# On a Dilemma of Conditional Obligation

Ho Ngoc Duc

Conditional obligations play an important role in normative argumentation. A large number, if not the majority of norms have a conditional form: if $p$ then it ought to be that $q$. An important class among the conditional obligations are the so-called commitments, that is, sentences of the form 'the performance of (the action described by) $p$ commits one person to perform $q$'. The concept of conditional obligation has caused deontic logicians much trouble. Many proposals were made how to formalize the notion properly, but none of them has proved satisfactory. The dilemma can be summarized as follows: it seems necessary to have a form of detachment rule so that one can proceed from conditional to unconditional obligation, if the condition is satisfied. On the other hand, if the rule of detachment is allowed, then in many cases one can deduce too much: from intuitively consistent premisses one can deduce contradictory consequences. The best-known example for such counter-intuitive results is Chisholm's contrary-to-duty paradox. The aim of my paper is to propose a simple solution to this problem. I shall show that Chisholm's contrary-to-duty paradox can be solved in a very simple way.

In monadic deontic logic, there are two natural candidates for formalizing the notion of conditional obligation. Some authors suggest the formula "$O(p \supset q)$", others the formula "$p \supset Oq$" as the formalization of the sentence: "$q$ ought to be under the condition $p$". The first formula is favored, e.g., by von Wright ([7], [8]), and the second by, e.g. Chisholm ([4]) and Weinberger ([6].) Both suggestions have their own drawbacks. It is well-known that the following formulae are valid in standard deontic logic (SDL):

(1) $Fp \supset O(p \supset q)$

(2) $Oq \supset O(p \supset q)$

Thus, if "$O(p \supset q)$" is to be read as "$p$ commits you to do $q$" then formula (1) suggests that doing something forbidden commits one to do anything, and formula (2) that anything commits one to do one's duty. One may try otherwise and use "$p \supset Oq$" to formalize the notion of commitment. However, similar problems arise: the following formulae are valid for SDL:

(3) $\neg p \supset (p \supset Oq)$

(4) $Oq \supset (p \supset Oq)$

Formula (3) seems to say that whatever is not done (or is not the case) commits one to do anything, and formula (4) says that anything commits one to do one's duty. The above formulae (1–4) are quite unobjectionable as such. More serious problems arise when one tries to formalize concrete normative circumstances adequately. The following paradox was detected by R. Chisholm ([4]) and has been discussed extensively in the literature. I shall discuss it in the form presented by Åqvist ([2]). Let us consider the following four sentences:

(I) It ought to be that John does not impregnate Suzy Mae.

(II) Not-impregnating Suzy Mae commits John to not marrying her.

(III) Impregnating Suzy Mae commits John to marry her.

(IV) John impregnates Suzy Mae.

Let "$p$" stand for the sentence "John impregnates Suzy Mae" and "$q$" for the sentence "John marries Suzy Mae". The first and the fourth sentence seem to have a canonical formalization. They do not cause any difficulty: (I) and (IV) can be formalized as follows:

(Ia) $O\neg p$

(IVa) $p$

The sentences (II) and (III) are conditional obligations. In the language of SDL, (II) can be formalized by either "$\neg p \supset O\neg q$" or "$O(\neg p \supset \neg q)$. The former formula is a logical consequence of (IVa), therefore it is not a suitable formalization of (II), because intuitively (II) is not a consequence of (IV). Analogously, from (Ia) one can prove (in SDL) the formula "$O(p \supset q)$", therefore the latter formula is not a suitable formalization of (III), because (III) is not a consequence of (I). Thus, the only suitable candidates for (II) and (III) are:

(IIa) $O(\neg p \supset \neg q)$

(IIIa) $p \supset Oq$

In standard deontic logic the following distribution principle is valid:

(5) $O(A \supset B) \supset (OA \supset OB)$

Thus, from (Ia) and (IIa) we can deduce "$O\neg q$". From (IIIa) and (IVa) we can deduce $Oq$, by detachment. From "$O\neg q$" one can infer "$\neg Oq$". Thus, one can deduce (in standard deontic logic) a contradiction from (Ia)–(IVa), although (I)–(IV) are perfectly consistent intuitively. Note that if the principle $O\neg A \supset \neg OA$ is not valid then we cannot deduce a formal contradiction from (Ia)–(IVa). But in this case

we can infer $O\neg q$ and $Oq$, which is also paradoxical, and (I)–(IV) do not seem to suggest that both $q$ and $\neg q$ are obligatory.

The discussed paradox suggests that it is not possible to formalize adequately the notion of commitment in standard monadic deontic logic. Some authors have proposed a new primitive dyadic operator $O(q/p)$ with the intended meaning "$q$ is obligatory in circumstances $p$". Unconditional obligation can be defined in the following way: let $true$ be any tautology. Then one defines:

$$Op =_{def} O(p/true)$$

With this new operator, the sentences (I)–(IV) above can be formalized as follows:

(Ib) $O\neg p$

(IIb) $O(\neg q/\neg p)$

(IIIb) $O(q/p)$

(IVb) $p$

Åqvist ([2]) has shown that (Ib)–(IVb) are consistent in his strongest dyadic system. However, the price one has to pay for this consistency is too high: there is no rule of detachment available for conditional obligation. From $O(q/p)$ and $p$ one cannot deduce $Oq$. This lack of detachment imposes a heavy restriction on the application of deontic logic: how can conditional obligations play their important role in normative argumentation if they cannot, by way of detachment, lead to unconditional obligations?

The difficulties with Chisholm's paradox leads Åqvist to the question: "We seem to feel that detachment should be possible after all. But we cannot have things both ways, can we? This is the dilemma on commitment and detachment." ([2], p. 658.)

My claim is that we can have things both ways, commitment and detachment. I shall argue that both formalizations above fail because they do not pay attention to the temporal (or situational) aspect of the sentences (I)–(IV). I shall suggest a simple temporal deontic logic and show that the discussed paradox can be solved in a simple way. In this system TDL it is possible to formalize these sentences correctly, and the intuitively valid inferences are also TDL-valid. The logic TDL turns out to be the bimodal system $S5/KD$, where $S5$ is the logic of the modality 'at any time' and $KD$ is the familiar standard logic of 'ought'.

Let us take a closer look at the sentences (I)–(IV). We consider what can be inferred from them by common-sense. We want to see which of these intuitively correct inferences are also formally correct in SDL and in dyadic deontic logic, and which of the formally correct inferences are unintuitive from the standpoint of common-sense. Intuitively, from (III) and (IV) it follows "John ought to marry Suzy Mae", that is, $Oq$. This conclusion can be drawn from $(IIIa)$ and $(IVa)$ in SDL. However, it is not possible to infer $Oq$ from $(IIIb)$ and $(IVb)$ in Åqvist dyadic systems. On the other

hand, the set $\{(I), (II), (III), (IV)\}$ is perfectly consistent from an intuitive point of view. Thus, the sentence "John ought not to marry Suzy Mae" ($O\neg q$) should not be deducible from this set, because in case it is deducible we would have a contradiction. Clearly, it seems very odd if this inference should be valid: we would have the norm that John ought not to marry Suzy Mae, although he has impregnated her, and at the same time we have another norm saying that impregnating Suzy Mae commits John to marrying her. We saw above that $O\neg q$ is a consequence from $\{(Ia), \ldots, (IVa)\}$, making the set inconsistent. This inference is intuitively unacceptable. Thus, a solution to the paradox must make the inference based on the distribution principle (5) invalid. In the formalization in dyadic system (formulae (Ib)–(IVb),) however, $O\neg q$ is derivable from (Ib) and (IIb), in conflict with our intuitions. Thus, this formalization avoids inconsistency, but at the price that an intuitively valid inference is formally not valid, and an intuitively invalid inference is formally valid. So it is not a satisfactory solution either.

How does common-sense work to infer $Oq$ from (III) and (IV)? If we examine (III) and (IV) carefully, we find that they have very different structures: (IV) refers to a single time point (or situation.) Its structure is: at time $t$ it is the case that John impregnates Suzy Mae. On the other hand, (III) involves a temporal (or situational) quantification: (III) should be read as follows:

(IIIt) At any time, impregnating Suzy Mae commits John to marrying her.

From (IIIt) it follows: if John impregnates Suzy Mae at time $t$, then he ought to marry her (at some time later than $t$.) Because the premise "John impregnates Suzy Mae at t" is satisfied, one can deduce "(After t,) John ought to marry Suzy Mae".

For dealing with this temporal quantification formally, an easy way is to introduce a modal operator. Let '$L$' stand for 'for all time points'. Then (IIIt) can be formalized by the formula '$LO\neg p$', for example. Our temporal logic will be very simple: we do not care about the structure of time and we do not need the usual temporal operators of past and future. But it should be sufficient for our purposes. Formally, we add the operator "$L$" into the language of deontic logic. The definition of formulae is modified accordingly. We assume that the logic of "ought" is $KD$ and the logic of (temporal) necessity is $S5$. Thus, we need to extend SDL by the axioms and inference rules of $S5$ (for the operator $L$.) Thus, our temporal-deontic logic has the following axiom schemata:

(PC) All axioms of propositional logic

(O1) $O(p \supset q) \supset (Op \supset Oq)$

(O2) $Op \supset \neg O\neg p$

(L1) $L(A \supset B) \supset (LA \supset LB)$

(L2) $LA \supset A$

(L3) $LA \supset LLA$

(L4) $\neg LA \supset L\neg LA$

(MP) Modus Ponens

(OR) O-rule: if $A$ is a theorem, so is $OA$

(LR) L-rule: if $A$ is a theorem, so is $LA$

As we know from multi-modal logic, this logic $S5/KD$ is characterized by the class of models with two binary relations: an equivalence relation $R_L$ (corresponding to $L$) and a serial relation $R_O$ (corresponding to $O$). Note that a world in such a model should be interpreted now as a total state of a world at a given time.

Some words about the truth-condition of formulae involving the operator $L$. The logic of temporal necessity (in the sense of "eternity": truth at all past, present, and future time points) is $S5$. If necessity is used in the above sense, then it is often reasonable to require $R_L$ to be the universal relation, that is, $R_L = W \times W$. However, in many cases we are not interested in all possible time points, but only in those which are in some sense compatible with the immediate time point. Thus, we use $LA$ in the sense "$A$ is true at all relevant time points". To achieve the distinction between relevant and irrelevant time points we require $L_R$ to be an equivalent relation. To illustrate this point consider the following example. If we want to formalize the norm "It is forbidden to kill" then we shall use the formula "$LFq$" (where $q$ describes the action "kill"). In this case we would be tempted to require $R_L$ to be the universal relation, because all possible time points are relevant. On the other hand, the situation is different if we want to formalize the sentence "John ought not to impregnate Suzy Mae." We could formalize the sentence by means of the formula "$LO\neg p$", where "$p$" stands for "John impregnates Suzy Mae". The temporal operator is necessity, because the norm is clearly not intended to say that John ought not to impregnate Suzy Mae *now.* But it also sounds odd to say, for example, that John ought not to impregnate Suzy Mae after he has married her. Thus, the relevant situations for this norm do not exhaust all possible time points. It includes, for example, only those time points where John can impregnate Suzy Mae, where they are still unmarried, et cetera.

With the formal apparatus developed so far, let us reconsider the above example. Let us fix a time point $t$ for evaluating the sentences (I)–(IV). The sentence (IV) can be formalized by a propositional letter $p$ as before. Sentences like (I) are ambiguous: (I) can mean that John ought not to impregnate Suzy Mae at time $t$, or it can mean that John ought not to impregnate Suzy Mae at any (relevant) time. In this context, the second reading is more plausible. The two sentences (II) and (III) express conditional obligations: whenever something is the case, then some action ought to be done. This 'whenever' must be formalized by the operator $L$. Thus, we can formalize (I)-(IV) as follows:

(Ic) $LO\neg p$

(IIc) $L(\neg p \supset O\neg q)$

(IIIc) $L(p \supset Oq)$

(IVc) $p$

We easily see that the set $\{(Ic), (IIc), (IIIc), (IVc)\}$ is consistent. The construction of a $S5/KD$-model for this set of formulae is an easy exercise. Moreover, from $(IIIc)$ and $(IVc)$ we can infer $Oq$, as desired. We can also find models of the above set in which $O\neg q$ is false, that is, $O\neg q$ is not a consequence of $(Ic)$–$(IVc)$, in accordance with our informal considerations above.

My solution to the paradox of contrary-to-duty imperatives is much simpler than the other attempts to solve the problem by introducing time to deontic logic, e.g., the "tree" system by Åqvist and Hoepelman ([3]). The system TDL has a very simple syntax and a clear semantics. Moreover, it is decidable. These facts make TDL a suitable candidate for formalizing normative circumstances. My formalization of the notion of commitment is similar to that of A. R. Anderson ([1]), but the differences cannot be overlooked. First, the interpretation of the modalities are quite different. Second, Anderson operates with one single underlying modal (alethic) system, and defines both notions of conditional and unconditional obligations with the aid of this modal operator. This leads to undesirable consequences (see the remarks at the end of his paper.) Third, the problems we have to solve are not the same: Anderson wrote his paper long before Chisholm discovered the paradox of contrary-to-duty imperatives.

The formalization of conditional obligation "if $p$ then it shall be that $q$" by means of the formula $L(p \supset Oq)$ provides not only a simple solution to the dilemma of conditional obligation and detachment. It also solves other paradoxes of commitments. One can check easily that all of Anderson's adequacy criteria ([1], see also [2]) are satisfied.

Besides the paradox of contrary-to-duty imperatives there are a number of other paradoxes of deontic logic. Our system TDL contains some theorems of standard deontic logic that seem paradoxical under certain interpretations, for example, the formulae "$Op \supset O(p \vee q)$" (Ross's paradox) and "$Fp \supset F(p \wedge q)$" (paradox of the good samaritan.) Moreover, the TDL-version of the so-called amalgation law is valid: "$L(p \supset Oq) \supset L(p \wedge r \supset Oq)$". This formula is problematic because it seems to say that if "$p$" commits you to do "$q$" then "$p$ and $r$" also commit you to do "$q$". I have not discussed these paradoxes because they posses a structure which is very different from that the dilemma of conditional norms, as I have argued elsewhere ([5]). If standard deontic logic is to be interpreted as a logic of the concept of Ought-to-be (Sein-sollen), then these theorems have nothing paradoxical. In contrast, the difficulty with conditional norms arises simply because SDL does not have the resources to formalize the temporal aspect of normative statements. This difficulty

can be overcome by introducing the time factor, as in TDL. In order to solve the other paradoxes which arise from the improper interpretation of deontic logic we need another type of deotic logic, namely logics of the concept of Ought-to-do (Tun-Sollen). In a logic of this type, the connectives "not", "and", and "or", when used as action connectives to form complex actions from simpler ones, have a logical structure very different from that of the familiar truth functional connectives with the same names. Most paradoxes of SDL disappear as soon as we develop our deontic logic on the basis of a logic of action and interpret the mentioned connectives properly, namely as action connectives. The solution to the paradoxes of conditional norms, however, still depends heavily on the temporal factor. Therefore, a comprehensive treatment of the paradoxes of deontic logic must combine the solutions of the problems of both groups and must be able to treat the aspects action and time correctly. Here I cannot go into details and must refer the reader to my paper [5].

**Acknowledgements.** I acknowledge my long-standing teacher Peter Steinacker for guidance over the last several years. The present paper has also profited much from my discussions with Werner Wolff and with Professor Georg Henrik von Wright.

# References

[1] Anderson, Alan Ross, On the Logic of 'Commitment', *Philosophical Studies* 10 (1959) 23-29

[2] Åqvist, Lennart, Deontic Logic. In: *Handbook of Philosophical Logic* (ed. D. Gabbay and F. Guenthner,) Reidel, Dordrecht 1984.

[3] Åqvist, Lennart; Hoepelman, J., Some Theorems about a "Tree" System of Deontic Tense Logic. In: New Studies in Deontic Logic, ed. R. Hilpinen, Reidel, Dordrecht, 1981, pp. 187-221

[4] Chisholm, Roderick M., Contrary-to-duty Imperatives and Deontic Logic. Analysis 24 (1963) 33-36

[5] Ho Ngoc Duc, Semanticals Investigations in the Logic of Action and Norms. Typescript, Institute of Logic and Philosophy of Science, University of Leipzig, 1994

[6] Weinberger, Ota, Normenlogik und logische Bereiche. In: Deontische Logik und Semantik (ed. by A. G. Conte, R. Hilpinen and G. H. von Wright,) Athenaion, Wiesbaden, 1977, 176–212

[7] von Wright, Georg Henrik, Deontic Logic, Mind 60 (1951) 1-15. Reprinted in: Logical Studies (by G. H. von Wright,) Routledge and Kegan Paul, London 1957, pp. 58-74

[8] von Wright, Georg Henrik, Bedingungsnormen — Ein Prüfstein für die Normenlogik. In: Normen, Werte und Handlungen (by G. H. von Wright,) Suhrkamp, Frankfurt 1994, 44–55