Logique & Analyse **147-148** (1994), 295-302

AN INDEFINABILITY ARGUMENT FOR CONDITIONAL OBLIGATION*

Heikki-Pekka INNALA

The search for logical form derives its fascination from its tendency to produce surprises every now and then. A concept that seems to be unproblematic turns out to be something else altogether when we look into its semantics. Conditionals are an especially tricky subject matter in this respect. In this paper we shall argue that the conditional moral obligation of deontic logic is a very peculiar sort of conditional indeed, for it cannot be expressed in terms of a conditional and a deontic element. Our formal indefinability argument shows it to be a primitive concept, contrary to common belief.[1]

1. *Some Background*

Possible worlds semantics has greatly enhanced our understanding of deontic logic. If we simplify things a little, its development can be described as progression to more and more articulated deontic languages and their greater expressive power necessitated by a single paradox.[2] The key idea is to treat the deontic operators as restricted or relative modalities. The first systems had a monadic obligation operator *O* which was semantically interpreted as a universal quantifier over deontically perfect or ideal worlds. Then came the dyadic systems with their two-place operator *O*(/) designed to express conditional obligations or as they have been aptly called, iffy oughts. The corresponding semantic adjustment was to allow worlds with different degrees of deontic ideality into the models. Next came the tempo-

* This article is based on my Ph.D thesis "Is Deontic Logic Possible?", Victoria University of Wellington 1989. I am indebted to my supervisor Max Cresswell, and to Graham Oddie for a criticism of an earlier version of my proof. I thank the Jenny and Antti Wihuri Foundation for financial support.

[1] I.e. the belief that the problem of conditional obligation is simply the problem of finding the right conditional.

[2] This is the famous Chisholm-paradox. Nowadays there are several versions around but the original was presented in R. Chisholm: "Contrary-to-duty-imperatives and deontic logic", Analysis 24, 1963, pp. 33-36.

ral systems with their triadic operator $O_t(/)$. Semantically the introduction of a temporal subscript meant restricting the accessibility relation: deontic alternative worlds are now a subset of the worlds with duplicate histories up to but not including t.[3] And finally, state-of-the-art deontic logic has expressive resources also for the important fact that obligations are personally relative. So the operator looks like $Ox, t, (/)$. This novelty amounts to another restriction for the accessibility relation: different persons have access to different worlds even in the same situation. After this brief introduction we are now ready to state our main problem.

2. *The Problem*

The logician's quest for the logical form of conditional obligation is not of purely technical interest. As usual in philosophical logic, the technical and philosophical aspects of the problem are intertwined. Roughly, the technical part concerns (i) the symbolic (syntactic) representation of the concept and (ii) the corresponding decisions at the semantic level. The philosophical part is then about the adequacy conditions that (i) and (ii) must fulfil. By looking at the many proposals that have been put forward one cannot but end up with deontic pluralism: there are many different conditional obligations.[4] But there are also many different deontic systems with different theorems and semantics for syntactically identical obligation concepts. Let us consider the different possible roles that the basic dyadic formula $O(/)$ might have in a deontic system. The following four come to mind:

(1) We have a system belonging to the axiomatic research tradition. $O(/)$ is introduced together with some axioms for it. It remains an open question if this concept is definable in terms of monadic O plus a conditional.
(2) $O(/)$ might be used as an abbreviation for some combination of monadic O and a conditional i.e. it would be definable/expressible in

[3] This seems to be required unless we are willing to accept the necessity of the present. A pioneer temporal system is the California Theory in B. Chellas: The Logical Form of Imperatives, Perry Lane Press, Stanford 1969. Regarding applications read J. van Eck: A System of Temporally Relative Modal and Deontic Predicate Logic and Its Philosophical Applications, Doctoral dissertation, University of Groningen 1981, reprinted in Logique et Analyse 99/100, 1982. See also the discussion in R. Thomason: Combinations of Tense and Modality, in Gabbay and Guenther (eds.), The Handbook of Philosophical Logic II, Reidel, Dordrecht 1984, pp. 135-165.

[4] Practically all possible combinations of oughts and conditionals have been proposed in the literature. See e.g. Feldman (1986).

terms of them. In the object language one could have $O(/)$, but at the semantic level it would not be primitive.

(3) $O(/)$ might be introduced as a semantic primitive but its semantics might still be equivalent to a semantics for a concept built from monadic O + conditional.

(4) $O(/)$ is introduced with a semantics that is inexpressible with the aforesaid combination. The question still remains whether there is some new not yet discovered conditional that might do the job.

These considerations should make it clear that we cannot get very far with the axiomatic method. Indefinability questions require semantic methods. It is important to realize that due to the philosophical adequacy conditions we may be able to avoid the threat of relativism inherent in deontic pluralism. Not all of the many proposed formalizations of conditional obligation are of equal importance, some are simply more adequate than others. A candidate formula must pass the philosophical adequacy test created by the many formalization problems also known as the paradoxes of deontic logic.[5] Typically, there are two ways for a formalization to fail: either it does not preserve the logical properties (consistency, independence etc.) of the original natural language statements or else it gets the truth-values wrong. Another intuitively crucial requirement is that the proposed semantics for conditional obligation must be able to offer a workable account of detachment: under which conditions can we detach an unconditioned obligation from a conditional one? Before proceeding to our actual indefinability argument we shall survey earlier relevant work and outline the most promising semantics so far.

3. *Earlier Results and Opinions*

The idea that $O(/)$ might be primitive is not so new.[6] A look into the literature reveals that the idea has also been contested. In Counterfactuals[7] Lewis has noticed that you could get a semantics for the conditional obligation operator by replacing the nearness ordering by a deontic betterness ordering and then proceeding analogously to counterfactuals. The resulting dyadic

[5] For a typology of the paradoxes see L. Åqvist: Deontic Logic, in The Handbook of Philosophical Logic II, especially pp. 621-664.

[6] The early constructors of dyadic systems got their stimulus from probability theory, e.g. von Wright, the pioneer of modern deontic logic.

[7] Blackwell, Oxford 1973.

deontic logics belong to the same tradition as earlier semantic proposals by Danielsson[8] and van Fraassen[9] in the sense that their truth-condition says that $O(q/p)$ is true at w just in case the set of best p-worlds is a *non-empty* subset of the extension of q. Lewis argues for the primitivity of $O(/)$ in this framework which has only deontic detachment (i.e. the principle $(Op \; \& \; O(q/p)) \supset Oq$ is valid) and codifies an impersonal sense of obligation. This claim has been contested by McKinney[10] on the philosophical grounds that factual detachment (i.e. $(p \; \& \; O(q/p)) \supset Oq$) is needed to make sense of conditional obligation. The trouble with the usual dyadic systems is that if they contain both deontic and factual detachment then they lead to inconsistent formalizations of consistent situations.[11] This difficulty was avoided by introducing temporal factors to deontic systems.[12] Regarding factual detachment the move meant transition to *unalterability* detachment proposed first by Greenspan.[13] The new principle requires that p has become unalterable before detachment of Oq is allowed. Why trust in the kind of ordering semantics outlined above? Simply because its score in the adequacy test described in section 2 is so impressive. Let us look into its most promising version.

8 S. Danielsson: Preference and Obligation. Studies in the Logic of Ethics, Filosofiska Föreningen, Uppsala 1968.

9 B. van Fraassen: "The Logic of Conditional Obligation", Journal of Philosophical Logic 1, 1972, pp. 417-438.

10 A. McKinney: Conditional Obligation and Temporally Dependent Necessity: a study in conditional deontic logic, Doctoral dissertation, University of Pennsylvania 1977.

11 Chisholm-paradox once again.

12 E.g. van Eck (ibid.) has temporal constants, variables and quantifiers in his object-language. Thomason in "Deontic Logic as Founded on Tense Logic" (in R. Hilpinen ed. New Studies in Deontic Logic, Reidel, Dordrecht 1981) uses Priorean tense operators instead. However, the cost of formal elegance seems too high. Consider the following pairs of formulae: $OFp \; \& \; OF{\sim}p$, $FOp \; \& \; FO{\sim}p$. We might arrive at them by way of detachment. The problem is that such formulae do not specify the exact moments of time they pertain to. In Priorean deontic tense logic the formula $F(Op \; \& \; O{\sim}p) \equiv (FOp \; \& \; FO{\sim}p)$ would not be valid. More precisely, the invalidity of its right-to-left conjunct means that the above formulae do not reveal whether we have a case of conflicting obligations or not! (Above F is the future tense operator.)

13 In "Conditional Oughts and Hypothetical Imperatives", Journal of Philosophy 72, 1975, pp. 259-276.

4. *Feldman-semantics*

In Doing the Best We Can[14] Feldman presents his theory of moral obligation which amounts to an agent- and time-relativized version of DFL-type semantics with an explicitly chosen intrinsic value ordering IV.[15] This DBWC-system has both deontic and unalterability detachment. These improvements open the way for simple solutions to the paradoxes and many other fruitful applications.[16] The truth-conditions are as follows.

(1) The basic operator; Os,t,p is true at w iff $\exists w'(As,t,w',w$ & p is true at w' & $\sim\exists w''(As,t,w'',w$ & $\sim p$ is true at w'' & $IV(w'') \geq IV(w')))$.
(2) The conditional operator; $Os,t,(q/p)$ is true at w iff $\exists w'(As,t,w',w$ & $(p$ & $q)$ is true at w' & $\sim\exists w''(As,t,w'',w$ & $(p$ & $\sim q)$ is true at w'' & $IV(w'') \geq IV(w')))$.

As is easily seen, the first existential quantifier occurring in (2) amounts to the Limit Assumption, again familiar from counterfactual semantics. The accessibility clause ($As,t,w'',w = w''$ is accessible for s at t from w) has been relativized with respect to persons and times. This is the case for unalterability too (it is symbolized as Us,t,p).

5. *The Argument*

Our problem concerning the primitivity / indefinability / independence / inexpressibility of conditional obligation can now be solved.[17] To decide its logical status we ought to use model-theoretic methods. In that framework

[14] F. Feldman: Doing the Best We Can; an essay in informal deontic logic, Reidel, Dordrecht 1986.

[15] DFL is after Danielsson, van Fraassen and Lewis. For a slightly different approach by Hansson, Hilpinen and Åqvist, see Åqvist (1984). In our view, the ordering idea is the best way to deal with the indexicality of oughts. Different contexts presuppose different orderings and these may change the truth-value. Consider the sentence 'John ought not to pour boiling water over Katie'. In most context it has to be interpreted as a true statement of moral obligation determined by the intrinsic value ordering. But perhaps John and Katie are textile engineers testing some sort of protective clothing for housewives. Now another ordering is presupposed and the truth-value changes.

[16] See e.g. F. Feldman: "A Simpler Solution to the Paradoxes of Deontic Logic", Philosophical Perspectives 4, (ed.) J. Tomberlin, 1990.

[17] We are not aware of earlier proofs to this effect. The argument stems from my (1989) where the supplementary conjecture C is missing.

we can construct proofs that show the logical independence of axioms, and in general, of formulae.

We shall demonstrate that where $\Rightarrow$ is any conditional satisfying the following natural requirement:

(R) $(p \Rightarrow q) \therefore (p \supset q)$

we can make $O(q/p)$ true but $p \Rightarrow Oq$ false (Argument A). Furthermore, we shall show that $O(q/p)$ is not equivalent to $O(p \supset q)$ or $O(p \rightarrow q)$, where $\rightarrow$ is the Stalnaker/Lewis counterfactual conditional (Argument B). And finally, by examining the structure of Argument B, we present Conjecture C to the effect that a similar argument can be constructed for any conditional.

In order to prove our claims we must first set out the assumptions we make about the behaviour of the conditional and monadic obligation operators. We assume that (i) $O(q/p)$ has Feldman-type semantics; (ii) the monadic Oq is $O(q/T)$ i.e. there is an accessible q-world better than any accessible ~q-world or (iii) it has Hintikka's semantics i.e. q is true in every deontically perfect world.

Argument A

Here we shall construct a model where $O(q/p)$ is true but $p \Rightarrow Oq$ is false according to both (ii) and (iii). Suppose that we have a model consisting of worlds *w0*-*w4*. Let *w0* be a base world where p is true (i.e. *w1*-*w4* are its deontic alternatives in a general sense, relevant for determining the truth-values of deontic sentences at *w0*). Let *w1* be a ~p & ~q-world, *w2* a ~p & q-world, *w3* a p & q-world and *w4* a p & ~q-world. Let the deontic ordering be such that *w1* is a deontically perfect world better than *w2* which is better than *w3* which is better than *w4*. Now $O(q/p)$ is true at *w0* because there is an accessible p & q-world, *w3*, better than any accessible p & ~q-world, *w4*. Furthermore, according to both (ii) and (iii) Oq is false at *w0*. Thus $p \supset Oq$ is false at *w0*, and consequently, $p \Rightarrow Oq$ is false at *w0*. Hence this model shows that there are situations where $O(q/p)$ and $p \Rightarrow Oq$ differ with respect to their truth-value and so - by the so-called most certain principle of semantics - they have different meanings. This completes argument A.

Argument B

To prove that $O(q/p)$ is logically independent from both $O(p \supset q)$ and $O(p \rightarrow q)$ we shall construct a model where the monadic formulae are both true and the dyadic formula false. This decision is necessitated by the validity of $O(q/p) \supset O(p \supset q)$. In other words, $O(p \supset q)$ will be true in all models for $O(q/p)$. By contraposition, this fact also removes the possibility of a general model of the type A against the definability of $O(q/p)$ as $O(p \Rightarrow q)$.

Hence in the following model we cannot operate on requirement (R) only. The simplest model satisfying the requirements outlined above consists of worlds *w0-w3*. Again, let *w0* be the base world. Let *w1* be a $\sim p$ & $\sim q$-world, *w2* a p & $\sim q$-world and *w3* a p & q-world. Let the deontic ordering be such that *w1* is a deontically perfect world better than *w2* better than *w3*. And finally, let the nearness ordering for $\rightarrow$ be such that $p \rightarrow q$ is true at *w1* (e.g. the nearest p-world to *w1* is *w3*, a q-world). Now $O(p \supset q)$ and $O(p \rightarrow q)$ are true at *w0* according to both (ii) and (iii). However, $O(q/p)$ is false at *w0* because *w2* is better than *w3*. Hence our model shows that $O(q/p)$ is logically independent from both $O(p \supset q)$ and $O(p \rightarrow q)$. This completes argument B.

Conjecture C
The fact that argument B could not be formulated in the same general form as argument A leaves open a formal possibility that some other conditional (possibly a yet to be discovered intensional conditional) might suffice for the definition of $O(q/p)$. Formal proof to the contrary would require a separate B-type argument for each possible conditional. Does it follow that our arguments fail to prove what we promised? We think not. Let us return to argument B and consider its structure. The crucial question is this: what sort of semantics the possible conditional would have to have in order to guarantee the failure of our proof strategy? The answer is easy to see. Remember that first we made the monadic formulae true and only then chose the deontic ordering in a way that falsified $O(q/p)$ at *w0*. Therefore, the possible conditional would have to have a semantics based on an ordering equivalent to the deontic ordering with respect to worlds *w2* and *w3*. In other words, the truth of $(p \Rightarrow q)$ at *w1* would have to guarantee that *w3* is deontically better than *w2* in order to make $O(q/p)$ true at *w0* for any p and q. Clearly this is not plausible. The very idea of defining $O(q/p)$ presupposes that we can separate its conditional element from its deontic element, i.e. that we can use a non-deontic conditional. Based on these considerations we conjecture that B-type arguments can be found also for conditionals other than $\supset$ and $\rightarrow$. And therefore, we conclude that $O(q/p)$ really is primitive or indefinable.

6. *Conclusion*

Given that $O(q/p)$ is indefinable we may still ask if it is indispensable. Åqvist has a result to the effect that when p is permissible, we can manage without the dyadic operator.[18] And Thomason has argued that we can for-

[18] Åqvist (ibid.) pp. 655-656.

mate obligation kinematics without dyadic operators in the object-language.[19] The answer depends on our theoretical aims. We might be interested in the semantics of natural language. In that language conditional obligations of various sorts do nevertheless occur quite frequently and with varying semantics. The answer depends also on the underlying ethics.[20] Our question might be whether the true theory of moral obligation can be formulated without the concept of conditional moral obligation. Many principles (perhaps even all) of deontological theories are conditional in an irreducible way. Thus Thomason's claim could be correct only in a framework like Feldman's where the particular moral facts of the form *Os,t,p* are determined by the value of accessible worlds. But even on a such axiological approach the epistemological difficulties concerning accessibility and the value ordering are such that unless we were syntactic fanatics we would better keep $O(q/p)$ at our disposal. The simple justification for including conditional obligation and detachment principles in our conceptual machinery is that they are needed to explain under which conditions our obligations change through time.

University of Helsinki

19 Thomason (1984) p. 154. For a counterargument, see note 12 above.

20 One of the central lessons of my (1989) is that deontic logic is not, and cannot be, a neutral tool of metaethics but is inextricably tied to ethics proper. At least two other scholars have independently come to similar conclusions; G. Sayre-McCord: "Deontic Logic and the Priority of Moral Theory", Nous 20, 1986 and C. Pigden: "Logic and the Autonomy of Ethics", Australasian Journal of Philosophy 67, 1989. My own views on this issue are presented more fully in the forthcoming papers "Against von Wright's Rationality Interpretation of Deontic Logic" and "Against Non-cognitivistic Interpretations of Deontic Logic".