ALTERNATIVE REDUCTIONS FOR DYNAMIC DEONTIC LOGICS

ALBERT J.J. ANGLBERGER
University of Salzburg

Abstract
A recent approach to deontic logic – the logic of permission, obligation and prohibition – places deontic logic into a dynamic framework. In dynamic logics we differentiate between actions and assertions. For every action term α, an execution operator '$[\alpha]\ldots$' is introduced, which is read as 'every execution of α leads to a state in which … holds'. Enriching our language by a violation constant V, allows us to reduce deontic predicates in two obvious ways: (i) An action is forbidden iff every execution leads to a violation, (ii) an action is forbidden iff at least one execution leads to a violation. Both reductions lead – besides being somewhat coarse grained – to implausible theorems. In our paper we will address the question: where and how one may find more sophisticated reductions.

1 INTRODUCTION: STANDARD DEONTIC LOGIC AND ANDERSONIAN REDUCTIONS

In the standard systems of deontic logic – the logic of obligations, permissions and forbiddances, originating from [von Wright 1951] – the language of propositional logic is extended by (at least one of) the normative operators 'O', 'P' and 'F' for 'it ought to be the case that', 'it is permitted that it is the case that' and 'it is forbidden that it is the case that', respectively. At least one of these operators is primitive in those systems. Taken 'O' as primitive, the other operators are usually introduced via the definitions

(DF) $F\varphi \leftrightarrow O\neg\varphi$, where φ is an arbitrary formula

and

(DP) $P\varphi \leftrightarrow \neg O\neg\varphi$, where φ is an arbitrary formula.

A system of deontic logic is said to have the property of strong interdefinability iff (DP) and (DF) hold in it. Furthermore we call a deontic logic *re-*

ductive iff every formula containing a deontic operator is equivalent to a formula without deontic operators. In a reductive deontic logic normative operators can therefore be eliminated.

The proof theory of the simplest deontic logic is the modal logic D and its semantics therefore a Kripke-semantics with serial frames. Additionally it contains (DP) and (DF). From now on we shall call this logic the *standard system of deontic logic* (short: *SDL*). Hence, we may call SDL non-reductive, and SDL has the property of strong interdefinability.

In [Anderson 1958] Alan Ross Anderson presented a reduction of deontic logic to alethic modal logic. He enriched the modal base language by the propositional constant '*S*', which he interpreted as occurrence of a sanction. His reduction was

$$\text{(AR) } O\varphi \leftrightarrow \Box(\neg\varphi \rightarrow S),$$

meaning that it ought to be the case that φ iff it is necessary that $\neg\varphi$ implies a sanction. The other deontic operators were introduced via (DF) and (DP). Obviously, Anderson's deontic logic is reductive and has the property of strong interdefinability.

(AR) may be criticized for several reasons. E.g. the formula

$$\text{(PR) } P\varphi \leftrightarrow \Diamond(\varphi \wedge \neg S)$$

can be derived from (AR), (DP) and its modal base logic. In words: It is permitted that φ iff it is possible that φ but no sanction occurs. (PR) suggests a very weak notion of permission, maybe a notion too weak. However, it can be shown that

$$\text{(PA) } Op \rightarrow \Box Op$$

and

$$\text{(PA*) } \Box(p \rightarrow q) \rightarrow (Op \rightarrow Oq)$$

are theorems in Anderson's logic too. (PA) excludes obligations that are not necessary and (PA*) yields the (in)famous *Good Samaritan Paradox*. Additionally**,** Anderson's deontic logic contains *SDL*, and with it all its paradoxes. This shows – as it has been argued – that the reduction (AR) was not successful.

In 1950 a similar reduction has already been proposed by Stig Kanger in an unpublished manuscript. Kanger discusses in [Kanger 1971] his reduction

(KR) $O\varphi \leftrightarrow \Box(Q \to \varphi)$,

where 'Q' is a propositional constant stating what is morally prescribed. (KR) hence reads 'φ ought to be the case iff it is necessary that the morally prescribed implies φ'. Defining 'Q' as '$\neg S$', Kanger's reduction (KR) turns out to be equivalent to Anderson's reduction (AR).

The implausible theorems depend on the specific reduction and on the modal base logic. Is it possible to avoid them in a different framework? In the next sections we will explore more recent approaches to such reductions. These reductions are formulated in a different formal language, hence, in a different base logic: in the language of a dynamic logic.

2. DYNAMIC DEONTIC LOGICS: KNOWN REDUCTIONS

2.1 Dynamic Logics: Execution Operators

In the language of a dynamic logic we differentiate between *action terms* and *assertions*. The language of propositional logic is extended by the undefined *execution operator* '[..]…' which applied to an action term α and a formula φ yields the formula

'$[\alpha]\varphi$'

which reads: 'every execution of α leads to a state in which φ holds'.

The dual operator '$\langle\alpha\rangle\varphi$' is defined as '$\neg[\alpha]\neg\varphi$', which reads 'there is an execution of α that leads to a state in which φ holds'. Instead of '$[\alpha]\varphi$' it is sometimes said that α *leads* to φ, and '$\langle\alpha\rangle\varphi$' is also read as '$\alpha$ *may* lead to φ'.

With the operator '$\langle..\rangle…$' we can elegantly define that an action is possible in a certain state: An action is possible iff there is a way to execute that action such that after its execution a tautology holds. In our formal language:

(DPoss) $Possible(\alpha) \leftrightarrow \langle\alpha\rangle\top$,

where '⊤' represents an arbitrary tautology. In its simplest form, the proof theory of '[..]…' is the modal logic K, extended by special axioms for complex actions.

2.2 Dynamic Logics: Actions

We start with a denumerable set of *atomic action terms*. A very natural way to interpret actions is to assign to each atomic action a set of ordered pairs of states (i.e. possible worlds). The first element of such a pair is a *start state*, the second element a *result state*.

On the basis of atomic actions we may construct several *complex actions*, for example

(Negation)	$\sim\alpha$	… not-α
(Choice)	$\alpha\cup\beta$	… α or β
(Conjunction)	$\alpha\&\beta$	… α and β
(Sequence)	$\alpha;\beta$	… α followed by β

One natural way to interpret these complex actions is as *set theoretic complement (w.r.t. to the set of all ordered pairs of possible worlds), union, intersection and relative product,* respectively. Depending on what kind of complex actions are allowed in the language, we get different proof theories. Unfortunately, a proof theory for & interpreted as intersection gets already very complicated (because & interpreted as intersection is not modally definable, see [Balbiani 2003]). Moreover, the question of how to interpret action negation remains a field of research in its own right (see for example [Broersen 2004]).

2.3 Dynamic Deontic Logics: Reductions

Looking back to the beginning of dynamic logic, Krister Segerberg has already developed a reduction of deontic logic to a dynamic logic in [Segerberg 1980]. Segerberg defined

(SR) $P\alpha \leftrightarrow [\alpha]OK$,

where '*OK*' is a propositional constant, expressing that a state is "deontically satisfactory". (SR) therefore calls an action permitted iff every execution of that action leads to a state which is deontically satisfactory.

Nevertheless, the first systematic contribution to dynamic deontic logic is to be found in Meyer's famous paper [Meyer 1988]. Meyer enriched the dynamic base language by the propositional constant V – standing for *violation* – which is said to be true in a state iff a violation occurs in that state. He then proposed the following reduction:

(R1) $F\alpha \leftrightarrow [\alpha]V$,

which calls an action α forbidden iff every execution of α leads to a violation. Additionally, we find the strong interdefinabilities

(DOD) $O\alpha \leftrightarrow F\sim\alpha$ and
(DPD) $P\alpha \leftrightarrow \neg F\alpha$

in Meyer's deontic logic PD_eL. (R1), (DOD) and (DPD) imply that PD_eL is reductive too. (R1) and (DPD) lead – together with some basic principles of dynamic logic – to

(MP) $P\alpha \leftrightarrow \langle\alpha\rangle\neg V$,

meaning that it is permitted to do α iff there is a way to execute α that leads to a state in which no violation occurs. (MP) and (R1) look very similar to Anderson's reduction (AR), and (PR). This justifies calling (R1) – as Meyer did himself – an *Andersonian reduction*. The crucial question now is: Does the shift from a "classical" modal base logic to a dynamic logic prevent all implausible theorems – usually called *paradoxes* – from being provable? Meyer had great hopes in his reduction. He claims that "reducing deontic logic to dynamic logic kills two birds with one stone. Most importantly, in this way we get rid of most (if not all) of the nasty paradoxes that have plagued traditional deontic logic." [Meyer 1993, p.11] Unfortunately, it was shown by R. van der Meyden in [vd Meyden 1990] and [vd Meyden 1996], and by myself in [Anglberger 2008], that some pretty nasty paradoxes are provable in PD_eL. Van der Meyden showed that the paradoxical formula

(vdM) $\langle\alpha\rangle P\beta \rightarrow P(\alpha;\beta)$

is a theorem in Meyer's logic. For its proof only (R1) and some basic principles of dynamic logic are needed. (vdM) has the highly implausible instantiation:

(vdMi) If there is an execution of shooting the president after which it is permitted to remain silent, then it is permitted to shoot the president followed by remaining silent.

In [Anglberger 2008] I was able to prove another implausible theorem in PD_eL, the formula

(T3) $F\alpha \rightarrow [\alpha]F\beta$.

(T3) says that after a forbidden action has been done, any arbitrary action is forbidden. (T3) may be regarded as the dynamic pendant of the *Paradox of Derived Obligation* (since $F\beta \leftrightarrow O\sim\beta$). This last paradox also has consequences for the treatment of contrary to duty imperatives, see [Anglberger 2008, p.430]. Additionally, if we extend Meyer's logic with the dynamic pendant of the axiom **D**

(Dd) $O\alpha \rightarrow P\alpha$

from classical deontic logic (which may be desired), we will be able to derive a theorem stating that no possible action is forbidden (see [Anglberger 2008, p.432]). But van der Meyden's paradox still seems to be the more interesting one. This is because in (T3)'s proof Meyer's full action algebra (and of course the dynamic base logic) was needed. The preconditions of (vdM)'s proof are therefore logically weaker and show with much more distinctness that (R1) needs to be given up, if we want to keep the dynamic framework.

In his paper [Broersen 2004] on action negation, J. Broersen proposed a different reduction. His reduction was

(R2) $P\alpha \leftrightarrow [\alpha]\neg V_p$,

where V_p is a propositional constant representing a special kind of violation ("permission-violation"). Broersen gave up strong interdefinability, he rather suggested a separate reduction for each deontic predicate. But his logic is still reductive; every deontic predicate can be eliminated. Furthermore, if 'OK' is defined as '$\neg V_p$' Broersen's 2004 reduction (R2) is equivalent to Segerberg's 1980 reducion (SR).

It was observed by Broersen himself that (R2) presupposes a special reading of actions. (R2) implies

(R2T) $P\alpha \rightarrow P(\alpha\&\beta)$,

which at first sight looks highly implausible. But under the so-called *open reading of actions* (R2T) becomes valid. The open reading of actions means that, whenever we talk about an action, we talk about *every way to perform* this action. Since $\alpha\&\beta$ is a way to perform α, $P(\alpha\&\beta)$ is implied by $P\alpha$. This validates (R2T). However, there definitely is a reason, why (R2T) looks paradoxical. Though quite interesting and elegant, the open reading of actions is simply not the one used in our ethical and juridical discourse. That is why, from an intuitive point of view, nearly everyone would immediately reject (R2T). Should not a deontic logic somehow model our (or at least: *one* of our) usage(s) of deontic expressions? Since a deontic logic is designed to be applied in ethical and juridical discourse, it seems justified to look for more "sophisticated" reductions.

3 ALTERNATIVE REDUCTIONS

3.1 A Bit of History

The paradoxes of classical deontic logics are usually regarded as a touchstone for any new system of deontic logic. As quoted above, Meyer hoped that his system would be able to solve all (or nearly all) of these paradoxes. However, when working out a solution for – as he calls it – "the deepest paradox in deontic logic" (i.e. the Good Samaritan Paradox) in [Meyer 1987], he recognized that his simple reduction was not sophisticated enough. He had to introduce more than just one violation constant. Therefore he extended the language of the dynamic base logic by n many violation constants $V_1, V_2,…, V_n$ where the index i ($1 \leq i \leq n$) of V_i indicates a violation of the i-th degree. Furthermore, he proposed n new reductions of the form

(R1n) $F_i(\alpha) \leftrightarrow [\alpha]V_i$, where $i = 1, 2,…, n$.

With (R1n) he was able to express certain *degrees of forbiddance*: '$F_1(\alpha)$' was interpreted as 'α is forbidden to degree one', '$F_3(\beta)$' as 'β is forbidden

to degree three' etc. In this example β is thus "more forbidden" than α. But this reduction is not able to solve the problems mentioned above: (R1n) is simply (R1) for every level $i = 1, 2,\ldots, n$. It is therefore possible to prove the aforementioned paradoxical formulae on every level. But in our opinion, it hints at a possible way for constructing new reductions.

3.2 New Reductions

3.2.1 Ordering violations

Let $L_{\mathrm{dyn}(n)}$ be a language of dynamic logic extended by a finite number n of violation constants $V_1, V_2,\ldots, V_n$. V_n is true in a state iff a violation of the n-th degree occurs in this state. Interestingly, at the end of his paper [Meyer 1987] Meyer discusses the possibility of ordering these violations:

> The extension of the number of propositional variables related to sanctions also raises the question whether it is meaningful to put an ordering (e.g. a partial ordering) on them. $V_i \leq V_j$ for instance, would then express that sanction j is more severe than sanction i. Perhaps even $V_j \supset V_i$ [in our terminology: $V_j \to V_i$] can be chosen as an ordering. In this case, $[\alpha]V_j \wedge V_i \equiv [\alpha]V_j$ [in our terminology: $[\alpha]V_j \wedge V_i \leftrightarrow [\alpha]V_j$], since V_j comprises V_i entirely. [Meyer 1987, p.89]

Meyer did not work out this stimulating thought, though it opens up very interesting possibilities. In fact, we will use Meyer's own suggestion – the ordering of violations – as our only axiom for violations:

> (AxV) $V_i \to V_{i-1}$, for all i (*1 < i ≤ n*), where n is the number of violation constants in $L_{dyn(n)}$.

This axiom may be motivated along various lines. For example, consider the following a bit more concrete application: V_n expresses that one has to stay in prison for n days – a possible measure of a violation. If one has to stay in prison for n days, one also has to stay in prison for n–1 days. In this application (AxV) is obviously valid. In general: Whenever a state is "bad" to a certain degree, it also is "bad" to every degree below.

To achieve model theoretic validity of (AxV) we just have to postulate the according condition in the definition of a Kripke-model $\langle W, R, I \rangle$. In addition to the usual requirements on W (the set of possible worlds), R

(interpretation of actions) and *I* (valuation) the following condition has to be met:

(CAxV) $I(V_i) \subseteq I(V_{i-1})$, for all violation constants V_i ($1 < i \leq n$)

3.2.2 Evaluating the deontic status of actions

Consider the following situation: An airplane is about to crash because of some technical failure. Its pilot cannot prevent the airplane from crashing. Suppose further, that only two actions are available to the pilot: The first option is to crash-land the plane and by that maybe saving some people, the second option is to refrain from crash-landing the plane and by that killing all passengers with certainty. Though both actions would lead to an undesirable state of affairs – to a violation –, we would intuitively only call the second one *forbidden*. The reason for this is quite simple: In judging actions we usually relate them to other available options; if we intend to blame a person for having done something, we ask beforehand, what *she could have done instead*. Relating to our "airplane example": If the pilot tried to crash-land the plane, we could not blame her for having done something wrong, because it was the best action among its alternatives (actually, there was only one alternative to the crash-landing).

   Summing this up: When evaluating the deontic status of an action we usually relate this action to other available alternatives. A (realistic) deontic logic should also somehow consider this idea. At this point one question naturally arises: Can this be expressed within the language $L_{dyn(n)}$ of a dynamic (deontic) logic?

3.2.3 Introducing new reductions

In dynamic logic an action can easily be compared with one of its alternatives: We just have to compare an action with its negation. Of course, what exactly is compared here depends on how action negation is modelled. If action negation is modelled as set theoretic complement – as mentioned above – it is not clear, whether it is useful in ethical and juridical discourse (for a more extensive critique see [Broersen 2004]). So this model of action negation does probably not suit our purposes. But there is a more realistic account: Broersen's relativized action negation from [Broersen 2004] models refraining from an action. Within Broersen's account doing ~α means refraining from doing α. And refraining from doing α is a choice

among (all) other available actions, i.e. (in a way) α's alternatives. With Broersen's relativized action negation we can express the comparison needed. This leads to the question: In what respect do we have to compare α with $\sim\alpha$? In our formal language we can compare them with respect to their *outcomes* (i.e. result states). Consider the formula

$$[\alpha]V_3 \wedge [\beta]\neg V_3.$$

The ordering implies that this formula expresses that doing β never leads to a greater violation than V_2 (otherwise the right conjunct would have to be false). This in turn means that every execution of α is worse – i.e. leads to a greater violation – than every execution of β. Furthermore, when comparing two actions, the ordering of violations allows us to measure the upper bound of their outcomes w.r.t. violations. E.g. the formula

$$\exists k([\alpha]V_k \wedge [\beta]\neg V_k)$$

expresses that there is a certain degree of violation such that every execution of α leads to that violation whereas every execution of β does not. The index k is the upper bound of these violations that may be led to by α (but not by β). The existential quantifier is justified by the fact that only finitely many violation constants are allowed in $L_{dyn(n)}$. There is no need for quantification over propositional variables as developed in [Shilov 1997] – it is a mere technical convenience: Let's suppose the number of violation constants is n. The formula above may then be rewritten (without the existential quantifier) as

$$([\alpha]V_1 \wedge [\beta]\neg V_1) \vee ([\alpha]V_2 \wedge [\beta]\neg V_2) \vee \ldots \vee ([\alpha]V_n \wedge [\beta]\neg V_n)$$

i.e. a disjunction of n conjunctions.

Substituting '$\sim\alpha$' for 'β' in the above formula, we get the formula

$$\exists k([\alpha]V_k \wedge [\sim\alpha]\neg V_k),$$

which says that $\sim\alpha$ leads necessarily to a better state than α – every execution of α leads to a greater violation than every execution of $\sim\alpha$. One might now suggest the following reduction:

(NR1) $F\alpha \leftrightarrow \exists k([\alpha]V_k \wedge [\sim\alpha]\neg V_k),$

which calls an action α forbidden iff $\sim\alpha$'s outcome is always better than α's outcome. This seems to be too strong. A more useful concept could be defined by

$\quad$ (NR1*) $F\alpha \leftrightarrow \exists k([\alpha]V_k \wedge \langle\sim\alpha\rangle\neg V_k)$.

(NR1*) defines an action α as forbidden iff $\sim\alpha$ *may lead* to a better state than every execution of α. Consider our "airplane example" once again: There is an execution of refraining from doing nothing – e.g. crash landing the plane – which *may lead* to a better state (some people may be saved) than every "execution" of doing nothing (all people are going to die with certainty). That is why we would call the "action" of doing nothing forbidden. A second relevant factor seems to be, that, if an action α is forbidden, $\sim\alpha$ has to be possible (the meaning of an action being possible was defined above). This is a version of the *Principle of Alternate Possibilities* (= PAP) and – at least in the dynamic formalization with strong interdefinabilities – a version of the *Ought Implies Can Principle*. This leads to a second and better reduction:

$\quad$ (NR1**) $F\alpha \leftrightarrow Possible(\sim\alpha) \wedge \exists k([\alpha]V_k \wedge \langle\sim\alpha\rangle\neg V_k)$

However, there is still one relevant condition missing: If we call an action α forbidden, it may lead to a violation. If every execution of α leads solely to states where no violation whatsoever occurs, α would not be called forbidden. One natural way to express this is the formula

$\quad$ $\langle\alpha\rangle\,(V_1 \vee V_2 \vee \ldots \vee V_n)$.

Though it is a sound formalization of this thought, the formula still contains redundancies. Since we may use our ordering axiom, we can easily verify that

$\quad$ $\langle\alpha\rangle V_1$

is a shorter but equivalent formulation. This leads us to our final proposal:

$\quad$ (NR1F) $F\alpha \leftrightarrow \langle\alpha\rangle V_1 \wedge Possible(\sim\alpha) \wedge \exists k([\alpha]V_k \wedge \langle\sim\alpha\rangle\neg V_k)$

In a nutshell, (NR1F) calls an action α forbidden iff the following conditions are satisfied:

(i) α may lead to a violation,
(ii) it has to be possible to refrain from doing α,
(iii) if one refrains from doing α, we may end up in a better state than what is reachable by doing α.

This reduction is much more complex than (R1) and (R2), and obviously the mentioned paradoxes vanish. Whether (NR1F) leads to new paradoxes has to be further investigated. Anyway, (NR1F) seems to be a more realistic concept of *forbiddance*. A fact that justifies our hope of (possible) implausible theorems not exceeding a certain (reasonable) extent.

3.3 Further Research

The language $L_{dyn(n)}$ can easily be extended. For example, we could add *m reward constants $R_1,...,R_m$* to $L_{dyn(n)}$. A similar idea was already suggested by Meyer in his paper [Meyer 1988, p.125], though Meyer had only one reward constant in mind. The method developed in our paper – measuring the upper bounds of outcomes – makes it possible to define several concepts of *praiseworthiness*. This would also allow us to make a distinction between actions that ought to be done and actions that are praiseworthy. Praiseworthy actions are "better" than obliged actions; an action might be too good to be obliged.

REFERENCES

[Anglberger 2008] Anglberger, A.J.J.: Dynamic Deontic Logic and its Paradoxes. In: *Studia Logica* 89 (2008), pp.427–435.
[Anderson 1958] Anderson A.R.: A Reduction of Deontic Logic to Alethic Modal Logic. In: *Mind* 67 (1958), pp.100–103.
[Balbiani 2003] Balbiani et al.: PDL with Intersection of Programs: A Complete Axiomatization. In: *Journal of Applied Non-Classical Logics* 13, No.3–4 (2003), pp.231–276.
[Broersen 2004] Broersen, J.: Action Negation and Alternative Reductions for Dynamic Deontic Logics. In: *Journal of Applied Logic* 2 (2004), pp.153–168.
[Kanger 1971] Kanger, S.: New Foundations for Ethical Theory. In: Føllesdal D., Hilpinen R. (eds): *Deontic Logic: Introductory and Systematic Readings*. Dordrecht : D. Reidel Publishing Company, 1971, pp.36–58.

190

[Meyer 1987] Meyer, J.-J.C.: A Simple Solution to the "Deepest" Paradox in Deontic Logic. In: *Logique et Analyse* 117–118 (1987), pp.81–90.

[Meyer 1988] Meyer, J.-J.C.: A Different Approach to Deontic Logic: Deontic Logic Viewed as a Variant of Dynamic Logic. In: *Notre Dame Journal of Formal Logic* 29 (1988), pp. 109–136.

[Meyer 1993] Meyer, J.-J.C.: Deontic Logic: A Concise Overview. In: Meyer, J.-J.C., Wieringa, R.J. (eds.): *Deontic Logic in Computer Science.* Chichester: John Wiley and Sons Ltd., 1993, pp.3–16.

[Segerberg 1980] Segerberg, K.: Applying Modal Logic. In: *Studia Logica* 39 (1980), pp.275–295.

[Shilov 1997] Shilov, N.V.: Program Schemata vs. Automata for Decidability of Program Logics. In: *Theoretical Computer Science* 175 (1997), pp.15–27.

[vdMeyden 1990] Meydern, R. van d.: The Dynamic Logic of Permission. In: *Proc. 5th IEEE Conf. on Logic in Computer Science*, Philadelphia (1990), pp.72–78.

[vdMeyden 1996] Meyden, R. van d.: The Dynamic Logic of Permission. In: *Journal of Logic and Computation* 6 (1996), Nr.3, pp.465–479.

[von Wright 1951] Wright, G.H. von.: Deontic Logic. In: *Mind* 60 (1951) pp.1–15.