

J.-J. Ch. Meyer, R.J. Wieringa, and F.P.M. Dignum. The Role of Deontic Logic in the Specification of Information Systems. In J. Chomicki and G. Saake (eds.) *Logics for Databases and Information Systems*, pages 71-115, Kluwer Academics Publishers, 1998.

# The Role of Deontic Logic in the Specification of Information Systems

*J.-J. Ch. Meyer*<sup>¶</sup>

Utrecht University, Dept. of Comp.Sc., P.O. Box 80.089, NL-3508 TB Utrecht

*R.J. Wieringa*<sup>¶</sup>

Free University, Amsterdam, Dept. of Comp.Sc.

*F.P.M. Dignum*

Technical University Eindhoven, Dept. of Comp.Sc.

## Abstract

In this paper we discuss the role that deontic logic plays in the specification of information systems, either because constraints on the systems directly concern norms or, and even more importantly, system constraints are considered ideal but violable (so-called ‘soft’ constraints). To overcome the traditional problems with deontic logic (the so-called paradoxes), we first state the importance of distinguishing between ought-to-be and ought-to-do constraints and next focus on the most severe paradox, the so-called Chisholm paradox, involving contrary-to-duty norms. We present a multi-modal extension of standard deontic logic (**SDL**) to represent the ought-to-be version of the Chisholm set properly. For the ought-to-do variant we employ a reduction to dynamic logic, and show how the Chisholm set can be treated adequately in this setting. Finally we discuss a way of integrating both ought-to-be and ought-to-do reasoning, enabling one to draw conclusions from ought-to-be constraints to ought-to-do ones, and show by an example the use(fulness) of this.

## 1. Introduction: Soft Constraints and Deontic Logic

### 1.1 Integrity Constraints for Information Systems

An information system is a system that stores data about a part of the real world called the *Universe of Discourse* (UoD). In order to specify an information system, we must specify a conceptual model of the UoD that fixes the meaning of the data in the information system. The conceptual model is in fact a set of meaning conventions for the symbols manipulated by the system. For example, a personnel information system stores data about the employees of a company. The UoD of the system is the set of employees of the company and a conceptual model of this set describes the properties of employees represented by the system. It is this conceptual model that determines the meaning of the data; an interpretation of the data that goes beyond the conceptual model may be valid but is not guaranteed to be correct by the system.

---

<sup>¶</sup> Partially supported by the ESPRIT III BRWG Project No. 8319 ‘MODELAGE’.

Traditionally, conceptual data models represented the types of entities that can exist in the UoD and the relationships that can exist between these types of entities. With the advent of richer specification languages for conceptual models, it becomes possible to specify various kinds of actual and desirable properties of entities in the UoD. For example, predicate logic can be used to specify that all employees are persons and that no employee can have a temporary and permanent employment contract with the same employer at the same time. Temporal logic can be used to specify that the age of a person can never decrease or that an employment contract will eventually be terminated. These properties of the UoD are (*hard*) *constraints* on the system, because they specify properties that the structure and behaviour of the data in the system must have. If we assume, for the sake of the argument, that it is not possible that an employee violates one of these properties, then a state or behaviour of the system that violates one of these properties is wrong. Hence, any property known to be true of the UoD is a constraint that must be satisfied by the information system. These properties are called *integrity constraints* for the system.

In order to be able to reason about integrity constraints, it is useful to use logic as a specification language for conceptual models. For example, order-sorted predicate logic can be used to specify taxonomic structures, temporal logic can be used to specify temporal constraints, dynamic logic to specify system transactions, etc. Once we have a logic specification of a conceptual model, we can use the inference system of the logic to derive properties of the model and we can validate the model by checking whether these properties do, as a matter of fact, hold in the UoD.

Recently, it has been recognized that it is useful to be able to specify *desirable* properties of the UoD, that may be violated by the UoD ([WMW89]). A popular example in this respect is the UoD constraint that the salary of an employee never decreases. Other examples studied in the literature include the UoD constraint that a book that is borrowed from a library should be returned within three weeks and that a bank account should have a non-negative balance. All of these properties may be violated by a UoD. This means that these properties must not be translated into constraints on the system. Rather, the system must be able to represent violations of these constraints. Representing these violations is precisely one of the functions of the system. Nevertheless, it would be useful if the system would be able to represent these desirable properties of the UoD and would be able to represent violations of these properties *as violations*. We call [such a desirable property of the UoD](#) that is used in the specification of a conceptual model a *soft constraint*, because the system must be able to represent and signal violations of these constraints. The desire to specify soft constraints motivates the use of deontic logic in the specification of conceptual models.

## 1.2 Deontic logic and violations of constraints

Deontic logic is a logic to reason about ideal versus actual states or behaviour. These can be expressed in the logic by means of operators for prohibition (F), permission (P) and obligation (O). In deontic logic one can express that ideally constraints are complied to, but actually violations may occur. Moreover, one has the ability to specify what should happen in case of a violation (in order to restore integrity or take some other compensating action, for instance). As such it is a very useful tool for the specification of ‘soft’ integrity constraints for computer systems, and information systems in particular.

For instance, consider an information system (knowledge base) for a library. Specifying such a system involves a lot of ‘soft’ constraints expressing ideal situations or behaviour of the agents involved, which may be represented by formulas with deontic operators. For example, one may use the following clause in a Library KB:

$$[(\text{borrow}(p, b)] \text{O}(\text{return}(p, b))_{\leq 3\text{weeks}}$$

expressing the ‘soft’ (violable) constraint that when a person  $p$  borrows a book  $b$ , he should return it within 3 weeks. Typically, since a library does not have any control over the borrower, it is very well conceivable that the constraint is violated. In our particular logic we shall signal such a situation of violation by means of a special propositional atom; e.g. in this particular case by the atom  $V_{\text{return}(p, b)}$ . This atom can then be employed by specifying a compensating action such as paying a fine, e.g. as follows:

$$V_{\text{return}(p, b)} \rightarrow \text{O}(\text{pay}(p, \$2, b),$$

expressing that when a person fails to fulfil his obligation to return the borrowed book, he has to pay a fine of \$2.

Sometimes also database integrity constraints *per se* are regarded as deontic (e.g. [Kwa93]):

$$\text{DB} \text{ } \text{O}(\text{IC})$$

stating that the database must satisfy (ideally satisfies) the integrity constraints IC. In effect this approach treats every IC as a ‘soft’ constraint, the violation of which can be represented, so that, when violated, integrity-recovering actions might be specified.

Recently deontic logic has indeed been taken up as a specification tool for a wide variety of software systems, ranging from the specification of fault-tolerant behaviour for use in advanced

software to electronic contracting systems for use in organisations, as well as integrity constraints in databases and legal expert systems (cf. [MW93b], in particular [WM93]).

### 1.3 The Paradoxes of Deontic Logic

(Standard) deontic logic (**SDL**) is not without problems, however: it has been plagued by many so-called paradoxes since its inception. These paradoxes signify more or less mismatches between the formal deontic concepts and their everyday use. More specifically, the paradoxes in deontic logic are logical expressions that are valid in a (or even most) well-known logical system(s) for deontic reasoning, but which are counterintuitive in a common-sense reading (cf. [Åqv84], Section II). Of course, matching theorems and intuitions is a general problem of measuring a logic (or even more generally, a formal approach) against the purpose it was devised for. (This is sometimes called the *validation* problem.) But it is remarkable that in the realm of deontic logic these problems appear to be much more serious and persistent through the years than in, for instance, other modal logics such as temporal and epistemic logics. Some of these paradoxes appear and re-appear again in the literature for several decades now without the community of researchers seemingly arriving at a consensus. What is even more surprising is the relative simplicity of the paradoxes. Most of them can be explained to a complete layman in deontic logic in a couple of minutes. Yet they have been haunting deontic logicians for many years now. To mention some of the best known ones:

- Ross's Paradox: ought-to-mail-a-letter implies ought-to-mail-a-letter-or-burn-it.
- Free Choice Paradox: allowed-to-mail-a-letter implies allowed-to-mail-a-letter-or-burn-it.
- No Conflicting Obligations: ought implies permitted
- Good Samaritan Paradox: ought-to-help-Jones-who-is-robbed implies Jones-ought-to-be-robbed
- Chisholm's Paradox: ought-to-do-A, ought-to-do-(A-implies-B), not-A implies ought-to-do-not-B and actually doing-not-A is inconsistent.

Not all paradoxes are considered equally serious. Of the small list above the contrary-to-duty-imperatives such as Chisholm's paradox are generally considered by far the most awkward and even embarrassing by deontic logicians.

When (proposing to) using deontic logic for specification of concrete systems, these should be dealt with one way or another, either by resolving the paradoxes or by explaining why they are harmless in the context of application. Our claim in this paper is that all of the deontic paradoxes in the literature are harmless, except for the ones concerning the so-called contrary-to-duty (CTD) imperatives, such as the Chisholm paradox. Moreover, CTD imperatives are omnipresent in daily life as well as in more technical circumstances. An instance of the Chisholm paradox comprises the following assertions:

1. there should be no error.

2. if there is an error, there should be a warning.
3. if there is no error, there should be no warning.
4. there is an error.

In **SDL** this set of assertions cannot be represented consistently (without one of the assertions being redundant), while intuitively the set is both consistent and nonredundant. On the other hand, it is absolutely imperative to be able to specify situations of this kind.

Another important claim in this paper is that because of their conceptual difference one should distinguish between the notions of ought-to-be ('Seinsollen') and ought-to-do ('Tunsollen'), to the extent of using really different logics for these two notions. The former notion can be employed to specify what should hold ideally in the states of a system (or, put in other words, what the ideal states of the system are), while the latter notion can express the ideal behaviour of a system.

So, when we try to solve the Chisholm paradox we first have to distinguish between ought-to-be and ought-to-do versions of this paradox. In order to represent the ought-to-be version of the Chisholm set properly we present a multi-modal extension **S5O<sub>(n)</sub>** of **SDL**. For solving the ought-to-do variant we employ a reduction to dynamic logic, and show how the (ought-to-do version of the) Chisholm set can be treated adequately in this setting. Although we advocate distinguishing between logics for ought-to-be and ought-to-do reasoning, on the other hand we claim that in practice one needs both forms and hence a way to deal with these in an integrated framework. In Section 7 we discuss such a way of integrating both ought-to-be and ought-to-do reasoning, enabling one to draw conclusions from ought-to-be constraints to ought-to-do ones, and show the use(fulness) of this by an example.

## 2. Standard Deontic Logic (SDL)

One of the first systems for deontic logic that really was a serious attempt to capture deontic reasoning was the now so-called "Old System" of Von Wright ([vW51]), of which a modal (Kripke-style) version has become known as Standard Deontic Logic (**SDL**).

**SDL** consists of the following axioms and rules:

- (K<sub>O</sub>)  $O(\varphi \rightarrow \psi) \rightarrow (O\varphi \rightarrow O\psi)$
- (D<sub>O</sub>)  $\neg O\perp$
- (N<sub>O</sub>)  $\varphi / O\varphi$
- (P)  $P\varphi \leftrightarrow \neg O\neg\varphi$
- (F)  $F\varphi \leftrightarrow O\neg\varphi$
- (Taut) the tautologies of propositional logic (or just enough of them)

(MP)  $\varphi, \varphi \rightarrow \psi / \psi$

Here  $\perp$  stands for falsum.

**SDL** has a Kripke-style modal semantics based on a set of possible worlds(, a truth assignment function of primitive propositions per possible world) and an accessibility relation associated with the O-modality (cf. e.g. [MW93a]). This accessibility relation points to “ideal” or “perfect deontic alternatives” of the world under consideration. The crux behind this is that in some possible world something (say  $\varphi$ ) is obligated, if  $\varphi$  holds in all the perfect alternatives of this world, as indicated by the accessibility relation.

So, formally these models have the following form:  $M = (S, \pi, R_O)$ , where  $S$  is the set of states,  $\pi$  is a truth assignment function, and  $R_O$  is the deontic accessibility relation, which are assumed to be serial, i.e. for all  $s \in S$  there is a  $t \in S$  such that  $R_O(s, t)$ .

The operator  $O$  is interpreted by means of the relation  $R_O$ :  $M, s \vDash O\varphi$  iff  $M, t \vDash \varphi$  for all  $t$  with  $R_O(s, t)$ . The system **SDL** can be shown sound and complete with respect to validity in this class of models via a standard argument. (The system **SDL** coincides with the system **KD** in the classification of Chellas [Che80].)

A few theorems of **SDL**:

(O $\wedge$ )  $O(\varphi \wedge \psi) \leftrightarrow (O\varphi \wedge O\psi)$

(P $\wedge$ )  $P(\varphi \wedge \psi) \rightarrow (P\varphi \wedge P\psi)$

(F $\wedge$ )  $(F\varphi \vee F\psi) \rightarrow F(\varphi \wedge \psi)$

(O $\vee$ )  $(O\varphi \vee O\psi) \rightarrow O(\varphi \vee \psi)$

(P $\vee$ )  $P(\varphi \vee \psi) \leftrightarrow (P\varphi \vee P\psi)$

(F $\vee$ )  $F(\varphi \vee \psi) \leftrightarrow (F\varphi \wedge F\psi)$

(Cfl)  $\neg(O\varphi \wedge O\neg\varphi)$

### 3. The Paradoxes of Deontic Logic

The paradoxes of deontic logic are logical expressions (in some logical language) that are *valid* in a (or even most, as is often the case) well-known logical system for deontic reasoning, but which are *counterintuitive in a common-sense reading*. Of course, there is the general problem of measuring a formal approach (logic) against the purpose it was devised for: do formal theorems in the logic match the intuition? But in the realm of deontic logic this problem seems to be more serious and persistent than in other (modal) logics such as temporal and epistemic logic (although in epistemic logic there is, for instance, the problem of logical omniscience, see, e.g., [MH95]).

Traditionally, deontic logic has been proposed and used for reasoning about ethical and legal aspects. This has resulted in a critical evaluation of the logic in its capacity of supposedly giving an adequate representation and reasoning mechanism for these applications. Originally, also the paradoxes were discovered and judged against this background. The question arises naturally whether the paradoxes are equally problematic when deontic logic is used for the specification of advanced information systems, e.g. knowledge based or intelligent systems where norms (or rather normative versus nonnormative behaviour) play a role. We will argue that the computer science (or AI) view may be different in at least two ways:

- it may offer a *dynamic* perspective, taking notions of action and time into (more explicit) account.
- it may also offer a *pragmatic* perspective: computer science has a definite engineering aspect and therefore computer scientists are interested in things that work (sufficiently) rather than deep philosophical issues.

With respect to the former, computer science may really give some new insight and help the traditional area of deontic logic by offering new interpretations and logics. We shall in fact our ought-to-do logic on one of these in Section 6. With respect to the latter, we admit that computer scientists will mainly help themselves, (ab?)using deontic logic the way they need, not being bothered by old and profound questions. So, on the one hand, one has deontic logicians being plagued by long-standing paradoxes without seemingly arriving at a consensus, while on the other hand, one has computer scientists wanting to apply deontic logic to their field of interest. What should be their attitude towards deontic logic, and the paradoxes in particular? Do these present insurmountable problems for them, too? Let us first have a look at some of the most infamous paradoxes from the literature.

### 3.1. Some Well-Known Paradoxes

The following is a list of the most well-known paradoxes from the deontic logic literature:

- |   |   |
|---|---|
| 1. Empty normative system               | $O\phi$ ( $\phi$ a tautology), e.g. $O(\phi \vee \neg\phi)$   |
| 2. Ross' paradox                        | $O\phi \rightarrow O(\phi \vee \psi)$   |
| 3. No free choice permission            | $(P\phi \vee P\psi) \leftrightarrow P(\phi \vee \psi)$  |
| 4. Penitent's paradox                   | $F\phi \rightarrow F(\phi \wedge \psi)$   |
| 5. Good Samaritan paradox               | $\phi \rightarrow \psi \quad f \quad O\phi \rightarrow O\psi$   |
| 6. Chisholm's paradox                   | $(O\phi \wedge O(\phi \rightarrow \psi) \wedge (\neg\phi \rightarrow O\neg\psi) \wedge \neg\phi) \rightarrow \perp$ |
| 7. Forrester's paradox of gentle murder | $(F\phi \wedge (\phi \rightarrow O\psi) \wedge (\psi \rightarrow \phi) \wedge \phi) \quad f \quad \perp$            |
| 8. Conflicting obligations              | $\neg(O\phi \wedge O\neg\phi)$ , or equivalently,<br>$O\phi \rightarrow P\phi$                                      |
| 9. Derived obligation                   | $O\phi \rightarrow O(\psi \rightarrow \phi)$  |

	$F\psi \rightarrow O(\psi \rightarrow \phi)$
	$\neg\phi \rightarrow (\phi \rightarrow O\psi)$
10. Deontic detachment	$(O\phi \wedge O(\phi \rightarrow \psi)) \rightarrow O\psi$
11. Kant's ought implies can	$O\phi \rightarrow \phi$
12. Epistemic obligation	$OK\phi \rightarrow O\phi$

**Remarks.** 1. says that every tautology is obligated. This is by some authors (including Von Wright [vW51]) viewed as an undesirable property of a deontic logic, since necessary and therefore inevitable things cannot be obligated in a true sense. 2.'s strangeness is often illustrated with an instance like "if one is obliged to mail a letter, one is obliged to mail the letter or burn it." This sounds paradoxical in ordinary language usage. 3. implies that if  $\phi$  is permitted but  $\psi$  is not, nevertheless " $\phi$  or  $\psi$ " is permitted. This is sometimes felt as counter-intuitive since " $\phi$  or  $\psi$  is permitted" suggests that one is free to choose either doing  $\phi$  or doing  $\psi$ . 4. is analogous to 3. It has as an instance that if someone is forbidden to do a crime, one is also forbidden to do a crime and do penitence. 5. says that every logical consequence of something that is obligated is obligated itself. This is amenable to debate, as in an instance like "if Jones helps Smith who has been injured implies Smith has been injured, so if it is obliged that Jones helps Smith who has been injured, it is obliged that Smith has been injured". 6. is an abstract version of the set "if you are obliged to go to a party; it is obliged that, if you go, you tell you're coming; but if you do not go, you are obliged to not tell you're coming; and, in fact, you do not go to the party". This is intuitively consistent, but inconsistent in **SDL** (see Section 3.2). 7.'s paradoxical nature is exemplified by the macabre instance: "one is forbidden to murder; still, if one murders someone, one has to do it gently (i.e. one has to commit a gentle murder); moreover, a gentle murder implies a murder; one murders someone". This set is inconsistent again in **SDL**, whereas it makes perfect (common) sense. 8. states that there is no conflict of duties, which is manifestly not in line with daily-life situations. 9. are deontic versions of the paradoxes of material implication in classical logic. 10. states that obligations are closed under implication. This is controversial in the same sense as 5. with which it is strongly related. 11. states that only things (actions) that *can* be (done) might be obliged. This denies the possibility to specify that one is really obligated to do the impossible. 12. is puzzling as shown by the instance: "if you ought to know that your spouse is committing adultery, it ought to be that your spouse is committing adultery."

### 3.2 The Paradoxes in SDL

In **SDL** all paradoxes 1. - 10. are theorems. (We leave 11. and 12. out of our discussion since they involve modalities, viz. knowledge and ontological possibility, which are not present in **SDL**.) Moreover, it is worthwhile to note what axioms are responsible for them (discarding the principles P and F, which are viewed in this context as mere abbreviations. Of course, one is also free to tempt these as proper definitions of permission and prohibition.) Paradox 7 depends on

(K), (N) and (D), while paradoxes 6 and 8 depend on (K) and (D). (K) and (N) are responsible for the paradoxes 2, 3, 4, 5 and the first two of 9. The third one of 9. is even valid in propositional logic. Paradox 1 is an immediate consequence of the necessity rule (N) and paradox 10. is just a consequence of the modal axiom (K). So, apart from the paradoxes 6, 7 and 8, which can be avoided by denying the principle (D), the rest of the paradoxes are direct and unavoidable consequences of viewing the obligation operator  $O$  as a (“normal” in the sense of [Che80]) Kripke-style modality, which necessarily satisfies (K) and (N)! (Moreover, denying principle (D) is also not done painlessly, since, historically, it is the very axiom in modal logic that is associated with deontics!) Do we have to conclude from this that this is not a good approach? Well, this depends on the situation at hand. In fact, we shall argue below that sometimes it is not, but in other situations this style of semantics is sufficient in adequacy.

### 3.3 Contrary-to-Duty Imperatives

One of the most serious paradoxes in (standard) deontic logic involves the notion of *contrary-to-duty imperatives*. These have to do with the specification of norms in case some other norm(s) have already been violated. The best-known example is the one given by Chisholm ([Chi63]): consider the following statements in natural language:

- (i) You ought to go to the party
- (ii) If you go to the party, you ought to tell you’re coming
- (iii) If you don’t go, you ought not to tell you’re coming
- (iv) You don’t go to the party.

Intuitively, this set of statements is perfectly understandable and consistent, and none of the four statements seems to be redundant in the set. However, if we try to represent this set in **SDL**, we run into serious trouble. A more or less natural way to represent the Chisholm set in **SDL** is the following:

- (i')  $O p$
- (ii')  $O(p \rightarrow q)$
- (iii')  $\neg p \rightarrow O\neg q$
- (iv')  $\neg p$

In **SDL** this set (i') - (iv') is inconsistent, contrary to the intuitions about (i) - (iv):

In **SDL** we have as a theorem  $O(p \rightarrow q) \rightarrow (O p \rightarrow O q)$ , which together with (ii') yields  $O q$ . On the other hand, (iii') and (iv') give  $O\neg q$ . And in **SDL**  $O q \wedge O\neg q$  is inconsistent.

Note furthermore that it is not really coincidental that (ii) and (iii) are represented in a different way (which may be questionable, of course). If we would, for instance, replace (iii') by

(iii'')  $O(\neg p \rightarrow \neg q)$

we have that (iii'') is derivable from (i') in **SDL**, which does no justice to the intuition that all of (i) - (iv) are independent from each other. Similarly, if we would replace (ii') by (ii'')  $p \rightarrow Oq$ , this statement (ii'') would be derivable from (iv'), contrary to the intuitions concerning independence.

#### 4. A Diagnosis of the Problems

At least part of the problems arise from the following four confusions:

##### (i) confusion between *ought-to-be* and *ought-to-do*.

Since the inception of deontic logic there has been some confusion about the meaning of  $O\phi$ , or rather of the meaning of  $\phi$  in this context: is it a description of a state-of-affairs or does it denote an act(ion)? We will argue that if one views  $\phi$  as an action description (i.e., one is interested in so-called *ought-to-do*'s) viewing the  $O$ -modality as a modal operator in **SDL**-style is not adequate, while for *ought-to-be*'s where  $\phi$  denoted a state-of-affairs this might be adequate for concrete applications. Here we may adhere to some form of *pragmatics*: we simply look at what works in a given simple context without having any pretensions or claims about a logic of norms, ethics or morality in general. In any case what is very important to notice is that in our view this results in having distinct (and really different) logics for *ought-to-be* and *ought-to-do*.

##### (ii) confusion between the formal interpretation (of the logical operators $\wedge$ , $\vee$ and $\rightarrow$ , for example) and the natural language (commonsense) reading of these.

Here we have again a form of *pragmatics*: once we realise that the formal interpretation of the operators used do not match the commonsense one, we just use them in their formal interpretation without trying to solve the mismatch. The only thing one should be aware of if one follows this line of action, is that one should be very careful when translating informal (natural language-style) specifications into formal ones or vice versa.

##### (iii) confusion between ideality and actuality; and, in particular, an overestimation of ideality and the notion of perfect alternative as it appears in the formal semantics. In particular, norms can conflict in reality!

In formal terms:  $O\perp$  is not equivalent with  $\perp$ ! (But, of course, in practice, when we encounter conflicting duties, we must try and resolve these.) This is very much related to *defeasible* (nonmonotonic) reasoning in AI research (cf. e.g. [Luk90], [MT93] as excellent introductions to

the area of nonmonotonic reasoning; in [LHM95] the emphasis is on the dynamics of default reasoning viewed from the stand-point of a reasoning agent. However, the dynamics of reasoning is also studied in a more general context of complex reasoning patterns like in meta-level reasoning architectures, cf. e.g. [HMT94]). In the subject of defeasible or nonmonotonic reasoning one considers (sequences of) steps while having the disposal of incomplete and overcomplete (inconsistent) information and “jumping to conclusions”. Here we encounter dynamics in the epistemic state of the agent, like belief revision and truth (reason) maintenance. Actually, this also plays a role when encountering normative information that is incomplete and inconsistent. Historically some of the first work on defeasibility was done in the realm of normative reasoning ([AM81]), where it is imperative to keep one’s duties straight in a situation with inconsistent norms. One might speak of “*duty maintenance*” in this case. Recently, mainly inspired by work done in AI on defeasibility over the last decade or so, there have been developing a rapidly increasing interest in the defeasibility approach to deontic logic and legal reasoning more in general (cf. e.g. [MW91a], [MW93a], [Rya93], [Hor93], [Pra93, 94], [Jon93], [PS94], [vdT94]). Clearly, this topic has now been taken up very seriously by researchers in the AI & Law Community, and it is to be expected that a lot of this line of research will dominate deontic logic and reasoning about norms for the years to come.

**(iv) confusion between normative notions necessary in general abstract contexts (such as ethics) and those needed (and sufficient) for a concrete practical application.**

When one is interested in the specification of concrete systems one can imagine to be interested in different problems than one is analyzing deep philosophical problems. This appears also to be the case when one is dealing with normative issues. Specifying normative systems, systems in which norms play an important role, requires a more *pragmatic* view on matters pertaining to reasoning about norms. It has now been recognized by several authors that deontic logic may be useful for normative system specification, but which deontic logic is best suited for this purpose? It appears to be the case that the most important role of the logic to be used is to accommodate for the need to distinguish between *ideal* and *actual* behaviour of the system. So, in the ought-to-be setting, we would use  $O\phi$  to mean “ $\phi$  is desirable” or “ideally  $\phi$  is the case” rather than “ $\phi$  is obliged”. We would then have the possibility to specify what is to be done if the actual behaviour deviates from the ideal or desirable one. This enables us to specify adequately fault-tolerant behaviour of the software or system at hand. When  $O\phi$  is read as “ideally  $\phi$ ” even the standard system **SDL** or Anderson’s variant of this is useful: all paradoxes but Chisholm’s lose their bite and become perfectly understandable and intuitive properties of “ideality”. So in this case we need not bother too much about deep problems (even changing the reading the O-operator slightly, as we please), and we may use **SDL** as a logic for ought-to-be constraints. The only thing that remains to be done is to provide a (pragmatic) solution of the Chisholm paradox, which we shall offer next.

### 5. A Solution to the ‘Ought-to-Be’ Version of the Chisholm Paradox: $S5O_{(n)}$

In the previous section we have seen that a natural representation of the Chisholm set in **SDL** becomes inconsistent. The crux of the matter is that the inconsistency arises because of two norms  $O_q$  and  $O_{\neg q}$  that together are inconsistent because of the D-axiom in **SDL** and, *a fortiori* the fact that both norms are expressed by the same obligation operator  $O$ . We now propose to extend **SDL** with multiple  $O$ -operators to overcome this problem (without resorting to much more sophisticated things like dyadic deontic logic or conditional logic ([vW64], Chapter 10 of [Che80])). To this end we introduce the logic  $S5O_{(n)}$  which has as modalities the universal necessity operator  $\zeta$  and a number of distinct operators  $O_i$  which express obligation with respect to a frame of reference  $i$ .

The system  $S5O_{(n)}$  consists of the following axioms and rules (here  $\perp$  is used to abbreviate  $\neg\zeta\neg$ ):

$$(K_{\zeta}) \quad \zeta(\varphi \rightarrow \psi) \rightarrow (\zeta\varphi \rightarrow \zeta\psi)$$

$$(T_{\zeta}) \quad \zeta\varphi \rightarrow \varphi$$

$$(5_{\zeta}) \quad \varphi \rightarrow \zeta\varphi$$

$$(K_i) \quad O_i(\varphi \rightarrow \psi) \rightarrow (O_i\varphi \rightarrow O_i\psi)$$

$$(D_i) \quad \neg O_i\perp$$

$$(P_i) \quad P_i\varphi \leftrightarrow \neg O_i\neg\varphi$$

$$(F_i) \quad F_i\varphi \leftrightarrow O_i\neg\varphi$$

$$(\supseteq) \quad \zeta\varphi \rightarrow O_i\varphi$$

and the rules (Taut), (MP) and

$$(N_{\zeta}) \quad \varphi / \zeta\varphi$$

Note that the rule

$$(N_i) \quad \varphi / O_i\varphi$$

is derivable.

(Kripke) Models for this system are of the form  $M = (S, \pi, R, \{R_i \mid i = 1, \dots, n\})$ , where  $S$  is the set of states,  $\pi$  is a truth assignment function,  $R = S \times S$  is the (universal) possibility relation,

and the  $R_i$  are the deontic accessibility relations, which are assumed to be serial. Note that in this set-up it holds that  $R_i \subseteq R$ .

The operator  $\varepsilon$  is now interpreted by means of the relation  $R$ , and the operators  $O_i$  are interpreted by the relation  $R_i$ . So,  $M, s \vDash \varepsilon\phi$  iff  $M, t \vDash \phi$  for all  $t$  with  $R(s, t)$ , and  $M, s \vDash O_i\phi$  iff  $M, t \vDash \phi$  for all  $t$  with  $R_i(s, t)$ . The system **S5O<sub>(n)</sub>** can be shown sound and complete with respect to validity in this class of models via a standard argument.

Note that the following are *non*-validities in this logic:  $O_i\phi \rightarrow O_iO_i\phi$ ,  $\neg O_i\phi \rightarrow O_i\neg O_i\phi$ ,  $O_i\phi \rightarrow O_jO_i\phi$ , and  $\neg(O_i\phi \wedge O_j\neg\phi)$  ( $i \neq j$ ). The nonvalidity of the last formula expresses that it is now indeed possible to represent contrary norms in a consistent way, using distinct frames of reference, although within *one* frame of reference  $i$  the formula  $O_i\phi \wedge O_i\neg\phi$  is still inconsistent.

The Chisholm set can now be consistently (and nonredundantly) be represented as:

- (i'')  $O_1p$
- (ii'')  $\varepsilon(p \rightarrow O_2q)$
- (iii'')  $\varepsilon(\neg p \rightarrow O_3\neg q)$
- (iv'')  $\neg p$

Moreover, note that in this representation the strange incongruence between (ii'') and {iii''} has disappeared. From this representation one can derive (iii'')  $\varepsilon(\neg p \rightarrow O_3\neg q) \Rightarrow \neg p \rightarrow O_3\neg q \Rightarrow_{(iv'')} O_3\neg q$ , and (ii'')  $\varepsilon(p \rightarrow O_2q) \Rightarrow O_1(p \rightarrow O_2q) \Rightarrow O_1p \rightarrow O_1O_2q \Rightarrow_{(i'')} O_1O_2q$ , which gives a precise record of how the contrary norms come about (in frame 3, and in frame 2 via frame 1 on the other hand) without being inconsistent.

## 6. Ought-to-Do: The Dynamic Perspective

As said before, computer science provides a dynamic perspective. Basically this is because in this area one is interested in a way to compute (constructively) things, and since this process generally takes time, this naturally involves changes of computer states over time. We can distinguish at least the following three particular issues that are studied in computer science:

1. First and foremost computer science has to do with (computer) *programs*. Since its inception computer science has to do with algorithms to achieve certain goals and computer programs written in some computer language to enable one to execute these algorithms. These programs contain statements or instructions to express what has to be done. In other words programs express how certain actions involving (the hardware of) the computer should be performed. To reason about programs, special logics have been developed such as Hoare's logic and dynamic

logic ([Hoa69], [Har79]). (Also temporal logic has been employed for this purpose (see e.g. [Krö87], [MP92]). In some way one might view temporal logic as a dynamic logic in which one abstracts away from the particular actions that take place, and only considers the flow of time while executing a program. On the other hand, one might also view dynamic logic as a kind of temporal logic where records are kept of what exactly happens in a time step.) In these logics it is possible to express precisely what the effect is of the execution of (parts of) programs. One can use them to reason about pre- and postconditions of (executions of) programs. Since especially in dynamic logic this is done in a way that abstracts from the basic programming actions, it is an easy and straight-forward step to abstract away from the particular application to computer programs and consider general actions, whether they are supposed to be executed by computers or by humans. In this way, dynamic logic may be viewed as a general logic for reasoning about actions.

We ([Mey87, Mey88], [MWW89], [MWM89], [DM90], [WWMD91], [DMW94a,b], [DMW96]) have employed this idea in order to get a deontic logic for ought-to-do (i.e. obliged *actions*). We will sketch this approach below in section 5.1. Admittedly, also in the philosophical literature on deontic logic there have been proposals to distinguish between actions and assertions ([vW81], [Cas81], also cf. [Hil93]), but we believe that the explicit connections with formalisms to reason about programs such as dynamic logic has made things much more concrete. Interestingly, recent proposals for a philosophical theory of action as put forward by Krister Segerberg (e.g. [Seg89]), are also strongly influenced by the dynamic logic approach stemming from computer science and are very close in spirit to our approach.

2. More generally, computer science is concerned with the study of *processes*. Processes might be viewed as generalizations of executions of standard (sequential) programs, where also parallel (and nondeterministic) execution is catered for. In fact, there is a whole branch of theoretical computer science dealing with so-called process theory, which ranges from concurrency semantics in which models for concurrent or parallel (nondeterministic) programming are investigated (e.g. [Win82], [BM88], [vG90]) to process algebra, in which one tries to give an algebraic calculus for this kind of programming ([Mil80], [BK86], [BW90]). In these algebraic calculi (or *process algebras*, as they are usually called) one is actually interested in calculating equivalences of processes on the basis of certain observational criteria: if two processes are indistinguishable with respect to some observational criterion, they are regarded as equal, so that one can reduce terms denoting complex processes to those representing simpler ones. This appears to be very useful when considering correctness issues of these processes: by algebraic calculation one aims at verifying certain properties of processes at hand such as protocols for communication between ‘agents’ (computers or processors within a multi-processor computer system).

We employed the dynamic perspective in the sense of using dynamic logic over process (algebraic) terms to treat the paradox of free choice permission. In [MW91], [WM91], [WM93a] we used an algebraic approach on process terms to distinguish between two choice operators representing free and imposed choice and obtained a framework (albeit complicated) in which both could be used intertwined. Here we mixed techniques from process algebra (and universal algebra) with dynamic logic to get a solution for this well-known problem in deontic logic. (In [DMW94a, DMW96] we gave a less involved solution to this problem using another (stronger) definition of permission together with admissible contexts for actions, but also here dynamic logic is a crucial ingredient of the approach.)

### 6.1 A Logic of Ought-to-Do: a Deontic Logic Based on Dynamic Logic

**PDeL**, introduced in [Mey88], is a version of dynamic logic especially tuned to use as ought-to-do style deontic logic. It is based on the idea of Anderson's reduction of ought-to-be style deontic logic to alethic modal logic, but instead it reduces ought-to-do deontic logic to dynamic logic ([Har79]). The basic idea is very simple: some action is forbidden if doing the action leads to a state of violation. In a formula:  $\underline{F}\alpha \leftrightarrow_{\text{def}} [\alpha]\mathbf{V}$ , where the dynamic logic formula  $[\alpha]\varphi$  denotes that execution / performance of the action  $\alpha$  leads (necessarily) to a state (or states) where  $\varphi$  holds, and  $\mathbf{V}$  is a special atomic formula denoting violation. Formally, we say that the meaning of action  $\alpha$  is captured by an accessibility relation  $R_\alpha \subseteq S \times S$  associated with  $\alpha$ , where  $S$  is the set of possible worlds. This relation  $R_\alpha$  describes exactly what possible moves (state transitions) are induced by performance of the action  $\alpha$ :  $R_\alpha(s, t)$  says that from  $s$  one can get into state  $t$  by performing  $\alpha$ . (In concurrency semantics and process algebra this is often specified by a so-called transition system which enables one to derive (all) transitions of the kind  $s \xrightarrow{\alpha} t$ , which in fact defines the relation  $R_\alpha$  for all possible actions  $\alpha$ .) Now the formal meaning of the formula  $[\alpha]\varphi$  is given by:  $[\alpha]\varphi$  is true in a state (possible world)  $s$  iff all states  $t$  with  $R_\alpha(s, t)$  satisfy  $\varphi$ . This then provides the formal definition of the  $\underline{F}$ -operator, as given above.

The other deontic modalities are derivatives of  $\underline{F}$ : permission is not-forbidden ( $\underline{P}\alpha \leftrightarrow \neg \underline{F}\alpha$ ), and obligation is forbidden-not-to ( $\underline{O}\alpha \leftrightarrow \underline{F}\bar{\alpha}$ ), where  $\bar{\alpha}$  has the meaning of "not- $\alpha$ ". The formal semantics of this negated action is non-trivial, especially in case one considers composite actions, cf. [Mey88], [WMW89], [DM90], [MW91], [WM93a]. In these papers we considered connectives for composing non-atomic actions, such as ' $\cup$ ' (choice, the dynamic analogue of disjunction in a static setting), '&' (parallel, the analogue of conjunction), ' $\neg$ ' (non-performance, the analogue of negation), and ';' (sequential composition, which has no analogue in a static setting). Without giving a formal semantics here (see the papers mentioned above for that), the meaning of these are as follows:  $\alpha_1 \cup \alpha_2$  expresses a choice between  $\alpha_1$  and  $\alpha_2$  (this—roughly—corresponds to taking  $R_{\alpha_1 \cup \alpha_2}$  as the set-theoretic union of  $R_{\alpha_1}$  and  $R_{\alpha_2}$ ),  $\alpha_1 \& \alpha_2$  a parallel performance of  $\alpha_1$  and  $\alpha_2$  (this amounts to more or less taking  $R_{\alpha_1 \& \alpha_2}$  to be

the intersection of  $R_{\alpha_1}$  and  $R_{\alpha_2}$ ),  $\bar{\alpha}$  (we will also often write  $-\alpha$ ) the non-performance of  $\alpha$ , as stated above (it more or less amounts to taking  $R_{\bar{\alpha}}$  to be some complement of  $R_{\alpha}$ , but see also the discussion below), and  $\alpha_1 ; \alpha_2$  the performance of  $\alpha_1$  followed by that of  $\alpha_2$ . For a full account of the semantics of particularly negated actions we refer to [Mey88, DM90, DMW94a, DMW96].

The logic **PDeL** now consists of the following axioms and rules:

- (K<sub>[ $\alpha$ ]</sub>)  $[\alpha](\varphi \rightarrow \psi) \rightarrow ([\alpha]\varphi \rightarrow [\alpha]\psi)$
- (;)  $[\alpha ; \beta]\varphi \leftrightarrow [\alpha][\beta]\varphi$
- ( $\cup$ )  $[\alpha \cup \beta]\varphi \leftrightarrow ([\alpha]\varphi \wedge [\beta]\varphi)$
- (&)  $[\alpha]\varphi \rightarrow [\alpha \& \beta]\varphi$
- (-;)  $[-(\alpha ; \beta)]\varphi \leftrightarrow ([-\alpha]\varphi \wedge [\alpha][-\beta]\varphi)$
- (- $\cup$ )  $[-\alpha]\varphi \rightarrow [-(\alpha \cup \beta)]\varphi$
- (-&)  $[-(\alpha \& \beta)]\varphi \leftrightarrow ([-\alpha]\varphi \wedge [-\beta]\varphi)$
- (F)  $\underline{F}\alpha \leftrightarrow [\alpha]V$
- (P)  $\underline{P}\alpha \leftrightarrow \neg \underline{O}(\alpha) (\leftrightarrow \langle \alpha \rangle \neg V)$
- (O)  $\underline{O}\alpha \leftrightarrow \underline{F}(-\alpha) (\leftrightarrow [-\alpha]V)$

This results in the following theorems concerning the deontic operators:

- (O;)  $\underline{O}(\alpha ; \beta) \leftrightarrow (\underline{O}\alpha \wedge [\alpha]\underline{O}\beta)$
- (P;)  $\underline{P}(\alpha ; \beta) \leftrightarrow \langle \alpha \rangle \underline{P}\beta$
- (F;)  $\underline{F}(\alpha ; \beta) \leftrightarrow [\alpha]\underline{F}\beta$
- (O&)  $\underline{O}(\alpha \& \beta) \leftrightarrow (\underline{O}\alpha \wedge \underline{O}\beta)$
- (P&)  $\underline{P}(\alpha \& \beta) \rightarrow (\underline{P}\alpha \wedge \underline{P}\beta)$
- (F&)  $(\underline{F}\alpha \vee \underline{F}\beta) \rightarrow \underline{F}(\alpha \& \beta)$
- (O $\cup$ )  $(\underline{O}\alpha \vee \underline{O}\beta) \rightarrow \underline{O}(\alpha \cup \beta)$
- (P $\cup$ )  $\underline{P}(\alpha \cup \beta) \leftrightarrow (\underline{P}\alpha \vee \underline{P}\beta)$
- (F $\cup$ )  $\underline{F}(\alpha \cup \beta) \leftrightarrow (\underline{F}\alpha \wedge \underline{F}\beta)$

## 6.2 The Paradoxes in PDeL

A nice feature of **PDeL** is that most of the paradoxes appearing in **SDL** are either not expressible or, if they are, not valid (cf. [Mey88], [MW93a]). To be more specific, if we consider (dynamic variants of) the paradoxes 1 – 10 of section 2, we get:

- 1'.  $\underline{O}(\alpha \cup \bar{\alpha})$  *not* valid in **PDeL**
- 2'.  $\underline{O}\alpha \rightarrow \underline{O}(\alpha \cup \beta)$  valid in **PDeL**

- 3'.  $(\underline{P}\alpha \vee \underline{P}\beta) \leftrightarrow \underline{P}(\alpha \vee \beta)$  valid in **PDeL**  
 4'.  $\underline{F}\alpha \rightarrow \underline{F}(\alpha \ \& \ \beta)$  valid in **PDeL**  
 5'. not expressible in **PDeL**  
 6'. Chisholm's paradox: see the discussion in the next section.  
 7'. Forrester's paradox of gentle murder: see the discussion in the next section.  
 8'.  $\neg(\underline{O}\alpha \wedge \underline{O}\bar{\alpha})$  *not* valid in **PDeL**  
      $\underline{O}\alpha \rightarrow \underline{P}\alpha$  *not* valid in **PDeL**  
 9'. the first two not expressible in **PDeL**  
      $\neg\varphi \rightarrow (\varphi \rightarrow \underline{O}\alpha)$  valid in **PDeL**  
 10'. not expressible in **PDeL**

Thus we see that 1' and 8' are not valid in **PDeL**; we still are left with Ross' paradox and the related ones of free choice permission and the penitent 2' - 4') (However, many authors do not really see them as paradoxes once one realizes the mismatch between the natural language reading and the formal semantics of these expression, cf. section 6. Nevertheless, for readers interested in attempts how to overcome even these paradoxes formally we refer to [Mey92], [WM93a], [DMW94a], [DWW96].) We can say something sensible about contrary-to-duty imperatives involved in the paradoxes of Chisholm's and Forrester's (see next subsection). The third expression of 9' is valid as it is just an instance of propositional logic (and particularly the paradox of the material implication) again. Finally, some words about the paradoxes of **SDL** that not even have direct counterparts in **PDeL** (5', 9' (first two), 10'). It is perhaps slightly too easy to just say these are not expressible, so there is no paradox any more. One might, for instance, introduce a dynamic counterpart of the logical implication, as follows:

$$\alpha \gg \beta$$

with as reading "action  $\alpha$  involves action  $\beta$ ", and as formal semantics:  $R_\alpha \subseteq R_\beta$ : all states reachable by doing action  $\alpha$  are also reachable by doing action  $\beta$ . This means, of course, that all results of action  $\beta$  are also results of action  $\alpha$ , so that we have the following proposition:

PROPOSITION.  $\alpha \gg \beta$  implies the validity of  $[\beta]\varphi \rightarrow [\alpha]\varphi$  for all assertions  $\varphi$ .

To give an informal example, "murdering someone gently" involves "murdering someone". (Cf. the discussion of Forrester's paradox in the next subsection.) We refer to **PDeL** extended with the operator  $\gg$  as **PDeL**( $\gg$ ).

Now we can express dynamic counterparts of 5':

5'.  $\alpha \gg \beta \not\vdash \underline{O}\alpha \rightarrow \underline{O}\beta$ . Whether this is a validity in **PDeL**( $\gg$ ) depends on the interpretation of the negated action  $\bar{\alpha}$ . As we have discussed in e.g. [WM93] there is not an obvious unique choice for this. A possible choice—which, by the way, we purposely did *not* follow in our original paper [Mey88]—would be that the accessibility relation  $R_{\bar{\alpha}}$  associated with  $\bar{\alpha}$  is the set-theoretical complement of that ( $R_{\alpha}$ ) associated with  $\alpha$ , i.e.,  $R_{\bar{\alpha}} = (S \times S) \setminus R_{\alpha}$ , where  $S$  is the set of all possible worlds. In this case we have that  $\alpha \gg \beta$  implies  $\bar{\beta} \gg \bar{\alpha}$  (since now  $\alpha \gg \beta \Leftrightarrow R_{\alpha} \subseteq R_{\beta} \Leftrightarrow R_{\bar{\alpha}} \supseteq R_{\bar{\beta}} \Leftrightarrow \bar{\beta} \gg \bar{\alpha}$ ). In its turn this implies that under this interpretation of negated action we have that 5' is a validity (since  $\alpha \gg \beta \Rightarrow \bar{\beta} \gg \bar{\alpha} \Rightarrow \nexists [\bar{\alpha}]V \rightarrow [\bar{\beta}]V \Rightarrow \nexists \underline{O}\alpha \rightarrow \underline{O}\beta$ ). Note, however, that this is only true in this particular interpretation of  $\bar{\cdot}$ .

However, we have that the first two cases of 9' and that of 10' are still not expressible. Note that the obvious attempts, viz.  $\underline{O}\alpha \rightarrow \underline{O}(\beta \gg \alpha)$ ,  $\underline{E}\beta \rightarrow \underline{O}(\beta \gg \alpha)$ , and  $(\underline{O}\alpha \wedge \underline{O}(\alpha \gg \beta)) \rightarrow \underline{O}\beta$  are not well-formed in **PDeL**( $\gg$ ), since  $\underline{O}(\alpha_1 \gg \alpha_2)$  is not ( $\underline{O}$  should have an action as an argument). So to deal with this we should really construct a hybrid logic of both ought-to-do and ought-to-be operators. In Section 7 we shall discuss a proposal into this direction, which nevertheless will *not* have the above formulas as validities.

### 6.3 A Solution to the ‘Ought-to-Do’ Version of the Chisholm Paradox in PDeL

Here we like to show how the dynamic perspective and particularly **PDeL** can help us solve the problems with the Chisholm set. To analyse the problem in **PDeL** we need to be a little more specific about the order in which the actions take place. (This distinction does not really occur in the **SDL** representation since here everything is formalized in a static way.) We can in fact distinguish three versions of the Chisholm set. We here use rather abstract versions. See the thesis of Tina Smith ([Smi94]) for some very nice daily-life instances.

1. The “*forward*” version of the Chisholm set.

- (li) it is obligatory to do  $\alpha$
- (lii) if you do  $\alpha$ , you have to do  $\beta$  *afterwards* (i.e., *after*  $\alpha$ )
- (liii) if you don’t do  $\alpha$ , you have to do non- $\beta$
- (liv) you don’t do  $\alpha$

This version is the easiest one to formalize in **PDeL**. We immediately can represent it as follows:

- (li')  $\underline{O}\alpha$
- (lii')  $[\alpha]\underline{O}\beta$

(1iii')  $[\bar{\alpha}] \underline{O}\bar{\beta}$ .

(Remark: the fourth premise of the set, which expresses that some action is actually performed, cannot be represented in **PDeL**. In some sense, statements of actions in **PDeL** and the underlying dynamic logic is of a hypothetical nature: “if one (would) perform the action, the following holds”. The implication implicit in an formula  $[\alpha]\varphi$  is therefore more like a conditional in conditional logic. As such, it is not really important what actually happens. Here and in the sequel we shall just ignore the fourth assertion in the formal representation.)

In **PDeL** one may derive from this representation that it holds that  $\underline{O}(\alpha ; \beta) \wedge [\bar{\alpha}] (\vee \wedge F\beta)$ , in other words: it obligatory to perform the sequence  $\alpha$  followed by  $\beta$ , and moreover, if  $\alpha$  has not been done, one is in a state of violation but nevertheless also forbidden to do  $\beta$  (see [Mey88]). This is exactly as one would expect.

## 2. The “parallel” version of the Chisholm set

- (2i) it is obligatory to do  $\alpha$
- (2ii) you have to do  $\beta$  *while*  $\alpha$  is being done
- (2iii) if you don't do  $\alpha$ , you have to do non- $\beta$
- (2iv) you don't do  $\alpha$

In fact, this version is very much related to another infamous paradox involving contrary-to-duties, viz. Forrester's. This is also known as the paradox of the gentle murderer:

- (F1) One is forbidden to commit murder
- (F2) Still, if one murders someone, one should do so gently
- (F3) Jones murders someone.

In **SDL** this is a big problem again. There one would use a formalization like (using  $m$  for committing murder and  $g$  for committing a murder gently):

- (F1')  $Fm$
  - (F2')  $m \rightarrow Og$
  - (F3')  $m$
- together with the implicit necessary truth
- (F4')  $g \rightarrow m$

Now we can derive: (F4')  $g \rightarrow m \quad f_{\mathbf{SDL}} \quad O(g \rightarrow m) \quad f_{\mathbf{SDL}} \quad Og \rightarrow Om$ , so that by (F3') and (F2') we arrive at  $Om$ , which, apart from being quite absurd, is inconsistent (in  $\mathbf{SDL}$ ) with (F1')! Castañeda dubbed Forrester's paradox as the "deepest" in deontic logic ([Cas84]): the main reason being that distinguishing between assertions and actions ("*practitions*") did not help in his approach to deal with it. He placed emphasis on the fact that there is an *adverb* ("*aspect*"), viz. gently, involved. However, we believe that the problem is the *simultaneity* of actions: one should be gentle (or rather act gently) *while* murdering (see also [Mey87]). if we would consider a "forward" variant of Forrester's set, we get something which is very easy to formulate in  $\mathbf{PDeL}$ :

e.g.

- (a) You are forbidden to go
- (b) Yet, if you go, you have to close the door *afterwards*

In  $\mathbf{PDeL}$ , using obvious abbreviations:

- (a')  $\underline{F}g$
- (b')  $[\underline{g}]\underline{O}c$

In fact, this is "one half" of the forward Chisholm set (viz. isomorphic to the set  $\{(1i'), (1iii')\}$ , since  $\underline{F}g$  is equivalent with  $\underline{O}\bar{g}$ , so that taking  $\alpha = \bar{g}$  results in the aforementioned set). Now, however, we have a parallel version, for which the above representation is erroneous. As we stated in [Mey87], in some sense the Forrester paradox (but also the parallel version of the Chisholm paradox) is the hardest to represent in  $\mathbf{PDeL}$ . As was shown there the intuitive representation

- (F1'')  $\underline{F}m$                     it is forbidden to murder
- (F2'')  $\underline{O}(\bar{m} \cup g)$         one ought to not-murder or murder gently

is not adequate either, since in  $\mathbf{PDeL}$  the latter (which is equivalent with  $\underline{F}(m \ \& \ \bar{g})$ ) is derivable from the first, which contradicts the intuition. But, as we indicated in [Mey87] too, there is an easy way out by using multiple violation atoms (as we have also been using frequently in subsequent papers). In short, assuming that we have a set  $\{V_i\}_i$  of distinct violation atoms and using the abbreviations  $\underline{F}_k\alpha \leftrightarrow_{\text{def}} [\alpha]V_k$  and  $\underline{O}_k\alpha \leftrightarrow_{\text{def}} \neg\underline{F}_k\alpha$  we can now adequately represent Forrester's set as

- (F1''')  $\underline{F}_1m$
- (F2''')  $\underline{F}_2(m \ \& \ \bar{g})$

which has as a desirable consequence  $[m \ \& \ \bar{g}](V_1 \wedge V_2)$ , i.e., after murdering someone non-gently one is guilty of two offences: having murdered someone and having been not gentle while murdering. Note that ‘just’ murdering someone gets one into a state where  $V_1$  holds, while murdering someone in a non-gentle way yields a state where  $V_1 \wedge V_2$ . One might thus view the latter state as even less ideal than the former. In [DMW94b] we have elaborated on the notion of (sub-)ideality. In that paper we went even a step further in analyzing the effect of actions on violations by not only considering whether by performing an action a state *results* in a violation state, but also whether the *transition* brought about by the action *changes* the state of violation. We will not pursue this here.

Finally, in **PDeL** with multiple violations the parallel version of the Chisholm set can be represented in a similar way as the Forrester one:

- (2i')  $\underline{O}_1\alpha$                     one ought to do  $\alpha$
- (2ii')  $\underline{F}_2(\bar{\beta} \ \& \ \alpha)$     it is forbidden to do both  $\alpha$  and non- $\beta$  at the same time
- (2iii')  $\underline{F}_3(\beta \ \& \ \bar{\alpha})$     it is forbidden to do both  $\beta$  and non- $\alpha$  at the same time

3. The “backward” version of the Chisholm set. This is the original one, as formulated by Chisholm:

- (3i)    it is obligatory to do  $\alpha$
- (3ii)    if you do  $\alpha$ , you have to do  $\beta$  *first* (i.e., *before*  $\alpha$ )
- (3iii)    if you don’t do  $\alpha$ , you have to do non- $\beta$  (first)
- (3iv)    you don’t do  $\alpha$

This is perhaps the more difficult one to express in **PDeL**, since **PDeL** does not contain an operator of the kind “if you do  $\alpha$ , you have to do  $\beta$  *first*”. Surely again, if one would use the same representation as for the forward version, one does not get the intended result, e.g. (2ii) would be misrepresented. In [Mey87] it is argued that the best representation in **PDeL** is:

- (3i')     $\underline{O}_1\alpha$                     one ought to do  $\alpha$
- (3ii')     $\underline{F}_2(\bar{\beta} \ ; \ \alpha)$             it is forbidden to do non- $\beta$  followed by  $\alpha$
- (3iii')     $\underline{F}_3(\beta \ ; \ \bar{\alpha})$             it is forbidden to do  $\beta$  followed by non- $\alpha$

See [Mey87] for a more elaborate discussion on this representation. We admit that it would be far nicer to have a representation closer to the natural language representation, but this would call for a non-trivial extension of **PDeL**, in which one also can reason “backward” directly.

We have seen how we can reason about ought-to-be and ought-to-do constraints in separate logics. However, in the practice of system specification one naturally needs both notions intermingled, so that an integrated logic for both these notions is called for.

## 7. An Integrated Logic of Ought-to-Be and Ought-to-Do Constraints

In [AMW93] we have proposed a unifying framework for reasoning about ought-to-be and ought-to-do constraints. There we employed **PDeL** as in Section 6 for ought-to-do constraints and Anderson's reduction ([And58]) for ought-to-be constraints. Since in this paper we started out from **SDL** with normal deontic modalities (which are not defined by a reduction to alethic modal logic as in Anderson's reduction) we have to do a little more work to integrate our logics for ought-to-do and ought-to-be.

### 7.1 Anderson's Reduction to Modal Alethic Logic Related to SDL

As mentioned before, Anderson [And58] reduced deontic modalities to alethic ones by employing the special propositional atom  $V$  denoting 'something bad' or undesirable, which we refer to as 'violation'. In a logic with an S5-type (alethic) necessity operator  $\zeta$  we can define the obligation operator  $O$  by  $O\phi = \zeta(\neg\phi \rightarrow V)$ , expressing that something is obligated iff it is necessarily the case that the falsity of  $\phi$  implies being in a state of violation. (Actually, Anderson used the modal system KT as a basis, but we will use S5 instead.) Semantically, we employ (simple) S5-models for the modality  $\zeta$ : such a model consists of a non-empty set  $S$  of worlds and a universal accessibility relation on  $S$ ; in a world the formula  $\zeta\phi$  holds iff  $\phi$  holds in all worlds of  $S$ .

We now show that this idea is related to the standard definition of obligation in **SDL** (cf. also Part IV of [Åqv84]). Recall that **SDL**-models are Kripke models  $M = (S, \pi, R)$ , where  $S$  is a non-empty set of possible worlds,  $\pi$  is a truth assignment function of the propositional atoms per world, and  $R$  is an accessibility relation pointing at deontically ideal alternatives:  $R(s, t)$  represents that the world  $t$  is a deontically ideal alternative for  $s$ . In a world  $s \in S$ , the formula  $O\phi$  is said to hold iff  $\phi$  holds in all deontically ideal alternatives of  $s$ :  $(M, s) \vDash O\phi$  iff  $(M, t) \vDash \phi$  for all  $t$  with  $R(s, t)$ .

We can reconcile Anderson's approach with **SDL** if we consider the worlds where  $V$  does not hold as ideal worlds, independent of the world where we are viewing from.

Define  $R(s, t) \Leftrightarrow (M, t) \vDash \neg V$ .

Now we see:  $(M, s) \vDash O\phi \Leftrightarrow$   
 for all  $t$ :  $[R(s, t) \Rightarrow (M, t) \vDash \phi] \Leftrightarrow$   
 for all  $t$ :  $[(M, t) \vDash \neg V \Rightarrow (M, t) \vDash \phi] \Leftrightarrow$   
 for all  $t$ :  $[(M, t) \vDash \neg V \Rightarrow (M, t) \vDash \phi] \Leftrightarrow$

$$\begin{aligned}
& \text{for all } t : [(M, t) \vDash \neg V \rightarrow \varphi] \Leftrightarrow \\
& \text{for all } t : [(M, t) \vDash \neg\varphi \rightarrow V] \Leftrightarrow \\
& (M, s) \vDash \zeta(\neg\varphi \rightarrow V) \Leftrightarrow \\
& (M, s) \vDash O'\varphi
\end{aligned}$$

That is to say, if we take this definition of the relation  $R$ , the **SDL**-based obligation operator coincides with Anderson's operator.

Note that from this definition of  $R$  it follows immediately that the relation  $R$  is an equivalence relation, so that the operator  $O$  ( $= O'$ ) satisfies the well-known  $S5$ -axioms as well:

$$\begin{aligned}
& \vDash O\varphi \rightarrow OO\varphi \text{ and} \\
& \vDash \neg O\varphi \rightarrow O\neg O\varphi
\end{aligned}$$

and in fact the stronger formulas  $\vDash O\varphi \leftrightarrow OO\varphi$  and  $\vDash \neg O\varphi \leftrightarrow O\neg O\varphi$

So, actually what we get is what is called the system *Deontic S5* by Chellas ([Che80]). As we see from the validities above, nestings are trivial in Deontic  $S5$ . One way to look at this, is that if one uses a system as Deontic  $S5$  for obligations, one is not really interested in nested obligations (since the commonsense idea of such nested obligations are unlikely to satisfy properties like the above).

Furthermore, it also immediately apparent from the definition of  $R$  that  $R(s, t)$  does not depend on the world  $s$ . In fact the set  $\{t \mid R(s, t)\} = \{t \mid (M, t) \vDash V\}$  is a fixed set, *the set of ideal worlds* (the set *opt* of optimal worlds in terms of Åqvist ([Åqv84])). This also has some further consequences of nestings between the modal operators  $O$  and  $\zeta$ : we have as validities:

$$\begin{aligned}
& \vDash O\varphi \leftrightarrow \zeta O\varphi \text{ and } \vDash \neg O\varphi \leftrightarrow \zeta \neg O\varphi \\
& \vDash \zeta\varphi \leftrightarrow O\zeta\varphi \text{ and } \vDash \neg\zeta\varphi \leftrightarrow O\neg\zeta\varphi
\end{aligned}$$

and, of course, we have that

$$\vDash \zeta\varphi \rightarrow O\varphi$$

since the set of ideal worlds are a subset of the set of all possible worlds.

In fact, when we add (P) and (F) to express permission and prohibition in terms of obligation again, we obtain an extension of **S5O**<sub>(1)</sub> from Section 5 (where we drop the subscript 1 from the operators): besides the axioms and rules of that system

$$(K_{\exists}) \quad \exists(\varphi \rightarrow \psi) \rightarrow (\exists\varphi \rightarrow \exists\psi)$$

$$(T_{\exists}) \quad \exists\varphi \rightarrow \varphi$$

$$(5_{\exists}) \quad \varphi \rightarrow \exists \varphi$$

$$(K_O) \quad O(\varphi \rightarrow \psi) \rightarrow (O\varphi \rightarrow O\psi)$$

$$(D_O) \quad \neg O\perp$$

$$(P) \quad P\varphi \leftrightarrow \neg O\neg\varphi$$

$$(F) \quad F\varphi \leftrightarrow O\neg\varphi$$

$$(\supseteq) \quad \exists\varphi \rightarrow O\varphi$$

and the rules (Taut), (MP) and

$$(N_{\exists}) \quad \varphi / \exists\varphi$$

we have additionally:

$$(4_O) \quad O\varphi \rightarrow OO\varphi$$

$$(5_O) \quad \neg O\varphi \rightarrow O\neg O\varphi$$

$$(\text{opt!}) \quad O\varphi \rightarrow \exists O\varphi$$

The axiom (opt!) expresses that there is actually only one (a *unique*) set of ideal or optimal states, independent from the world one is looking from.

This system can be shown sound and complete with respect to the class of Kripke models we have defined above. (It is in fact a special case of the system **S5P** in [MH95] and can also be viewed as a subsystem of the system of Kraus & Lehmann for knowledge and belief for one agent [KL86].) It can be shown that all the other validities we have listed above are derivable in the system. (The system, however, still contains some redundancy: one can omit (4<sub>O</sub>) and (5<sub>O</sub>) from the system without loss of inferential power. E.g. (4<sub>O</sub>) is derivable from (opt!) and ( $\supseteq$ ). We have nevertheless included (4<sub>O</sub>) and (5<sub>O</sub>) to expose the link with **Deontic S5** explicitly.)

Whether the system **S5O**<sub>(1)</sub> is suitable for normative reasoning is again a matter of pragmatics. Of course, (4<sub>O</sub>) and (5<sub>O</sub>) are highly implausible if one wants to interpret nested obligation in a meaningful way. (This has been the traditional objection against **Deontic S5**, cf. [Che80].) On the other hand, if one is simply not interested in these nested obligations this logic provides a way of eliminating them in any expression, and the system becomes very useful indeed.

## 7.2 Integrating $S5O_{(n)}$ with PDeL

The view unifying Anderson's reduction with **SDL** as expounded above provides us with the opportunity to give a natural integration of our ought-to-be and ought-to-do frameworks. For the ought-to-do approach we employ multiple violation atoms  $V_i$  (in order to represent contrary-to-duty constraints adequately). We now simply employ a multiple  $V_i$ -version of the logic above for the ought-to-be part, which then amounts to using the system  $S5O_{(n)}$  together with the axioms:

$$(4_i) \quad O_i\phi \rightarrow O_iO_i\phi$$

$$(5_i) \quad \neg O_i\phi \rightarrow O_i\neg O_i\phi$$

$$(opt_i!) \quad O_i\phi \rightarrow \zeta_i O_i\phi$$

The discussion of relating Anderson's reduction to **SDL** generalizes easily to the case of  $S5O_{(n)}$  where multiple violation atoms and modalities are employed by defining  $R_i(s, t) \Leftrightarrow (M, t) \vdash V_i$ . Now we can add the ought-to-do part of Section 6 without any difficulty:

$$(\mathcal{K}_{[\alpha]}) \quad [\alpha](\phi \rightarrow \psi) \rightarrow ([\alpha]\phi \rightarrow [\alpha]\psi)$$

$$(\cdot) \quad [\alpha ; \beta]\phi \Leftrightarrow [\alpha][\beta]\phi$$

$$(\cup) \quad [\alpha \cup \beta]\phi \Leftrightarrow ([\alpha]\phi \wedge [\beta]\phi)$$

$$(\&) \quad [\alpha]\phi \rightarrow [\alpha \& \beta]\phi$$

$$(-\cdot) \quad [-(\alpha ; \beta)]\phi \Leftrightarrow ([-\alpha]\phi \wedge [\alpha][-\beta]\phi)$$

$$(-\cup) \quad [-\alpha]\phi \rightarrow [-(\alpha \cup \beta)]\phi$$

$$(-\&) \quad [-(\alpha \& \beta)]\phi \Leftrightarrow ([-\alpha]\phi \wedge [-\beta]\phi)$$

$$(\underline{E}_i) \quad \underline{E}_i\alpha \Leftrightarrow [\alpha]V_i$$

$$(\underline{P}_i) \quad \underline{P}_i\alpha \Leftrightarrow \neg \underline{F}_i\alpha \quad (\Leftrightarrow \langle \alpha \rangle \neg V_i)$$

$$(\underline{O}_i) \quad \underline{O}_i\alpha \Leftrightarrow \underline{F}_i(-\alpha) \quad (\Leftrightarrow [-\alpha]V_i)$$

We finally mention that in this integrated logic the formulas as mentioned at the end of Section 6.2, i.e.  $\underline{O}_i\alpha \rightarrow O_i(\beta \gg \alpha)$ ,  $\underline{E}_i\beta \rightarrow O_i(\beta \gg \alpha)$ , and  $(\underline{O}_i\alpha \wedge O_i(\alpha \gg \beta)) \rightarrow \underline{O}_i\beta$ , are *not* valid, which may be viewed as a desirable feature of this approach.

## 8. Applications

Library regulations are a popular source of examples for deontic logicians. We have specified the rules for borrowing and returning documents from a library in an earlier paper [WMW89]. Jones [Jon90] points out that the Chisholm set appears in library regulations as follows. If  $p$  identifies a person and  $d$  a document, then the following rules are quite common:

- $p$  shall return  $d$  by date due.
- If  $p$  does not return  $d$  by data due, then disciplinary action shall be taken against  $p$ .

- If  $p$  returns  $d$  by date due, then disciplinary action is not taken against  $p$ .

Other applications of deontic logic involve bank accounts. For example, we have shown that in a combined logic of ought-to-be and ought-to-do such as the one in the previous section, it can be proven that having a negative balance of a bank account (a forbidden state) implies the obligation to deposit money (an obligated action) [AMW93, AMW96]. The popularity of examples like these may be taken as evidence of the fact that deontic logicians have ample experience with libraries and negative bank accounts. Although the examples are quite realistic (especially the negative bank account), we here take another interesting case study, the specification of procedures for overseas trade. We use an analysis of this case by Bons, Lee and Wagenaar [BLW94] as point of departure. It turns out that the overseas trade procedures illustrate all three variants of the Chisholm set that we identified above. In order to understand the example, we must digress a little in overseas trade procedures. During this explanation, it will become clear that deontics play an essential role in the specification of this procedure.

The problem with overseas trade is that the seller does not want to ship his goods before he is paid but the buyer does not want to pay before he receives the goods. To get out of this deadlock, buyer and seller each contact a bank in their own country, called the *issuing bank* and *corresponding bank*, respectively.

The buyer and seller negotiate a contract in which the buyer promises to buy goods from the seller at a certain price and the seller promises to sell these goods at this price to the buyer. The buyer then enters a contract with the issuing bank, which results in a *letter of credit* (LoC) in which the buyer specifies which documents the seller must produce in evidence of his performance according to contract. The LoC may be subject to further negotiations between the buyer and seller. Once the LoC is produced, the seller is obliged to produce the documents specified in the LoC, evidencing his performance according to the sales contract.

For the purpose of this example we consider a simplified trade procedure in which the LoC determines that the seller must produce a bill of lading (BoL) as evidence of performance according to contract. The seller receives the BoL from the carrier when he transfers the goods to the carrier. There is only one copy of the BoL that is authentic; this copy is evidence of ownership of the goods. The seller transfers the BoL to the corresponding bank, who pays the seller for the goods. The corresponding bank sells the BoL to the issuing bank, who sells it to the buyer. The buyer now uses the BoL as evidence of ownership in order to receive the goods from the carrier.

In this situation we have a number of (ought-to-be) constraints that must be satisfied in order for the procedure to work. We will show that these constraints entail ought-to-do constraints specifying constraints on the actions executed in the procedure. For example, there is an (ought-to-be) constraint which asserts that it ought to be the case that if the carrier C has the goods then the seller S has the BoL:

$$O_1(\text{Has\_Goods}(C) \rightarrow \text{Has\_BoL}(S))$$

Here and in the sequel the letters C and S are used as identifiers of the carrier and seller, respectively. A transfer from C to S will be denoted C2S and the reverse transfer is denoted S2C, with the transferred item as parameter. Similar notations will also be used involving the corresponding bank CB.

From the above ought-to-be constraint we can derive a related ought-to-do constraint, if we assume some further obvious things like the fact that the seller's transferring the goods to the carrier results in the carrier having the goods, i.e.,

$$[\text{S2C}(\text{Goods})]\text{Has\_Goods}(C)$$

and the fact that if the carrier has no goods then after *not* transferring the goods by the seller to the carrier the carrier still has no goods, viz.

$$\neg\text{Has\_BoL}(S) \rightarrow [\neg\text{C2S}(\text{BoL})]\neg\text{Has\_BoL}(S).$$

For we can now derive an ought-to-do constraint in our logic expressing that in a state where the seller has no BoL it is forbidden that the seller transfers the goods to the carrier while the carrier does not send the BoL to the seller:  $\neg\text{Has\_BoL}(S) \rightarrow \underline{F}_1(\text{S2C}(\text{Goods}) \ \& \ \neg\text{C2S}(\text{BoL}))$ , as follows:

$$\begin{aligned} & \text{O}_1(\text{Has\_Goods}(C) \rightarrow \text{Has\_BoL}(S)) \ f \\ & \exists(\neg(\text{Has\_Goods}(C) \rightarrow \text{Has\_BoL}(S)) \rightarrow V_1) \ f \\ & \exists((\text{Has\_Goods}(C) \wedge \neg\text{Has\_BoL}(S)) \rightarrow V_1). \end{aligned}$$

Thus, using this together with the above assumptions and axiom (&), we have that:

$$\begin{aligned} & \neg\text{Has\_BoL}(S) \rightarrow [\text{S2C}(\text{Goods}) \ \& \ \neg\text{C2S}(\text{BoL})] (\text{Has\_Goods}(C) \wedge \neg\text{Has\_BoL}(S)) \ f \\ & \neg\text{Has\_BoL}(S) \rightarrow [\text{S2C}(\text{Goods}) \ \& \ \neg\text{C2S}(\text{BoL})] V_1 \ f \\ & \neg\text{Has\_BoL}(S) \rightarrow \underline{F}_1(\text{S2C}(\text{Goods}) \ \& \ \neg\text{C2S}(\text{BoL})) \ (1) \end{aligned}$$

Likewise, from the assumptions

$$[\text{C2S}(\text{BoL})]\text{Has\_BoL}(S)$$

and

$$\neg\text{Has\_Goods}(C) \rightarrow [\neg\text{S2C}(\text{Goods})]\neg\text{Has\_Goods}(C)$$

one can show from the ought-to-be constraint  $\text{O}_2(\text{Has\_BoL}(S) \rightarrow \text{Has\_Goods}(C))$  that

$$\neg\text{Has\_Goods}(C) \rightarrow \underline{F}_2(\text{C2S}(\text{BoL}) \ \& \ \neg\text{S2C}(\text{Goods})) \ (2)$$

Together with an obligation that in some situation the seller should supply the goods to the carrier  $\underline{O}_3(\text{S2C}(\text{Goods}))$  (3)—as required in a contract, the constraints (1), (2) and (3) constitute an instantiation of the parallel version of the Chisholm set, which proves again the importance of this set in practice.

In fact, we can also easily derive instantiations of the other versions. For instance, from the ought-to-be constraint that it should be so that if the seller S has got the money then the corresponding bank CB has the BoL:

$O_4(\text{Has\_Money}(S) \rightarrow \text{Has\_BoL}(\text{CB}))$

it follows—using similar assumptions as before such as

$\neg \text{Has\_BoL}(\text{CB}) \rightarrow [\neg \text{S2CB}(\text{BoL})] \neg \text{Has\_BoL}(\text{CB})$

and its consequence

$\neg \text{Has\_BoL}(\text{CB}) \rightarrow [\text{CB2B}(\text{Money}) \ \& \ \neg \text{S2CB}(\text{BoL})] \neg \text{Has\_BoL}(\text{CB})$

—that it is forbidden that the seller gets the money from the corresponding bank without the seller having sent the corresponding bank the BoL first:  $\neg \text{Has\_BoL}(\text{CB}) \rightarrow F_4(\neg \text{S2CB}(\text{BoL}); \text{CB2S}(\text{Money}))$ :

$\neg \text{Has\_BoL}(\text{CB}) \Rightarrow$

$[\neg \text{S2CB}(\text{BoL})] \neg \text{Has\_BoL}(\text{CB}) \Rightarrow$

$[\neg \text{S2CB}(\text{BoL})][\text{CB2S}(\text{Money}) \ \& \ \neg \text{S2CB}(\text{BoL})](\neg \text{Has\_BoL}(\text{CB}) \wedge \text{Has\_Money}(S)) \Rightarrow$

$[\neg \text{S2CB}(\text{BoL})][\text{CB2S}(\text{Money}) \ \& \ \neg \text{S2CB}(\text{BoL})]V_4 \Rightarrow$

$F_4(\neg \text{S2CB}(\text{BoL}); (\text{CB2S}(\text{Money}) \ \& \ \neg \text{S2CB}(\text{BoL})))$

Together with the ought-to-do constraint that can be derived in an analogous manner from the ought-to-be constraint  $O_5(\text{Has\_BoL}(\text{CB}) \rightarrow \text{Has\_Money}(S))$  (using similar additional assumptions), viz  $\neg \text{Has\_Money}(S) \rightarrow F_5((\text{S2CB}(\text{BoL}) \ \& \ \neg \text{CB2S}(\text{Money})); \neg \text{CB2S}(\text{Money}))$ , and a constraint stating that there should be a transfer of money from the corresponding bank to the seller, viz.  $O_6(\text{CB2S}(\text{Money}))$ , we obtain a hybrid instantiation in between the parallel and the backward Chisholm set. Moreover, if we assume additionally:

1. that in the action  $\text{CB2S}(\text{Money})$  the non-performance of  $\text{S2CB}(\text{BoL})$  is included, i.e. it holds that  $\text{CB2S}(\text{Money}) \gg \neg \text{S2CB}(\text{BoL})$  (we would have this if we would consider the action *only*( $\text{CB2S}(\text{Money})$ ) from [DMW94a, DMW96] rather than  $\text{CB2S}(\text{Money})$  which we normally endow with an *open* interpretation leaving open what happens concurrently with it), and

2. likewise,  $\text{S2CB}(\text{BoL})$  involves the non-performance of  $\text{CB2S}(\text{Money})$ , i.e.,  $\text{S2CB}(\text{BoL}) \gg \neg \text{CB2S}(\text{Money})$ ,

we end up with the plain version of the backward Chisholm set, since in this case  $CB2S(Money)$  &  $\neg S2CB(BoL)$  comes down to a mere  $CB2S(Money)$ , and  $S2CB(BoL)$  &  $\neg CB2S(Money)$  to just  $S2CB(BoL)$ :

$$\begin{aligned} &\neg Has\_BoL(CB) \rightarrow F_4(\neg S2CB(BoL); CB2S(Money)) \\ &\neg Has\_Money(S) \rightarrow F_5(S2CB(BoL); \neg CB2S(Money)) \\ &O_6(CB2S(Money)) \end{aligned}$$

Furthermore, we can derive, assuming for simplicity that  $S2CB(BoL) \gg \neg CB2S(Money)$ :

$$\begin{aligned} &\neg Has\_Money(S) \wedge \neg Has\_BoL(CB) \Rightarrow \\ &[S2CB(BoL)](Has\_BoL(CB) \wedge \neg Has\_Money(S)) \Rightarrow \\ &[S2CB(BoL)][\neg CB2S(Money)](Has\_BoL(CB) \wedge \neg Has\_Money(S)) \Rightarrow \\ &[S2CB(BoL)][\neg CB2S(Money)]V_5 \Rightarrow \\ &[S2CB(BoL)]O_5CB2S(Money) \end{aligned}$$

Analogously, we can derive  $\neg Has\_Money(S) \wedge \neg Has\_BoL(CB) \rightarrow [\neg S2CB(BoL)]O_4(\neg CB2S(Money))$ , so that with a ought-to-be constraint  $O_7(S2CB(BoL))$  (effective after S's transferring goods to C, for example) we obtain an instance of the forward Chisholm set.

In passing we observe that it holds that from  $[S2CB(BoL)](Has\_BoL(CB) \wedge \neg Has\_Money(S))$ , we obtain  $[S2CB(BoL)]V_5$  and thus  $F_5S2CB(BoL)$ , which means that even if it is obliged for the seller to send the BoL to the corresponding bank on the grounds of the obligation  $O_7(S2CB(BoL))$ , it is nevertheless forbidden to stop just here without completing the transaction (and do  $CB2S(Money)$ )! Thus this provides an interesting example where we have on the one hand an obligation to do an action  $\alpha$  in order to execute a transaction  $\alpha ; \beta$ , while on the other hand doing (just)  $\alpha$  is forbidden. In our framework we can express this consistently by means of our multiple violation atoms without any difficulty. Note, by the way, that in this example the assumption that  $S2CB(BoL) \gg \neg CB2S(Money)$  is crucial! If it is allowed that  $S2CB(BoL)$  may also involve a concurrent transfer of money from CB to S doing just is (and should) not be forbidden, of course.

## 9. Discussion and Conclusion

In this paper we have seen how deontic logic may be employed for specifying normative integrity constraints for information systems. A key claim of our paper is that ought-to-be and ought-to-do constraints follow different logics, but also that these logics can be integrated enabling representing and reasoning about both types of constraints in one single framework. A side issue, but nevertheless an important one as to the topic of this paper, is that in order to deal

with problematic cases of representation, known from the philosophical literature, one may adhere to a pragmatic view and use logics that are adequate for concrete situations though not necessarily so in general: to specify concrete constraints for practical systems one need not solve all profound problems that philosophy poses for the general abstract case!

## References

- [AM81] C.A. Alchourrón & D. Makinson, Hierarchies of Regulations and Their Logic, in: *New Studies in Deontic Logic* (R. Hilpinen, ed.), Reidel, Dordrecht / Boston, 1981, pp. 125-148.
- [AMW93] P. d'Altan, J.-J. Ch. Meyer & R.J. Wieringa, An Integrated Framework for Ought-to-Be and Ought-to-Do Constraints (extended abstract), in *Working Notes AAAI Spring Symposium on Reasoning about Mental States* (J. Horty, Y. Shoham, J. Doyle, H. Levesque & M. Pollack, eds.), Stanford University, Palo Alto, pp. 48-57;
- [AMW96] P. d'Altan, J.-J. Ch. Meyer & R.J. Wieringa, An Integrated Framework for Ought-to-Be and Ought-to-Do Constraints, *Artif. Intell. and Law* 4, 1996, pp. 77-111.
- [And58] A.R. Anderson, A Reduction of Deontic Logic to Alethic Modal Logic, *Mind* 67, 1958, pp. 100-103.
- [BM88] J.W. de Bakker & J.-J. Ch. Meyer, Metric Semantics for Concurrency, *BIT* 28, 1988, pp. 504-529.
- [BW90] J.C.M. Baeten & W.P. Weijland, *Process Algebra*, Cambridge University Press, 1990.
- [BK86] J.A. Bergstra & J.W. Klop, Algebra of Communicating Processes, in: *Math. & Comp. Sc. I* (J.W. de Bakker, M. Hazewinkel & J.K. Lenstra, eds.), CWI Monographs 1, North-Holland, Amsterdam, 1986, pp. 89-138.
- [BLW94] R.W.H. Bons, R.M. Lee, and R.W. Wagenaar, Implementing the Electronic Bill of Lading, Technical Report, EURIDIS, Rotterdam, December 1994.
- [Cas81] H.-N. Castañeda, The Paradoxes of Deontic logic: The Simplest Solution to All of Them in One Fell Swoop, in: R. Hilpinen (ed.), *New Studies in Deontic Logic: Norms, Actions and the Foundation of Ethics*, Reidel, Dordrecht, 1981, pp. 37-85.
- [Cas84] H.-N. Castañeda, Aspectual Actions and the Deepest Paradox of Deontic Logic, Davidson Conference, Rutgers University, April 28, 1984.
- [Che80] B.F. Chellas, *Modal Logic: An Introduction*, Cambridge University Press, Cambridge / London, 1980.
- [Chi63] R.M. Chisholm, Contrary-to-Duty Imperatives and Deontic Logic, *Analysis* 24, 1963, pp. 33-36.
- [DM90] F.P.M. Dignum & J.-J.Ch. Meyer, Negations of Transactions and Their Use in the Specification of Dynamic and Deontic Integrity Constraints, in: *Semantics for Concurrency*, Leicester 1990, (M.Z. Kwiatkowska, M.W. Shields & R.M. Thomas, eds.), Springer-Verlag, London/Berlin, 1990, pp. 61-80.

- [DMW94a] F. Dignum, J.-J. Ch. Meyer & R.J. Wieringa, Contextual Permission: A Solution to the Free Choice Paradox, in Proc. 2nd Int. Workshop on Deontic Logic in Computer Science (DEON'94), (A.J. Jones & M. Sergot, eds.), Tano A.S., Oslo, 1994, pp. 107-135.
- [DMW94b] F. Dignum, J.-J. Ch. Meyer & R.J. Wieringa, A Dynamic Logic for Reasoning about Sub-Ideal States, in Proc. ECAI'94 Workshop "Artificial Normative Reasoning" (J. Breuker, ed.), Amsterdam, 1994, pp. 79-92.
- [DMW96] F. Dignum, J.-J. Ch. Meyer & R.J. Wieringa, [Free Choice and Contextually Permitted Actions](#), *Studia Logica* 57(1), pp. 193-220, 1996.
- [For84] J.W. Forrester, Gentle Murder, or the Adverbial Samaritan, *J. of Philosophy* 81(4), 1984, pp. 193-197.
- [vG90] R.J. van Glabbeek, Comparative Concurrency Semantics, with Refinements of Actions, Ph. D. Thesis, Free University, Amsterdam, 1990.
- [Har79] D. Harel, First-Order Dynamic Logic, Springer-verlag, Berlin, 1979.
- [Hil93] R. Hilpinen, Actions in deontic Logic, in: J.-J. Ch. Meyer & R.J. Wieringa (eds.), *Deontic Logic in Computer Science: Normative System Specification*, John Wiley & Sons Ltd., Chichester, 1993, pp. 85-100.
- [Hoa69] C.A.R. Hoare, An Axiomatic Basis for Computer Programming, *Communications ACM* 12, 1969, pp. 576-580.
- [HMT94] W. van der Hoek, J.-J. Ch. Meyer & J. Treur, Formal Semantics of Temporal Epistemic Reflection, in [Logic Program Synthesis and Transformation - Meta-Programming in Logic](#), 4th Int. Workshops, LOPSTR'94 and META'94, Pisa, 1994, (L. Fribourg & F. Turini, eds.), LNCS 883, 1994, pp. 332-352.
- [Hor93] J.F. Horty, Deontic Logic as Founded on Nonmonotonic Logic, *Annals of Math. & Artif. Intell.* 9(1, 2), 1993, pp. 69-91.
- [Jon90] A.J.I. Jones, Deontic Logic and Legal Knowledge Representation, *Ratio Juris* 3, 1990, pp. 237-244.
- [Jon93] A.J.I. Jones, Towards a Formal Theory of Defeasible Deontic Conditionals, *Annals of Math. & Artif. Intell.* 9(1, 2), 1993, pp. 151-166.
- [KL86] S. Kraus & D. Lehmann, Knowledge, Belief and Time, in: L. Kott (ed.), *Proceedings of the 13th Int. Colloquium on Automata, Languages and Programming*, Rennes, LNCS 226, Springer, Berlin, 1986.
- [Krö87] F. Kröger, *Temporal Logic of Programs*, Springer-Verlag, Berlin, 1987.
- [Kwa93] K.L. Kwast, A Deontic Approach to Database Integrity, *Annals of Math. and Artif. Intell.* 9(1, 2), 1993, pp. 205-238.
- [LHM95] B. van Linder, W. van der Hoek & J.-J. Ch. Meyer, The Dynamics of Default Reasoning, in: *Symbolic and Quantitative Approaches to Reasoning and Uncertainty (Proc. ECSQARU'95, Fribourg, Switzerland, July 1995)* (Chr. Froidevaux & J. Kohlas, eds.), LNCS 946, Springer, Berlin, 1995, pp. 277-284.

- [Luk90] W. Lukasiewicz, *Non-Monotonic Reasoning, Formalization of Commonsense Reasoning*, Ellis Horwood, New York, 1990.
- [MP92] Z. Manna & A. Pnueli, *The Temporal Logic of Reactive and Concurrent Systems: Specification*, Springer-Verlag, Berlin, 1992.
- [MT93] V.W. Marek & M. Truszczyński, *Nonmonotonic Logic, Context-Dependent Reasoning*, Springer-Verlag, Berlin, 1993.
- [Mey87] J.-J.Ch. Meyer, A Simple Solution to the 'Deepest' Paradox of Deontic Logic, in *Logique et Analyse* 117-118 (1987), pp. 81-90; also in *Proc. 6th Amsterdam Colloquium* (1987), pp. 281-290, under the title "The Ethics of Murder".
- [Mey88] J.-J.Ch. Meyer, A Different Approach to Deontic Logic: Deontic Logic Viewed as a Variant of Dynamic Logic, *Notre Dame J. of Formal Logic* 29(1), (1988), pp. 109-136.
- [Mey92] J.-J.Ch. Meyer, Free Choice Permissions and Ross's Paradox: Internal vs External Nondeterminism, in *Proc. 8th Amsterdam Colloquium* (1991) (P. Dekker & M. Stokhof, eds.), ILLC, Univ. of Amsterdam, Amsterdam, 1992, pp. 367-380.
- [MH95] J.-J. Ch. Meyer & W. van der Hoek, *Epistemic Logic for AI and Computer Science*, Cambridge Tracts in Theoretical Computer Science 41, Cambridge University Press, 1995.
- [MWW89] J.-J.Ch. Meyer, H. Weigand & R.J. Wieringa, A Specification Language for Static, Dynamic and Deontic Integrity Constraints, in *Proc. MFDBS 89*, J. Demetrovics & B. Thalheim (eds.), Visegrád, Hungary, LNCS 364, Springer, Berlin, 1989, pp. 347-366.
- [Mil80] R. Milner, *A Calculus of Communicating Systems*, Lect. Notes in Comp. Sc. 92, Springer-Verlag, Berlin, 1980.
- [MW91] J.-J. Ch. Meyer & R.J. Wieringa, Actor-Oriented System Specification with Dynamic Logic, in *Proc. TAPSOFT/CCPSD '91* (S. Abramsky & T.S.E. Maibaum, eds.), Brighton, LNCS 494, Springer, 1991, pp. 337-357,
- [MW91a] J.-J. Ch. Meyer & R.J. Wieringa (eds.), *Proceedings of the First International Workshop on Deontic Logic in Computer Science (DEON'91)*, VU Amsterdam, 1991, 412 p.
- [MW93a] J.-J. Ch. Meyer & R.J. Wieringa, Deontic Logic: A Concise Overview, in: J.-J. Ch. Meyer & R.J. Wieringa (eds.), *Deontic Logic in Computer Science: Normative System Specification*, John Wiley & Sons Ltd., Chichester, 1993, pp. 3-16.
- [MW93b] J.-J. Ch. Meyer & R.J. Wieringa (eds.), *Deontic Logic in Computer Science: Normative System Specification*, John Wiley & Sons Ltd., Chichester, 1993, xiv + 317 p.
- [Pra93] H. Prakken, An Argumentation Framework in Default Logic, *Annals of Math. & Artif. Intell.* 9(1, 2), 1993, pp. 93-132.
- [Pra94] H. Prakken, Two Approaches to Defeasible Reasoning, in: *Proc. DEON'94* (A.J.I. Jones & M.J. Sergot, eds.), Tano A.S., Oslo, 1994, pp. 281-295.
- [PS94] H. Prakken & M.J. Sergot, Contrary-to-Duty Imperatives, Defeasibility and Violability, in: *Proc. DEON'94* (A.J.I. Jones & M.J. Sergot, eds.), TANO, Oslo, 1994, pp. 296-318.
- [Rya93] M. Ryan, Towards Specifying Norms, *Annals of Math. & Artif. Intell.* 9(1, 2), 1993, pp. 49-67.

- [Seg89] K. Segerberg, Bringing It About, *J. of Philos. Logic* 18(4), 1989, pp. 327-347.
- [Smi94] T. Smith, *Legal Expert Systems: Discussion of Theoretical Assumptions*, Ph.D. Thesis, Utrecht University, 1994.
- [vdT94] L.W.N. van der Torre, Violated Obligations in a Defeasible Deontic Logic, *Proc. ECAI'94* (A.G. Cohn, ed.), John Wiley & Sons Ltd., Chichester, 1994, pp. 371-375.
- [WM91] R.J. Wieringa & J.-J. Ch. Meyer, Actor-Oriented Specification of Deontic Integrity Constraints, in *Proc. MFDBS '91* (B. Thalheim, J. Demetrovics & H.-D. Gerhardt, eds.), Rostock, LNCS 495, Springer, 1991, pp. 89-103.59.
- [WM93] R.J. Wieringa & J.-J. Ch. Meyer, Applications of Deontic Logic in Computer Science: A Concise Overview, in J.-J. Ch. Meyer & R.J. Wieringa (eds.), *Deontic Logic in Computer Science: Normative System Specification*, John Wiley & Sons Ltd., Chichester, 1993, pp. 17-40.
- [WM93] R.J. Wieringa & J.-J. Ch. Meyer, Actors, Actions, and Initiative in Normative System Specification, *Annals of Mathematics and Artificial Intelligence* 7, 1993, pp. 289-346.
- [WMW89] R.J. Wieringa, J.-J. Ch. Meyer & H. Weigand, Specifying Dynamic and Deontic Integrity Constraints, *Data & Knowledge Engineering* 4(2), 1989, pp. 157-190.
- [WWMD91] R.J. Wieringa, H. Weigand, J.-J. Ch. Meyer & F.P.M. Dignum, The Inheritance of Dynamic and Deontic Integrity Constraints, *Annals of Math. and Artif. Intell.* 3, 1991, pp. 393-428.
- [Win82] G. Winskel, Event Structure Semantics for CCS and Related Languages, in: *Proc. 9th ICALP*, Aarhus (M. Nielsen & E.M. Schmidt, eds.), *Lect. Notes in Comp. Sc.* 140, Springer-Verlag, Berlin, 1982, pp. 561-576.
- [vW51] G.H. von Wright, Deontic Logic, *Mind* 60, 1951, pp. 1-15.
- [vW64] G.H. von Wright, A New System of Deontic Logic, *Danish Yearbook of Philosophy* 1, 1964, pp. 173-182.
- [vW81] G.H. von Wright, On the Logic of Norms and Action, in: R. Hilpinen (ed.), *New Studies in Deontic Logic: Norms, Actions and the Foundation of Ethics*, Reidel, Dordrecht, 1981, pp. 3-35.
- [Åqv84] L. Åqvist, Deontic Logic, in: *Handbook of Philosophical Logic*, Vol. II (D. Gabbay & F. Guenther, eds.), Reidel, Dordrecht, 1984, pp. 605-714.