Reactive Standard Deontic Logic

Dov Gabbay[†] and Christian Straßer[‡]
[†]Bar-Ilan University
[†]King's College London,
[†]University of Luxembourg
[‡]Ghent University

Abstract

We introduce a reactive variant of **SDL** (standard deontic logic): **SDLR1** (reactive standard deontic logic). Given a Kripkean view on the semantics of **SDL** in terms of directed graphs where arrows $\rightarrow$ represent the accessibility relation between worlds, reactive models add two elements: arrows $\rightarrow$ are labelled as "active" or "inactive", and double arrows $\twoheadrightarrow$ connect arrows, e.g., $(x_1 \rightarrow x_2) \twoheadrightarrow (x_3 \rightarrow x_4)$. The idea is that passing through $x_1 \rightarrow x_2$ activates a switch represented by $\twoheadrightarrow$ which inverts the label of $x_3 \rightarrow x_4$ and hence activates resp. deactivates this arrow. This allows to introduce two modalities: $\square$ is the usual **KD**-modality of **SDL** and operates on the Kripkean graph where all labels and double arrows are ignored, while $\oslash$ takes them into account.

We demonstrate that **RSDL1** allows for an intuitive interpretation of 'ought'. The logic can handle contrary-to-duty cases such as several instantiations of the Chisholm set in a paradox-free way by means of using double arrows and annotations to block and give access to ideal worlds.

1 Standard Deontic Logic and its Problems

A logic with modality $\square$ is **KD** modality if we have the axioms

K0   All substitution instances of classical tautologies

K1 $\square(p \wedge q) \equiv (\square p \wedge \square q)$

K2 $\vdash A \Rightarrow \vdash \square A$

D $\neg \square \bot$

It is complete for frames of the form (S, R, a) where $S \neq \oslash$ is a set of possible worlds, $a \in S, R \subseteq S \times S$ is serial (i.e., for all x there exists a y such that xRy).

Standard Deontic Logic **SDL** is a **KD** modality O. We read $u \vDash Op$ as saying p holds in all ideal worlds relative to u, i.e. for all t we have: if uRt then $t \vDash p$. So the set of ideal worlds relative to u is the set

$$I(u) = \{t \mid uRt\}.$$

The **D** condition says

$$I(x) \neq \emptyset \text{ for } x \in S.$$

Following [18], let us quickly review some of the difficulties facing **SDL** in formalizing certain examples. These paradoxes are as follows, see [4, 23, 22, 1].

The Chisholm paradox

Consider the following statements:

1. It ought to be that a certain man go to the assistance of his neighbour.

2. It ought to be that if he does go he tell them he is coming.

3. If he does not go then he ought not to tell them he is coming.

4. He does not go.

It is agreed that intuitively (1)–(4) of Chisholm are consistent and totally independent of each other. Therefore it is expected that their formal translation into logic **SDL** should retain these properties.

Let us semantically write the Chisholm set in semiformal English, with p and q as follows, p means HELP and q means TELL.

1. Obligatory p.

2. $p \rightarrow$ Obligatory q.

3. $\neg p \rightarrow$ Obligatory $\neg q$.

4. $\neg p$.

Consider also the following:

5. p.

6. Obligatory q.

7. Obligatory $\neg q$.

We intuitively accept that (1)–(4) are consistent and logically independent of each-other. Also we accept that (3) and (4) imply (7), and that (2) and (5) imply (6). Note that some authors would also intuitively expect to conclude (6) from (1) and (2).[1]

Now suppose we offer a logical system **L** and a translation τ of (1), (2), (3), (4) of the Chisholm set into **L**.

For example **L** could be Standard Deontic Logic or **L** could be a modal logic with a dyadic modality $\bigcirc_Y X$ (X is obligatory in the context of Y).

We now list coherence conditions for the translation τ and for **L**.

We expect the following to hold.

(**a**)   (2) and (3) are translated the same way, i.e., we translate the form:
(23): $X \rightarrow$ Obligatory Y
to be $\psi(X, Y)$ and the translation does not depend on the fact that we have (4)
$\neg p$ as opposed to (5) p.

---

[1]The principle that allows for such inferences is called *deontic detachment* (see [19]). See also [24] for a discussion of its role in the context of the Chisholm set.

(**b**)   The translations of (1)–(4) remain logically independent in **L**.

(**c**)   The translated system maintains its properties under reasonable substitution in **L**.

   The notion of reasonable substitution is a tricky one. Let us say for the time being that if we offer a solution for one paradox, say $\Pi_1(p, q, r, \ldots)$ and by substitution for $p, q, r, \ldots$ we can get another well known paradox Π_2, then we would like to have a solution for Π_2. This is a reasonable expectation from mathematical reasoning. We give a general solution to a general problem which yields specific solutions to specific problems which can be obtained from the general problem.

(**d**)   The translation is essentially linguistically uniform and can be done item by item in a uniform way depending on parameters derived from the entire database. To explain what we mean consider in classical logic the set

   (1) p

   (2) $p \rightarrow q$.

   To translate it into disjunctive normal form we need to know the number of atoms to be used. Item (1) is already in normal form in the language of $\{p\}$ but in the language of $\{p, q\}$ its normal form is $(p \wedge q) \vee (p \wedge \neg q)$. If we had another item

   (3) r

   then the normal form of p in the language of $\{p, q, r\}$ would be $(p \wedge q \wedge r) \vee (p \wedge q \wedge \neg r) \vee (p \wedge \neg q \wedge r) \vee (p \wedge \neg q \wedge \neg r)$.

   The moral of the story is that although the translation of (1) is uniform algorithmically, we need to know what other items are in the database to set some parameters for the algorithm.

   Jones and Pörn, for example, examine in [18] possible translations of the Chisholm (1)–(4) into **SDL**. They make the following points:

   (1) If we translate according to, what they call, option a:

   (1a) Op

   (2a) $O(p \rightarrow q)$

   (3a) $\neg p \rightarrow O \neg q$

   (4a) $\neg p$

   then we do not have consistency, although we do have independence.[2]

   (2) If we translate the Chisholm item (2) according to what they call option b:

   (2b) $p \rightarrow Oq$

   then we have consistency but not independence, since (4a) implies logically (2b).

   (3) If (3a) is replaced by

---

[2]Note that this translation is also not uniform in its treatment of conditional obligations (see (2a) and (3a)) and hence violates our coherence condition (a).

(3b) $O(\neg p \rightarrow \neg q)$

then we get back consistency but lose independence, since (1a) implies (3b).

(4) Further, if we want (2) and (5) to imply (6), and (3) and (4) to imply (7) then we cannot use (3b) and (2a).

The translation of the Chisholm set is a "paradox" because known translations into Standard Deontic Logic (the logic with O only) are either inconsistent or dependent.

The fence paradox

The following example is formulated in [21]. Compare with items (1)–(4) of the Chisholm paradox.

1. There ought to be no fence.

2. [We are not dealing with this item here.]

3. If there is a fence it should be white.

4. There is a fence.

The fence paradox is a variation of the original gentle murderer paradox of Forrester [6].

1. It is obligatory that Smith not murder Jones.

2. [we are not dealing with this item here.]

3. It is obligatory that, if Smith murders Jones, Smith murders Jones gently.

4. Smith murders Jones.

Of course the fence can be taken down, but the murder cannot be undone. This difference can be significant when we consider our options in a long chain of violations and contrary to duties.

The problems with **SDL** are so serious that nowadays researchers in the field call it 'silly deontic logic' instead of 'standard deontic logic'.

There have been many proposals for better systems to accommodate Contrary to Duties ranging from slight variations of **SDL** (see, for example, [18] and [5]) to considerably different systems (see, for example, [25, 21]), mainly the dyadic preferential approach and the temporal approach. See also [3, 15]. Our own reactive approach is new, see [8, 10, 11], and bridges the gap between the two extremes, as this paper will show.

The community seems quite happy with a dyadic modal operator $\bigcirc_A B$, reading B is obligatory in the context of A. Thus we have for the fence example

1. $\bigcirc \neg$ fence

2. $\bigcirc_{\text{fence}}$ white-fence.
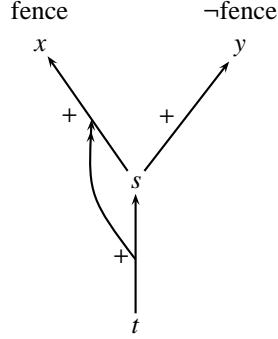
fence   ¬fence

Figure 1: A reactive graph

The dyadic deontic logic satisfies some axioms which can basically turn it into a preferential system. We need not go into details here. See [20, 25, 21].

Our aim is to show that **SDL** is not so silly, by turning it *reactive* and by looking at reactive standard deontic logic, **SDLR1**, we can save the situation. See [10, 11] for a more general reactive approach.

Consider the situation in Figure 1. Its graph describes a possible world model, of a reactive kind, for a special reactive modality $\oslash$.

The nodes $\{t, s, x, y\}$ are possible worlds. The syntactical word 'fence' represents an atomic proposition and by writing 'fence' next to x we inform the reader that in some given assignment h (fixed for our discussion) we have $x \vDash$ fence under h. Similarly we have $y \vDash \neg$ fence under h. The arrows $s \to y, t \to s$ and $s \to x$ are annotated with either '+' or '−'. The '+' annotation means the connection is active (or connected) and the '−' annotation means the connection is not active or not connected. The double arrow from $t \to s$ to $s \to x$ is a switch. It says that if we pass from t to s along the arrow $t \to s$, then the double arrow gets activated and disconnects the arrow from s to x, if its active, and connects it, if it is not active.

Suppose we want to evaluate $t \vDash \oslash \oslash \neg$ fence.

We have $t \vDash \oslash \oslash \neg$ fence iff (since $t \to s$ is active) $s \vDash \oslash \neg$ fence. Now since we passed through $t \to s$ to get to s, the connection $s \to x$ is switched off and we continue (since $s \to y$ is active):

$s \vDash \oslash \neg$ fence iff $y \vDash \neg$ fence

which indeed holds.

Thus, in the model of Figure 1 we indeed have that $t \vDash \oslash \oslash \neg$ fence.

We can add a connective $\square$ which ignores all the double arrows and all annotations. Thus we have

$t \vDash \square\square\neg$fence iff $s \vDash \square\neg$fence iff ($x \vDash \neg$fence and $y \vDash \neg$fence)

Since in Figure 1 $x \vDash$ fence we get that $t \nvDash \square\square\neg$fence.

Our logic **SDLR1** is the logic with $\square$ and $\oslash$. We introduce it formally in the next section. Then, in sections 3– 5 we shall translate the deontic paradoxes (such as clauses (1)–(4) of the Chisholm set and (HC1) and (HC2) of Figure 3 below) into it.

2 Formal properties of SDLR1

This section develops the formal technical machinery for our paper. We give formal definitions of the reactive models we use. We also investigate an axiomatisation of our logic relative to the proposed semantics.

2.1 Introducing and discussing the semantics

Definition 2.1 (Reactive **K** model)**.**

1. *Let S be a non-empty set. A reactive relation on S has the form*

$$R = R_1 \cup R_2$$

where $R_1 \subseteq S^2$ and $R_2 \subseteq S^2 \times S^2$ such that the following holds

$$\text{If } ((x, y), (w, z)) \in R_2, \text{ then } (x, y) \in R_1 \text{ and } (w, z) \in R_1.$$

(This means $R_2 \subseteq R_1 \times R_1$).

We can represent the elements of R_1 as '$x \to y$' and the elements of R_2 as '$(x \to y) \twoheadrightarrow (w \to z)$'. The elements of R_1 are called arrows and those of R_2 are called double arrows.

2. *A function $\mathbf{f} : R_1 \to \{0, 1\}$ is called an activity function. When $\mathbf{f}(x, y) = 1$ we say $x \to y$ is 'on' or is 'active' and we write $x \xrightarrow{+} y$. When $\mathbf{f}(x, y) = 0$, we say $x \to y$ is 'off' or 'not active' and we write $x \xrightarrow{-} y$.*

3. *An assignment h is a function giving each atomic q a subset $h(q) \subseteq S$.*

4. *A model has the form $\mathbf{m} = (S, R_1, R_2, \mathbf{f}, a, h)$ where $a \in S$ is the initial world.*

Definition 2.2 (Satisfaction)**.** *Our language contains the classical connectives $\neg, \wedge, \vee, \to$, $\top, \bot$ and the two modalities $\square$ and $\oslash$. We define satisfaction in a model as follows.*

1. *Let $(S, R_1, R_2, \mathbf{f})$ be given. Where $(x, y) \in R_1$ and $x \xrightarrow{+} y$, we define a new function, called $\mathbf{f}[x, y]$, which is derived using $\mathbf{f}$ and (x, y), and is defined by indicating what values this function gives to arbitrary pairs (u, v) in R_1 as follows:*

$$\mathbf{f}[x, y](u, v) = \begin{cases} \mathbf{f}(u, v) \text{ if } ((x, y), (u, v)) \notin R_2 \\ 1 - \mathbf{f}(u, v)) \text{ if } ((x, y), (u, v)) \in R_2 \end{cases}$$

2. *Given a model $\mathbf{m} = (S, R_1, R_2, \mathbf{f}, a, h)$ we define the notion of $\mathbf{m} \vDash A$ by induction as follows:*

   - *$\mathbf{m} \vDash q$ iff $a \in h(q)$, for q atomic*
   - *$\mathbf{m} \vDash A \wedge B$ iff $\mathbf{m} \vDash A$ and $\mathbf{m} \vDash B$*
   - *$\mathbf{m} \vDash \neg A$ iff $\mathbf{m} \nvDash A$*
   - *Similarly for the other classical connectives*
   - *$\mathbf{m} \vDash \square A$ iff for all y such that $(a, y) \in R_1$ we have that $\mathbf{n} \vDash A$, for $\mathbf{n} = (S, R_1, R_2, \mathbf{f}, y, h)$*
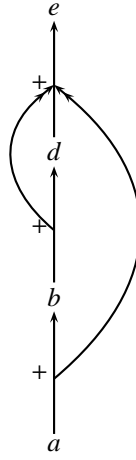
6

Figure 2: A model for **KR**

- **m** $\vDash \oslash A$ *iff for all y such that* $(a, y) \in R_1$ *and* $\mathbf{f}(a, y) = 1$ *we have that* $\mathbf{n}' \vDash A$ *for* $\mathbf{n}' = (S, R_1, R_2, \mathbf{f}[a, y], y, h)$.

3. *The logic* **KR** *(reactive* **K***) is defined as the set of all wffs A such that for all models* **m** *we have* **m** $\vDash A$.

4. *The logic* **SDLR1** *is defined analogously just that* R_1 *is required to be serial.*

5. *Note that we have frames of the form* $(S, R_1, R_2, \mathbf{f})$. *The logic is normal, as a and h can be chosen arbitrarily.*

Remark 2.3. Note that the function **f** introduced in the previous Definition 2.1 actually defines another relation R_0 which is a subset of R_1: namely the active arrows in R_1. When **f** changes to $\mathbf{f}[x, y]$, as in item 1 of Definition 2.2, then equivalently R_0 changes to $R_0[x, y]$. We use the notation with the function **f** to stress the fact that all reactivity changes arise from arrow connections from R_1.

We now address the task of axiomatising the logic **KR**. The axiomatisation of **SDLR1** strengthens **KR** by **D**. To be able to do that effectively, let us familiarise ourselves better with **KR** and its unique features.

Consider Figure 2.

Let our starting point be a. As we traverse the arc $a \to b$, there are two possibilities.

1. We ignore all double arrows and what they do.

2. We take account of what the double arrows say and what they do.

Let possibility 1 be formalised by $\lozenge$ and let possibility 2 be formalised by $\mathbb{P}$. So we can write: $a \vDash \lozenge q$ to mean traverse from a to the accessible world b and ignore all double arrows and $a \vDash \mathbb{P}q$ to mean traverse from a to the accessible world b and take account of the double arrows.

Consider now

$$a \vDash \mathbb{P}\lozenge\lozenge\top$$

Here we move from a to b and activate the double arrow which disconnects the arc $d \to e$. We check:

$$b \vDash \lozenge\lozenge\top$$

7

and then
$$d \vDash \Diamond\top$$

Now the arc $d \to e$ is disconnected by a double arrow but as we are evaluating $\Diamond$ at node d and $\Diamond$ ignores all double arrows and their effects, as far as $\Diamond$ is concerned $d \to e$ is connected and hence $d \vDash \Diamond\top$ holds and therefore $a \vDash \mathbb{P}\Diamond\Diamond\top$ also holds.

The situation is different with

$$a \vDash \mathbb{P}\Diamond\mathbb{P}\top$$

Evaluating here takes us to $d \vDash \mathbb{P}\top$, but $\mathbb{P}$ is affected and takes account of double arrows and so $d \vDash \mathbb{P}\top$ does not hold and hence

$$a \vDash \mathbb{P}\Diamond\mathbb{P}\top$$

does not hold.

Consider now

$$a \vDash \mathbb{P}\mathbb{P}\mathbb{P}\top.$$

We can go from a to e taking account of double arrows. Moving from a to b switches the connection $d \to e$ off but continuing from b to d switches it on again and therefore $d \vDash \mathbb{P}\top$ holds.

Note that we cannot evaluate any $x \vDash \mathbb{P}q$ without knowing how we got to x and which double arrows we activated along the path.

So asking does $d \vDash ?\mathbb{P}\top$ hold cannot guarantee an answer. We need to ask something like

$$a \xrightarrow{\Diamond} b \xrightarrow{\mathbb{P}} d \vDash ?\mathbb{P}\top$$

or

$$a \xrightarrow{\mathbb{P}} b \xrightarrow{\Diamond} d \vDash ?\mathbb{P}\top$$

or in general

$$(x_0 \xrightarrow{M_1} x_1 \xrightarrow{M_2} x_2 \to \ldots \xrightarrow{M_n} x_n) \vDash ?A$$

where $M_i \in \{\Diamond, \mathbb{P}\}$.

Note that we need a notion of *legitimate sequences*. The sequence

$$a \xrightarrow{\mathbb{P}} b \xrightarrow{\Diamond} d \xrightarrow{\mathbb{P}} e$$

is not legitimate because by the time we get to d, $d \to e$ is disconnected. So to get to e from d we need to use $\Diamond$.

We thus get a bimodal logic with two modalities $\Diamond$ and $\mathbb{P}$ with evaluation of formulas depending on paths $(x_0, \ldots, x_n)$ but not ordinary paths. We also need to know how we traverse them. So our paths have the form $\mathbf{t} = (x_0, M_1x_1, M_2x_2, \ldots, M_nx_n)$, where $\mathbf{t}$ is a legitimate sequence.

Our task in this section is to find axioms and prove completeness for **KR**. We might have thought that it would be useful to consider **KR** as a bimodal logic with two accessibility relations $R_{\mathbb{P}}$ and $R_{\Diamond}$, and see what properties they have. The problem is that the accessible worlds are the same for both relations, the difference being the way we traverse them. So really our most reasonable option is to consider a transformed Kripke model where the points/worlds are legitimate paths of the form

$$\mathbf{t} = (a, M_1x_1, \ldots, M_nx_n)$$

such that $aR_1x_1R_1x_2, \ldots, x_{n-1}R_1x_n$ holds.

Thus we start with

$$\mathbf{m} = (S, R_1, R_2, a, h)$$

for example (S, R_1, R_2, a) is as in Figure 2, and then transform it to a model

$$\mathbf{m}^* = (S^*, R_\diamond, R_\mathbb{P}, (a), h)$$

where S^* is the set of all legitimate sequences of the form $\mathbf{t} = (a, M_1x_1, \ldots, M_nx_n)$ where $aR_1x_1 \wedge x_1R_1x_2 \wedge, \ldots, \wedge x_{n-1}R_1x_n$.

We define $R_\diamond$ and $R_\mathbb{P}$ by

$$\mathbf{t}R_\diamond\mathbf{t} * (\diamond y), \text{ when } x_nR_1y,$$
$$\mathbf{t}R_\mathbb{P}\mathbf{t} * (\mathbb{P}y), \text{ when } x_nR_1y$$

$*$ is concatenation of sequences.

Let $\mathbf{t} \vDash q$ iff $x_n \vDash q$, for q atomic.

Note that R_1 and R_2 are replaced by $R_\diamond$ and $R_\mathbb{P}$. R_1 is used directly in the definition of the sequences $\mathbf{t}$ which go into S^* and R_2 is used in the notion of legitimate sequences.

So $\mathbf{m}$ is replaced by $\mathbf{m}^*$.

If we do that we can get an axiomatisation and a completeness theorem after some hard work (see [12]).

3   Contrary to duty in SDLR1

This section presents our reactive model. We do it in stages, presenting several intermediate models and modifying them.

We use a story (Example 3.1) to motivate our model. We offer the model of Figure 4 below as a first approximation, and then we move to the better model of Figure 7 and then, for technical reasons, we adopt the slightly better pre-final model of Figure 9. This is not the final model, but we stick with it for a while.

We then express the fence example in this model, and discuss it at length.

The final model, which is a slight modification of the pre-final model, is given formally in Section 5.

We want to alert the reader to some disadvantages of our model, and urge her or him to nevertheless continue to read.

The translation turns out to be syntactically complicated, even though the idea is semantically very simple. This can be a disappointment for the syntactially/Hilbert system minded reader, but there are two reasons for it:

1. The subtle complexity of the CTD paradox itself. The obligation not to have a fence is in one context while the CTD kicks in in another context. This subtle complexity is unavoidable.

2. The second reason for the complexity is the fact that semantically we are relying on the reactivity of the double arrows to solve the paradoxes while the syntax does not give us connectives which talk directly about the double arrows. This makes the model more complex, both in translating the CTD expressions and, as already mentioned, in axiomatising the logic.

The perceptive reader is sure to detect another problem. The translation of the Deontic modality "Obligatory" is not a pure connective but is done as a combination of the two modalities □ and ⊘ of the system. This is necessitated by the lack of direct connective in the logic representing the double arrows, despite the fact that semantically the obligation is expressed by the double arrows!

We are thus forced to go in a roundabout way to express the obligations and thus make a mixed use of the modalities of the system. The weakness of the translation is compounded by the need to make it dependent on the structure of the model.

So why consider our system at all?

The answer is that it is semantically very intuitive and has the advantages of offering a systematic solution to the so called paradoxes and can do multiple CTD sentences all at once. I put forward to the reader that the call for axiomatisation (more specifically, a Hilbert style axiom system) is partly a cultural remnant from the time when modal logic had no satisfactory semantics and when most logics were presented Hilbert style. See however paper [14], where a long chain of Chisholm type sets is considered and modelled in a Hilbert axiom system with a reactive semantical interpretation.

Furthermore, the above difficulties would disappear once we allow a direct syntactic counterpart to our double arrows in the axiomatic system. We can in fact give a really simple axiomatisation as a Labelled Deductive System, see [11] and [7].

We shall address these problems in the concluding section.

We now motivate, through examples, how we treat CTDs in **SDLR1**.

Example 3.1 (UK gas boiler). Many homes in the UK have heating with gas boilers. The technology of such boilers keeps on improving both in efficiency and safety. British Gas inspects each boiler once a year and makes recommendations. In many cases a slightly older model boiler no longer complies with the new safety regulations, but of course one cannot ask for the boiler to be replaced, just on the strength of that, so some additional safety measures are recommended. However, if one wants to install a new boiler then only the new models can be installed.

Consider now a family who has a very old boiler. This boiler breaks down. It is dead and finished. They want to get a new boiler. The law says only the new models are allowed to be installed. However, the slightly older models are still good, they cost a lot less and still have some good years of service in them. They are perfectly safe with the new additional safety measures added to them. The only problem is that the law does not allow certified engineers to install them. They can only modify older existing boilers but not install them!

Our family a wants to buy a boiler. They go to an agent e, who offers them options. They can get an old model or a ¬old model (= new model). The old model cannot be installed by a certified engineer but unofficially it can be installed and then the family can pretend that this is an existing old model and ask a certified engineer to add the safety modifications.

The clever agent e solves the problem of how to work around the regulations as follows: the sale of new models he assigns to a special department e_2 and the sale of old models he assigns to a separate department e_1. If e_1 is investigated and is caught trying to go around the rules, then that is a local problem. So in practice the agent will sell both boilers himself through different departments. The new boilers department and the old boilers department. Figure 3 shows the analogy between the fence example and the boiler example.

The only difference between the two examples is the incentive to the agent e to work through specialised departments e_1 and e_2. There is no such incentive in fences.

10

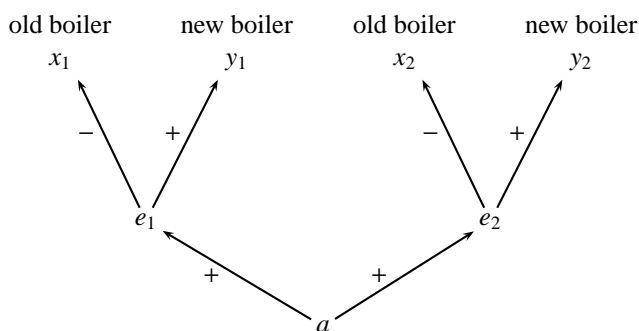|	Fence	Boiler
HC1.	There should be no fence	There should be no old boiler (installed)
HC2.	If there is a fence it should be white	If there is an existing old boiler it should be modified with additional safety measures.

Figure 3: Fence and boiler problems compared

Figure 4: Agents showing only new boilers

So let us go back to our family who wants to buy a boiler.

They approach agents $e_1, \ldots, e_m$ and each one of these agents offers a list of possible boilers for sale. Let $x^i_1, \ldots, x^i_{m(i)}, y^i_1, \ldots, y^i_{n(i)}$ be the list of boilers offered by e_i. Assume that x^i_j are old boilers, and y^i_j are not old boilers.

The family may observe the rule (HC1). Therefore they may consider only the new boilers. In fact, the agent will probably show them only the new boilers. This means that the story is represented as in Figure 4. This figure shows only two agents e_1, e_2, each showing one old boiler and one new boiler. They may be showing the same boilers in which case we have Figure 6a. The node a represents the family. The arrow from a to e_1 symbolises the family a going to e_1. (Note that according to Figure 4, the family goes directly to agents e_1, e_2. This scenario shall be modified in Figure 7, in which the family goes to agent e and agent e refers them to departments e_i.) The nodes x_i, y_i denote boilers for sale (x_i old boilers and y_i new boilers).

The '−' label on $e_1 \to x_1$ switches the link from e_1 to x_1 off. This blockage of access to the old boiler is in accordance with (HC1). Similarly, the arrow $e_2 \to x_2$ is labeled by '−' and hence switched off. Both dealers only present new boilers. Although they have the old boilers (x_1, x_2), they do not show them. Now if our family wants to be presented old boilers they need to explicitly ask the dealer for old boilers. This is realized by means of double arrows and illustrated in Figure 5.

Hence, if our family has the wish to see the old boilers despite (HC1) they can force their way to x_1 and x_2 by changing the '−' label of the arcs via double arrows. They pass via $\mathbb{P}$ from a to the dealer e_i. This activates the double arrows $(a \to e_i) \twoheadrightarrow (e_i \to x_i)$ and $(a \to e_i) \twoheadrightarrow (e_i \to y_i)$. The double arrows invert the labels or the arcs $e_i \to x_i$ and $e_i \to y_i$ resulting in the '+'-labeling of $e_i \to x_i$ and the '−'-labeling of $e_i \to y_i$. Hence, when our family arrives at e_i they see Figure 6b.
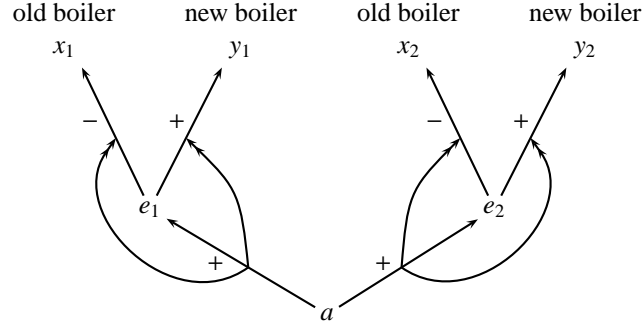
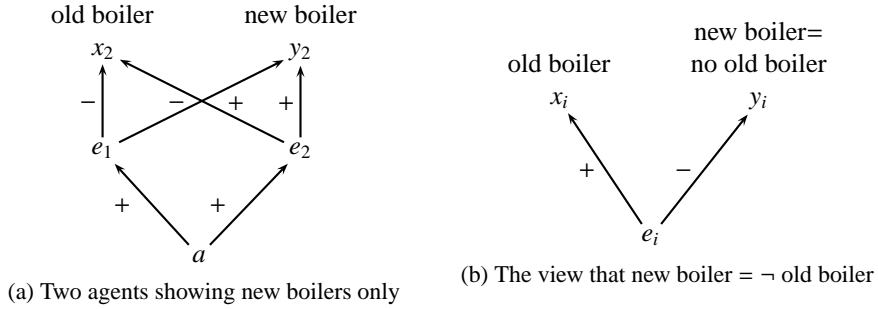Figure 5: The family asks the dealers to show only old boilers.

(a) Two agents showing new boilers only

(b) The view that new boiler = $\neg$ old boiler

Figure 6

We now consider Figure 5 as a reactive Kripke model.

By evaluating $a \vDash \Box \oslash \neg$ old boiler in Figure 5 we are saying that at all agents $e_i, \oslash \neg$ old boiler holds. Hence, (HC1) holds and the agents do not offer any old boiler. This concerns the case when the double arrows do not fire since we move from a to e_i via $\Diamond$. We take the wish of our family to see only the old boilers into account when we move from a to e_i via $\mathbb{P}$. In this case the double arrows fire and the dealers show only the old models. This is expressed by $a \vDash \oslash\oslash$ old boiler.

Thus we may translate

(HC1):   There should be no old boiler

as

(HC1)*: $\Box \oslash \neg$ old boiler.

The fact 'old boiler', which we interpret as what the family wants to buy, may be expressed by $\oslash\oslash$ old boiler.

To make the translation clear, let us evaluate $a \vDash \oslash\oslash$ old boiler.

We have $a \vDash \oslash\oslash$ old boiler iff for all e_i that we access via $\mathbb{P}$, $e_i \vDash \oslash$ old boiler.

Since we are arriving at e_i along the arrow $a \to e_i$ via $\mathbb{P}$ the double arrows fire (in Figure 5) and hence what we see at e_i is Figure 6b.

We now evaluate $e_i \vDash \oslash$ old boiler, and evaluating $\oslash$ does take account of the annotations and so we get:

$$e_i \vDash \oslash \text{ old boiler iff for all } x_i, x_i \vDash \text{ old boiler,}$$

old boiler with   old boiler   new
added safety   no safety   boiler

x' x y

$+$ $-$ $+$
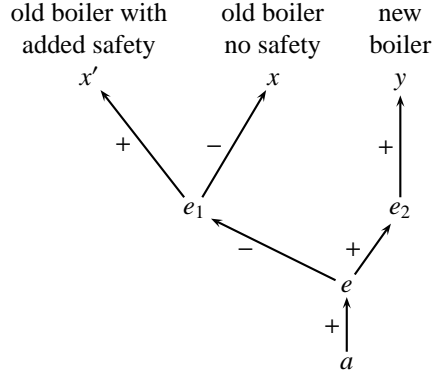
e_1 e_2

$-$ $+$

e

$+$

a

Figure 7: You ought to have a new boiler but if not then you should have a safe old boiler

which holds.

How do we translate (HC2)? It is not convenient to use Figure 5 to model (HC2). The scenario of Figure 7 is better. The difference is that we use only one agent node e.

In Figure 4 we have two agents, each agent is selling both types of boilers. In Figure 7, there is one agent, agent e, who has two departments, one selling old boilers and one selling new boilers, these two departments are now denoted by e_1 and e_2.

The path from e to the department that sells the old boilers e_1 is labeled '$-$' and the path to the department that sells the new boilers e_2 is labeled '$+$'. This is in accordance with (HC1). That means, by default our family is referred to the department that sells only new boilers. (HC1) may be translated according to Figure 7 as $\square \oslash \oslash \neg$ old boiler. However, we will fine-tune this further below.

As before, were our family to wish to see old boilers instead, they would have to request so. This is represented by means of double arrows that emanate from $a \to e$ and target both $e \to e_1$ and $e \to e_2$ (see Figure 8a). As the family passes through $a \to e$, a signal is sent to invert the labels of $e \to e_1$ and $e \to e_2$. Note that if the family moves through $a \to e$ via $\mathbb{P}$ then the switching takes place. However, if they move through $a \to e$ via $\lozenge$ then no switching takes place. The latter represents the situation in which we disregard the wishes of the family: then the labels that represent the obligations determine the available paths. The translation of the fact "old boiler" which we interpret as what the family wants to buy is according to Figure 8a $\oslash \oslash \oslash$ old boiler.

Now suppose our family is at department e_1 that sells the old boilers. (HC2) is realized by means of the labeling of the arcs $e_1 \to x'$ and $e_1 \to x$. The former leads to the node with the boilers that have added safety x' and hence it is labeled '$+$' in accordance with (HC2). The latter arc leading to the boilers without added safety x are labeled by '$-$'. Hence this path is blocked. So the translation of (HC2) will involve $\square \oslash \oslash$ old boiler with safety.

Example 3.1 motivated a sub-tree of the reactive tree that can be used for any contrary to duty situation. We will enrich it stepwise by adding to the complexity of the considered sets of conditional obligations.

We needed an explanation for our choice of points e, e_1, e_2. If we were to say directly without explanation that we want semantics where every accessibility configuration of the form aRx_i and aRy_i is replaced by aRe, eRe_1, eRe_2,
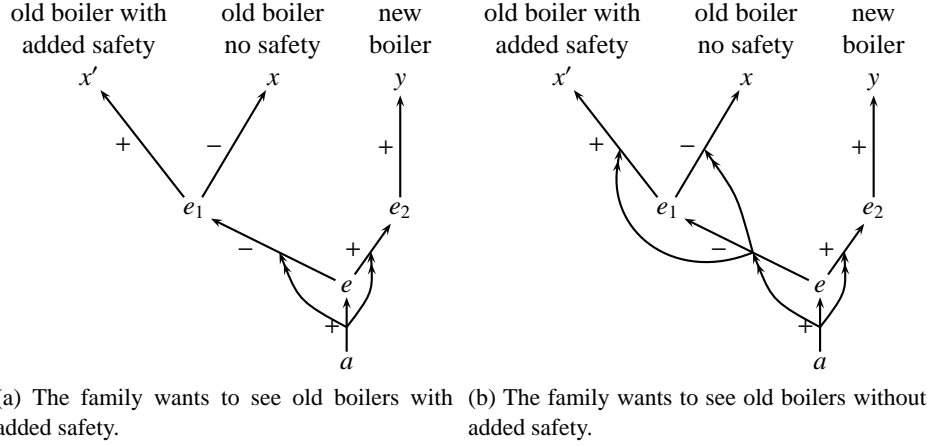
(a) The family wants to see old boilers with added safety.

(b) The family wants to see old boilers without added safety.

Figure 8

$e_1 R x_i, e_2 R y_j$, the reader would have wondered where and why these auxiliary e, e_1, e_2 pop up from! We have our explanation now.

We now summarise our graph model and analyse it. Once we do that we present the logical system formally and study its properties. We shall see that for technical reasons the pre-final model we want to use is Figure 9. This figure is the same as Figure 8a, except that we made the top nodes of the tree reflexive. The reasons for that are technical, having to do with substitution, as we shall discuss below.

Remark 3.2 (Summary of the proposed model). We have the CTD obligations of the form below (think of A=fence, B = white fence), where A and B are classical logic formulas:

(A1)   Obligatory that $\neg A$

(A2)   If it is the case that A then we should have $A \wedge B$ (in the context of A)

(A3)   It is the case that A

Figure 9 is the graph corresponding to (A1)–(A3) which, when viewed as a reactive Kripke model, is supposed to encode the logical information expressed by the set (A1)–(A3) and thus model it. We first explain Figure 9 and then describe how it is done.

The story behind Figure 9 is that we have a family standing at node a intending to buy a boiler. We have also an agent e selling boilers, having two departments, department e_1 selling old boilers A, which can be safe B, or not safe $\neg B$, and department e_2 selling new boilers $\neg A$.

The arcs represent paths which the family can take moving from their initial position to the boiler they want to get.

We want to use this setup to represent the contrary to duty set (A1), (A2) and (A3). The representation must be intuitive and make (A1),(A2) and (A3) all completely independent of one another.

(A1) is what is obligatory. We represent it by labeling the arcs leading to the departments e_1 and e_2: $e \rightarrow e_1$ is labeled '$-$' and $e \rightarrow e_2$ is labeled '$+$'. This intuitively directs the family from node a to the node e_2 from which $\neg A$ –the new boiler– is accessible.
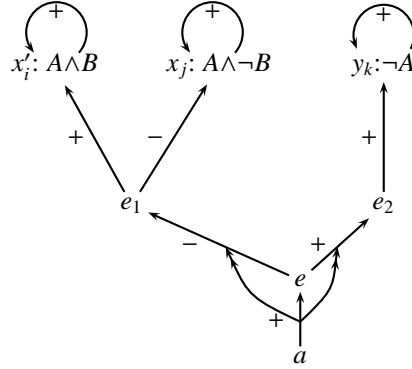
Figure 9: A model for Remark 3.2

Similarly, the contrary to duty obligation (A2) is represented by labeling the arcs. (A2) concerns what boilers are presented by the department e_1. In accordance with (A2) the arc $e_1 \rightarrow x'_i$ is labeled '+' while the arc $e_1 \rightarrow x_j$ is labeled '−'.

(A3) are the facts, which we interpret as what the family wants to buy, which means the path they want to follow. We give the family control of the double arrows so that they can force their way through the tree even when it means breaking obligations. In that case they need to invert the '−'-label of some arc by means of a double arrow. See Figures 8a and 8b. In both cases their wish is in conflict with (HC1). Hence, they use the double arrows to block the way to the dealer e_2 with the new boilers and instead free their way to the dealer e_1 who sells the old boilers. In Figure 8a their wish is to see old boilers with added safety and hence in coherence with (HC2). Thus, no additional double arrows are needed. The situation is different in Figure 8b where they want to see old boilers without added safety. Here they have to use the double arrows to invert the labels of $e_1 \rightarrow x$ and $e_1 \rightarrow x'$.

$x'_i, i = 1, 2, 3, \ldots$ are all points where $A \wedge B$ holds.

$x_j, j = 1, 2, 3, \ldots$ are all points where $A \wedge \neg B$ holds.

$y_k, k = 1, 2, 3, \ldots$ are all points where $\neg A$ holds.

Note that we added reflexive arrows to the top points, x'_i, x_j and y_k. The reason is that Standard Deontic Logic has the condition that every point has a world accessible to it. So we made the top points reflexive.

Let us now examine how we can express the ideas presented graphically by means of formulas. Let Q be a finite the set of atoms, and $n = |Q|$. In the following we presuppose a semantic frame in form a tree analogous to the model in Figures 9 which has a maximal depth of $n+1$ from the root node a to the top node. The top nodes are reflexive, as remarked above. In Section 5 we will show how this frame can be characterised axiomatically.

We define:

$$\Diamond X =_{\mathrm{df}} \neg \Box \neg X$$

$$\mathbb{P} X =_{\mathrm{df}} \neg \oslash \neg X$$

The position of a node in the graph indicates what is to be considered as being settled when our family reaches the node. For instance, the position of e_1 indicates that the

family settled for buying an old boiler, or, in the fence example, it indicates that there is a fence. We can express this as follows: where

$$\text{context}(A) =_{\text{df}} \Box^n A,$$

we have e.g., $e_1 \vDash \text{context}(A)$, $e_2 \vDash \text{context}(\neg A)$, $x_i' \vDash \text{context}(B)$, and $x_j \vDash \text{context}(\neg B)$. Note that the context of a node is established independent of the labels of the arcs. Note that due to the reflexivity of the top nodes the evaluation of $w \vDash \text{context}(A)$ always checks whether A is valid at all accessible top models from w. There is no need to keep track at which level of the tree the evaluation takes place.

Another important property of nodes is that they provide us with specific choices. For instance, at node e our family has the choice between settling for buying old (e_1) resp. new boilers (e_2). Similarly, at e_1 they have the choice between going for a boiler with or without added safety. This is expressed as follows: where

$$\text{switch}(A) =_{\text{df}} \Diamond\text{context}(\neg A) \land \Diamond\neg\text{context}(\neg A),$$

we have e.g., $e \vDash \text{switch}(A)$, $e_1 \vDash \text{switch}(B)$, and $e_1 \vDash \text{switch}(\neg A \lor B)$.

A conditional obligation $\bigcirc_A B$ expresses that whenever our family stands at a node where A is considered to be settled and they have the choice between a node in a context that is consistent with B and a node where $\text{context}(\neg B)$ holds, they should opt for the former node.

Translated in our language this can be expressed by in the following way:

$$\pi^i =_{\text{df}} \Box \ldots \Box \ (i \text{ times}) \text{ where } \pi \in \{\Box, \Diamond, \oslash, \mathbb{P}\}$$

$$\text{opt}(A) =_{\text{df}} \mathbb{P}\neg\text{context}(\neg A) \land \neg\mathbb{P}\text{context}(\neg A),$$

$$\text{obliged}_A B =_{\text{df}} \left[\bigwedge_{i=1}^{n} \Box^i(\text{context}(A) \land \text{switch}(B) \to \text{opt}(B))\right] \land$$
$$\left[\bigvee_{i=1}^{n} \Diamond^i(\text{context}(A) \land \text{switch}(B))\right], \text{ and}$$

$$\text{obliged}B =_{\text{df}} \text{obliged}_\top B$$

The conditional obligation is expressed by $a \vDash \text{obliged}_A B$. Note that in Figure 9 we have, for instance,

(*A1) $a \models \text{obliged}_\top \neg A$, and

(*A2) $a \vDash \text{obliged}_A A \land B$.

The first conjunct of $\text{obliged}_A B$, namely $\bigwedge_{i=1}^{n} \Box^i(\text{context}(A) \land \text{switch}(B) \to \text{opt}(B))$ expresses that, whenever our family is at a node where A is settled and where they have the choice between a path that is consistent with B and a path that leads to $\neg B$, they are supposed to opt for the former.

Note that a given conditional obligation $\bigcirc_A B$ imposes the following succession: if A is considered as settled, B is obliged. This is mirrored in the tree: first it splits between A (node e_1) and $\neg A$ (node e_2), then it splits between B (node x_i') and $\neg B$ (node x_j). The second conjunction of $\text{obliged}_A B$, namely $\bigvee_{i=1}^{n} \Diamond^i(\text{context}(A) \land \text{switch}(B))$, enforces the correct splitting in the tree. Were we to skip the second conjunct we would get pragmatic oddities. For instance, the model in Figure 9 would also verify $\text{obliged}_B A$ and $\text{obliged}_B \neg A$.

We have already mentioned that facts represent the wishes of our family. That means, whenever our family has the choice between A and $\neg A$ they "use" the double arrow to block the way to $\neg A$ and opt for A if they wish for A. For instance, when

our family approaches the boiler shop e, they activate the double arrows $(a \to e) \twoheadrightarrow (e \to e_1)$ and $(a \to e) \twoheadrightarrow (e \to e_2)$ which inverts the labels on the arcs $e \to e_1$ and $e \to e_2$. After all, their wish is to buy an old boiler (despite (HC1)). This is expressed as follows:

(*A3) $\oslash^{n+1} A$

For instance we have $a \vDash \oslash^3 A$ and $a \vDash \oslash^3 B$ in Figure 9. In Section 5 it is demonstrated how we can axiomatically ensure that in each model our family takes a unique path to exactly one of the models on top of the tree.

Remark 3.3.   1. It is very important to notice that if we substitute (A3*) for A inside (A1)* we get (A1)* back. This is because in Figure 9 the points x_i' and x_j and y_k are all reflexive, i.e. they are R accessible to themselves and hence all modalities collapse.

Therefore, for example, $\Box\Box\oslash(\oslash\Box\oslash A) \equiv \Box\Box\oslash A$, and similarly $\oslash\oslash\oslash(\oslash\Box\oslash A \wedge \oslash\Box\oslash B)$ is equivalent to $\oslash\oslash\oslash(A \wedge B)$. Therefore we can say that first we translate A and B and then substitute the translation in (A2) and (A1) and get (A2)* and (A1)*.

2. Also note that the translations of $A \to \bigcirc B$ is taken essentially as the translation of $\bigcirc_A B$. Therefore we represent the linguistic CTD 'If A then obligatory B' in our logic as saying $\bigcirc_A B$. To see this consider Figure 9. The labels implementing $A \to \bigcirc B$ are blocking the arc $e_1 \to x_j$ to the point where $\neg B$ holds by '$-$' and give access to the point where B holds by means of the '$+$'-label on $e_1 \to x_i'$. Compare this with the original obligation $\bigcirc\neg A$. The labels that implement this obligation are the '$-$'-label on $e \to e_1$ blocking access to e_1 and hence to the points where A holds, and the '$+$'-label on $e \to e_2$ giving access to e_2 where $\neg A$ holds.

3. Note that the assumption is that A and B are classical wffs and are considered complete units for the purpose of translation. Thus for example if X is a wff of classical logic and is translated into say $\oslash\Box\oslash X$ for whatever reason, then when $\neg X$ needs to be translated for the same reason, then $\neg X$ is translated as a unit into $\oslash\Box\oslash\neg X$ and not into $\neg\oslash\Box\oslash X$.

4. Note that the translation of the simple fact

(A3) It is the case that A

is rather complex. This happens because the node in the model where (A3) is "committed" and thus a violation of (A1) becomes inevitable, is the node e_2. The node where the obligation (A1) is put forward is the node a. One dimensional Kripke semantics allows us to evaluate at one point only. So we are forced to describe the entire set up from the context of a single point. It may be natural to take this point as e_2, where the violation is, but then we have no backward connective to take us back to point a, where the obligation is. Thus we describe all aspects of the model from point a, because our modalities allow us to go forward. So from point a we must isolate and talk about point e_2 where the violation occurs and make sure it is not "confused" with point e_1.

Expressing all of this becomes complicated.

Note that if we do not make these distinctions we get the paradoxes.
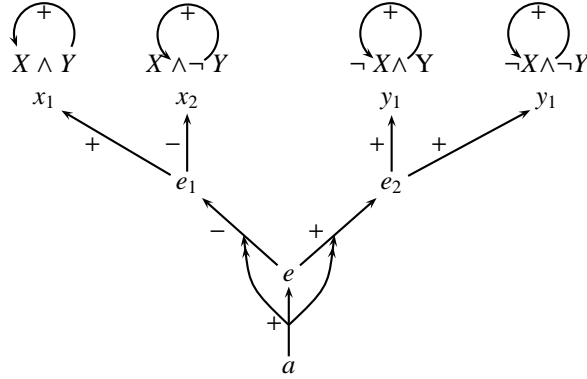
So the naive reader should NOT ask:

Figure 10: Modelling the dog example

- "why is the translation of (A3) complex, after all (A3) says is just a simple fact?!"

My answer would be:

"yes, it is because you keep it simple that you get paradoxes"

Note that in our paper [5] we keep it slightly more simple by using two dimensional modal logic, and so we can keep track of both points , the point of obligation and the point of violation.

Remark 3.4 (Dog example, see [21]). The CTD with the dog is slightly different from that of the fence. It is

(D1)   There should be no dog

(D2)   If there is a dog, there should be a sign

(D3)   There is a dog.

Both examples have the form

(X1) $\bigcirc\neg X$

(X2) $X \to \bigcirc Y$

(X3) X

In the fence case we have $\vdash Y \to X$, which we do not have in the dog case.

We have *not* used the special feature $\vdash Y \to X$ in our solution in Remark 3.2. The translation is the same. Figure 9 becomes Figure 10.

4   Checking for Paradoxes

Remark 4.1 (Rationality conditions). When we offer a translation τ into some deontic logic **L** and hope to show that we have no paradoxes, we must satisfy some rationality conditions for **L** and τ.

(ℝ1)   The semantics for **L** must be compatible in its spirit with our intuitive grasp of obligations and contrary to duties.

(ℝ2)   The translation of Δ into **L** must be consistent in **L** with no pragmatic oddities (i.e. the translations of Obligatory A and Obligatory $\neg A$ are consistent in **L**). Furthermore, the translation must be modular and incremental. For example, if Δ_1 is a set of obligations (which might be a paradox) and Δ_2 is another set, then

   (a) $\tau(\Delta_1 \cup \Delta_2) = \tau(\Delta_1) \cup \tau(\Delta_2)$

   (b) If $\Delta\theta$ is obtained from Δ by substitution θ of classical formulas to the atoms in Δ then $\tau(\Delta\theta) = \tau(\Delta)\theta$.[3]

Example 4.2 (Chisholm paradox and rationality of translation). We illustrate our rationality conditions by applying them to the Chisholm set.

1. It ought to be that a certain man goes to the assistance of his neighbour. Written as $\bigcirc H$.

2. It ought to be that if he does go he tells him he is coming, written as $H \rightarrow \bigcirc T$.

3. If he does not go then he ought not to tell him he is coming, written $\neg H \rightarrow \bigcirc \neg T$.

4. He does not go, written $\neg H$.

Note that if we take in this Chisholm set the subset comprised of (1), (3) and (4) and substitute for
$$\neg H = F$$
$$\neg T = F \wedge W$$

i.e. we let θ be the substitute $H = \neg F$ and $T = \neg(F \wedge W)$, we get the fence example:

1. $\theta : \bigcirc \neg F$

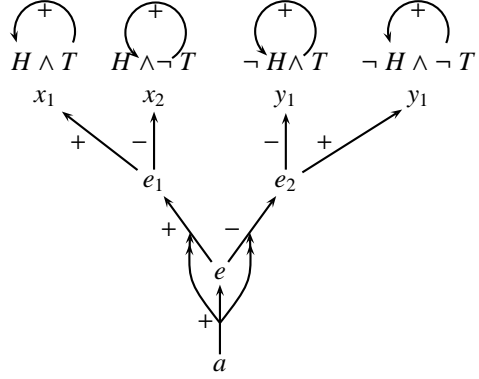2. $\theta : F \rightarrow \bigcirc(F \wedge W)$

3. $\theta : F$

Therefore any solution to the Chisholm paradox (1)–(4) via a logic **L** and a translation τ must also give a solution to the fence paradox by looking at $\tau(1)\theta, \tau(3)\theta$ and $\tau(4)\theta$.

   We shall see that the Jones-Pörn solution in [4] does not satisfy the rationality condition.

Example 4.3 (Translation of the Chisholm paradox). We now give the reader an intuitive idea for a rational solution to the Chisholm paradox. A proper solution will be given in Section 5, once we have a precise mathematical basis for our modelling.

   Take the set

1. $\bigcirc H$

2. $H \rightarrow \bigcirc T$

3. $\neg H \rightarrow \bigcirc \neg T$

4. $\neg H$

Figure 11a	Figure 9
H	$\neg A$
T	$\neg B$
x_1	y_k
y_1	x_j
y_2	x'_i
x_2	does not exist

(b)

(a) Chisholm paradox

Figure 11

We use Figure 11a for the translation. Note that it satisfies the rationality conditions.

Let us now reduce Figure 11a to Figure 9 (the fence example) by looking at clauses (1), (3), (4) of the Chisholm set and by using the substitution θ, where $\theta(H) = \neg F = \neg$fence, ($A$ in Figure 9) and $\theta(T) = \neg(F \wedge W)$, ($\neg(A \wedge B)$ in Figure 9). Note that this substitution complies with $\mathbb{R}2b$ of the Rationality Condition.

We get

$$x_1 = \neg F \wedge \neg(F \wedge W) = \neg F$$
$$x_2 = \neg F \wedge F \wedge W = \bot$$
$$y_1 = F \wedge \neg(F \wedge W) = F \wedge \neg W$$
$$y_2 = F \wedge (F \wedge W) = F \wedge W$$

We therefore get that after substitution Figure 11a becomes Figure 9 with the correspondence table as given in Figure 11b. Note that we use in Figure 11a not only the substitution θ, but we also drop out clause (2).

We have seen that our translation model makes the fence paradox and the Chisholm paradox disappear.

5 General theory of CTD's in SDLR1

This section describes a final general model for a general characterization of single CTD sets of arbitrary size (such as all previously considered examples, see the definition below). In the next section we give the most generic characterization that is able to deal with any sets of conditional obligations and facts.

Definition 5.1 (Generalised single CTD set). *Let $Q = \{p_1, \ldots, p_n\}$ be a finite set of atoms and $x_i \in \{p_i, \neg p_i\}$ for each $i \leq n$.*

A general single CTD set Δ of depth n contains

- $\Delta_0 = \{\bigcirc x_1\}$

---

[3]Carmo and Jones [3, p. 275] give the same rationality conditions except that they do not ask for condition $\mathbb{R}2b$. Indeed their system **DL** of [18] is not closed under substitution. See Example 4.2 below, and the discussion in Section 6 below, and see [3]. Our system does satisfy $\mathbb{R}2(b)$ as shown in Remark 3.3.

- *a non-empty subset Δ_{CTD}^i of $\{x_i \to \bigcirc x_{i+1}, \text{co} x_i \to \bigcirc \text{co} x_{i+1}\}$, where $\text{co} x_i = \neg p_i$ if $x_i = p_i$ and $\text{co} x_i = p_i$ else*

- *a set of facts $\Delta_F \subseteq \{\tau x_i \mid \tau \in \{\varepsilon, \text{co}\}, i \leq n\}$ where ε is the empty string*

Hence, Δ is a triple $\langle \Delta_0, \Delta_{CTD}, \Delta_F \rangle$ where $\Delta_{CTD} = \bigcup_{i=1}^{n-1} \Delta_{CTD}^i$.

Example 5.2. The the dog example and the Chisholm set fall under Definition 5.1.

Remark 5.3. For some CTD sets there are constraints. For instance in the fence example we have that w implies f (a white fence is a fence). We will first disregard such constraints but will later come back to them (see Remark 5.9).

When constructing a model for a CTD set Δ we have two degrees of freedom. On the one hand Δ_F may be incomplete in the sense that for some $p \in Q$ neither p nor $\neg p$ is in Δ_F. We want our models factually complete (so, for each $p \in \Delta_F$ either p is modeled or $\neg p$). Recall that the facts are taken to be the wishes of our family: they reflect which of the top models they want to reach. Hence, in each model M our family will reach a unique top model which determines the facts that are modeled by M. That means that for instance for the Chisholm set with fact g there are two corresponding models: one with fact t (where our family reaches the top model that verifies g and t), and one with $\neg t$ (where our family reaches the top model that verifies g and $\neg t$).

The second degree of freedom concerns cases in which for some $i < n$ e.g. $x_i \to \bigcirc x_{i+1} \notin \Delta_{CTD}$. There are different options to deal with this situation in a model of Δ: e.g., we could '+'-mark both arcs, the one leading from an x_i-node to an x_{i+1}-node and the one leading from an x_i-node to a $\text{co} x_{i+1}$-node. Or we could validate $\text{obliged}_{x_i} \text{co} x_{i+1}$ in the model (in which case the former arc is labeled '$-$' and the latter is labeled '+').

For the sake of simplicity we will in the following presuppose that Δ is complete in the sense that for each $i < n$, $x_i \to \bigcirc x_{i+1}, \text{co} x_i \to \bigcirc \text{co} x_{i+1} \in \Delta_{CTD}$.

Definition 5.4 (CTD model template for Δ). *Let Q be a finite set of atoms, let Δ be a generalised single CTD set based on Q. Let $\Delta_F^* \supseteq \Delta_F$ such that for each p_i, $\{p_i, \neg p_i\} \cap \Delta_F^* = \emptyset$. We define a CTD reactive model template for Δ and Δ_F^* as follows:*

1. *The frame has the form $(S, R_1, R_2, \mathbf{f}, a)$*

2. *Let $\mathsf{Q}_\Delta = \{\langle \tau_1 x_1, \ldots, \tau_k x_k \rangle \mid k \leq n, \tau_i \in \{\varepsilon, \text{co}\}\}$. For each $s \in \mathsf{Q}_\Delta$, $\overline{s}$ is the set of all members of s.*

3. *$S = \{a, e, e_s \mid s \in \mathsf{Q}_\Delta\}$*

4. *R_1 contains the following pairs:*

    - *(a, e)*

    - *(e, e_{x_1}) and $(e, e_{\text{co} x_1})$*

    - *$(e_s, e_{s \circ x_i})$ where $s, s \circ x_i \in \mathsf{Q}_\Delta$ and $\circ$ is the concatenation function*

    - *$(e_s, e_{s \circ \text{co} x_i})$ where $s, s \circ \text{co} x_i \in \mathsf{Q}_\Delta$*

    - *(e_s, e_s) where s has length n*

5. *the graph has the following labels (where $s \in \mathsf{Q}_\Delta$):*

    - *$\mathbf{f}(e \to e_{\text{co} x_1}) = 0$*

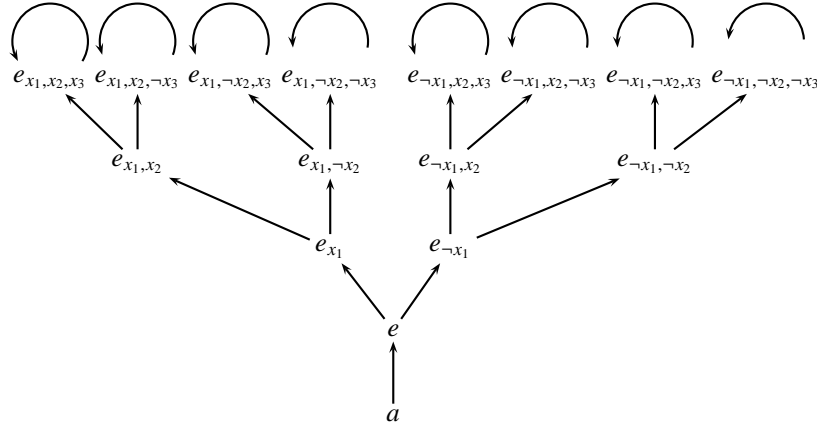   - *$\mathbf{f}(e_{s \circ x_i} \to e_{s \circ x_i \circ \text{co} x_{i+1}}) = 0$*

Figure 12: The tree template with $n = 3$

- $\mathbf{f}(e_{s \circ \mathsf{co} x_i} \to e_{s \circ \mathsf{co} x_i \circ x_{i+1}}) = 0$
- *all other labels are 1*

6. *R_2 contains the double arrows as follows:*

   - $(a \to e) \twoheadrightarrow (e \to e_{x_1})$ *and* $(e \to e) \twoheadrightarrow (e \to e_{\mathsf{co} x_1})$ *iff* $\mathsf{co} x_1 \in \Delta_F^*$

   - $(e_s, e_{s'}) \twoheadrightarrow (e_{s'} \to e_{s' \circ y})$ *iff* $\overline{s'} \subseteq \Delta_F^*$ *and (i)* $\mathbf{f}(e_{s'} \to e_{s' \circ y}) = 1$ *and* $\mathsf{co} y \in \Delta_F^*$, *or (ii)* $\mathbf{f}(e_{s'} \to e_{s' \circ y}) = 0$ *and* $y \in \Delta_F^*$.

7. *for the assignment function $v : S \to 2^Q$ we have the following requirement for the models e_s on top of our tree $v(e_s) = \overline{s} \cap Q$*

Remark 5.5. Item 4 in Definition 5.4 defines a modal frame in form of a tree. Figure 12 illustrates this for $n = 3$. Figure 13 applies the labeling as in item 5 and the double arrows as in item 6 for the CTD set

$$\Delta_0 = \{\bigcirc x_1\},$$
$$\Delta_{CTD} = \{x_1 \to \bigcirc x_2, \mathsf{co} x_1 \to \bigcirc \mathsf{co} x_2, x_2 \to \bigcirc x_3, \mathsf{co} x_2 \to \bigcirc \mathsf{co} x_3\},$$
$$\Delta_F = \{\neg x_1, \neg x_2, x_3\}$$

Example 5.6 (Chisholm set). Compare with Example 4.3. We have $Q = \{g, t\}$.

$$\Delta_0 = \{\bigcirc g\}$$
$$\Delta_{CTD} = \{g \to \bigcirc t, \neg g \to \bigcirc \neg t\}$$
$$\Delta_F = \{\neg g\}.$$

Our models are illustrated in Figures 14a and 14b (the former for $\Delta_F^* = \{\neg g, \neg t\}$ and the latter for $\Delta_F^* = \{\neg g, t\}$).

Some authors use $O(g \to t)$ in the Chisholm set. This is no problem for us, we can translate it, as we explained in item 5 of Remark 3.3.

Example 5.7 (Chisholm with $n = 3$). We now enhance our Chisholm set. g stands for "going", t stands for "telling", o stands for "dressing your overall". We enhance the usual Chisholm set $\{\bigcirc g, g \to \bigcirc t, \neg g \bigcirc \neg t\}$ by $\{t \to \bigcirc o, \neg t \to \bigcirc \neg o\}$: if you tell you should dress your overall, if you don't tell you should not dress your overall.

Obviously this set is in accordance with Definition 5.1. In Figure 15 we have a model for $\Delta_F^* = \{\neg g, t, \neg o\}$ constructed as in Definition 5.4.
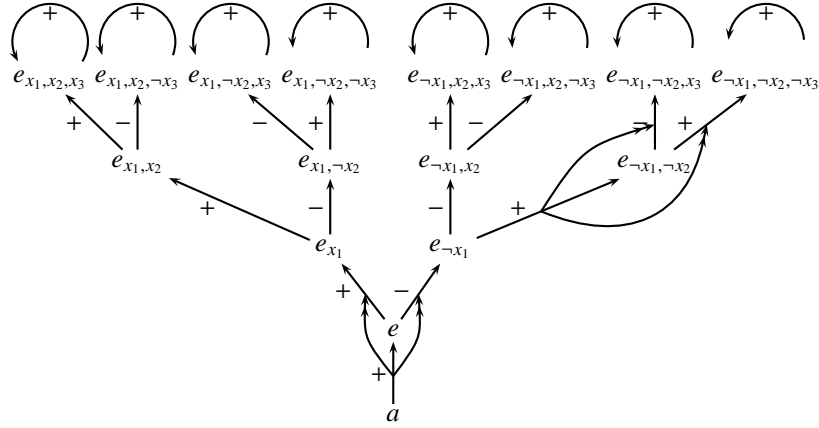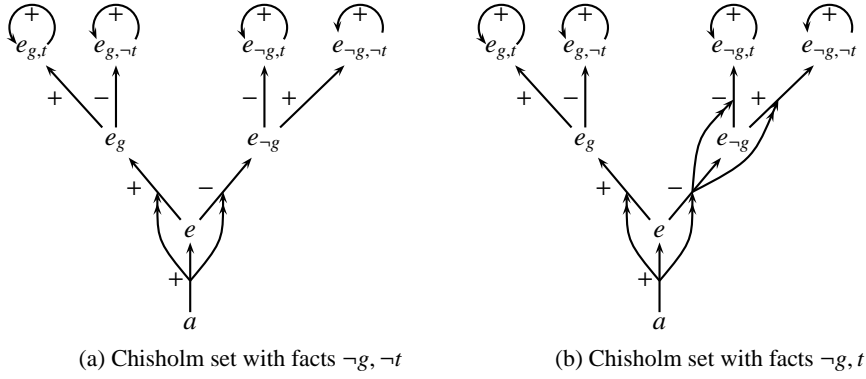
Figure 13: The model template with $n = 3$ (incl. labels and double arrows)

(a) Chisholm set with facts $\neg g, \neg t$

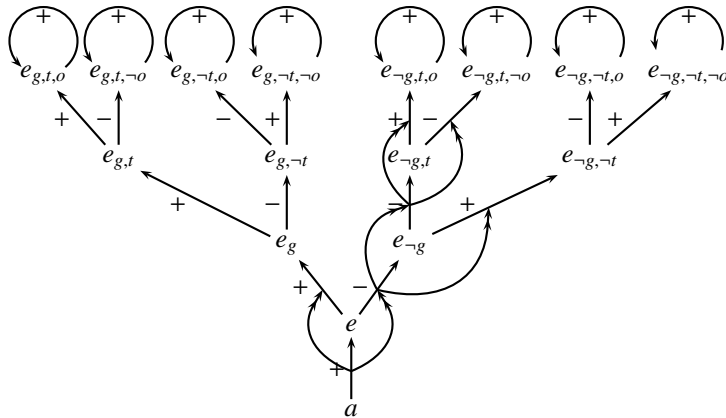(b) Chisholm set with facts $\neg g, t$

Figure 14

Figure 15: Chisholm with $n = 3$

Example 5.8 (Expressive power of $\square$ and $\oslash$)**.** We try to characterise the graph axiomatically (as a Kripke frame) using $\square$ by writing modal axioms holding at a. The logic is non-normal, so the axioms hold at a only. If we want an axiom to hold at level two points such as e_g and $e_{\neg g}$ in Figure 15 we need to put a $\square$ in front. The wffs X in the axioms do not contain $\oslash$.

The following are the axiom schemas:

(*1) $\lozenge X \to \square X$

(*2) $\square X \to \lozenge X$

(*1) and (*2) say that a has exactly one accessible possible world e

(*3) $\bigwedge_{j=1}^{n} \square^j (\bigwedge_{i=1}^{3} \lozenge X_i \to \bigvee_{i \neq j \leq 3} \lozenge(X_i \wedge X_j))$

(*3) says that each e_s has at most 2 accessible worlds.

(*4) $\square^{n+1}(X \leftrightarrow \square X)$

(*4) expresses the reflexivity of our top nodes.

We now turn to our translation of Δ. While discussing the various parts of Δ, namely Δ_{CTD} and Δ_F, we will still add some more axioms in order to characterise our frame.

Now we make sure that the labels representing our obligations are placed correctly and the splitting of the branches of the tree is in correspondence to the violation order imposed by Δ. Each $x \to \bigcirc y \in \Delta_{CTD}$ is translated by:

$$\left[\bigwedge_{i=1}^{n} \square^i(\mathsf{context}(x) \wedge \mathsf{switch}(y) \to \mathsf{opt}(y)) \right] \wedge \left[\bigvee_{i=1}^{n} \lozenge^i(\mathsf{context}(x) \wedge \mathsf{switch}(y)) \right] \quad (\delta_{CTD})$$

(δ_{CTD}) expresses that whenever at some point e_s at which $\mathsf{context}(x)$ holds there is an arc towards a node where $\mathsf{context}(\neg y)$ holds and another one towards a node where $\neg\mathsf{context}(\neg y)$ holds, then the path towards $\mathsf{context}(\neg y)$ is blocked and the path towards $\neg\mathsf{context}(\neg y)$ is accessible. This realizes the obligation to y in the context x. The second conjunct of (δ_{CTD}) makes sure that the splitting of the branches of the tree are as expected (i.e., in accordance with the violation order).

The facts are modeled by means of the following formulas. Where x is a fact we add:
$$\oslash^{n+1} x \qquad\qquad (\delta_F)$$

Furthermore we add:

(*5) $\neg \mathbb{P}^{n+1} \bot$

(*6) $\bigwedge_{i=1}^{n+1} (\mathbb{P}^i X \wedge \mathbb{P}^i X' \to \mathbb{P}^i(X \wedge X'))$

These make sure that our family takes one unique path to a top node in each model.

(*7) $\bigwedge_{i=0}^{n-2} \square^i \neg \oslash \bigvee_{j=1}^{n-1-i} \lozenge^j (\lozenge X \wedge \neg \mathbb{P} X)$

This formula makes sure that all double arrows are of the type $(e_s \to e_{s'}) \twoheadrightarrow (e_{s'} \to e_{s''})$ (i.e., double arrows always target arcs in the next level in the graph).

Hence, where (*1)–(*7) axiomatize our frame, (δ_{CTD}) and (δ_F) provide the translation of our CTD set Δ.

24

Remark 5.9 (Constraints). Suppose we have a set of constraints $\Delta_C = \{\neg \bigwedge \mathcal{E} \mid \mathcal{E} \subset Q^l\}$ such as $\neg(w \wedge \neg f)$ in the fence example (a white fence is a fence). In this case we have to change the construction of our model template slightly. For instance we adjust $Q_\Delta = \{\langle \tau_1 x_1, \ldots, \tau_k x_k \rangle \mid k \leq n, \tau_i \in \{\varepsilon, \text{co}\}$, there is no $\mathcal{E} \subseteq \{\tau_i x_i \mid i \leq k\}$ such that $\neg \bigwedge \mathcal{E} \in \Delta_C\}$.

Constraints can be translated into formulas by $\neg \Diamond^{n+1} \bigwedge \mathcal{E}$.

Remark 5.10 (Independence). Both (a) the translations of the members of CTD-sets and (b) the instructions for attaching the labels to the arcs and for drawing the arrows are fully independent in our treatment. There are no logical dependencies between the translations of the obligations and facts of a given CTD set.[4]

Remark 5.11 (Solution to Chisholm). Consider the Chisholm set of Examples 4.3 and 5.6. We are seeking a logic **L** and a translation τ of the Chisholm sentences into **L** such that

1. All translations are logically independent in **L**.

2. The translation is consistent in **L**.

3. The translation satisfies the rationality conditions of Remark 4.1.

Let the logic **L** be characterised as in Example 5.8, then (δ_{CTD}) and (δ_F) give the translation of the Chisholm set.

6 Towards more generality: Permissions, separable conjunctive obligations, specificity cases and CTD cycles

While in the last section we gave a detailed account of how we model a generalized single CTD set in **SDLR1**, in this section we allow for less restrictive configurations. Namely, we will also allow for cycles in CTD sequences and specifically for specificity cases. Moreover, we show how different concepts of obligations as well as permissions can be expressed within our logic.

Example 6.1 (Extended Chisholm set with cycle). We now add $\bigcirc_t g$ and $\bigcirc_{\neg t} g$ to our Chisholm configuration. Hence, we have a cycle: $\bigcirc g, \bigcirc_g t, \bigcirc_t g$. The new conditional obligations $\bigcirc_t g$ and $\bigcirc_{\neg t} g$ introduce a new progression: namely the one where we first

---

[4]Note that the placement of the double arrows as presented in Definition 5.4 item 6 is a function of the labeling of the tree. This should not be taken as an indicator for dependence. The reason is that the instruction is generic and independent of the concretely given (obligations that are represented by) labels.

Our perspective is that the facts reflect the wishes of our family. The double arrows enable them to force their way to achieve what they want even if it violates obligations. If what they want violates an obligation the path will be labeled '−'. In this case a double arrow will free the way for them. This can be phrased in a generic way that is independent of the concrete obligations that are in place. E.g., if they want A such a generic instruction which is independent of whether we have $\bigcirc \neg A$ or not is: if A is illegal and hence the path is barred by '−' then use a double arrow to force your way. For more complex facts that instruction is the generic case distinction offered in item 6 of our Definition.

Another more subtle worry is the following. By means of the double arrows we allow our family to react to the given obligations and force their way to reach what they want. One may argue that the norm giver may use a similar strategy: depending an what the individuals that are subjected to the norms do he may change the norms. E.g., if there is a certain loophole in the tax system and too many individuals use it in order to avoid taxation the norm giver may adjust the law accordingly in order to fill the loophole. Our reply is that in this paper we consider the norms as statically given and we do not study **RSDL1** as a system to model norm change.
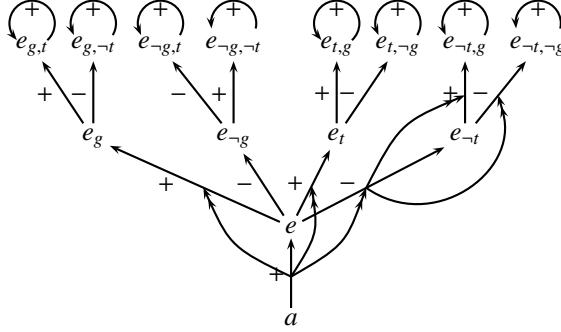
Figure 16: Extended Chisholm with facts $\neg t, \neg g$ and obligations $\bigcirc_{\neg t} g, \bigcirc_t g$

consider whether t resp. $\neg t$ is settled and then are concerned with the question whether we should do g. This is modeled semantically by introducing this sequence in the tree. See Figure 16. While the left hand side of the tree is as before, we add the progressions on the right hand side (i.e., the sub-trees $\langle a, e, e_t, \ldots \rangle$ and $\langle a, e, e_{\neg t}, \ldots \rangle$).[5]

Example 6.2 (Cycle with $n = 3$). We consider the following configuration

$$\{\bigcirc p_1, \bigcirc_{p_1} p_2, \bigcirc_{\neg p_1} \neg p_2, \bigcirc_{p_2} p_3, \bigcirc_{\neg p_2} \neg p_3, \bigcirc_{p_3} p_1, \bigcirc_{\neg p_3} p_1\}$$

See Figure 17 for the modeling which is in analogy to the previous example.

Example 6.3 (Makinson's Moebius example). Makinson gave the following cyclic example:

$$\{\bigcirc_{p_1} p_2, \bigcirc_{p_2} p_3, \bigcirc_{p_3} \neg p_1\}$$

See Figure 18 for a model of this set.

Remark 6.4 (Parallel CTD sets, the problem with specificity). Another question is: how to deal with parallel CTD sets such as

$$\{x_i^j \to \bigcirc x_{i+1}^j, \text{co}x_i^j \to \bigcirc \text{co}x_{i+1}^j \mid i \leq n(j) - 1, j \leq m\}$$

where for each $j \leq m$ the set $\{x_i^j \to \bigcirc x_{i+1}^j, \text{co}x_i^j \to \bigcirc \text{co}x_{i+1}^j \mid i \leq n(j) - 1\}$ is a CTD set as in Definition 5.1.

If some of the CTD sets share a member, it may give rise to specificity cases. A case of specificity occurs whenever we have a conditional obligation $\bigcirc_A B$ and a more specific context $A \wedge A'$ (where $A \wedge A' \vdash A$ and $A \nvdash A \wedge A'$) such that $\neg \bigcirc_{A \wedge A'} B$, and A' is consistent with B ($\nvdash \neg(A' \wedge B)$). It is hence a specific failure of monotonicity in the condition of the obligation.

For instance, we can give another twist to the Chisholm set by introducing a case of specificity. While in Example 6.1 we had $\bigcirc_{\neg t} g$, we could instead add $\bigcirc_{\neg t} \neg g$ (after all, as a rule of politeness, you are supposed never to visit unannounced). Note that this set cannot be modeled by means of the translation we offered for conditional obligations in Sections 4 and 5. The reason is that: (a) $\text{obliged}_{\neg t} \neg g$ enforces that at some node e_s in the tree we have $e_s \vDash \text{context}(\neg t) \wedge \text{switch}(\neg g)$, and (b) that $e_s \vDash \text{opt}(\neg g)$. However, by $\text{obliged} g$ we also have $e_s \vDash \text{opt}(g)$ (in contradiction to (b)) since trivially $e_s \vDash \text{context}(\top) \wedge \text{switch}(\neg g)$.

---

[5]Hence, if we want our logic to model these more general cases we have to drop (*3) from our axiomatisation in Example 5.8.
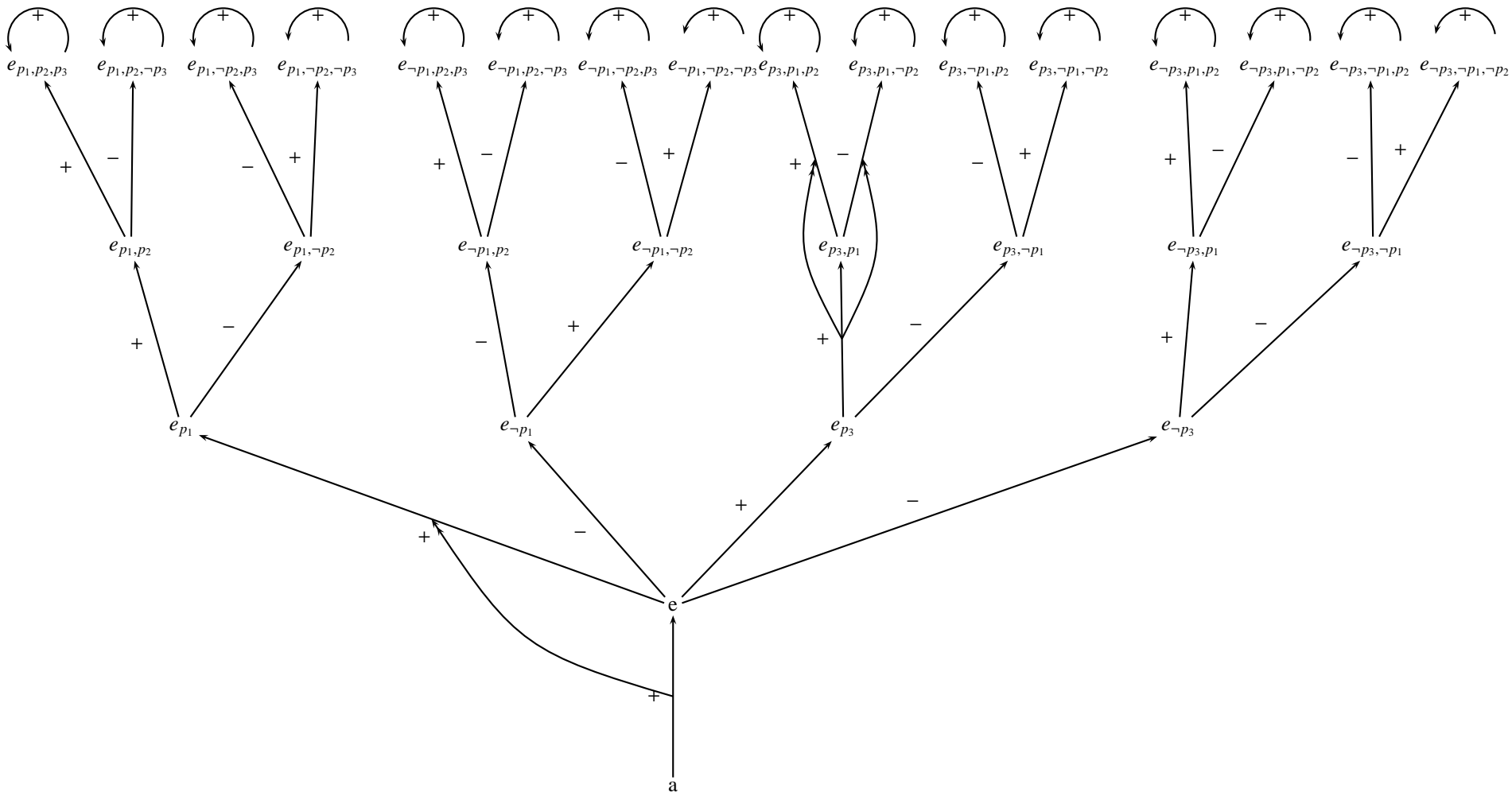
e_{p_1,p_2,p_3} $e_{p_1,p_2,\neg p_3}$ $e_{p_1,\neg p_2,p_3}$ $e_{p_1,\neg p_2,\neg p_3}$ $e_{\neg p_1,p_2,p_3}$ $e_{\neg p_1,p_2,\neg p_3}$ $e_{\neg p_1,\neg p_2,p_3}$ $e_{\neg p_1,\neg p_2,\neg p_3}$ e_{p_3,p_1,p_2} $e_{p_3,p_1,\neg p_2}$ $e_{p_3,\neg p_1,p_2}$ $e_{p_3,\neg p_1,\neg p_2}$ $e_{\neg p_3,p_1,p_2}$ $e_{\neg p_3,p_1,\neg p_2}$ $e_{\neg p_3,\neg p_1,p_2}$ $e_{\neg p_3,\neg p_1,\neg p_2}$

e_{p_1,p_2} $e_{p_1,\neg p_2}$ $e_{\neg p_1,p_2}$ $e_{\neg p_1,\neg p_2}$ e_{p_3,p_1} $e_{p_3,\neg p_1}$ $e_{\neg p_3,p_1}$ $e_{\neg p_3,\neg p_1}$

e_{p_1} $e_{\neg p_1}$ e_{p_3} $e_{\neg p_3}$

e

a

Figure 17: A model with a cycle for $n = 3$ and with facts $p_3, p_1, \neg p_2$
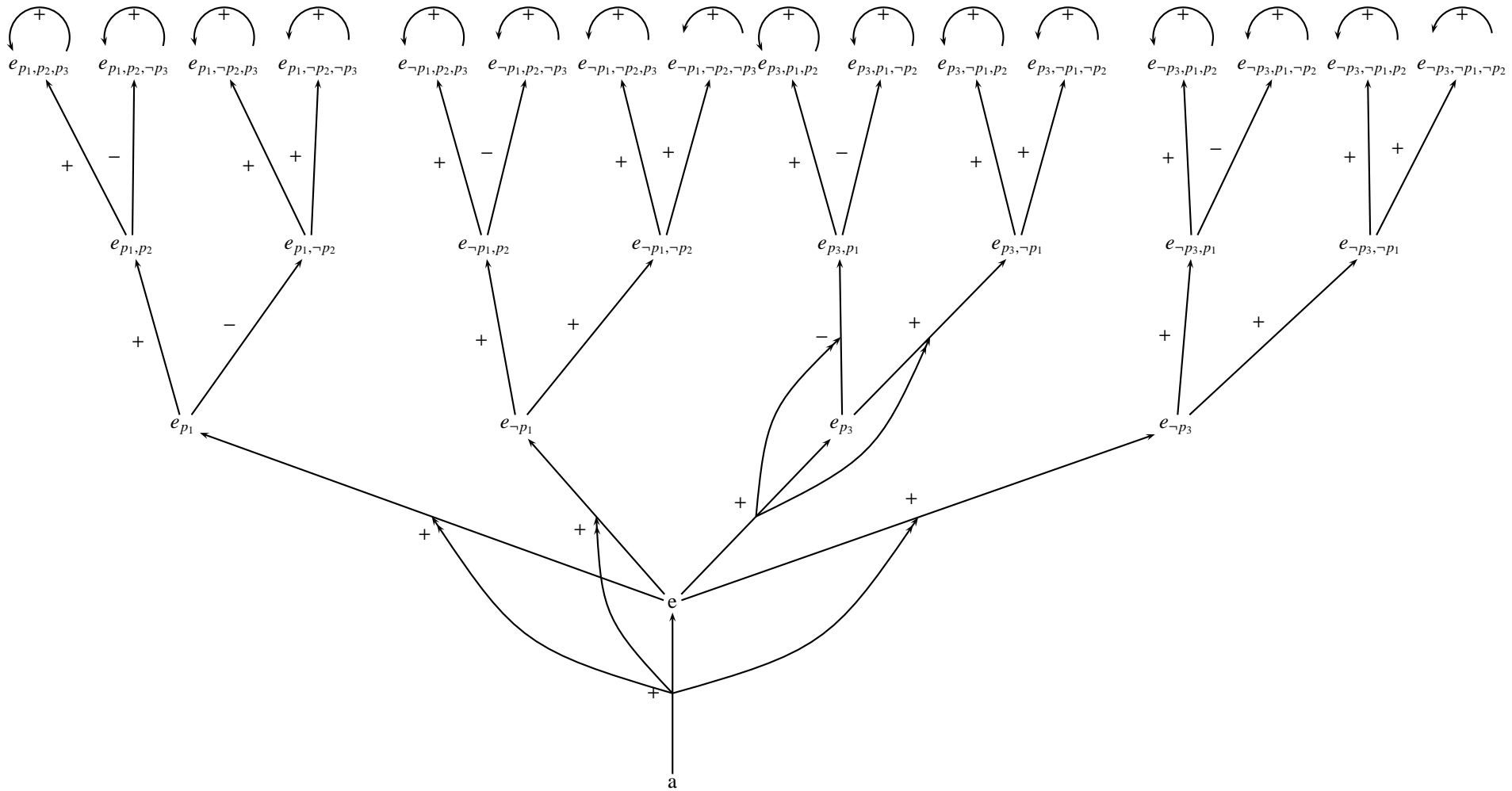
Figure 18: A model for Makinson's Moebius example with facts p_3, p_1, p_2

Specificity can easily arise when we combine two CTD sets that share members. Take as a simple example the two CTD sets $\Delta_1 = \{\bigcirc a, a \to \bigcirc b, \neg a \to \bigcirc \neg b\}$ and $\Delta_2 = \{\bigcirc c, c \to \bigcirc b, \neg c \to \bigcirc \neg b\}$. Now suppose $a \wedge \neg c$: due to $a \to \bigcirc b$ we expect $(a \wedge \neg c) \to \bigcirc b$, however, due to $\neg c \to \bigcirc \neg b$ we expect $(a \wedge \neg c) \to \bigcirc \neg b$. Hence, we either get a violation of the monotonicity of $a \to \bigcirc b$ or a violation of the monotonicity of $\neg c \to \bigcirc \neg b$. Since so far we have not introduced means to deal with specificity cases, we cannot represent such combinations of CTD sets.

However, if the various CTD sets do not share any literals (i.e., $\{x_i^j, \text{co}x_i^j \mid i \leq n(j)\} \cap \{x_i^k, \text{co}x_i^k \mid i \leq n(k)\} = \emptyset$ for all $j \neq k$ where $j, k \leq m$) we can easily construct models such as the one in Figure 19 for the combination of the two CTD sets $\Delta_1 = \{\bigcirc a, a \to \bigcirc b, \neg a \to \bigcirc \neg b\}$ and $\Delta_2 = \{\bigcirc c, c \to \bigcirc d, \neg c \to \bigcirc \neg d\}$.

Example 6.5 (Ross' Paradox)**.** The Ross paradox concerns the inference from $\bigcirc l$ ("You ought to post the letter.") to $\bigcirc(l \vee b)$ ("You ought to post the letter or burn it."). By many authors this is considered as being not in coherence with our every-day usage of disjunctive obligations. In our semantics we can embed the Ross set paradox-free.

Figure 20 features a model for $\{\bigcirc l, \neg \bigcirc (l \vee b), \bigcirc \neg b, \neg l, \neg b\}$.

Remark 6.6 (Deliberative and Non-deliberative obligations)**.** Note that our translation of obligations has a deliberative flavor. Take for instance the fence example. Here we have the constraint $\neg(w \wedge \neg f)$ (a white fence is a fence, see Remark 5.9). Hence, in Figure 9 (where A is substituted by f and B is substituted by w), all the worlds above e_2 verify $\neg w \wedge \neg f$. Nevertheless, we do not have $\text{obliged}_{\neg f} \neg w$. The reason is as follows. We have $e_2 \vDash \text{context}(\neg f)$. However, we do not have $e_2 \vDash \text{switch}(\neg w)$ since we do not have $e_2 \vDash \Diamond w$. Thus, from e_2 on $\neg w$ is necessary.

We can also express non-deliberative obligations in our framework:

$$\text{obliged}_A^* B =_{\text{df}} \left[\left(\bigwedge_{i=1}^n \Box^i(\text{context}(A) \wedge \text{switch}(B) \to \text{opt}(B))\right) \wedge \right.$$
$$\left(\bigvee_{i=1}^n \Diamond^i(\text{context}(A) \wedge \text{switch}(B))\right)\right] \vee$$
$$\left[\bigwedge_{i=1}^n \Box^i(\text{context}(A) \to \text{context}(B)) \wedge \bigvee_{i=1}^n \Diamond^i\text{context}(A)\right]$$

What is new in comparison to the definition of $\text{obliged}_A B$ is the third line. It gives the non-deliberative flavor to obligations and expresses that $\text{obliged}_A^* B$ holds also if any path leading to A leads also to B.

Note that the model in Figure 9 verifies $\text{obliged}_{\neg f}^* \neg w$.

Remark 6.7. Note that $\text{obliged}_C(A \to B)$ is not equivalent to $\text{obliged}_{C \wedge A} B$ in our logic. This is demonstrated in Figure 21.

Remark 6.8 (Permissions)**.** A permission $P_A B$ can be expressed by means of

$$\text{permitted}_A^* B =_{\text{df}} \left[\left(\bigwedge_{i=1}^n \Box^i(\text{context}(A) \wedge \text{switch}(B) \to \mathbb{P}\neg\text{context}(\neg B))\right) \wedge \right.$$
$$\left(\bigvee_{i=1}^n \Diamond^i(\text{context}(A) \wedge \text{switch}(B))\right)\right] \vee$$
$$\left[\bigwedge_{i=1}^n \Box^i(\text{context}(A) \to \text{context}(B)) \wedge \bigvee_{i=1}^n \Diamond^i\text{context}(A)\right]$$

The first line expresses that whenever our family has the choice between $\neg B$ and a path that is compatible with B, they are free to go for the second since the corresponding arc is labeled by '+'.

The second line enforces the correct succession in the tree (see also the discussion of the definition of obliged in Section 3). Were we to skip this requirement we would face pragmatic oddities: e.g., the model in Figure 14a would verify $\text{permitted}_t \neg g$.
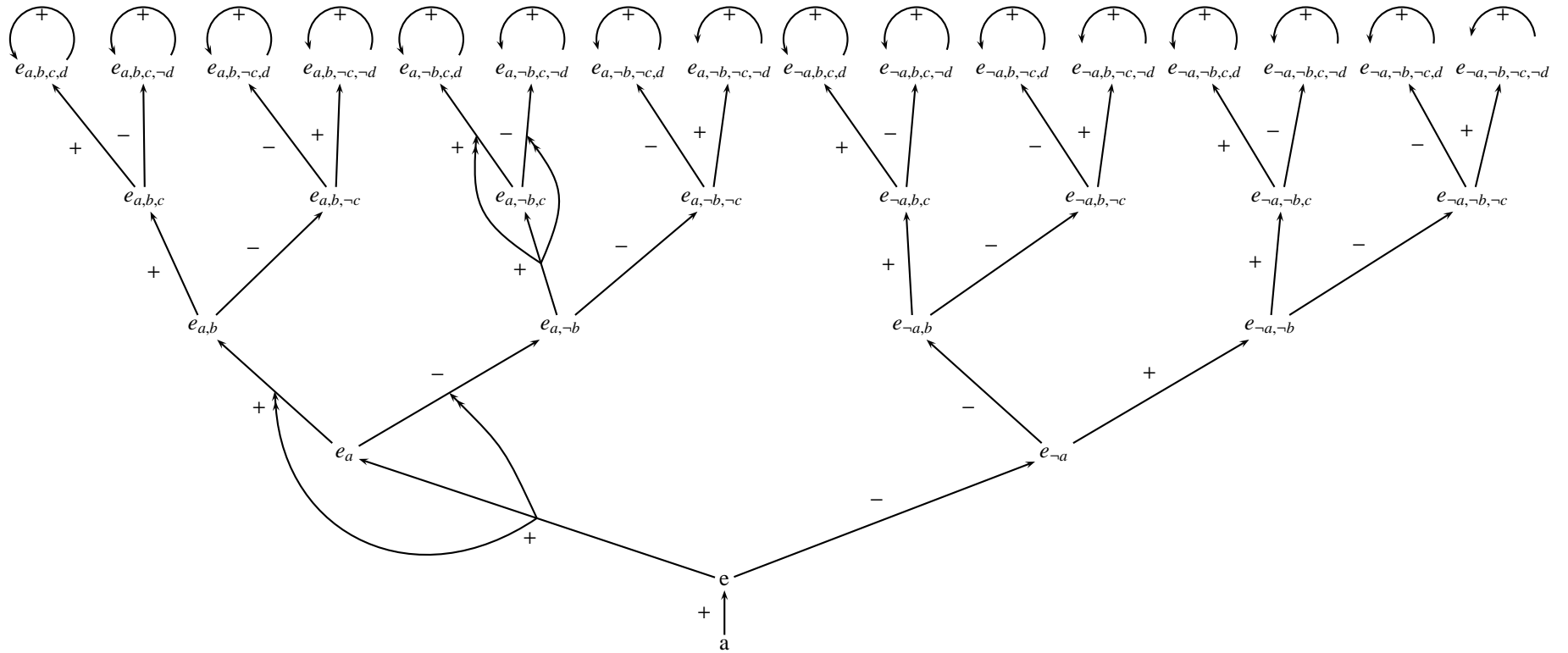
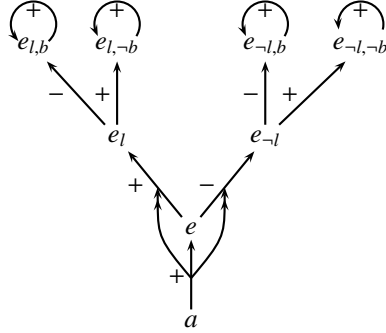Figure 19: Modeling multiple CTDs with facts $a, \neg b, c, \neg d$
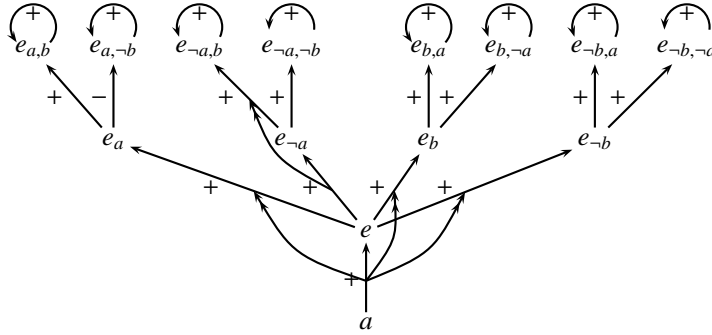
Figure 20: Ross' Paradox with facts $\neg l, \neg b$

Figure 21: A model that verifies $\neg\mathsf{obliged}(a \to b)$ and $\mathsf{obliged}_a b$ for the facts $\neg a, \neg b$

The third line expresses that in case every path leading to A is also leading to B, our family is allowed to bring about B in the context A (after all, in this context they have no other choice).

Note that weak permissions $PA =_{\mathrm{df}} \neg \bigcirc \neg A$ are problematic for logics that block the Ross-inference from $\bigcirc l$ to $\bigcirc(l \vee b)$ (such as **SDLR1**). The reason is that in such logics $\bigcirc l, \neg\bigcirc(l \vee b)$ is consistent. Now given that $\bigcirc(l \vee b)$ is equivalent to $\bigcirc\neg(\neg l \wedge \neg b)$, we get $P(\neg l \wedge \neg b)$, which is obviously counter-intuitive. Hence, our permission is not weak.

Note that the model in Figure 20 verifies: $\mathsf{permission}_\top(l \wedge \neg b)$, $\mathsf{permission}_{\neg l}\neg b$, $\neg\mathsf{permission}_\top(\neg l \wedge b)$, $\neg\mathsf{permission}_\top(\neg l \wedge \neg b)$, and $\neg\mathsf{permission}_\top\neg l$.

Remark 6.9 (Separable and inseparable conjunctive obligations)**.** Note that in our translation of conditional obligations we do not get

$$\mathsf{obliged}_A(B \wedge C) \equiv \mathsf{obliged}_A B \wedge \mathsf{obliged}_A C. \qquad\qquad \text{(Sep)}$$

Conjunctive norms of that kind are often referred to as "inseparable conjunctive norms". Consider the following example:

Figure 22 is a model of

$$\{\mathsf{obliged}(a \wedge b), \mathsf{obliged}_{\neg a}\neg b, \mathsf{obliged}a, \neg\mathsf{obliged}b, \mathsf{obliged}_{\neg b}a, \neg\mathsf{obliged}_{\neg b}\neg a\}$$

Suppose we are supposed to buy ingredients for a strawberry cake which is to be baked tonight for the visit of a good friend. a stands for "buying flour" and b stands for "buying strawberry". Suppose the shop is out of flour and we have no time left to go to
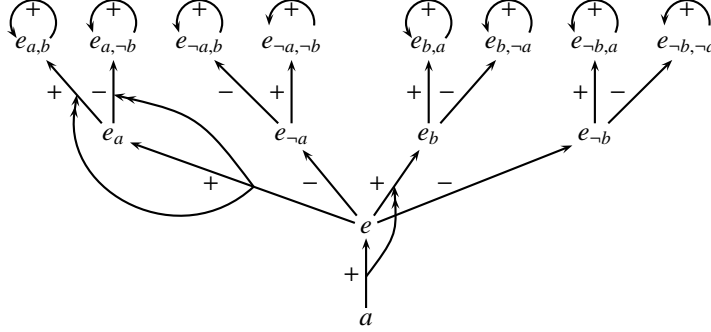
Figure 22: Inseparable conjunctive norms: a model of $\mathsf{obliged}_\top(a \wedge b)$, $\neg\mathsf{obliged}_\top^{\mathsf{sep}}(a \wedge b)$, $\mathsf{obliged}_\top a$, $\neg\mathsf{obliged}_\top b$, and facts $a, \neg b$.

another shop. In this case we have to cancel our plan to bake the cake and are supposed not to buy strawberries since we have no use for them: hence $\bigcirc_{\neg a}\neg b$. Now suppose the shop ran out of strawberries but still has flour. Since the flour can still be used for baking fresh bread, we are still supposed to buy it: $\bigcirc_{\neg b}a$. This situation is modeled by the model in Figure 22. Note that this cannot be modeled if we have (Sep).

Now the question arises whether our logic is expressible enough also to represent separable conjunctive norms for which (Sep) holds. This can be achieved as follows: where $\mathcal{E} \subseteq Q^l$, we define $\mathsf{obliged}_A^{\mathsf{sep}} \bigvee \mathcal{E}$ just as $\mathsf{obliged}_A \bigvee \mathcal{E}$ by

$$
\left[\bigwedge_{i=1}^{n} \Box^i (\mathsf{context}(A) \wedge \mathsf{switch}(B) \to \mathsf{opt}(B)) \right] \wedge \left[\bigvee_{i=1}^{n} \Diamond^i (\mathsf{context}(A) \wedge \mathsf{switch}(B)) \right]
$$

Moreover, we define:

$$\mathsf{obliged}_A^{\mathsf{sep}}(B \wedge C) =_{\mathrm{df}} \mathsf{obliged}_A^{\mathsf{sep}} B \wedge \mathsf{obliged}_A^{\mathsf{sep}} C$$

$$\mathsf{obliged}_A^{\mathsf{sep}} \neg\neg B =_{\mathrm{df}} \mathsf{obliged}_A^{\mathsf{sep}} B$$

$$\mathsf{obliged}_A^{\mathsf{sep}} \neg(B \vee C) =_{\mathrm{df}} \mathsf{obliged}_A^{\mathsf{sep}} (\neg B \wedge \neg C)$$

$$\mathsf{obliged}_A^{\mathsf{sep}} \neg(B \wedge C) =_{\mathrm{df}} \mathsf{obliged}_A^{\mathsf{sep}} (\neg B \vee \neg C)$$

Figure 23 is a model of $\mathsf{obliged}_\top^{\mathsf{sep}}(a \wedge b)$.

Remark 6.10 (Temporal distinctions)**.** The sequences encoded in the tree structure of our models offer means to express temporal successions. There are various interpretations. For instance we could interpret the fact that our family stands at a node with context A by A being the case. Another, less strict reading is that at this point our family considers the fact that they are going to realize A as being settled (without it necessarily already being realized). So in the Chisholm example they may stand at node e_g which means that they decided to go (for instance they all agreed that they will go). In this case they may have next to consider whether they tell or whether they do not tell that they are going.

The second reading is more apt for Forrester's gentle murderer example.

- $\bigcirc \neg k$ "Smith ought not to kill Jones."

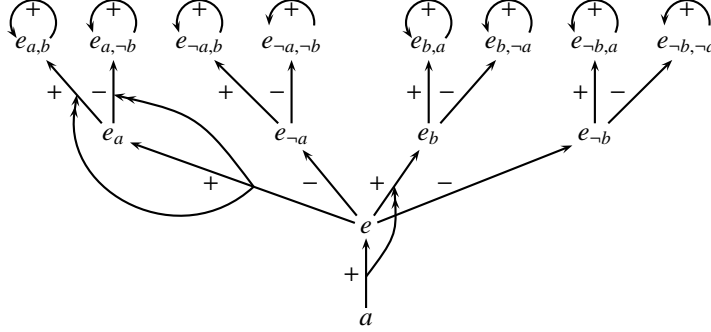- $k \to \bigcirc g$ "If Smith kills Jones he ought to do it gently."

Figure 23: Separable conjunctive norms: a model of $\mathsf{obliged}_\top(a \wedge b)$, $\mathsf{obliged}^{\mathsf{sep}}_\top(a \wedge b)$, and facts $a, \neg b$.

See Figure 24. Of course, if the killing already took place it is also already a matter of fact whether the killing was gentle or not. However, if we go with the second reading we only consider it as being settled whether Smith decided to kill Jones. So if Smith reaches node e_k it makes still sense to make a decision between $e_{k,g}$ and $e_{k,\neg g}$ and hence to consider the CTD obligation to kill gently.

Altogether, this gives us additional expressivity which can be exploited. For instance, we can express conditional obligations that are sensitive to the exact sequence of decisions taken by the agent in question:

$$
\begin{aligned}
\mathsf{obliged}_{(A_1,\ldots,A_m)} B =_{\mathrm{df}} & \left[\bigwedge_{i_1=1}^{n+1-m} \Box^{i_1} \left(\mathsf{context}(A_1) \wedge \bigwedge_{1<j\leq m} \neg\mathsf{context}(A_j) \to \right. \right. \\
& \left(\bigwedge_{i_2=1}^{n+1-m-(i_1-1)} \Box^{i_2} \left(\mathsf{context}(A_2) \wedge \bigwedge_{2<j\leq m} \neg\mathsf{context}(A_j) \to \right. \right. \\
& \left. \left(\ldots \left(\bigwedge_{i_m=1}^{n+1-m-(i_1-1)-\ldots-(i_{m-1}-1)} \Box^{i_m} (\mathsf{context}(A_m) \wedge \mathsf{switch}(B) \to \mathsf{opt}(B)) \right) \ldots \right) \right) \right) \right) \wedge \\
& \left[\bigwedge_{i_1=1}^{n+1-m} \Diamond^{i_1} \left(\mathsf{context}(A_1) \wedge \bigwedge_{1<j\leq m} \neg\mathsf{context}(A_j) \to \right. \right. \\
& \left(\bigwedge_{i_2=1}^{n+1-m-(i_1-1)} \left(\Diamond^{i_2} \mathsf{context}(A_2) \wedge \bigwedge_{2<j\leq m} \neg\mathsf{context}(A_j) \to \right. \right. \\
& \left. \left(\ldots \left(\bigwedge_{i_m=1}^{n+1-m-(i_1-1)-\ldots-(i_{m-1}-1)} \Diamond^{i_m} (\mathsf{context}(A_m) \wedge \mathsf{switch}(B)) \right) \ldots \right) \right) \right) \right]
\end{aligned}
$$

The first "[...]"-conjunct expresses that whenever our family traverses the sequence $A_1, \ldots, A_m$ and faces the choice between an arc leading to $\neg B$ and an arc that leading to a node that is compatible with B, the former is labeled '$-$' while the latter is labeled '$+$'. The second "[...]"-conjunct makes sure that the each model offers a path with the sequence $A_1, \ldots, A_m$ and a choice between $\neg B$ and a path that is consistent with B.

Of course, we can also give a temporal account of the facts. For instance our family may settle first A_1, then A_2, ..., then A_n. This can be expressed by

$$
\oslash^2 A_1 \wedge \ldots \wedge \oslash^{n+1} A_n
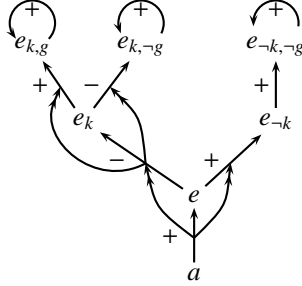$$

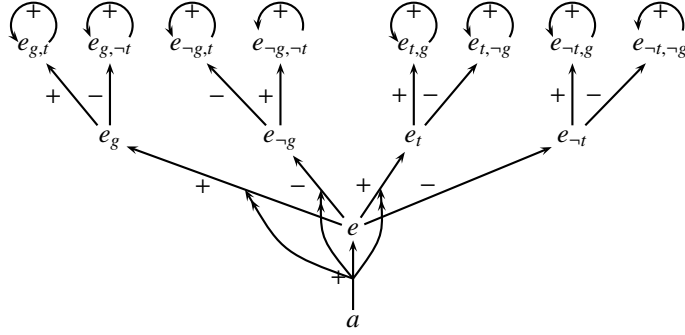Figure 24: The Forrester set with facts $k, \neg g$

Figure 25: Extended Chisholm with facts $\neg t, \neg g$ in the sequence $(\neg g, \neg t)$

Compare for instance Figures 16 and 25. Both models verify $\neg t$ and $\neg g$. However, in the former we have the sequence $(\neg t, \neg g)$ and hence $\oslash^2 \neg t \wedge \oslash^3 \neg g$ while in the latter we have the sequence $(\neg g, \neg t)$ and hence $\oslash^2 \neg g \wedge \oslash^3 \neg t$.

Example 6.11 (Extended Chisholm with specificity). We have already discussed why our translation cannot model specificity cases (see Remark 6.4).

Hence, in order to model specificity cases we need to alter our translation of conditional obligations. As the examples in Remark 6.4 indicate, the translation need to be more sensitive concerning the specificity of the given context. This can be achieved as follows. Where $\mathcal{E} \subseteq Q^l$ and

$$\mathsf{scontext}(\bigwedge \mathcal{E}) =_{\mathrm{df}} \mathsf{context}(\bigwedge \mathcal{E}) \wedge \bigwedge_{p \in Q^l \setminus \{q, \mathsf{co}(q) | q \in \mathcal{E}\}} \neg \mathsf{context}(\bigwedge \mathcal{E} \wedge p)$$

we translate $\bigcirc_{\bigwedge \mathcal{E}} A$ by

$$\mathsf{obliged}^{\mathsf{spe}}_{\bigwedge \mathcal{E}} A =_{\mathrm{df}} \bigwedge_{i=1}^{n} \square^i (\mathsf{scontext}(\bigwedge \mathcal{E}) \wedge \mathsf{switch}(A) \to \mathsf{opt}(A))$$

and define

$$\mathsf{obliged}^{\mathsf{spe}}_{A \vee B} C =_{\mathrm{df}} \mathsf{obliged}^{\mathsf{spe}}_{A} C \wedge \mathsf{obliged}^{\mathsf{spe}}_{B} C$$

$$\mathsf{obliged}^{\mathsf{spe}}_{\neg \neg A} B =_{\mathrm{df}} \mathsf{obliged}^{\mathsf{spe}}_{A} C$$

$$\mathsf{obliged}^{\mathsf{spe}}_{\neg (A \wedge B)} C =_{\mathrm{df}} \mathsf{obliged}^{\mathsf{spe}}_{\neg A \vee \neg B} C$$
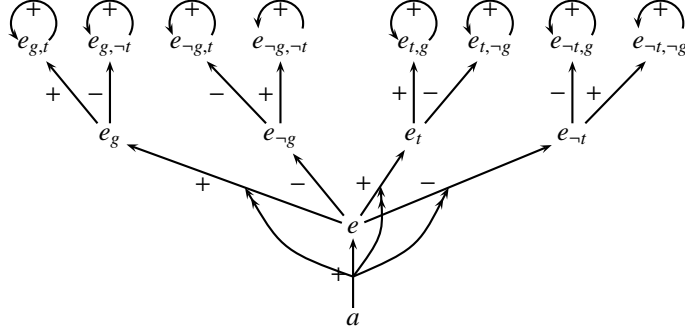
34

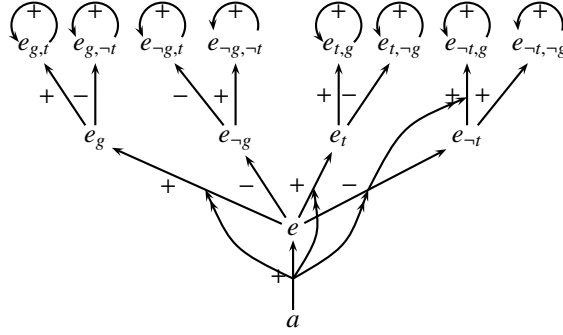Figure 26: Extended Chisholm with facts $\neg t$, $\neg g$ and $\bigcirc_{\neg t}\neg g$

Figure 27: Extended Chisholm with facts $\neg t$, $\neg g$ and $\neg \bigcirc_{\neg t}\neg g$, $\neg \bigcirc_{\neg t} g$

$$\text{obliged}^{\text{spe}}_{\neg(A\vee B)}C =_{\text{df}} \text{obliged}^{\text{spe}}_{\neg A \wedge \neg B}C$$

Given this translation, the model in Figure 26 is for the set

$$\{\text{obliged}^{\text{spe}}_{\top}g, \text{obliged}^{\text{spe}}_{\top}t, \text{obliged}^{\text{spe}}_{\neg g}\neg t, \text{obliged}^{\text{spe}}_{g}t, \text{obliged}^{\text{spe}}_{t}g, \text{obliged}^{\text{spe}}_{\neg t}\neg g, \neg t, \neg g\}$$

Note that this model does not verify $\text{obliged}_{\top}g$. The reason is that the arc $e_t \rightarrow e_{t,g}$ is '+'-labeled and the arc $e_t \rightarrow e_{t,\neg g}$ is '$-$'-labeled. Since we have $e_t \vDash \text{context}\top$, if $\text{obliged}_{\top}g$ were to hold we would have $e_t \vDash \text{opt}(g)$. However, due to the labels of the two arcs that emanate from e_t we have $e_t \nvDash \text{opt}(g)$. Note also that despite the labels of these arcs we have $\text{obliged}^{\text{spe}}_{\top}g$. The reason is that $e_t \nvDash \text{scontext}\top$. Indeed, the only node where $\text{scontext}\top$ holds is e. Since we have $e \vDash \text{opt}(g)$, we also have $\text{obliged}^{\text{spe}}_{\top}g$.

Another possibility is to alter the Chisholm set by adding $\neg \bigcirc_{\neg t} \neg g$ and $\neg \bigcirc_{\neg t} g$. This expresses that in case you don't tell, you are neither obliged to come nor are you obliged not to come. Neither option is preferable to the other. The model in Figure 27 is for the set:

$$\{\text{obliged}^{\text{spe}}_{\top}g, \text{obliged}^{\text{spe}}_{\top}t, \text{obliged}^{\text{spe}}_{\neg g}\neg t, \text{obliged}^{\text{spe}}_{g}t, \text{obliged}^{\text{spe}}_{t}g,$$
$$\neg\text{obliged}^{\text{spe}}_{\neg t}\neg g, \neg\text{obliged}^{\text{spe}}_{\neg t}g, \neg t, \neg g\}$$

Example 6.12 (Horty's Asparagus, Specificity). Another well-known case of specificity is Horty's example $\bigcirc\neg f$ and $\bigcirc_a f$ where f is "eating with fingers" and a is "eating asparagus". The model in Figure 28 is for the set

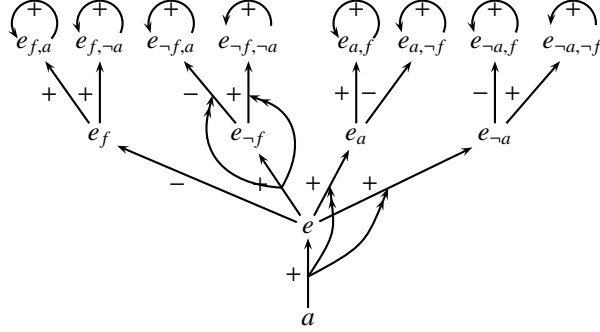$$\{\text{obliged}^{\text{spe}}\neg f, \text{obliged}^{\text{spe}}_{a}f, \neg f, a\}$$
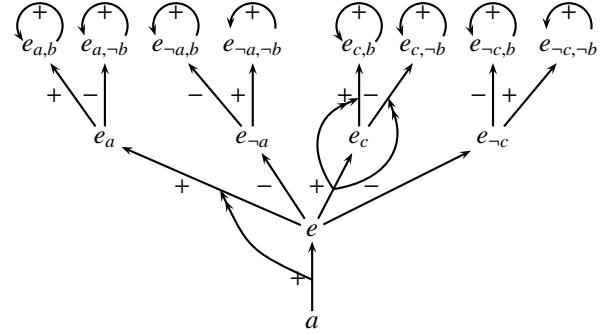
Figure 28: The asparagus example with facts $\neg f, a$.

Figure 29: Parallel CTDs with facts $c, \neg b$.

Example 6.13 (Parallel CTD sets). By means of obliged we can only model combinations of CTD sets that do not include any cases of specificity (see Remark 6.4). However, since obliged$^{\text{spe}}$ is able to tolerate specificity cases, we are now in the position to model combinations of any number of CTD sets. E.g., given

$$\Delta = \left\{ \left(\bigwedge_{k \leq i} x_k^j \right) \to \bigcirc x_{i+1}^j, \left(\bigwedge_{k \leq i} \text{co} x_k^j \right) \to \bigcirc \text{co} x_{i+1}^j \mid i \leq n(j) - 1, j \leq m \right\}$$

we only need to merge the semantic trees for each $\{(\bigwedge_{k \leq i} x_k^j) \to \bigcirc x_{i+1}^j, (\bigwedge_{k \leq i} \text{co} x_k^j) \to \bigcirc \text{co} x_{i+1}^j \mid i \leq n(j)\}$ so that they share the nodes a and e in order to get a model of Δ.

Figure 29 is a model of the two CTD sets $\Delta_1 = \{\bigcirc a, a \to \bigcirc b, \neg a \to \bigcirc \neg b\}$ and $\Delta_2 = \{\bigcirc c, c \to \bigcirc b, \neg c \to \bigcirc \neg b\}$. As noted in Remark 6.4 this example cannot be modeled by means of obliged since it includes a specificity case. The model in Figure 29 verifies

$$\{\text{obliged}^{\text{spe}} a, \text{obliged}_a^{\text{spe}} b, \text{obliged}_{\neg a}^{\text{spe}} \neg b\} \cup \{\text{obliged}^{\text{spe}} c, \text{obliged}_c^{\text{spe}} b, \text{obliged}_{\neg c}^{\text{spe}} \neg b\}$$

7 Comparison with the Jones–Pörn system DL[6]

It is illuminating to compare our system with the system **DL** presented in the classic paper [18] of Jones and Pörn in 1985.

They focus in their paper on the Chisholm paradox and offer on page 278 semantics of the form $(S, R_0, R_{0'}, h)$, where $R_0, R_{0'} \subseteq S^2$ are serial and

$$C1 : \quad R_0 \cap R_{0'} = \varnothing$$
$$C2 : \quad \{(x, x)|x \in S\} \subseteq R_0 \cup R_{0'}$$

The two relations correspond to modalities O and O'.

Let $\Box A =_{\mathrm{df}} \bigcirc A \wedge \bigcirc' A$. After some serious and illuminating discussion Jones and Pörn translate the Chisholm set into their system as follows (see page 284):

Let Ought $p =_{\mathrm{df}} \bigcirc p \wedge \neg \bigcirc' p$. Then the Chisholm set becomes (note that Ought p is also equivalent to $\bigcirc p \wedge \neg \Box p$, and thus the translation below does not use $\bigcirc'$).

1. Ought p

2. $\Box(p \rightarrow$ Ought $q)$

3. $\Box(\neg p \rightarrow$ Ought $\neg q)$

4. $\neg p$

Their proposal was criticised by Hanson [16] and replied to immediately in [17].

We will continue the criticism in the spirit of Hanson and show that the solution proposed by Jones and Pörn to the Chisholm paradox is not closed under substitution.

Remark 7.1 (Substitution in the Jones-Pörn system)**.** We show that the solution of the Chisholm paradox proposed in the Jones-Pörn system **DL** in [18], is not closed under substitution.

We observed that if we take the Chisholm set (see example 4.3) and substitute $H = \neg F$ and $T = \neg(F \wedge W)$, we get the fence paradox if we ignore the CTD linguistic formula $\neg H \rightarrow O(\neg T)$. Thus if the Jones-Pörn solution is consistent and independent it must solve the fence paradox under the above substitution. However, it does not. We get an inconsistent set under the above substitution.

The Jones-Pörn translation is into the modal logic with $R_\Box$ and R_0 as we have shown in the beginning of this section. By substituting the fence variables we get

1. Ought¬fence

2. $\Box(\neg \text{fence} \rightarrow \text{Ought}(\neg(\text{fence} \wedge \text{white})))$

3. $\Box(\text{fence} \rightarrow \text{Ought}(\text{fence} \wedge \text{white}))$

4. fence

Recall that Ought $X \equiv OX \wedge \neg \Box X$.
Let $F =$ fence, $W =$ white we get

---

[6]The perceptive reader might ask why are we comparing with the Jones and Pörn system, when there are many other systems to discuss, especially those which use preferential ordering, a device which can also be simulated by reactivity. The answer is simply pragmatic: Proper systematic comparison with other systems requires technical development and can be the subject of a subsequent paper. The comparison with Jones and Pörn is simple and quick and we already studied this system in [5].
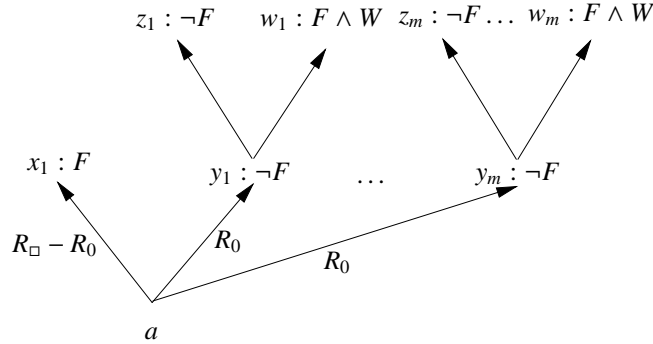
$$z_1 : \neg F \qquad w_1 : F \wedge W \quad z_m : \neg F \ldots \quad w_m : F \wedge W$$

$$x_1 : F \qquad\qquad y_1 : \neg F \qquad \ldots \qquad y_m : \neg F$$

$$R_\square - R_0 \qquad\qquad R_0 \qquad\qquad R_0$$

$$a$$

Figure 30

1. $O\neg F \wedge \Diamond F$

2. $\square(\neg F \rightarrow O(\neg(F \wedge W)) \wedge \Diamond(F \wedge W))$

3. $\square(F \rightarrow O(F \wedge W) \wedge \Diamond\neg(F \wedge W))$

4. F

Let us check whether the above is consistent.

Let $(S, R_\square, R_0, a, h)$ be a model of (1)–(4) above, with $R_0 \subseteq R_\square$ and $R_\square$ reflexive and R_0 a **KD** modality.

So under h assume (1)–(4) hold at a. We claim (1), (3) and (4) (the fence example) is already inconsistent.

Use Figure 30 for help.

From (1) we get that $a \vDash \Diamond F \wedge O\neg F$, hence for some x_1 such that $aR_\square x_1 \wedge \neg aR_0 x_1$ we have $x_1 \vDash F$. We say $\neg aR_0 x_1$ because $a \vDash O\neg F$. Let $y_1, \ldots, y_m$ in the figure be all R_0 accessible points to a. $m \geq 1$ because O is a **D** modality we have $y_i \vDash \neg F$.

From (4) we get $a \vDash F$. From (3) and reflexivity we get that $O(F \wedge W)$ must hold at a but this is not possible since $m \geq 1$ and $\neg F$ holds at y_1. Note that (2) can hold in the model if we add points z_i, w_i as indicated!

Note that Remark 7.1 could have been written in 1985. It has nothing to do with reactivity. It is a straightforward calculation in the same spirit as Hansson 1989, see [16]. For a serious analysis of the Jones and Pörn 1985 system using two-dimensional modal logic and other pre-1985 methods, see [5].

8 Conclusion and comparison

We would like to compare this paper with several other papers we wrote which make use of the reactive idea. We use the Chisholm paradox as a case study for comparison.

In Section 6 (see Remark 6.4) we have offered the following generic characterisation of combinations of various Chisholm sets of arbitrary depth:

$$\{x_i^j \rightarrow \bigcirc x_{i+1}^j, \mathrm{co}x_i^j \rightarrow \bigcirc \mathrm{co}x_{i+1}^j \mid i \leq n(j) - 1, j \leq m\}$$

The following are problems to be solved:

GP   Give a logic and reactive semantics for the general problem, for m and n arbitrary.

We have two options for the logic we use:

1. a traditional Hilbert System with Reactive Kripke Semantics for it, or

2. a most general Labelled Deductive System for General Reactive Kripke Semantics

We can break the general problem **GP** into several particular ones **PP**, which we may solve for the time being, using simpler models.

PP1 Solve **GP** for the case $m = 1$, and n arbitrary, i.e., a single arbitrary long chain of Chisholm like contrary to duties, only paying attention to the violation ordering while neglecting temporal considerations.

PP2 Solve **GP** for the case n and m arbitrary, i.e. for several Chisholm like sets, taking into account both the violation order (indicated by the sequence $1, \ldots, n(j)$ for each $j \leq m$) and temporal considerations.

PP3 Solve **GP** by giving a general Labelled Deductive system semantically, without giving any proof theory for it.

PP4 Add a labelled proof system to **PP3**.

PP5 Offer a solution with a new type of modelling, like argumentation or automata theory.

We now make our comments:

1. In the present paper we first offered a solution to **PP1** (see Section 5) by using a semantic approach. We offered a reactive model where the annotations to arcs mirror the obligations and where there are double arrows which can change the annotations. The latter represent the given facts. Using such semantics we can translate the traditional Chisholm set into a language with the two modalities $\Box$ and $\oslash$.

   In Section 6 we demonstrated how the present approach can be used to tackle **PP2**. We discussed combinations of CTD sets. We pointed out how interactions between various CTD sets can lead to complications such as specificity cases and how to deal with them in our semantics. We also demonstrated how **RSDL1** can be used to model temporal considerations.

2. We now compare with other papers using the reactive idea.

   (a) Paper [9] is a general paper, introducing for the first time the idea of reactivity.

   (b) Papers [10, 11] use a similar semantic approach to deal with the single Chisholm paradox with two letters (go, tell) where the temporal aspect is also taken into account. Part A ([10]) deals with the semantics, i.e. it solves **PP3**. Part B ([11]) deals with the proof theory, i.e. papers [10, 11] address **PP4** by means of the Labelled Deductive System approach. In fact, we can deal with several letters. Once we have semantics, we can deal with a family of general Chisholm sets, because we can evaluate them individually. The problem is that they may interact and so we need to process them first and this requires proof theory.

(c) Paper [14] uses an axiomatic Hilbert type approach, axiomatising the reactive semantics and is capable of translating any single Chisholm set with arbitrary number of letters ($n > 2$), including an arbitrary temporal order (different from the obligation/violation order).

We believe the axiomatic approach can handle several Chisholm sets (m, n arbitrary), but a complete temporal order for all sets should be given. So for example the case of $m = 1, n = 2$ would require a complete temporal order, say $x_1 < x_1 < z_1 < z_2$, and the resulting translation and modelling will be different from what we proposed in (1) above.

(d) Paper [5] is based on Jones and Pörn [18] and addresses the problem discussed in Remark 7.1 of the present paper, the solution of which requires a modified model. We use axiomatic two-dimensional temporal logic to model the Chisholm case ($m = 1, n = 2$). Strangely enough we do not model the temporal order. We do not think it is possible to do so in this approach. We use the obligation/violation progression (go < tell) as our "virtual time", and it is not clear how to account for real time as well.

We believe that combining the approach of the present paper with that of [14] will allow us to model the most general family of Chisholm sets (i.e. m, n aribtrary with arbitrary temporal order) without using the theory of Labelled Deductive Systems . This is probably possible by using multi-modal logic. However, this way we may not achieve an intuitive model but rather a purely technical one.

3. Completely new ideas tackling **PP5** can be found in papers [13, 2]. The direction in which they go is clear from their titles.

Acknowledgements

References

[1] L. Aqvist. A new approach to the logical theory of interrogatives, Part I. *Analysis*, 1965.

[2] J. Broersen, D. M. Gabbay, and L. van der Torre. Norms as automata and norm change as reactive automata. In preparation.

[3] J. Carmo and A. J. I. Jones. Deontic logic and contrary to duties. In D. M. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic, Vol 8*, pages 265–344. Springer, 2002.

[4] R. M. Chisholm. Contrary-to-duty imperatives and deontic logic. *Analysis*, (24), 1963.

[5] M. de Boer, D. M. Gabbay, X. Parent, and M. Slavkovic. Two-dimensional standard deontic logic. *Synthese*, pages 1–38, 2011. Published online: DOI: 10.1007/s11229-010-9866-4.

[6] J. W. Forrester. Gentle murder, or the adverbial samartian. *The Journal of Philosophy*, 81(4):193–197, 1984.

[7] D. M. Gabbay. *Labelled Deductive Systems*. OUP, 1996.

[8] D. M. Gabbay. Reactive Kripke models and contrary-to-duty obligations. In Ron van der Meyden and Leendert van der Torre, editors, *DEON 2008, Deontic Logic in computer science, LNAI 5076*, pages 155–173. Springer, 2008.

[9] D. M. Gabbay. Reactive Kripke semantics and arc accessibility. pages 292–341. Springer-Verlag, Berlin, 2008. Pillars of Computer Science: Essays Dedicated to Boris (Boaz) Trakhtenbrot on the Occasion of His 85th Birthday, Lecture Notes in Computer Science, vol. 4800.

[10] D. M. Gabbay. Reactive Kripke models and contrary to duty obligations. Part A: Semantics. 2009. To appear in *Journal of Applied Logic*.

[11] D. M. Gabbay. Reactive Kripke models and contrary to duty obligations. Part B: Proof theory. 2009. To appear in *Journal of Applied Logic*.

[12] D. M. Gabbay. Completeness theorems for reactive modal logics. 2010. Paper 392, July 2010. To appear in special issue of AMAI 2012.

[13] D. M. Gabbay. Bipolar argumentation frames and contrary to duty obligations, a position paper. In M. Fisher et al., editor, *Proceedings of CLIMA 2012, LNAI 7486*, pages 1–24. Springer-Verlag Berlin Heidelberg, 2012.

[14] D. M. Gabbay. Temporal deontic logic for the Generalised Chisholm set of contrary to duty obligations. In *DEON 2012, LNAI 7393*, pages 91–107. Springer, Heidelberg, 2012.

[15] D. M. Gabbay and K. Schlechta. Critical analysis of the Carmo–Jones model of contrary to duty obligations. 2009. Draft paper 358, `http://arxiv.org/PS_cache/arxiv/pdf/1002/1002.3021v1.pdf`.

[16] S. O. Hansson. A note on the deontic system DL of Jones and Pörn. *Synthese*, (80):427 –428, 1989.

[17] A. I. J. Jones and I. Pörn. A rejoinder to Hansson. *Synthese*, (80):429–432, 1989.

[18] A. J. I. Jones and I. Pörn. Ideality, sub-ideality and deontic logic. *Synthese*, 1985.

[19] B. Loewer and M. Belzer. Dyadic deontic detachment. *Synthese*, (54), 1983.

[20] H. Prakken and M. Sergot. Contrary to duty obligations. *Studia Logica*, pages 91–115, 1996.

[21] H. Prakken and M. Sergot. Dyadic deontic logic and contrary to duty obligations. In D. Nute, editor, *Defeasible Deontic Logic, Synthese Library, Vol 263*, pages 223–262. Kluwer, 1997.

[22] A. N. Prior. Escapism: The logical basis of ethics. In A. I. Melden, editor, *Essays in Moral Philosophy*, pages 135–146. University of Washington Press, 1958.

[23] A. Ross. Imperatives and logic. *Theoria*, (7):53–71, 1941.

[24] C. Straßer. A deontic logic framework allowing for factual detachment. *Journal of Applied Logic*, 9(1):61–80, 2010.

[25] L. van der Torre and Y.-H. Tan. Contrary to duty reasoning with preference-based dyadic obligations. *Annals of Mathematics and Artificial Intelligence*, (27):49–78, 1999.