Legal knowledge based systems
JURIX '95
Telecommunication and AI & Law

The Foundation for Legal Knowledge Systems
Editors:
J.C. Hage
T.J.M. Bench-Capon
M.J. Cohen
H.J. van den Herik

Lambèr Royakkers and Frank Dignum, The idea of obligation; or: How to interpret *O*(*p*)?, in: J.C. Hage, T.J.M. Bench-Capon, M.J. Cohen, H.J. van den Herik (eds.), *Legal knowledge based systems JURIX '95: Telecommunication and AI & Law*, Lelystad: Koninklijke Vermande, 1995, 105-112, ISBN 90 5458 252 9.

More information about the JURIX foundation and its activities can be obtained by contacting the JURIX secretariat:

Mr. C.N.J. de Vey Mestdagh
University of Groningen, Faculty of Law
Oude Kijk in 't Jatstraat 26
P.O. Box 716
9700 AS Groningen
Tel: +31 50 3635790/5433
Fax: +31 50 3635603
Email: sesam@rechten.rug.nl

THE IDEA OF OBLIGATION OR: HOW TO INTERPRET $O(p)$?[1]

Lambèr Royakkers
Center for Law and Formalization
Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands
email: l.m.m.royakkers@kub.nl

Frank Dignum
Department of Mathematics and Computer Science
Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands
email: dignum@win.tue.nl

Abstract

Most deontic logics disregard that obligations are thought of as obliging *some particular individual*. Obligatory acts are considered as impersonal: related to one and the same individual all the time. It seems impossible to express that some acts are obligated for one or for some, but not for all, individuals. In this paper, we discuss the formalizations of relativized deontic modalities of Bailhache (1991), and Herrestad and Krogh (1995). Moreover, we modify of Herrestad and Krogh's theory by introducing collective obligations.

1 Introduction

Conventional approaches to deontic logic employ impersonal deontic operators. These approaches assume an implicit reference to one and the same individual all the time; an explicit reference to this person is then unnecessary (Hintikka, 1970). A sentence such as " It is obligatory for John that the window is closed and it is obligatory for Paul that the light is on" cannot be expressed within these approaches and, hence, leads to a problem. These approaches are limited, since they cannot deal with individuals.

Another problem treated below is how we can interpret $O(p)$ differently from an obligation for one and the same individual all the time. Possible interpretations are (1) an obligation for all individuals (the *general* obligation)[2], (2) an obligation for some unspecific individual (an *unspecific* obligation), (3) a strong obligation that implies the general obligation and (4) a weak obligation that is implied by the unspecific obligation[3]. In this paper, we first investigate which axioms of the standard deontic logic hold for the distinct interpretations on the basis of the theories by Bailhache (1981, 1991) and Herrestad and Krogh (1995). Then we develop a semantical model-structure for all these interpretations, providing a systematic analysis of how these interpretations are related to each other.

Furthermore, we modify the theory by Herrestad and Krogh to avoid a serious problem, viz. a principle they do not want to hold, which creeps back into the theory by extending it with a weaker notion of the obligation than the unspecific obligation. Our solution is to define this weaker notion of the obligation as a collective obligation.

This article is structured along the following lines. In section 2, we present the syntax and the semantics of the standard deontic logic. In section 3, we discuss the

1 This research was sponsored by the foundation for Law and Public Administration (Reob), which is part of the Netherlands Organization for Scientific Research (NWO).
2 Cf. Hansson (1970).
3 Cf. Hilpinen (1973).

theory of Bailhache. The theory of Herrestad and Krogh is described and modified in section 4. The modification consists of the introduction of the notion of a *collective obligation.* In the last section, we provide some conclusions.

2 Standard deontic logic and individuals

By standard deontic logic (SDL), a modal (Kripke-style) version of the now so-called “ Old System” of Von Wright (1951), we mean the system D^* based on propositional logic and axiomatized by the rule of inference[4]

(ROM) $\quad p \rightarrow q \,/\, O(p) \rightarrow O(q)$

together with the following axiom schemata

(OC) $\quad (O(p) \quad O(q)) \rightarrow O(p \quad q)$

(ON) $\quad O(p \quad \neg p)$

(OD) $\quad \neg O(p \quad \neg p)$

($Df.P$) $\quad P(p) \equiv \neg O(\neg p)$

The semantics of this system can be given using the model structure $\boldsymbol{M} = (W, \boldsymbol{R}, V)$ consisting of three elements:

1. The set of possible worlds $W = \{w, w_1, w_2, ...\}$
2. The accessibility function $R \quad \boldsymbol{R}$, which takes a world and returns a subset of W:
   $R: W \rightarrow 2^W$.
3. A valuation function V, which assigns one of the values true or false to any proposition at a world in W.

The intuition behind the function R is that it yields the deontically ideal worlds relative to a given world. Formally:

$$\boldsymbol{M}, w \text{ J } O(p) \text{ iff } R(w) \quad [[p]] \tag{1}$$

$$\boldsymbol{M}, w \text{ J } P(p) \text{ iff } R(w) \cap [[p]] \neq \tag{2}$$

with the function $[[\ ]] \quad L \rightarrow 2^W$ and L the set of well-formed formulas (wffs) of the propositional calculus.[5]

The following constraint (which gives the schema OD)

$$R(w) \neq \quad \text{for all } w \quad W \tag{3}$$

will be added to validate the schema (OD). The truth conditions (1) and (2) are sufficient to validate the rule and all other schemata. Thus D^* is sound.[6]

On the basis of this system D^* we can give the semantics of the system D_i^* to relativize deontic operators to individuals using the following model-structure $\boldsymbol{M} = (W, \boldsymbol{R}, I, V)$, with the function $R_i \quad \boldsymbol{R}$ on W, returning the deontically ideal worlds for individual i given a world (R_i: $W \rightarrow 2^W$) and the set of individuals $I = \{i, i_1, i_2, ...\}$.

The truth conditions for O_i and P_i are defined as follows

$$\boldsymbol{M}, w \text{ J } O_i(p) \text{ iff } R_i(w) \quad [[p]] \tag{4}$$

---

[4] The system D^* is the smallest normal *KD*-system of modal logic (cf. Chellas (1980)).

[5] $[[p]] = \{w \mid P(w, p) = \mathit{true}\}$. It is easy to see that the following properties hold: $[[p \quad q]] = [[p]] \quad [[q]]$, $[[p \quad q]] = [[p]] \cap [[q]]$ and $[[\neg p]] = \overline{[[p]]}$.

[6] A system is sound iff for all wffs p it holds that if H p then J p.

$$\boldsymbol{M}, w \text{ J } P_i(p) \text{ iff } R_i(w) \cap [[p]] \neq \qquad (5)$$

The following constraint gives the schema *OD* for O_i:

$$i \quad I\text{: } R_i(w) \neq \quad \text{, for all } w \quad W. \qquad (6)$$

From the semantics of O_i and P_i, it follows that the rule and all schemata for O of the system D^* are also valid for O_i.

The question arises whether there is a relation between the individuals. Hilpinen (1973) and Bailhache (1981, 1991) want a coherent system, without conflicting obligations such as $O_i(p) \quad O_j(\neg p)$. To accomplish this, they add the following axiom to the system D_i^*:

$$O_{i_1}(p) \quad O_{i_2}(q) \quad \ldots \quad O_{i_n}(v) \to P_{i_k}(p \quad q \quad \ldots \quad v) \qquad (7)$$

which is validated by adding the following constraint

$$\cap_i \quad {}_I R_i(w) \neq \quad \text{, for all } w \quad W. \qquad (8)$$

Nowadays, this axiom (7) is controversial, because it states that there is no conflict of personal duties, which is manifestly not in line with daily life situations where we can often find conflicts between legal rules, moral codes, promises, etc. It removes the possibility to express conflicts between personal obligations of different individuals, with the result that the systems of norms must be consistent (i.e., without conflicting obligations). That is why we do not enforce the axiom (7).

With the semantics of the personal obligation and personal permission, we can formalize the general and unspecific obligation and permission:

- The *general* obligation and permission: $i \quad I$: $O_i(p)$ and $i \quad I$: $P_i(p)$;
- The *unspecific* obligation and permission: $i \quad I$: $O_i(p)$ and $i \quad I$: $P_i(p)$.

3 Bailhache

Bailhache (1981, 1991) allows two different notions of obligation in his system:

1. the general obligation $O^+(p)$ ($O^+(p) \equiv \quad i \quad I\text{: } O_i(p)$), with its dual the unspecific permission $P^+(p)$ ($P^+(p) \equiv \quad i \quad I\text{: } P_i(p)$);
2. the unspecific obligation $O^-(p)$ ($O^-(p) \equiv \quad i \quad I\text{: } O_i(p)$), with its dual the general permission $P^-(p)$ ($P^-(p) \equiv \quad i \quad I\text{: } P_i(p)$).

According to the semantics given so far, all the axioms of *SDL* hold for O^+. However, for O^- the axiom (*OC*) is not valid:

$$i \quad I\text{: } O_i(p) \qquad i \quad I\text{: } O_i(q) \not\to \quad i \quad I\text{: } O_i(p \quad q). \qquad (9)$$

This is in accordance with our intuition. For instance, " A janitor might be obliged that the floor in a building is swept clean every morning, and a financial minister might be obliged that the rate of inflation is as low as possible. That there is a person for whom it is obligatory that both the floor is swept and that the inflation rate is as low as possible, we find strange." (Herrestad and Krogh, 1995, p. 462)

Bailhache (1981) wants the following schema to hold

$$i \quad I\text{: } O_i(p) \to \quad i \quad I\text{: } P_i(p) \qquad (10)$$

which is equivalent to $O^-(p) \to P^-(p)$. We denote this schema as (OD^*). This schema is valid by adding constraint (8). Furthermore, the following four relations between O^- and O^+ are valid: $O^+(p) \to O^-(p)$, $P^-(p) \to P^+(p)$, $O^+(p) \to P^-(p)$ and $O^-(p) \to P^+(p)$.

Figure 1 depicts the logical relations between the operators O^+, O^-, P^+, P^-, O_i and P_i.

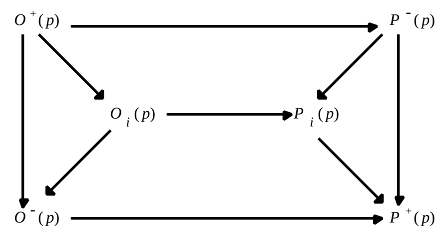

Figure 1: The logical relations between six operators

In contrast to this approach, we do not enforce the schema (OD^*) for O^- as we have discussed, by giving up the constraint (8). Then the valid properties can be summarized as in Figure 2.

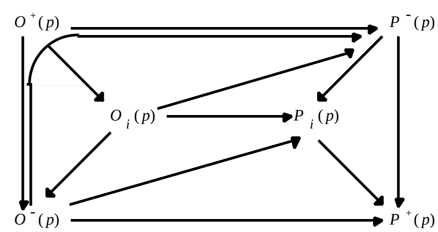

Figure 2: The valid properties of the logical operators

4 Herrestad and Krogh

Herrestad and Krogh (1995) introduce a stronger notion of the general obligation $O^+(p)$ and general permission $P^-(p)$, and a weaker notion of the unspecific obligation $O^-(p)$ and unspecific permission $P^+(p)$. With the general obligation, we can only express obligations for any particular individual in a group; with the unspecified obligation, we can only express obligations for some particular individual in the group.

However, Herrestad and Krogh do not exactly define the weaker and the stronger notion of the obligation (O^{--} and O^{++} respectively) and the permission (P^{++} and P^{--} respectively). They only aim at a situation where the following properties hold for these notions:

$O^{++}(p) \rightarrow \quad i \quad I: O_i(p)$,

$i \quad I: P_i(p) \rightarrow P^{++}(p)$,

$i \quad I: O_i(p) \rightarrow O^{--}(p)$, and

$P^{--}(p) \rightarrow \quad i \quad I: P_i(p)$.

The truth conditions for O^{++} and P^{++} are the same as (1) and (2), with the accessibility function R^{++}:

$$_{i \quad I} R_i(w) \quad R^{++}(w) \text{ for all } w \quad W. \tag{11}$$

From (6) and (11) it follows that the following constraint is valid:

$$R^{++}(w) \neq \quad \text{for all } w \quad W. \tag{12}$$

Thus, $O^{++}(p) \rightarrow P^{++}(p)$.

The truth conditions for O^{--} and P^{--} are also the same as (1) and (2), with the accessibility function R^{--}:

$$R^{--}(w) = \cap_{i \quad I} R_i(w) \text{ for all } w \quad W. \tag{13}$$

Herrestad and Krogh (1995) do not want the schema (OD^*) for O^{--}, since otherwise (10) is derivable. Thus they do not add the principle

$$R^{--}(w) \neq \quad \text{for all } w \quad W. \tag{14}$$

The rule and all other schemata of SDL are valid for O^{--}.[7]

From (11) and (13) we can derive the constraint

$$R^{--}(w) \quad R^{++}(w) \text{ for all } w \quad W, \tag{15}$$

which provides us with the principles

$$O^{++}(p) \rightarrow O^{--}(p), \text{ and} \tag{16}$$

$$P^{--}(p) \rightarrow P^{++}(p). \tag{17}$$

4.1 The problem of accepting $O^{++}(p) \rightarrow P^{--}(p)$

Just as if the general obligation implies the general permission and the weak obligation implies the weak permission, Herrestad and Krogh (1995) aim at holding these principles for O^{++} and O^{--}:

$$O^{++}(p) \rightarrow P^{--}(p), \text{ and} \tag{18}$$

$$O^{--}(p) \rightarrow P^{++}(p). \tag{19}$$

But this does not follow from the semantics given so far. We cannot add the constraint

$$\text{If } R^{++}(w) \quad [[p]] \text{ then } R^{--}(w) \cap [[p]] \neq \quad \text{for all } w \quad W \tag{20}$$

which validates the principle (18) and (by contraposing (18) and substituting $\neg p$ for p) principle (19), since now constraint (14) holds.[8] This provides the schema (OD^*) for O^{--}. Thus, the consequence of this schema (OD^*) for O^{--} is that it requires that no obligations are in conflict in order to be applicable.

According to Herrestad and Krogh (1995), the problem is due to the validity of (ON) for O^{++}, i.e., $O^{++}(p \quad \neg p)$. The solution they provide is to block the inference of $P^{--}(p)$ from $O^{++}(p)$ when p is a tautology. Instead of (20), they offer the principle[9]

$$(\Diamond p \quad \Diamond \neg p) \rightarrow (O^{++}(p) \rightarrow P^{--}(p)). \tag{21}$$

However, the solution is not quite satisfactory, because we can derive the formula $P^{--}(p) \rightarrow (O^{--}(q) \rightarrow P^{--}(q))$, since $R^{--}(w) \neq$ if $R^{--}(w) \cap [[p]] \neq$, and therefore $O^{++}(p) \rightarrow (O^{--}(q) \rightarrow P^{--}(q))$. Thus, if there is a proposition p (p is not a tautology) for which $P^{--}(p)$ or $O^{++}(p)$ is true, the schema (OD) is valid for O^{--}. We can solve this problem by stating that $P^{--}(p)$ and $O^{++}(p)$ are always false for every p which is no tautology. But the question thus rises why, for example, one introduces a stronger notion of the obligation, where this obligation actually has the value false.

We do not think that the problem is the validity of (*ON*) for O^{++}. In the following subsection, we provide a far more simple solution.

4.2 Our solution to the problem

Obligations can also aim at a group: collective obligations (cf. Royakkers and Dignum, 1995a, 1995b). An example of a collective obligation is

"John and Tom are obligated that the table is moved".

---

[7] Notice that the schema (*OC*) does hold for O^{--} but not for the unspecific obligation O^{-} (see previous section). We discuss this in more detail in subsection 4.2.

[8] Since $R^{++}(w) \quad [[p \quad \neg p]]$ is true for all $w \quad W$, then by means of (15) and modus ponens $R^{--}(w) \cap [[p \quad \neg p]] \neq$ is also true, which is equivalent with $R^{--}(w) \neq$.

[9] $\Diamond$ is the possibility sign. $\Diamond p$ is true just in case p is true at some possible world.

We cannot express this example by means of the three notions of obligation which we have discussed far: the unspecific, personal and general notion. From this example, it does not follow that

" John is obligated that the table is moved".[10]

Our solution to the problem (that the schema (OD^*) creeps back in for the weak obligation) is that we define the stronger notion of the general obligation as an obligation for all groups (a general collective obligation), and the weaker notion of the unspecific obligation as an obligation for some group (an unspecific collective obligation). We call them respectively the *strong* and the *weak* obligation:

1. the strong obligation $O^{++}(p)$ ($O^{++}(p) \equiv \quad X \quad I$: $O_X(p)$), with its dual the weak permission $P^{++}(p)$ ($P^{++}(p) \equiv \quad X \quad I$: $P_X(p)$);
2. the weak obligation $O^{--}(p)$ ($O^{--}(p) \equiv \quad X \quad I$: $O_X(p)$), with its dual the general permission $P^{--}(p)$ ($P^{--}(p) \equiv \quad X \quad I$: $P_X(p)$).

The justification of these definitions is that we claim that the strong obligation is the strongest notion of the obligation, and that the weak obligation is the weakest notion of the obligation.

The relativized obligation $O_X(p)$ means that it is obligated for the group X that p. In other words, X, as a group, has to accomplish that p. Thus, this does not mean that $O_X(p)$ is an obligation for every individual in the group X. The truth conditions for O_X and P_X are defined as follows

$$\boldsymbol{M}, w \text{ J } O_X(p) \text{ iff } R_X(w) \quad [[p]] \qquad (22)$$

$$\boldsymbol{M}, w \text{ J } P_X(p) \text{ iff } R_X(w) \cap [[p]] \neq \qquad (23)$$

with the additional constraint (which gives the schema (OD) for O_X)

$$X \quad I: R_X(w) \neq \quad , \text{ for all } w \quad W. \qquad (24)$$

The function $R_X \quad \boldsymbol{R}$ on W returns the deontically ideal worlds for the group X given a world: R_X: $W \rightarrow 2^W$[11] and $X \quad I$.

Now it follows that the truth conditions for O_X and P_X provide us with the principles (16), (17), (18) and (19), without needing an extra constraint, such as (21). Furthermore, all the schemata of the system SDL are valid for O^{++}. For O^{--}, the schemata (OD^*) and (OC) are not valid, which corresponds to our intuition.

The valid properties are summarized in Figure 3.

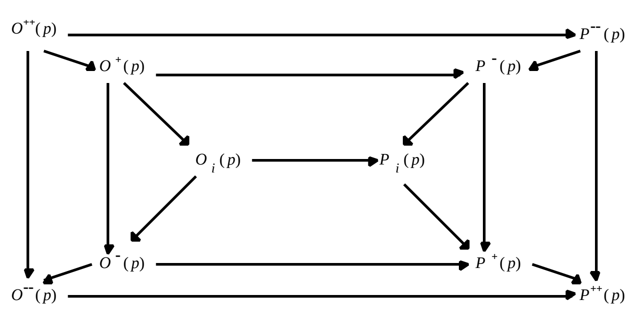

Figure 3: The valid properties

[10] Cf. Kordig (1975).

[11] We assume that $R_{\{i\}}(w) = R_i(w)$ for all $w \quad W$.

From the consideration of collective obligations as given above and the semantical model-structure that validates the properties of Figure 3, we can conclude that the given solution is satisfactory, since (OD^*) is not valid and all properties Herrestad and Krogh (1995) want to hold are valid.

4.3 The schema (OC)

A reason for adding the schema (*OC*) to a system is that " if there is a group for which it is obligated that *p*, say *X*, and there is a group for which it is obligated that *q*, say *Y*, then it is obligated for *at least* the group $X \quad Y$ to accomplish $p \quad q$."

If one wants to hold this schema, one must add the constraint

$$R_X(w) \quad R_Y(w) \text{ if } Y \quad X \text{ for all } w \quad W, \tag{25}$$

which also validates the principle[12]

$$O_X(p) \rightarrow O_{X \quad Y}(p). \tag{26}$$

Notice that now the collective obligation is defined as a (restricted) weak obligation:

$$O_X(p) \equiv \quad Y \quad X: O_Y(p). \tag{27}$$

This means that the collective obligation $O_X(p)$ is not an obligation for the group *X* as a whole (as in the previous subsection), but an obligation for some subgroup in *X*.

However, now we can derive the schema (OD^*) from (*OC*) and (*OD*). To reject the schema (OD^*), we have to reject (*OD*), if we want to keep (*OC*). This can be accomplished by giving up the constraint (24).

5 Conclusion

The result of the theory presented in this paper is a distinction between three levels of notions of obligation and permission:

- The first level is the level of the personal notions: O_i and P_i ;
- The second level is the level of the general and unspecific notions: O^+ and P^-, respectively O^- and P^+ ;
- The last level we discussed is the level of the general and unspecific collective notions: O^{++} and P^{--}, respectively O^{--} and P^{++}.

The relations between the three levels are summarized in Figure 3.

The extension by including relativized deontic modalities in the standard deontic logic provides us with the possibility to express, for example, that *p* is obligatory for an individual, but not for everyone: $O_i(p) \quad \neg \quad j \quad I: O_j(p)$, which is not expressible within a non-relativized deontic logic.

A problem arises in the theory of Herrestad and Krogh (1995) when we add constraint (20) to validate (18) and (19). However, this gives the schema (OD^*) for O^{--} and for O^-. They propose a restricted bridge principle (21) instead of (20). But we have shown that this solution is not quite satisfactory. Instead, we suggest a stricter definition of O^{++} and O^{--} by means of collective obligations, thus not modifying the schema (*ON*) for O^{++}.

Finally, we have rejected (*OC*) for O^{--} for the same reason as the rejection of (*OC*) for O^-. However, this schema is not counter-intuitive if we define the collective

[12] This principle seems paradoxial (" If John is obligated that *p*, then John and Paul are obligated that *p*") corresponding to the Ross' paradox (" If it is obligated to post the letter, then it is obligated to post or burn the letter".)

obligation as a restricted weak obligation. To keep this schema we have to give up (OD) for O^{--}, since otherwise the schema (OD^{*}) would be valid.

References

Bailhache, P. (1981). *Analytical Deontic Logic: Authorities and Addressees.* Logique et Analyse, Vol. 93, pp. 65-80.

Bailhache, P. (1991). Authorities and Addressees in Deontic Logic: Indexed Operators and Action. *ΔEON'91, Proceedings of the First International Workshop on Deontic Logic in Computer Science* (eds. J.-J.Ch. Meyer and R.J. Wieringa), pp. 72-88. Amsterdam.

Chellas, B.F. (1980). *Modal Logic, an Introduction.* Cambridge University Press, Cambridge.

Hansson, B. (1970). Deontic Logic and Different Levels of Generality. *Theoria*, Vol. 36, pp. 241-248.

Herrestad, H. and Krogh, C. (1995). Deontic Logic Relativised to Bearers and Counterparties. *Anniversary Anthology in Computer and Law* (eds. J. Bing and O. Torvund), pp. 453-522. COMPLEX - TANO.

Hilpinen, R. (1973). On the Semantics of Personal Directives. *Ajatus*, Vol. 35, pp. 140-157.

Hintikka, J. (1970). Some Main Problems of Deontic Logic. *Deontic Logic* (ed. R. Hilpinen), pp. 59-104. D. Reidel, Dordrecht.

Kordig, C.R. (1975). Relativised Deontic Modalities. *The Logical Enterprise* (eds. A.R. Anderson *et al.*), pp. 221-257. Yale University Press, New Haven.

Royakkers, L.M.M., and Dignum, F. (1995a). Actors and Norms. *Information Systems, Correctness and Reusability* (eds. R.J. Wieringa and R.B. Feenstra), pp. 225-241. World Scientific Publishing.

Royakkers, L.M.M., and Dignum, F. (1995b). Collective Obligation. *Proceedings Seventh Dutch Conference on Artificial Intelligence (NAIC'95)* (eds. J.C. Bioch and Y.H. Tan), pp. 351-360. Erasmus Universiteit Rotterdam.

Wright, G.H. von (1951). Deontic Logic. *Mind* 60, pp. 1-15.