NO revision and NO contraction

Gregory Wheeler and Marco Alberti

CENTRIA - Artificial Intelligence Center
Department of Computer Science
Universidade Nova de Lisboa, FCT/UNL
2829-516 Caparica, Portugal
{grw,m.alberti}@fct.unl.pt

Please cite the final version published in *Minds and Machines*, 2011.

Abstract. One goal of normative multi-agent system theory is to formulate principles for normative system change that maintain the rule-like structure of norms and preserve links between norms and individual agent obligations. A central question raised by this problem is whether there is a framework for norm change that is at once specific enough to capture this rule-like behavior of norms, yet general enough to support a full battery of norm and obligation change operators. In this paper we propose an answer to this question by developing a bimodal logic for *norms* and *obligations* called *NO*. A key to our approach is that norms are treated as propositional formulas, and we provide some independent reasons for adopting this stance. Then we define norm change operations for a wide class of modal systems, including the class of *NO* systems, by constructing a class of modal revision operators that satisfy all the AGM postulates for revision, and constructing a class of modal contraction operators that satisfy all the AGM postulates for contraction. More generally, our approach yields an easily extendable framework within which to work out principles for a theory of normative system change.

1 Introduction

One goal of normative multi-agent systems theory is to formulate principles for normative system change that maintain the structure of system norms and the links between norms and agent obligations.

In simplest terms, a norm may be viewed as a rule intended to promote or inhibit agent actions within some context, whereas an obligation is a morally necessary requirement for an agent to take, or refrain from taking, some course of action. Typically, at least within this literature, a collection of norms are conceived as a set of "conditionals", each responsible for imposing an obligation on an agent given some context, and norm change is understood as the one-shot subtraction or addition of a conditional norm [3,28,10].

Even so, there is disagreement over the proper structure of norms and the correct principles for governing norm change. One approach represents conditional norms as pairs of formulas, where each coordinate of the pair is instantiated in propositional logic and the pair itself is interpreted within an input/output logic [19]. The problem of norm change then is reduced to the problem of adapting *some* of the AGM postulates [1] to an input/output logic setting. Although input/output logics capture the conditional structure of norms, defining change operators for this class of logics is problematic. To date the characterization results for contraction [3] are limited to the first six AGM contraction postulates, and even less is established for revision. These limitations are also reflected in [28], where "derogation" is defined as the norm-theoretic analogue of contraction and serves as a basis for a theory of permission.

One reason for the paucity of results on this approach is that the conditional behavior of norms within input/output logic is treated as a *metalinguistic relation*, on analogy with provability, rather than as "conditional" formulas within some object language [17]. Makinson and van der Torre [21] argue that this feature of input/output logics captures the correct structure of norms, so the difficultly input/output logic has in accommodating robust change operators is thought to be some reflection of the problem of norm change rather than an artifact of the choice in modeling language. Even so, we are not persuaded. There is little to recommend input/output logic for modeling norms in general, and much to recommend against adopting the framework for modeling norm change in particular. We return to discuss why we take this position in section 2.

Another approach to modeling norm change [10] treats norms as rules within a defeasible logic [22]. Governatori and Rotolo are interested in two types of change operations that occur within legal reasoning, namely "abrogation", which is the introduction of an exception to a general rule, and "annulment", which is the complete repeal of a rule. Their approach faces a number of technical challenges as well, primarily to do with adapting AGM postulates to a logic featuring a non-monotonic consequence operator, and also a requirement to deal with sequences of state changes rather than a one-shot operation to handle a single change. Governatori and Rotolo report a number of very interesting results about their logic, but there are similar questions to those raised before about whether the structure of norms and norm change is driving the selection of the logical framework. In particular, we question whether their commitment to the Recovery Postulate is warranted given their adoption of Defeasible Logic as the base logic for building their contraction operator. We discuss this worry in section 3.

To summarize our view from the start, we are skeptical of the claim that norm change presents a new set of theoretical problems that fall outside the scope of known techniques for representing belief change. This is to say that we see the problem of norm change to be primarily a philosophical in nature—what is a norm, and what are the principles governing how they may change—which we think can be handled by off the shelf results given an expressive enough language. But to consider whether norm change represents a new *theoretical* problem or no problem at all, it is helpful to have a flexible and general framework within which to represent various proposals about norms and the operators imagined to regulate how they may change. Providing such a framework is one of the objectives of this paper.

Here we view norms in terms of a modal logic, called *NO*, based on a standard Kripke semantics but outfitted with two distinct types of unary modal operators, a standard deontic box operator for each agent's *obligations*, and a global box operator to express *norms* of the system. We then construct both revision and contraction operators for this class of logics by proving characterization results for full AGM revision and contraction operators with respect to a general class of modal systems that includes *NO*. What will become clear is that the methodology behind our modalization of the standard AGM operators lends itself to importing a wide range of the known operators for belief change. Thus, our system NO may be viewed as both a concrete normative system and a template for how to investigate the theory of normative system change. We maintain that this class of modal systems is rich enough to handle the theoretical problem of specifying a normative system which supports change operators. The upshot is that, even if you take issue with our NO solution to norm change, you might in the end agree with us that the problem of norm change presents no new problem after all.

Since we achieve our results largely because of the additional expressive capabilities of a bi-modal language, we begin in section 2 by providing some motivations for wanting a framework that represents norms as normative assertions. In section 3 we provide the semantics for our general framework, provide an axiomatization that is strongly complete with respect to the class of all standard Kripke frames, specify the bisimulation-invariant first-order correspondent to our modal language, and construct modal AGM revision and contraction operators by translating each canonical system into its corresponding first-order fragment, perform the norm change on this first-order correspondent, then translate back to the original modal space. In section 4 we define our modal logic for normative systems, where we explain what we mean by *NO*, define *NO* AGM revision, *NO* AGM contraction, and *NO* Levi/Harper identities. We also illustrate the benefits of our *NO* class of systems by giving *NO* examples.

2 Conditional Norms and Obligations

Informally, a standard deontic logic interprets the modal formula $\mathsf{O}_i\varphi$ as "agent i is obligated to satisfy φ," where a model of the agent's obligations will contain an accessibility relation R which associates every state of the world all morally acceptable variant states in which the agent satisfies φ.

There are refinements to this scheme to consider, but notice first that standard deontic logics can express conditional obligations but not conditional norms. A conditional obligation $\mathsf{O}_i(p \to q)$ is satisfied at a state w just when, for all w' accessible from w, either p is false or q is true. A conditional norm, $p \to q$, by contrast, asserts that whenever p is satisfied then so too is q. A conditional norm interpreted within the standard model corresponds to a set of satisfiability problems—namely for all states w in model $\mathcal{M}$, $\mathcal{M}, w \Vdash (p \to q)$. So, similar to input/output logic, a conditional norm within standard deontic logic is construed as a metatheoretic notion.

To bring norms down into the object language, we view the subjects of norms to be conditional formulae of a particular form, namely those in which the antecedent (p) is a proposition specifying a context or factual precondition and the consequent ($\mathsf{O}_i q$) is some agent i's obligation to satisfy q. The modal box operator $\boxdot$ then says that a conditional formula within its scope is a norm of the system. Thus, a norm expressing that "agent i is obligated to q if precondition p" is represented by:

$$\boxdot(p \rightarrow \mathsf{O}_i q). \tag{1}$$

Equation 1 is satisfied in a model just when, for every state w in the model, either p is false or $\mathsf{O}_i q$ is true. Our main focus in this paper is the relationship between norms and obligations in general, and we will return in section 4 to specify our language for normative multi-agent systems in greater detail. But for now let us focus on the basic idea.

A conditional norm is a rule which promotes or inhibits agent actions by imposing an obligation on an agent that requires him to take or refrain from taking some course of action. To illustrate, imagine that Billy the Boy Scout comes upon an old woman at a crosswalk and recognizes at once his obligation to help the woman across the street. The *norm* here is that Boy Scouts should help old women across the street—provided that they are willing and able to cross. The norm is not that Billy should help old people across the street. However, having signed up for the Boy Scouts and, let us assume, being a Scout in good standing, Billy is aware of the norm and recognizes that membership imposes on him an obligation to help the woman across the street. Similarly, whereas Billy is obligated to help this old woman, The World Organization of the Scout Movement, Inc., has no obligation to help anyone across the road; their obligations are to remain solvent, hire and fire staff, and maintain or change Scout norms.

Here only Scouts, Inc., may change a norm, but in general either norms or obligations may change. Billy might quit the scouts and thus be absolved of his Scouting obligations. Likewise, the Scouts may decide to get out of the business of helping old people across the road, which would likewise absolve Billy of this obligation—just as it would for every other Boy Scout.

This scenario allows us to imagine several types of normative claims that we might wish to reason with or change.

1. When outdoors, a Scout should be careful with fire.
2. If currently outdoors, Billy should be careful with fire.
3. Billy is obligated to be obligated to be careful with fire whenever he is outdoors.
4. Be prepared!
5. Under no circumstances is a Scout obligated to help if he is threatened.
6. A Scout is permitted to ignore his obligation to help if he is threatened by an old person.

Sentence 1 expresses a paradigmatic conditional norm, whereas sentence 2 expresses a contingent obligation of a single agent, Billy, given that he is outdoors. In contrast to sentence 1, sentence 3 expresses a specific agent's conditional obligation. This point is an awkward one to express in natural language without a context, but sentence 3 expresses the fact that Billy has adopted the general conditional norm expressed by sentence 1. Sentence 4 is the unconditional Scout motto. Sentence 5 is a norm expressing a categorical exemption from an obligation to help others, whereas sentence 6 expresses a norm that recognizes a permissible exception. We may represent these six claims by the corresponding formulas:

1′. $\boxdot(\mathit{outdoors} \to \mathsf{O}_i(\mathit{careful}))$, for all i Boy Scouts.
2′. $\mathit{outdoors} \to \mathsf{O}_b(\mathit{careful})$.
3′. $\mathsf{O}_b(\mathit{outdoors} \to \mathsf{O}_b(\mathit{careful}))$.
4′. $\boxdot(\top \to \mathsf{O}_i(\mathit{prepared}))$, for all i Boy Scouts.
5′. $\boxdot(\mathit{threat} \to \neg\mathsf{O}_i(\mathit{help}))$, for all i Boy Scouts.
6′. $\boxdot\neg\boxdot((\mathit{elderly} \land \mathit{threat}) \to \mathsf{O}_i(\text{help}))$, for any i Boy Scouts.

Sentence 6 and formula 6′ raise a question about negated norms, which are of the general form:

7′. $\neg\boxdot(p \to \mathsf{O}q)$.

The primary reason for including negated norms in our modal language is to facilitate norm revision, but there is a natural reading of 7′ which asserts there is at least one state in the model where the precondition p is satisfied but the agent is not obligated to satisfy q. In other words, negated norms express that, somewhere in the model, there is a true formula which is a concrete counterexample to the conditional norm.

Our discussion of conditional norms is far from exhaustive; instead, our intention is to give some independent motivations for viewing norms as a type of formula rather than a metatheoretic relation. Resistance to viewing norms as a statement is understandable. Norms, unlike statements, may be respected or flouted, and they may be judged from the standpoint of other norms but are not typically evaluated as "true" or "false". This difference is what motivates some to adopt input/output logic [19,20].

Input/output logic represents a norm as an ordered pair of formulas, and a normative system is a set G of such pairs. The task for an input/output logic is to prepare information to be passed into G, and to unpack the consequences from doing so. Abstractly, the set G is a transformation device for information, and we may characterize an "output" operator "Out" by logical properties typical of consequence operators. A formula x is a "simple-minded output" of G in context a, written $x \in Out(G, a)$, if there is a set of norms $(a_1, x_1), \ldots, (a_n, x_n) \in G$ such that each $a_i \in Cn(a)$ and $x \in Cn(x_1 \land \ldots \land x_n)$, where $Cn(a) = \{a_i \mid a \models a_i\}$ is the classical semantic consequence set of a that is a set of all contexts a_i such that every model of a is also a model of a_i [19].

$Out(G, a)$ satisfies three rules: writing (a, x) for $x \in Out(G, a)$, they are strengthening input (SI), conjoining output (AND), and weakening output (WO):

– (SI): From (a, x) to (b, x), whenever $a \in Cn(b)$.
– (AND): From $(a, x), (a, y)$ to $(a, x \wedge y)$.
– (WO) From (a, x) to (a, y), whenever $y \in Cn(x)$.

Strengthening input means that if context a leads to output x, then a stronger context b also leads to x. Similarly, weakening output means that if context a leads to output x, then context a also leads to a weaker output y, $x \models y$, that is $y \in Cn(x)$.

Here the idea is that, while classical logic may be used to "process" the input and also to "unpack" the output, the operator *Out* itself does not have either an associated language or a proof theory. Even so, *Out* may enjoy structural properties commonly attributed to consequence operators. Simple-minded output is the most general characterization of *Out*, but stronger input/output operators have been studied [19], including a characterization of Poole's default system [24] and a system similar to Reiter's default logic [25].

While we agree that norms should not be thought of as extensional statements, we do not find compelling the case for treating conditional norms as a metalinguistic relation, as advocates for input/output have suggested. For one thing, norms can be negated, and the recursive structure of the language allows for nested norms, i.e., normative statements which have as a precondition or a consequence another normative statement. It appears to us an artificial constraint to preclude these options from consideration. Moreover, and more importantly, there are considerable advantages to constructing norm change operators for an intensional logic that is able to treat conditional norms as object language statements. Indeed, given the problem at hand—formulating a theory of norm change—the selection of a modeling language should be flexible enough to handle the known constraints, while avoiding imposing artificial constraints on the problem from the logic. Otherwise limitations of the modeling language may be confused for real features of the problem.

The general point is that a modeling language is a tool, and which tool you select depends on the problem you wish to solve. Whereas an important issue is the apparent non-truth-functional character of norms, this feature alone tells us little about the type of representation language we should use to model a norm. An input/output logic does it by treating the conditional as a metalinguistic operator. Our modal logic treats the conditional as a statement within an expressive modal language. Both approaches manage to respect the intensional character of conditional norms, such as it is. But, input/output logic places severe restrictions on the structure of norms, whereas the expressive capacity of our approach does not unduly constrain the structure of norms yet provides greater leverage for constructing well-defined change operators.

Finally, another reason for treating norms as metalinguistic relations rather than object language statements is a wish to distinguish the dual role that norms can play, one as a speech act and the other as a declarative statement. Jorgensen's dilemma [15,21] turns on a distinction between "norms as propositions" and "norms as acts" that create a normative imperative—or which create an obligation, to use our terms. To illustrate the difference between normative propositions and normative speech acts, the sentence "Scouts are obligated to help others" is ambiguous between an assertion of the fact that Scouts are obligated to help, and an assertion made by an authority in the act of creating that obligation. We can imagine Billy asserting the sentence "Scouts are obligated to help others" to express the fact that he is obliged to help, whereas at some point in the history of the Scouting Movement someone in authority asserted "Scouts are obligated to help others" to create that obligation for all Boy Scouts. By viewing norms and obligations as formulas, does our approach run afoul of this basic distinction, thereby failing to make room for norms as performative acts?

No. The norm change operators—the operations of adding and subtracting norms from a normative system—should be thought of as the effect of an authority's speech act creating, changing, or eliminating a norm, whereas the body of norms, and their consequences, represent properties of the norms *qua* propositions. Thus, rather than conflate the distinction animating Jorgensen's dilemma, our approach treats the distinction between normative propositions and the act of "norming" as fundamental.

3 Framework

We now turn to developing a general framework for norms and obligations called *GO*. Although conditional norms are fully expressible in the language of *GO*, neither the full language of *GO* nor many of the *GO* classes of models are suitable for representing norms and obligations. The main problem is that the general language for *GO* places no restriction on the logical form of the sub-formula φ that may appear in $\boxdot\varphi$. This means that, in addition to conditional norms, the language of *GO* also includes formulas that do not admit a plausible interpretation as norms. In section 4 we address how to restrict admissible subformulas φ within $\boxdot\varphi$ to represent conditional norms—our *NO* class solution. But in this section we focus on developing a generic norm revision machinery that applies to the whole range of canonical *GO* logics, including *NO* logic.

3.1 Modal Logic for Norms and Obligations

Let Φ be a countable set of propositional atoms and $\boxdot$ and O be unary modal box operators. Then a formula in language $\mathcal{L}^{GO}(\Phi)$ is defined recursively as follows:

– if $p \in \Phi$, then p is a $\mathcal{L}^{GO}(\Phi)$ formula;
– if ϕ and ψ are $\mathcal{L}^{GO}(\Phi)$ formulæ, then $\neg\phi$, $\phi \rightarrow \psi$, $\mathsf{O}\phi$, and $\boxdot\phi$ are $\mathcal{L}^{GO}(\Phi)$ formulæ.[1]

Definition 1 (Satisfaction). *Suppose $\mathcal{L}^{GO}(\Phi)$ is defined as above and let w be a state in a standard Kripke model $\mathcal{M} = (W, R, V)$. We define when a formula ϕ is* ***satisfied*** *(or* ***true****) in $\mathcal{M}$ at w as follows:*

[1] Although a full theory of normative systems should be developed for languages containing n O_i operators, one for each $1 \leq i \leq n$ agents in the normative system, we focus in this paper on the basic case where $n = 1$ and thus omit the subscript in the remainder.

$\mathcal{M}, w \Vdash p$ *iff* $w \in V(p)$, *where* $p \in \Phi$,
$\mathcal{M}, w \nVdash \bot$,
$\mathcal{M}, w \Vdash \neg\phi$ *iff* $\mathcal{M}, w \nVdash \phi$,
$\mathcal{M}, w \Vdash \phi \to \psi$ *iff* $\mathcal{M}, w \nVdash \phi$ *or* $\mathcal{M}, w \Vdash \psi$,
$\mathcal{M}, w \Vdash \mathsf{O}\phi$ *iff for all* $w^* \in W$, *if* Rww^* *then* $\mathcal{M}, w^* \Vdash \phi$.
$\mathcal{M}, w \Vdash \boxdot\phi$ *iff for all* $w^* \in W$, $\mathcal{M}, w^* \Vdash \phi$.

Discussion: Although $\mathcal{L}^{GO}(\Phi)$ is a polymodal grammar, both of the language's monadic modal "necessity" box operators ($\boxdot, \mathsf{O}$) are defined here in terms of a standard Kripke model. O is a standard box operator, where $\mathsf{O}\varphi$ is interpreted to express that "agent i is obligated to (satisfy) φ." $\boxdot$ is a universal box operator with a fixed interpretation across all worlds $w \in \mathcal{M}$, and $\boxdot\varphi$ is interpreted to express the generic norm φ. That said, we could have instead held the standard satisfiability conditions fixed to define both O and $\boxdot$ and interpreted each modality through its own accessibility relation within a bimodal Kripke structure, $\mathcal{M}' = (W, R, R', V)$, where R' is (trivially) $W \times W$ and the interpretation of $\boxdot$ formulas above is replaced by $\mathcal{M}', w \Vdash \boxdot\phi$ iff for all $w^* \in W$, if $R'ww^*$ then $\mathcal{M}', w^* \Vdash \phi$. Having this variant in mind will help in understanding the axiomatization of *GO*, and we will put it to use in proving our AGM characterization result in Theorem 6.

Let $\diamonddot$ abbreviate $\neg \boxdot \neg$. Now consider the following three aixioms:

$\boxdot\varphi \to \varphi$. (**T**$\cdot$)
$\diamonddot\, \varphi \to \boxdot \diamonddot\, \varphi$ (**5**$\cdot$)
$\boxdot\varphi \to \mathsf{O}\varphi$. (**inclusion**)

(**T**$\cdot$) is valid with respect to the class of reflexive frames, and expresses that φ is a norm only if φ is true; (**5**$\cdot$) is valid respect to the class of euclidean frames, and expresses that if it is permissible that φ is satisfied at a world then there is a global guarantee (a norm) that no norm mandates $\neg\varphi$; finally, (**inclusion**) expresses that if there is a norm that φ, then the agent has an obligation to satisfy φ. So, given the conditional norm schema $\boxdot(p \to \mathsf{O}q)$ of Equation 1, (**inclusion**) says that it is the particular agent i's obligation to observe the norm, $\mathsf{O}_i(p \to \mathsf{O}_i q)$, in addition to observing that whenever c holds then i is obligated to p. The (**inclusion**) axiom would ensure that (3′) follows from (1′).

Theorem 1. *The set of valid* $\mathcal{L}^{GO}$ *formulas with respect to* $\mathcal{M}$ *is axiomatized by a minimal normal bimodal logic in* O *and* $\boxdot$ *plus (***T**$\cdot$*), (***5**$\cdot$*), and (****inclusion****). Call this system GO.*

Theorem 2. *GO is strongly complete with respect to the class of all frames.*

Theorem 3. *Suppose that* S *is a set of* $\mathcal{L}^{O}(\Phi)$ *formulas—the basic modal language without* $\boxdot$. *Let* $\mathcal{F} = (W, R)$ *be the class of Kripke frames that define* S. *If logic* $O.S$ *admits a canonical model, then logic* $GO.S$ *is strongly complete with respect to* $\mathcal{F}$.

Theorem 4. *The $\boxdot$ operator is a S5 box modality.*

Discussion: Close analogues of Theorem 1, Theorem 2 and Theorem 3 are proved for the two-diamond logic K_g with an alternative (but equivalent) axiomatization [2, Chapter 7]. Theorem 3 generalizes strong completeness for any canonical normal logic we might extend by $\boxdot$, which underpins our generality claim about GO. For example, if $O.D$ corresponds to the class of normal KD systems, then $O.D$ admits a canonical model and so too does $GO.D$. Theorem 4 follows since a normal modal logic is an S5-system iff it is a $KT5$ system, and every instance of those schemata involving $\boxdot$ are theorems of GO. Finally, a remark on (**inclusion**). Recall our alternative model $\mathcal{M}' = (W, R, R', V)$. Observe that $R \subseteq R'$, so all instances of $\mathsf{P}\varphi \rightarrow \Diamond\!\!\!\cdot\ \varphi$ are valid with respect to $\mathcal{M}$ when $\mathsf{P}\varphi$ iff $\neg\mathsf{O}\neg\varphi$. But this valid schema is simply the contrapositive of (**inclusion**).

3.2 Correspondence Languages

With these results in place we turn to the business of building revision operations. We follow the strategy laid out in [7], which has been extended for polymodal frames in [30]. The basic idea is that, since the AGM postulates for revision are defined for propositional languages but we are dealing with a bimodal propositional language, we effect revision by first translating our modal revision problems into first-order logic, solve them within a simulated modal theory running inside of first-order logic, then translate the solution back into the original modal logic. This approach stands in contrast to Dynamic Epistemic Logic [29], since here we show how the operators satisfy the AGM postulates rather than stipulate operators which satisfy the postulates.

This strategy depends upon having a first-order axiomatization of the semantics of the target logic. Applied to our case, that means that the satisfiability conditions of formulas in the modal language $\mathcal{L}^{GO}$ must be managed by first-order definable frames. We ensure this by focusing on *canonical* GO systems, which are based on first-order definable classes of frames that are definable in the language $\mathcal{L}^{GO}$ by an analogue of the Goldblatt-Thomason theorem [9].

A first-order correspondence language $\mathcal{L}^1(\Phi)$ is generated from first-order variables $x, y, z, \ldots$, unary predicates $P_0, P_1, \ldots$ for each propositional atom $p_0, p_1, \ldots \in \Phi$, and binary relation symbol R. The set of propositional atoms is constant, so we omit reference to Φ in the remainder.

First-order correspondence languages vary by the conditions imposed on the binary relations, and those conditions are determined by the interpretation supplied to O and $\boxdot$ in $\mathcal{L}^{NO}$ by a standard, canonical Kripke frame. Otherwise, the translation operations are homomorphic for non-modal formulas. Define a Kripke frame $\mathcal{F} = (W, R)$. We define a first-order correspondent for $\mathcal{L}^{GO}$ and

$\mathcal{L}^1$ characterized by t:

$$\mathcal{L}^{GO} \Leftrightarrow \mathcal{L}^1$$

$$\begin{aligned}
(\bot)^t(w) &= x \neq x \\
(p)^t(w) &= P(w) \\
(\neg\varphi)^t(w) &= \neg(\varphi^t(w)) \\
(\varphi \vee \psi)^t(w) &= \varphi^t(w) \vee \psi^t(w) \\
(\mathsf{O}\varphi)^t(w) &= \forall x(R(wx) \rightarrow \varphi^t(x)). \\
(\boxdot\varphi)^t(w) &= \forall w, \varphi^t(w).
\end{aligned}$$

where $R(ab)$ abbreviates $(a, b) \in R$.

For the correspondence between $\mathcal{L}^{GO}$ and $\mathcal{L}^1$, the local translation function t is defined in terms of the unary predicates $P_i \in \mathcal{L}^1$, which are interpreted by their corresponding propositional variables $p_i \in \Phi$ as follows. $\Vdash_w^{\mathcal{M}} p^t = P(w)$ expresses that p is satisfied at world w in model $\mathcal{M}$, and this assertion is translated into first-order logic by $P(w)$. $p^t(w)$ abbreviates $\Vdash_w^{\mathcal{M}} p^t = P(w)$; $\neg(p^t(w))$ abbreviates $\nVdash_w^{\mathcal{M}} p^t$. Hence, the expression $(\varphi)^t(w)$ translates the assertion that φ is satisfied at world w within a model. To translate that φ is valid with respect to a class of models, a *global translation* function T translates the assertion that φ is satisfied at all worlds with respect to that class of models. Finally, the translation of a modal theory Σ is just the set of translations of its sentences,

$$\Sigma^t(w) = \{\varphi^t(w) \mid \varphi \in \Sigma\}.$$

Derivability within the logic GO is as one expects, but the theorems of the logic need to be clearly marked to ensure the revision operators work correctly. For instance, all instances of the axiom characterizing the class of reflexive frames, $\boxdot\varphi \rightarrow \varphi$, are theorems of the logic, but their translation is not a theorem of first-order logic. In fact, the ($\mathbf{T}\cdot$) schema is a theorem because of the properties of GO frames. So, if we provide a sound and complete axiomatization of GO expressed by a set $\mathcal{A}_{GO}$ of $\mathcal{L}^1$ formulas, then we can ensure correspondence between the modal derivability and its simulation in first-order logic.

Theorem 5 (Correspondence). *Let Σ be a modal GO theory, $\varphi \in \mathcal{L}^{GO}$, and $\mathcal{A}_{GO}$ be the first-order sound and complete axiomatization of modal logic GO. Then,*

$$\Sigma \vdash_{GO} \varphi \textit{ iff } \mathcal{A}_{GO} \cup \Sigma^t(w) \vdash \varphi^t(w).$$

Discussion: Theorem 5 says that the function t identifies the first-order modal correspondent for GO, which is proved by showing that t characterizes the bisimulation invariant fragment of first-order logic.

We next turn to the task of defining modal revision and iterated modal revision operators for GO systems.

3.3 AGM Revision

The AGM postulates [1] are widely viewed to offer minimal conditions for revision operators, rather than sufficient conditions for rational belief change. To be sure, there are controversies about whether the AGM postulates are too strong. For instances, the Recovery Postulate states that removing a target belief from a belief set and then adding that very same belief back by expansion returns you directly to the same belief set. Although Recovery is presumed to *not* lose information, it appears that some applications of Recovery would lose information since some evidential relationships between beliefs are not deductive.

To illustrate, consider the plight of our poor colleague who lives in an old house that is infested with wood worms. He has hired a well-regarded exterminator, Carlos, to treat the house in order to kill the worms, and a subsequent test has reported that the house is free of wood worms. However, although the wood-worm test is very reliable, it is not fool-proof: sometimes a test yields a false negative, in which case Carlos is obliged to drive back to the house and start again with another treatment. If our colleague were to contract his belief set by "Carlos treats the house", the replacement belief set would omit "Carlos treats the house" *along with the belief that the house tested negative for wood-worms*, since it is Carlos's treatment that is the reason for believing the test's negative outcome. However, restoring the judgment that Carlos treated the house would not return the belief that the house tested negative since this is not a logical consequence of Carlos's treatment. Hence, in this case expansion after contraction does not result in the same set of beliefs.

Curiously, Governatori and Rotolo's criticism of AGM is not that AGM is too strong but rather that it is too weak. For even though they embrace defeasible logic as a means to handle reasoning with exceptions, and set out to adapt the AGM postulates for DL's non-monotonic consequence operator, they nevertheless embrace the Recovery postulate.

> This [AGM] procedure is not satisfactory unless more sophisticated measures are added...the contraction function [they define within defeasible logic] does not offer a suitable method for modeling annulment (and, in general, norm changes), even if it satisfies all AGM postulates.

Indeed, Governatori and Rotolo's main criticism of AGM is that revision and contraction operations regulate one-shot changes to a belief set rather than sequential or iterated changes. But this remark about AGM is not unique to norm change, and various proposals have been made to handle iterated belief change [27,5]. What's more, there is a large trove of results and techniques in the literature of belief change that a theory of norm change might appeal to—Rott's two-dimensional approach to rational belief change [26], which distinguishes between static AGM-style revision from dynamic updating [16]; belief bases [13]; weakened contraction [23,26]; Ramsey conditionals [18]—if only there was a way to export these operators to a language—or, better still: a *class* of languages—rich enough to represent conditional norms. To this question our reply is: all systems *GO*.

To define AGM revision on our correspondence languages we adapt a strategy for normal monomodal logic [7] that requires (i) a sound and complete axiomatization of each classical modal system, (ii) a classical AGM revision operator.

Recall the AGM postulates for the revision operator, $*$ [1], where $K = Cn(K)$, and φ, ψ are propositional formulas with respect to the propositional language $\mathcal{L}^{PL}$:

(K∗1) *Closure*: $K * \phi$ is a belief set.
(K∗2) *Success*: $\phi \in (K * \phi)$.
(K∗3) *Inclusion*: $(K * \phi) \subseteq Cn(K \cup \{\phi\})$.
(K∗4) *Vacuity*: If $\neg\phi \notin K$, then $Cn(K \cup \{\phi\}) \subseteq (K * \phi)$.
(K∗5) *Consistency*: $(K * \phi) = \mathcal{L}^{PL}$ only if $\phi \equiv \bot$.
(K∗6) *Extensionality*: If $\phi \equiv \psi$, then $(K * \phi) \equiv (K * \psi)$.
(K∗7) *Superexpansion*: $K * (\phi \wedge \psi) \subseteq Cn((K * \phi) \cup \{\psi\})$.
(K∗8) *Subexpansion*: If $\neg\psi \notin (K * \phi)$, then $Cn((K * \phi) \cup \{\psi\}) \subseteq K * (\phi \wedge \psi)$.

The AGM postulates hold for *any* consequence operation Cn defined on a classical propositional language that includes classical consequence, satisfies the Tarski closure conditions (idempotence, inclusion, and monotony), and satisfies disjunction in the premises. Classical modal consequence is supraclassical, but obviously is not expressed within a purely propositional language. However, since *local* translations offer sufficient expressivity for revision and contraction operators, the first-order correspondent admits a propositional representation if the set of worlds $W \in \mathbb{F}$ is finite and we restrict ourselves to canonical systems [30]. To ensure that closure and premise disjunction holds, we work with equivalent alternatives to (K∗3) and (K∗4), and (K*7) and (K*8).

($\mathrm{K}*_{3,4}$) If $K \cup \{\phi\}$ is consistent, then $K * \phi = Cn(K \cup \{\phi\})$;
($\mathrm{K}*_{7,8}$) $Cn((K * \phi) \cup \{\psi\}) = K * (\phi \wedge \psi)$, when ψ is consistent with $K * \phi$.

Let $\Sigma^t(w)$ be the first-order local translation into $\mathcal{L}^1$ of a GO theory, $\phi^t(w)$ and $\psi^t(w)$ first-order local translations of $\mathcal{L}^{GO}$ formulas ϕ and ψ, and $\mathcal{A}_{GO.S}$ the first-order characterization of a canonical GO modal system $GO.S$, where S denotes a set (possibly empty) of canonical modal schemata. Then:

$$\Sigma *_{\mathsf{go.s}} \psi = \{\phi : \Sigma^t(w) * \psi^t(w) \wedge \mathcal{A}_{GO.S} \vdash \phi^t(w)\}.$$

We now show that the revision operator for the minimal GO logic, $*_{\mathsf{go}}$, satisfies the AGM postulates.

(Σ∗1): $\Sigma *_{\mathsf{go}} \phi$ is a modal theory.

Since $\Sigma^t(w) * (\phi^t(w) \wedge \mathcal{A}_{GO})$ is closed under $\vdash$ by (K∗1), then $\Sigma *_{\mathsf{go}} \phi$ is closed under $\vdash_{\mathsf{GO}}$.

(Σ∗2): $\phi \in (\Sigma *_{\mathsf{go}} \phi)$.

$(\Sigma{*}2)$ states that $\phi \in (\Sigma *_{\mathsf{go}} \phi)$. We first show that $(\Sigma{*}2)$ holds for the revision operator for the smallest GO logic, $\phi \in (\Sigma *_{\mathsf{go}} \phi)$. From $(K{*}2)$ we have $\phi^t(w) \wedge \mathcal{A}_{GO} \in \Sigma^t(w) * (\phi^t(w) \wedge \mathcal{A}_{GO}$. Since $\Sigma^t(w) * (\phi^t(w) \wedge \mathcal{A}_{GO})$ is closed under $\vdash$, by (K∗1), and $\vdash$ is reflexive, then $\Sigma^t(w) * (\phi^t(w) \wedge \mathcal{A}_{GO}) \vdash \phi^t(w)$. So, $\phi \in (\Sigma *_{\mathsf{go}} \phi)$ by $(\Sigma{*}2)$.

$(\Sigma{*}3, 4)$: If sentence ϕ is GO-consistent with Σ, then $\Sigma *_{\mathsf{go}} \phi$ is equal to the closure of $\{\Sigma \cup \{\phi\}\}$ under $\vdash_{\mathsf{GO}}$, written $C_{\mathsf{go}}(\Sigma \cup \{\phi\})$.

First we make the following two observations.
Observation 1. Recall that if Σ is an GO-consistent modal theory, then $\Sigma \nvdash_{\mathsf{GO}} \perp$ and there exists a monotone neighborhood model for Σ.
Observation 2. If $\Sigma \cup \{\phi\}$ is consistent with respect to classical modal logic GO, then $\Sigma^t(w)$ is classically consistent with respect to its translation, $\phi^t(w) \wedge \mathcal{A}_{GO}$. Since by hypothesis $\Sigma \cup \{\phi\}$ has a monotone neighborhood model, by Observation 1, there exists a classical first-order model of its translation, $\Sigma^t(w) \cup \{\phi^t(w) \wedge \mathcal{A}_{GO}\}$.
Suppose that Θ denotes the classical provability closure of the first-order translation from Observation 2, $\Sigma^t(w){*}(\phi^t(w) \wedge \mathcal{A}_{GO})$. We now show that if $\psi^t(w) \in \Theta$, then $\Sigma{*}_{\mathsf{go}}\phi \vdash \psi$.
Suppose that $C_{\mathsf{go}}(\Sigma)$ is Σ closed under $\vdash_{\mathsf{GO}}$ and $\Sigma^t(w)$ is the first-order translation of Σ. We denote the corresponding $\mathcal{A}_{GO}$-simulated closure in classical logic of the first-order translation by $Cn(\Sigma^t)$. There are two parts.

1. First, for any $\alpha \in \mathcal{A}_{GO}$, if the corresponding canonical modal formula is γ and $\gamma^t \in Cn(\Sigma^t)$, then $\gamma \in \Sigma$. To see this, notice that $C_{\mathsf{go}}(\Sigma)$ is a maximally GO-consistent set, so $\gamma \in C_{\mathsf{go}}(\Sigma)$ iff $\Sigma \vdash_{\mathsf{GO}} \gamma$.
   Proof: Suppose that $\gamma \notin \Sigma$. Then, there is a Kripke model satisfying $\Sigma \cup \{\neg\gamma\}$ and a translation of this into first-order logic. But on the first-order model for this translation $\gamma^t \notin Cn(\Sigma^t)$, which falsifies the hypothesis.
2. Second, for a closed classical theory $Cn(\Sigma^t)$ s.t. $\mathcal{A}_{GO} \subseteq Cn(\Sigma^t)$ and $\{\gamma : \gamma^t \in \Sigma^t\}$, then $\Sigma \vdash \gamma$ only if $\gamma^t \in Cn(\Sigma^t)$.
   Proof: Suppose that $\gamma^t \notin Cn(\Sigma^t)$. Then there is a model of $\Sigma^t \cup \{\neg\gamma^t\}$, so there is a Kripke model satisfying $\Sigma \cup \{\neg\gamma\}$ which falsifies the hypothesis.

This concludes the proof of $(\Sigma{*}3, 4)$.

$(\Sigma{*}5)$: $\Sigma *_{\mathsf{go}} \phi = \mathcal{L}^{GO}$ only if $\phi \equiv \perp$.

Since Σ is an GO-consistent modal theory, $\Sigma \neq \mathcal{L}^{GO}$. So $\Sigma^t(w) \neq \mathcal{L}^1$. So if $\Sigma^t(w) * \phi^t(w) = \mathcal{L}^1$, then $\phi^t(w) = \perp$; thus $\phi \equiv \perp$.

$(\Sigma{*}6)$: If $\vdash_{\mathsf{GO}} \phi \equiv \psi$, then $\Sigma *_{\mathsf{go}} \phi \equiv \Sigma *_{\mathsf{go}} \psi$.

If $\vdash_{\mathsf{GO}} \phi \equiv \psi$, then $\vdash \phi^t \wedge \mathcal{A}_{GO}) \equiv \psi^t \wedge \mathcal{A}_{GO}$. So, by (K∗6), $\Sigma * (\phi^t \wedge \mathcal{A}_{GO}) \equiv \Sigma * (\psi^t \wedge \mathcal{A}_{GO})$. Therefore, $\Sigma *_{\mathsf{go}} \phi \equiv \Sigma *_{\mathsf{go}} \psi$.

(Σ∗7, 8): $\Sigma *_{\mathsf{go}} (\phi \wedge \psi) = C_{\mathsf{go}}((\Sigma *_{\mathsf{go}} \phi) \cup \{\psi\})$, when ψ is GO-consistent with $\Sigma *_{\mathsf{go}} \phi)$.

Now we proceed in two parts.

1. $\Sigma *_{\mathsf{go}} (\phi \wedge \psi) \subseteq C_{\mathsf{go}}((\Sigma *_{\mathsf{go}} \phi) \cup \{\psi\})$: By ($\Sigma$∗1), $\Sigma *_{\mathsf{go}} (\phi \wedge \psi) = C_{\mathsf{go}}(\Sigma *_{\mathsf{go}} (\phi \wedge \psi))$. Suppose that $\gamma \in C_{\mathsf{go}}(\Sigma *_{\mathsf{go}} (\phi \wedge \psi))$. Then by the correspondence theorem $\gamma^t \in Cn(\Sigma^t * (\phi^t \wedge \psi^t \wedge \mathcal{A}_{GO}))$. So $\gamma^t \in Cn(\Sigma^t * (\phi^t \wedge \mathcal{A}_{GO}) \cup \{\psi^t\})$, by (K∗7), and $\gamma \in C_{\mathsf{go}}((\Sigma *_{\mathsf{go}} \phi) \cup \{\psi\})$, by correspondence. Since γ is an arbitrary modal formula, $\Sigma *_{\mathsf{go}} (\phi \wedge \psi) \subseteq C_{\mathsf{go}}((\Sigma *_{\mathsf{go}} \phi) \cup \{\psi\})$.
2. $C_{\mathsf{go}}((\Sigma *_{\mathsf{go}} \phi) \cup \{\psi\}) \subseteq \Sigma *_{\mathsf{go}} (\phi \wedge \psi)$: Suppose that $\gamma \in C_{\mathsf{go}}(\Sigma *_{\mathsf{go}} \phi)$. Since γ is GO-consistent with $\Sigma *_{\mathsf{go}} \phi)$, $\gamma \in C_{\mathsf{go}}((\Sigma *_{\mathsf{go}} \phi) \cup \{\psi\})$. Thus, $\gamma^t \in Cn(\Sigma^t * (\phi^t \wedge \mathcal{A}_{GO}) \cup \{\psi^t\})$, by the correspondence theorem, and $\gamma^t \in Cn(\Sigma^t * (\phi^t \wedge \psi^t \wedge \mathcal{A}_{GO}))$, by (K∗8). So, $\gamma \in C_{\mathsf{go}}(\Sigma *_{\mathsf{go}} (\phi \wedge \psi))$, by correspondence. Since γ is an arbitrary modal formula, $C_{\mathsf{go}}((\Sigma *_{\mathsf{go}} \phi) \cup \{\psi\}) \subseteq \Sigma *_{\mathsf{go}} (\phi \wedge \psi)$.

This concludes the proof of (Σ∗7, 8).

This argument establishes that the operator $*_{\mathsf{go}}$ is an AGM revision operator for the smallest GO logic, which is the base case for a more general result about canonical $GO.S$ systems.

Theorem 6. *For any canonical GO.S system, there is a corresponding operator* $*_{\mathsf{go.s}}$ *that satisfies all 8 postulates of AGM revision.*

Discussion: The main difference between the basic proof for GO and various canonical systems $GO.S$ is controlled by the set $\mathcal{A}_{GO.S}$, which contains the first-order axiomatization of canonical $GO.S$ frames. Recall the model $\mathcal{M}'$ based on the frame (W, R, R'), were R controls the interpretation of O, and R' controls the interpretation of $\boxdot$. In the basic case for GO, $\mathcal{A}_{GO}$ contains the formulas $\forall w \in W(R'ww)$ for the (**T**·) schema; $\forall u, v, w \in W((R'uv \wedge R'uw) \rightarrow R'vw)$ for (**5**·); and $\forall v, w \in W(Rvw \rightarrow R'vw)$ for (**inclusion**). For a canonical system $GO.S$, we add the first-order frame definability condition that validates the new schema(s) S. Note that these schemas are added to vary the interpretation of the O operator. For example, for the class of $GO.D$ systems, $\mathcal{A}_{GO.S} = \mathcal{A}_{GO} \cup \{\forall v \exists w : Rvw\}$.

3.4 AGM Contraction

The AGM postulates for contraction, $\dot{-}$, give the minimal conditions for regulating the removal of information from a theory.

(K$\dot{-}$1) *Closure*: $K \dot{-} \phi$ is a belief set.
(K$\dot{-}$2) *Inclusion*: $(K \dot{-} \phi) \subseteq K$.
(K$\dot{-}$3) *Vacuity*: If $\phi \notin K$, then $K \subseteq K \dot{-} \phi$.
(K$\dot{-}$4) *Success*: If $\phi \in (K \dot{-} \phi)$, then $\phi \in Cn(\emptyset)$.

(K$\dot{-}$5) *Recovery*: $K \subseteq Cn((K \dot{-} \phi) \cup \{\phi\})$.
(K$\dot{-}$6) *Extensionality*: If $\phi \equiv \psi$, then $(K \dot{-} \phi) \equiv (K \dot{-} \psi)$.
(K$\dot{-}$7) *Conjunctive overlap*: $K \dot{-} \phi \cap K \dot{-} \psi \subseteq K \dot{-} (\phi \wedge \psi)$.
(K$\dot{-}$8) *Conjunctive inclusion*: If $\phi \notin (K \dot{-} \phi)$, then $K \dot{-} (\phi \wedge \psi) \subseteq K \dot{-} \phi$.

These postulates are also satisfied by a class of canonical *GO* contraction operators; the argument is analogous to the one above for revision, but with a twist. The trick is to block modal theorems from accidental removal when performing contraction within the first-order language. To do this, the translation of a target sentence may be composed of two disjoint parts, $\phi^t = \phi^t_{safe} \wedge \phi^t_{taboo}$, where ϕ^t_{taboo} is maximal in ϕ^t with respect to $\mathcal{A}_{GO}$.[2] Then define contraction for the 'safe' component of ϕ^t:[3]

$$\Sigma \dot{-}_{\mathsf{go.s}} \psi = \{\phi : (\Sigma^t(w) \cup \mathcal{A}_{GO} \dot{-} \psi^t_{safe}(w)) \vdash \phi^t(w)\}.$$

With this construction, we have the following result.

Theorem 7. *For any canonical GO.S system, there is a corresponding operator* $\dot{-}_{\mathsf{go.s}}$ *that satisfies all 8 postulates of AGM contraction.*

Discussion: Whereas the main role played by $\mathcal{A}_{GO.S}$ within the construction for revision operators is to ensure that revision within first-order logic is properly constrained by the modal semantics for the corresponding system *GO.S*, the main role played by $\mathcal{A}_{GO.S}$ is to satisfy the success postulate ($\dot{-}$4) holds, i.e., to ensure that theorems of *GO.S* aren't accidentally removed by a contraction.

It should be clear that revision operators can be defined in terms of a primitive contraction operator and that contraction can be defined in terms of a primitive revision operator, which is to say that both the Levi identity and the Harper hold, respectively. We remark that, like Levi's construction for belief contraction, our modal Levi identity for norm change does not rely upon the (controversial) recovery postulate, ($\dot{-}$5). This illustrates that a wide variety of techniques and constructions from the theory of belief revision may be carried over and explored for use in the theory of norm change.

4 Our NO Class Logics

The problem with *GO.S* systems is that the grammar for $\mathcal{L}^{GO}$ treats the universal modality $\boxdot$ like any other box modality, so φ in $\boxdot\varphi$ may take any form. But, we are principally interested in conditional norms of the form given by Equation 1, and we also want to express negated norms and conditional norms of arbitrary modal depth, for each modality. The philosophical motivations for these properties was addressed in section 2, and the technical motivations in section 3. Let us now explain why *NO* is the answer.

Let Φ be a countable set of propositional atoms. Then we define the language $\mathcal{L}^{NO}(\Phi)$ recursively as follows:

[2] A formula ϕ^t_{taboo} is maximal in ϕ^t (with respect to $\mathcal{A}_{GO}$) iff $\phi^t \vdash \phi^t_{taboo}$ and $\mathcal{A}_{GO} \vdash \phi^t_{taboo}$, and for all ϕ^{t*}_{taboo} such that $\phi^t \vdash \phi^{t*}_{taboo}$ and $\mathcal{A}_{GO} \vdash \phi^{t*}_{taboo}$, then $\phi^{t*}_{taboo} \nvdash \phi^t_{taboo}$.
[3] Thanks here to Choh Man Teng for saving our bacon.

– if $p \in \Phi$, then p is a $\mathcal{L}^{NO}(\Phi)$ formula;
– if ϕ and ψ are $\mathcal{L}^{NO}(\Phi)$ formulæ, then $\neg\phi$, $\phi \rightarrow \psi$, and $\mathsf{O}\phi$ are $\mathcal{L}^{NO}(\Phi)$ formulæ;
– if $\theta = (\bigwedge_i \pm p_i)$, where $\pm p_i$ is either p_i or $\neg p_i$, for $p_i \in \Phi$, and ϕ is a $\mathcal{L}^{NO}(\Phi)$ formula, then $\theta \rightarrow \mathsf{O}\phi$ and $\theta \rightarrow \neg\mathsf{O}\phi$ are *norm arguments*
– if ϕ^* and ψ^* are norm arguments, then $\phi^* \rightarrow \psi^*$ and $\boxdot\phi^*$ are norm arguments
– if ϕ^* is a norm argument, then $\boxdot\phi^*$ is a $\mathcal{L}^{NO}(\Phi)$ formula.

Discussion: The satisfiability conditions for $\mathcal{L}^{NO}$ are the same as for system *GO*, and all metatheoretic results we've shown for *GO* hold for *NO*. In essence, we have created canonical subsystems of *GO* simply by a restriction on the modal grammar: ϕ and ψ are arbitrary formulas, whereas ϕ^* and ψ^* are restricted formulas. Notice that nested conditionals are necessary for *NO* to remain normal. However, allowing the more general $\mathsf{O}\phi$ rather than $\mathsf{O}_j\theta$, for arbitrary modal depth j, is optional. With the more general grammar, we can interpret sentences that express a conditional norm which imparts an obligation to an agent to maintain a conditional norm. Such constructions might be helpful for representing democratic institutions which, unlike the Boy Scouts, allow member agents to govern themselves. Much more machinery would be necessary to model that sort of change behavior, but a more general language like $\mathcal{L}^{NO}$ is a start.

As should be clear, *NO* normative systems have *NO* revision operators and *NO* contraction operators:

Theorem 8. *For any canonical NO.S system,*

1. *there is a corresponding operator $*_{\mathsf{no.s}}$ that satisfies all 8 postulates of AGM revision, and*
2. *there is a corresponding operator $\dot{-}_{\mathsf{no.s}}$ that satisfies all 8 postulates of AGM contraction.*

A cherished principle for obligations is that 'ought implies can', which is expressed by the (D) schema, $\mathsf{O}\varphi \rightarrow \mathsf{P}\varphi$. We happily acknowledge this tradition with a *NO.D*, but stand behind our more general answer, *NO*, which we shall exercise throughout the following example.

Posh Pensioners: Adapting an example from [3], suppose there are two community norms: the state (s) is obligated to provide free health insurance (*insurance*) to the low-income agents (*poor*), and the state is obligated to provide free health insurance to pensioners (*elderly*). A *NO* system Σ for this example will include the following two sentences,

$$\boxdot(\mathit{elderly} \rightarrow \mathsf{O}_s(\mathit{insurance})), \tag{2}$$

$$\boxdot(\mathit{poor} \rightarrow \mathsf{O}_s(\mathit{insurance})). \tag{3}$$

There are several theorems of Σ, such as

$$\boxdot(\mathit{elderly} \wedge \mathit{poor} \rightarrow \mathsf{O}_s(\mathit{insurance})), \tag{4}$$

and likewise there are different sorts of contingent state descriptions that may satisfy the preconditions for a norm. For instance, an agent may be in a state w' where not having an income suffices to be poor, i.e., $\mathcal{M}, w' \Vdash \neg \mathit{income} \wedge (\neg \mathit{income} \rightarrow \mathit{poor})$. Then, it is easy to verify that $\mathsf{O}_s(\mathit{insurance})$ is satisfied at w', too.

Now imagine that a scandal hits the newspapers: 'State Subsidizing Posh Pensioners!', which causes the community to revise its declared norms in order to block the state from giving well-off pensioners free health insurance,

$$\Sigma *_{\mathsf{no}} \boxdot(\mathit{elderly} \wedge \neg \mathit{poor} \rightarrow \mathsf{O}_s(\neg \mathit{insurance})). \tag{5}$$

Performing the revision in (5) exposes a conflict with (2), and every rational possibility for (5) will involve removing (2) from Σ. Of course, to yield a unique result, we would need to impose an extra-logical structure on formulas constituting a normative system similar to techniques devised within the belief revision literature [8,11,27]. However, as our example hints at, *NO* norm revision operators ensure that no norm inconsistent with the target norm of a revision will appear in a candidate for rational norm change.

5 Closing Remarks

System *NO* is a normal modal logic, and there have long been doubts about whether normal modal systems correctly model moral obligations [4]. The primary concern is the strong distribution properties of the box operator over conjunction, which guarantee that there are no conflicting obligations: it is a theorem of any normal deontic logic that there are no moral dilemmas, which strikes many philosophers as too strong. That said, it is less clear whether this is a liability for modeling institutional norms. Nevertheless, there are results which bring AGM revision to supplemental neighborhood models [30], and the corresponding class monotone modal logics are ideal for modeling conflicting obligations.[4]

The logic based solely on $\boxdot$ is normal $S5$, which is decidable, and several normal modal logics are decidable, but unfortunately a *GO* logic composed of decidable components is not necessarily decidable. Even so, if a decidability proof for a normal $O.S$ system is established, via filtrations, which yields the finite model property, then *that* logic can be lifted to a decidable $GO.S$ system [2, p. 418]. Similarly, whereas the complexity of the satisfiability problem in normal modal logic O is PSPACE-complete, in *GO* the satisfiability problem is EXPTIME-complete. A line for future investigations is to explore the complexity

[4] Notice that, like normal modal logics but unlike classical (monotone) modal logics, the most general input/output logic, "simple-minded output", satisfies adjunction as well via the AND rule.

for the *NO* class of systems. One idea in particular would be to look at the effect of restricting the $S5$ modality to universal modal horn clauses.

Another topic for future exploration concerns implementations. An obvious problem is: given an initial normative theory and a narrative of required modifications (contraction, revisions, expansions), find the current normative state of the system. A possible implementation strategy would be to perform the translation of the initial normative theory and the norms to be contracted/revised into the modal first-order correspondence language as described in Sect. 3.2. The implementation of the actual revision step can rely on AGM revision systems for first order theories, such as [6]. If we assume a finite number of possible worlds, and restrict ourselves to canonical modal systems, the translated theory is propositional, possibly escaping the decidability problems faced by the aforementioned system. The result would then be translated back into a *NO* theory.

Finally, it is natural to explore iterated norm revision, and we can do this by adapting the basic translation techniques here to the DP axioms for iterated norm change [5,14]. Here the issue of entrenchment that we alluded to in section 4 becomes a central issue and, since a corollary of the DP axioms is a representation theorem for Spohn's negative ranking functions [27], having an iterated change operator for *NO* systems would provide a qualitative bridge to probability logics [12], which is an additional front to explore.

Acknowledgements: This research was supported in part by award LogiCCC/0001/2007, project DiFoS, from the *European Science Foundation*. Thanks to Erich Rast, Choh Man Teng, and three annonymous referees for very helpful comments on earlier drafts.

References

1. Carlos Alchourrón, Peter Gärdenfors, and David C. Makinson. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50(2):510–530, 1985.
2. Patrick Blackburn, Maarten de Rijke, and Yde Venema. *Modal logic*. Cambridge University Press, New York, 2001.
3. Guido Boella, Gabriella Pigozzi, and Leendert van der Torre. Normative framework for normative system change. In Decker, Sichman, Sierra, and Castelfranchi, editors, *Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2009)*, pages 169–176, 2009.
4. Brian Chellas. *Modal Logic*. Cambridge University Press, Cambridge, 1980.
5. Adnan Darwiche and Judea Pearl. On the logic of iterated belief revision. *Artificial Intelligence*, 89:1–29, 1997.
6. Simon Dixon and Wayne Wobcke. The implementation of a first-order logic AGM belief revision system. In *Proceedings of the Fifth IEEE International Conference on Tools in Artificial Intelligence*, pages 40–47. IEEE Computer Society Press, 1993.
7. Dov Gabbay, Odinaldo Rodrigues, and Alessandra Russo. Belief revision in non-classical logics. *The Review of Symbolic Logic*, 1(3):267–304, 2008.

8. Peter Gärdenfors and David C. Makinson. Revisions of knowledge systems using epistemic entrenchment. In *The 2nd Conference on Theoretical Aspects of Reasoning about Knowledge (TARK)*, pages 83–96, 1988.
9. Valentin Goranko and Solomon Passy. Using the universal modality: Gains and questions. *Journal of Logic and Computation*, 2:5–30, 1992.
10. Guido Governatori and Antonino Rotolo. Changing legal systems: Legal abrogations and annulments in defeasible logic. *Logic Journal of the IGPL*, 2010.
11. Adam Grove. Two modellings for theory change. *Journal of Philosophical Logic*, 17:157–170, 1988.
12. Rolf Haenni, Jan-Willem Romeyn, Gregory Wheeler, and Jon Williamson. *Probabilistic Logic and Probabilistic Networks*. Synthese Library. Springer, Dordrecht, 2010.
13. Sven Ove Hansson. *A Textbook of Belief Dynamics: Theory Change and Database Updating*. Kluwer Academic Publishers, 1999.
14. Yi Jin and Michael Thielscher. Iterated belief revision, revised. *Artificial Intelligence*, 171(1):1–18, 2007.
15. Jorgen Jorgensen. Imperatives and logic. *Erkenntnis*, 7:288–296, 1937.
16. Hirofumi Katsuno and Alberto Mendelzon. On the difference between updating a knowledge base and revising it. In *The 2nd International Conference on the Principles of Knowledge Representation and Reasoning (KR 1991)*, pages 387–394, 1991.
17. Henry E. Kyburg, Jr., Choh Man Teng, and Gregory Wheeler. Conditionals and consequences. *Journal of Applied Logic*, 5(4):638–650, 2007.
18. Isaac Levi. *Mild Contraction*. Clarendon Press, Oxford, 2004.
19. David C. Makinson and Leendert van der Torre. Input-output logics. *Journal of Philosophical Logic*, 30(2):155–185, 2000.
20. David C. Makinson and Leendert van der Torre. Constraints for input/output logics. *Journal of Philosophical Logic*, 30:155–185, 2001.
21. David C. Makinson and Leendert van der Torre. What is input/output logic? input/output logic, constraints, permissions. In Guido Boella, Leendert van der Torre, and Harko Verhagen, editors, *Normative Multi-agent Systems*, number 07122 in Dagstuhl Seminar Proceedings, Dagstuhl, Germany, 2007. Internationales Begegnungs und Forschungszentrum für Informatik (IBFI).
22. Donald Nute. Defeasible logic. In Dov Gabbay, C. Hogger, and J. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 3. Oxford University Press, New York, 1994.
23. Maurice Pagnucco and Hans Rott. Severe withdrawal—and recovery. *Journal of Philosophical Logic*, 28:501–547, 1999. (Re-printed with corrections to publisher's errors in February, 2000.).
24. David Poole. A logical framework for default reasoning. *Artificial Intelligence*, 36:27–47, 1988.
25. Ray Reiter. A logic for default reasoning. *Artificial Intelligence*, 13:81–132, 1980.
26. Hans Rott. *Change, Choice and Inference*. Oxford University Press, Oxford, 2001.
27. Wolfgang Spohn. Ordinal conditional functions: A dynamic theory of epistemic states. In William L. Harper and Brian Skyrms, editors, *Causation in Decision, Belief Change and Statistics*, volume 2, pages 105–134. Reidel, 1987.
28. Audun Stolpe. A theory of permission based on the notion of derogation. *Journal of Applied Logic*, 8:97–113, 2010.
29. Hans van Ditmarsch, Wiebe van der Hoek, and Barteld Kooi. *Dynamic Epistemic Logic*. Synthese Library. Springer, 2008.

30. Gregory Wheeler. AGM belief revision in monotone modal logics. In Ed Clarke and Andrei Voronkov, editors, *International Conference on Logic for Programming, Artificial Intelligence, and Reasoning (LPAR-16)*, Short paper proceedings, Dakar, Senegal, 2010.