Computing Default Extensions by Reductions on O^R

Espen H. Lian and **Arild Waaler**
Department of Informatics
University of Oslo
Norway
{elian,arild}@ifi.uio.no

Abstract

Based on a set of simple logical equivalences we define a rewriting procedure that computes extensions in the propositional fragment of the logic of O^R introduced by Lakemeyer and Levesque. This logic is capable of representing default logic with the advantage of itself being monotonic, with a clearly defined semantics and a separation of the object level and the meta level. The procedure prepares the ground for efficient implementations as it clearly separates the SAT-solving part of the reasoning problem from the modal aspects that are specifically caused by defaults. We sketch an extension of the logic to cover confidence levels and show that the resulting system can accommodate ordered default theories with a prescriptive interpretation of preference between defaults.

Introduction

In (Lakemeyer and Levesque 2005) a new logic of only-knowing is introduced which allows a faithful encoding of default logic. A default theory can be encoded as a formula of the form $O^R\varphi$, with roughly the same size as the default theory, and whose models exactly match the extensions of the default theory. For the propositional fragment the authors state a modal reduction theorem to the effect that a formula $O^R\varphi$ is logically equivalent to a disjunction $O\varphi_1 \vee \cdots \vee O\varphi_n$, where each φ_k is a propositional formula. Because each such disjunct $O\varphi_k$ has a unique model, it is possible, within the logic itself, to break down a formula $O^R\varphi$ into a form from which one can directly exhibit its models.

As an example, consider a simple supernormal default theory with two extensions:

$$(\{\neg(p \wedge q)\}, \{\tfrac{:p}{p}, \tfrac{:q}{q}\}).$$

To determine the set of extensions "the O^R-way", three steps must be carried out. The first step is to represent the default theory as a formula of the form $O^R\varphi$. Under the Konolige-style translation introduced in (Lakemeyer and Levesque 2005) the example above receives the representation

$$O^R(\,\neg(p \wedge q) \wedge (Mp \supset p) \wedge (Mq \supset q)\,)$$

where M is a possibility operator further discussed in the following sections. The second step is to carry out an equivalence-preserving reduction of the O^R-formula to a disjunction of modalized propositional formulae of the form $O\varphi_k$. The O^R-formula in the example reduces to $Op \vee Oq$. The third step is to determine the set of extensions of the default theory from the simpler formula obtained in the second step. This task is trivial, since each disjunct has a unique model. In our example, $Op \vee Oq$ represents two distinct extensions: The extension corresponding to Op is the set of consequences of p, whereas Oq corresponds to the set of consequences of q.

The logic of O^R builds on the contribution to only-knowing in (Levesque 1990) where the logic of the "All I Know"-operator O is first introduced. Like the logic of O^R, the original logic of O also admits a reduction theorem for the propositional fragment, stating that a formula $O\varphi$ is equivalent to a disjunction of modalized propositional formulae of the form $O\varphi_k$. Whereas O^R is designed for the representation of default theories, the logic of O allows a smooth representation of an autoepistemic theory. And as with default logic, it is possible to directly pick out the stable expansions of the autoepistemic theory from the equivalence $O\varphi \equiv O\varphi_1 \vee \cdots \vee O\varphi_n$, because each $O\varphi_k$ corresponds to a stable expansion. Default logic and autoepistemic logic do not differ in the way they treat the example discussed above, hence both the O^R-representation of the default theory and the O-representation of the corresponding autoepistemic theory are equivalent to $Op \vee Oq$.

The logics of only-knowing are themselves monotonic and have a clear separation between object level and meta level concepts, which is arguably a great conceptual advantage compared to the fixed-point definitions of extensions in both default and autoepistemic logics. Only-knowing logics have, moreover, a standard Kripke semantics. Encodings of default and autoepistemic theories into only-knowing logics thus provide the non-monotonic formalisms with formal semantics and conceptual clarity.

But the translations provide more than just semantics, they also provide another *model for computation*. By translating non-monotonic formalisms into only-knowing logics the problem of determining expansions is recast into the problem of finding a proof in the only-knowing logic of the equivalence of, say $O\varphi$, with $O\varphi_1 \vee \cdots \vee O\varphi_n$ for an appropriate n; for computing default logic extensions this is what we refer to as "the second step" above. The existence of a logical equivalence of this sort is guaranteed by the Modal Reduction Theorem, and algorithms for determining the equivalence can be extracted from proofs of that theorem.

There are three different proofs of the Modal Reduction Theorem for the original logic of O. The idea behind one proof is to take the set Γ of subformulae of $O\varphi$ and use this as a filtration set for the canonical model (using standard techniques from modal logic). Algorithmically, this boils down to computing all the maximal consistent subsets of Γ and use these to build models. This proof was introduced in (Waaler 1994), refined in (Segerberg 1995) and published in (Waaler et al. 2007). An implementation of this method can take advantage of the tableau method for only-knowing logics in (Rosati 2001), more precisely the tableau method may be used to efficiently check if a given subset of Γ is maximally consistent. Another proof, published in (Levesque and Lakemeyer 2001), is based on the idea of enumerating all ways of valuating modal atoms and use this to gradually approximate the models. From a computational point of view, the two methods amount to roughly the same kinds of computation. The two above-mentioned methods may be compared to a truth-table method for propositional logic: Traverse all interpretations and select those that are models of the formula at hand.

The third proof, also in (Waaler et al. 2007), is reminiscent of the proof in (Levesque and Lakemeyer 2001), but approximates the models in a more careful way. If the former algorithms are reminiscent of a truth-table method, the algorithm behind the third proof is reminiscent of a tableau method: Use information in the formula to constrain the set of potential models as much as possible. In worst-case scenarios, the two methods are of course equally bad, but the latter method is computationally superior in virtually any other situation.

The main contribution of this paper is to generalize the latter procedure to the logic of O^R. Because the procedure is what is needed to carry out the second step in the "O^R-way" of computing default extensions, as explained above, we thereby provide a new method for computing default extensions over propositional logic.

We do this by introducing a rewriting system, where each rewrite rule reflects a logical equivalence in O^R-logic. Presenting the procedure in this way provides us with a *calculus* for computing default extensions. The proposed calculus has a clear operational semantics, and it is, we believe, easy to use. An advantage of rewriting systems is that they are well understood; compared to more high-level (pseudo-code) algorithmic specifications they are easier to reason about and may be implemented more directly.

The calculus that we present is a formal system that is just strong enough for establishing the Modal Reduction Theorem: It is sound and complete for reductions of O^R-formulae into disjunctions of $O\varphi_k$'s of the appropriate type. It is, however, not complete for the logic of O^R itself. From the point of view of computing default extensions this is harmless, because only a subset of the logic of O^R is actually needed for the Modal Reduction Theorem. Our approach to formalizing default logic is complementary to the approach in (Lakemeyer and Levesque 2006), in which it is the logic of O^R that is axiomatized. Although this is, from the point of view of theoremhood, a stronger system than the rewriting system that we propose, it gives an indirect route to the Modal Reduction Theorem. The system in (Lakemeyer and Levesque 2006) is formulated as a Hilbert-style axiom system, which is natural given the author's focus on axiomatization. Although axiom systems of this kind give logical characterizations that are simple in terms of number of axioms and inference rules, they are of course not equally appropriate as bases for implementation, simply because their formal proofs do not enjoy the subformula property.

The rewrite rules introduced in this paper exhibit a sharp separation of the SAT-solving component and the modal-logic component. This will, we think, make it easier to adopt recent developments in SAT-solving technologies for use in default-logic applications. The treatment of modalities in terms of formula rewriting captures the part of the problem of computing default extension known as conflict resolution, and the procedure that we propose presents a solution to conflict resolution in a concise way. We thus believe that the rewriting system proposed in this paper can form the basis for efficient implementation, although empirical evidence for or against this claim remains to be established. At the time of writing, we do not know how this procedure for computing defaults compares, in terms of efficiency, to methods that are not based on translation into O^R (further addressed in the Conclusion).

In the first part of the paper we present a logic of O^R with essentially the same semantics as the one presented in (Lakemeyer and Levesque 2005). However, to facilitate the formulation of the rewrite rules, we slightly modify the syntax, most importantly by introducing a new modal operator to express minimality constraints, and then establish the Modal Reduction Theorem constructively.

Conflicts among defaults may be avoided by adding a partial order among defaults. If the order relation is used to constrain the generation of extensions, i.e. if the partial order is interpreted prescriptively (Delgrande and Schaub 2000), it may significantly prune the search space. In the last part of the paper, we add confidence layers to the logic of O^R along the lines of (Waaler et al. 2007) and sketch how to encode ordered default theories into this logic with a prescriptive interpretation of defaults. This builds on the encoding of this kind of default theories into a standard logic of O with confidence layers in (Engan et al. 2005). A shortcoming with the encoding in their work is that the O-representation of the default theory effectively enumerates all possible ways of constructing extensions, hence it is hopelessly intractable from the point of view of computation. The encoding sketched in this paper remedies the situation by mapping defaults into the context of O^R rather than O.

Syntax and Semantics

Since the proof system introduced in this paper is a rewriting system, it is natural to include a rich set of logical symbols in the formal language (rather than introducing as few symbols as possible and define the rest). The propositional connectives include the constants $\bot$ and $\top$ and usual symbols for negation, disjunction and conjunction. The conditional $\supset$ and biconditional $\equiv$ are taken as defined connectives. The set of propositional *objective* formulae is then defined over a finite set of propositional variables in the usual way. If $\varphi_1, \ldots, \varphi_n$ are propositional, $\mathsf{SAT}\{\varphi_1, \ldots, \varphi_n\}$ is the statement that $\varphi_1 \land \cdots \land \varphi_n$ is propositionally satisfiable.

There are in total six unary modal operators: B (belief), C (co-belief)[1], M (possibility), O (only knowing), O^R and $\boxdot$. M and O^R were introduced in (Lakemeyer and Levesque 2005); M is a possibility operator that may, or may not, be dual of B. It should not be confused with $\neg B \neg$, which is always the dual of B (and for which one could invent a defined symbol). O^R is an only-knowing operator that is stronger than O and capable of representing Reiter-style defaults. $\boxdot$ is a necessity operator introduced in this paper to express minimality constraints on models; $\diamonddot\varphi$ is $\neg\boxdot\neg\varphi$.

The set of formulae is generated from the propositional variables, propositional connectives and modal operators, with the following provisos. Firstly, in the formation of $\boxdot\varphi$, φ must be completely modalized (i.e. all propositional variables must occur within the scope of a modal operator). Secondly, $\boxdot\varphi$ and $O^R\varphi$ are not allowed to occur within the scope of any modal operator. Thus $\boxdot O^R\varphi$ is not a formula, neither is $\boxdot p$ for a propositional variable p.

Following the literature on only-knowing we call propositional formulae *objective* and completely modalized formulae *subjective*. A formula is *M-free* if it does not contain M. It is *M-basic* if it is subjective and only contains the modality M. It is *prime* if it is subjective and contains no nested modalities. A formula is a *modal atom* if it is of the form $B\varphi$, $C\varphi$ or $M\varphi$. A modal *literal* is a modal atom or its negation.

A default theory is a tuple (W, D), where W is a finite set of objective formulae and D is a finite set of defaults. The default $\alpha : \beta \,/\, \gamma$ is represented by its Konolige translation $B\alpha \land M\beta \supset \gamma$. If α or β are $\top$, i.e. the default is prerequisite-free or justification-free resp., we drop that conjunct. Hence $\top : \beta \,/\, \gamma$ translates to $M\beta \supset \gamma$, while $\alpha : \top \,/\, \gamma$ translates to $B\alpha \supset \gamma$. To translate the whole default theory, take the conjunction of all formulae in W and of the translations of the defaults in D, and put this conjunction in the context of O^R. We may define an autoepistemic translation of a default theory into only-knowing logic in essentially the same way as the translation into O^R-logic, except that that autoepistemic translation uses O instead of O^R and $\neg B \neg$ instead of M.

Example 1. *The theory $(\emptyset, \{p : \top \,/\, p\})$ has as its unique extension the set of all tautologies. The representation of the default theory is $O^R(Bp \supset p)$. The corresponding autoepistemic translation is $O(Bp \supset p)$, which has an additional autoepistemic expansion: The set of formulae following from p.*

Relative to the universal set $\mathcal{U}$ of all propositional valuations, a *model* is defined as a tuple (U, V) such that $V \subseteq U$ and $U \subseteq \mathcal{U}$. The $\boxdot$ modality quantifies over models by means of a binary relation $>$. If $\mathcal{M} = (U, V)$, $\mathcal{M}' > \mathcal{M}$ iff $\mathcal{M}' = (U', U)$ and U is a *proper* subset of U'. Note that $>$ is not an order, as it is not transitive; in fact it has the following property: if $\mathcal{M}_1 > \mathcal{M}_2$ and $\mathcal{M}_2 > \mathcal{M}_3$, then $\mathcal{M}_1 \not> \mathcal{M}_3$.

Example 2. *Let $\{a, b, c\} \subseteq \mathcal{U}$. Then*

$$(\{a, b, c\}, \{a, b\}) > (\{a, b\}, \{a\}) > (\{a\}, \{a\})$$

but $(\{a, b, c\}, \{a, b\}) \not> (\{a\}, \{a\})$, as $\{a, b\} \neq \{a\}$.

Truth is defined for a model relative to each point $x \in \mathcal{U}$. If φ is a propositional variable, $\mathcal{M} \models_x \varphi$ iff the valuation x makes φ true. Connectives are taken as truth functions in the usual way. In the following definition of truth conditions for modalities, $\mathcal{M} = (U, V)$:

- $\mathcal{M} \models_x B\varphi$ iff $\mathcal{M} \models_y \varphi$ for each $y \in U$
- $\mathcal{M} \models_x C\varphi$ iff $\mathcal{M} \models_y \varphi$ for each $y \in \mathcal{U} \setminus U$
- $\mathcal{M} \models_x M\varphi$ iff $\mathcal{M} \models_y \varphi$ for a $y \in V$
- $\mathcal{M} \models_x O\varphi$ iff for all $y \in \mathcal{U}$, $\mathcal{M} \models_y \varphi$ iff $y \in U$
- $\mathcal{M} \models_x \boxdot\varphi$ iff $\mathcal{M}' \models_x \varphi$ for every $\mathcal{M}' > \mathcal{M}$;
- $\mathcal{M} \models_x O^R\varphi$ iff $\mathcal{M} \models_x O\varphi$ and there is no $\mathcal{M}' > \mathcal{M}$ such that $\mathcal{M}' \models_x O\varphi$.

We write $\mathcal{M} \models \varphi$ if $\mathcal{M} \models_x \varphi$ for each $x \in \mathcal{U}$. Relative to a model $\mathcal{M}$, $\|\varphi\|$ denotes the set of points x in $\mathcal{U}$ such that $\mathcal{M} \models_x \varphi$. Note that if φ is objective, $\|\varphi\|$ is given independently of $\mathcal{M}$, as it only depends on the points in $\mathcal{U}$. Also note that in the clauses that define truth for the modal operators, the point x plays no active role in the definition. When φ is subjective, it is immediate that $\mathcal{M} \models_x \varphi$ iff $\mathcal{M} \models \varphi$, i.e. we can safely skip the reference to the point x. This is also the reason why the following observation holds.

Lemma 3. *If φ is subjective and $\mathcal{M}$ is any model, either $\mathcal{M} \models \varphi \equiv \top$ or $\mathcal{M} \models \varphi \equiv \bot$.*

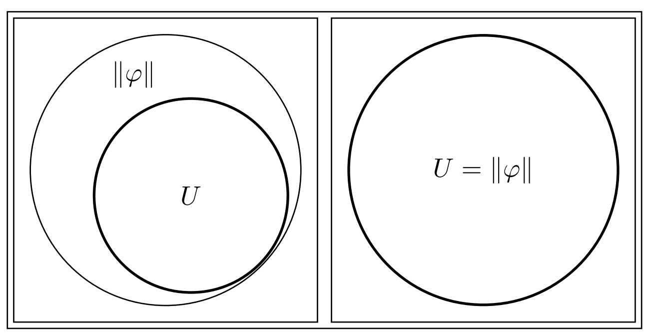

Figure 1: Left: $\neg O\varphi \land \diamonddot O\varphi$. Right: $O\varphi \land \neg\diamonddot O\varphi$.

Example 4. *Fig. 1 illustrates the truth conditions relative to a model $\mathcal{M} = (U, V)$ for an arbitrary $V \subseteq U$ and an objective φ.*

- *If $\mathcal{M}$ is the model to the left, $U \subset \|\varphi\|$. Then $C\neg\varphi$ does not hold in $\mathcal{M}$, hence neither does $O\varphi$. $O\varphi$ is, however, true in $\mathcal{M}' = (\|\varphi\|, U)$, and since $\mathcal{M}' > \mathcal{M}$, $\diamonddot O\varphi$ is true in $\mathcal{M}$.*

[1] The notion of co-belief is discussed at length in Sect. 3 of (Waaler et al. 2007).

- *If $\mathcal{M}$ is the model to the right, $U = \|\varphi\|$, in which case $O\varphi$ is true. But as there is no $\mathcal{M}' > \mathcal{M}$ that makes $O\varphi$ true, $\diamond O\varphi$ is not true. Hence $O^R\varphi$ holds in $\mathcal{M}$.*

A formula φ is *strongly valid*, written $\models\!\!\models \varphi$, if $\mathcal{M} \models \varphi$ for every model $\mathcal{M}$. There is also a weaker notion of validity, which is the notion of validity that we are primarily interested in. It is defined relative to the set of *weak models*: (U, V) is a weak model if $U = V$. φ is *valid*, written $\models \varphi$, if $\mathcal{M} \models \varphi$ for every weak model $\mathcal{M}$.

Example 5. *Let $\mathcal{M}_\top = (\mathcal{U}, \|p\|)$ and $\mathcal{M}_p = (\|p\|, \|p\|)$. If φ is $Bp \supset p$, then*

- *$\mathcal{M}_\top$ is a model of $O\varphi$ since Bp is false at every point;*
- *$\mathcal{M}_\top$ is a model of $O^R\varphi$ since there can be no $\mathcal{M}' > \mathcal{M}_\top$;*
- *$\mathcal{M}_p$ is a model of $O\varphi$ since $\|\varphi\| = \|Bp\| = \|p\|$;*
- *$\mathcal{M}_p$ is not a model of $O^R\varphi$ since $\mathcal{M}_\top$ is a model of $O\varphi$ and $\mathcal{M}_\top > \mathcal{M}_p$.*

$(\mathcal{U},\mathcal{U})$ is the only weak model of $O^R\varphi$, hence $O^R\varphi \equiv O\top$ is valid. $O\varphi$ has two weak models: $(\mathcal{U},\mathcal{U})$ and $(\|p\|, \|p\|)$. Hence $O\varphi \equiv (O\top \vee Op)$ is valid. These validities reflect the default and autoepistemic extensions of the theories in Example 1.

Clearly, strong validity implies validity, but not conversely. For formulae that do not contain M, the two notions coincide.

Lemma 6. *If φ is M-free, then $\models \varphi$ iff $\models\!\!\models \varphi$.*

A weak model corresponds directly to a model for the logic of O in (Levesque 1990). For formulae that do not contain M, O^R or $\boxdot$, the set of models is essentially the same as the set of models defined for the logic of O in (Levesque 1990).

For the M-free fragment of the language, the weak models of $O^R\varphi$ are exactly the models (U, U) with the largest belief state U that satisfy $O\varphi$. As pointed out in (Lakemeyer and Levesque 2005), these models correspond to Konolige-type minimization of the weak models of $O\varphi$, which in their system can be syntactically expressed by the O^K-modality. The function that O^K serves in their axiomatization is in our formulation taken over by $\boxdot$.

As explained in (Lakemeyer and Levesque 2005), the motivation behind the M-operator is that the possibility operator implicit in default theories is not the dual of the corresponding belief modality. The models that we in the end are interested in are the weak ones, in which M and B are duals. However, a number of inferences require that we do not limit ourselves to weak models. One can only in certain limited cases substitute $\neg B \neg$ for M, and the rewriting system in the next section is carefully designed to rewrite the input formula as much as required for substitutions of this kind to hold.

The $V \subseteq U$ condition on a model (U, V), which is required by the definition above, is not imposed on the models in (Lakemeyer and Levesque 2005). Whether or not this condition is imposed has no effect on the set of valid formulae. It does, however, affect the set of strongly valid formulae, and from the point of view of formula rewriting, it is desirable that the set of strongly valid formulae is as large as possible. The condition $V \subseteq U$ makes $M\varphi \supset \neg B\neg\varphi$ strongly valid. The converse implication is valid but not strongly. A more general result, which is essential for the rewriting system, is given in the following lemma.

Lemma 7. *Let φ and ψ be objective.*

1. *If not $\mathsf{SAT}\{\varphi, \psi\}$, then $\models\!\!\models O\varphi \supset \neg M\psi$.*
2. *If $\mathsf{SAT}\{\varphi, \psi\}$, then $\models O\varphi \supset M\psi$.*

Proof. 1. The assumption that not $\mathsf{SAT}\{\varphi, \psi\}$ implies that $\|\varphi\| \subseteq \|\neg\psi\|$. Let $(U, V) \models O\varphi$. Then $U = \|\varphi\|$. As $V \subseteq U$ in any model, $V \subseteq \|\neg\psi\|$, i.e. $(U, V) \models \neg M\psi$. 2. The assumption that $\mathsf{SAT}\{\varphi, \psi\}$, implies that there is a point x in $\|\phi\| \cap \|\psi\|$. Let $(U, V) \models O\varphi$. Then $U = \|\varphi\|$. As $U \subseteq V$ in any weak model, $x \in V$. Hence $(U, V) \models M\psi$. □

From the point of view of formula rewriting, the significance of strong validity is that it is required for general substitution of equivalents. To this end, $\varphi\langle\psi_1/\psi_2\rangle$ denotes the result of replacing every occurrence of ψ_1 in φ with ψ_2. The next lemmata are proved by induction on formulae. Note that the former addresses strong equivalence, while the latter addresses the weaker notion of equivalence.

Lemma 8. *If $\models\!\!\models \psi_1 \equiv \psi_2$, then $\models\!\!\models \varphi \equiv \varphi\langle\psi_1/\psi_2\rangle$.*

Lemma 9. *If $\models \psi_1 \equiv \psi_2$ and ψ_1 does not occur within the scope of $\boxdot$ or O^R, then $\models \varphi \equiv \varphi\langle\psi_1/\psi_2\rangle$.*

Substitution of strong equivalents makes the following equivalence strongly valid. This equivalence underlies a basic inference rule in the rewriting system.

Lemma 10. $\models\!\!\models O\varphi \equiv (O\varphi\langle\beta/\top\rangle \wedge \beta) \vee (O\varphi\langle\beta/\bot\rangle \wedge \neg\beta)$ *for a prime modal atom β.*

Proof. Let $\mathcal{M}$ be an arbitrary model. By Lemma 3, either $\mathcal{M} \models \beta \equiv \top$ or $\mathcal{M} \models \beta \equiv \bot$. By Lemma 8, either $\mathcal{M} \models O\varphi \equiv O\varphi\langle\beta/\top\rangle$ or $\mathcal{M} \models O\varphi \equiv O\varphi\langle\beta/\bot\rangle$, resp. In either case, we have $\mathcal{M} \models ((\beta \equiv \top) \wedge (O\varphi \equiv O\varphi\langle\beta/\top\rangle)) \vee ((\beta \equiv \bot) \wedge (O\varphi \equiv O\varphi\langle\beta/\bot\rangle))$, which is tautologically equivalent to the formula in the lemma. □

In the rest of the section, we identify some useful strong equivalents. The first two follow directly from the definitions.

Lemma 11. $\models\!\!\models O\varphi \equiv (B\varphi \wedge C\neg\varphi)$.

Lemma 12. $\models\!\!\models O^R\varphi \equiv (O\varphi \wedge \boxdot\neg O\varphi)$.

The idea underlying the next lemma can be illustrated with the help of Fig. 1 and Example 4. In the proof of Lemma 13 we argue that any model of $\diamond O\varphi$ must have the shape of the leftmost model in Fig. 1, in which there must be a point $x \notin U$ at which φ is true. As we show in Example 4, $B\varphi$ and $\neg C\neg\varphi$ are both true in the model. Conversely, any model of $B\varphi \wedge \neg C\neg\varphi$ must also have the shape of the leftmost model in Fig. 1, satisfying $\diamond O\varphi$.

Lemma 13. $\models\!\!\models \diamond O\varphi \equiv (B\varphi \wedge \neg C\neg\varphi)$ *if φ is objective.*

Proof. As φ is objective, $\mathcal{M} \models O\varphi$ iff $\mathcal{M}' \models \neg O\varphi$ for each $\mathcal{M}' > \mathcal{M}$. This is used in both directions below. ($\Rightarrow$) Assume that $\mathcal{M} \models \diamond O\varphi$. Then $\mathcal{M}' \models O\varphi$ for some $\mathcal{M}' > \mathcal{M}$, i.e. $\mathcal{M} \models \neg O\varphi$. Since $\mathcal{M}' \models B\varphi$, $\mathcal{M} \models B\varphi$. Thus $\mathcal{M} \models \neg C\neg\varphi$ by Lemma 11. ($\Leftarrow$) Assume that $\mathcal{M} \models$

$B\varphi \wedge \neg C\neg\varphi$. As $\mathcal{M} \models B\varphi$ and φ is objective, there is some $\mathcal{M}' \geqslant \mathcal{M}$ such that $\mathcal{M}' \models O\varphi$, and as $\mathcal{M} \models \neg C\neg\varphi$, $\mathcal{M} \models \neg O\varphi$, thus $\mathcal{M}' \neq \mathcal{M}$. Hence $\mathcal{M} \models \diamond O\varphi$. □

We let $[\cdot]$ denote the function that replaces M with $\neg B\neg$, and (for the service of the rewriting rules) puts the resulting formula on negation normal form, e.g. $[M\psi] = \neg B[\neg\psi]$, $[\neg M\psi] = B[\neg\psi]$ and $[\neg(\varphi \wedge \psi)] = [\neg\varphi] \vee [\neg\psi]$.

Lemma 14. *$\models \diamond\beta \supset [\beta]$ if β is M-basic.*

Proof. It is immediate that if $\mathcal{M}' \models \beta$ and $\mathcal{M}' > \mathcal{M}$, then $\mathcal{M} \models [\beta]$. The lemma follows from this. □

The set Ω contains formulae of a normal form wrt. one of rewriting relations introduced below. It is defined as the least set such that

- $O\varphi \in \Omega$ if φ is objective;
- $\varphi \wedge M\psi, \varphi \wedge \neg M\psi, \varphi \wedge \top \in \Omega$ if $\varphi \in \Omega$ and ψ is objective.

The next lemma is essential for the rewriting process, as it justifies a reduction of a formula which contains occurrences of $\boxdot$ and M to a formula which does not contain any of these modalities.

Lemma 15. *Let $O\psi \wedge \beta \in \Omega$. Then*

$$\models \boxdot\neg(O\psi \wedge \beta) \equiv ((B\psi \wedge \neg C\neg\psi) \supset [\neg\beta]).$$

Proof. By Lemma 13 and Lemma 8, we have to show that

$$\models \boxdot\neg(O\psi \wedge \beta) \equiv (\diamond O\psi \supset [\neg\beta]).$$

($\Rightarrow$) Assume that $\mathcal{M} \models \boxdot(O\psi \supset \neg\beta)$ and $\mathcal{M} \models \diamond O\psi$. Then $\mathcal{M} \models \diamond(O\psi \wedge \neg\beta)$, thus $\mathcal{M} \models \diamond\neg\beta$, hence $\mathcal{M} \models [\neg\beta]$ by Lemma 14. ($\Leftarrow$) We want to show that for each $\mathcal{M}$,

$$M \models (\boxdot\neg O\psi \vee [\neg\beta]) \supset \boxdot\neg(O\psi \wedge \beta)$$

Now there are two cases. If $\mathcal{M} \models \boxdot\neg O\psi$, then $\mathcal{M} \models \boxdot\neg(O\psi \wedge \beta)$. If $\mathcal{M} \models [\neg\beta]$, then $\mathcal{M} \models \neg[\beta]$, thus $\mathcal{M} \models \boxdot\neg\beta$ by Lemma 14, thus $\mathcal{M} \models \boxdot\neg(O\psi \wedge \beta)$. □

The Rewriting System

The rewriting system introduced in this section consists of two rewriting relations on formulae. The rules of the relation $\rightarrow$ are based on strong equivalences, whereas some of the equivalences underlying the relation $\hat{\rightarrow}$ are not strong. The rewriting process applies the $\rightarrow$ relation exhaustively before $\hat{\rightarrow}$ is applied.

We first define the reduction relation $\rightarrow$, and say that a *reduces* to b if $a \twoheadrightarrow b$, where $\twoheadrightarrow$ is the reflexive transitive closure of $\rightarrow$. Reduction can be performed on any subformula, e.g. if $p \rightarrow q$ is a rewrite rule, then $p \vee q \rightarrow q \vee q$. The same notation is used for the $\hat{\rightarrow}$ relation, i.e. $\hat{\twoheadrightarrow}$ denotes its reflexive transitive closure.

Reduction is performed modulo an equivalence relation $\sim$ under which $\wedge$ and $\vee$ are commutative and associative, and $\top$ and $\bot$ are the empty conjunction and disjunction resp., i.e. $\varphi \wedge \top \sim \varphi$ and $\varphi \vee \bot \sim \varphi$. Also $\neg\bot \sim \top$ and $\neg\top \sim \bot$. As we neglect confluence, we will also assume that some simple propositionally sound reductions, like $\varphi \vee \varphi \rightarrow \varphi$, are allowed. These are, however, not needed for correctness; their role is to hide low-level details and thus allow more readable reduction sequences.

Although we want to reduce formulae of the form $O^R\varphi$, there is only one rule where O^R occurs: $O^R\varphi$ is rewritten to $O\varphi \wedge \boxdot\neg O\varphi$. Hence we need rules to reduce boxed formulae and formulae of the form $O\varphi$, and whatever they are reduced to. A set of such rules for the language without $\boxdot$ and M is found in (Waaler et al. 2007). In this paper, we skip the rules that treat occurrences of O and C within the scope of O. We assume that B and M are the only modalities occurring in φ, and that C only occurs when generated from the rules. The resulting set of rules is sufficient for reducing encoded default theories.

Rules pertaining to O^R and $\boxdot$

$R_1:\ O^R\varphi \rightarrow O\varphi \wedge \boxdot\neg O\varphi$

$R_2:\ \boxdot\neg(\varphi \vee \psi) \rightarrow \boxdot\neg\varphi \wedge \boxdot\neg\psi$

$R_3:\ \boxdot\top \rightarrow \top$

$R_4:\ \boxdot\neg(O\varphi \wedge \beta) \rightarrow \neg B\varphi \vee C\neg\varphi \vee [\neg\beta]$ if $O\varphi \wedge \beta \in \Omega$

Since we assume that the input formula will always be either of the form $O^R\varphi$ or $O\varphi$, R_1 is the only rule that generates a formula with an occurrence of $\boxdot$ from a formula with no such occurrence. Observe that this has the pattern $\boxdot\neg$, a form which both R_2 and R_4 assume. R_2 generates a formula in which this pattern occurs twice, whereas R_4 removes the $\boxdot$.

Example 16. *Let us first examine the formula $\boxdot\neg O\varphi$. R_4 can be used to rewrite this formula modulo $\sim$:*

$$\begin{aligned} \boxdot\neg O\varphi &\sim \boxdot\neg(O\varphi \wedge \top) \\ &\rightarrow \neg B\varphi \vee C\neg\varphi \vee [\neg\top] \\ &\sim \neg B\varphi \vee C\neg\varphi. \end{aligned}$$

Thus $\boxdot\neg O\varphi \rightarrow \neg B\varphi \vee C\neg\varphi$.

Rules pertaining to O

$M_1:\ O\varphi \rightarrow (O\varphi\langle\beta/\top\rangle \wedge \beta) \vee (O\varphi\langle\beta/\bot\rangle \wedge \neg\beta)$
if β is a prime modal atom that occurs in φ

$M_2:\ \varphi \wedge \bot \rightarrow \bot$

$M_3:\ (\varphi \vee \mu) \wedge \psi \rightarrow (\varphi \wedge \psi) \vee (\mu \wedge \psi)$

$M_4:$ For objective φ and ψ,

- $O\varphi \wedge B\psi \rightarrow O\varphi$ if not $\mathsf{SAT}\{\varphi, \neg\psi\}$
- $O\varphi \wedge B\psi \rightarrow \bot$ if $\mathsf{SAT}\{\varphi, \neg\psi\}$
- $O\varphi \wedge \neg B\psi \rightarrow \bot$ if not $\mathsf{SAT}\{\varphi, \neg\psi\}$
- $O\varphi \wedge \neg B\psi \rightarrow O\varphi$ if $\mathsf{SAT}\{\varphi, \neg\psi\}$
- $O\varphi \wedge C\psi \rightarrow O\varphi$ if not $\mathsf{SAT}\{\neg\varphi, \neg\psi\}$
- $O\varphi \wedge C\psi \rightarrow \bot$ if $\mathsf{SAT}\{\neg\varphi, \neg\psi\}$
- $O\varphi \wedge M\psi \rightarrow \bot$ if not $\mathsf{SAT}\{\varphi, \psi\}$
- $O\varphi \wedge \neg M\psi \rightarrow O\varphi$ if not $\mathsf{SAT}\{\varphi, \psi\}$

M_1 is called the *expand rule*, M_2 the *contradiction rule*, M_3 the *distribution rule*, whereas the rules in the M_4 group are called *collapse rules*.

Example 17. *Let us once again address the default theory in Example 1, in which φ is $Bp \supset p$. In Example 5, we give* *a semantic analysis of the models of $O\varphi$ and $O^R\varphi$. Here we show the same results syntactically. To reduce $O\varphi$ we first apply the expand rule. Then we apply the first and the fourth rule in the M_4 group:*

$$\begin{aligned} O\varphi &\to (Op \wedge Bp) \vee (O\top \wedge \neg Bp) \\ &\to Op \vee (O\top \wedge \neg Bp) \\ &\to Op \vee O\top \end{aligned}$$

The same reductions also apply in a boxed context, after which one can apply R_2 and R_4 twice:

$$\begin{aligned} \boxdot\neg O\varphi &\twoheadrightarrow \boxdot\neg(Op \vee O\top) \\ &\to \boxdot\neg Op \wedge \boxdot\neg O\top \\ &\twoheadrightarrow (\neg Bp \vee C\neg p) \wedge (\neg B\top \vee C\bot) \end{aligned}$$

Having reduced $O\varphi$ and $\boxdot\neg O\varphi$, we reduce $O^R\varphi$.

$$\begin{aligned} O^R\varphi &\to O\varphi \wedge \boxdot\neg O\varphi \\ &\twoheadrightarrow (Op \vee O\top) \wedge (\neg Bp \vee C\neg p) \wedge (\neg B\top \vee C\bot) \\ &\to (Op \wedge (\neg Bp \vee C\neg p) \wedge (\neg B\top \vee C\bot)) \vee \\ &\quad\; (O\top \wedge (\neg Bp \vee C\neg p) \wedge (\neg B\top \vee C\bot)) \end{aligned}$$

Distributing conjunctions over disjunctions, and collapsing inconsistent conjuncts, we obtain

$$\begin{aligned} &\twoheadrightarrow (O\top \wedge \neg Bp \wedge C\bot) \vee (O\top \wedge C\neg p \wedge C\bot) \\ &\twoheadrightarrow O\top. \end{aligned}$$

Theorem 18. *If $\varphi \twoheadrightarrow \psi$ then $\models \varphi \equiv \psi$.*

Proof. By Lemma 8, it is sufficient to show that $\models l \equiv r$ for each rule $l \to r$, in which case we say that the rule is strongly valid. For R_1, this follows from Lemma 12, for R_2 from the fact that $\models \boxdot(\varphi \wedge \psi) \equiv (\boxdot\varphi \wedge \boxdot\psi)$, for R_3 from $\models \boxdot\top$, and for R_4 from Lemma 15. Strong validity of M_1 follows from Lemma 10. Strong validity of the last two rules in the M_4 group follows from Lemma 7; that the other rules are strongly valid can be proved by arguments similar to the proof of Lemma 7. The rest are trivial. □

The rules of the $\to$ relation reduce an only-knowing formula to a disjunction of formulae in Ω. We prove this, first for input of the form $O\varphi$, and then for input of the form $O^R\varphi$.

Lemma 19. *For some $n \geqslant 0$, there are $O\varphi_k \wedge \beta_k \in \Omega$ for $1 \leqslant k \leqslant n$ such that*

$$O\varphi \twoheadrightarrow (O\varphi_1 \wedge \beta_1) \vee \cdots \vee (O\varphi_n \wedge \beta_n).$$

Proof. For additional details, cf. (Lian, Langholm, and Waaler 2004; Waaler et al. 2007). The basic procedure is as follows.

1. Expand the formula by repeatedly applying M_1 until there are no modal subformulae left;
2. put the resulting formula on DNF using M_3;
3. collapse the conjunctions using M_4;
4. remove inconsistent conjunctions with M_2.

Let $O\varphi \to (O\varphi\langle\beta/\top\rangle \wedge \beta) \vee (O\varphi\langle\beta/\bot\rangle \wedge \neg\beta)$, and assume that the procedure has been applied to $O\varphi\langle\beta/\top\rangle$ and $O\varphi\langle\beta/\bot\rangle$, i.e. that there are conjunctions $\mu_1^\top, \ldots, \mu_m^\top$ and $\mu_1^\bot, \ldots, \mu_n^\bot$ such that

$$\begin{aligned} O\varphi\langle\beta/\top\rangle &\twoheadrightarrow \mu_1^\top \vee \cdots \vee \mu_m^\top, \text{ and} \\ O\varphi\langle\beta/\bot\rangle &\twoheadrightarrow \mu_1^\bot \vee \cdots \vee \mu_n^\bot. \end{aligned}$$

Then

$$\begin{aligned} O\varphi &\twoheadrightarrow ((\mu_1^\top \vee \cdots \vee \mu_m^\top) \wedge \beta) \vee \\ &\quad\; ((\mu_1^\bot \vee \cdots \vee \mu_n^\bot) \wedge \neg\beta) \\ &\twoheadrightarrow (\mu_1^\top \wedge \beta) \vee \cdots \vee (\mu_m^\top \wedge \beta) \vee \\ &\quad\; (\mu_1^\bot \wedge \neg\beta) \vee \cdots \vee (\mu_n^\bot \wedge \neg\beta). \end{aligned}$$

Now each $\mu_i^\top \wedge \beta$ and $\mu_k^\bot \wedge \neg\beta$ contains exactly one conjunct $O\psi$ for some objective ψ, while the other conjuncts are B-, C-, and M-literals, of which the B- and C-literals are collapsed. Hence we are left with a disjunction of formulae from Ω. □

Use of the distribution rule should be postponed whenever possible, as this may cause an exponential blowup. The proof of Lemma 19 reduces to DNF before the collapse rules are used; this strategy is simply used to make the proof easier. A more clever strategy would be to collapse as much as possible before using the distribution rule. In fact, there are cases where the distribution is not needed at all.

Example 20. *Let $\varphi = d_1 \wedge d_2$, where $d_1 = Bp \supset q$ and $d_2 = Bq \supset p$. Then*

$$\begin{aligned} O\varphi &\to (O(q \wedge d_2) \wedge Bp) \vee (Od_2 \wedge \neg Bp) \text{ by } M_1 \\ &\twoheadrightarrow [((O(q \wedge p) \wedge Bq) \vee (Oq \wedge \neg Bq)) \wedge Bp] \vee \\ &\quad\; [((Op \wedge Bq) \vee (O\top \wedge \neg Bq)) \wedge \neg Bp] \text{ by } M_1 \\ &\twoheadrightarrow [((O(q \wedge p) \wedge Bq) \vee \bot) \wedge Bp] \vee \\ &\quad\; [(\bot \vee (O\top \wedge \neg Bq)) \wedge \neg Bp] \text{ by } M_2 \\ &\sim [(O(q \wedge p) \wedge Bq) \wedge Bp] \vee [(O\top \wedge \neg Bq) \wedge \neg Bp] \\ &\twoheadrightarrow O(q \wedge p) \vee O\top \text{ by } M_4. \end{aligned}$$

M_3 was not needed at any point in the reduction.

Lemma 21. *For some $n \geqslant 0$, there are $O\varphi_k \wedge \beta_k \in \Omega$ for $1 \leqslant k \leqslant n$ such that*

$$O^R\varphi \twoheadrightarrow (O\varphi_1 \wedge \beta_1) \vee \cdots \vee (O\varphi_n \wedge \beta_n).$$

Proof. By Lemma 19, for some $n \geqslant 0$, there are $O\varphi_k \wedge \beta_k \in \Omega$ for $1 \leqslant k \leqslant n$ such that

$$\begin{aligned} O\varphi &\twoheadrightarrow (O\varphi_1 \wedge \beta_1) \vee \cdots \vee (O\varphi_n \wedge \beta_n), \text{ thus} \\ \boxdot\neg O\varphi &\twoheadrightarrow \boxdot\neg((O\varphi_1 \wedge \beta_1) \vee \cdots \vee (O\varphi_n \wedge \beta_n)) \\ &\twoheadrightarrow \boxdot\neg(O\varphi_1 \wedge \beta_1) \wedge \cdots \wedge \boxdot\neg(O\varphi_n \wedge \beta_n)) \\ &\twoheadrightarrow (\neg B\varphi_1 \vee C\neg\varphi_1 \vee [\neg\beta_1]) \wedge \cdots \wedge \\ &\quad\; (\neg B\varphi_n \vee C\neg\varphi_n \vee [\neg\beta_n]) = \mu, \end{aligned}$$

where $\mu = \nu_1 \wedge \cdots \wedge \nu_n$ and $\nu_k = \neg B\varphi_k \vee C\neg\varphi_k \vee [\neg\beta_k]$.

$$\begin{aligned} O^R\varphi &\to O\varphi \wedge \boxdot\neg O\varphi \twoheadrightarrow O\varphi \wedge \mu \\ &\twoheadrightarrow (O\varphi_1 \wedge \mu \wedge \beta_1) \vee \cdots \vee (O\varphi_n \wedge \mu \wedge \beta_n) \end{aligned}$$

Each $O\varphi_k \wedge \mu \wedge \beta_k$ reduces to either $O\varphi_k \wedge \beta_k$ or $\bot$. □

When we reach the situation of Lemma 21 in the rewriting process, we are left with disjunctions of elements in Ω. Note that the $\rightarrow$ relation has only two rules for collapsing M-formulae. The two collapse rules that are missing do not preserve strong equivalence and are hence not sound in all contexts. Hence we define a new reduction relation $\hat{\rightarrow}$ that includes the $\rightarrow$ relation defined above and extends it with the two M-collapsing rules that are missing in the $\rightarrow$ relation:

- $O\varphi \wedge M\psi \mathrel{\hat{\rightarrow}} O\varphi$ if $\mathsf{SAT}\{\varphi, \psi\}$
- $O\varphi \wedge \neg M\psi \mathrel{\hat{\rightarrow}} \bot$ if $\mathsf{SAT}\{\varphi, \psi\}$

Example 22. *In this example, we examine the prerequisite-free default theory $(\emptyset, \{\top : p \,/\, p\})$, which has the same unique expansion and extension. It translates into $O^R\varphi$, where φ is $(Mp \supset p)$. Note that in contrast to the previous example, the translation introduces an occurrence of M.*

$$\begin{aligned}
O\varphi &\rightarrow (Op \wedge Mp) \vee (O\top \wedge \neg Mp)\\
\boxdot\neg O\varphi &\twoheadrightarrow \boxdot\neg((Op \wedge Mp) \vee (O\top \wedge \neg Mp))\\
&\rightarrow \boxdot\neg(Op \wedge Mp) \wedge \boxdot\neg(O\top \wedge \neg Mp)\\
&\twoheadrightarrow (\neg Bp \vee C\neg p \vee [\neg Mp]) \wedge\\
&\quad\;\, (\neg B\top \vee C\bot \vee [\neg\neg Mp])\\
&= (\neg Bp \vee C\neg p \vee B\neg p) \wedge\\
&\quad\;\, (\neg B\top \vee C\bot \vee \neg B\neg p)\\
O^R\varphi &\twoheadrightarrow ((Op \wedge Mp) \vee (O\top \wedge \neg Mp)) \wedge\\
&\quad\;\, (\neg Bp \vee C\neg p \vee B\neg p) \wedge\\
&\quad\;\, (\neg B\top \vee C\bot \vee \neg B\neg p)\\
&\twoheadrightarrow (Op \wedge Mp) \vee (O\top \wedge \neg Mp)\\
&\mathrel{\hat{\twoheadrightarrow}} Op
\end{aligned}$$

The next example does not correspond to any default theory; it is included as a simple example of the case where $O^R\varphi$ has more than one weak model.

Example 23. *Let $\varphi = (\neg M\neg p \supset p)$.*

$$\begin{aligned}
O\varphi &\rightarrow (O\top \wedge M\neg p) \vee (Op \wedge \neg M\neg p)\\
\boxdot\neg O\varphi &\twoheadrightarrow \boxdot\neg((O\top \wedge M\neg p) \vee (Op \wedge \neg M\neg p))\\
&\rightarrow \boxdot\neg(O\top \wedge M\neg p) \wedge \boxdot\neg(Op \wedge \neg M\neg p)\\
&\twoheadrightarrow (\neg B\top \vee C\bot \vee [\neg M\neg p]) \wedge\\
&\quad\;\, (\neg Bp \vee C\neg p \vee [\neg\neg M\neg p])\\
&= (\neg B\top \vee C\bot \vee Bp) \wedge (\neg Bp \vee C\neg p \vee \neg Bp)\\
O^R\varphi &\twoheadrightarrow ((O\top \wedge M\neg p) \vee (Op \wedge \neg M\neg p)) \wedge\\
&\quad\;\, (\neg B\top \vee C\bot \vee Bp) \wedge (\neg Bp \vee C\neg p \vee \neg Bp)\\
&\twoheadrightarrow (O\top \wedge M\neg p) \vee (Op \wedge \neg M\neg p)\\
&\mathrel{\hat{\twoheadrightarrow}} O\top \vee Op
\end{aligned}$$

Lemma 24. *Let $\varphi = \bigvee\Phi$ for some $\Phi \subseteq \Omega$. If $\varphi \mathrel{\hat{\twoheadrightarrow}} \psi$ then $\models \varphi \equiv \psi$.*

Proof. We first show that all rules of $\hat{\rightarrow}$ are valid. That the rules of $\rightarrow$ are valid is obvious, given that they are strongly valid. It follows immediately that for an M-literal β,

$$O\varphi \wedge \beta \mathrel{\hat{\rightarrow}} \psi \text{ if } O\varphi \wedge [\beta] \rightarrow \psi,$$

from which validity of the rewrite rules listed above follows. The theorem follows from these observations and the fact that formulae in Ω do not contain $\boxdot$ and O^R, and Lemma 9. ☐

Theorem 25. *For some $n \geqslant 0$, there are objective $\varphi_1, \ldots, \varphi_n$ such that for some $\Phi \subseteq \Omega$,*

$$O^R\varphi \twoheadrightarrow \bigvee\Phi \mathrel{\hat{\twoheadrightarrow}} (O\varphi_1 \vee \cdots \vee O\varphi_n).$$

Proof. Follows from first applying Lemma 21, and then collapsing the conjunctions in Φ using $\hat{\rightarrow}$. ☐

Corollary 26. *For some $n \geqslant 0$, there are objective $\varphi_1, \ldots, \varphi_n$ such that $\models O^R\varphi \equiv (O\varphi_1 \vee \cdots \vee O\varphi_n)$.*

Proof. By Theorem 18, Lemma 24 and Theorem 25. ☐

Adding Simplification Rules

On the one hand, we wish to reduce $O\varphi$ to a formula on DNF; on the other, we wish to avoid applying the distribution rule unless strictly necessary. The fact that $\rightarrow$ cannot collapse all formulae in Ω means that we have to apply the distribution rule to a formula expanded to prime form which contains M. To illustrate this point, let

$$\varphi = (Mp \supset q) \wedge (Mq \supset q).$$

Reducing $O\varphi$, applying only the expand rule, we get

$$\begin{aligned}
&([(Oq \wedge Mq) \vee (Oq \wedge \neg Mq)] \wedge Mp) \vee\\
&([(Oq \wedge Mq) \vee (O\top \wedge \neg Mq)] \wedge \neg Mp).
\end{aligned}$$

The only rule that now applies is the distribution rule. But the formula as it stands can clearly be simplified more directly. We introduce three simplification rules to this end. For M-literals α and β,

$S_1 :\ O\varphi \vee (O\varphi \wedge \beta) \rightarrow O\varphi$

$S_2 :\ (O\varphi \wedge \beta) \vee (O\varphi \wedge \overline{\beta}) \rightarrow O\varphi$

$S_3 :\ (O\varphi \wedge \beta) \vee (O\varphi \wedge \overline{\beta} \wedge \alpha) \rightarrow (O\varphi \wedge \beta) \vee (O\varphi \wedge \alpha)$,

where $\overline{\beta}$ denotes the complement of β, i.e. $\overline{M\psi} = \neg M\psi$ and $\overline{\neg M\psi} = M\psi$. Now the above formula reduces to

$$(Oq \wedge Mp) \vee (Oq \wedge Mq) \vee (O\top \wedge \neg Mp \wedge \neg Mq).$$

The last example is taken from (Gottlob 1995), and is the translation of the default theory

$$(\emptyset, \{\tfrac{p \supset q \,:\, p}{p}, \tfrac{p \,:\, q}{q}\}).$$

As $W = \emptyset$ and both defaults have prerequisites, the only extension is the set of tautologies.

Example 27. *Let $\varphi = d_1 \wedge d_2$, where*

$$\begin{aligned}
d_1 &= B(p \supset q) \wedge Mp \supset p;\\
d_2 &= Bp \wedge Mq \supset q.
\end{aligned}$$

If we substitute the M's first, we can collapse the leaf nodes. Then we can use the simplification rules. $O\varphi$ reduces to

$$\begin{aligned}
&([((O\top \vee O(p \wedge q)) \wedge Mq) \vee (O\top \wedge \neg Mq)] \wedge Mp) \vee\\
&([(O\top \wedge Mq) \vee (O\top \wedge \neg Mq)] \wedge \neg Mp)
\end{aligned}$$

$\to ([((O\top \vee O(p \wedge q)) \wedge Mq) \vee (O\top \wedge \neg Mq)] \wedge Mp) \vee$
$(O\top \wedge \neg Mp)$ *by* S_2;
$\to ((O\top \vee O(p \wedge q)) \wedge Mq \wedge Mp) \vee$
$(O\top \wedge \neg Mq \wedge Mp) \vee (O\top \wedge \neg Mp)$ *by* M_3;
$\to ((O\top \vee O(p \wedge q)) \wedge Mq \wedge Mp) \vee$
$(O\top \wedge \neg Mq) \vee (O\top \wedge \neg Mp)$ *by* S_3;
$\to (O\top \wedge Mq \wedge Mp) \vee (O(p \wedge q) \wedge Mq \wedge Mp) \vee$
$(O\top \wedge \neg Mq) \vee (O\top \wedge \neg Mp)$ *by* M_3;
$\to (O\top \wedge Mq) \vee (O(p \wedge q) \wedge Mq \wedge Mp) \vee$
$(O\top \wedge \neg Mq) \vee (O\top \wedge \neg Mp)$ *by* S_3;
$\to O\top \vee (O(p \wedge q) \wedge Mq \wedge Mp) \vee (O\top \wedge \neg Mp)$ *by* S_2;
$\to O\top \vee (O(p \wedge q) \wedge Mq \wedge Mp)$ *by* S_1.

Hence

$$\begin{aligned} O\varphi &\to O\top \vee (O(p \wedge q) \wedge Mq \wedge Mp) \\ \boxdot\neg O\varphi &\twoheadrightarrow \boxdot\neg(O\top \vee (O(p \wedge q) \wedge Mq \wedge Mp)) \\ &\to \boxdot\neg O\top \wedge \boxdot\neg(O(p \wedge q) \wedge Mq \wedge Mp) \\ &\twoheadrightarrow (\neg B\top \vee C\bot) \wedge \\ &\quad (\neg B(p \wedge q) \vee C\neg(p \wedge q) \vee B\neg q \vee B\neg p) \\ O^R\varphi &\twoheadrightarrow O\top. \end{aligned}$$

Adding Confidence Levels

To enable the representation of ordered default theories, we extend the only-knowing system by introducing a partial order $(I, \preccurlyeq)$, intuitively representing confidence levels, and for each index $k \in I$, adding modal operators B_k, C_k, M_k, O_k, O^R_k, and $\boxdot_k$ to the signature of the logic. A formula of the form $\bigwedge_{k\in I} O_k\varphi_k$ is called an O_I*-block*. An O^R_I*-block* is defined similarly. Ω_I is defined as the least set such that

- $\varphi \in \Omega_I$ if φ is a prime O_I-block;
- $\varphi \wedge M_k\psi, \varphi \wedge \neg M_k\psi, \varphi \wedge \top \in \Omega_I$ if $k \in I$, $\varphi \in \Omega_I$ and ψ is objective.

A model for the logic with confidence levels is a set of tuples $\{(U_k, V_k) \mid k \in I\}$ such that $U_k \subseteq U_i$ for each $i \prec k$, and with a satisfaction relation which generalizes the satisfaction relation of the logic without confidence levels in the obvious way. To generalize the rewrite rules, it is in many cases sufficient to add subscripts to the modalities.

R'_1 : $O^R_k\varphi \to O_k\varphi \wedge \boxdot_k\neg O_k\varphi$

R'_2 : $\boxdot_k\neg(\varphi \vee \psi) \to \boxdot_k\neg\varphi \wedge \boxdot_k\neg\psi$

R'_3 : $\boxdot_k\top \to \top$

R'_4 : $\boxdot_k\neg(O_k\varphi \wedge \alpha \wedge \beta) \to \neg B_k\varphi \vee C_k\neg\varphi \vee [\neg\alpha] \vee [\neg\beta]$ if φ is propositional, and α and β are conjunctions of B- and M-literals resp.

The collapse rules are more intricate, as for a given $O_k\varphi$ and modal literal β, it might be the case that $O_k\varphi$ neither implies β nor its negation. Also $O_i\varphi \wedge O_k\psi$ might be inconsistent. The following rules are sufficient.

M'_4 : For objective φ and ψ,

- $O_i\varphi \wedge B_k\psi \to O_i\varphi$ if $i \preccurlyeq k$ and not $\mathsf{SAT}\{\varphi, \neg\psi\}$
- $O_i\varphi \wedge B_k\psi \to \bot$ if $k \preccurlyeq i$ and $\mathsf{SAT}\{\varphi, \neg\psi\}$
- $O_i\varphi \wedge \neg B_k\psi \to \bot$ if $i \preccurlyeq k$ and not $\mathsf{SAT}\{\varphi, \neg\psi\}$
- $O_i\varphi \wedge \neg B_k\psi \to O_i\varphi$ if $k \preccurlyeq i$ and $\mathsf{SAT}\{\varphi, \neg\psi\}$
- $O_i\varphi \wedge C_k\psi \to O_i\varphi$ if $k \preccurlyeq i$ and not $\mathsf{SAT}\{\neg\varphi, \neg\psi\}$
- $O_i\varphi \wedge C_k\psi \to \bot$ if $i \preccurlyeq k$ and $\mathsf{SAT}\{\neg\varphi, \neg\psi\}$
- $O_i\varphi \wedge O_k\psi \to \bot$ if $i \preccurlyeq k$ and $\mathsf{SAT}\{\neg\varphi, \psi\}$
- $O_i\varphi \wedge M_k\psi \hat{\to} \bot$ if $i \preccurlyeq k$ and not $\mathsf{SAT}\{\varphi, \psi\}$
- $O_i\varphi \wedge \neg M_k\psi \hat{\to} O_i\varphi$ if $i \preccurlyeq k$ and not $\mathsf{SAT}\{\varphi, \psi\}$

As before, the reduction relation $\to$ is extended to $\hat{\to}$ to deal with M_k.

- $O_i\varphi \wedge M_k\psi \hat{\to} O_i\varphi$ if $k \preccurlyeq i$ and $\mathsf{SAT}\{\varphi, \psi\}$
- $O_i\varphi \wedge \neg M_k\psi \hat{\to} \bot$ if $k \preccurlyeq i$ and $\mathsf{SAT}\{\varphi, \psi\}$

Lemma 28. *For any prime O_I-block and propositional ψ,*

1. *either*
   - $\varphi \wedge B_k\psi \twoheadrightarrow \varphi$ *and* $\varphi \wedge \neg B_k\psi \twoheadrightarrow \bot$*, or*
   - $\varphi \wedge B_k\psi \twoheadrightarrow \bot$ *and* $\varphi \wedge \neg B_k\psi \twoheadrightarrow \varphi$*;*
2. *either*
   - $\varphi \wedge M_k\psi \hat{\twoheadrightarrow} \varphi$ *and* $\varphi \wedge \neg M_k\psi \hat{\twoheadrightarrow} \bot$*, or*
   - $\varphi \wedge M_k\psi \hat{\twoheadrightarrow} \bot$ *and* $\varphi \wedge \neg M_k\psi \hat{\twoheadrightarrow} \varphi$*.*

Proof. For any $k \in I$, one of φ's conjuncts is of the form $O_k\mu$, and either $O_k\mu \wedge B_k\psi \to O_k\mu$ (in which case $O_k\mu \wedge \neg B_k\psi \to \bot$) or $O_k\mu \wedge B_k\psi \to \bot$ (in which case $O_k\mu \wedge \neg B_k\psi \to O_k\mu$). Similarly for M_k. ☐

Theorem 29. *For any O^R_I-block φ, there is a $\Phi \subseteq \Omega_I$ and prime O_I-blocks $\varphi_1, \ldots, \varphi_n$, $n \geqslant 0$, such that*

$$\varphi \twoheadrightarrow \bigvee\Phi \hat{\twoheadrightarrow} (\varphi_1 \vee \cdots \vee \varphi_n).$$

Proof. For each of φ's conjuncts $O^R_k\psi$,

$$O^R_k\psi \to O_k\psi \wedge \boxdot_k\neg O_k\psi.$$

Use the basic procedure given in the proof of Lemma 19 for showing that for each such $O_k\psi$, for some $n_k \geqslant 0$, there are

- propositional ψ_i,
- conjunctions of M-literals β_i, and
- conjunctions of B-literals α_i (as collapsing is not always possible when confidence levels has been added),

such that

$$O_k\psi \twoheadrightarrow (O_k\psi_1 \wedge \alpha_1 \wedge \beta_1) \vee \cdots \vee (O_k\psi_{n_k} \wedge \alpha_{n_k} \wedge \beta_{n_k}).$$

Let ω_k denote this reduct. Now

$$\boxdot_k\neg\omega_k \twoheadrightarrow \boxdot_k\neg\omega_{k,1} \wedge \cdots \wedge \boxdot_k\neg\omega_{k,n_k},$$

such that for $1 \leqslant i \leqslant n_k$,

$$\begin{aligned} \boxdot_k\neg\omega_{k,i} &= \boxdot_k\neg(O_k\psi_i \wedge \alpha_i \wedge \beta_i) \\ &\to \neg B_k\psi_i \vee C_k\psi_i \vee [\neg\alpha_i] \vee [\neg\beta_i]. \end{aligned}$$

Let $\tau_{k,i}$ denote this reduct. Using M_3, we can reduce $\bigwedge_{k\in I}\omega_k$ to a formula on DNF, whose disjuncts are of the form $\bigwedge_{k\in I} O_k\psi_k \wedge \alpha \wedge \beta$, where α is a conjunction of B-literals, and β is a conjunction of M-literals. By Lemma 28(1),

$$\textstyle\bigwedge_{k\in I} O_k\psi_k \wedge \alpha \wedge \beta \twoheadrightarrow \bigwedge_{k\in I} O_k\psi_k \wedge \beta,$$

which is in Ω_I. Hence for some $\Gamma \subseteq \Omega_I$,

$$O_k^R\psi \twoheadrightarrow \bigvee\Gamma \wedge \bigwedge_{k\in I}\bigwedge_{1\leqslant i\leqslant n_k}\tau_{k,i}$$

We are done by putting this formula on DNF using M_3, applying Lemma 28(1), then Lemma 28(2). □

Corollary 30. *For any O_I^R-block φ, there are prime O_I-blocks $\varphi_1,\ldots,\varphi_n$, $n\geqslant 0$, such that $\models \varphi \equiv (\varphi_1\vee\cdots\vee\varphi_n)$.*

Proof. Theorem 18 can be generalized along the lines of (Waaler et al. 2007). Lemma 24 is easily seen to generalize. Conclude by Theorem 29. □

With reasonable assumptions about $\sim$, both $\rightarrow$ and $\hat{\rightharpoonup}$ are terminating. They are not confluent as they stand, because this requires some additional rules and restrictions.

Example 31. *Let $\varphi_1 = (B_1p \supset p)$, $\varphi_2 = (q\wedge B_2p\supset p)$. Then*

$$\begin{aligned} O_1\varphi_1 &\twoheadrightarrow O_1\top \vee O_1 p\\ O_2\varphi_2 &\twoheadrightarrow O_2 q \vee O_2(p\wedge q)\\ O_1\varphi_1\wedge O_2\varphi_2 &\twoheadrightarrow (O_1\top\wedge O_2 q)\vee(O_1\top\wedge O_2(p\wedge q))\vee\\ &\quad\ (O_1p\wedge O_2q)\vee(O_1p\wedge O_2(p\wedge q))\end{aligned}$$

All but the third disjunct are consistent: $O_1p\wedge O_2q\rightarrow\bot$.

Complexity

Applying the distribution rule can cause an exponential increase in the size of the formula, and should as such be avoided unless strictly necessary.

After expanding, if collapsing is impossible, the distribution rule is the only applicable rule. Imagine a block where each conjunct is expanded until prime, and where no leaf node $O_k\psi\wedge\beta$ may be collapsed:

$$O_1\varphi_1\wedge O_2\varphi_2\wedge O_3\varphi_3 \twoheadrightarrow \psi_1\wedge\psi_2\wedge\psi_3$$

Then we may have to put each reduct ψ_k on DNF in order to collapse:

$$\psi_1\wedge\psi_2\wedge\psi_3 \twoheadrightarrow \psi_1^{\rm DNF}\wedge\psi_2^{\rm DNF}\wedge\psi_3^{\rm DNF}$$

Each $\psi_k^{\rm DNF}$ consists of disjuncts of the form

$$O_k\mu\wedge\beta_1\wedge\cdots\wedge\beta_n,$$

but even this formula may not be collapsable, hence we may have to put it on DNF:

$$\psi_1^{\rm DNF}\wedge\psi_2^{\rm DNF}\wedge\psi_3^{\rm DNF} \twoheadrightarrow (\psi_1^{\rm DNF}\wedge\psi_2^{\rm DNF}\wedge\psi_3^{\rm DNF})^{\rm DNF}$$

Now we may collapse, as each disjunct will be of the form

$$O_1\mu_1\wedge O_2\mu_2\wedge O_3\mu_3\wedge\beta_1\wedge\cdots\wedge\beta_n.$$

In some cases, collapsing is always possible; a singleton I is a trivial case. In other cases, a *strategy* is needed in order to be guaranteed that collapsing is possible. We examine one such case. We say that $K\subseteq I$ is a *downset* if for every $k\in K$, $i\prec k$ implies $i\in K$.

Definition 32 (Downset property). *If for every downset $K\subseteq I$, $\bigwedge_{k\in K}O_k\varphi_k$ only contains K-modalities, we say that $\bigwedge_{k\in I}O_k\varphi_k$ has the* downset property.

A formula that has the downset property can be reduced by reducing singleton blocks from "below." Assume that

$$O_1\varphi_1\wedge O_2\varphi_2\wedge O_3\varphi_3$$

has the downset property for $1\prec 2\prec 3$, and assume, for the sake of simplicity, that there are no occurrences of M_k. As $O_1\varphi_1$ only contains B_1-modalities, it may be reduced to a disjunction of prime O_1-formulae:

$$\begin{aligned} &O_1\varphi_1\wedge O_2\varphi_2\wedge O_3\varphi_3\\ \twoheadrightarrow\ &(O_1\mu_1\vee\cdots\vee O_1\mu_n)\wedge O_2\varphi_2\wedge O_3\varphi_3\end{aligned}$$

If we expand $O_2\varphi_2$, it may be impossible to collapse the leaf nodes – $O_2p\wedge B_1p$ is an example of this – but if we *distribute in* the reduct of $O_1\varphi_1$, we get

$$\begin{aligned} &(O_1\mu_1\vee\cdots\vee O_1\mu_n)\wedge(O_2p\wedge B_1p)\\ \twoheadrightarrow\ &(O_1\mu_1\wedge B_1p\wedge O_2p)\vee\cdots\vee(O_1\mu_n\wedge B_1p\wedge O_2p),\end{aligned}$$

and each $O_1\mu_k\wedge B_1p$ may be collapsed. Hence

$$\begin{aligned} &O_1\varphi_1\wedge O_2\varphi_2\wedge O_3\varphi_3\\ \twoheadrightarrow\ &((O_1\mu_1'\wedge O_2\nu_1)\vee\cdots\vee(O_1\mu_m'\wedge O_2\nu_m))\wedge O_3\varphi_3.\end{aligned}$$

The argument may be repeated for $O_3\varphi_3$. Thus, when we have the downset property, we avoid the DNF reductions described above.

The default logic translation given in the next section has the downset property.

Example 33. *Let $I=\{1,2,3,4\}$, and*

$$1\prec 2\prec 4 \text{ and } 1\prec 3\prec 4.$$

The non-empty downsets are $\{1\}$, $\{1,2\}$, $\{1,3\}$ and I. We want to reduce

$$O_1\varphi_1\wedge O_2\varphi_2\wedge O_3\varphi_3\wedge O_4\varphi_4.$$

Assume that this formula has the downset property, i.e.

- *φ_1 contains only 1-modalities;*
- *φ_2 contains only 1- and 2-modalities;*
- *φ_3 contains only 1- and 3-modalities;*
- *φ_4 may contain any modality.*

Then we may reduce $O_1\varphi_1$, to let us say O_1p. If $\varphi_2 = B_1p\supset p$, then

$$\begin{aligned} O_2\varphi_2 &\rightarrow (O_2p\wedge B_1p)\vee(O_2\top\wedge\neg B_1p)\\ &\rightarrow (O_2p\wedge B_1p)\vee O_2\top.\end{aligned}$$

This cannot be reduced further but

$$\begin{aligned} O_1\varphi_1\wedge O_2\varphi_2 &\twoheadrightarrow O_1p\wedge((O_2p\wedge B_1p)\vee O_2\top)\\ &\rightarrow (O_1p\wedge O_2p\wedge B_1p)\vee(O_1p\wedge O_2\top)\\ &\twoheadrightarrow O_1p\wedge O_2p.\end{aligned}$$

We could also have reduced $O_1\varphi_1\wedge O_3\varphi_3$ but in any case we may have to reduce $O_1\varphi_1\wedge O_2\varphi_2\wedge O_3\varphi_3$ before $O_4\varphi_4$, since $1\prec 4$, $2\prec 4$ and $3\prec 4$.

Ordered Default Logic

The procedure of translating ordered default theories into a standard only-knowing logic given in (Engan et al. 2005) is as follows. Given an ordered default theory $(W, D, <)$, where $n = |D|$, the index set of the signature is defined to be $I = \{0, \ldots, n\}$ with $\preccurlyeq$ equal to the usual number order on I. This defines the modalities of the only-knowing logic that the theory translates into. Then, for every topological ordering σ of $(D, <)$, let

$$\varphi_\sigma = O_0\varphi_0 \wedge \cdots \wedge O_n\varphi_n \wedge \text{IC}(\sigma),$$

where

- $\varphi_0 = \bigwedge W$;
- $\varphi_{k+1} = \varphi_k \wedge \text{tr}_{k+1}(\delta_{k+1})$;
- $\text{tr}_k(\delta) = B_k\alpha \wedge \neg B_k\neg\beta \supset \gamma$ (as M_k is not a part of the standard language, $\neg B_k\neg$ is used instead);
- $\text{IC}(\sigma)$ is a formula that expresses an *integrity constraint*, assuring a prescriptive interpretation on the preference order.

The encoding is the disjunction of *every* such φ_σ. Note that the representation spans all ways in which extensions can potentially be generated as long as the partial order on defaults is respected. This very large formula can then be collapsed into a much smaller one, which directly reflects all the models. Also note that the representation contains massive amounts of redundancy. E.g., if two defaults δ and δ' are unrelated, then for each topological ordering, there will be another that is equal except that δ and δ' have swapped positions. These two orderings will generate the same extension but both of them will have to be reduced in isolation.

In the reduction process, reducing a conjunct $O_{k+1}(\varphi_k \wedge \text{tr}_{k+1}(\delta_{k+1}))$ (in conjunction with the $O_i\varphi_i$'s for $i \prec k$) basically amounts to reducing with *exactly* one default. In this way conflicts are avoided, which is the reason why the translation works for O. Using O^R instead of O provides a correct treatment of conflicting defaults which, as we shall see, allows a much more economical representation.

Example 34. *Let* $W = \{\kappa\}$, $D = \{\delta_1, \delta_2, \delta_3\}$, *and* $\delta_3 < \delta_1$. *Then there are three topological orderings of* $(D, <)$, *one of which is* $\delta_3\delta_2\delta_1$, *whose translation is*

$$O_0\kappa \wedge O_1(\kappa \wedge \delta_3) \wedge O_2(\kappa \wedge \delta_2 \wedge \delta_3) \wedge$$
$$O_3(\kappa \wedge \delta_1 \wedge \delta_2 \wedge \delta_3) \wedge IC(\delta_3\delta_2\delta_1).$$

Example 35. *Let* $D = \{1, 2, 3, 4, 5\}$, *and*

$$1 < 2, \quad 1 < 3, \quad 2 < 4, \quad \textit{and} \quad 2 < 5.$$

See Fig. 2 for the tree of topological orderings of $(D, <)$.

The New Translation

Departing from the encoding given above, the new translation first collapses the tree of topological orders and then encodes the resulting tree using O^R rather than O. The general procedure for collapsing the tree is as follows.

1. For any node a, replace it with $\{a\}$.
2. For any node Γ, let $\Gamma_1, \ldots, \Gamma_m$ denote its children. Now if every $a \in \Gamma$ is unrelated to every $b \in \Gamma_k$ for every child node Γ_k, replace Γ and its children with the new node $\Gamma \cup \Gamma_1 \cup \cdots \cup \Gamma_m$.

Let C_D denote the collapsed tree and $\sqsubset$ the corresponding order on C_D induced from the tree of topological orders.

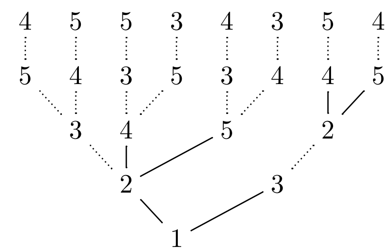

Figure 2: The tree of topological orderings of $(D, <)$ from Example 35. A dotted line between a and b denotes that neither $a < b$ nor $b < a$. If every line from a node to its children is dotted, we may collapse that node and its children.

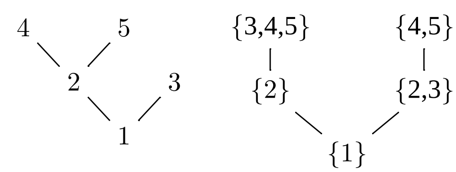

Figure 3: $(D, <)$ and $(C_D, \sqsubset)$ from Example 35.

Two particularly interesting cases are these.

- When $<$ is empty, we may collapse the entire tree, hence we get the single element $\{D\}$ with an empty order, which means that we do not need confidence levels to represent (unordered) default logic.
- When $<$ is linear, there is only one branch in the tree, $1 < \cdots < n$, hence we get $\{1\} \sqsubset \cdots \sqsubset \{n\}$.

Using O_i^R and M_k instead of O_i and $\neg B_k\neg$, a new translation can be specified as follows. Given a theory $(W, D, <)$, for every collapsed branch $\sigma = \Gamma_1 \cdots \Gamma_n$, generate a formula $O_0^R\varphi_0 \wedge \cdots \wedge O_n^R\varphi_n \wedge \text{IC}(\sigma)$, where

- $\varphi_0 = \bigwedge W$,
- $\varphi_{k+1} = \varphi_k \wedge \bigwedge\{\text{tr}_{k+1}(\delta) \mid \delta \in \Gamma_{k+1}\}$,
- $\text{tr}_k(\delta) = B_k\alpha \wedge M_k\beta \supset \gamma$, and
- $\text{IC}(\sigma) = \bigwedge_{k<n}\bigwedge\{P_{k,n}(\delta) \wedge J_{k,n}(\delta) \mid \delta \in \sigma(k)\}$, s.t.
  - $P_{k,n}(\delta) = B_n\alpha \supset B_k\alpha$,
  - $J_{k,n}(\delta) = B_k\alpha \wedge M_k\beta \supset M_n\beta$.

The following example is also found in (Engan et al. 2005; Brewka and Eiter 2000; Delgrande and Schaub 2000).

Example 36. *Let* $W = \emptyset$ *and* $D = \{\delta_1, \delta_2, \delta_3\}$, *where*

$$\delta_1 = \frac{\top : q}{q}, \quad \delta_2 = \frac{\top : p}{p}, \quad \textit{and} \quad \delta_3 = \frac{\top : \neg q}{p}.$$

These translate to

$$tr_k(\delta_1) = M_k q \supset q;$$
$$tr_k(\delta_2) = M_k p \supset p;$$
$$tr_k(\delta_3) = M_k \neg q \supset p.$$

If the order is empty, this translates to

$$O^R(tr(\delta_1) \wedge tr(\delta_2) \wedge tr(\delta_3)) \twoheadrightarrow \cdot \hat{\twoheadrightarrow} O(p \wedge q).$$

Note that in this particular case, O^R and O give the same expansions. If we let $\delta_3 < \delta_1$, we get two branches in the collapsed tree: $\sigma_1 = \{\delta_2, \delta_3\}\{\delta_1\}$ and $\sigma_2 = \{\delta_3\}\{\delta_1, \delta_2\}$. The translation is now, if we let d_i^k denote $tr_k(\delta_i)$:

$$(O_0^R \top \wedge O_1^R(d_2^1 \wedge d_3^1) \wedge O_2^R(d_1^2 \wedge d_2^1 \wedge d_3^1) \wedge IC(\sigma_1)) \vee$$
$$(O_0^R \top \wedge O_1^R d_3^1 \wedge O_2^R(d_1^2 \wedge d_2^2 \wedge d_3^1) \wedge IC(\sigma_2)).$$

Reducing this formula yields

$$(O_0 \top \wedge O_1 p \wedge O_2(p \wedge q) \wedge$$
$$(\neg M_1 p \vee M_2 p) \wedge (\neg M_1 \neg q \vee M_2 \neg q)) \vee$$
$$(O_0 \top \wedge O_1 p \wedge O_2(p \wedge q) \wedge (\neg M_1 \neg q \vee M_2 \neg q)).$$

Both disjuncts are inconsistent as

$$O_1 p \wedge \neg M_1 \neg q \hat{\rightarrow} \bot \text{ and } O_2(p \wedge q) \wedge M_2 \neg q \hat{\rightarrow} \bot.$$

Hence $(W, D, \delta_3 < \delta_1)$ has no extension.

Conclusion

We have in this paper established a Modal Reduction Theorem for the propositional only-knowing logic of O^R by means of a rewriting system. Since the logic is capable of representing default theories as O^R-formulae, the rewriting system can be used as a calculus to determine the extensions of a default theory. A novelty of the rewriting system is that it clearly separates SAT-solving parts of the algorithm from the modal parts that deal with conflict resolution, and that it thus makes logical structures explicit that are only implicitly present in default theories.

We have also generalized the only-knowing logic to cover confidence levels and sketched a way to represent prescriptively ordered default theories. It is beyond the scope of this paper to treat the representation of ordered default logic in full depth; this is planned in a follow-up paper.

The present work is entirely within the "only-knowing camp" of non-monotonic reasoning. Even though we believe that the rewriting system that we propose gives a better way of computing default extensions than previously known methods within this tradition, we do of course not claim that our method is superior to other approaches to computing defaults.

Future work includes an implementation of the rewriting system in the Rewriting Logic tool Maude, thereby providing a high-level prototype implementation of default logic. This requires a terminating and confluent rewriting system. A more low-level implementation that exploits state of the art incremental SAT-solving techniques is also planned. An implementation of this sort can be compared to implementations using complementary approaches, which may give an indication of how well-suited the proposed rewriting system is for the task of computing default extensions.

Moreover, the logic of $\boxdot$ has not yet been axiomatized. The propositional fragment of the original logic of only-knowing is very well understood, both model-theoretically in terms of, e.g. the finite model property, and proof-theoretically in terms of cut-elimination results (Waaler 2005). It would be of interest to reach the same level of understanding for the logic addressed in this paper.

References

Brewka, G., and Eiter, T. 2000. Prioritizing Default Logic. In *Intellectics and Computational Logic, Papers in Honor of Wolfgang Bibel*, volume 19 of *Applied Logic Series*. Kluwer Academic Publishers. 27–45.

Delgrande, J. P., and Schaub, T. 2000. Expressing Preferences in Default Logic. *Artificial Intelligence* 123:41–87.

Engan, I.; Langholm, T.; Lian, E. H.; and Waaler, A. 2005. Default Reasoning with Preference within Only Knowing Logic. In *Proceedings of LPNMR 2005*, 304–316.

Gottlob, G. 1995. Translating Default Logic into Standard Autoepistemic Logic. *J. ACM* 42(4):711–740.

Lakemeyer, G., and Levesque, H. J. 2005. Only-Knowing: Taking It Beyond Autoepistemic Reasoning. In Veloso, M. M., and Kambhampati, S., eds., *AAAI*, 633–638. AAAI Press AAAI Press / The MIT Press.

Lakemeyer, G., and Levesque, H. J. 2006. Towards an axiom system for default logic. In *AAAI*. AAAI Press.

Levesque, H. J., and Lakemeyer, G. 2001. *The Logic of Knowledge Bases*. MIT Press.

Levesque, H. J. 1990. All I Know: A Study in Autoepistemic Logic. *Artificial Intelligence* 42:263–309.

Lian, E. H.; Langholm, T.; and Waaler, A. 2004. Only Knowing with Confidence Levels: Reductions and Complexity. In *Proceedings of JELIA'04*, volume 3225 of *Lecture Notes in Artificial Intelligence*, 500–512.

Lifschitz, V. 1994. Minimal Belief and Negation as Failure. *Artif. Intell.* 70(1–2):53–72.

Rosati, R. 2001. A Sound and Complete Tableau Calculus for Reasoning about Only Knowing and Knowing at Most. *Studia Logica* 69(1):171–191.

Segerberg, K. 1995. *Some modal reduction theorems in autoepistemic logic*. Uppsala Prints and Preprints in Philosophy. Uppsala University.

Waaler, A.; Klüwer, J. W.; Langholm, T.; and Lian, E. H. 2007. Only knowing with degrees of confidence. *J. Applied Logic* 5(3):492–518.

Waaler, A. 1994. *Logical Studies in Complementary Weak S5*. Doctoral thesis, University of Oslo.

Waaler, A. 2005. Consistency proofs for systems of multi-agent only knowing. *Advances in Modal Logic* 5:347–366.