the impossible that attempting

[16] Helmut Otto Pappe (born 1907) was educated in Germany, escaped in 1939 the Nazi regime and lived for over 10 years in New Zealand as an industrial machinery seller. From 1949 he applied for university posts in Australia and England and was eventually appointed at the University of Sussex.

the impossible is unlawful also. We can still, however, make for the man who attempts the impossible without knowing that it is impossible the same allowance that we always make for a man who acts in ignorance of some morally relevant fact; and the sin, if it be such, of attempting what is known to be impossible would seem to be itself impossible, for if we know a thing to be impossible, no action or omission that we perform can be envisaged as a step towards its achievement. Alternatively it might be said that attempting what is known to be impossible is not impossible, but that it is not unreasonable to condemn it, for there is in it a kind of proud rebellion against 'the nature of things' – something of the spirit, let us say, of Ahab in Moby Dick.[17] (On this view, of course, we have deduced a significant moral precept from the nature of obligation). Or we might deny the additional premiss suggested by Dr. Pappe. These are lines of discussion which clearly need to be pursued further; but they will be neglected here.

2. G.H. von Wright, An Essay in Modal Logic (Amsterdam 1951) pp. 38–39[18].
3. These three formulae are used as axioms for the propositional calculus in C and N by J. Łukasiewicz (Aristotle's Syllogistic, Oxford 1951, §23)[19]. The symbolism I have used for representing proofs is a modification of Łukasiewicz's taken from I.M. Bocheński ("On the Categorical Syllogism", Dominican Studies I, {36} 1948, and Précis de Logique Mathématique, Bussum, 1949, §8)[20].
4. This is substantially one of a group of interpretations discussed by Ivo Thomas in Dominican Studies IV (1951), pp. 77–78.[21]
5. Cf. Łukasiewicz, op.cit., §24.
6. This definition of 'Not' is due to C.S. Peirce (Collected Papers, 4.402 & 454.)[22]
7. In Łukasiewicz's paper "The Shortest Axiom of the Implicational Calculus of Propositions" (Proceedings of the Royal Irish Academy, 52 A 3, 1948)[23] this axiom is shown to be sufficient for that part of the propositional calculus in which no operators but C occur; and by applying Π1 to Cpp we obtain CΠppp, which is known to yield the full propositional calculus if subjoined to a complete basis for its 'implicational' part, with the definition of Np as CpΠpp.

---

[17] In Herman Melville's 1851 novel, Moby Dick, Melville (1851) Ahab is the captain of the whaleship Pequod. Moby Dick is the name of a legendary white whale of enormous size and ferocity that in an earlier encounter with Ahab had destroyed his ship and bit one of his legs off. On the Pequod Ahab is trying the impossible, namely to take revenge on Moby Dick and kill a creature of supernatural abilities that no mortal man is able to catch and kill.

[18] In the passage to which Prior refers, von Wright writes: "It should, however, be observed that if there really existed an act, say A , which is such that P(A &~A) expresses a true proposition, then every act would be permitted. For, since A &~A and A &~A & B are names of acts which necessarily have the same performance-value, we should in virtue of the Principle of p-Extensionality, have P(A &~A & B) and from this PB follows. The assumption that there existed an act, the tautology-act of which is not obligatory, would then lead to 'moral anarchy' or 'moral nihilism'." In a footnote referring to this passage von Wright writes: "I am indebted for this observation to Mr. Hintikka." See von Wright (1951).

[19] See Łukasiewicz (1951a).

[20] See Bocheński (1948, 1949).

[21] See Thomas (1951).

[22] See Peirce (1933). Prior refers to *Collected Papers*, sections 4.402 and 4.454. These sections are in included in Peirce's discussion of existential graphs.

[23] See Łukasiewicz (1948).

8. A comment may be in order about the use here of the word ‘entails’. The C of our calculus of course expresses something much weaker than entailment, but we may take it that α entails β if Cαβ is a T-formula.
9. This is a characteristic thesis of the Lewis system S5, to which the system here given is equivalent.

10. For a deduction of this thesis in a different but equivalent system, and a discussion of the questions it raises, see my “Modality, Quantification and Time” (forthcoming)[24] [25].
11. Cf. my Formal Logic (Oxford, 1955) pp. 205–207[26].
12. This system is due to R. Feys, in ‘Les systèmes formalisés des modalités aristotéliciennes’, Revue Philosophique de Louvain Nov. 1950, note 13.[27]See also B. Sobociński, “On a Modal System of Feys-von Wright”, Journal of Computing Systems, July 1953[28].
13. This view is taken in Jonathan Bennett’s “Iterated Modalities”, {37} Philsophical Quarterly[29] [30].
14. At this point the system of deontic logic here sketched differs from that of von Wright, as von Wright’s deontic operators form statements by being attached, not to other statements, but to predicates. (“An Essay in Modal Logic”, pp. 40–41). It is a simple matter, however, to interpret von Wright’s system within the present one, provided that we introduce predicate-variables ϕ, ψ etc. and name-variables x, y, etc. The form ϕx could then be read either as “The individual x is ϕ–ing” or as “The individual act x is an act of ϕ-ing”, depending on how we wish to proceed with further systematisation (for present purposes, either interpretation will do). Similarly <u>Oϕx</u> may be read either as “The individual x ought to be ϕ-ing” or as “The individual act x (the next thing I do, say) ought to be an act of ϕ-ing”. And using O’ for von Wright’s obligation-operator, we may treat his form O’ϕ (“ϕ-ing is obligatory”) simply as an abbreviation for <u>Oϕx</u>. He is able to ignore the individual upon whom it is obligatory to ϕ, or the individual act that is or ought to be one of ϕ-ing, because in the sorts of T-formulae which he considers we can take it that the same individual or action is intended throughout the formula. We assert, for example, to his <u>CO’NϕNO’ϕ</u>, “If it is obligatory not to ϕ then it is not obligatory to ϕ”, because we assume some one individual x as intended throughout, and read it as “If x has an obligation not to ϕ then he has no obligation to ϕ”, or we may assume some action x (my next, say) as intended throughout, and read it as “If it be morally necessary that what I next do be not a ϕ-ing, then it is not morally necessary that it be a ϕ-ing”. Von Wright’s formalisation thus covers just that fragment of the logic of {38} obligation which is capable of being thus abbreviated, just as his ‘quantified logic of properties’ covers that fragment of the theory of quantifiers in which we may work with a similar abbreviation (of <u>Σxϕx</u> to a form that we could write as <u>Σ’ϕ</u>, and of <u>Πxϕx</u> to Π’ϕ).

---

[24] The article became published under the somewhat altered title “Modality and quantification in S5” in *The Journal of Symbolic Logic*, vol. 21, March 1956, pp. 60–62.
[25] See Prior (1956).
[26] See Prior (1955b).
[27] See Feys (1950).
[28] See Sobociński (1953).
[29] The complete reference is *The Philosophical Quarterly*, Vol. 5, No. 18, Jan. 1955, pp. 45–56.
[30] See Bennett (1955).

Professor G.E. Hughes[31] has pointed out to me that if, in the logic of obligation, we proceed in vonWright's manner, we will sooner or later find it necessary to supplement his deontic operators by a 'performance-operator', say D, with 'Dϕ' read as "ϕ-in is done". We would need this in order to formulate for example, the principle that if a certain action (in von Wright's sense of 'action') is obligatory, then if the doing of that would imply that a second thing is obligatory, then that second thing is obligatory too. Using O' and D, this would be
CO'ϕCCDϕO'ψO'ψ

But in the symbolism used here we can express this quite simply as:
COpCCpOqOq

15. Some examples are given in my Formal Logic, pp.[32] (One of them is the formula just given.)
16. An example of this sort would be the law of K. Grelling discussed in my Formal Logic pp.[33]
17. In von Wright's system (see note 14), the result that the obligatoriness of an action is not itself capable of being or not being obligatory – i.e. that O'O'ϕ is not a P-formula – is secured by the mere fact that his deontic operators do not form functions of the same syntactical sort as their arguments. They do not form {39} statements out of statements, or predicates out of predicates, but statements out of predicates. O'O'ϕ is thus not a P-formula just because O' ϕ is a P-formula (and so not, as it would need to be for O'O'ϕ to be a P-formula, a predicate). In our earlier translation of O' this result is preserved. If O'ϕ is short for Oϕx, then O'O'ϕ would be short for OOϕxx; and this is not a P-formula because (i) Oϕx, being not a predicate but a statement, does not form a P-formula when attached to a second x, so that (ii) the result of prefixing O to Oϕxx is not a P-formula either. On the other hand, with our O we could construct the form OOϕx, unless we exclude this by making a syntactical distinction between A -formulae and other P-formulae. It should be noted, however, that while von Wright's procedure does eliminate such forms as O'O'ϕ, and that without an elaborate mechanism of A-formulae, it does not eliminate such forms as O'Mϕ, "Being-capable-of-ϕ-ing is obligatory". (Since M does form predicates out of predicates).
18. This is the procedure of Łukasiewicz, in his Aristotle's Syllogistic.
19. This result is due to C.A. Meredith, as reported by J. Łukasiewicz in "A System of Modal Logic", Journal of Computing Systems, July 1953[34]. See also J. Łukasiewicz, "On Variable Functions of Propositional Arguments",[35] and C.A. Meredith, "On an Extended System of Propositional Calculus" (Proceedings of the Royal Irish Academy, 54 A 2 and 3, 1951).[36] In the latter, the 'functorial' variable ζ is introduced into that version of the propositional calculus in which Np is defined as CpΠpp.
20. On this discipline see B. Sobociński, "L'Analyse de l'Antinomie Russellienne par Leśniewski", Methodos 1949 and 1950[37], {40} J. Łukasiewicz, "The Principle of Individuation". Aristotelian Society Supplementary Volume 27(1953) p. 77ff.[38]; C. Lejewski,

[31] Prior refers to G. E. Hughes (1919–1994).

[32] Prior omitted the page numbers. See pp. 224–225 of *Formal Logic* (Second Edition 1962, Oxford: Clarendon Press) is the most obvious proposal (Prior 1962).
[33] The findings of Grelling are discussed on p. 227 of *Formal Logic* (Second Edition 1962, Oxford: Clarendon Press).
[34] See Łukasiewicz (1953b).
[35] See Łukasiewicz (1951b).
[36] See Meredith (1951).
[37] See Sobociński (1949 a, b, c, 1950).
[38] See Łukasiewicz (1953a).

"Logic and Existence", British Journal for the Philosophy of Science, August 1954[39]; A.N. Prior, "English and Ontology"[40], and Formal Logic III.iii.4. On a syllogistic system with is more easily interpretable within ontology than Łukasiewicz's, Suszko's review of Słupecki, Journal of Symbolic Logic, Vol. XIII (1948), p. 166[41], and my Formal Logic, Appendix I, systems 11. 1–2.

21. The even wilder contention that self-contradictory formulae in general do not express 'propositions', is very sensibly discussed by A. Church in a review of M. Lazerowicz in the Journal of Symbolic Logic, Vol V (1940), p. 82[42], and by E. Nagel in a review of A. Ambrose, ibid, Vol. IX (1944) p. 47[43]. And on the complementary contention that tautologous formulae do not express 'propositions', see P. Henle, reviewing Lazerowicz, ibid. Vol II (1937) pp. 141–142[44].
22. This formation rule would be liberal enough to allow OCOaa to be a P-formula, and therefore, if we wished to make it so, a T-formula; but restrictive enough to prevent us from passing from OCOaa by means of 30 (COCabCOaOb) with the substitutions a/Oa, b/a, to COOaOa. For since Oa would still not be an A-formula, it would not be substitutable for a in 30, that is, OCOaa, though a genuine statement-form, would not exemplify the form OCab, and COOaOa would no more be a statement-form (P-formula) than on the narrower supposition.
23. I owe this very simple but very diplomatic suggestion to Mr. W.W. Sawyer[45]. For a somewhat similar solution to the corresponding problem about syllogistic logic, see my Formal Logic p.[46]
{41}
24. If we replace 34 (NΠbPb) by the equivalent axiom ΣbOb, we may obtain this rule from RO directly (i.e. without passing through Hintikka's), Oα being provable from α by means of RO as follows:
   I. α
   II. CpCqp
   III. ΣbOb
   II p/α, q/b = CI – IV
   IV. Cbα
   IV x RO x Σ1b = CIII - V
   V. Oα
25. The alternative formula ΣbOb is obtainable still more simply, thus:
   I. Cpp
   I p/Ob x Σ2b = II
   II. CObΣbOb

---

[39] See Lejewski (1954).
[40] See Prior (1955a).

[41] See Suszko (1948).
[42] See Church (1940).
[43] See Nagel (1944).
[44] Henle (1937).
[45] W.W. Sawyer (born 1911) was one of Prior's other colleagues from New Zealand. He was especially interested in mathematics.
[46] Prior omitted the page number. It is likely that the passage which Prior has in mind is pp. 168–169 in *Formal Logic* (Second Edition 1962, Oxford: Clarendon Press).

I p/a x RO’ = III
III. OCaa
II b/Caa = CIII – IV
IV. ΣbOb

References

Bennett, J. (1955). Iterated modalities. The Philosophical Quarterly, 5(18), 45–56.
Bocheński, I., (1948). On the categorical syllogism. Dominican Studies, I, 35–57.
Bocheński, I. (1949). Précis de Logique Mathématique. Bussum, Netherlands: F.G. Kroonder.
Church, A. (1940). Review: M. Lazerowitz, self-contradictory propositions. Journal of Symbolic Logic, 5, 81–82.
Feys, R. (1950). Les systèmes formalisés des modalités aristotéliciennes. Revue Philosophique de Louvain, 48, 478–509.
Henle, P. (1937). Review: Morris Lazerowitz, tautologies and the matrix method. Journal of Symbolic Logic, 2, 141–142.
Lejewski, C. (1954). Logic and existence. British Journal for the Philosophy of Science, 5, 104–119.
Łukasiewicz, J. (1948). The shortest axiom of the implicational calculus of propositions. Proceedings of the Royal Irish Academy, 52 A, 25–33.
Łukasiewicz, J. (1951). Aristotle’s syllogistic from the standpoint of modern formal logic. Oxford: Clarendon Press.
Łukasiewicz, J. (1951b). On variable functions of propositional arguments. Proceedings of the Royal Irish Academy, Section A: Mathematical and Physical Sciences, Vol. 54, (1951/1952), pp. 25–35.
Łukasiewicz, J. (1953a). The principle of individuation. Aristotelian Society Supplementary, 27, 77 ff.
Łukasiewicz, J. (1953b). A system of modal logic. Journal of Computing Systems, 1, 111–149.
Melville, H. (1851). Moby Dick. Oxford: Oxford University Press.
Meredith, C. (1951). On an extended system of propositional calculus. Proceedings of the Royal Irish Academy, 54, 37–47.
Nagel, E. (1944). Review: Alice Ambrose, self-contradictory suppositions. Journal of Symbolic Logic, 9, 46–47.
Peirce, C. (1933). In C. Hartshorne & P. Weiss (Eds.), Collected Papers (Vol. 4). Cambridge, MA: Harvard University Press.
Prior, A.N. (1955a). English and ontology. British Journal for the Philosophy of Science, 6, 64–65.
Prior,A.N. (1955b). Formal logic. Oxford: Clarendon Press.
Prior, A.N. (1956). Modality and quantification in S5. The Journal of Symbolic Logic, 21, 60–62.
Prior, A.N. (1962). Formal logic (2nd ed.). Oxford: Clarendon Press.
Sobociński, B. (1949a). ‘L’Analyse de l’Antinomie Russellienne par Leśniewski’. Methodos, 1, 99–104.
Sobociński, B. (1949b). ‘L’Analyse de l’Antinomie Russellienne par Leśniewski’. Methodos, 1, 220–228.
Sobociński, B. (1949c). ‘L’Analyse de l’Antinomie Russellienne par Leśniewski’. Methodos, 1, 308–316.
Sobociński, B. (1950). ‘L’Analyse de l’Antinomie Russellienne par Leśniewski’. Methodos, 2, 237–257.

Sobociński B. (1953). On a modal system of Feys-von Wright. Journal of Computing Systems, 1, 171–178.
Suszko, R. (1948). Review: Jerzy Slupecki, Uwagi o Sylogistyce Arystotelesa (Les Remarques sur la Syllogistique d'Aristote). Journal of Symbolic Logic, 13, 166.
Thomas, I. (1951). Farrago Logica. Dominican Studies, 4, 69–79.
von Wright, G. (1951). An essay in modal logic. Amsterdam: North-Holland Publishing Company.