The Paradoxes of Deontic Logic: Alive and Kicking

Jörg Hansen

Institut für Philosophie
Universität Leipzig
Beethovenstraße 15, D-04107 Leipzig
jhansen@uni-leipzig.de

Abstract. In a recent paper, Sven Danielsson argued that the 'original paradoxes' of deontic logic, in particular Ross's paradox and Prior's paradox of derived obligation, can be solved by restricting the modal inheritance rule. I argue that this does not solve the paradoxes.

In a recent paper, Sven Danielsson [8] argued that that the 'original paradoxes' of deontic logic, in particular Ross's paradox and Prior's paradox of derived obligation, can be solved by restricting the modal inheritance rule.[1] Since these paradoxes have troubled deontic logic for three generations, and called the whole enterprise of deontic logic into question,[2] a solution would be extremely welcome. However, I argue that Danielsson's proposal has not succeeded in solving the paradoxes.

1 Ross's Paradox

Can there be logical relations between imperatives, entities that are usually considered to be neither true nor false? Yes, answers Dubislav [9] and Jørgensen [17]: identify the indicative parallel-sentence, which describes the state that obtains when the imperative is satisfied; from this descriptive sentence derive another descriptive sentence 'in the usual manner'; then the imperative that demands what this descriptive sentence describes can be called derivable from the first imperative. Fine, says Alf Ross [24], take the imperative 'Post the letter!' then use this method to derive the imperative 'Post the letter or burn it!'

Ross's paradox appears perhaps most paradoxical in its original setting, as an argument against imperative inferences. If the inference is accepted (as it is e.g. by Hare [14]), then it must be accepted that by burning the letter some

[1] Another new proposal to restrict this rule can be found in Goble [13], where, however, the aim is a deontic logic that admits normative conflicts.

[2] Cf. Kamp [18] p. 281: "Ross's paradox and the problem of formulating conditional obligation are not only a reason to question the adequacy of this or that deontic-logical rule or axiom, but provide sufficient reason to drop the deontic-logical approach altogether."

imperative is satisfied: not the original imperative to post the letter, obviously, but the imperative derived from it. In fact, given the existence of one imperative, anything, even the worst behavior, will satisfy (an infinite number of) derived imperatives. But this seems to impair the ability of derived imperatives to direct human behavior and renders them a somehow lesser sort of imperatives.

The deontic logic formula OA, for 'it ought to be the case that A', is usually not interpreted as a norm,[3] but rather a true or false statement about the existence of a (moral or legal) obligation. The following is an axiomatization of von Wright's [29] first deontic logic:[4]

(A1) $OA \to \neg O \neg A$
(A2) $OA \,\&\, OB \to O(A \,\&\, B)$
(A3) $O(A \,\&\, B) \to OA$
(R) If $A \leftrightarrow B$ is a theorem, then $OA \leftrightarrow OB$ is a theorem.

(A3) and (R) derive the monotonicity, or inheritance, rule

(C) if $A \to B$ is a theorem, then $OA \to OB$ is a theorem,

according to which anything that is logically implied something obligatory is likewise obligatory. This makes the following deontic version of Ross's paradox a theorem:

(RP) $OA \to O(A \vee B)$.

So if it ought to be that you post the letter, it ought to be that you post it or burn it. There is much debate whether (RP) is truly paradoxical: von Wright [31] termed it "odd", but later, in his re-interpretation of deontic logic as rules for rational norm-giving [32], called it "not in the least paradoxical". Føllesdal/Hilpinen [11] think it "no more paradoxical than the fact that $p \vee q$ is a logical consequence of p", Carmo & Jones [6], in the Handbook of Philosophical Logic, call it "peculiar", and for Åqvist [3], in the same volume, it is not a serious threat but a useful reminder of "the ambiguity of normative phrases in natural language as a possible source of error and confusion".

2 Prior's Paradox

In his 'old' system of deontic logic [29], von Wright used '$O(A \to B)$', where '$\to$' is the usual arrow of material implication, to formalize the notion of conditional obligation, i.e. that doing one thing commits us to doing something else. In [23], Prior pointed out the deontic analogues of the paradoxes of strict implication, of which he thought one, namely

(PP) $O\neg A \to O(A \to B)$,

[3] But cf. Åqvist [4].
[4] I use the axiomatization and names that are employed in Danielsson [8], and take the background, i.e. alphabet, language and additional rules (modus ponens, substitution into tautologies, uniform substitution) to be as usual.

troublesome, since it can be understood as stating that the doing of what is forbidden commits us to the doing of anything whatsoever, e.g. the forbidden act of stealing commits us to committing adultery. He writes that this suggests a 'might-as-well-be-hanged-for-a-sheep-as-a-lamb morality', and that it goes against the ordinary notion of commitment, where if committing a wrong act α commits us to the (reparational) act β, the omission of β is – unlike the omission of adultery – usually considered an additional wrong.

Prior's paradox had the most dramatic consequences. Von Wright [30] considered it a 'real difficulty' and not only retracted his earlier opinion that $O(A \to B)$ is an adequate expression of conditional obligation,[5] but also introduced the first system of dyadic deontic logic, which formalizes 'it ought to be that A under condition C' as '$O(A/C)$'. The paradoxical character of Prior's paradox was later underlined by Chisholm's [7] paradox of contrary-to-duty imperatives. Chisholm presented four apparently consistent and independent statements about conditional obligations that, regardless of which formalization of monadic deontic logic is chosen, $O(A \to B)$ or $A \to OB$, cannot be formalized without their being either inconsistent or derivable from each other.

3 Danielsson's Solution

Danielsson's solution to the paradoxes requires nothing more than the temporal indexing of the deontic operator, where 'O_tA' reads as 'at the time t, it ought to be that A', and the introduction of normal modal operators for temporal necessity 'N_i' and temporal possibility 'M_i', where 'N_tA' means that at time t, A has become temporally inevitable, and 'M_tA' means that at time t, $\neg A$ has not (yet) become temporally inevitable. This is much like the temporally relative modal and deontic logic introduced by van Eck [10], but the details shall not worry us here. Danielsson's proposal is then to employ

(A6) $\quad O_i(A \,\&\, B) \;\&\; M_i(B \,\&\, \neg A) \;\&\; M_i(\neg B \,\&\, \neg A) \to O_iA$

instead of (A3), together with the uniformly temporally indexed versions of (A1), (A2) and (R). So the solution allows one to infer O_tA from $O_t(A \,\&\, B)$ only if A and B are independent, in the sense that B can both be true and false without A being true.

This solution solves Ross's paradox because (RP) was previously obtained by chaining

(1) $\quad OA \leftrightarrow O((A \vee B) \,\&\, A)$,

which is an instance of (R), and

(2) $\quad O((A \vee B) \,\&\, A) \to O(A \vee B)$,

which is an instance of (A3). But with (A3) no longer available and having only (A6) instead, for the detachment of $O_t((A \vee B) \,\&\, A) \to O_t(A \vee B)$ in

[5] Later he returned to his original formulation, calling Prior's paradox "but a variant of the Ross Paradox" in [32].

(3) $O_t((A \vee B)\ \&\ A)\ \ \&\ \ M_t(A\ \&\ \neg(A \vee B))\ \ \&\ \ M_t(\neg A\ \&\ \neg(A \vee B)) \rightarrow \rightarrow O_t(A \vee B)$

we would need to establish $M_i(A\ \&\ \neg(A \vee B))$. But $A\ \&\ \neg(A \vee B)$ is equivalent to $A\ \&\ \neg A\ \&\ \neg B$, which is clearly impossible, temporally or otherwise.

The solution of Prior's paradox works similarly. (PP) was previously obtained by chaining

(4) $O\neg A \leftrightarrow O((A \rightarrow B)\ \&\ \neg A)$,

which is an instance of (R), and

(5) $O((A \rightarrow B)\ \&\ \neg A) \rightarrow O(A \rightarrow B)$,

which is an instance of (A3). With just (A6) available, it would be necessary to detach $O_t((A \rightarrow B)\ \&\ \neg A) \rightarrow O_t(A \rightarrow B)$ from

(6) $O_t((A \rightarrow B)\ \&\ A)\ \ \&\ \ M_t(A\ \&\ \neg(A \rightarrow B))\ \ \&\ \ M_t(\neg A\ \&\ \neg(A \rightarrow B)) \rightarrow \rightarrow O_t(A \rightarrow B)$,

for which $M_t(\neg A\ \&\ \neg(A \rightarrow B))$ must be true, but $\neg A\ \&\ \neg(A \rightarrow B)$ equals $\neg A\ \&\ A\ \&\ \neg B$, which is clearly impossible.

4 The Paradoxes: Alive and Kicking

For Ross's paradox, suppose that posting the letter A is obligatory, and let B stand for my burning the letter. Furthermore, assume that I have the additional obligation to write the letter C. Hence we have

(7) $O_tA\ \ \&\ \ O_tC$,

from which we derive

(8) $O_t(A\ \&\ C)$,

by applying (A2). An application of (R) provides both (9)

(9) $O_t(((A \vee B)\ \&\ (A \vee \neg B))\ \&\ C)$, and
(10) $O_t((A \vee B)\ \&\ ((A \vee \neg B)\ \&\ C))$.

It is possible for me to write the letter without posting or burning it, and also to not write the letter (and hence not posting or burning it), i.e. we have

(11) $M_t(\neg A\ \&\ \neg B\ \&\ C)$ and
(12) $M_t(\neg A\ \&\ \neg B\ \&\ \neg C)$.

N_i and M_i are normal, so the rule of extensionality (R) holds also for M_i:

(M_i-R) If $A \leftrightarrow B$ is a theorem, then $M_iA \leftrightarrow M_iB$ is a theorem.

As is easily checked, (11) and (12) are then equivalent to

(13) $M_t(((A \vee \neg B)\ \&\ C)\ \&\ \neg(A \vee B))$ and
(14) $M_t(\neg((A \vee \neg B)\ \&\ C)\ \&\ \neg(A \vee B))$, respectively.

Applying (A6) to (10), (13) and (14) we arrive at

(15) $O_t(A \vee B)$.

So from being obliged to write and post the letter I can derive that I must post or burn the letter.

For Prior's paradox, assume I am forbidden to steal A, and let B stand for my committing adultery. Furthermore, assume that I am still obliged to write the letter C. Formalizing these assumptions we obtain

(16) $O_t\neg A \;\&\; O_tC$,

which using (A2) we can combine into

(17) $O_t(\neg A \;\&\; C)$.

Applying (R), (17) is equivalent to

(18) $O_t(((A \to B) \;\&\; (A \to \neg B)) \;\&\; C)$ and
(19) $O_t((A \to B) \;\&\; ((A \to \neg B) \;\&\; C))$

We may assume that I can be a faithful thief while either writing or not writing my letters. So we obtain

(20) $M_t(A \;\&\; \neg B \;\&\; C)$ and
(21) $M_t(A \;\&\; \neg B \;\&\; \neg C)$.

Using (M_i-R), it is easily checked that (20) and (21) are equivalent to

(22) $M_t(((A \to \neg B) \;\&\; C) \;\&\; \neg(A \to B))$ and
(23) $M_t(\neg((A \to \neg B) \;\&\; C) \;\&\; \neg(A \to B))$, respectively.

So applying (A6) to (19), (22) and (23) we arrive at

(24) $O_t(A \to B)$.

Hence from the obligation not to steal and some unrelated obligation, like having to write a letter, it can be derived that stealing commits to committing adultery, which means we are back at square one.

Is it possible to save Danielsson's proposal in spite of these results? The idea behind (A6) is that

> "we can detach the obligation that A and the obligation that B from the obligation that A-and-B provided that A and B are alternatives which are independent of each other, in the sense that each can be realized without the other, and none of them is realized just because the other is not, in the situation under consideration."

So one might try to weaken (A6) to (A6*):

(A6*) $O_i(A \;\&\; B) \;\&\; M_i(B \;\&\; \neg A) \;\&\; M_i(\neg B \;\&\; \neg A) \;\&\; M_i(B \;\&\; A) \;\&\; M_i(\neg B \;\&\; A) \to O_iA$

thus ensuring that A and B are completely independent from each other. However, even with this change, (10), (13) and (14) derive the unwanted (15) if it is

both possible to write the letter and post it, and not to post the letter but burn it, since then

(25) $M_t(((A \vee \neg B)\ \&\ C)\ \&\ (A \vee B))$ and
(26) $M_t(\neg((A \vee \neg B)\ \&\ C)\ \&\ (A \vee B))$

are true and so (A6*) can be applied. Likewise, (19), (22) and (23) continue to derive the unwanted (24) if it is both possible to refrain from stealing and write one's letters, and to be an adulterous thief, since then

(27) $M_t(((A \to \neg B)\ \&\ C)\ \&\ (A \to B))$ and
(28) $M_t(\neg((A \to \neg B)\ \&\ C)\ \&\ (A \to B))$

are both true and permit the application of (A6*). So weakening (A6) to (A6*) will not be enough to ban the paradoxes from deontic logic.

5 Further Complications

Von Wright [28], often called the founder of the subject of deontic logic, attributed that property to Leibniz [20] who wrote in 1671 that all the complications, transpositions and oppositions of the Aristotelian modal logic can be transferred to the *Iuris Modalia* of the obligatory (debitum), permitted (licitum), prohibited (illicitum) and the facultative (indifferens). Since the relation between the different modalities is such an integral part of the study of deontic logic, it is worthwhile to consider what happens to these other modalities when (A6) replaces (A3).

Permission is usually defined as the absence of a prohibition to the contrary:

(Df-P) $PA \leftrightarrow \neg O\neg A$.

(C) is equivalent to (P-C)

(P-C) If $A \to B$ is a theorem, then $PA \to PB$ is a theorem.

So being told that I am permitted to bring one carton of cigarettes and 3 quarts of liquor into the US, I can with (P-C) conclude that I am also permitted to bring 3 quarts of liquor into the US – great news for an ex-smoker. However, when (A6) replaces (A3), this leaves

(P-A6) $P_iA\ \ \&\ \ M_i(A\ \&\ B)\ \ \&\ \ M_i(A\ \&\ \neg B) \to P_i(A \vee B)$.

The following is an instance of (P-A6):

(29) $P_i(A\ \&\ B)\ \ \&\ \ M_i(A\ \&\ B)\ \ \&\ \ M_i(A\ \&\ B\ \&\ \neg B) \to P_iB$.

But in contrast to the case of the O-operator, $P_t(A\ \&\ B) \to P_tB$ cannot be detached even if A and B are independent, since $A\ \&\ B\ \&\ \neg B$ is necessarily false. So if the non-smoking drinkers want to make sure they are allowed into the US, it seems they should better bring gifts for their smoking friends.

Prohibition is an obligation to the contrary:

(Df-F) $FA \leftrightarrow O\neg A$.

(C) is equivalent to (F-C)

(F-C) If $A \to B$ is a theorem, then $FB \to FA$ is a theorem.

So according to (C), if stealing is forbidden, then it is also forbidden to steal and be a reader of Proust. When (A3) is replaced with (A6), (C) and (F-C) are no longer derivable. The following is then an instance of (A6):

(30) $F_iA \ \ \& \ \ M_i(\neg A \,\&\, A \,\&\, B) \ \ \& \ \ M_i(A \,\&\, A \,\&\, B) \to F_i(A \,\&\, B),$

but we cannot detach $F_iA \to F_i(A \,\&\, B)$ since $\neg A \,\&\, A \,\&\, B$ is impossible.

Of course this last result is just as Danielsson desires it, for he wants to exclude cases like the paradox of the good Samaritan, where the prohibition to rob includes the prohibition to help the robbed. But the price for the solution of this paradox seems too high when it affects the 'ordinary cases' as well: $FA \vee PA$ is a tautology, so not being sure about $F(A \,\&\, B)$ means believing $P(A \,\&\, B)$ might be true, so that people who know there is a prohibition to steal may wonder (since they cannot conclude the contrary) whether they are permitted to steal when they are readers of Proust.[6]

6 The Missing Paradox

Danielsson's formulation of (A6) rests on the intuition that (A3)

(A3) $O(A \,\&\, B) \to OA$

is to be retained for the 'great majority of cases' in which A and B are independent alternatives. He provides the example that when it ought to be that the door is closed and the light turned on, one should be able to conclude that the door ought to be closed. Yet (A3) has been thought paradoxical in just such cases where A and B may be logically or temporally independent, but otherwise function like interdependent commodities. This so called 'Paradox of the window' has been used by Weinberger to argue against (A3):[7] when it ought to be that I leave the window closed and play the piano, it does not follow that I must play the piano, since playing the piano when the window is not closed might even be forbidden. But (A6) permits the derivation of 'it ought to be that I play the piano' if I am able to refrain from playing the piano while the window is open or closed. Thus this paradox is not solved by Danielsson's proposal.[8]

This is not necessarily blaming Danielsson's proposal for not providing a solution to a paradox he does not consider, or perhaps does not consider paradoxical – he would, indeed, be in good company (cf. [1] p. 459). But what seems desirable is some 'objective' ground on which it can be decided which paradoxes may be dismissed as harmless and which must be taken seriously. As Åqvist remarked,

[6] And witches fighting wizards might believe themselves permitted to turn into purple dragons, when the rules say that no dragons are allowed.

[7] Cf. Weinberger [26] p. 248, [27] p. 303. I adopt the name from Stranzinger [25].

[8] Stranzingers [25] system PF solves this and the other paradoxes by dropping (A3) altogether. But such a deontic logic seems much too weak.

ordinary language seems not to be a reliable guide in normative matters – maybe lawyers have used it too abundantly for their clients' purposes. However, I think two positions can be made out that provide a clearer view of the matter. One is the intuition that Menger [21] and Bohnert [5] have first appealed to and that was later studied in an alethic modal context by Anderson [2]. According to that intuition, to state that A is obligatory is the statement that if A is not the case, something unpleasant – a sanction – will occur. From this point of view, neither Ross's paradox nor Prior's paradox seems troublesome; if not posting the letter incurs the sanction, then not posting it will (usually) do the same even if one does not additionally burn it. And if thieves are punishable, they are (usually) so even when they are faithful to their spouses. The other[9] intuition appeared e.g. in Kanger [19] and is perhaps most clearly expressed in Alchourrón & Bulygin [1]: that OA means that the truth of A is necessary to satisfy all the demands of some authority or normative system (Kanger terms it a 'welfare program for the universe of discourse'). Again, neither Ross's paradox nor Prior's appear harmful. If my posting the letter is necessary to satisfy the demands, then so must be my posting-the-letter-or-burning-it. Similarly, if I cannot fulfill all demands unless I refrain from stealing, my refraining-from-stealing-or-committing-adultery is likewise necessary to ensure norm satisfaction. And while my playing the piano will be necessary to satisfy a norm that demands that I leave the window closed and play the piano, it does not imply that my playing the piano alone is also a sufficient condition to satisfy this norm.

It seems to me that more agreement on such intuitions about the meaning of the deontic operators might reduce the need for a change in the axiomatic basis of deontic logic.

References

1. Alchourrón, C. E. and Bulygin, E., "Pragmatic Foundations for a Logic of Norms", *Rechtstheorie*, **15**, 1984, 453–464.
2. Anderson, A. R., "The Formal Analysis of Normative concepts", in: Rescher, N. (ed.), *The Logic of Decision and Action*, Pittsburgh: University Press, 1967, 147–213.
3. Åqvist, L., "Deontic Logic", in [12] 147-264.
4. Åqvist, L., "Interpretations of Deontic Logic", *Mind*, **73**, 1964, 246–253.
5. Bohnert, H. G., "The Semiotic Status of Commands", *Philosophy of Science*, **12**, 1945, 302–315.
6. Carmo, J. and Jones, A. J. I., "Deontic Logic and Contrary-to-duties", in [12] 265-343.
7. Chisholm, R. M., "Contrary-to-Duty Imperatives and Deontic Logic", *Analysis*, **24**, 1963, 33–36.
8. Danielsson, S., "Taking Ross's Paradox Seriously. A Note on the Original Problems of Deontic Logic", *Theoria*, **71**, 2005, 20–28.

[9] A closer examination of the notions of 'incurring a sanction' or 'being punishable' might reveal that these are to be identified with simply 'necessitating a failure to fulfill all of ones obligations', and so the two intuitions might be identical.

9. Dubislav, W., "Zur Unbegründbarkeit der Forderungssätze", *Theoria*, **3**, 1938, 330–342.
10. van Eck, J., "A System of Temporally Relative Modal and Deontic Predicate Logic and Its Philosophical Applications", *Logique & Analyse*, **25**, 1982, 249–290, 339–381.
11. Føllesdal, D. and Hilpinen, R., "Deontic Logic: An Introduction", in [15] 1–35.
12. Gabbay, D. M. and Guenthner, F. (eds.), *Handbook of Philosophical Logic*, 2nd ed., Dordrecht: Kluwer, 2002.
13. Goble, L., "A Logic for Deontic Dilemmas", *Journal of Applied Logic*, *forthcoming*.
14. Hare, R. M., "Some Alleged Differences Between Imperatives and Indicatives", *Mind*, **76**, 1967, 309–326.
15. Hilpinen, R. (ed.), *Deontic Logic: Introductory and Systematic Readings*, Dordrecht: Reidel, 1971.
16. Hilpinen, R. (ed.), *New Studies in Deontic Logic*, Dordrecht: Reidel, 1981.
17. Jørgensen, J., "Imperatives and Logic", *Erkenntnis*, **7**, 1938, 288–296.
18. Kamp, G., *Logik und Deontik*, Paderborn: Mentis, 2001.
19. Kanger, S., "New Foundations for Ethical Theory: Part 1", duplic., 42 p., 1957, reprinted in [15] 36–58.
20. Leibniz, G. W., "Elementa Juris Naturalis [2nd half 1671, holographic concept B (Hannover)]", in: Preussische Akademie der Wissenschaften (ed.), *Gottfried Wilhelm Leibniz: Sämtliche Schriften und Briefe*, vol. 1 of *6. Reihe: Philosophische Schriften*, Darmstadt: Otto Reichl, 1930, 465–485.
21. Menger, K., "A Logic of the Doubtful. On Optative and Imperative Logic", *Reports of a Mathematical Colloquium*, **2**, 1939, 53–64, reprinted in [22] 91–102.
22. Menger, K., *Selected Papers in Logic and Foundations, Didactics, Economics*, Dordrecht: Reidel, 1979.
23. Prior, A. N., "The Paradoxes of Derived Obligation", *Mind*, **63**, 1954, 64–65.
24. Ross, A., "Imperatives and Logic", *Theoria*, **7**, 1941, 53–71.
25. Stranzinger, R., "Ein paradoxienfreies deontisches System", in: Tammelo, I. and Schreiner, H. (eds.), *Strukturierungen und Entscheidungen im Rechtsdenken*, Wien: Springer, 1978, 183–193.
26. Weinberger, O., *Rechtslogik*, 2nd ed., Berlin: Duncker&Humblot, 1989.
27. Weinberger, O., "Logical Analysis in the Realm of Law", in: Meggle, G. (ed.), *Actions, Norms, Values. Discussions with Georg Henrik von Wright. Papers presented at a colloquium held April 1996 at the Center for Interdisciplinary Research, University of Bielefeld*, Berlin: de Gruyter, 1999, 291–304.
28. von Wright, G. H., "On the Logic of Norms and Actions", in [16] 3–35.
29. von Wright, G. H., "Deontic Logic", *Mind*, **60**, 1951, 1–15.
30. von Wright, G. H., "A Note on Deontic Logic and Derived Obligation", *Mind*, **65**, 1956, 507–509.
31. von Wright, G. H., "Deontic Logics", *American Philosophical Quarterly*, **4**, 1967, 136–143.
32. von Wright, G. H., "Norms, Truth and Logic", in: von Wright, G. H. (ed.), *Practical Reason: Philosophical Papers vol. I*, Oxford: Blackwell, 1983, 130–209.