# A blueprint for deontic logic in three (not necessarily easy) steps

Krister Segerberg*
(Uppsala University)

A famous paper by Alchourrón, Gärdenfors and Makinson opened up a new avenue of research into the logic of belief and belief change [1]. One of the later extensions is dynamic doxastic logic (DDL), which channels develops the AGM approach as a modal logic. Work in this area continues.

It is noteworthy that the erstwhile interest in theory change of one of the founding fathers of AGM was not in belief change but in normative change. What the late Carlos Alchourrón, professor of jurisprudence, had originally wanted was, it seems, a logic of norms and norm change. Many years later it makes sense to ask whether there is a dynamic deontic logic (DΔL, say) that pursues the ambition that Alchourrón seems to have had. In this note we outline a blueprint of an answer. As mentioned in the title of this note, we proceed in three steps.

Already Georg Henrik von Wright, the founder of modern deontic logic, found that deontic logic must be built on a logic of action. Accordingly, in step 1 we outline a (fairly meagre) logic of action. It avoids a number of important but difficult topics, such as agency, causality and intentionality.

In step 2 we develop a deontic logic which is dynamic in the sense of allowing for what we call real actions. However, it is only in step 3 that also provides for what we call deontic actions. The treatment is sketchy throughout, in particularly towards the end. This is not a finished paper. It is not even a proper abstract of an almost finished paper. It is what it says: a blueprint—and an uneven one at that.[1]

## Step 1: A temporal logic of action

### Model theory

Without giving rigorous explanations, let us outline some key concepts. The fundament of any model will be a set (universe) $U$ of points called the *environment.* Sequences of points will be called paths; they can be either finite or infinite in one or two directions. Two paths $p$ and $q$ can be combined into one path, denoted by $pq$, if $p$ has a last

*Some of the work reported in this note was carried out when the author was a fellow-in-residence at N.I.A.S., the Netherlands Institute of Advanced Study at Wassenaar.

[1]So why publish? One reason is that it gives me a welcome opportunity to publicly thank the organizers and the Dagstuhl staff for a very well organized workshop—very informative, very enjoyable!



Dagstuhl Seminar Proceedings 07351
Formal Models of Belief Change in Rational Agents
http://drops.dagstuhl.de/opus/volltexte/2007/1218

element $p(\#)$ and $q$ has a first element $q(*)$ and the two are the same (if not, we regard the notation $pq$ as meaningless).

Another fundamental model theoretical ingredient is that of a given set $E$ of *actions* or *events* in $U$.[2] An *event* in $U$ is a set of finite paths in $U$. If $a$ is an event and $p \in a$, we say that $p$ *realizes* $E$ ($p$ is a *realization* of $a$). One can think of a number of set-theoretical operations on events under which $E$ is closed, for example, the sum $a \cup b$, the relative product $a \,|\, b$ and the difference $a - b$. (But universal complement is not one of them, nor is the Kleene star.)

Yet another fundamental concept is a given set $H$ of *(complete) histories* in $H$: paths in $U$ that are complete in the sense that if $f$ is a proper subpath of a history $h$, then $f$ is not itself a history. If a history is of the form $hg$, where thus the last element $h(\#)$ of $h$ is also the first element $p(*)$ of $p$, then we will refer to $(h, g)$ as an *articulated history*. One may say that $(h, g)$ represents a particular way of looking at $hg$ with $h$ as the past, $g$ as the future and the point $h(\#) = p(*)$ as the present.

We say that $h$ is a *(possible) past* if $hg \in H$ for some $g$, while $g$ is a *(possible) future* if $hg \in H$ for some $h$, If $h$ is a past, then we write $\mathrm{cont}(h)$ for the set $\{g : hg \in H\}$ of possible continuations (possible futures) of $h$. If $S \subseteq \mathrm{cont}(h)$ we say that $S$ is a *possible open future* of $h$. We refer to $(h, S)$ as a *possible situation*; if $S = \mathrm{cont}(h)$ we say that $(h, S)$ is a possible *actual situation.* If $g \in S$, we refer to $(h, g, S)$ as a *possible scenario.*

If $f$ is a history or a past or a future we say that $f$ *includes* an event $a$ if $f$ contains a subpath that realizes $a$, and that $f$ *excludes* $a$ if there is no such subpath.

## Syntax and meaning conditions

Our object languages must contain a denumerable set of propositional letters (primitive formulæ) $\mathsf{P}_0, \mathsf{P}_1, \cdots, \mathsf{P}_n, \cdots$ and a disjoint denumerable set $\mathsf{e}_0, \mathsf{e}_1, \cdots, \mathsf{e}_n, \cdots$ of event letters primitive terms). In addition there will be an adequate supply of Boolean (truth-functional) connectives as well as special operators to be mentioned; the latter will include at least the sum operator (+) and the catenation operator (;). Whatever the details, our language will contain both formulæ and terms.

A *basic frame* is a triple $(U, E, H)$ such that $U$ is a universe, $E$ is a set of events (with certain closure conditions) and $H$ is a set of complete histories. A *valuation* is a function $V$ from the set of propositional letters into the power set of $U$ and from the set of event letters into $E$. This function is extended in a natural way to all pure Boolean formulæ and to all terms. We will write $[\![\phi]\!]$ for the value assigned to a pure Boolean formula $\phi$ and $[\![\alpha]\!]$ for the value assigned to a term $\alpha$. Examples of meaning conditions ($\phi$ and $\psi$ are pure Boolean formulæ, $\alpha$ and $\beta$ are terms):

$$[\![\neg\phi]\!] = U - [\![\phi]\!],$$

$$[\![\phi \wedge \psi]\!] = [\![\phi]\!] \cap [\![\psi]\!],$$

$$[\![\phi \vee \psi]\!] = [\![\phi]\!] \cup [\![\psi]\!],$$

[2]Many philosophers distinguish between actions and events, as they should. But for the limited purposes of this note it is not important.

$[\![\alpha + \beta]\!] = [\![\alpha]\!] \cup [\![\beta]\!]$,

$[\![\alpha ; \beta]\!] = [\![\alpha]\!] \mid [\![\beta]\!]$.

Relative to such a model it is easy to give meaning conditions also for temporal operators. For example:

$(h, g) \models [\mathsf{F}]\theta$ iff $(h', g') \models \theta$, for all $p, h', g'$ such that $hp = h'$ and $pg' = g$ (and therefore $h'g' = hg$),

$(h, g) \models [\mathsf{P}]\theta$ iff $(h', g') \models \theta$, for all $p, h', g'$ such that $h'p = h$ and $g' = pg$ (and therefore $h'g' = hg$),

$(h, g) \models [\mathsf{H}]\theta$ iff $(h, g') \models \theta$, for all $g' \in \mathrm{cont}(h)$.

Here $[\mathsf{F}]$ and $[\mathsf{P}]$ correspond to Prior's $G$ and $H$, respectively, while $[\mathsf{H}]$ is the operator which has been read variously as "historically necessary", "unavoidably" and "settled true". In a similar way it would be easy to add unary proposition-forming propositional operators such as $[\mathsf{NEXT}]$, $[\mathsf{LAST}]$, $[\mathsf{UNTIL}\,\phi]$ and $[\mathsf{SINCE}\,\phi]$ (for all formulæ $\phi$) with obvious intuitive meanings (we omit the details).

We also want an operator of a more complicated kind, one that involves the consideration of another model. In the notation so far we have suppressed the reference to the model $\mathfrak{M} = (U, E, H, V)$, which has been taken for granted. To be explicit we could have written something like $(h, g) \models^{\mathfrak{M}} \theta$ in stead of just $(h, g) \models \theta$. This perspective is necessary for the definition of $[\mathsf{H} : \phi]$, where again $\phi$ must be a formula:

$(h, g) \models^{\mathfrak{M}} [\mathsf{H} : \phi]\,\theta$ iff $(h, g') \models^{\mathfrak{M}'} \theta$, for all $g' \in \mathrm{cont}_{\mathfrak{M}'}(h)$,
where $\mathfrak{M}' = (U, E, H^{\phi}, V)$ and $H^{\phi} = \{f : (h, f) \models^{\mathfrak{M}} \phi\}$.

This new operator can of course be viewed as—is!—a conditional operator. But noting the validity of the schema $[\mathsf{H}]\,\theta \leftrightarrow [\mathsf{H} : \top]\,\theta$, we think of $[\mathsf{H} : \phi]$ as a kind of "focus" operator: the operator $[\mathsf{H}]$ restricted to or focussed on $\phi$. More specifically, $[\mathsf{H} : \phi]$ focusses on the set of futures described by $\phi$.

The dynamic operators that we need are less common. First there are three proposition-forming term operators $\mathsf{occurs}$, $\mathsf{occurring}$ and $\mathsf{occurred}$:

$(h, g) \models \mathsf{occurs}\,\alpha$ iff $g = pg'$, for some finite path $p \in [\![\alpha]\!]$ and (unique) future $g'$,

$(h, g) \models \mathsf{occurring}\,\alpha$ iff $h = h'p$ and $g = qg'$, for some finite nonempty paths $p$ and $q$ such that $pq \in [\![\alpha]\!]$, (unique) past $h'$ and (unique) future $g'$,

$(h, g) \models \mathsf{occurred}\,\alpha$ iff $h = h'p$, for some finite path $p \in [\![\alpha]\!]$ and (unique) past $h'$.

The dynamic operators also include three complex formula-making operators $[\mathsf{after}\,\alpha]$, $[\mathsf{during}\,\alpha]$ and $[\mathsf{before}\,\alpha]$, where $\alpha$ must be a real term:

$(h, g) \models [\mathsf{after}\,\alpha]\,\phi$ iff $(h', g') \models \phi$, for all finite paths $p$ such that $p \in [\![\alpha]\!]$ and $h' = hp$ and $g = pg'$,

$(h,g) \models [\mathsf{during}\,\alpha]\,\phi$ iff $(h',g') \models \phi$, for all finite paths $p,q$ such that $pq \in [\![\alpha]\!]$ and $h' = hp$ and $g = qg'$,

$(h,g) \models [\mathsf{before}\,\alpha]\,\phi$ iff $(h',g') \models \phi$, for all finite paths $p$ such that $p \in [\![\alpha]\!]$ and $h = h'p$ and $g' = pg$.

Note that Pratt's well-known after-operator $[\alpha]$ is rendered in our idiom as $[\mathsf{H}][\mathsf{after}\,\alpha]$. Similar remarks relate to $[\mathsf{during}\,\alpha]$ and $[\mathsf{before}\,\alpha]$.

*Truth* in a model and *validity* in a frame are defined along traditional lines.

## The result operator

Minimality is a concept that surfaces in connexion with concepts such as conditionals and belief revision. Ramsey was happy to accept a conditional "if A then B" if B would be true in a situation in which things had been changed just enough to make A true. In AGM type belief revision, a new piece of information is incorporated into one's set of beliefs by making a certain minimal adjustment. Makinson has given other examples of conceptual analysis where minimality shows its face.

One such example is offered by what may be called *resultative* actions or events. In everyday life there may be many ways in which a certain state of affairs can result, but talking about them we automatically filter out from consideration ways that are extraordinary or inappropriate. Thus we need to try to capture the notion of "bringing it about that $P$" or "the coming about that $P$", where $P$ is a proposition: the "paradigmatic" or "standard" event resulting in its being the case that $P$. And this is where minimality comes in.

To proceed more formally, say that $f$ is a *selection function* for $U$ if $f$ is defined on the set of subsets of $U$ and the following three conditions are satisfied: for all $P$ and $Q$,

$fP \subseteq P$ ( ),

if $P \subseteq Q$, if $fP \neq \emptyset$ then $fQ \neq \emptyset$ ( ),

if $P \subseteq Q$ and $P \cap fQ \neq \emptyset$ then $fP = P \cap fQ$ ( ).

Let $(U, E, H)$ be a basic frame and $F$ a function defined on $U$ such that, for each $u \in U$, $F_u$ is a selection function for $U$. Assume that $E$ is closed under $F$ in the sense that $F_u P \in E$, for each point $u \in U$ and proposition $P$. Then $F$ is called the *result functor* while $(U, E, F)$ is called a *result frame.*

On the syntactic side we add a new term-forming propositional operator $\partial$ and a new meaning-condition:

$$[\![\partial\phi]\!] = \{(u, v) : v \in F_u[\![\phi]\!]\}.$$

The notions relating to the idea that some actions have results is important, and it is possible to develop it along with other ideas in this note. However, for reasons of simplicity we will not pursue this topic further here.

## Step 2: Deontic logic with real actions

### Pre-theoretical remarks

A norm draws a distinction between what is acceptable and what is not: what is in accordance with the norm and what is not. Legal codes separate legal from illegal, moralities right from wrong and good from bad, conventions correct from incorrect, fashion what is "in" from what is "out", and so on. In real life norms are never sharp enough or complete enough to settle all questions, but in the norms that appear in this note are supposed to be both sharp and complete.

Thus we may think of a complete norm as a norm-giver (legislator, moral genius, arbiter, God) who can answer all questions as to what is normal (= in accordance with the norm) in any given situation. With respect to any given past, however irregular from a normative point of view, the norm-giver should be able to delineate a subset consisting of exactly those futures that are still possible at that time and that are in accordance with the norm.[3]

Traditionally the major deontic notions are obligation, permission and prohibition. In the dominant *Seinsollen* ("ought-to-be") tradition they are treated as concepts applying to propositions. In our modelling, as presented so far, it seems more natural to follow the *Tunsollen* ("ought-to-do") tradition, in which they are treated as concept applying to actions or events. Thus here we classify the deontic status of an action or event according to whether it is must or may be done or omitted. Furthermore, we limit ourselves to one special case among many: "be done" shall mean "be done at least once", and "omitted" shall mean "be omitted altogether".[4]

> $a$ is obligatory iff $a$ must be done,
>
> $a$ is permitted iff $a$ may be done,
>
> $a$ is forbidden iff $a$ must be omitted,
>
> $a$ is non-obligatory iff $a$ may be omitted.

The fourth notion, non-obligation, is not standard, and our terminology is not ideal. However, it is not easy to think of a term that is really apt.[5] In any case pre-theoretical intuitions dictate that no action is ever forbidden and permitted at the same time. By the same token, an obligatory action is always not non-obligatory. Hence if we were

[3]Thus in this modelling we are committed to the view that there are no "moral dilemmas". However irregular or illegal your past, there will always be a possible legal future. This commitment is of course not of a logical nature. It would be possible to modify our modelling so as to accommodate philosophers who believe in the existence of moral dilemmas.

[4]Actually also the given conditions are still quite general—further conditions can characterize different varieties of the general case. For example, if $a$ is permitted, will $a$ still be permitted if done once? If $a$ is forbidden, will $a$ still be forbidden if done once?

[5]There are also two other instances of linguistic awkwardness. In order to follow the pattern *permit* / *permitted* / *permission* we will usually choose *forbid* / *forbidden* / *forbiddance*, rather than, for example, *prohibit* / *prohibited* / *prohibition*. Furthermore, we accept *order* / *obligatory* / *obligation* rather than insisting that *order* be replaced by *obligate*.

to limit ourselves to so-called closed systems—systems in which every action is either permitted and forbidden, and in which actions that are not non-obligatory are obligatory—then we would have

*a* is permitted iff *a* is not forbidden,

*a* is non-obligatory iff *a* is not obligatory.

In other words, in closed systems we could begin with the two notions of obligation and forbiddance and define the other two. But in general we need all four.

In addition to these unconditional concepts there are numerous conditional ones: *a* is *obligatory* / *permitted* / *forbidden* / *non-obligatory* relative to certain condition.

## Model theory, syntax and meaning-conditions

A *norm* for a basic frame $(U, E, H)$ is a function $N$ defined on the set of situations which, for any situation $(h, S)$, where $S \subseteq \text{cont}(h)$, assigns a set $N(h, S)$ (intuitively, the set of normal futures, that is, normal from the point of view of this situation). There are four conditions on $N$:

(i) $N(h, S) \subseteq S$ (      );

(ii) if $S \subseteq T$ then $N(h, S) \neq \emptyset$ only if $N(h, T) \neq \emptyset$ (      );

(iii) if $S \subseteq T$ then $N(h, S) = S \cap N(h, T)$ (      );

(iv) if $g = pg'$, for a finite path $p$, then $g \in N(h, S)$ only if $g' \in N(hp, S')$, where $S' = \{f \in \text{fut}(p(\#)) : pf \in S\}$ (      ).

Notice that a norm-giver must be able to handle not only situations in which $S = \text{cont}(h(\#))$—in order to be complete, the norm must govern every imaginable situation. Note that in this modelling there are no degrees of non-normality.

The major new operators are four unary formula-forming term operators ob, pm, fb and no. Given a basic frame $(U, E, H)$ and a norm $N$, truth-conditions of formulæ can be given with respect to articulated histories:

$(h, g) \vDash \mathsf{ob}\,\alpha$ iff for all finite paths $p$ such that $h(\#) = p(*)$, if $p$ excludes $[\![\alpha]\!]$ then

- $pf$ includes $[\![\alpha]\!]$,
  for all $f \in N(hp, \text{cont}(hp))$,
- if $f' = qf''$, for any $q \in [\![\alpha]\!]$ and $f' \in \text{cont}(hp)$,
  then $k$ excludes $[\![\alpha]\!]$, for some $k \in N(hpq, \text{cont}(hpq))$.

$(h, g) \vDash \mathsf{pm}\,\alpha$ iff for all finite paths $p$ such that $h(\#) = p(*)$, if $p$ excludes $[\![\alpha]\!]$ then

- $pf$ includes $[\![\alpha]\!]$,
  for some $f \in N(hp, \text{cont}(hp))$,

- if $f' = qf''$, for any $q \in [\![\alpha]\!]$ and $f' \in \mathrm{cont}(hp)$, then $k$ excludes $[\![\alpha]\!]$, for some $k \in N(hpq, \mathrm{cont}(hpq))$.[6]

$(h, g) \vDash \mathsf{fb}\,\alpha$ iff for all finite paths $p$ such that $h(\#) = p(*)$, if $p$ excludes $[\![\alpha]\!]$ then

- $pf$ excludes $[\![\alpha]\!]$, for all $f \in N(hp, \mathrm{cont}(hp))$,
- if $f' = qf''$, for any $q \in [\![\alpha]\!]$ and $f' \in \mathrm{cont}(hp)$, then $k$ includes $[\![\alpha]\!]$, for some $k \in N(hpq, \mathrm{cont}(hpq))$.

$(h, g) \vDash \mathsf{no}\,\alpha$ iff for all finite paths $p$ such that $h(\#) = p(*)$, if $p$ excludes $[\![\alpha]\!]$ then

- $pf$ excludes $[\![\alpha]\!]$, for some $f \in N(hp, \mathrm{cont}(hp))$,
- if $f' = qf''$, for any $q \in [\![\alpha]\!]$ and $f' \in \mathrm{cont}(hp)$, then $k$ includes $[\![\alpha]\!]$, for some $k \in N(hpq, \mathrm{cont}(hpq))$.[7]

As explained in the following section, the definition of $\mathsf{ob}\,\alpha$ owes much to Ross [4]. The definitions of the other three operators have been designed to "harmonize" with that of $\mathsf{ob}\,\alpha$.

We omit meaning-conditions for the conditional deontic operators $\mathsf{ob}(\alpha/\phi)$, $\mathsf{pm}(\alpha/\phi)$, $\mathsf{fb}(\alpha/\phi)$ and $\mathsf{no}(\alpha/\phi)$.

## *Seinsollen* and *Tunsollen*

A question sometimes aired in the philosophical literature concerns the relative primacy of *Seinsollen* and *Tunsollen*. Three views are possible: (i) that *Seinsollen* is the basic concept and that *Tunsollen* can be defined in terms of it and non-deontic concepts; (ii) that *Tunsollen* is the basic concept and that *Seinsollen* can be defined in terms of it and non-deontic concepts; and (iii) that both concepts are basic and that neither is definable in terms of the other. In this note we are not taking a stand on this issue. For our (limited) purposes we find the *Tunsollen* approach congenial, but it would certainly be possible to introduce a deontic propositional operator in terms of which our deontic term operators would be definable.

Here is one way of doing it. Let $[\mathsf{D}]$—D for *deontic*—be a new unary proposition-forming propositional operator with the truth-condition

$$(h, g) \vDash [\mathsf{D}]\phi \text{ iff } (h, g') \vDash \phi, \text{ for all } g' \in N(h).$$

This operator may perhaps be read as "it is deontically necessary that" or "ideally". But it is too weak to be identified with "it ought to be the case" or "it is obligatory that".

[6] This defines a "weak" concept of permission. It must be possible to define "strong" concepts of permission as well.

[7] This defines a "weak" concept of non-obligation. Cf. the previous footnote!

With the help of this new operator our four deontic term operators can now be at least implicitly defined since the following schemata are logically valid:[8]

$$\mathsf{ob}\,\alpha \leftrightarrow [\mathsf{UNTIL}(\mathsf{occurred}\,\alpha)]$$
$$([\mathsf{D}]\langle\mathsf{F}\rangle(\mathsf{occurred}\,\alpha)\ \wedge\ [\mathsf{H}](\mathsf{occurred}\,\alpha \rightarrow [\mathsf{after}\,\alpha]\neg\mathsf{ob}\alpha),$$

$$\mathsf{pm}\,\alpha \leftrightarrow [\mathsf{UNTIL}(\mathsf{occurred}\,\alpha)]$$
$$(\langle\mathsf{D}\rangle\langle\mathsf{F}\rangle(\mathsf{occurred}\,\alpha)\ \wedge\ [\mathsf{H}](\mathsf{occurred}\,\alpha \rightarrow [\mathsf{after}\,\alpha]\neg\mathsf{pm}\alpha),$$

$$\mathsf{fb}\,\alpha \leftrightarrow [\mathsf{UNTIL}(\mathsf{occurred}\,\alpha)]$$
$$([\mathsf{D}]|\mathsf{F}]\neg(\mathsf{occurred}\,\alpha)\ \wedge\ [\mathsf{H}](\mathsf{occurred}\,\alpha \rightarrow [\mathsf{after}\,\alpha]\neg\mathsf{fb}\alpha),$$

$$\mathsf{no}\,\alpha \leftrightarrow [\mathsf{UNTIL}(\mathsf{occurred}\,\alpha)]$$
$$(\langle\mathsf{D}\rangle\langle\mathsf{F}\rangle(\neg\mathsf{occurred}\,\alpha)\ \wedge\ [\mathsf{H}](\mathsf{occurred}\,\alpha \rightarrow [\mathsf{after}\,\alpha]\neg\mathsf{no}\alpha).$$

Thus in this sense the deontic term operators are definable in terms of a deontic propositional operator and other non-deontic operators. On the other hand, given the deontic term operators as well as the resultative operator $\partial$ mentioned above, we would be able to define the usual deontic propositional operators. For example, obligation operators $\mathsf{O}$ and $\mathsf{O}'$ can be defined and seem natural under certain circumstances:

$$\mathsf{O}\phi \leftrightarrow \mathsf{ob}\ \partial\phi,$$

$$\mathsf{O}'\phi \leftrightarrow \mathsf{ob}\,\partial([\mathsf{F}]\phi).$$

In a similar way the other terms operators also give rise to propositional operators.

## Interludium: three so-called paradoxes

There is a family of conundrums in the literature on deontic logic, known as paradoxes. Most or all of them have been raised in order to make a certain point: that the paradox in question cannot be formalized within any of the then current systems of deontic logic. When new systems of deontic logic are presented it is therefore a good idea to see if and how they can handle these "paradoxes". The modelling presented in this section allows us to deal with some but not all of them. We give three examples.

***Chisholm's Paradox.*** This well-known paradox, which was first formulated by R. M. Chisholm in [2], turns on the difficulty of finding a model for four propositions of the following kind:

($C_1$) It ought to be the case that if X will do B then X does A (as soon as possible).

($C_2$) If X will not do B, then X should not do A (ever).

($C_3$) X ought to do B.

($C_4$) X will not do B.

[8] To repeat what has already been said: there are other ways of defining formal operators that may be claimed to correspond to pre-theoretical intuitive concepts.

If one tries to formalize these propositions in Standard Deontic Logic (neglecting the tense-logical aspect), the first two are naturally rendered on the format $\mathsf{O}(\phi \rightarrow \psi)$ and $\phi \rightarrow \mathsf{O}\psi$, respectively; and contradiction results.

In one familiar version of this example, B stands for X's going to see his grandmother, while A is notifying her in advance. The situation described is well-known: X will fail to do his duty. But with this understanding of the situation it is not clear that ($C_1$) and ($C_2$) are the only way to formalize the human predicament facing X. The propositions

($C_1^*$) It ought to be the case that if X will not do B then X does not do A.

($C_2^*$) If X will do B, then X should not do A.

would also be true of the hypothesized situation. Of course, Chisholm chose his formulations with an eye to bringing out the limitation of SDL. Our concern would perhaps disappear if we could find a new, binary connective $\mathsf{O}(\phi, \psi)$ (different from the ordinary unary $\mathsf{O}$, although we use the same letter for both operators), meaning something like "$\phi$ commits to $\psi$" or "$\phi$ makes obligatory that $\psi$". Then ($C_1$) and ($C_1$) could be rendered as $\mathsf{O}(\phi, \psi)$ and $\mathsf{O}(\neg\phi, \neg\psi)$, respectively. Moreover, if $\mathsf{O}(\phi, \psi) \rightarrow (\phi \rightarrow \mathsf{O}\psi)$ were generally valid, everybody could be happy. (There are such solutions in the literature.)

In our present setting our "solution" to Chisholm's Paradox has to be different, but in principle it follows the same line. If $\alpha$ and $\beta$ are two distinct event letters, our recommended translation is:

($C_1'$) $[\mathsf{H} : \langle\mathsf{F}\rangle\, \mathsf{occurs}\, \beta]\, \mathsf{ob}\, \alpha$.

($C_2'$) $[\mathsf{H} : \neg\langle\mathsf{F}\rangle\, \mathsf{occurs}\, \beta]\, \neg\mathsf{ob}\, \alpha$.

($C_3'$) $\mathsf{ob}\, \beta$.

($C_4'$) $\neg\langle\mathsf{F}\rangle\, \mathsf{occurs}\, \beta$.

The problem is then reduced to finding a model and an index $(h, g, H)$ at which all four formulæ are true. One would have to find sets

$$H_1 = \{g' \in \mathrm{cont}_H(h) : (h, g', H) \vDash \langle\mathsf{F}\rangle\, \mathsf{occurs}\, \beta\},$$

$$H_2 = \{g'' \in \mathrm{cont}_H(h) : (h, g'', H) \vDash \neg\langle\mathsf{F}\rangle\, \mathsf{occurs}\, \beta\}$$

such that

(i) $(h, H_1) \vDash \mathsf{ob}\, \alpha$,

(ii) $(h, H_2) \vDash \neg\mathsf{ob}\, \alpha$,

(iii) $(h, H) \vDash \mathsf{ob}\, \beta$,

(iv) $(h, g, H) \vDash \neg\langle\mathsf{F}\rangle\, \mathsf{occurs}\, \beta$.

But this task is easily solved. Hence our system may be said to pass the Chisholm test.

***The Ross Paradox.*** The Ross Paradox is Alf Ross's challenge in [4] to imperative logic provide a plausible formalization of the imperative "Post this letter!" that does not imply the imperative "Post this letter or burn it!" The parallel challenge to deontic logic is of course to provide a system in which "Posting the letter is obligatory" does not imply "Posting the letter or burning it, is obligatory". Standard Deontic Logic of course fails to meet Ross's challenge since it validates the schema

($R_1$) $\mathsf{O}\phi \to \mathsf{O}(\phi \vee \psi)$.

Ross's own advice was to distinguish between what he called the logic of validity and the logic of satisfaction. According to him there are two sides to the concept of obligation: it is one thing for an obligation to be in force (valid, in his terminology), another to be discharged (satisfied, in his terminology). We can rephrase his insight by saying that a (one-time) obligation remains in force as long as it has not been discharged. But once discharged, that particular obligation is no longer in force.

In the logic presented in this paper (which follows the analysis first given in [5]) Ross's example is formalized in a different way:

($R_2$) $\mathsf{ob}\,\alpha \to \mathsf{ob}\,(\alpha + \beta)$.

It is easy to see that this is not a valid schema. Hence our system passes the Ross test.

***Forrester's Paradox.*** Forrester's Paradox, first presented in [3], is the challenge to formalize sentences like

($F_1$) "Don't kill her! But if you do, do it gently!"

With respect to some model, let *a* be the event (action) of killing, and let *b* be any sub-event of *a*. In other words, $b \subseteq a$. We might then re-state the situation by saying that, while *a* is obligatory, given that *a* will not be done, *b* is obligatory. Calling up our focus operator, we might try the formula

$$\mathsf{fb}\,\alpha \wedge [\mathsf{H} : \langle \mathsf{F} \rangle\, \mathsf{oocurs}\,\alpha]\, \mathsf{ob}\,\beta,$$

or perhaps

$$\mathsf{fb}\,\alpha \wedge [\mathsf{H} : \langle \mathsf{F} \rangle\, \mathsf{oocurs}\,\alpha]\, (\mathsf{ob}\,\beta \wedge \mathsf{fb}\,(\alpha - \beta)),$$

where $a = [\![\alpha]\!]$ and $b = [\![\beta]\!]$. This formalization goes some of the way towards catching the structure of the further example

($F_2$) Don't kill her! But if you do, do it by giving her enough sleeping-pills (and not in any other way)!

at least if we consider that feeding someone sufficiently many sleeping-pills is a way of killing someone. But killing-gently is not in the same sense a sub-event of killing. Every element *p* of *a* is a particular realization of *a*. But, in a different sense of realization, *p* itself can be realized in different ways, depending on what *a* is—perhaps gently, perhaps quickly, perhaps carefully. And this concern performance, an aspect that the present formalism cannot do justice to. Thus Forrester's "paradox" marks one limitation of the present modelling.

# Step 3: Deontic logic with both real and deontic actions

## Pre-theoretical remarks

In addition to real actions, there are deontic actions. Corresponding to each of the basic deontic categories obligation, permission, forbiddance and non-necessitation there is a type of deontic action. The norm giver may order an action, making it obligatory. He may permit it, making it permitted. He may forbid it, making it forbidden. He may non-obligate it,[9] making it non-obligatory. How are we to represent those obviously crucially important actions?

One thing to keep in mind is the rôles played in our formal semantics by the primitive technical concepts $E$ and $H$. The former identifies the sets of finite paths that are recognizable as event types. The latter tells us which complete histories are really possible (where "really" means 'really'!).

The deontic actions we are primarily interested in in this note are ordering, permission, prohibition, and non-necessitation: for any action or event $a$,

- to order $a$: to make $a$ obligatory ("$a$ must be done!")
- to permit $a$: to make $a$ permissible ("$a$ may be done!")
- to prohibit $a$: to make $a$ forbidden ("$a$ must be omitted!")
- to declare $a$ non-obligatory: to make $a$ non-obligatory ("$a$ may be omitted!")

Here we shall be content to single out one specific explication for each of them. There are other possibilities that deserve to be considered, but the interest here is in the general problem of formalization rather than in a philosophical discussion of particular definitions of deontic conepts.

This said, here is our semi-formal understanding of the four actions that we want to formalize:

- ordering $a$: as long as $a$ has not been realized, every legal future includes $a$,
- permitting $a$: as long as $a$ has not been realized, some legal future includes $a$,
- forbidding $a$: as long as $a$ has not been realized, every legal future excludes $a$,
- non-obligating $a$: as long as $a$ has not been realized, some legal future excludes $a$.

## Model theory, syntax and meaning-conditions

Let $(U, E, H)$ be a basic frame. Norms are still functions $N$ defined on the set of situations $(h, S)$ such that $N(h, S)$ is a subset of $S$. As before, real actions are sets of finite sequences of points in $U$.[10] What is new now is that deontic actions are relations in the

[9] Yes, it is linguistically awkward!

[10] More generally, deontic actions of order $n$ may be viewed as relations in the set of $n$-order norms. However, this idea is not developped here.

set of norms. In fact, for simplicity we will assume that they are binary relations in the set of norms. In particular, our four special deontic actions are analysed as follows:

- The deontic action of ordering an event $a$, *ordering*($a$), is the set of all ordered couples $(N, N')$ such that, for all real situations $(h, S)$,

$$N'(S) = N\{f \in S : f \text{ includes } a\}.$$

- The deontic action of permitting an event $a$, *permitting*($a$), is some set of ordered couples $(N, N')$ such that, for all real situations $(h, S)$,

$$N'(S) = N(S) \cup N\{f \in S : f \text{ includes } a\}.$$

- The deontic action of prohibiting an event $a$, *forbidding*($a$), is the set of all ordered couples $(N, N')$ such that, for all real situations $(h, S)$,

$$N'(S) = N\{f \in S : f \text{ excludes } a\}.$$

- The deontic action of non-obligating an event $a$, *non-obligating*($a$), is some set of ordered couples $(N, N')$ such that, for all real situations $(h, S)$,

$$N'(S) \subseteq N(S) \cup N\{f \in S : f \text{ excludes } a\}.$$

Next we introduce four term-forming term operators !!, !, §§ and §. If $(U, E, H)$ is a basic frame, then the meaning-definition for terms is extended by the following clauses. where $\alpha$ is a term and $[\![\alpha]\!]$ is a real action or event:

$$[\![!!\alpha]\!] = \textit{ordering}([\![\alpha]\!]),$$

$$[\![!\alpha]\!] = \textit{permitting}([\![\alpha]\!]),$$

$$[\![\S\S\alpha]\!] = \textit{forbidding}([\![\alpha]\!]),$$

$$[\![\S\alpha]\!] = \textit{non-obligating}([\![\alpha]\!]).$$

Time for truth-conditions. The plot thickens! Unfortunately there is time only for some very brief remarks.

When agents are capable of deontic actions, the notions of events and histories much be generalized. Therefore frames will have to be more complicated than before. Let us start with a basic frame $(U, E, H)$. Now that we have deontic actions as well as real actions, we need a new category $D$ of deontic actions. Just as a real action or event is a set of finite paths in $U$, so a deontic action ought to be a set of finite paths in $M$, where $M$ is a motley of norms (in the sense of norm as defined above).

Furthermore, real histories are sequences of points in $U$. But now we need a more inclusive category! Let us use the word *chronicle* for sequences of pairs $(h, N)$, where $h$ is a past history and $N$ is a norm, and let us write $K$ for the set of all chronicles. The notions of *maximal chronicle*, *articulated chronicle*, *past chronicle* and *future chronicle* can be defined in analogy with the corresponding historical concepts. If $c$ is a past chronicle. the *real past history* of $c$ can be retrieved; and if $c(\#) = (u, N)$ is the last element of $c$ then we call $N$ the *legal* or *normative position after h*. It is clear that corresponding to each chronicle in $K$ there is a unique history in $H$.

Hence our new frames become ordered tuples $(U, E, H, M, D, K)$, with various conditions regulating the primitives. Given a model on such a frame, truth-conditions can be given with respect to articulated chronicles $(c, d)$. All our old truth-conditions have to be generalized, but we will spare readers the details except for the conditions pertaining to our four favourite deontic actions:

$(c, d) \vDash [\mathsf{after}\, !!\alpha]\theta$ iff there are $h, N, N'$ such that $c = (h, N)$ and $d = (h, N')$ and $(N, N') \in [\![ !!\alpha ]\!]$,

$(c, d) \vDash [\mathsf{after}\, !\alpha]\theta$ iff there are $h, N, N'$ such that $c = (h, N)$ and $d = (h, N')$ and $(N, N') \in [\![ !\alpha ]\!]$,

$(c, d) \vDash [\mathsf{after}\, \S\S\alpha]\theta$ iff there are $h, N, N'$ such that $c = (h, N)$ and $d = (h, N')$ and $(N, N') \in [\![ \S\S\alpha ]\!]$,

$(c, d) \vDash [\mathsf{after}\, \S\alpha]\theta$ iff there are $h, N, N'$ such that $c = (h, N)$ and $d = (h, N')$ and $(N, N') \in [\![ \S\alpha ]\!]$.

## Ability and competence

In order to drive a car you need to know how to manœuvre it (real ability). But in order to do it legally, you also need a driver's licence (legal competence). The ability you acquire by learning. The competence can be bestowed upon you by the Department of Motor Vehicles, which under certain circumstances will issue a licence to you (a deontic action). That authority has itself been established by another higher-order authority (a higher-order deontic action). Which in turn draws its authority from somewhere (from some even higher-order deontic action). And so on. (In human affairs, this kind of regression of authority is always finite.)

Let us see how ability and competence can be analysed within the formalism developped here. Suppose $(U, E, H, N, D, K)$ is a new frame. Let $S_i$ be a function on the set of possible pasts $h$ such that always $S_i(h) \subseteq \mathrm{cont}_H(h)$. The informal intuition is that $S_i(h)$ is the set of possible futures after $h$ that the agent $i$ controls in the sense that by his action he can make sure that the actual future will turn out to be one of the elements of $S_i(h)$. There are two obvious concepts of ability:

$i$ is *weakly able* after $h$ to realize $a$ iff some $f \in S_i(h)$ includes $a$,

$i$ is *strongly able* after $h$ to realize $a$ iff every $f \in S_i(h)$ includes $a$.[11]

Assume that $S_i^*$ is a function on the set of possible past chronicles $c$ such that always $S_i^*(c) \subseteq \mathrm{cont}_C(c)$. Then in analogy with the definitions of two concepts of ability there are the two definitions that follows. Let $(c, d)$ be an articulated chronicle in $K$ and let $a$ be a deontic action in $D$.

[11]There are related concepts that can be defined in a similar vein, such as the opportunity to realize an event at once.

$i$ is *weakly competent* after $c$ to realize $a$ iff some $e \in S_i^*(c)$ includes $a$,

$i$ is *strongly able* after $c$ to realize $a$ iff every $e \in S_i^*(c)$ includes $a$.

When $a$ is one of our four simple deontic actions, weak and strong competence coincide.

What is important in definitions of ability is the interplay between the primitives $E$ and $H$: both are needed. For definitions of competence, $D$ and $C$ play similar rôles.

## And then . . . ?

Let us call the basic frames $(U, E, H)$ that we defined in step 1 *zero-order frames*. They can be written on the form $(U, E_0, H_0, M_0)$, where $M_0 = \varnothing$ is the set of zero-order norms (there aren't any!). In step 2 we considered zero-order frames with a norm $N$ that regulated the real actions in $E_0$. In step 3 we met with frames that may be called *first-order frames*: frames of the form $(U, E_0, H_0, M_0, E_1, H_1)$, where $E_1$ is a set of deontic actions (previously known as $D$) and $H_1$ is a set of chronicles ($K$). Then why not next add a norm $N'$ regulating the deontic actions in $E_1$? And then . . . ? This is obviously not the end, but rather the beginning of a long regress.

In principle we could define sets $E_n$ of actions of order $n$ (the actions of order 0 being the real ones); sets $H_n$ of histories of order $n$, and sets $M_n$ of norms of order $n$ (regulating actions of order $n - 1$, if $n > 0$). This would give us frames of the form $(U, (E_n, H_n, M_n)_{n=0}^m)$, where $m$ should be a natural number (or possibly $\omega$).

Even more general frames would be of the form $(U, ((E_i, H_i, M_i)_{i \in I}))$ where $I$ is an index set with some structure to it. Certain families of frames of this kind might be of interest in connexion with the study of hierarchies.

## References


[1] A ´ , C , P G¨ & D M . "On the logic of theory change." *The journal of symbolic logic*, vol. 50 (1985), pp. 510-530.

[2] C , R M. "Contrary-to-duty imperatives and deontic logic." *Analysis*, vol. 24 (1963), pp. 33-36.

[3] F , J . "Gentle murder, or the adverbial Samaritan." *Journal of philosophy*, vol. 81 (1984), pp. 193-197.

[4] R , A . "Imperatives and logic." *Theoria*, vol. 7 (1941), pp. 53-71.

[5] S , K . "Trying to meet Ross's challenge." In *Logic and philosophy in Italy: some trends and perspectives. Essays in honor of Corrado Mangione on his 75th birthday*, edited by Edoardo Ballo & Miriam Franchella, pp. 155-166. Monza, Italy: Polimetrica International Scientific Publisher, 2006.