Extensionality vs Intensionality: A Perspectival Account of Conditional Ought with Definite Descriptions

Xavier Parent* and Dominik Pichler†
TU Wien
{xavier,dominik}@logic.at

Abstract

The theme of extensionality in first-order deontic logic has been thoroughly studied in the past, but not in the context of a combination of different types of modalities. An operator is extensional if it allows substitution *salva veritate* of co-referential terms within its scope and intensional if it does not. It can be argued that one distinctive feature of "ought" (as opposed to the other modalities) is that it is extensional. The question naturally arises as to whether it is possible to combine extensionality and intensionality of different modal operators in the same semantics without creating the deontic collapse. We answer this question within a particular framework, Åqvist's system **F** for conditional obligation. We develop in full detail a perspectival account of obligation (and related notions), as was done for Standard Deontic Logic (SDL) by Goble. It is called "perspectival", because one always evaluates the content of an obligation in one world from the perspective of another one. This requires using some form of cross-world evaluation to handle non-rigid terms like definite descriptions. The proposed framework allows for a more nuanced way of approaching first-order deontic principles.

Keywords: First-order reasoning, extensionality, conditional obligation, 2-dimensional semantics, preferences, perspectivism, definite description

The present paper supersedes [33], which was based on [32] and was presented at DEON 2023. We thank Lou Goble and Paul McNamara for their valuable feedback. We also thank two anonymous reviewers for their valuable feedback.

*Supported by the Austrian Science Fund (FWF) through the ANCoR project [M-3240-N, doi: 10.55776/I2982] and the LoDEX project [doi: 10.55776/I6372].

†Supported in part by the Austrian Science Fund (FWF) through the LoDEX project [doi: 10.55776/I6372].

1 Introduction

The past 15 years have seen a renewed interest in so-called relativism or perspectivism in the philosophy of language. Relativist or perspectivist accounts have been put forth to explain discourse about knowledge, epistemic possibility, matters of taste, contingent future events, modalities (including the deontic ones) and the like. Here relativism is usually taken to be, or to presuppose, a semantic thesis. Understanding how some discourses function requires recognizing that speakers express propositions whose truth or falsity are relative to parameters or perspectives in addition to a possible world—see Kölbel [25] for a thorough defense of this view, and also MacFarlane [27]. The approach is often called "perspectivism" as it has a less negative connotation than "relativism", and we will stick to this term.

The purpose of the present paper is to show some of the usefulness of this view for normative reasoning. We believe it may shed light on a topic that has been overlooked in the recent papers devoted to first-order deontic reasoning, e.g. [10, 11, 37]. This is the topic of extensionality of "ought". We do not claim to be original, as we will pick up on a proposal made long ago by Goble [15, 16, 17]. It can be summarized thus. An operator is extensional if it allows substitution *salva veritate* of co-referential terms within its scope, and intensional if it does not. It can be argued that one distinctive feature of "ought" (as opposed to the other modalities) is that it is extensional. The problem is: a deontic logic in which "ought" is extensional can be shown to collapse to triviality. Goble developed his own solution to this problem, and we will refer to it as the original "perspectival" account. The basic idea is that the content of an obligation at one world is to be evaluated from the perspective of another one, so that some form of cross-world evaluation is made possible. This idea of cross-world evaluation is familiar from the literature on multi-dimensional modal logic (see e.g. [3, 14, 22, 38]). Other works in multi-dimensional deontic logic we are aware of focus on the propositional case [7, 12, 13, 21]. The novelty lies in linking the perspectival idea to first-order (FO) considerations.

Our goal is to improve the original account in two ways. By doing so, we hope to strengthen the case for the perspectival idea, and provide more credibility to it.

- The original account is cast within the framework of Standard Deontic Logic (SDL) [41], which is known to be plagued by the deontic paradoxes, in particular the paradox of contrary-to-duty (CTD) obligation [4]. We will recast the account within the framework of preference-based dyadic deontic logic [1, 6, 18, 20, 30]. Dyadic deontic logic is the logic for reasoning with dyadic obligations "it ought to be the case that ψ if it is the case that φ" (notation: $\bigcirc(\psi/\varphi)$). Its semantics is in terms of a betterness relation. Initially devised to resolve the CTD paradox, dyadic deontic logic is a recognized standard for normative reasoning. The idea of making it two-dimensional is not entirely new: Lewis [26, p. 63] suggested to analyze conditionals within the framework of two-dimensional modal logic, but his motivations were different.

- The original account does not allow for different types of modalities to interact. We will lift this restriction, and look at the question of whether it is possible to combine extensionality and intensionality of different modal operators in the same semantics without creating the collapse. Åqvist's mixed alethic-deontic systems **E**, **F** and **G** [1, 30, 31] are obvious candidates for this study. Their language includes an additional modal operator, $\Box$ ("It is settled that"), enabling the capture of fundamental principles of normative reasoning, such as "strong factual detachment". Factual detachment, as referred to by Decew [8], is the principle that allows one to infer $\bigcirc\psi$ from $\bigcirc(\psi/\varphi)$ and the mere truth of φ. Van Eck emphasized the significance of factual detachment for normative reasoning by asking:

  > "How can we take a conditional obligation seriously if it cannot, by way of detachment, lead to an unconditional obligation?" [9, p. 263]

  First discussed by Greenspan [19], strong factual detachment (the name is Prakken and Sergot [34]'s) requires that $\Box\varphi$ holds, rather than just φ. There is a widespread agreement among deontic logicians that, for CTD obligations, strong factual detachment is more appropriate than factual detachment. Consider a primary obligation of the form $\bigcirc\neg\varphi$ and its associated CTD obligation $\bigcirc(\psi/\varphi)$. Example: a person ought to breast-feed her baby, and if she does not, she ought to use instant formula. Prakken and Sergot write:

  > "It is only if the violation of the primary obligation $\bigcirc\neg\varphi$ is unavoidable if $\Box\varphi$ holds [she *cannot* breast-feed], that the [associated] CTD obligation comes into full effect, and [is detached]" [34, §5.1][1]

Of the three systems mentioned above, we choose to focus on **F**, because it is the weakest one in which the collapse arises. The first-order extension of **F** will be called $\mathbf{F}^{\forall}$. One could object that, in **F**, $\Box$ is a *soi disant* modality, definable in terms of $\bigcirc(-/-)$. For that reason, $\mathbf{F}^{\forall}$ will contain two alethic modal operators $\boxdot$ and $\boxtimes$. The former will be definable in terms of $\bigcirc(-/-)$, but not the latter.

The paper is organized as follows. Sec. 2 sets the stage, and defines a list of basic requirements to be met by the logic. Sec. 3 develops in full semantic detail the perspectival account of obligation (and related notions) alluded to above. Sec. 4 shows how the requirements are met. Sec. 5 concludes.

[1] For further discussion, we refer the reader to [28, 34]

2 Setting the stage

We give a list of basic requirements that we think an adequate first-order (FO) deontic logic should meet. The problem dealt with in this paper will be to devise a framework meeting them. For ease of readability, we formulate the requirements within the language of a monadic deontic logic. Our list is not meant to be exhaustive.

2.1 Requirements

Requirement 1 (Extensionality for "ought")**.** *$\bigcirc$ ("It ought to be the case that ...") should validate the principle of substitution salva veritate (E-$\bigcirc$), where φ is a formula, t and s are terms, and $\varphi_{t\hookrightarrow s}$ is the result of replacing zero up to all occurrences of t, in φ, by s:*

$$t = s \rightarrow (\bigcirc\varphi \leftrightarrow \bigcirc\varphi_{t\hookrightarrow s}) \tag{E-$\bigcirc$}$$

Intuitively: two co-referential terms may be interchanged without altering the truth-value of the deontic formula in which they occur.

A modal operator is usually said to be referentially transparent, when it satisfies the principle of substitution *salva veritate*, and referentially opaque otherwise. As pointed out by Castañeda [2] there are good reasons to believe that deontic operators are referentially transparent. For instance, the inference from (1-a) and (1-b) to (1-c) is intuitively valid:[2]

(1) a. It ought to be that the Pope blesses the pregnant woman
b. Jose is the Pope
c. It ought to be that Jose blesses the pregnant woman

Formally:

$$\begin{array}{l} \bigcirc\, B(\imath x P(x), \imath y(W(y) \wedge Pr(y))) \\ j = \imath x P(x) \\ \bigcirc\, B(j, \imath y(W(y) \wedge Pr(y))) \end{array}$$

$\imath x P(x)$ is a so-called definite description, and is read "the x that is P" ("the Pope"). Definite descriptions are used to refer to what a speaker wishes to talk about. It is hard to find counter-examples to the principle of substitution *salva veritate* in the

[2] Our original example in [33] was misleading. It used "The Pope ought to live a life of exceptional sanctity" as a first premise. This is a generic statement about Popes, and not a singular statement.

deontic domain. Castañeda (rightly) says: "a man's obligations are *his* [the author's emphasis] regardless of his characterizations". In other words, they are independent of the way he is referred to. In daily conversations, one casually switches between a proper name and a definite description, or between different definite descriptions (the Pope, the direct successor of St Peter, ...). When using one instead of the other, we are still talking about the same individual. This would just not be possible if "ought" was not referentially transparent.

The above point applies to the bearer of an obligation, but also to the party to whom the obligation is owed. In other words, it applies to anyone affected by the consequences of the obligation, whether those consequences are positive or negative. Consider:

(2) a. It ought to be that the Pope blesses the pregnant woman
  b. Marie is the pregnant woman
  c. It ought to be that the Pope blesses Marie

Intuitively, (2-c) is derived from (2-a) and (2-b) in a similar manner to how (1-c) is derived from (1-a) and (1-b).

We adopt a "directly referential" approach to definite descriptions that is similar to Kaplan's concept of "dthat" [23]. Thus, the meaning of a definite description lies in what it points out in the world. This "directly referential" take allows us to put aside putative counterexamples like this one:[3]

(3) a. Jose is the (actual) Pope: $j = \imath x P(x)$
  b. It ought to be that Joey is the Pope: $\bigcirc(j' = \imath x P(x))$
  c. It ought to be that Joey is Jose: $\bigcirc(j' = j)$.

We add "actual" between brackets, a more accurate reading of our "$\imath x \varphi(x)$" being "the x: actually $\varphi(x)$". (3-c) follows from (3-a) and (3-b), by substitution *salva veritate*. Suppose (3-a) is true. Suppose also that Jose rigged his own election as a Pope and that Joey is, in fact, the Cardinal who got the most votes. (3-b) is, then, true. But intuitively (3-c) is false. The move to (3-c) is not warranted, not because the principle of substitution fails, but because it rests on an equivocation on "the Pope", the use of which in (3-b) is not directly referential, but descriptive. The meaning of "the Pope" in the sentence (3-b) is not the individual it points at in the actual world, namely Jose.

Our interest is really in definite descriptions, and not proper names. Following Kripke [24], a proper name is often taken to be a rigid designator, and assumed to refer to the same individual in all possible worlds in which that individual exists.

[3] We owe this point (and this example) to Paul McNamara.

Obviously, (E-$\bigcirc$) holds if t and s are rigid designators. We build on the insights of Kaplan and others, who, while accepting Kripke's main argument about proper names, observed that certain uses of definite descriptions appear to be directly referential. In these cases (E-$\bigcirc$) also applies. Thus, the question becomes: how can we account for the validity of (E-$\bigcirc$), when one of s and t (maybe both) is a definite description used in this manner? In this study, we do not assume the rigidity of proper names; however, introducing this assumption would not affect our arguments.

Requirement 2 (Intensionality for "necessarily")**.** *$\Box$ ("It is necessary that ...") should not validate the principle of substitution salva veritate, where t and s are terms (either a constant or a definite description):*[4]

$$t = s \rightarrow (\Box\varphi \leftrightarrow \Box\varphi_{t\hookrightarrow s}) \tag{E-$\Box$}$$

This requirement is best motivated using the following well-known example:

(4) a. Number of planets $= 8$
 b. $\Box(8 = 8)$
 c. $\Box$(Number of planets $= 8$)

Consider the Aristotelian/Megarian tensed interpretation of $\Box$, which takes the past as well as the future into consideration [36, p. 125]. Under this interpretation, $\Box\varphi$ is taken as a shorthand for $\varphi \wedge H\varphi \wedge G\varphi$, where H and G mean "always in the past" and "always in the future", respectively. This interpretation of $\Box$ is plausible. Under such an interpretation, the move from (4-a) and (4-b) to (4-c) is not warranted.[5]

We believe this requirement makes sense for the most common (non-deontic) interpretations of $\Box$. However, we reckon there are also less common (non-deontic) readings of $\Box$ for which this requirement does not apply. For example, it does not apply to "historical necessity" [40]. Noticeably, $\Box\varphi$ is equivalent with φ, if φ does not contain the modality of the future (see law (6) in [40]), in which case the principle of *salva veritate* holds trivially.

Requirement 3 (No collapse)**.** *The logic should avoid the deontic collapse. That is, the formula $\varphi \leftrightarrow \bigcirc\varphi$ should not be derivable.*

A separate section is devoted to this requirement, taken from Goble [15, 16, 17].

The *raison d'être* of our last requirement is this: obligations are there to make the world a better place; they are constantly violated, but should not be so. Therefore,

[4]Quine argues for this requirement in his [35].

[5]The number of planets happened to be 9. In 2006, the International Astronomical Union (IAU) redefined the criteria for counting as a planet, resulting in Pluto losing its status as the ninth planet. See `https://science.nasa.gov/dwarf-planets/pluto/`.

our account should make the notion of definite description well-behaved with respect to negation. That is to say:

Requirement 4 (Self-negation)**.** *The logic should be able to account for the meaningfulness of a deontic statement denying a property of an individual identified using that very same property.*

Consider the following instantiation of the principle of substitution *salva veritate*:

(5) a. Jose is the (actual) Pope: $(j = \imath x P(x))$
b. It ought to be that Jose is not Pope: $\bigcirc\neg P(j)$
c. It ought to be that the (actual) Pope is not Pope: $\bigcirc\neg P(\imath x P(x))$

(5-b) is true, if the election was rigged. (5-c) makes perfect sense. Self-negation like the one in (5-c) cannot be accounted for in (a straightforward FO extension of) SDL. The reason is that $\bigcirc\neg P(\imath x P(x))$ is not satisfiable, assuming there exists a Pope in the best of all possible worlds. Such an obligation tells us that in the best of all possible worlds, the Pope x is not Pope. But this is a contradiction (assuming that such an x exists). Of course, the claim is not that in the best of all possible worlds there is an x that is Pope and is not Pope. Rather−to anticipate our solution−the claim is that the individual x who is Pope in the actual world (viz. Jose) is not Pope in any of the best worlds. This is a relation among objects in possible worlds that cannot be captured in the standard possible world semantics. The semantic analysis of (5-c) calls for a "cross-world" mode of evaluation.

We emphasize that the use of an "actually" operator in discussions concerning the *a priori* has been motivated by very similar considerations. Here is the kind of example commonly discussed (see, e.g., [5, p. 350]):

(6) It might have been that everyone actually happy was sad

As observed by Hughes and Cresswell, (6) cannot be formalized as

$$\Diamond\forall x(Hx \to Sx)$$

They write: "that envisages a possible world in which all happy are sad, and this can only be so if no one at all is happy" [5, ibidem]. For this reason, it has been suggested to translate (6) as

$$\Diamond\forall x(\mathcal{A}Hx \to Sx) \qquad (\#)$$

where $\mathcal{A}$ is a modal operator read as "actually", whose semantics is defined in terms of "truth on the diagonal". Intuitively, this sentence holds in world w if there is an accessible world v such that everybody who is happy in w is sad in v. There are similarities between the two approaches. A more thorough comparison should be postponed to another occasion.

One could object that (5-c) is better formalized as

$$\exists x(P(x) \wedge \bigcirc\neg P(x)) \tag{\$}$$

(\$) is unproblematic. But note that as a spin-off of the extensionality of the deontic operator the principles of universal instantiation (UI) and existential generalisation (EG) hold unrestrictedly, that is, even if t is within the scope of $\bigcirc$:[6]

$$\exists x(x = t) \rightarrow (\forall x\varphi(x) \rightarrow \varphi(t)) \tag{UI}$$

$$\exists x(x = t) \rightarrow (\varphi(t) \rightarrow \exists x\varphi(x)) \tag{EG}$$

Given the assumption $\exists x(x = \imath y P(y))$, (\$) and (5-c) are inter-derivable. Thus the principle of extensionality turns an apparently unproblematic formula, $(\exists x(P(x) \wedge \bigcirc\neg P(x))$, into a problematic one, $(\bigcirc\neg P(\imath x P(x)))$. Our task is to account of the meaningfulness of the latter formula.

We show below the inter-derivability of (\$) and (5-c). $\exists! x\varphi$ is a shorthand for $\exists x\forall y(\varphi \leftrightarrow y = x)$.

(a) $\exists x(x = \imath y P(y))$	(Hypothesis)
(b) $\exists x(P(x) \wedge \bigcirc\neg P(x))$	(Hypothesis)
(c) $\exists! x P(x)$	(a)
(d) $\exists! x(P(x) \wedge \bigcirc\neg P(x))$	(FO + b + c)
(e) $\forall x(P(x) \rightarrow \bigcirc\neg P(x))$	(FO + d)
(f) $P(\imath y P(y)) \rightarrow \bigcirc\neg P(\imath y P(y))$	(e + UI)
(g) $P(\imath y P(y))$	(a)
(h) $\bigcirc\neg P(\imath y P(y))$	(f + g)

Derivation 1

(a) $\exists x(x = \imath y P(y))$	(Hypothesis)
(b) $\bigcirc\neg P(\imath y P(y))$	(Hypothesis)
(c) $P(\imath y P(y))$	(a)
(d) $P(\imath y P(y)) \wedge \bigcirc\neg P(\imath y P(y))$	(b + c)
(e) $\exists x(P(x) \wedge \bigcirc\neg P(x))$	(d + EG)

Derivation 2

[6]This is in general not the case in first-order modal logic.

2.2 Collapse

We explain in more detail how the collapse mentioned in requirement 3 arises. The discussion draws on Goble [15, 16, 17]. We say the deontic collapse arises in a logic if the formula $\varphi \leftrightarrow \bigcirc\varphi$ is derivable in this logic. This means that everything that is true is obligatory and vice versa. Goble pointed out that, if the principle of substitution *salva veritate* holds, then the deontic collapse follows. We reiterate and amplify his main points.

The derivation of $\bigcirc\varphi \to \varphi$ appeals to the derivation of $\varphi \to \bigcirc\varphi$, so we begin with the latter. As originally given by Goble, derivation 3 invokes the law of contraposition, the law of double $\neg$-elimination, and the **D** axiom for $\bigcirc$:

(a) $\bigcirc\varphi$	(Hypothesis)
(b) $\neg\bigcirc\neg\varphi$	(**D** axiom)
(c) $\neg\neg\varphi$	($\varphi \to \bigcirc\varphi$ and contraposition)
(d) φ	(Double $\neg$-elimination)

Derivation 3

One may be tempted to block this derivation by just abandoning the principle of contraposition or the principle of double $\neg$-elimination. However, this would not block the derivation of $\varphi \to \bigcirc\varphi$, which is already counter-intuitive. We turn to this implication. We do not give the original argument,[7] but a variant one, which highlights the role of $\Box$.

Proposition 1. *Consider a deontic logic containing (i) the usual principles of FO logic, (ii) the principle of substitution* salva veritate *for "ought" (E-$\bigcirc$), $t = s \to (\bigcirc\varphi \leftrightarrow \bigcirc\varphi_{t\hookrightarrow s})$, (iii) the principle $\Box\varphi \to \bigcirc\varphi$ ($\Box 2\bigcirc$) and (iv) the principle of inheritance "If $\vdash \varphi \to \psi$ then $\vdash \bigcirc\varphi \to \bigcirc\psi$" (In). Then $\varphi \to \bigcirc\varphi$ is derivable from $\Box\exists y(y = t)$.*

Proof. In this derivation we assume that x and y do not occur free in φ:

[7] Goble's derivation can be found in [15].

(a) φ	(Hypothesis)
(b) $\Box\exists y(y = t)$	(Hypothesis)
(c) $t = \imath x(x = t \wedge \varphi)$	(FO + a)
(d) $\bigcirc\exists y(y = t)$	($\Box 2 \bigcirc$ + b)
(e) $\bigcirc\exists y(y = \imath x(x = t \wedge \varphi))$	(E-$\bigcirc$ + c + d)
(f) $\bigcirc\varphi$	(In + e)

Derivation 4

$\Box$

Some comments are in order:

- We show $\varphi \to \bigcirc\varphi$, where the original argument shows $\bigcirc\psi \to (\varphi \to \bigcirc\varphi)$.
- Our derivation starts from the supposition $\Box\exists y(y = t)$. This may be read as t necessarily denotes. We take this supposition to be harmless. We do not want the collapse even under this assumption.
- Line (c) "drags" φ inside the scope of the definite description. (c) is read: t is the-unique-x-identical-with-t-and-φ. Line (f) "drags" φ outside the scope of the definite description. Both moves are allowed in FO logic.
- The principle (E-$\bigcirc$) is applied on line (e); t is replaced by the co-referential term "the-unique-x-identical-with-t-and-φ". The intuitive meaning of (e) may seem difficult to grasp. As explained in Sec. 4.3, the two-dimensional semantics presented in this paper can be of help.
- Line (f) is obtained by applying (In). This final move will be discussed further in a moment (see derivation 5).

To avoid the deontic collapse, the following ways out suggest themselves:

Option 1: revise the laws of first-order logic;

Option 2: abandon ($\Box 2 \bigcirc$);

Option 3: abandon (In), or restrict its application.

We will go with option 3. Thus, in derivation 4, the move from (e) to (f) will be blocked. A good reason for choosing this path is that option 2 alone would not block the original derivation of the collapse in a mono-modal setting, which uses (In) and the laws of FO logic. Note that, in Åqvist's system **F**, (In) is not a primitive rule, but is derivable from ($\Box 2 \bigcirc$) and two extra principles:

- the principle of necessitation for $\Box$: "If $\vdash \varphi$, then $\vdash \Box\varphi$" (N-$\Box$)
- the K axiom for $\bigcirc$: $\bigcirc(\varphi \to \psi) \to (\bigcirc\varphi \to \bigcirc\psi)$ (K-$\bigcirc$)

In a system in which (In) is not primitive, like Åqvist's one, the move from (e) to (f) is explained thus:

$$
\begin{array}{ll}
\text{(a)} \vdash \exists y(y = \imath x(x = t \wedge \varphi)) \to \varphi & \text{(FO)} \\
\text{(b)} \vdash \Box[\exists y(y = \imath x(x = t \wedge \varphi)) \to \varphi] & \text{(N-}\Box\text{)} \\
\text{(c)} \vdash \bigcirc[\exists y(y = \imath x(x = t \wedge \varphi)) \to \varphi] & (\Box 2 \bigcirc) \\
\text{(d)} \vdash \bigcirc\exists y(y = \imath x(x = t \wedge \varphi)) \to \bigcirc\varphi & \text{(K-}\bigcirc\text{)}
\end{array}
$$

Derivation 5

Ultimately, the solution will consist in restricting the application of (N-$\Box$) so as to block the move from (a) to (b). However, the final effect will be the same: (In) will go away in its plain form.[8]

Proposition 2 indicates that the extensionality of $\Box$ can lead to a collapse, regardless of the stance on whether $\bigcirc$ is extensional.[9]

Proposition 2. *Consider the same deontic logic as in Prop. 1, but with (E-$\bigcirc$) replaced with (E-$\Box$). In such a logic, $\varphi \to \bigcirc\varphi$ is derivable from $\Box\exists y(y = t)$.*

Proof. As before we assume that x and y do not occur free in φ:

$$
\begin{array}{ll}
\text{(a) } \varphi & \text{(Hypothesis)} \\
\text{(b) } \Box\exists y(y = t) & \text{(Hypothesis)} \\
\text{(c) } t = \imath x(x = t \wedge \varphi) & \text{(FO + a)} \\
\text{(d) } \Box\exists y(y = \imath x(x = t \wedge \varphi)) & \text{(E-}\Box\text{ + b + c)} \\
\text{(e) } \bigcirc\exists y(y = \imath x(x = t \wedge \varphi)) & (\Box 2 \bigcirc) \\
\text{(f) } \bigcirc\varphi & \text{(In)}
\end{array}
$$

Derivation 6

$\Box$

3 The perspectival account

In this section, we develop in full detail our perspectival account. The basic idea is that the content of an obligation at one world is to be evaluated from the perspective of another one. What we mean by this is the following. Formulas will

[8] This is an adaptation of Goble's solution to our bi-modal setting. Goble uses an axiomatization of SDL in which (In) is primitive. One could have used instead an axiomatization of SDL in which the rule of necessitation for $\bigcirc$ is primitive, and (In) is derivable from it. Whatever axiomatization is chosen, the effect is the same: both rules hold in a restricted form.

[9] This observation is new to the literature. Again, we will block the last step of the derivation.

be evaluated with respect to two dimensions, or pair of worlds (v, w). World v is where the evaluation takes place, and world w is the one from the perspective of which formulas are evaluated (call it the reference or actual world, if you wish). Throughout the paper the reference world will be represented as an upper index in the notation $v \models^w$. What is meant by "φ is evaluated in v from w's perspective" is this: when determining the truth-value of φ in v, the terms occurring in φ get the same denotation as in w.

To keep the logic as close as possible to the original **F**, we use two alethic modal operators, $\boxdot$ and $\boxtimes$. The first is extensional, and the second intensional. $\boxdot$ is definable in terms of $\bigcirc$ (see Appendix B), and is thus dispensable.

Definition 1. *The language $\mathcal{L}$ contains:*

- *A countable set of variables $V := \{x, y, z, ...\}$*
- *A countable set of constants $C := \{c, d, e, ...\}$*
- *Two propositional connectives $\wedge, \neg$*
- *Three first-order logic symbols $\forall, \imath, =$*
- *A binary obligation operator $\bigcirc(-/-)$*
- *Two unary alethic operators $\boxdot$ and $\boxtimes$*
- *For each $n \in \mathbb{Z}^+$ a countable set of n-place predicate symbols $\mathbb{P}^n := \{A^n, B^n, ...\}$, we define $\mathbb{P} := \bigcup_{n\in\mathbb{N}} \mathbb{P}^n$*

We can now define inductively the well-formed terms and formulas used in our logic and their respective complexity ($\ulcorner ... \urcorner$).

Definition 2 (Terms and formulas)**.**

- ***Terms:***
  - *Every element of $V \cup C$ is a term of complexity* 0
  - *If φ is a formula and $x \in V$ then $\imath x\varphi$ is a term with $\ulcorner \imath x\varphi \urcorner := \ulcorner \varphi \urcorner + 1$*
- ***Formulas:***
  - *If $R^n \in \mathbb{P}$ is a n-place predicate symbol and $t_1, ..., t_n$ are terms then $R^n(t_1, ..., t_n)$ is a formula with $\ulcorner R^n(t_1, ..., t_n) \urcorner := \sum_{i=1}^{n} \ulcorner t_i \urcorner$*
  - *If φ is a formula and $x \in V$ then $\forall x\varphi$ is a formula with $\ulcorner \forall x\varphi \urcorner := \ulcorner \varphi \urcorner + 1$*
  - *If t_1 and t_2 are terms then $t_1 = t_2$ is a formula with $\ulcorner t_1 = t_2 \urcorner := \ulcorner t_1 \urcorner + \ulcorner t_2 \urcorner + 1$*
  - *If φ is a formula then $\neg\varphi$ is a formula with $\ulcorner \neg\varphi \urcorner := \ulcorner \varphi \urcorner + 1$*
  - *If φ is a formula then $\boxdot\varphi$ is a formula with $\ulcorner \boxdot\varphi \urcorner := \ulcorner \varphi \urcorner + 1$*
  - *If φ is a formula then $\boxtimes\varphi$ is a formula with $\ulcorner \boxtimes\varphi \urcorner := \ulcorner \varphi \urcorner + 1$*

– *If φ and ψ are formulas then $\varphi \wedge \psi$ is a formula with* $\ulcorner \varphi \wedge \psi \urcorner := \ulcorner \varphi \urcorner + \ulcorner \psi \urcorner + 1$
– *If φ and ψ are formulas then $\bigcirc(\psi/\varphi)$ is a formula with* $\ulcorner \bigcirc(\psi/\varphi) \urcorner := \ulcorner \varphi \urcorner + \ulcorner \psi \urcorner + 1$
– *Nothing else is a formula*

Definition 3 (Derived connectives)**.** *Let t be a term. We define $E(t)$ as $\exists x(x = t)$, where x is the first element of V not appearing in t. The symbols $\vee, \bot, \top, \rightarrow, \leftrightarrow, \Diamond\varphi, \circledDiamond\varphi, P(./.), \exists, \exists!$ and $\neq$ are introduced the usual way.*

Definition 4 (Frames)**.** *$\mathcal{F} = \langle W, \succeq, D \rangle$ is called a frame, where*

- *$W \neq \emptyset$ is a set of worlds.*
- *$\succeq \subseteq W \times W$ is a binary relation called the betterness relation. When $w \succeq v$, we say that world w is at least as good as world v.*
- *D is a function which maps every world $w \in W$ to a non-empty set D_w.*

D is called the domain function, and D_w is called the domain of w.
$\mathbb{D} := \bigcup_{w \in W} D_w$ is called the "actual" domain and $\mathbb{D}^+ := \mathbb{D} \cup \{\mathbb{D}\}$ the (whole) domain.

The individual domains $(D_w)_{w \in W}$ contain all objects which are within the range of the universal quantifier at a world w. The actual domain $\mathbb{D}$ is not contained in the domain of any world[10] and is used as the value assigned to definite descriptions that do not designate (uniquely).

Definition 5 (Models)**.** *$\mathcal{M} = \langle W, \succeq, D, I \rangle$ is called a model (on the frame $\mathcal{F} = \langle W, \succeq, D \rangle$), where I is a function (called interpretation function) such that:*

- *for $c \in C$ and $w \in W$: $I(c, w) \in \mathbb{D}^+$*
- *for $R^n \in \mathbb{P}$ and $w \in W$: $I(R^n, w) \subseteq (\mathbb{D}^+)^n$*

$I(c, w) = a$ says that a is the denotation of c in w.

Definition 6 (Variable assignment)**.** *Given a model $\mathcal{M} = \langle W, \succeq, D, I \rangle$ we call a function $g : V \times W \mapsto \mathbb{D}^+$ a variable assignment (of $\mathcal{M}$).*

Roughly speaking, $g(x, w) = a$ says that a is the denotation of x in w. Note that $g(x, w)$ does not have to be an element of the domain of w.[11] Note also that the variable assignment and the interpretation function are world-dependent. This

[10] $\mathbb{D} \notin \mathbb{D}$.
[11] The element a does not even have to be contained in the actual domain.

is because we do not assume rigidness of terms, as mentioned at the beginning of Sec. 2. To adopt a more mainstream approach using rigid constants and world-independent variable assignments, one must add an assumption: $I(c, w) = I(c, v)$ and $g(x, w) = g(x, v)$ for every constant c and variable x, and every world w and v.

We amend the usual notion of an x-variant as follows. An *x-variant of some variable assignment g at a world w* is a variable assignment h that agrees with g on all values except for x, whose value in every world remains constant, and an element of D_w. Formally:

Definition 7 (x-variant)**.** *Assume a model $\mathcal{M} = \langle W, \succeq, D, I\rangle$, a variable assignment g of $\mathcal{M}$ and an element of the whole domain $d \in \mathbb{D}^+$. We write $g_{x \Rightarrow d}$ for the variable assignment which replaces the value assigned to x at any world by d:*

$$g_{x\Rightarrow d}(z, v) := \begin{cases} d & \textit{if} \ \ (z, v) \in \{x\} \times W \\ g(z, v) & \textit{otherwise} \end{cases}$$

A variable assignment h is an x-variant of g at w iff $h = g_{x\Rightarrow d}$ for some $d \in D_w$.

"Best", in terms of which the truth-conditions for $\bigcirc(-/-)$ are cast, is defined by:

Definition 8 (best)**.** *Given a model $\mathcal{M} = \langle W, \succeq, D, I\rangle$ and a set of worlds $X \subseteq W$ we define*

$$\text{best}(X) := \{w \in X : \forall v \in W(v \in X \Rightarrow w \succeq v)\}$$

$\text{best}(X)$ *is the set of worlds in X that are at least as good as every member of X.*

Remark 1. *We define "best" using the concept of optimality, following the terminology of [29]. This is in keeping with Åqvist [1]'s own proposal. Whether other notions, like maximality, would make a significant difference or only result in minimal changes to the logic as in the propositional case [30] remains an open question for future research.*

The construct "$\mathcal{M}, v \models_g^w \varphi$" can be read as "$\varphi$ holds at v under g if looked at from the point of view of (an inhabitant of) w".

A non-denoting definite description is assigned the value $\mathbb{D}$. This element can verify a given property in a given model. For example, nothing rules out the possibility that $\mathbb{D} \in I(R, w)$. In this case $R(\imath x B(x))$ holds in w even if $\imath x B(x)$ does not denote in w. The only thing that cannot happen is that $\mathbb{D} \in D_w$. Intuitively, one may want to be able to talk about the properties of non-existing individuals, as in "Santa Claus has a beard" or "Santa Claus is not giving gifts to bad children".

Definition 9. *Let $\mathcal{M} = \langle W, \succeq, D, I \rangle$ be a model, g a variable assignment, $x \in V$ and $c \in C$. We define*

- $I_g^w(x) := g(x, w)$
- $I_g^w(c) := I(c, w)$
- $I_g^w(\imath x\varphi) := \begin{cases} h(x,w) & \text{if there exists a \textbf{unique} } x\text{-variant } h \text{ of } g \text{ at } w \\ & \text{such that } \mathcal{M}, w \models_h^w \varphi \\ \mathbb{D} & \text{otherwise} \end{cases}$

The forcing relation $\models$ can be defined inductively as follows:

- $\mathcal{M}, v \models_g^w R^n(t_1, ..., t_n) :\Leftrightarrow \langle I_g^w(t_1), ..., I_g^w(t_n)\rangle \in I(R^n, v)$
- $\mathcal{M}, v \models_g^w \neg\varphi :\Leftrightarrow \mathcal{M}, v \not\models_g^w \varphi$
- $\mathcal{M}, v \models_g^w \varphi \wedge \psi :\Leftrightarrow \mathcal{M}, v \models_g^w \varphi$ *and* $\mathcal{M}, v \models_g^w \psi$
- $\mathcal{M}, v \models_g^w \forall x\varphi :\Leftrightarrow \mathcal{M}, v \models_h^w \varphi$ *for all x-variants h of g at v*
- $\mathcal{M}, v \models_g^w t_1 = t_2 :\Leftrightarrow I_g^w(t_1) = I_g^w(t_2)$
- $\mathcal{M}, v \models_g^w \boxdot\varphi :\Leftrightarrow \forall u \in W \ \mathcal{M}, u \models_g^w \varphi$
- $\mathcal{M}, v \models_g^w \boxtimes\varphi :\Leftrightarrow \forall u \ \forall v' \in W \ \mathcal{M}, u \models_g^{v'} \varphi$
- $\mathcal{M}, v \models_g^w \bigcirc(\psi/\varphi) :\Leftrightarrow best(||\varphi||_{g,w}^{\mathcal{M}}) \subseteq ||\psi||_{g,w}^{\mathcal{M}}$

We drop the reference to $\mathcal{M}$ when it is clear what model is intended.

Definition 10 (Truth in $\mathbf{F}^\forall$)**.** *Given a model $\mathcal{M} = \langle W, \succeq, D, I \rangle$, a variable assignment g, a formula φ and a world w we define what it means that φ is true in $\mathcal{M}$ at w under g (in symbols: $\mathcal{M}, w \models_g \varphi$) as*

$$\mathcal{M}, w \models_g \varphi :\Leftrightarrow \mathcal{M}, w \models_g^w \varphi$$

The meaning of $\boxdot$, $\boxtimes$ and $\bigcirc$ is easier to explain using the following derived truth conditions.

Remark 2 (Derived truth conditions)**.**

- $\mathcal{M}, w \models_g \boxdot\varphi \Leftrightarrow \forall v \in W \ \mathcal{M}, v \models_g^w \varphi$
- $\mathcal{M}, w \models_g \boxtimes\varphi \Leftrightarrow \forall u \ \forall v \in W \ \mathcal{M}, u \models_g^v \varphi$
- $\mathcal{M}, w \models_g \bigcirc(\psi/\varphi) \Leftrightarrow best(||\varphi||_{g,w}^{\mathcal{M}}) \subseteq ||\psi||_{g,w}^{\mathcal{M}}$

When evaluating the truth-value of $\boxdot\varphi$ at w, one moves to an arbitrary world v, and determines the truth-value of φ in v from w's perspective. This means giving to the terms occurring in φ the denotation they have in w. When evaluating the truth-value of $\boxtimes\varphi$ at w, one moves to an arbitrary world u, and evaluates φ in u from every other world's v perspective. As a consequence, we have $\mathcal{M}, w \models_g \boxtimes\varphi \to \boxdot\varphi$ for every w, g and formula φ.

For obligation, the idea is similar. The standard evaluation rule puts $\bigcirc(\psi/\varphi)$ as true at w whenever all the best φ-worlds are ψ-worlds. The φ-worlds and the ψ-worlds in question are those according to w's perspective. This is how the principle of substitution *salva veritate* will be validated for $\bigcirc$ and $\boxdot$, and invalidated for $\boxtimes$.

Definition 11. *Given a model $\mathcal{M} = \langle W, \succeq, D, I\rangle$. $\succeq$ is reflexive if $\forall w \in W(w \succeq w)$, and $\succeq$ fulfils the limitedness condition if for every φ, g and $w \in W$ we have*

$$||\varphi||^{\mathcal{M}}_{g,w} \neq \emptyset \Rightarrow best(||\varphi||^{\mathcal{M}}_{g,w}) \neq \emptyset$$

$\mathcal{U}$ is the class of models in which $\succeq$ is reflexive and fulfils limitedness.

Intuitively, the limitedness condition validates the dyadic version of the **D** axiom (with $\Diamond$ replaced with $\diamondsuit$) involved in derivation 3 of the collapse (see Subsec. 2.2).

Definition 12 (Validity in $\mathbf{F}^{\forall}$)**.** *We set:*

- *φ holds at w in a model $\mathcal{M}$ (notation: $\mathcal{M}, w \models \varphi$) if for every variable assignment g, we have that $\mathcal{M}, w \models_g \varphi$;*
- *φ is valid in a model $\mathcal{M}$ (notation: $\mathcal{M} \models \varphi$) if for every world w we have $\mathcal{M}, w \models \varphi$;*
- *φ is valid in a class $\mathbb{M}$ of models (notation: $\mathbb{M} \models \varphi$) if for every model $\mathcal{M} \in \mathbb{M}$ we have $\mathcal{M} \models \varphi$;*
- *φ is valid (notation: $\models \varphi$) if φ is valid in the class $\mathcal{U}$ as defined above.*

4 Benchmarking

We test the account introduced in Sec. 3 against the requirements discussed in Sec. 2.

4.1 Extensionality / intensionality / self-negation

A proof of the principle of extensionality in its general form is given in Subsec. 4.2. For simplicity's sake, we only discuss the examples considered in Sec. 2.

Proposition 3 (Extensionality of $\bigcirc$, requirement 1)**.** *We have:*

$$(\bigcirc B(\imath x P(x), \imath y(W(y) \wedge Pr(y))) \wedge j = \imath x P(x)) \to \bigcirc B(j, \imath y(W(y) \wedge Pr(y)))$$

Proof. When a formula does not contain a free variable its truth-conditions do not depend on which variable assignment is assumed. Therefore for this and all future proofs (in which no free variable is involved) we always deal with an arbitrary variable assignment. Now, if $w \models_g^w j = \imath x P(x)$, then for every $u \in best(||\top||_{g,w}^{\mathcal{M}})$

$$u \models_g^w B(\imath x P(x), \imath y(W(y) \wedge Pr(y))) \Leftrightarrow u \models_g^w B(j, \imath y(W(y) \wedge Pr(y)))$$

This is because the terms on both sides of $\Leftrightarrow$ get the denotation they have in w. Therefore:

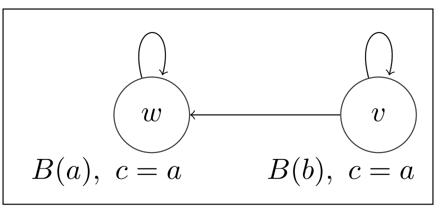
$$\begin{aligned} & best(||\top||_{g,w}^{\mathcal{M}}) \subseteq ||B(\imath x P(x), \imath y(W(y) \wedge Pr(y)))||_{g,w}^{\mathcal{M}} \\ \Leftrightarrow\ & best(||\top||_{g,w}^{\mathcal{M}}) \subseteq ||B(j, \imath y(W(y) \wedge Pr(y)))||_{g,w}^{\mathcal{M}} \end{aligned}$$

This proves the claim. □

Proposition 4 (Intensionality of $\boxtimes$, requirement 2)**.** *We do not have:*

$$c = \imath x B(x) \rightarrow (\boxtimes(c = c) \leftrightarrow \boxtimes(c = \imath x B(x)))$$

Proof. Put $\mathcal{M} = \langle W, \succeq, I, D \rangle$ with (an arrow from v to w means $v \succeq w$, and no arrow from from w to v means $w \not\succeq v$):



$W := \{w, v\}$
$\succeq :=$ the reflexive closure of $\{(v, w)\}$
$D_w := \{a\}, \quad D_v := \{a, b\}$
$I(B, w) := \{a\}, \quad I(B, w) := \{b\}$
$I(c, w) := a, \quad I(c, v) := a$

The condition of limitedness is fulfilled. We have:

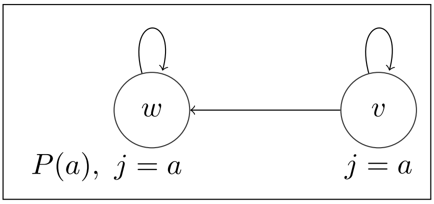
- $w \models_g^w c = \imath x B(x)$ since c and $\imath x B(x)$ denote a in w
- $w \models_g^w \boxtimes(c = c)$ since $c = c$ is a tautology
- $w \not\models_g^w \boxtimes(c = \imath x B(x))$ since $w \not\models_g^v c = \imath x B(x)$[12]

□

Proposition 5 (Self-negation, requirement 4)**.** *The sentences (5-a), (5-b) and (5-c) are simultaneously satisfiable.*

[12] c and $\imath x B(x)$ do not have the same denotation in v.

Proof. We give a model which satisfies all three formulas in the same world.



$W := \{w, v\}$
$\succeq :=$ the reflexive closure of $\{(v, w)\}$
$D_w := \{a\}, \quad D_v := \{a\}$
$I(P, w) := \{a\}, \quad I(P, v) := \emptyset$
$I(j, w) := a, \quad I(j, v) := a$

As before $\succeq$ is limited. We have:

- $w \models_g^w j = \imath x P(x)$ since j and $\imath x P(x)$ denote a in w
- $w \models_g^w \bigcirc \neg P(j)$ since a is not P in v
- $w \models_g^w \bigcirc \neg P(\imath x P(x))$ since a (=the unique P in w) is not P in v

The paradox is resolved by having Jose, who is the pope in the actual world w, not be the pope in the best world v. Therefore $\bigcirc \neg P(\imath x P(x))$ can be satisfied. $\square$

4.2 Extensionality (general form)

We show the principle of extensionality in its general form. Where φ is a formula and s and t terms, let $\varphi_{t \hookrightarrow s}$ be the result of replacing zero up to all unbound occurrences of t,[13] in φ, by s. We may re-letter bound variables, if necessary, to avoid rendering the new occurrences of variables in s bound in φ.

Proposition 6. *Consider some g and some w in $\mathcal{M}$ such that $w \models_g^w t = s$. Then, for all v in $\mathcal{M}$,*

$$v \models_g^w \varphi \leftrightarrow \varphi_{t \hookrightarrow s} \qquad (\#)$$

provided t is not contained in the scope of the $\boxtimes$ operator in φ.

Proof. By induction on the complexity n of a formula φ. The base case, if φ is $R(t_1, ..., t_m)$ with $\ulcorner R(t_1, ..., t_m) \urcorner = 0$, follows from the definitions involved. For the inductive case, we assume (#) holds for all $k < n$, and *for all* v in $\mathcal{M}$. We only consider three cases−the other ones are left to the reader:

- $\varphi := \forall x\ \psi$. Given the restrictions put on t and s, we have the following chain of equivalences:

$$v \models_g^w \forall x\ \psi \text{ iff } v \models_h^w \psi \quad \text{for all } x\text{-variants } h \text{ at } v$$

[13] By an unbounded occurrence of t, we mean that no variables in t are in the scope of a quantifier or a definite description *not* in t.

$$v \models_h^w \psi_{t \hookrightarrow s} \quad \text{for all } x\text{-variants } h \text{ at } v \text{ (by IH)}$$
$$v \models_g^w \forall x\, \psi_{t \hookrightarrow s}$$

- $\varphi := \bigcirc(\chi/\psi)$.

$$\begin{aligned} v \models_g^w \bigcirc(\chi/\psi) \text{ iff } & best(||\psi||_{g,w}^{\mathcal{M}}) \subseteq ||\chi||_{g,w}^{\mathcal{M}} \\ & best(||\psi_{t \hookrightarrow s}||_{g,w}^{\mathcal{M}}) \subseteq ||\chi_{t \hookrightarrow s}||_{g,w}^{\mathcal{M}} \text{ (by IH)} \\ & v \models_g^w \bigcirc(\chi_{t \hookrightarrow s}/\psi_{t \hookrightarrow s}) \\ & v \models_g^w \bigcirc(\chi/\psi)_{t \hookrightarrow s} \end{aligned}$$

- $\varphi := R(t_1, ..., t_m)$. Assume $v \models_g^w R(t_1, ..., t_m)$. If t appears only as one of the t_i's, then we are done. So let us suppose that t appears *in* one (or more) of the t_i's. W.l.o.g. let t only appear in $t_1 = \imath x \psi$. By the IH $w \models_g^w \psi \leftrightarrow \psi_{t \hookrightarrow s}$, so $I_g^w(\imath x \psi) = I_g^w(\imath x \psi_{t \hookrightarrow s})$. Consider some $v \in W$. We have $\langle I_g^w(\imath x \psi), ..., I_g^w(t_m)\rangle \in I(R, v)$, so $\langle I_g^w(\imath x \psi_{t \hookrightarrow s}), ..., I_g^w(t_m)\rangle \in I(R, v)$. Hence $v \models_g^w R(t_1, ..., t_m)_{t \hookrightarrow s}$ as required. For the converse implication, the argument is the same.

□

Corollary 1 (Extensionality)**.** *The principle (E) is valid:*

$$\models t = s \rightarrow (\varphi \leftrightarrow \varphi_{t \hookrightarrow s}) \quad \textit{if } t \textit{ is not in the scope of } \boxtimes \tag{E}$$

Proof. This follows from Prop. 6 putting $v = w$. □

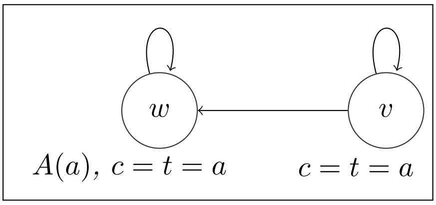
Remark 3. *We draw the reader's attention to the proviso "if t is not in the scope of $\boxtimes$". At first, it may seem that all terms, including definite descriptions, are rigid. However, this is not the case. As the proviso indicates, the terms do not exhibit a rigid behavior by themselves. It is the operators $\bigcirc$ and $\boxdot$ that treat the terms rigidly, ensuring they remain tied to the original world.*[14] *By contrast, the modal operator $\boxtimes$ does not treat terms rigidly, as shown in Prop. 4.*

4.3 Deontic collapse

We start by explaining how the collapse is avoided semantically. We define a model in which the formulas at steps (a)-(e) in derivation 4 are true in the actual world w but the formula at step (f) is not.

[14] Goble [15, p. 347] makes a similar point.

Example 1. *Put $\varphi := A(c)$. $\mathcal{M}$ is defined by*



$$
\begin{aligned}
&W := \{w, v\}\\
&\succeq := \text{ the reflexive closure of } \{(v, w)\}\\
&D_w := \{a\}, \quad D_v := \{a\}\\
&I(c, w) := I(c, v) := a\\
&I(t, w) := I(t, v) := a\\
&I(A, w) := \{a\}, \quad I(A, v) := \emptyset
\end{aligned}
$$

We have

(a) $w \models_g A(c)$ *since* $I(c, w) = a \in I(A, w)$
(b) $w \models_g \boxtimes \exists y(y = t)$ *since* $I(t, w) = I(t, v) = a \in D_w$
and $I(t, w) = I(t, v) = a \in D_v$
(c) $w \models_g t = \iota x(x = t \wedge A(c))$ *since* $I(t, w) = a = I^w_g(\iota x(x = t \wedge A(c)))$
(d) $w \models_g \bigcirc \exists y(y = t)$ *since* $I(t, w) = a \in D_v$[15]
(e) $w \models_g \bigcirc \exists y(y = \iota x(x = t \wedge A(c)))$ *since* $I^w_g(\iota x(x = t \wedge A(c))) = a \in D_v$
not-(f) $w \not\models_g \bigcirc A(c)$ *since* $I(c, w) = a \notin I(A, v)$

We briefly explain the last two steps:

- *(e) holds, because* $v \models^w_g \exists y(y = \iota x(x = t \wedge A(c)))$*, where v is the best world. This formula picks up the unique x for which $x = t \wedge A(c)$ holds in w, namely a. It tells us that such an x exists in v, but not which properties it has* **in** *v (existence not being a property).*
- *(e) does* **not** *imply* $v \models^w_g A(c)$*; hence not-(f). The clause* $v \models^w_g A(c)$ *says that the denotation of c in w, namely a, has property A in v–this is not the case.*

In other words, statements $v \models^w_g \exists y(y = \iota x(x = t \wedge A(c)))$ *and* $v \models^w_g A(c)$ *talk about the same individual c, but in two different worlds,*[16] *where it can evidently verify different properties.*

This model gives a counter-example to the principle of inheritance. The implication $\exists y(y = \iota x(x = t \wedge A(c))) \to A(c)$ is valid, but the implication $\bigcirc \exists y(y = \iota x(x = t \wedge A(c))) \to \bigcirc A(c)$ is not valid.

[15] By definition $v \models^w_g \exists y(y = t)$ holds if there exists an y-variant h of g at v such that $h(y, w) = I(t, w)$. This is equivalent to $I(t, w)$ being an element of D_v.

[16] w in the former case, and v in the latter case.

To explain proof-theoretically how the deontic collapse is avoided, we introduce the notion of "variable only" version φ^* of a formula φ. Intuitively, φ^* is obtained by substituting, in φ, a new variable for every definite description and constant occurring in φ. This ensures that φ^* contains only variables, making it impossible to apply the rule of inheritance (and necessitation) from which the collapse follows. Formally:

Definition 13 (Variable only version, Goble [16])**.** *Given a formula φ, we define φ^* as the formula in which all terms $t_1, ..., t_n$, which are not variables and are occurring in the formula φ, have been replaced by $x_1, ..., x_n \in V$ respectively. The variables $x_1, ..., x_n$ are the first, pairwise different, elements of V such that $x_1, ..., x_n$ do not occur in φ.*

Example 2. *Let A, B and C be predicate symbols, $x, y, z \in V$ the first three variables of V, $c \in C$ a constant and $\varphi \in WF$ a well-formed formula:*

- $A(\imath y\varphi, c)^* = A(x, z)$
- $\forall x A(\imath y B(y, d), x)^* = \forall x A(z, x)$
- $A(\imath y B(\imath x C(x, y)), y)^* = A(z, y)$
- $A(y, y)^* = A(y, y)$

In Sec. 2.2, we mentioned that the collapse will be avoided by restraining the application of the rule of necessitation for $\boxtimes$. We are now in a position to define formally our new rule:

$$\text{If } \models \varphi^* \text{ then } \models \boxtimes\varphi \tag{N*-$\boxtimes$}$$

Like in Goble's original solution, N*-$\boxtimes$ entails the following restricted form of inheritance:

$$\text{If } \models (\psi_1 \to \psi_2)^* \text{ then } \models \bigcirc(\psi_1/\varphi) \to \bigcirc(\psi_2/\varphi) \tag{In*}$$

Before showing the validity of these two rules, we observe that the other law involved in the collapse, $\boxtimes\psi \to \bigcirc(\psi/\varphi)$, still holds. This follows at once from the following:

Proposition 7. *We have*

$$\models \boxtimes\psi \to \boxdot\psi \tag{$\boxtimes$2$\boxdot$}$$

$$\models \boxdot\psi \to \bigcirc(\psi/\varphi) \tag{$\boxdot$2$\bigcirc$}$$

Proof. ($\boxtimes$2$\boxdot$) is straightforward, and may be left to the reader. For ($\boxdot$2$\bigcirc$), let us assume $w \models_g \boxdot\psi$ holds for a fixed model $\mathcal{M} = \langle W, \succeq, D, I\rangle$, a world $w \in W$ and a variable assignment g. This is equivalent to $||\psi||^{\mathcal{M}}_{g,w}$ being equal to the whole set of worlds W. Hence we can infer that for any formula φ we have $best(||\varphi||^{\mathcal{M}}_{g,w}) \subseteq W = ||\psi||^{\mathcal{M}}_{g,w}$, which, by definition, means $w \models_g \bigcirc(\psi/\varphi)$. □

In Sec. 1, we pointed out that combining alethic and deontic modalities allows us to express fundamental principles, such as the law of strong factual detachment (SFD). We note that the distinction between extensional and intensional contexts has no bearing on the validity of this law, as this one continues to hold for both types of alethic modal operators.

Proposition 8 (Strong factual detachment)**.** *We have:*

$$\models \bigcirc(\psi/\varphi) \wedge \boxtimes\varphi \rightarrow \bigcirc\psi \qquad (\boxtimes\text{-SFD})$$

$$\models \bigcirc(\psi/\varphi) \wedge \boxdot\varphi \rightarrow \bigcirc\psi \qquad (\boxdot\text{-SFD})$$

Proof. ($\boxtimes$-SFD) follows from ($\boxdot$-SFD) and ($\boxtimes$2$\boxdot$), so we concentrate on ($\boxdot$-SFD). Assume a model M, a world w and a variable assignment g such that

$$w \models_g \bigcirc(\psi/\varphi) \qquad (1)$$

$$w \models_g \boxdot\varphi \qquad (2)$$

By Def. 10,

$$w \models^w_g \bigcirc(\psi/\varphi) \qquad (3)$$

$$w \models^w_g \boxdot\varphi \qquad (4)$$

Let $v \in best(||\top||^{\mathcal{M}}_{g,w})$.

- By (4), $v \models^w_g \varphi$
- Let $u \models^w_g \varphi$. Clearly, $u \models^w_g \top$, so that $u \in ||\top||^{\mathcal{M}}_{g,w}$. Hence $v \succeq u$.

This shows that $v \in best(||\varphi||^{\mathcal{M}}_{g,w})$. By (3), $v \in ||\psi||^{\mathcal{M}}_{g,w}$, and so by Def. 9 $w \models^w_g \bigcirc\psi$. By Def. 10, $w \models_g \bigcirc\psi$. By Def. 12, $\models$ $\boxdot$-SFD.

□

We now show that the rules (N*-$\boxtimes$) and (In*) preserve validity. To show this we need the following two lemmas.

Lemma 1. *Given a formula φ and a model $\mathcal{M}$, then*

$$\mathcal{M} \models \varphi^* \Rightarrow \mathcal{M} \models \boxtimes(\varphi^*)$$

Proof. Let φ be a formula and $\mathcal{M} = \langle W, \succeq, D, I\rangle$ a model. If for every world $w \in W$ and every variable assignment g of $\mathcal{M}$ it holds that $w \models_g \varphi^*$, it follows that $w \models_g^w \varphi^*$ holds for every world $w \in W$ and every variable assignment g of $\mathcal{M}$. Now let us take two arbitrary but fixed worlds $v, w \in W$ and an arbitrary but fixed variable assignment g and define a new variable assignment $h : V \times W \to \mathbb{D}^+$ of $\mathcal{M}$ as:

$$h(x,u) := \begin{cases} g(x,w) & \text{if} \quad u = v \\ g(x,v) & \text{if} \quad u = w \\ g(x,u) & \text{otherwise} \end{cases}$$

Since h and g only swap how they see the variables at w and v, and φ^* does not contain constants or definite descriptions, we get $\forall u(u \models_g^w \varphi^* \Leftrightarrow u \models_h^v \varphi^*)$. Therefore from $v \models_h^v \varphi^*$, which holds by assumption, we can infer $v \models_g^w \varphi^*$. Since $v, w \in W$ and g were arbitrary we can conclude $\mathcal{M} \models \boxtimes\varphi^*$. $\square$

Lemma 2. *Given a formula φ and a model $\mathcal{M}$, then*

$$\mathcal{M} \models \varphi^* \Rightarrow \mathcal{M} \models \varphi$$

Proof. This proof is done by contraposition. Suppose there are $\mathcal{M} = \langle W, \succeq, D, I\rangle$, $w \in W$ and g such that $w \not\models_g^w \varphi$. Let $t_1, ..., t_n$ be all terms in φ which are replaced by the corresponding variables $x_1, ..., x_n$ in φ^*. Then for the variable assignment

$$h(x,v) := \begin{cases} I_g^v(t_i) & \text{if} \quad (x,v) \in \{x_i\} \times W \quad \text{where} \quad i \in \{1, ..., n\} \\ g(x,v) & \text{otherwise} \end{cases}$$

we have $w \not\models_h^w \varphi^*$. $\square$

Putting those two lemmas together, we can prove the soundness of (N*-$\boxtimes$):

Lemma 3. *Given a formula φ and a model $\mathcal{M}$ then*

$$\mathcal{M} \models \varphi^* \quad \textit{implies} \quad \mathcal{M} \models \boxtimes\varphi$$

Proof. $\mathcal{M} \models \varphi^* \Rightarrow \mathcal{M} \models \boxtimes(\varphi^*) \Leftrightarrow \mathcal{M} \models (\boxtimes\varphi)^* \Rightarrow \mathcal{M} \models \boxtimes\varphi$. $\square$

Theorem 1. *We have*

$$\textit{If} \ \models \varphi^* \ \textit{then} \ \models \boxtimes\varphi \qquad (\text{N}^*\text{-}\boxtimes)$$

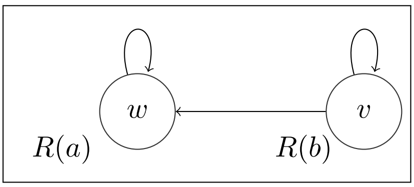
$$\textit{If} \ \models (\psi_1 \to \psi_2)^* \ \textit{then} \ \models \bigcirc(\psi_1/\varphi) \to \bigcirc(\psi_2/\varphi) \qquad (\text{In}^\star)$$

Proof. The first rule follows at once from Lem. 3. The second rule follows from the first one and Prop. 7. $\square$

We end this section by showing that the rule of necessitation in its plain form fails for $\boxtimes$. Here is a counter-example. The formula $\exists y(y = \imath x R(x)) \to R(\imath x R(x))$ is valid in any model. To see why, fix a model $\mathcal{M} = \langle W, \succeq, D, I \rangle$, a variable assignment g, and a world $w \in W$. Assume $w \models_g \exists y(y = \imath x R(x))$. Hence, there exists a y-variant h of g at w such that $h(y, w) = I_h^w(\imath x R(x))$. This means that $h(y, w) = a$ for some $a \in D_w$. By definition of $\imath x R(x)$, a is the unique element in D_w s.t. $a \in I(R, w)$. So $w \models_h R(\imath x R(x))$. Since y does not occur in $R(\imath x R(x))$ we conclude $w \models_g R(\imath x R(x))$ as required.

Now we define a model in which $\boxtimes[\exists y(y = \imath x R(x)) \to R(\imath x R(x))]$ is not valid:

Example 3. *Consider the model* $\mathcal{M} := \langle W, \succeq, D, I \rangle$ *with*



$W := \{w, v\}$
$\succeq :=$ *the reflexive closure of* $\{(v, w)\}$
$D_w := \{a, b\}, \quad D_v := \{a, b\}$
$I(R, w) := \{a\}, \quad I(R, v) := \{b\}$

We have $v \models_g^w \exists y(y = \imath x R(x))$, *as* $I_g^w(\imath x R(x)) = a \in D_v$. *But* $v \not\models_g^w R(\imath x R(x))$ *because* $I_g^w(\imath x R(x)) = a \notin I(R, v)$. *So* $\mathcal{M} \not\models \boxtimes[\exists y(y = \imath x R(x)) \to R(\imath x R(x))]$.

5 Concluding remarks

We have defined and studied a new perspectival account of conditional obligation. A number of requirements were identified, and shown to be met by the framework. The framework allows for a more nuanced way of approaching first-order deontic principles. Topics for future research include:

(i) to investigate variant candidate truth-conditions for $\boxtimes$;

(ii) to find a suitable axiomatic basis;

Ad (i): the truth-conditions for $\boxtimes$ in Def. 9 allowed us to make the minimal changes to the axiomatic basis of **F**. The most significant change is that Lewis's absoluteness principle $\bigcirc(\psi/\varphi) \to \boxtimes \bigcirc (\psi/\varphi)$, stipulating that obligations are necessary, goes away. This may be considered good news. But ($\boxtimes 2 \bigcirc$) remains, and this law may be considered counter-intuitive. The following alternative truth-conditions may be used:

$$w \models_g \boxtimes\varphi \text{ iff } \forall v:\ v \models_g^v \varphi$$

Intuitively: $w \models_g \boxtimes\varphi$ holds, if φ holds at all v under the hypothesis that the terms occurring in φ take the reference they have in this very same world. With this definition of $\boxtimes$, ($\boxtimes 2\bigcirc$) goes away, and the rule of necessitation holds without any restriction.

Ad (ii): we have identified a sound axiomatic basis for the logic. This logic is defined in Appendix C. Completeness is left as a topic for future research.

Appendix A: Universal instantiation

In Sec. 2, we observed that the principles of universal instantiation (UI) and existential generalisation (EG) hold in our logic even if the replaced term appears within a deontic operator. We are now going to prove this observation and discuss what these principles tell us about our logic. They are inter-derivable, so we focus on (UI).

In the following $\varphi_{x \Rightarrow t}$ denotes the result of replacing all free occurrences of the variable x, in a formula φ, by the term t.

Proposition 9. *Consider some φ, some g, some world w in $\mathcal{M}$ and a term t such that no bound variable in φ appears free in t. Then for $d := I_g^w(t)$ and for all v the equivalence*

$$v \models_g^w \varphi_{x \Rightarrow t} \Leftrightarrow v \models_{g_{x \Rightarrow d}}^w \varphi \qquad (\#')$$

holds, provided x is not contained in the scope of the $\boxtimes$ operator in φ.

Proof. The proof is by induction on the complexity n of a formula φ and in a similar fashion to the proof of Prop. 6. The base case, if φ is $R(t_1, ..., t_m)$ with $\ulcorner R(t_1, ..., t_m) \urcorner = 0$, follows from the definitions involved. For the inductive case, we assume #' holds for all $k < n$, and *for all* v in $\mathcal{M}$. We again only consider three cases:

- $\varphi := \forall y\ \psi$. In the case that y is the variable x, the formulas $\varphi_{x \Rightarrow t}$ and φ are the same [17] and the evaluation via the variable assignments g and $g_{x \Rightarrow t}$ coincide. In the case that y is not x, then given the restrictions put on t, we have that y does not appear free in t. Therefore we get the following chain of equivalences:

  $v \models_g^w (\forall y\ \psi)_{x \Rightarrow t}$ iff $v \models_g^w \forall y\ (\psi_{x \Rightarrow t})$

[17] since all x in φ are bound by $\forall$

$$v \models^w_h \psi_{x\Rightarrow t} \quad \text{for all } y\text{-variants } h \text{ of } g \text{ at } v$$
$$v \models^w_{h_{x\Rightarrow d}} \psi \quad \text{for all } y\text{-variants } h \text{ of } g \text{ at } v \text{ (by IH)}$$
$$v \models^w_{h'} \psi \quad \text{for all } y\text{-variants } h' \text{ of } g_{x\Rightarrow d} \text{ at } v \text{ (since } x \neq y\text{)}$$
$$v \models^w_{g_{x\Rightarrow d}} \forall y\, \psi$$

- $\varphi := \bigcirc(\chi/\psi)$.

$$v \models^w_g \bigcirc(\chi/\psi)_{x\Rightarrow t} \text{ iff } best(||\psi_{x\Rightarrow t}||^{\mathcal{M}}_{g,w}) \subseteq ||\chi_{x\Rightarrow t}||^{\mathcal{M}}_{g,w}$$
$$best(||\psi||^{\mathcal{M}}_{g_{x\Rightarrow d},w}) \subseteq ||\chi||^{\mathcal{M}}_{g_{x\Rightarrow d},w} \text{ (by IH)}$$
$$v \models^w_{g_{x\Rightarrow d}} \bigcirc(\chi/\psi)$$

- $\varphi := R(t_1, ..., t_m)$. Assume $v \models^w_g R(t_1, ..., t_m)_{x\Rightarrow t}$. If x appears only as one of the t_i's, then we are done. So let us suppose that x appears *in* one (or more) of the t_i's. W.l.o.g. let x only appear in $t_1 = \imath y\psi$. By the IH $w \models^w_g \psi_{x\Rightarrow t} \Leftrightarrow w \models^w_{g_{x\Rightarrow d}} \psi$, so $I^w_g(\imath y\psi_{x\Rightarrow t}) = I^w_{g_{x\Rightarrow d}}(\imath y\psi)$. Consider some $v \in W$. We have $\langle I^w_g(\imath y\psi_{x\Rightarrow t}), ..., I^w_g((t_m)_{x\Rightarrow t})\rangle \in I(R, v)$, so $\langle I^w_{g_{x\Rightarrow d}}(\imath y\psi), ..., I^w_{g_{x\Rightarrow d}}(t_m)\rangle \in I(R, v)$. Hence $v \models^w_{g_{x\Rightarrow d}} R(t_1, ..., t_m)$ as required. For the converse implication, the argument is the same.

□

Corollary 2 (Universal instantiation)**.** *The principle of (UI) is valid:*

$$\models E(t) \to (\forall x\varphi \to \varphi_{x\Rightarrow t}) \quad \textit{if } x \textit{ is not in the scope of } \boxtimes$$

Proof. This follows from the fact that $w \models^w_g E(t)$ holds if and only if $d := I^w_g(t) \in D_w$ holds. Therefore $w \models^w_g E(t)$ implies that $g_{x\Rightarrow d}$ is a x-variant of g at w. Now using Prop. 9 and putting $v = w$ we obtain $w \models^w_g \varphi_{x\Rightarrow t}$ from $w \models^w_g E(t)$ and $w \models^w_g \forall x\varphi$. □

(UI) highlights the difference between $\forall x\bigcirc(\psi(x)/\varphi)$ and $\bigcirc(\forall x\psi(x)/\varphi)$. (UI) can be applied to the former but not to the latter formula. This is because the Barcan formula, as well as the converse Barcan formula, do not hold.[18] $\forall x\bigcirc(\psi(x)/\varphi)$ states that ψ is an obligation for each existing individual under condition φ. $\bigcirc(\forall x\psi(x)/\varphi)$ states that $\forall x\psi(x)$ is obligatory under condition φ. This means that in an optimal φ-world everyone fulfills ψ. This does not imply that someone currently existing has to fulfill ψ. As an example let us contrast the two sentences:

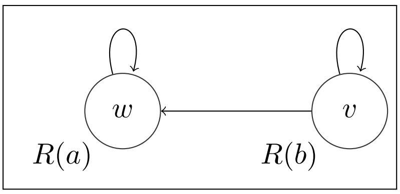
(7) a. Everyone should live eco-friendly: $\forall x \bigcirc\varphi(x)$

[18] Also, $\forall x \bigcirc(\psi/\varphi(x))$ and $\bigcirc(\psi/\forall x\varphi(x))$ do not imply each other.

b. It should be that everyone live eco-friendly: $\bigcirc\forall x\varphi(x)$

Unlike (7-b), (7-a) describes an obligation binding each existing individual. From (7-a) and $E(t)$ one gets $\bigcirc\varphi(t)$. By contrast (7-b) does not warrant the move to $\bigcirc\varphi(t)$ even in the presence of $E(t)$. This is as it should be.

Example 4. *Consider the model $\mathcal{M} := \langle W, \succeq, D, I\rangle$ with*



$W := \{w, v\}$
$\succeq :=$ *the reflexive closure of* $\{(v, w)\}$
$D_w := \{a\}, \quad D_v := \{b\}$
$I(t, w) := I(t, v) := a$
$I(R, w) := \{a\}, \quad I(R, v) := \{b\}$

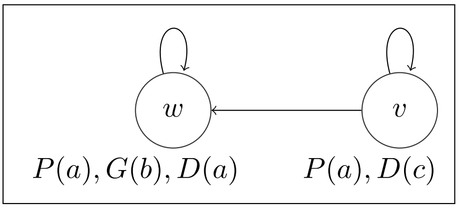
We have $v \models^w_g \forall x R(x)$, as $g_{x\Rightarrow b}$ is the only x-variant of g at v and $I^w_{g_{x\Rightarrow b}}(x) = b \in I(R, v)$. Furthermore $w \models^w_g E(t)$ holds, since $I^w_g(t) = a \in D_w$. But $v \not\models^w_g R(t)$ because $I^w_g(t) = a \notin I(R, v)$. Hence $\mathcal{M} \not\models E(t) \to (\bigcirc\forall x R(x) \to \bigcirc R(t))$.

For our final example, we take a look at the application of existential generalisation when the term appears in the antecedent:

(8) a. There exists a (current) Pope: $E(\imath x P(x))$
b. It ought to be that John says grace if the Pope joins him for dinner: $\bigcirc(G(j)/D(\imath x P(x)))$
c. There exists someone such that John ought to say grace if this person joins him for dinner: $\exists y \bigcirc (G(j)/D(y))$

The inference from (8-a) and (8-b) to (8-c) is both intuitively and formally valid. On the other hand, $\bigcirc(G(j)/\exists y D(y))$ does not follow from (8-a) and (8-b), as shown below.

Example 5. *Consider the model* $\mathcal{M} := \langle W, \succeq, D, I \rangle$ *with*



$$
\begin{aligned}
&W := \{w, v\}\\
&\succeq := \textit{the reflexive closure of } \{(v, w)\}\\
&D_w := \{a, b, c\}, \quad D_v := \{a, b, c\}\\
&I(j, w) := I(j, v) := b\\
&I(P, w) := \{a\}, \quad I(P, v) := \{a\}\\
&I(G, w) := \{b\}, \quad I(G, v) := \{\}\\
&I(D, w) := \{a\}, \quad I(D, v) := \{c\}
\end{aligned}
$$

- $w \models_g^w E(\imath x P(x))$ *since* $I_g^w(\imath x P(x)) = a \in D_w$
- $w \models_g^w \bigcirc(G(j)/D(\imath x P(x)))$ *since* $best(||D(\imath x P(x))||_{g,w}^M) = \{w\} \subseteq \{w\} = ||G(j)||_{g,w}^M$
- $w \not\models_g^w \bigcirc(G(j)/\exists y D(y))$ *since* $best(||\exists y D(y)||_{g,w}^M) = \{v\} \not\subseteq \{w\} = ||G(j)||_{g,w}^M$

Appendix B: Inclusion of F

The following is an axiomatization of Åqvist's system **F**.

Axioms:

All truth-functional tautologies

S5-schemata for $\Box$ and $\Diamond$	(S5)
$\bigcirc(\psi \to \chi/\varphi) \to (\bigcirc(\psi/\varphi) \to \bigcirc(\chi/\varphi))$	(COK)
$\bigcirc(\psi/\varphi) \to \Box \bigcirc(\psi/\varphi)$	(Abs)
$\Box\psi \to \bigcirc(\psi/\varphi)$	(O-nec)
$\Box(\varphi \leftrightarrow \psi) \to (\bigcirc(\chi/\varphi) \leftrightarrow \bigcirc(\chi/\psi))$	(Ext)
$\bigcirc(\varphi/\varphi)$	(Id)
$\bigcirc(\chi/\varphi \wedge \psi) \to \bigcirc(\psi \to \chi/\varphi)$	(Sh)
$\Diamond\varphi \to (\bigcirc(\psi/\varphi) \to P(\psi/\varphi))$	(D*)

Rules:

If $\vdash \varphi$ and $\vdash \varphi \to \chi$ then $\vdash \chi$	(MP)
If $\vdash \varphi$ then $\vdash \Box\varphi$	(N)

An explanation of the axioms can be found in [31]. The distinctive axiom of the system is (D*). This is the dyadic version of the **D** axiom. We now show that the system $\mathbf{F}^\forall$ is a first-order extension of **F**:

Theorem 2. *The rule (MP) and all the axioms of* **F***, where* $\Box$ *is replaced with* $\boxdot$ *and* $\Diamond$ *is replaced with* $\Diamond$*, are valid in* $\boldsymbol{F}^\forall$*.*

Proof. This proof works very similarly to the propositional case. We therefore limit ourselves to (D*). Suppose $w \models_g^w \Diamond\varphi$. Then, there is some $v \in W$ such that $v \models_g^w \varphi$. Suppose $w \models_g^w \bigcirc(\psi/\varphi)$. Then, $best(||\varphi||_{g,w}^{\mathcal{M}}) \subseteq ||\psi||_{g,w}^{\mathcal{M}}$. By limitedness, there is $v' \in W$ such that $v' \in best(||\varphi||_{g,w}^{\mathcal{M}})$. Combining the two, it immediately follows that $best(||\varphi||_{g,w}^{\mathcal{M}}) \cap ||\psi||_{g,w}^{\mathcal{M}} \neq \emptyset$, which is equivalent to $w \models_g^w P(\psi/\varphi)$. $\Box$

As mentioned in Sect. 3 the operator $\boxdot$ can be defined in terms of $\bigcirc$ in $\mathbf{F}^\forall$, in the same way as $\Box$ can be defined in terms of $\bigcirc$ in **F**. Formally:

Theorem 3. $\models \boxdot\varphi \leftrightarrow \bigcirc(\bot/\neg\varphi)$.

Proof.

$$
\begin{array}{lll}
w \models_g^w \boxdot\varphi \text{ iff} & ||\varphi||_{g,w}^{\mathcal{M}} = W & \text{truth conditions for } \boxdot \\
& ||\neg\varphi||_{g,w}^{\mathcal{M}} = \emptyset & \text{truth conditions for } \neg \\
& best(||\neg\varphi||_{g,w}^{\mathcal{M}}) = \emptyset & \text{by limitedness} \\
& best(||\neg\varphi||_{g,w}^{\mathcal{M}}) \subseteq ||\bot||_{g,w}^{\mathcal{M}} & \\
& w \models_g^w \bigcirc(\bot/\neg\varphi) & \text{truth conditions for } \bigcirc
\end{array}
$$

$\Box$

In Sec. 1, we mentioned that $\boxtimes$ is a primitive modality in $\mathbf{F}^\forall$. This directly follows from the fact that, unlike $\bigcirc$, $\boxtimes$ is an intensional modality. Consequently, we do not have the validity of all the axioms of **F** with $\Box$ and $\Diamond$ replaced with $\boxtimes$ and $\diamondsuit$, respectively. For a full list of valid axioms, see Appendix C.

Appendix C: Axiomatisation of $\mathbf{F}^{\forall}$

A sound Hilbert axiomatic system of the logic proposed in this paper is shown below. An explanation of the axioms and rules of FO logic with definite descriptions can be found in [39].

Axioms:

All truth functional tautologies	
All axioms of system **F**	with $\Box$ replaced with $\boxdot$ and $\Diamond$ with $\Diamonddot$
S5-schemata for $\boxtimes$ and $\Diamondtimes$	
$\boxtimes\varphi \to \boxdot\varphi$	
$\boxtimes\psi \to \boxtimes\bigcirc(\psi/\varphi)$	
$\boxtimes(\varphi \leftrightarrow \psi) \to \boxtimes(\bigcirc(\chi/\varphi) \leftrightarrow \bigcirc(\chi/\psi))$	
$t = s \to (\varphi \leftrightarrow \varphi_{t \hookrightarrow s})$	if t is not in the scope of $\boxtimes$
$E(t) \to (\forall x\varphi \to \varphi_{x \Rightarrow t})$	if x is not in the scope of $\boxtimes$
$\exists x \exists y(x = y)$	
$t = t$	
$t \neq s \to \boxdot t \neq s$	
$\forall y((\forall x(\varphi \leftrightarrow x = y)) \to y = \imath x\varphi)$	
$E(\imath x\varphi) \to \exists! x\varphi$	
$\forall x(E(x) \to \varphi) \to \forall x\varphi$	
$(\forall x\varphi \wedge \forall x\psi) \leftrightarrow \forall x(\varphi \wedge \psi)$	

Rules:

If $\vdash \varphi$ and $\vdash \varphi \to \chi$ then $\vdash \chi$	
If $\vdash \varphi^*$ then $\vdash \boxtimes\varphi$	
If $\vdash \bigcirc(\psi/\varphi)$ then $\vdash \boxtimes\bigcirc(\psi/\varphi)$	
If $\vdash \varphi \to t \neq x$ then $\vdash \neg\varphi$	where $x \notin \mathrm{free}(\varphi)$
If $\vdash \varphi \to \psi$ then $\vdash \varphi \to \forall x\psi$	where $x \notin \mathrm{free}(\varphi)$
If $\vdash \varphi \to \boxdot\psi$ then $\vdash \varphi \to \boxdot\forall x\psi$	where $x \notin \mathrm{free}(\varphi)$
If $\vdash \varphi \to \boxtimes\psi$ then $\vdash \varphi \to \boxtimes\forall x\psi$	where $x \notin \mathrm{free}(\varphi)$

References

[1] L. Åqvist. *An Introduction to Deontic logic and the Theory of Normative Systems*. Bibliopolis, Naples, 1987.

[2] H.-N. Castañeda. The paradoxes of deontic logic: The simplest solution to all of them in one fell swoop. In R. Hilpinen, editor, *New Studies in Deontic Logic*, pages 37–85. Springer, 1981.

[3] D. Chalmers. Epistemic two-dimensional semantics. *Philosophical Studies*, 118:153–226, 2004.

[4] R. Chisholm. Contrary-to-duty imperatives and deontic logic. *Analysis*, 24(2):33–36, 1963.

[5] M.J. Cresswell and G.E. Hughes. *A New Introduction to Modal Logic*. Routledge, 1996.

[6] S. Danielsson. *Preference and Obligation*. Filosofiska Färeningen, Uppsala, 1968.

[7] M. de Boer, D. Gabbay, X. Parent, and M. Slavkovic. Two dimensional standard deontic logic. *Synthese*, 187(2):623–660, 2012.

[8] J. Decew. Conditional obligation and counterfactuals. *Journal of Philosophical Logic*, 10(1):55–72, 1981.

[9] J. Van Eck. A system of temporally relative modal and deontic predicate logic and its philosophical applications. *Logique & Analyse*, 25(99):249–290, 1982.

[10] S. Frijters. *All Doctors Have an Obligation to Care for Their Patients*. PhD thesis, Ghent, 2021.

[11] S. Frijters and T. De Coninck. The Manchester twins: Conflicts between directed obligations. In F. Liu, A. Marra, P. Portner, and F. Van De Putte, editors, *Deontic Logic and Normative Systems - 15th International Conference, DEON 2020/21*, pages 166–182. College publications, 2021.

[12] M. Fusco. A two-dimensional logic for diagonalization and the a priori. *Synthese*, 198(9):8307–8322, 2021.

[13] M. Fusco and A. W. Kocurek. A two-dimensional logic for two paradoxes of deontic modality. *Review of Symbolic Logic*, 15(4):991–1022, 2022.

[14] D. M. Gabbay, A. Kurucz, F. Wolter, and M. Zakharyaschev. *Multi-dimensional Modal Logic*. Elsevier, 2003.

[15] L. Goble. Opacity and the ought-to-be. *Noûs*, pages 407–412, 1973.

[16] L. Goble. Quantified deontic logic with definite descriptions. *Logique & Analyse*, 37(147/148):239–253, 1994.

[17] L. Goble. 'Ought' and extensionality. *Noûs*, 30(3):330–355, 1996.

[18] L. Goble. Axioms for Hansson's dyadic deontic logics. *Filosofiska Notiser*, 6(1):13–61, 2019.

[19] P. S. Greenspan. Conditional oughts and hypothetical imperatives. *Journal of Philosophy*, 72(10):259–276, 1975.

[20] B. Hansson. An analysis of some deontic logics. *Noûs*, pages 373–398, 1969.

[21] J. Horty. Perspectival act utilitarianism. In Patrick Girard, Olivier Roy, and Mathieu

Marion, editors, *Dynamic Formal Epistemology*, pages 197–221. Springer Netherlands, Dordrecht, 2011.

[22] L. Humberstone. Two-dimensional adventures. *Philosophical Studies*, 118(1-2):17–65, 2004.

[23] D. Kaplan. Dthat. In P. Cole, editor, *Syntax and Semantics*, pages 221–243. Academic Press, 1978.

[24] S. Kripke. *Naming and Necessity.* Harvard University Press, Cambridge, 1980.

[25] M. Kölbel. *Truth without Objectivity.* Routledge, 2002.

[26] D. Lewis. *Counterfactuals.* Blackwell, Oxford, 1973.

[27] J. MacFarlane. *Assessment Sensitivity: Relative Truth and its Applications.* Oxford University Press, 2014.

[28] X. Parent. Why be afraid of identity? In A. Artikis, R. Craven, N. K. Cicekli, B. Sadighi, and K. Stathis, editors, *Logic Programs, Norms and Action - Essays in Honor of Marek J. Sergot on the Occasion of His 60th Birthday*, volume 7360 of *Lecture Notes in Artificial Intelligence*, pages 295–307, Heidelberg, 2012. Springer.

[29] X. Parent. Maximality vs. optimality in dyadic deontic logic. *Journal of Philosophical Logic*, 43:1101–1128, 2014.

[30] X. Parent. Completeness of Åqvist's systems E and F. *Review of Symbolic Logic*, 8(1):164–177, 2015.

[31] X. Parent. Preference-based semantics for Hansson-type dyadic deontic logic. In D. Gabbay, J. Horty, X. Parent, R. van der Meyden, and L. van der Torre, editors, *Handbook of Deontic Logic and Normative Systems*, pages 7–70. College Publications, 2021. Volume 2.

[32] D. Pichler. Extensionality of obligations in Åqvist's system F. Master's thesis, TU Wien, 2022.

[33] D. Pichler and X. Parent. Perspectival obligation and extensionality in an alethic-deontic setting. In J. Maranhão, C. Peterson, C. Straßer, and L. van der Torre, editors, *Deontic Logic and Normative Systems, DEON 2023*, pages 57–77. College Publications, 2023.

[34] H. Prakken and M. Sergot. Dyadic deontic logic and contrary-to-duty obligations. In D. Nute, editor, *Defeasible Deontic Logic*, pages 223–262. Kluwer, Dordrecht, 1997.

[35] W. Quine. Notes on existence and necessity. *Journal of Philosophy*, 40(5):113–127, 1943.

[36] N. Rescher and A. Urquhart. *Temporal Logic.* Springer Vienna, Vienna, 1971.

[37] T. Sawasaki and K. Sano. Term-sequence-dyadic deontic logic. In F. Liu, A. Marra, P. Portner, and F. Van De Putte, editors, *Deontic Logic and Normative Systems, DEON 2020/21*, pages 376–393. College publications, 2021.

[38] K. Segerberg. Two-dimensional modal logic. *Journal of Philosophical Logic*, 2(1):77–96, 1973.

[39] R. Thomason. Some completeness results for modal predicate calculi. In *Philosophical Problems in Logic*, pages 56–76. Springer, 1970.

[40] R. Thomason. Combinations of tense and modality. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume 2, pages 205–234. Springer Netherlands, Dordrecht, 2002.

[41] G.H. von Wright. Deontic logic. *Mind*, 60(237):1–15, 1951.