ELSEVIER

Artificial Intelligence 101 (1998) 1–34

Artificial Intelligence

On the relation between default and modal nonmonotonic reasoning

Alexander Bochman [1]

Department of Computer Systems, Center for Technological Education, 52 Golomb Str., 58102, Holon, Israel

Received 29 July 1996; received in revised form 26 September 1997

Abstract

The notion of a default consequence relation is introduced as a generalization of both default and modal formalizations of nonmonotonic reasoning. It is used to study a general problem of correspondence between these two formalisms. © 1998 Elsevier Science B.V. All rights reserved.

Keywords: Nonmonotonic reasoning; Default logic; Modal nonmonotonic logics; Autoepistemic logic

1. Introduction

Investigations in the field of nonmonotonic and "commonsense" reasoning have given raise to a bewildering diversity of approaches and constructions. It seems that this diversity is gradually becoming a burden for the subsequent development in this field. One of the main purposes of this paper is to show that a number of such approaches are actually different representations of the same basic ideas and intuitions. By showing this, we hopefully pave the way to a future general theory of nonmonotonic reasoning.

The present study [2] pertains mainly to two approaches to nonmonotonic reasoning. One is a default logic, suggested by Raymond Reiter in [25], the other is a modal approach to nonmonotonic reasoning, initiated by McDermott and Doyle in [22]. Default logic is based on the notion of a default rule holding between ordinary classical propositions, and the intended nonmonotonic theories representing plausible "views of the world" are defined via a certain fixed-point construction (see below). The approach

[1] Email: bochman@macs.biu.ac.il.

[2] A preliminary version of the paper has appeared as in the Proceedings of the 4th International Conference on Principles of Knowledge Representation and Reasoning, KR'94, [1].

PII: S0004-3702(98)00011-3

of McDermott and Doyle was based instead on the language of modal logic, though the intended nonmonotonic objects were defined using a similar fixed-point construction. This and the subsequent paper [23] gave rise to a study of autoepistemic logic [24] and of a whole range of modal nonmonotonic logics based on "negative introspection" in the works of Marek, Truszczyński and Schwarz (see, e.g., [16–18,21,27,30]).

Despite the difference in the underlying language, the similarity between the above two approaches to nonmonotonic reasoning were noticed early, and a number of attempts has been made to clarify their relationship. Thus, Konolige [13] attempted to translate default logic into Moore's autoepistemic logic, though the attempt was only partly successful. As will become clear in what follows, the reason for this is that Moore's stable expansions directly correspond not to Reiter's extensions, but to nonmonotonic objects of a different kind. An adequate translation of default logic into a broad range of modal nonmonotonic logics has been proposed by Truszczyński [32,33].

In this paper we will attempt a systematic study of the relation between default and modal formalizations of nonmonotonic reasoning. However, instead of translating these formalisms into each other, we will consider a reformulation of both default logic and modal nonmonotonic logics in the framework of a certain *monotonic* inference system that is based on rules, or sequents, with the form

$$a : b \Vdash A,$$

where a and b are finite sets of propositions. An informal interpretation of such sequents will be

> If all propositions from a are assumed (or believed) to hold and no proposition from b is assumed to hold, then infer A.

We will call such sequents *default rules*, due to their correspondence with the rules of Reiter's default logic [25]. Sets of default rules satisfying certain conditions will be called *default consequence relations*. The main distinctive feature of our formalization, as compared with that of Reiter's, is an explicit use of "meta-rules" allowing to infer new default sequents from given ones. It turns out that for extensions and other "preferred" objects relevant to our study there are important structural rules that preserve these objects. Default consequence relations obtained by adding such rules can be seen as providing a *logical* (*monotonic*) *basis* for different kinds of nonmonotonic reasoning. In addition, they will allow us to give a very simple characterization of extensions and their relatives.

As the next step, we introduce the notion of a *modal* default consequence relation. These consequence relations will be defined in a language with a modal operator, but otherwise will involve the same rules as general default consequence relations. Modal default consequence relations will turn out to be an especially suitable tool for studying modal nonmonotonic reasoning. Thus, both autoepistemic reasoning [24] and reasoning with "negative" introspection [16,18,23] will acquire a natural characterization in this framework.

As has been shown in [2], under certain reasonable conditions modal consequence relations are reducible to the associated nonmodal default consequence relations in a way that preserves the associated nonmonotonic objects. These results will be used here in order to establish a two-way correspondence between modal and default formalizations. Thus, for a number of modal nonmonotonic logics that appear in the literature, we will give a representation in terms of modal default consequence relations. We will show also how and under what conditions objective default consequence relations can be faithfully embedded into modal ones. These latter results provide a natural generalization of the above mentioned Truszczyński's results concerning modal translation of defaults. Finally, we will demonstrate that the modal logic **K45**, associated with autoepistemic reasoning, is in some strong sense equivalent to a certain nonmodal default consequence relation. As a result, we obtain a non-modal, default-type formalization of autoepistemic logic.

2. Default consequence relations

In this section we will introduce the notion of a default consequence relation. It will be defined as a set of default sequents satisfying certain rules that allow to infer new sequents from given ones. In fact, it is these rules that make a set of defaults a proof system. Defaults as such do not bear information about when and how they can be applied on their heads. For ordinary inference rules, this information can be embodied in the form of "meta-rules" that produce new inference rules from given ones (this is actually the main idea behind various sequent calculi). Default consequence relation is an attempt to extend this idea on inference rules that involve as their premises not only what is assumed to hold, but also what is assumed not to hold.

We will assume in this paper that default consequence relations are defined in a classical propositional language with a predefined classical entailment. The corresponding classical consequence operator will be denoted, as usual, by Th.

Definition 2.1. A set of default sequents will be called a *default consequence relation* if it satisfies the following two rules:

(Monotonicity) If $a : b \Vdash A$ and $a \subseteq a'$, $b \subseteq b'$, then $a' : b' \Vdash A$.

(Deductive Closure) If $A \in \mathrm{Th}(c)$ and $a : b \Vdash C_i$, for any $C_i \in c$, then $a : b \Vdash A$.

Default consequence relations can indeed be considered as relations, that is relations between pairs of premise sets, on the one hand, and propositions in conclusions, on the other. Propositions from the first premise set of a default sequent will be called *positive premises*, while those from the second premise set—*negative premises*.

Monotonicity and Deductive Closure provide a primary constraint on our understanding of default sequents. Deductive Closure seems obvious; it says that deductive consequences of provable propositions are also provable. It implies also that if A is a classical tautology, then $: \Vdash A$ belongs to any default consequence relation.

Monotonicity says, in effect, that default sequents are applicable in *all* contexts in which their premises hold. This makes a default consequence relation a fully monotonic (though somewhat unusual) inference system, and immediately distinguish it from, e.g., cumulative and preferential nonmonotonic inference (see [14]) that are not monotonic in this sense. As we will see later, the nonmonotonicity is introduced in our framework not by the default rules themselves, but rather by choosing some "intended" models as representing the meaning of a default consequence relation. It will be shown, however, that Monotonicity is an admissible rule with respect to such models.

As for the classical sequent calculus, the definition of a default consequence relation can be extended to arbitrary sets of propositions in premises of default rules by stipulating that for any possibly infinite sets of propositions u and v,

$$u : v \Vdash A \quad \text{if and only if} \quad a : b \Vdash A,$$

for some finite a, b such that $a \subseteq u$, $b \subseteq v$. This stipulation also ensures that default consequence relations will satisfy the *compactness property*.

The general notion of a default consequence relation is rather uninformative. It is only a frame that can be "filled" with additional rules that would provide a more tight description of our intuitions about nonmonotonic reasoning. As we will see, there is no single system that reflects adequately *all* our intuitions. In fact, different nonmonotonic constructions admit different, even incompatible, reasoning paradigms. Below we will consider a number of rules and conditions that will form the basis for a subsequent classification of various kinds of default reasoning.

To begin with, we introduce the following rule:

$$\text{(Cut)} \qquad \frac{a : b \Vdash A \qquad a, A : b \Vdash B}{a : b \Vdash B}.$$

The rule Cut is nothing other than the usual Cut rule for the classical sequent calculus, though extended to default sequents. It permits the use of inferred propositions as additional positive premises in the proof. As we will see, the rule allows to avoid explicit iterative constructions commonly used in defining nonmonotonic objects (see, e.g., [17,25]). Accordingly, a default consequence relation will be called *iterative* if it satisfies Cut.

The following axiom states that no proposition can serve as both a positive and negative premise in a proof:

$$\text{(Consistency)} \qquad A : A \Vdash \bot,$$

where $\bot$ denotes the proposition "False". The axiom implies that consistent pairs of premise sets must be disjoint. Though this requirement is not universally acceptable (it does not hold, for example, in some semantics for logic programming—see [3]), it will hold for all systems we will consider in this paper.

The following pair of rules reflect the requirement of deductive closure for positive and negative premises, respectively.

$$\text{(Positive Closure)} \qquad \frac{A \in \mathrm{Th}(a) \quad a, A : b \Vdash B}{a : b \Vdash B}.$$

(Negative Closure) If $B \in \mathrm{Th}(c)$ and $a : b, C_i \Vdash A$, for any $C_i \in c$, then $a : b, B \Vdash A$.

Positive Closure implies that deductive consequences of positive premises can be used as additional positive premises, while Negative Closure says that if we reject a proposition, we must reject at least one proposition from any set of propositions that implies it deductively. Negative Closure is reducible to the following rule:[3]

$$\frac{a : b, A \Vdash C \qquad a : b, A \to B \Vdash C}{a : b, B \Vdash C}.$$

One of the main consequences of the above closure rules is the possibility of replacement of deductively equivalent formulas both in positive and in negative premises. In addition, Positive Closure allows to replace sets of positive premises by their conjunctions, while Negative Closure implies the following rule:

$$\frac{a : b, A \Vdash C \qquad a : b, B \Vdash C}{a : b, A \wedge B \Vdash C}.$$

The rule allows to conjoin different sets of negative premises leading to the same conclusion.

A default consequence relation will be called *basic* if it satisfies the above four rules. In the next section we will provide a semantic characterization for such consequence relations.

2.1. Semantics

We introduce first some notation. For a set of propositions u, we will denote by $\overline{u}$ the complement of u. For a given default consequence relation, we will denote by $\mathbb{C}\mathrm{n}(u, v)$ the set of all consequences of the pair of sets (u, v), that is, the set $\{A \mid u : v \Vdash A\}$.

By a *semantics* for a default consequence relation we will mean a set of models. A *model* is a triple (w, u, v), where w, u, v are sets of propositions and w is closed with respect to classical consequence. We will not give an informal interpretation of such models at this point, mainly because of the diversity of their potential interpretations used in the paper.

Let $\mathbb{S}$ be a semantics. Then a default sequent $a : b \Vdash A$ will be said to be *valid* with respect to $\mathbb{S}$ if and only if

$a \subseteq u$ and $b \subseteq \overline{v}$ imply $A \in w$, for any model (w, u, v) from $\mathbb{S}$.

For any semantics $\mathbb{S}$, we will denote by $\Vdash_{\mathbb{S}}$ the set of all default sequents valid with respect to $\mathbb{S}$. Then the following theorem shows that default consequence relations are complete with respect our semantics.

Theorem 2.2 (Completeness Theorem). *$\Vdash$ is a default consequence relation if and only if there is a semantics $\mathbb{S}$ such that $\Vdash$ coincides with $\Vdash_{\mathbb{S}}$.*

The above Completeness Theorem will serve as a basis for a characterization of various extensions of the notion of a default consequence relation, described in what

[3] Similar rules can also be given for Positive Closure.

follows. To begin with, it is easy to show that Consistency amounts to the requirement that in any model (w, u, v), u is included in v, while Cut corresponds to the condition that $w \subseteq u$. Similarly, Positive Closure and Negative Closure correspond, respectively, to conditions that u and v must also be deductively closed sets of propositions. Combining all these conditions, we will obtain a semantic characterization of basic default consequence relations.

A model (w, u, v) will be called *basic*, if w, u, v are deductively closed sets and $w \subseteq u \subseteq v$. By a *basic semantics* we will mean a set of basic models. It is easy to show that all the rules of a basic default consequence relation are valid with respect to such a semantics. Moreover, the following result shows that basic default consequence relations are complete with respect to such a semantic interpretation.

Corollary 2.3. *$\Vdash$ is a basic default consequence relation if and only if there is a basic semantics $\mathbb{S}$ such that $\Vdash$ coincides with $\Vdash_{\mathbb{S}}$.*

Note. A more standard, though equivalent, description of a basic model were obtained, if we would "label" the relevant deductively closed sets by sets of worlds associated with them, just as this was done in [14] for cumulative consequence relations. However, in the context of this paper the suggested formulation seems to be more simple and illuminating (cf. also [7] for a discussion on the relationship between these two forms of representation).

2.2. *Kinds of default reasoning*

A common feature of both default logic and modal nonmonotonic logics is that they have two components. The first component is a logical framework, e.g., some modal logic or a set of defaults. The second, nonmonotonic, component involves a stipulation what set of potential models we should consider as intended, or "preferred" ones. For Reiter's default logic these are extensions, while for autoepistemic logic it is stable expansions. The relation between these two components is usually more complex than in the monotonic case. In usual, monotonic, logical systems the set of all theories (that is, sets of propositions closed with respect to the rules of the system) determines in turn the source provability relation. Unfortunately, this useful property of mutual determination holds neither for default logic nor for modal nonmonotonic formalisms. In both these cases different systems may determine the same set of "preferred" objects and hence the same nonmonotonic inference. What complicates matters still further is that, in general, the set of such objects does not change monotonically with a change of the underlying system. However, we will show that both for default logic and modal nonmonotonic logics there are rules that preserve "preferred" models. Such rules can be considered as providing a logical basis for the corresponding systems of nonmonotonic reasoning.

All the rules and conditions for a basic default consequence relation will hold in all systems discussed in the paper. Now we are going to consider rules that will make a difference. The first is the following Reflexivity axiom:

(Reflexivity) $A : \emptyset \Vdash A$.

Despite its apparent plausibility, Reflexivity does not hold for some natural interpretations of default sequents (e.g., when the premises represent propositions that are or are not *believed*, while conclusions are assumed to be true). The semantic condition corresponding to this rule is that, in any model (w, u, v), w coincides with u (see below).

There is an instance of the above axiom that will be assumed to hold in all cases:

(Positive Consistency) $\bot : \Vdash \bot$.

Positive Consistency implies that consistent pairs of premise sets must include consistent sets of positive premises. The corresponding semantic condition is that w is consistent for any model (w, u, v).

The second controversial rule is a rule that permits "reasoning by cases":

$$\text{(Factoring)} \qquad \frac{a, B : b \Vdash A \qquad a : b, B \Vdash A}{a : b \Vdash A}.$$

The rule implies, in effect, that contexts of reasoning are complete (two-valued) with respect to positive and negative assumptions. It turns out to be characteristic of autoepistemic reasoning (see below). This rule can be characterized semantically by a requirement that, in any model (w, u, v), u coincides with v.

Again, there is an important weaker form of "factoring" that will hold for all systems considered in the paper.

$$\text{(Negative Factoring)} \qquad \frac{a, B : b \Vdash \bot \qquad a : b, B \Vdash A}{a : b \Vdash A}.$$

The rule says that if it is inconsistent to assume a proposition as a positive premise, then it can be assumed as an additional negative premise. In fact, the rule can be seen as a realization of the "negation as inconsistency" principle suggested in [10]. The semantic condition corresponding to this rule is that if (w, u, v) is a model, then there exists a set w' such that (w', v, v) is also a model.

What will happen if we accept all the rules given above? Before we answer this question, let us introduce the following definition.

Definition 2.4. A default consequence relation will be called *stable* if it satisfies Cut, Consistency, Reflexivity and Factoring.

It can be shown that the above four rules imply both Positive Closure and Negative Closure. Hence, stable consequence relations are basic. Moreover, as follows from the semantic characterization of Reflexivity and Factoring, models of such a consequence relation should have the form (u, u, u), that is, all three their components are identical. The following theorem shows that a stable default consequence relation constitutes a limit case—it is already equivalent to an ordinary sequent calculus.

A binary relation $a \vdash b$ between sets of propositions is called a *Scott consequence relation* (see [9]) if it satisfies the following conditions:

(Reflexivity) $A \vdash A$;

(Monotonicity) If $a \vdash b$ and $a \subseteq a'$, $b \subseteq b'$, then $a' \vdash b'$;

(Cut) $$\frac{a \vdash b, A \qquad a, A \vdash b}{a \vdash b}.$$

Theorem 2.5. *Let $\Vdash$ be a stable consequence relation. Define the following consequence relation between sets of propositions:*

$$a \vdash_{\Vdash} b \equiv a : b \Vdash \bot.$$

Then $\vdash_{\Vdash}$ is a Scott consequence relation and $a : b \Vdash A$ if and only if $a \vdash_{\Vdash} b, A$.

A distinctive feature of stable consequence relations, a feature that makes them inappropriate as a basis for nonmonotonic reasoning systems, is the validity of the following rule:

(Symmetry) $$\frac{a : b, B \Vdash A}{a : b, A \Vdash B}.$$

It is this rule that actually reduces default-type sequents to disjunctive, or "multiple-conclusion", rules. Nevertheless, we will see that stable consequence relations constitute an important "upper bound" on reasonable default-type systems. In other words, for reasons that will become clear from what follows, all such systems should contain only rules that are also valid for stable consequence relations.

Thus, the main lesson from the theorem is that in order to obtain nontrivial default consequence relations, we must reject, or weaken, one of the four rules constituting the definition of a stable consequence relation. As we will show below, default logic and modal nonmonotonic logics give rise to two basic kinds of reasoning. One of them, which is associated with autoepistemic logic, involves rejection of Reflexivity. The second kind of reasoning is associated with default logic and modal nonmonotonic logics based on "negative introspection"; it is characterized by rejecting Factoring. Below we will give a brief description of the corresponding systems.

2.3. Autoepistemic consequence relations

We will begin with autoepistemic logic. Though the following description does not involve modal operators, we will see that it provides an adequate characterization of autoepistemic reasoning.

Definition 2.6. A basic default consequence relation will be called *autoepistemic* if it satisfies Factoring and *strongly autoepistemic* if it also satisfies Positive Consistency.

As we mentioned above, Factoring can be characterized semantically by a requirement that, in any model (w, u, v), u must coincide with v. This means, in fact, that such models are reducible to pairs of sets (w, u). This suggests the following alternative definition of the semantics for autoepistemic consequence relations.

A pair of sets of propositions (u, v) will be called a *bimodel* if u and v are deductively closed and $u \subseteq v$.

Definition 2.7. A default sequent $a : b \Vdash A$ will be said to be *A-valid* with respect to a bimodel (u, v) if and only if $A \in u$ whenever $a \subseteq v$ and $b \subseteq \overline{v}$.

The above notion of validity reflects an *autoepistemic interpretation* of default rules according to which $a : b \Vdash A$ says that if all propositions from a are believed, while all propositions are not believed, then A should be true.

By a *binary semantics* we will mean a set of bimodels; a default sequent will be said to be *A-valid* with respect to a binary semantics if it is A-valid with respect to all its bimodels. As before, for any binary semantics $\mathbb{S}$, we will denote by $\Vdash^{A}_{\mathbb{S}}$ the set of all default sequents that are A-valid with respect to $\mathbb{S}$. It is easy to check that this set forms an autoepistemic consequence relation. Moreover, the following result shows that autoepistemic consequence relations are characterized by this semantics.

Theorem 2.8. *$\Vdash$ is an autoepistemic consequence relation if and only if there is a binary semantics $\mathbb{S}$ such that $\Vdash$ coincides with $\Vdash^{A}_{\mathbb{S}}$.*

It can be shown that binary semantics restricted to bimodels of the form (α, v), where α is a world (maximal deductively closed set) are still adequate for autoepistemic consequence relations. This observation allows to establish a correspondence between our semantics and autoepistemic interpretations suggested by Moore in [24], since the latter are also defined as pairs consisting of a world and a theory.

It turns out that autoepistemic consequence relations provide an adequate logical basis for reasoning about the key concepts involved in autoepistemic reasoning. The latter are described in the following definition:

Definition 2.9. Let $\Vdash$ be a default consequence relation.

(1) A set of propositions u will be called *stable* in $\Vdash$ if it is deductively closed and $\mathbb{C}\mathrm{n}(u, \overline{u}) \subseteq u$;

(2) u will be called an *expansion* in $\Vdash$ (or $\Vdash$-expansion) if $u = \mathbb{C}\mathrm{n}(u, \overline{u})$.

Recall that a default consequence relation is a certain set of default sequents. Moreover, since all the rules for such consequence relations, described earlier, have a "Horn" form, for any set of default sequents Γ there always exists a least consequence relation (in the sense of set inclusion) that includes Γ and is closed with respect to such rules.

For an arbitrary consequence relation $\Vdash$, we let $\Vdash^{ae}$ ($\Vdash^{sae}$, $\Vdash^{s}$) denote the least autoepistemic (respectively, the least strongly autoepistemic and the least stable) consequence relation containing $\Vdash$. These consequence relations can be described alternatively as consequence relations obtained from $\Vdash$ by adding the appropriate rules and axioms.

The following theorem (proved in [2]) gives a characterization of default consequence relations that are appropriate for "intended" autoepistemic objects.

Theorem 2.10. *For any default consequence relation* $\Vdash$,

(1) $\Vdash^{\mathrm{s}}$ *has the same stable sets as* $\Vdash$;

(2) $\Vdash^{\mathrm{ae}}$ *has the same stable sets and expansions as* $\Vdash$;

(3) $\Vdash^{\mathrm{sae}}$ *has the same stable sets and consistent expansions as* $\Vdash$.

An immediate consequence of this result is that addition of any of the rules involved in the definition of an autoepistemic consequence relation does not change stable sets and expansions. Similarly, Positive Consistency does not change stable sets and consistent expansions, and Reflexivity does not change stable sets.

It is easy to show that, in any consequence relation satisfying Reflexivity, expansions coincide with stable sets (since for such relations $u \subseteq \mathbb{C}\mathrm{n}(u, v)$, for any u, v). Thus, Reflexivity does not preserve expansions. Moreover, even a weaker form of reflexivity, Positive Consistency, always forces the set of all propositions to be an expansion, though it preserves consistent expansions.

The main conclusion that can be made from the above results is that autoepistemic consequence relations provide an admissible framework for reasoning about stable sets and expansions. In the next section we will present a similar result for default logic.

2.4. *Default logic and reflexive consequence relations*

Reiter [25] defines a default theory as a pair $\Delta = (W, D)$, where W is a set of propositions and D a set of default rules of the form[4] $A : B_1, \ldots, B_k / C$. The connection between default theories and default consequence relations can be established by representing propositions from W as sequents $: \Vdash A$ and default rules from D as sequents

$$A : \neg B_1, \ldots, \neg B_k \Vdash C.$$

This translation will be denoted by $\mathrm{tr}(\Delta)$. As can be seen, it agrees with the informal meaning of default sequents given in the Introduction. Note also that the translation is reversible: a default sequent $A_1, \ldots, A_n : B_1, \ldots, B_k \Vdash C$ is representable by a default rule

$$A_1 \wedge \cdots \wedge A_n : \neg B_1, \ldots, \neg B_k / C.$$

Reiter's default logic is based on the notion of extension. The latter can be defined using a certain iterative construction (see [25, Theorem 2.1]). It turns out that this construction can be captured in our system through the use of the rule Cut given above. The following definition gives a formalization of the notion of extension in the framework of iterative consequence relations, that is, default consequence relations satisfying Cut.

Definition 2.11. Let $\Vdash$ be an iterative consequence relation. A set of propositions u will be called an *extension* in $\Vdash$ (or $\Vdash$-extension) if

$$u = \mathbb{C}\mathrm{n}(\emptyset, \overline{u}).$$

[4] It is interesting to note that a "pseudo-modal" operator **M** that appears in the original formulation of default rules in [25] was later eliminated by Reiter as unnecessary.

Thus, extensions are sets of propositions that are provable in an iterative consequence relation by taking their complements as a set of negative premises. It can be shown (see [2]) that if $\Vdash$ is an iterative consequence relation, then any $\Vdash$-extension is a $\Vdash$-expansion (and hence a $\Vdash$-stable set). The following theorem shows that iterative consequence relations and $\Vdash$-extensions provide a proper generalization of Reiter's default logic.

Theorem 2.12. *Let Δ be a default theory and $\Vdash_\Delta$ the least iterative consequence relation containing* $\mathrm{tr}(\Delta)$. *Then extensions of Δ coincide with $\Vdash_\Delta$-extensions.*

The above result can be informally described as follows. First, we translate a default theory Δ into a set of default sequents $\mathrm{tr}(\Delta)$. Then, for a set of propositions u, we check whether all propositions from u can be proved from $\mathrm{tr}(\Delta)$ by taking $\overline{u}$ as a set of auxiliary negative premises and applying the two rules of a default consequence relation and Cut. If they are, u is an extension of the default theory Δ.

Now we are going to describe a stronger default consequence relation that is appropriate for Reiter's default logic. This consequence relation will play an important role in establishing a correspondence between default logic and modal nonmonotonic logics.

Definition 2.13. A basic default consequence relation will be called *reflexive* if it satisfies Reflexivity and Negative Factoring.

As we mentioned, the semantic condition corresponding to Reflexivity is that, in any model (w, u, v), w coincides with u. Thus, models in this case are also reducible, in effect, to pairs of sets of propositions. This suggests that a semantics for such consequence relations can be given a simpler formulation. To this end, we will employ the notion of a bimodel used earlier in defining semantics for autoepistemic consequence relations.

Definition 2.14. A default sequent $a : b \Vdash A$ will be said to be *R-valid* with respect to a bimodel (u, v) if and only if $A \in u$ whenever $a \subseteq u$ and $b \subseteq \overline{v}$.

As can be seen, this definition of validity differs from the corresponding definition for A-validity only by the condition for positive premises a: while A-validity requires this set to be included in v, R-validity requires its inclusion in the smaller set, u.

As before, by a *binary semantics* we will mean a set of bimodels, and a default sequent will be said to be *R-valid* with respect to a binary semantics if it is R-valid with respect to all its bimodels. However, in this case the presence of Negative Factoring imposes restrictions on admissible sets of bimodels. A binary semantics $\mathbb{S}$ will be called *reflexive* if it satisfies the following additional condition: if $(u, v) \in \mathbb{S}$, then $(v, v) \in \mathbb{S}$.

For a reflexive binary semantics $\mathbb{S}$, we will denote by $\Vdash^{\mathrm{R}}_{\mathbb{S}}$ the set of all default sequents that are R-valid with respect to $\mathbb{S}$. It is easy to show that such a set always forms a reflexive consequence relation. The following theorem shows that reflexive consequence relations are complete with respect to this semantics.

Theorem 2.15. *$\Vdash$ is a reflexive consequence relation if and only if there is a reflexive binary semantics $\mathbb{S}$ such that $\Vdash$ coincides with $\Vdash^{\mathrm{R}}_{\mathbb{S}}$.*

Let $\Vdash^{\mathrm{r}}$ denote the least reflexive consequence relation containing $\Vdash$. The following result has been proved in [2].

Theorem 2.16. *For any iterative consequence relation $\Vdash$, $\Vdash^{\mathrm{r}}$ has the same stable sets and extensions as $\Vdash$.*

Thus, reflexive consequence relations are appropriate for reasoning about extensions. This means, in particular, that the semantics for reflexive consequence relations, described earlier, can thereby be considered as an adequate semantics for Reiter's default rules.

It can be shown that the rule Factoring does not preserve extensions. On the other hand, as we already mentioned, Reflexivity obliterates the distinction between stable sets and expansions. This indicates that expansion- and extension-based kinds of nonmonotonic reasoning are in some sense incompatible—each admits inference steps that are inadmissible in the other. However, the rules common to both autoepistemic and reflexive consequence relations clearly preserve all the objects we have considered, i.e., stable sets, expansions and extensions. This suggests the following definition that will be used in what follows.

Definition 2.17. A basic default consequence relation will be called *introspective* if it satisfies Positive Consistency and Negative Factoring.

It is easy to see that all the rules of an introspective consequence relation belong both to autoepistemic and to reflexive consequence relations. Hence, it follows from the results, stated above, that introspective consequence relations form a representative class with respect to all three kinds of objects we have considered. So, they can be considered as a natural "common part" of autoepistemic and reflexive reasoning.

3. Modal default consequence relations

In this section we will introduce the notion of a modal default consequence relation. As we will see, modal default consequence relations provide a natural logical basis for modal nonmonotonic logics.

Let $\mathcal{L}_L$ be the set of all propositions in a classical propositional language with a modal operator L. For any set of propositions u from $\mathcal{L}_L$, we let Lu denote the set of all propositions of the form LA, where $A \in u$. The notation $\neg u$ will have a similar meaning.

Definition 3.1. A default consequence relation in the language $\mathcal{L}_L$ will be called *modal* if it satisfies the following two modal axioms:

$$A : \Vdash LA \qquad : A \Vdash \neg LA.$$

In the context of modal nonmonotonic logics, the operator L is usually interpreted as an *epistemic* operator, that is, either as an operator of belief (see, e.g., [24]) or as an operator of knowledge (e.g., in [30]). In fact, the belief interpretation is appropriate for autoepistemic logic, while the knowledge interpretation is more appropriate for reflexive consequence relations. Still, to ease the presentation, we will stick for the time being to the term "belief" in describing the meaning of the operator. Then the above two axioms of a modal default consequence relation imply, in effect, that positive premises of any sequent include propositions that are believed and negative premises include propositions that are not believed. [5] Consequently, the following understanding of default sequents $a : b \Vdash A$ in modal default consequence relations will be appropriate:

> If all propositions from a are believed and all propositions from b are *not* believed, then infer A.

This interpretation is in agreement with the following strengthening of the notion of a modal consequence relation:

Definition 3.2. A modal default consequence relation will be called *regular* if it satisfies the following two rules:

$$\frac{a : b \Vdash LA \qquad a, A : b \Vdash B}{a : b \Vdash B},$$

$$\frac{a : b \Vdash \neg LA \qquad a : b, A \Vdash B}{a : b \Vdash B}.$$

The first rule says that believed propositions can serve as additional positive premises, while those that are not believed can be used as additional negative premises.

Now we are going to give a semantic characterization of regular consequence relations. A model (w, u, v) in the language $\mathcal{L}_L$ will be said to be *regular*, if it satisfies the following two conditions:

$$u = \{B \mid LB \in w\}, \qquad v = \{B \mid \neg LB \notin w\}.$$

Accordingly, a semantics will be called to be *regular*, if it consists of regular models. As the following proposition shows, regular semantics provide an adequate characterization of regular default consequence relations.

Theorem 3.3. *$\Vdash$ is a regular modal default consequence relation if and only if there is a regular semantics $\mathbb{S}$ such that $\Vdash$ coincides with $\Vdash_{\mathbb{S}}$.*

As can be seen, the last two components of any regular model are fully determined by modal formulas that belong to its first component, w. This fact can be used to give regular default consequence relations a natural semantic interpretation similar to an autoepistemic interpretation proposed for autoepistemic logic (see [13,24]). By an

[5] Note that any modal default consequence relation satisfies Consistency.

MD-model we will mean any consistent deductively closed set in $\mathcal{L}_L$. For any MD-model w we define the following two sets:

$$w_L = \{B \mid LB \in w\}, \qquad w^L = \{B \mid \neg LB \in w\}.$$

The set w_L can be naturally interpreted as the set of propositions that are *believed* in w, while w^L as the set of propositions that are *not believed* in w. Note that, in contrast to autoepistemic logic, the interpretation is partial with respect to modal propositions. Now, a default sequent $a : b \Vdash A$ will be said to be *valid* with respect to an MD-model w if and only if A belongs to w whenever all propositions from a are believed in w and all propositions from b are not believed in w, that is,

$$a \subseteq w_L \wedge b \subseteq w^L \quad \Rightarrow \quad A \in w.$$

As before, for a set of MD-models $\mathbb{M}$, we will define $\Vdash_{\mathbb{M}}$ as the set of all modal default sequents that are valid with respect to all MD-models from $\mathbb{M}$. It is easy to check that $\Vdash_{\mathbb{M}}$ is a regular default consequence relation. Moreover, it should be clear that any MD-model w can be identified with a regular model $(w, w_L, \overline{w^L})$. Consequently, the above semantic interpretation turns out to be adequate for regular default consequence relations.

Corollary 3.4. *A default consequence relation $\Vdash$ in the language $\mathcal{L}_L$ is regular if and only if there exists a set of MD-models $\mathbb{M}$ such that $\Vdash = \Vdash_{\mathbb{M}}$.*

Thus, regular default consequence relations are determined, in effect, by arbitrary deductively closed sets of modal formulas. This implies, in particular, that modal default consequence relations in general have no "modal content" in the sense that they impose no restriction whatsoever on the modal operator. However, we will see that additional rules of the kind described earlier correspond to well-known modal axioms for L.

Now we are going to show how modal nonmonotonic logics are representable in this formalism.

To begin with, note that the two modal axioms characterizing a modal default consequence relation imply that $\Vdash$-stable sets in such consequence relations are stable sets in the usual sense, that is, they are deductively closed sets satisfying the following two conditions:

- If $A \in u$, then $LA \in u$.
- If $A \notin u$, then $\neg LA \in u$.

Let $\Vdash_u$ denote the least modal default consequence relation containing a set of (modal) propositions u (that is, $\Vdash A$, for any $A \in u$). Clearly, $\Vdash_u$ is simply the set of all sequents obtained from u by applying the two axioms and two rules of modal default consequence relations. The following simple lemma was proved in [2]:

Lemma 3.5. $a : b \Vdash_u A$ *if and only if* $A \in \mathrm{Th}(u \cup La \cup \neg Lb)$.

As an immediate consequence of this lemma, we obtain that $\Vdash_u$-stable sets are those deductively closed sets of propositions s that satisfy the condition

$\mathrm{Th}(u \cup Ls \cup \neg L\bar{s}) \subseteq s.$

Thus, $\Vdash_u$-stable sets are exactly stable sets containing u (see [24]). Similarly, $\Vdash_u$-expansions are sets satisfying the condition

$$s = \mathrm{Th}(u \cup Ls \cup \neg L\bar{s}),$$

and hence they coincide with Moore's stable expansions of u. Consequently, stable sets and stable expansions of any modal theory u coincide with the corresponding objects of the generated modal default consequence relation $\Vdash_u$. Thus, Moore's autoepistemic logic can be adequately translated into the framework of modal default consequence relations. Furthermore, applying now Theorem 2.10, we can infer that autoepistemic logic can be faithfully represented by means of modal autoepistemic consequence relations. In the next section we will complete the picture by demonstrating that the latter exactly correspond to consequence relations based on the modal logic **K45**.

Now we will turn to modal nonmonotonic logics in general. It is easy to show that the rule Cut in modal default consequence relations implies the following rule:

$$(\text{Necessitation}) \qquad \frac{a : b \Vdash A}{a : b \Vdash LA}.$$

(In fact, it can be shown that for regular consequence relations the two rules are equivalent.) Thus, Cut captures the effect of the necessitation rule A/LA in modal logics.

Let $\mathcal{S}$ be a modal logic containing the necessitation rule. We will say that a modal default consequence relation is an *$\mathcal{S}$-consequence relation* if it is an iterative consequence relation such that if A is an instance of a modal axiom from $\mathcal{S}$, then $\Vdash A$. For any set of propositions u, let $\Vdash_u^{\mathcal{S}}$ be the least $\mathcal{S}$-consequence relation containing u. This consequence relation can also be described as the set of all sequents obtained from u by using the axioms of $\mathcal{S}$, the axioms and rules of modal default consequence relations and the Cut rule. The following lemma was also proved in [2]:

Lemma 3.6. *$a : b \Vdash_u^{\mathcal{S}} A$ if and only if $A \in \mathrm{Cn}_{\mathcal{S}}(u \cup La \cup \neg Lb)$.*

As a consequence of the lemma, we obtain that $\Vdash_u^{\mathcal{S}}$-extensions are sets of propositions satisfying the following condition:

$$s = \mathrm{Cn}_{\mathcal{S}}(u \cup \neg L\bar{s}).$$

Thus, $\Vdash_u^{\mathcal{S}}$-extensions coincide with $\mathcal{S}$-expansions of u as defined in [18] (see also [23]). It follows that a modal nonmonotonic reasoning based on "negative introspection" can be also represented in terms of modal default consequence relations and the notion of $\Vdash$-extension. Moreover, Theorem 2.16 implies that modal reflexive consequence relations provide an adequate framework for reasoning of this kind. It is interesting to note that, in the modal case, the inappropriateness of autoepistemic consequence relations for reasoning about extensions follows immediately from the fact (proved in [2]) that for modal autoepistemic consequence relations extensions collapse to expansions. In view of what will be shown below, this result is actually a generalization of the well-known result of Schwarz [26] saying that stable expansions are **K45**-expansions.

In the next section we will consider how and to what extent various modal axioms influence modal nonmonotonic reasoning.

3.1. Modal consequence relations versus modal nonmonotonic logics

In this section we will consider the correspondence between modal nonmonotonic logics and their associated default consequence relations. It follows from the results described above that both autoepistemic logic and modal nonmonotonic logics are representable via modal default consequence relations of a special kind, namely those generated by a set of modal formulas. Consequently, we can restrict our attention to the latter. The following definition provides a formal characterization for such consequence relations.

Definition 3.7. A modal default consequence relation $\Vdash$ will be called *prime* if it coincides with the least iterative consequence relation containing $\mathbb{C}\mathrm{n}(\emptyset, \emptyset)$.

It is easy to show that a modal consequence relation is prime if and only if it is the least iterative consequence relation containing some set of propositions. As has been said, the rule Cut, that characterizes iterative consequence relations, is equivalent to the modal necessitation rule. Consequently, Lemma 3.6 could be replaced by a more general lemma:

Lemma 3.8. *$\Vdash$ is a prime consequence relation if and only if, for any a, b and A, $a : b \Vdash A$ is equivalent to*

$$A \in \mathrm{Cn}_{\mathbf{N}}(\mathbb{C}\mathrm{n}(\emptyset, \emptyset) \cup La \cup \neg Lb),$$

where $\mathrm{Cn}_{\mathbf{N}}$ *denotes the provability operator of the modal logic* **N**.

The lemma shows that, in general, prime modal default consequence relations correspond to modal nonmonotonic logics based on the pure logic of necessitation **N**, that is, a modal logic that has no proper modal axioms and the necessitation rule as the only additional modal rule (see [8]).

An important consequence of the lemma is the following corollary.

Corollary 3.9. *Any prime modal default consequence relation is regular.*

The set $\mathbb{C}\mathrm{n}(\emptyset, \emptyset)$ may include all instances of modal axioms characterizing various modal logics. An important question that arises here is to what extent different modal axioms appearing in $\mathbb{C}\mathrm{n}(\emptyset, \emptyset)$ influence the general properties of the corresponding consequence relation, since, as is well known, different modal logics may determine the same nonmonotonic inference—see [16]. In the rest of this section we will give representation results for a number of well-known modal nonmonotonic logics. It will turn out that most of them possess a simple and natural characterization in terms of different structural rules that hold in the associated default consequence relations.

We begin with demonstrating that prime **K4**-consequence relations can be characterized as consequence relations that satisfy certain *deduction rules*.

Theorem 3.10. *$\Vdash$ is a prime* **K4**-*consequence relation if and only if it is iterative, satisfies* Positive Closure *and the following two modal deduction rules*:

(Positive Deduction) $$\frac{A, a : b \Vdash B}{a : b \Vdash LA \to B},$$

(Weak Negative Deduction) $$\frac{a : A, b \Vdash B}{a : b \Vdash \neg L\bot \wedge L\neg LA \to B}.$$

The theorem gives an example of a correspondence between rules for default consequence relations and usual modal axioms. Note that, given Positive Deduction, the Positive Closure rule can be shown to be equivalent to the modal **K** axiom. In addition, it can be shown that a prime $\mathcal{S}$-consequence relation in general satisfies Positive Deduction if and only if the modal logic $\mathcal{S}$ satisfies the rule

$$\frac{LA \to B}{LA \to LB}.$$

The two deduction rules are rules that permit propositions to be transferred from premises to conclusions. Note that these rules are reversible. Consequently, by successive applications of these rules, any sequent can be transformed to a provable proposition:

Corollary 3.11. *For prime* **K4**-*consequence relations, any default sequent $A_1, \ldots, A_n : B_1, \ldots, B_m \Vdash C$ is equivalent to a provable formula*

$$LA_1 \wedge \cdots \wedge LA_n \wedge [\wedge \neg L\bot] \wedge L\neg LB_1 \wedge \cdots \wedge L\neg LB_m \to C$$

(*where the conjunct $\neg L\bot$ is present only if the set of negative premises is not empty*).

The above result shows that, for any modal logic $\mathcal{S}$ including **K4**, a prime $\mathcal{S}$-consequence relation can be seen simply as an alternative representation of a modal $\mathcal{S}$-theory. This alternative representation, however, will make vivid the structural rules that characterize the corresponding nonmonotonic logic.

It can be shown that prime **K4**-consequence relations satisfy all the rules of introspective consequence relations, except Positive Consistency and Negative Closure. Adding the first rule amounts to addition of the modal **D** axiom $LA \to \neg L\neg A$:[6]

Theorem 3.12. *The following conditions are equivalent*:

(1) *$\Vdash$ is a prime* **KD4**-*consequence relation.*

[6] The next three theorems provide a correct replacement for Theorems 3.8 and 3.10 from [1], which are wrong as stated.

(2) $\Vdash$ *is iterative, satisfies* Positive Closure, Positive Deduction *and the following rule*:

$$\text{(Negative Deduction)} \qquad \frac{a : A, b \Vdash B}{a : b \Vdash L\neg LA \rightarrow B}.$$

(3) $\Vdash$ *satisfies* Positive Deduction *and all the rules of an introspective consequence relation except* Negative Closure.

As can be seen from the above result, prime **KD4**-consequence relations satisfy a stronger rule of negative deduction that does not include the conjunct $\neg\bot$. Consequently, we have:

Corollary 3.13. *For prime* **KD4***-consequence relations, any default sequent* $A_1, \ldots, A_n : B_1, \ldots, B_m \Vdash C$ *is equivalent to a provable formula*

$$LA_1 \wedge \cdots \wedge LA_n \wedge L\neg LB_1 \wedge \cdots \wedge L\neg LB_m \rightarrow C.$$

As can be seen, taking into account the correspondence between default sequents and ordinary default rules described earlier as $\mathrm{tr}(\Delta)$, the above transformation of default sequents into modal formulas is in fact identical with the modal translation of defaults suggested by Truszczyński in [33]. We will return to this translation below when studying the possibility of embedding default consequence relations into modal ones.

Negative Closure corresponds in our context to the Geach axiom **G**:

$$\neg L\neg LA \rightarrow L\neg L\neg A.$$

This axiom imposes a directionality condition on the corresponding accessibility relation. Consequently, modal introspective consequence relations correspond to modal nonmonotonic logics based on **KD4G**:

Theorem 3.14. $\Vdash$ *is a prime* **KD4G***-consequence relation if and only if it is introspective and satisfies* Positive Deduction.

Replacing the **D** axiom by the more strong **T** axiom $LA \rightarrow A$ amounts to adding the (Reflexivity) rule. Note that **KT4G** is nothing other than a well known modal logic **S4.2**. Consequently, we obtain the following characterization of prime **S4.2**-consequence relations:

Theorem 3.15. $\Vdash$ *is a prime* **S4.2***-consequence relation if and only if it is a reflexive consequence relation satisfying* Positive Deduction.

Now we will consider autoepistemic consequence relations. It turns out that for prime **K4**-consequence relations the rule Factoring, which is characteristic of autoepistemic reasoning, is equivalent to the modal **5** axiom $\neg LA \rightarrow L\neg LA$. Moreover, we have that prime **K45**-consequence relations actually *coincide* with modal autoepistemic consequence relations.

Theorem 3.16. *The following conditions are equivalent:*
(1) $\Vdash$ *is a prime* **K45**-*consequence relation.*
(2) $\Vdash$ *is iterative and satisfies* Positive Deduction *and the following rule:*

$$\text{(Strong Negative Deduction)} \qquad \frac{a : A, b \Vdash B}{a : b \Vdash \neg LA \rightarrow B}.$$

(3) $\Vdash$ *is a modal autoepistemic consequence relation.*

Prime **K45**-consequence relations validate a still more strong rule of negative deduction. Note, however, that, since prime **K45**-consequence relations coincide with modal autoepistemic consequence relations, this time both Positive Deduction and the new negative deduction rule are consequences of the rules of a modal autoepistemic consequence relation.

Corollary 3.17. *For prime* **K45**-*consequence relations, any default sequent* $A_1, \ldots, A_n : B_1, \ldots, B_m \Vdash C$ *is equivalent to a provable formula*

$$LA_1 \wedge \cdots \wedge LA_n \wedge \neg LB_1 \wedge \cdots \wedge \neg LB_m \rightarrow C.$$

Viewed from the standpoint of our framework, the replacement of $L\neg L$ by a simpler $\neg L$ in the modal translation of default rules amounts precisely to admitting Factoring as an additional rule appropriate for autoepistemic reasoning.

Since Positive Consistency is equivalent to the **D** axiom and Reflexivity is equivalent to the modal reflexivity axiom, an immediate consequence of the last theorem is the following characterization of strongly autoepistemic and stable modal consequence relations.

Corollary 3.18. *A* $\Vdash$ *is a prime* **KD45**-*consequence relation if and only if it is a strongly autoepistemic modal consequence relation.*

Corollary 3.19. $\Vdash$ *is a prime* **S5**-*consequence relation if and only if it is a stable modal consequence relation.*

The equivalence of prime **S5** and stable consequence relations can now be combined with Theorem 2.5, and we obtain that prime **S5**-consequence relations are equivalent to Scott consequence relations. This fact can be seen as the source of nonmonotonic degeneration of modal nonmonotonic reasoning based on **S5** (cf. [31]).

To end this section, we introduce still another important consequence relation.

As was proved by Schwarz in [27], nonmonotonic modal logics based on **KD45** and **SW5** (known also as **S4.4**) are maximal nonmonotonic logics satisfying certain natural conditions. Schwarz proposed to treat nonmonotonic **SW5** as a plausible candidate for nonmonotonic logic of knowledge.

As we have demonstrated, the first nonmonotonic logic corresponds to strongly autoepistemic consequence relations. It turns out that the characteristic axiom of **SW5**, the so-called "weak" **5** axiom,

$$A \wedge \neg LA \rightarrow L\neg LA,$$

is equivalent in our system to the following "non-modal" rule:

(Conditional Factoring) $$\frac{a, B : b \Vdash A \qquad a : b, B \Vdash A}{a : b \Vdash B \rightarrow A}.$$

We will say that a default consequence relation is *strongly reflexive* if it is reflexive and satisfies Conditional Factoring. We have the following result:

Theorem 3.20. *$\Vdash$ is a prime* **SW5**-*consequence relation if and only if it is a regular strongly reflexive consequence relation.*

The results of this section show that there is a remarkable correspondence between major structural types of default consequence relations and well-known modal nonmonotonic logics. This correspondence can also be considered as a justification of the claim that particular modal axioms, as distinct from ordinary modal propositions, are important for modal nonmonotonic reasoning only to the extent they influence the structural properties of the associated default consequence relations. This claim will find further justification in the next section, where we consider the relationship between modal and objective default consequence relations.

4. Modal versus objective consequence relations

We let $\mathcal{L}_o$ denote the subset of $\mathcal{L}_L$ consisting of all propositions without occurrences of L; such propositions will be called *objective*. For any set of propositions u from $\mathcal{L}_L$, we let u_o denote the set $u \cap \mathcal{L}_o$ and $\overline{u}_o$ the set $\mathcal{L}_o \setminus u$. Note that, for any modal default consequence relation $\Vdash$, its restriction to $\mathcal{L}_o$ is clearly an objective default consequence relation having the same structural rules as $\Vdash$. We will denote this objective subrelation by ${}_o\Vdash$.

All nonmonotonic objects we have considered in this paper are stable sets, and it is well known that the latter are uniquely determined by their objective subsets (kernels). This suggests a possibility of reducing modal nonmonotonic reasoning to nonmodal one. The only question here is whether the reasoning about the kernels can be accomplished entirely in a nonmodal framework. This was the question we considered in [2]. The main result proved there amounts to demonstrating that if $\Vdash$ is a modal introspective consequence relation and u a stable set, then

- u is a $\Vdash$-stable set iff u_o is a ${}_o\Vdash$-stable set,
- u is a $\Vdash$-expansion iff u_o is a ${}_o\Vdash$-expansion,
- u is a ground $\Vdash$-extension if and only if u_o is a ${}_o\Vdash$-extension.

(*Ground* extensions are extensions which are stably minimal, that is, there is no $\Vdash$-stable set v such that $v_o \subset u_o$.)

These results can be reformulated as saying that, for consequence relations that are introspective, the reduction of modal consequence relations to their objective subrelations provides an adequate translation with respect to stable sets, expansions and ground extensions. As to extensions in general, it was shown that, for any introspective consequence relation $\Vdash$, we can construct an objective strongly autoepistemic consequence relation such that its stable sets coincide with kernels of $\Vdash$-stable sets and its expansions are exactly kernels of $\Vdash$-extensions.

Since introspective consequence relations form a representative class of consequence relations with respect to the key nonmonotonic objects, the above reduction provides, in fact, the crucial step in a general translation from modal nonmonotonic logics to default logics. It shows that objective subrelations of modal introspective consequence relations embody all the essential information about the corresponding modal nonmonotonic objects.

We will consider below the reverse problem, namely the problem translating, or embedding, objective default consequence relations into corresponding (prime) modal consequence relations.

It turns out that the maximal "host" modal logic for objective introspective consequence relations is the logic determined by the following Kripke frames: the set of worlds M is the union of three disjoint sets M_1, M_2 and M_3 (where $M_3 \neq \emptyset$) and the accessibility relation is $[(M_1 \cup M_2) \times (M_2 \cup M_3)] \cup (M_3 \times M_3)$. In other words, the corresponding frames are directional frames of depth less or equal 3. We will denote this logic by **KD4I**. This logic contains **KD4** and is included in both **S4F** (see [32]) and **KD45**. It is in fact equivalent to the logic **KD4** · **3B₃** in the classification of [6].

Theorem 4.1. *Any objective introspective consequence relation coincides with the objective subrelation of some prime* **KD4I***-consequence relation.*

In view of the above mentioned results, the embedding is faithful with respect to stable sets, expansions and ground extensions. Note also that by Corollary 3.13 objective sequents in **KD4I**-consequence relations are equivalent to their Truszczyński's translations. Thus, it can be said that Truszczyński's translation of defaults generates an exact translation of objective introspective consequence relations into prime **KD4I**-consequence relations. Moreover, it follows from Theorem 3.14 that prime **KD4G**-consequence relations are already introspective. Consequently, any modal logic in the range (**KD4G**–**KD4I**) can serve as a host logic for such a translation.

The importance of the above theorem lies not only in demonstrating that defaults can be translated into modal formulas. What is especially important for our present study is that it can be used to show that extension of objective introspective consequence relations to modal **KD4I**-consequence relations is *conservative* with respect to provability of objective sequents. In other words, addition of the modal axioms of **KD4I** to an objective introspective consequence relation cannot result in provability of some new objective sequents. Formally, we have:

Theorem 4.2. *Let $\mathfrak{s}$ be an arbitrary set of objective sequents, $\Vdash^{o}_{\mathfrak{s}}$ the least objective introspective consequence relation containing $\mathfrak{s}$, and $\Vdash^{I}_{\mathfrak{s}}$ the least modal* **KD4I***-consequence relation containing $\mathfrak{s}$. Then*

$$\Vdash^{o}_{\mathfrak{s}} = {}_{o}\Vdash^{I}_{\mathfrak{s}}.$$

The theorem says that, given a set of objective sequents $\mathfrak{s}$, an objective sequent is provable from $\mathfrak{s}$ using all the rules and modal axioms that hold in prime **KD4I**-consequence relations if and only if it is provable from $\mathfrak{s}$ using only the basic rules of introspective consequence relations. This result complements the results about reduction of introspective consequence relations to their objective subrelations, discussed earlier in this section, by showing that the latter are autonomous with respect to provability of objective sequents.

Another consequence of the above embedding theorem is the following result for reflexive consequence relations:

Theorem 4.3. *Any objective reflexive consequence relation coincides with the objective subrelation of some prime* **S4F***-consequence relation.*

Again, Theorem 3.15 implies that any modal logic in the interval (**S4.2**–**S4F**) is appropriate for such an embedding.

It is useful to dwell upon the construction used in the proof of the above theorem (see below). As we have shown, reflexive consequence relations are characterized semantically by pairs of deductively closed sets (u, v) such that $u \subseteq v$. We have mentioned also that such a bimodel can be represented alternatively by a pair of sets of *worlds* (U, V) (where $V \subseteq U$) that "label" the components of the bimodel. Now, any such frame can be transformed into an **S4F**-model by way of *defining* an accessibility relation on it as follows:

$$\alpha R \beta \equiv \alpha \in U \setminus V \text{ or } \beta \in V.$$

In some sense, the resulting models can be seen as a *definitional extension* of the notion of an objective bimodel. Note also that such models coincide with (unimodal) knowledge and belief models (K-B-models) from [30]. As follows from the results stated in that paper (see [30, Theorem 4.7]), **S4F** is a maximal logic that admits an adequate translation of defaults. Moreover, the results from [32, 33] can be used to show that there is a *one-to-one correspondence* between extensions of an objective reflexive consequence relation and (objective kernels of) modal extensions of the corresponding modal consequence relation.

The next theorem shows that, as can be expected, the logic **SW5** is a modal counterpart of strongly reflexive consequence relations.

Theorem 4.4. *Any objective strongly reflexive consequence relation coincides with the objective subrelation of some prime* **SW5***-consequence relation.*

Theorem 3.20 can be used this time to show that **SW5** is the only logic that permits the embedding.

Note that the above two theorems also imply that the corresponding objective consequence relations are conservative with respect to their associated modal consequence relations.

Finally, we will consider autoepistemic consequence relations.

Theorem 4.5. *Any objective (strongly) autoepistemic consequence relation coincides with the objective subrelation of some prime* **K(D)45**-*consequence relation.*

As the following theorem shows, for autoepistemic consequence relations we already have a perfect match between objective and modal variants.

Theorem 4.6. *Two modal autoepistemic consequence relations having the same objective subrelations coincide.*

The theorem implies, in particular, that there is a one-to-one correspondence between prime **K45**-consequence relations and objective autoepistemic consequence relations. In other words, we have a full-fledged equivalence between autoepistemic logic and a particular kind of objective default consequence relations. In fact, we have even more. As Konolige demonstrated, for any set of modal propositions there exists a **K45**-equivalent set of disjunctive clauses without nested occurrences of L (see [13, Proposition 3.9]). Now, taking into account the deduction rules that hold for prime **K45**-consequence relations (see Corollary 3.17), any such clause

$$\neg LA_1 \vee \cdots \vee \neg A_n \vee LB_1 \vee \cdots \vee LB_m \vee C$$

can be transformed into an *objective* sequent

$$A_1, \ldots, A_n : B_1, \ldots, B_m \Vdash C.$$

Thus, any set of modal propositions can be assigned an "autoepistemically equivalent" set of objective sequents. For a set of propositions a, let a^s denote the corresponding set of objective sequents. The following theorem shows that provability in **K45** is reducible to provability of objective sequents in autoepistemic consequence relations.

Theorem 4.7. *For any set of modal propositions a and any proposition A, A is provable from a in* **K45** *if and only if any sequent from $\{A\}^s$ is provable from a^s using the rules of an (objective) autoepistemic consequence relation.*

This result shows, in fact, that *the modal logic* **K45** *itself is reducible to objective autoepistemic consequence relations.* As a corollary, we also have a reduction of modal logics **KD45** and **S5** to objective strongly autoepistemic and stable consequence relations, respectively.

Summing up the results described in this section, we can say that there is a two-way correspondence between objective default consequence relations and their modal counterparts. This correspondence establishes, in particular, an equivalence between objective stable sets, expansions and extensions, on the one hand, and modal stable sets, stable expansions and ground modal extensions, on the other. Note, however, that there is no direct correspondence between modal extensions in general and objective extensions. The results of Truszczyński [32, 33] stating that Reiter's extensions are representable as modal $\mathcal{S}$-expansions of the corresponding modal translation do not change this fact, since the relevant modal translation can be shown to generate modal theories *that have* *only ground modal expansions*. In fact, this absence of correspondence should be expected if we notice that objective kernels of modal extensions can be included into one another, which is impossible for ordinary Reiter's extensions. In this respect, modal extensions behave rather as objective *expansions*. And indeed, it was shown in [2] that the set of (modal) extensions of any modal default consequence relation can be always represented as a set of expansions of some autoepistemic consequence relation. A nonmodular translation of a modal theory into an autoepistemic theory has been given, in fact, in [12] as part of a general translation of default logic into an autoepistemic logic. Unfortunately, this translation is far from being simple or illuminating. Actually, an adequate objective representation of the reasoning about modal extensions requires an extension of the very framework of default consequence relations (see below).

5. Conclusions and further issues

We see the notion of a default consequence relation as the main contribution of the paper. As the results presented above demonstrate, it can be considered as a natural generalization of default logic, on the one hand, and modal nonmonotonic logics, on the other. Moreover, we show in [3] that default consequence relations can also serve as a proper logical basis of normal logic programs and hence provide, in effect, a unified framework for all these areas of nonmonotonic reasoning. Using this formalism, various translations and correspondences established between these fields can be recast in the form of straightforward results about equivalence between different nonmonotonic constructions in a single framework.

Default consequence relations have given us a convenient common ground for studying the relationship between (objective) default and modal formalizations of nonmonotonic reasoning. It should be noted, that the suggested translation, or embedding, of different kinds of objective default consequence relations into the corresponding modal logics, as well as the reverse reductions described in [2], have an advantage over earlier attempts in that they are not restricted as such to particular "preferred" nonmonotonic objects. Rather, they establish a direct correspondence between modal and default-based formalizations of different kinds of nonmonotonic *reasoning*.

Both default and modal nonmonotonic formalisms have advantages of their own. For nonmodal default systems it is mainly conceptual simplicity and avoidance of nested layers of modalities. For modal formalisms it is convenience of working with familiar modal constructions, for which the underlying theory and semantics already exist. As the results of the paper show, in most cases we can freely choose each of these formalisms. Note, however, that the results presented here (as well as corresponding results in [16]) show that a modal formalism often introduce distinctions that are irrelevant for nonmonotonic reasoning about the associated constructions; different modal axioms influence such a reasoning only to the extent they change the structural properties of the corresponding default consequence relation. The above results also show why such otherwise esoteric modal logics as **KD4G**, **K(D)45**, **S4.2**, **S4F** and **SW5** occupy a special place in studying modal nonmonotonic reasoning.

There is at least one aspect of the relation between modal and objective formalizations of nonmonotonic reasoning that points out to the need of extending the very formalism of default consequence relations. As we have mentioned, there is no direct correspondence between modal extensions, on the one hand, and objective extensions, on the other. We have shown, however, that modal extensions are naturally representable in the framework of objective *biconsequence relations* introduced in [4]. The latter involve sequents of the form $a : b \Vdash c : d$ that permit multiple conclusions, both positive and negative ones. Such a formalism has been shown to be adequate for representing a modal logic of belief and negation as failure (MBNF), suggested by Lifschitz in [15]. It allows also to represent default theories based on *disjunctive defaults* from [11]. Default consequence relations can be seen as an important special case of biconsequence relations based on sequents that involve only singular positive conclusions. All this suggest that such biconsequence relations can serve as a plausible candidate on the role of a general theory of nonmonotonic reasoning. A preliminary description of such a general theory is presented in [5].

Appendix A. Proofs of the main results

We give here proofs of the new results presented in the paper.

Proof of Theorem 2.2. To begin with, it is easy to check that any set $\Vdash_{\mathbb{S}}$ forms a default consequence relation. For the other direction, we construct a canonical semantics for $\Vdash$.

A pair of sets of propositions (x, y) will be called *saturated* with respect to $\Vdash$ if, for some proposition A, (x, y) is a maximal (with respect to inclusion) pair such that $x : y \nVdash A$. Then a triple (w, u, v) of sets of propositions will be called a *canonical model,* if $(u, \bar{v})$ is a saturated pair and $w = \mathbb{C}\mathrm{n}(u, \bar{v})$. It is easy to see that any canonical model is a model in our sense.

Finally, we define a *canonical semantics* of a default consequence relation $\Vdash$ to be a set of all its canonical models. Now we are going to prove the following lemma:

Auxiliary Lemma. *If* $\Vdash$ *is a default consequence relation, then a default sequent belongs to* $\Vdash$ *if and only if it is valid in the canonical semantics of* $\Vdash$.

Proof. If $a : b \Vdash A$, for some a, b and A, then, due to Monotonicity $A \in w$, for any canonical model (w, u, v) such that $a \subseteq u$ and $b \subseteq \bar{v}$. In the other direction, if $a : b \nVdash A$, then, due to compactness, it is easy to show that (a, b) can be extended to a saturated pair $(u, \bar{v})$ that does not imply A. Consequently, for a canonical model (w, u, v), where $w = \mathbb{C}\mathrm{n}(u, \bar{v})$, we have $a \subseteq u$, $b \subseteq \bar{v}$, but $A \notin w$. $\square$

If $\mathbb{S}_{\Vdash}$ denotes the canonical semantics of $\Vdash$, then the above lemma shows, in effect, that the default consequence relation generated by $\mathbb{S}_{\Vdash}$ coincides with $\Vdash$. This completes the proof of the theorem. $\square$

Proof of Corollary 2.3. We only need to show that the canonical semantics (see the proof of the completeness theorem) of any basic default consequence relation is basic. Let (w,u,v) be a canonical model of a basic default consequence relation $\Vdash$, that is, $w = \mathbb{C}\mathrm{n}(u,\overline{v})$ and $(u,\overline{v})$ is a saturated pair. Consistency implies that, for any saturated pair (x,y), x and y must be disjoint. Hence $u \subseteq v$. Cut implies that if (x,y) is a saturated pair that does not imply B, then, for any proposition A, either A belongs to y or $x : y \nVdash A$. Consequently, $w \subseteq u$. Positive Closure implies that if A is a deductive consequence of x and (x,y) is a saturated pair, then A must belong x. Hence, u is a deductively closed set. Similarly, Negative Closure implies that v must be deductively closed. Thus, (w,u,v) is a basic model, and we are done. □

Proof of Theorem 2.5. We show first that $\vdash_{\Vdash}$ is a Scott consequence relation. Indeed, Reflexivity for $\vdash_{\Vdash}$ follows from Consistency for $\Vdash$, while Cut for $\vdash_{\Vdash}$ follows from Factoring for $\Vdash$. Now assume that $a : b \Vdash A$ holds. Since $a, A : A, b \Vdash \bot$ by Consistency, we obtain $a : A, b \Vdash \bot$ by Cut. Thus, $a \vdash_{\Vdash} b, A$ holds. In the other direction, if $a : A, b \Vdash \bot$, then $a : A, b \Vdash A$ by Deductive Closure. But $A, a : b \Vdash A$ by Reflexivity and hence $a : b \Vdash A$ by Factoring. Thus, $a : b \Vdash A$ if and only if $a \vdash_{\Vdash} b, A$. □

Proof of Theorem 2.8. Let $\Vdash$ be an autoepistemic consequence relation and (w,u,v) its canonical model. Note that, due to Factoring, if (x,y) is a saturated pair with respect to B, then, any proposition A belongs to either x or to y. Hence, $x = \overline{y}$. Consequently, any canonical model has the form (u,v,v). Note also that, since any consistent pair of premises of the form $(v,\overline{v})$ is already saturated, any such canonical model is determined by pairs (u,v) such that $u = \mathbb{C}\mathrm{n}(v,\overline{v})$ and v is a deductively closed set such that $\mathbb{C}\mathrm{n}(v,\overline{v}) \subseteq v$. In other words, canonical models in our case are determined by pairs $(\mathbb{C}\mathrm{n}(v,\overline{v}),v)$, where v is a $\Vdash$-stable set.

We will define a canonical binary semantics for $\Vdash$ as the set of all pairs (u,v) such that (u,v,v) is a canonical model. Note that a default sequent is valid with respect to a canonical model (u,v,v) iff it is A-valid with respect to the corresponding bimodel (u,v). Consequently, due to the Completeness Theorem, a sequent belongs to an autoepistemic consequence relation iff it is A-valid in its canonical binary semantics. □

Proof of Theorem 2.12. Extensions of a default theory are defined as fixed points of a certain operator Γ (see [25], Definition 1). For any set of propositions u, $\Gamma(u)$ is defined as the least deductively closed set of propositions that includes W and such that if $A : b/C$ is a default rule, $A \in \Gamma(u)$ and $\neg b \subseteq \overline{u}$, then C belongs to $\Gamma(u)$. Now let $\mathbb{C}\mathrm{n}_\Delta$ denote the provability operator corresponding to $\Vdash_\Delta$. We will show that, for any set of propositions u, $\mathbb{C}\mathrm{n}_\Delta(\emptyset,\overline{u})$ coincides with $\Gamma(u)$.

Clearly, $\mathbb{C}\mathrm{n}_\Delta(\emptyset,\overline{u})$ is a deductively closed set. Moreover, it contains W, since we have $: \Vdash A$, for any $A \in W$. In addition, if $A : b/C$ is a default rule such that $A \in \mathbb{C}\mathrm{n}_\Delta(\emptyset,\overline{u})$ and $\neg b \subseteq \overline{u}$, then we have $\emptyset : \overline{u} \Vdash_\Delta C$ by (Cut) and consequently $\mathbb{C}\mathrm{n}_\Delta(\emptyset,\overline{u})$ is closed with respect to the default rules of Δ. Now, since $\Gamma(u)$ is the least such set, we obtain $\Gamma(u) \subseteq \mathbb{C}\mathrm{n}_\Delta(\emptyset,\overline{u})$. In order to prove the inverse inclusion, we will make use of the notion of a strong proof introduced in [17] (see also [21]). Generalizing slightly the original definition, we will define a *strong proof* of a proposition A from a set of propositions a in a default theory Δ with respect to the context u as a finite sequence of formulas $A_1, A_2, \ldots, A_n$ such that A_n is A and, for every $1 \leqslant i \leqslant n$, one of the following holds:

(1) A_i belongs to $W \cup a$ or is a tautology,

(2) A_i follows from the preceding formulas using modus ponens,

(3) there is a default rule $A_j : b/A_i$ such that $j < i$ and $\neg b \subseteq \overline{u}$.

It follows from the results stated in [17] that, for any set of propositions u, $\Gamma(u)$ is exactly the set of propositions possessing a strong proof with respect to the context u (that is, when a is empty). Now we will define the following default consequence relation:

$a : b \Vdash^{\mathrm{d}}_{\Delta} A$ iff A has a strong proof from a with respect to the context $\overline{b}$.

The above result can now be formulated as saying that $\Gamma(u)$ coincides with $\mathbb{C}\mathrm{n}^{\mathrm{d}}_{\Delta}(\emptyset, \overline{u})$.

It is easy to show that $\Vdash^{\mathrm{d}}_{\Delta}$ is an iterative consequence relation containing $\mathrm{tr}(\Delta)$ (checking Cut is the only nontrivial step). Since $\Vdash_{\Delta}$ is the least such consequence relation, we have $\Vdash_{\Delta} \subseteq \Vdash^{\mathrm{d}}_{\Delta}$ and consequently $\mathbb{C}\mathrm{n}_{\Delta}(\emptyset, \overline{u}) \subseteq \mathbb{C}\mathrm{n}^{\mathrm{d}}_{\Delta}(\emptyset \overline{u})$, for any u. Hence $\mathbb{C}\mathrm{n}_{\Delta}(\emptyset, \overline{u})$ is included in $\Gamma(u)$.

Now, since $\mathbb{C}\mathrm{n}_{\Delta}(\emptyset, \overline{u}) = \Gamma(u)$, we immediately obtain that fixed points of Γ coincide with $\Vdash_{\Delta}$-extensions. $\square$

Proof of Theorem 2.15. It is easy to check that any default consequence relation of the form $\Vdash^{\mathrm{R}}_{\mathbb{S}}$ is reflexive. Assume now that $\Vdash$ is a reflexive consequence relation and (w, u, v) its canonical model. Due to Reflexivity, if (x, y) is a saturated pair that does not imply B, then, $x \subseteq \mathbb{C}\mathrm{n}(x, y)$. Hence, any canonical model has the form (u, u, v). Moreover, in this case $(\overline{y}, y)$ is also a saturated pair, namely a maximal pair that does not imply $\bot$, and consequently (v, v, v) is also a canonical model. Indeed, in the opposite case $\overline{y} : y \Vdash \bot$ and by Compactness there is a finite set $a \subseteq \overline{y}$ such that $a, x : y \Vdash \bot$. Let A denotes the conjunction of all propositions from a. Since $\overline{y}$ is deductively closed, we have $A \in \overline{y}$ and $A, x : y \Vdash \bot$. In addition, we have $x : y, A \Vdash B$, since (x, y) is a maximal pair that does not imply B. But then Negative Factoring implies $x : y \Vdash B$, which is impossible.

We will define a canonical reflexive binary semantics for $\Vdash$ as the set of all pairs (u, v) such that (u, u, v) is a canonical model. Note that a default sequent is valid with respect to a canonical model (u, v, v) if and only if it is R-valid with respect to the corresponding bimodel (u, v). Consequently, due to the Completeness Theorem, a sequent belongs to a reflexive consequence relation if and only if it is R-valid in its canonical reflexive binary semantics. $\square$

Proof of Theorem 3.3. Let (w, u, v) be a canonical model of a regular modal default consequence relation $\Vdash$. The first modal axiom implies that if $A \in u$, then $LA \in w$. Moreover, by the first regularity rule, either A belongs to u or $u : \overline{v} \nVdash LA$. Consequently, $A \in u$ if and only if $LA \in w$. Similarly, using the second modal axiom and the second regularity rule, it can be shown that in this case $A \in v$ if and only if $\neg LA \notin w$. Hence the canonical semantics of $\Vdash$ is regular, and the assertion follows from the Completeness Theorem. $\square$

Proof of Lemma 3.8. We will show that a modal default consequence relation

$$a : b \Vdash_{\mathrm{m}} A \equiv A \in \mathrm{Cn}_{\mathbf{N}}(\mathbb{C}\mathrm{n}(\emptyset, \emptyset) \cup La \cup \neg Lb)$$

is a least iterative consequence relation containing $\mathbb{C}\mathrm{n}(\emptyset, \emptyset)$. Indeed, for any modal default consequence relation $\Vdash_1$ we have

$$\mathrm{Th}(\mathbb{C}\mathrm{n}(\emptyset, \emptyset) \cup La \cup \neg Lb) \subseteq \mathbb{C}\mathrm{n}_1(a, b)$$

due to Deductive Closure and the modal axioms. Moreover, modal iterative consequence relations satisfy Necessitation and hence $\mathbb{C}\mathrm{n}(a, b)$ is closed with respect to $\mathrm{Cn}_{\mathbf{N}}$. Thus

$$\mathrm{Cn}_{\mathbf{N}}(\mathbb{C}\mathrm{n}(\emptyset, \emptyset) \cup La \cup \neg Lb) \subseteq \mathbb{C}\mathrm{n}(a, b).$$

Hence, the result follows from the definition of a prime default consequence relation. □

Proof of Corollary 3.9. It is easy to see that $\Vdash_{\mathrm{m}}$, as defined in the proof of the preceding lemma, satisfies the two regularity rules. □

Proof of Theorem 3.10. (*From left to right*) If $\Vdash$ is a prime **K4**-consequence relation, then Positive Closure and Cut obviously hold. For Positive Deduction, assume $a : b \nVdash LA \to B$. Then, in view of Lemma 3.6, there exists a **K4**-model M such that all formulas from $La \cup \neg Lb$ are valid in M, but there is a world $\alpha \in M$ such that $LA \to B$ is false in α. Let M_α be a submodel of M generated by α. Clearly, A and all formulas from $La \cup \neg Lb$ are valid in M_α, while B is still not valid. Consequently we have $A, a : b \nVdash B$. Therefore, Positive Deduction is satisfied. For Weak Negative Deduction the proof is similar.

(*From right to left*) We prove first the axioms of **K4**. By the first modal axiom, $A, A \to B, B \Vdash LB$, and hence $A, A \to B \Vdash LB$ by Positive Closure. Applying Positive Deduction, we obtain $\emptyset : \emptyset \Vdash LA \wedge L(A \to B) \to LB$. Thus, the **K**-axiom is satisfied. Similarly, since $A : \emptyset \Vdash LA$, we have $A : \emptyset \Vdash LLA$ by Necessitation and consequently $\emptyset : \emptyset \Vdash LA \to LLA$ by Positive Deduction. Thus, the **4**-axiom also holds.

In order to prove that $\Vdash$ is a prime consequence relation, it is sufficient to show that both deduction rules are reversible, that is, the following rules are also valid:

$$\frac{a : b \Vdash La \to B}{A, a : b \Vdash B}, \qquad \frac{a : b \Vdash \neg L\bot \wedge L\neg LA \to B}{a : b, A \Vdash B}.$$

The first rule immediately follows from the first modal axiom. For the second rule, we need to show that $: A \Vdash \neg L\bot$ is a valid sequent. We have $\emptyset : \emptyset \Vdash L(\bot \to A)$ by Deductive Closure and Necessitation, and hence $\Vdash \neg LA \to \neg\bot$, since $\Vdash$ satisfies the **K** axiom. Using the second modal axiom and Deductive Closure, we finally infer $\emptyset : A \Vdash \neg L\bot$. Now, the second rule above follows from the validity of this sequent and the sequent $\emptyset : A \Vdash L\neg LA$ (the latter follows by Necessitation from the second modal axiom). Note also that the above proof actually holds for all **K**-consequence relations.

Now let $\Vdash'$ be an arbitrary modal iterative consequence relation such that $\mathbb{C}\mathrm{n}(\emptyset, \emptyset) \subseteq \Vdash'$, and assume that $A_1, \ldots, A_n : B_1, \ldots, B_m \Vdash C$ holds. Applying the two deduction rules, we obtain

$$\emptyset : \emptyset \Vdash [\neg L\bot\wedge] LA_1 \wedge \cdots \wedge LA_n \wedge L\neg LB_1 \wedge \cdots \wedge L\neg LB_m \rightarrow C$$

and consequently

$$\emptyset : \emptyset \Vdash' [\neg L\bot\wedge] LA_1 \wedge \cdots \wedge LA_n \wedge L\neg LB_1 \wedge \cdots \wedge L\neg LB_m \rightarrow C.$$

Note that $\Vdash'$ includes the **K** axiom, and hence the inverse deduction rules hold (see the proof above). Consequently $A_1, \ldots, A_n : B_1, \ldots, B_m \Vdash' C$ and hence $\Vdash \subseteq \Vdash'$. Thus, $\Vdash$ is a prime consequence relation. □

Proof of Theorem 3.12. The proof will proceed in the following order: (1) $\Rightarrow$ (3) $\Rightarrow$ (2) $\Rightarrow$ (1).

(1) $\Rightarrow$ (3) By Theorem 3.10, prime **K4**-consequence relations satisfy Positive Closure and Positive Deduction. The rest of the rules can be easily checked using appropriate Kripke models of **KD4**. We omit details.

(3) $\Rightarrow$ (2) We only need to show that Negative Deduction holds. To begin with, we have $A, \neg LA : \emptyset \Vdash LA$ by the first modal axiom, and $LA, A, \neg LA : \emptyset \Vdash \bot$ by Consistency. Applying Cut to these two sequents, we obtain $A, \neg LA : \emptyset \Vdash \bot$. Now assume that $a : b, A \Vdash B$ holds. Then by Negative Factoring $\neg LA, a : b \Vdash B$, and hence by Positive Deduction $a : b \Vdash L\neg LA \rightarrow B$. Thus, Negative Deduction is a valid rule.

(2) $\Rightarrow$ (1) Weak Negative Deduction clearly follows from Negative Deduction. Hence, in view of Theorem 3.10, we only need to show that the **D** axiom holds. Since **K** holds, we have both $\Vdash L\bot \rightarrow LA$ and $\Vdash L\bot \rightarrow L\neg LA$. But $\Vdash L\neg LA \rightarrow \neg LA$ follows by Negative Deduction from the second modal axiom. Hence $\Vdash \neg L\bot$ by Deductive Closure, which is equivalent to the **D** axiom. □

Proof of Theorem 3.14. In view of the previous theorem, we only need to show that in the framework of prime **KD4**-consequence relations, Negative Closure is equivalent to the **G** axiom. It is easy to check that prime **KD4G**-consequence relations satisfy Negative Closure. In the other direction, the second modal axiom and Deductive Closure imply both $: A \Vdash L\neg LA \vee L\neg L\neg A$ and $: \neg A \Vdash L\neg LA \vee L\neg L\neg A$. Applying Negative Closure, we obtain $: \bot \Vdash L\neg LA \vee L\neg L\neg A$ and by Negative Deduction $\Vdash L\neg L\bot \rightarrow (L\neg LA \vee L\neg L\neg A)$. But $L\neg L\bot$ is provable in **KD4**, and hence we have $\Vdash L\neg LA \vee L\neg L\neg A$. Thus, the **G** axiom holds. □

Proof of Theorem 3.15. The result follows immediately from Theorem 3.14 if we notice that, given Positive Deduction, $A : \emptyset \Vdash A$ is equivalent to $\Vdash LA \rightarrow A$. □

Proof of Theorem 3.16. (1) $\Rightarrow$ (3) Theorem 3.10 implies that prime **K45**-consequence relations satisfy Cut and Positive Closure. Thus, we need only to check the validity of Factoring; this can be easily done by constructing an appropriate Kripke model.

(3) $\Rightarrow$ (2) We must show that the corresponding deduction rules hold for autoepistemic consequence relations. Assume that $A, a : b \Vdash B$ holds. Then $A, a : b \Vdash LA \rightarrow B$ by Deductive Closure. Note also that the sequent $a : b, A \Vdash LA \rightarrow B$ follows from the second modal axiom using Monotonicity and Deductive Closure. Applying Factoring to

these two sequents, we obtain $a : b \Vdash LA \to B$. Thus, Positive Deduction is satisfied. In the same way, $a : b, A \Vdash B$ implies $a : b, A \Vdash \neg LA \to B$, while $A, a : b \Vdash \neg LA \to B$ follows from the first modal axiom. Applying Factoring, we obtain $a : b \Vdash \neg LA \to B$. Hence, Strong Negative Deduction also holds.

(2) $\Rightarrow$ (1) In view of Theorem 3.10, we only need to prove the **5** axiom. Applying Necessitation to the second modal axiom, we obtain $: A \Vdash L\neg LA$, and hence by Strong Negative Deduction $\Vdash \neg LA \to L\neg LA$. $\square$

Proof of Theorem 3.20. (*From left to right*) In view of Corollary 3.9, any prime **SW5**-consequence relation is regular. Moreover, in view of Theorem 3.15, we only need to show that such consequence relations satisfy Conditional Factoring. As before, this can be done using appropriate Kripke models.

(*From right to left*) We prove first that Positive Deduction holds. Assume $A, a : b \Vdash B$. Since $LA, a : b \Vdash LA$ by Reflexivity, we can use the first of the regularity rules to obtain $LA, a : b \Vdash B$ and Deductive Closure to obtain $LA, a : b \Vdash LA \to B$. Now we will show that $a : b, LA \Vdash LA \to B$ also holds. We have $A : LA \Vdash LLA$ (by the first modal axiom and Necessitation) and $A : LA \Vdash \neg LLA$ (by the second modal axiom). Hence $A : LA \Vdash \bot$ by Deductive Closure. In addition, the second modal axiom implies $: A, LA \Vdash \neg LA$. Applying now Negative Factoring to these two sequents, we obtain $: LA \Vdash \neg LA$. Finally, applying Deductive Closure, we have $a : b, LA \Vdash LA \to B$.

Applying Conditional Factoring to $LA, a : b \Vdash LA \to B$ and $a : b, LA \Vdash LA \to B$, we obtain $a : b \Vdash LA \to B$. Thus, Positive Deduction holds and consequently $\Vdash$ is a prime **S4**-consequence relation. Now we only need to prove the "weak" **5** axiom, $A \wedge \neg LA \to L\neg LA$. We have $A : \Vdash LA \vee L\neg LA$ by the first modal axiom and Deductive Closure and $: A \Vdash LA \vee L\neg LA$ by the second modal axiom, Necessitation and Deductive Closure. Applying Conditional Factoring to these two sequents, we obtain $\Vdash LA \to (LA \vee L\neg LA)$, which is equivalent to **W5**. $\square$

Proof of Theorem4.1. Let $\Vdash$ be an objective introspective consequence relation and Tr the set of Truszczyński's modal counterparts of the sequents from $\Vdash$, that is, Tr is a set of all formulas of the form

$$LA_1 \wedge \cdots \wedge LA_n \wedge L\neg LB_1 \wedge \cdots \wedge L\neg LB_m \to C,$$

where $A_1, \ldots, A_n : B_1, \ldots, B_m \Vdash C$. Define $\Vdash_{\mathrm{Tr}}$ to be the least **KD4I**-consequence relation containing Tr. Clearly, $\Vdash_{\mathrm{Tr}}$ is a prime consequence relation and $\Vdash \subseteq \Vdash_{\mathrm{Tr}}$.

To show the reverse inclusion, assume that $a : b \nVdash A$, for some (objective) a, b and A. Then there is a basic canonical model (w, u, v) such that $a \subseteq u$, $b \subseteq \overline{v}$ and $A \notin w$. Moreover, due to Negative Closure, there is w' such that (w', v, v) is also a canonical model.

For any set of propositions x, let $\widehat{x}$ be the set of all maximal consistent deductively closed sets ("worlds") containing x. Now, for the above model (w, u, v) we define the following **KD4I**-model $\langle M, R \rangle$: the set of worlds $M = \widehat{w}$, while R is defined as follows for any $\alpha, \beta \in M$:

$$\alpha R \beta \equiv \beta \in \widehat{u} \,\&\, (\alpha \in \widehat{v} \Rightarrow \beta \in \widehat{v}).$$

It is easy to see that R is the intended accessibility relation if M is taken to be the union of the following three disjoint sets: $M_1 = \widehat{w} \setminus \widehat{u}$, $M_2 = \widehat{u} \setminus \widehat{v}$ and $M_3 = \widehat{v}$.

It immediately follows from the above definition that all formulas from $Lu \cup \neg Lv$ are valid in $\langle M, R \rangle$ and that A is not valid in $\langle M, R \rangle$. Now we will show that all Formulas from Tr are valid in $\langle M, R \rangle$. Assume that $C_1, \ldots, C_n : D_1, \ldots, D_m \Vdash E$, but $LC_1 \wedge \cdots \wedge LC_n \wedge L\neg LD_1 \wedge \cdots \wedge L\neg LD_m \to E$ is false in some $\alpha \in M$. Two cases will be considered:

Case 1: *α does not belong to $\widehat{v}$.* Since all LC_i must be true in α, the definition of the accessibility relation gives us that all C_i belong to u. Similarly, since all $L\neg LD_j$ are also true in α, we have that all D_j belong to $\bar{v}$. But then $u : \bar{v} \Vdash E$ and consequently E is valid in $\langle M, R \rangle$, which is impossible, since E must be false in α.

Case 2: *α belongs to $\widehat{v}$.* In this case the same considerations give us that all C_i belong to v and all D_j belong to $\bar{v}$. But $\mathrm{Cn}(v, \bar{v}) \subseteq v$ and consequently $E \in v$—a contradiction with the assumption that E is false in α.

Thus, all formulas from Tr are valid in $\langle M, R \rangle$. An immediate consequence of this fact is that $u : \bar{v} \nVdash_{\mathrm{Tr}} A$ and hence $a : b \nVdash_{\mathrm{Tr}} A$. Therefore, all objective sequents of $\Vdash_{\mathrm{Tr}}$ belong to $\Vdash$ and consequently $\Vdash = {}_0{\Vdash_{\mathrm{Tr}}}$. □

Proof of Theorem 4.2. Since $\mathfrak{s} \subseteq {}_0{\Vdash^I_{\mathfrak{s}}}$ and all rules of $\Vdash^o_{\mathfrak{s}}$ are also rules of ${}_0{\Vdash^I_{\mathfrak{s}}}$, we have $\Vdash^o_{\mathfrak{s}} \subseteq {}_0{\Vdash^I_{\mathfrak{s}}}$. Moreover, by the above embedding result, there exists a prime **KD4I**-consequence relation such that its objective part coincides with $\Vdash^o_{\mathfrak{s}}$. Consequently, ${}_0{\Vdash^I_{\mathfrak{s}}} \subseteq \Vdash^o_{\mathfrak{s}}$, since $\Vdash^I_{\mathfrak{s}}$ is the least prime **KD4I**-consequence relation containing $\mathfrak{s}$. □

Proof of Theorem 4.3. The proof is the same as for Theorem 4.1, but now Reflexivity implies that canonical models has the form (u, u, v), and the corresponding accessibility relation is definable on $\widehat{u}$ as follows:

$$\alpha R \beta \equiv \alpha \in \widehat{v} \Rightarrow \beta \in \widehat{v}.$$

As can be seen, the resulting Kripke models are **S4F** models. □

Proof of Theorem 4.4.. Again, the proof proceeds along the lines of that for Theorem 4.1. The only difficult point consists in demonstrating that possible worlds models generated by canonical bimodels are now **SW5** models. In order to show this, we need the following lemma:

Auxiliary Lemma. *If (u, v) is a canonical bimodel of an objective strongly reflexive consequence relation $\Vdash$, then there exists a maximal deductively closed set α such that $u = \alpha \cap v$.*

Proof. Since $\Vdash$ is reflexive, if (u, v) is a canonical bimodel of $\Vdash$, then (v, v) is also a canonical bimodel. We show first that in our case there are no other bimodels (u', v') such that $u \subseteq u'$ and $v' \subseteq v$. Indeed, an instance of Conditional Factoring is the following rule

$$\frac{A \to B, a : b \Vdash A \qquad a : b, A \to B \Vdash A}{a : b \Vdash A}.$$

Consequently, if $(u, \overline{v})$ is a maximal pair that does not imply A, then, for any B, either $A \to B \in u$ or $A \to B \in \overline{v}$. Now let $(u', \overline{v}')$ be another saturated pair, distinct from (u, v), such that $u \subseteq u'$ and $v' \subseteq v$. Then $u' : \overline{v}' \Vdash A$ (since $(u, \overline{v})$ is a maximal pair not implying A) and hence $A \in u'$. Now if $B \in v$, then $A \to B \in v$ and hence $A \to B \in u$. Consequently $A \to B \in u'$ and therefore $B \in u'$ (since u' is deductively closed). Thus, $v \subseteq u'$, and consequently (u', v') coincides with (v, v).

Let us consider the following set: $u \cup \neg(v \setminus u)$. If this set is consistent, it can be extended to a maximal consistent set, say β. It is easy to see that β is a required set, that is, $u = \beta \cap v$. Indeed, the inclusion from left to right is obvious. Assume that $C \in \beta \cap v$, but $C \notin u$. Then $C \in (v \setminus u)$ and hence $\neg C \in \neg(v \setminus u)$. Consequently $\neg C \in \beta$, which is impossible, since β is consistent.

Assume now that an appropriate α does not exist. The above considerations imply that in this case the set $u \cup \neg(v \setminus u)$ must be inconsistent. This means that $v \setminus u$ contains a finite set of propositions $\{A_1, \ldots, A_n\}$ such that their disjunction belongs to u. Clearly, this set contains at least one pair of propositions such that $A_i \to A_j \notin u$ (otherwise all the set is included in u). Now by (Cut) we have $u : \overline{v} \nVdash A_i \to A_j$, and Conditional Factoring implies that either $A_i, u : \overline{v} \nVdash A_j$ or $u : \overline{v}, A_i \nVdash A_j$. In the first case we have that there exists a saturated pair (x, y) that extends $(u, \overline{v})$ and such that $A_i \in x$ but $A_j \notin x$. However, this is impossible since neither A_i nor A_j belong to u and both belong to v. Similarly, in the second case we would have $A_i \in \overline{v}$. Thus, an appropriate α does exist. □

It follows from the above lemma that $\widehat{u} = \{\alpha\} \cup \widehat{v}$, and hence the possible worlds model corresponding to (u, v) (defined in the proof of the preceding theorem) will be an **SW5** model. □

Proof of Theorem 4.5. Let $\Vdash$ be an objective autoepistemic consequence relation and $\mathbb{S}$ the set of its canonical bimodels, as defined in the proof of Theorem 2.8. For any deductively closed set of objective propositions x, we let x_{m} denote the modal stable set having x as its kernel.

Define the following modal consequence relation:

$$a : b \Vdash_{\mathrm{m}} A \equiv \forall (u, v) \in \mathbb{S}(a \subseteq v_{\mathrm{m}} \,\&\, b \subseteq \overline{v_{\mathrm{m}}} \Rightarrow A \in \mathrm{Th}(u \cup Lv_{\mathrm{m}} \cup \neg L\overline{v_{\mathrm{m}}})).$$

It is easy to show that $\Vdash_{\mathrm{m}}$ is a modal autoepistemic consequence relation. Indeed, the modal axioms and Factoring are obvious, Positive Closure follows from deductive closure of v_{m}, while Cut follows from the fact that $\mathrm{Th}(u \cup Lv_{\mathrm{m}} \cup \neg L\overline{v_{\mathrm{m}}})$ is included in v_{m} (this, in turn, follows from the stability of v_{m} and from the fact that $u \subseteq v$). Moreover, the objective subrelation of $\Vdash_{\mathrm{m}}$ coincides with $\Vdash$. First, it is easy to see that $\Vdash$ is included in $\Vdash_{\mathrm{m}}$. Assume that $a : b \nVdash A$, for some a, b, A from $\mathcal{L}_{\mathrm{o}}$. Since $\Vdash$ is autoepistemic, there exists a canonical bimodel (u, v) such that $A \notin u$, $a \subseteq v$ and $b \subseteq \overline{v}$. Clearly, $A \notin \mathrm{Th}(u \cup Lv_{\mathrm{m}} \cup \neg L\overline{v_{\mathrm{m}}})$, and consequently $a : b \nVdash A$ by the definition of $\Vdash_{\mathrm{m}}$. Thus, $\Vdash$ coincides with ${}_{\mathrm{o}}\Vdash_{\mathrm{m}}$.

Now the theorem follows from Theorem 3.16 which says that prime **K45**-consequence relation coincide with modal autoepistemic consequence relations. The proof for strongly autoepistemic consequence relations is the same except that instead of all canonical bimodels we must consider only bimodels in which v is consistent. □

Proof of Theorem 4.6. Let $\Vdash_1, \Vdash_2$ be modal autoepistemic consequence relations such that, for some a, b, A from $\mathcal{L}_L$, $a : b \Vdash_1 A$ and $a : b \nVdash_2 A$. Using appropriate deduction rules, we obtain that there is a proposition, say B, such that $\Vdash_1 B$ and $\nVdash_2 B$. Now B is **K45**-equivalent to a conjunction of disjunctive clauses without nested L's. Hence there is at least one disjunctive clause C such that $\Vdash_1 C$ and $\nVdash_2 C$. Since the deduction rules are reversible, we finally obtain that there is an objective sequent that distinguish $\Vdash_1$ from $\Vdash_2$. □

Proof of Theorem 4.7. For any set of propositions a let a^{d} denote a set of disjunctive sequents without nested occurrences of L such that a is **K45**-equivalent to a^{d}. Then $a \vdash_{\mathbf{K45}} A$ if and only if all clauses from $\{A\}^{\mathrm{d}}$ are provable from a^{d} in **K45**. This means, in turn, that any clause from $\{A\}^{\mathrm{d}}$ is provable in the least prime **K45**-consequence relation, containing a^{d}. But prime **K45**-consequence relations coincide with modal autoepistemic ones, and in the latter any disjunctive clause of the above kind is equivalent to an objective sequent. This means, in turn, that all sequents from $\{A\}^{\mathrm{s}}$ are provable in the least modal autoepistemic consequence relation containing a^{s}. Moreover, since all sequents in a^{s} and $\{A\}^{\mathrm{s}}$ are objective, sequents from $\{A\}^{\mathrm{s}}$ are provable already in the least objective autoepistemic consequence relation containing a^{s}. In other words, any sequent from $\{A\}^{\mathrm{s}}$ is provable from a^{s} using the characteristic rules of autoepistemic consequence relations. □

References

[1] A. Bochman, On the relation between default and modal consequence relations, in: Proceedings 4th International Conference on Principles of Knowledge Representation and Reasoning, KR'94, Bonn, Germany, Morgan Kaufmann, San Francisco, CA, 1994, pp. 63–74.

[2] A. Bochman, Modal nonmonotonic logics demodalized, Ann. Math. Artif. Intell. 15 (1995) 101–123.

[3] A. Bochman, Default consequence relations as a logical framework for logic programs, in: Proceedings 3rd International Conference on Logic Programming and Nonmonotonic Reasoning, LPNMR'95, Lecture Notes in Artificial Intelligence, Vol. 928, Springer, Berlin, 1995, pp. 245–258.

[4] A. Bochman, On bimodal nonmonotonic logics and their unimodal and nonmodal equivalents, in: Proceedings IJCAI'95, Montreal, Quebec, 1995, pp. 1518–1524.

[5] A. Bochman, Biconsequence relations for nonmonotonic reasoning, in: Proceedings 5th International Conference on Principles of Knowledge Representation and Reasoning, KR'96, Vancouver, BC, Morgan Kaufmann, San Francisco, CA, 1996.

[6] R.A. Bull, K. Segerberg, Basic modal logic, in: D. Gabbay, F. Guenthner (Eds.), Handbook of Philosophical Logic, Reidel, Dordrecht, 1984, pp. 1–88.

[7] J. Dix, D. Makinson, The relationship between KLM and MAK models for nonmonotonic inference operations, J. Logic and Comput. 1 (1992) 131–140.

[8] M.C. Fitting, W. Marek, M. Truszczyński, The pure logic of necessitation, J. Logic and Comput. 2 (1992) 349–373.

[9] D.M. Gabbay, Investigations in Modal and Tense Logics, Reidel, Dordrecht, 1976.

[10] D.M. Gabbay, H.J. Sergot, Negation as inconsistency, J. Logic Programming 1 (1986) 1–35.

[11] M. Gelfond, V. Lifschitz, H. Przymusińska, M. Truszczyński, Disjunctive defaults, in: Proceedings 3rd International Conference on Principles of Knowledge Representation and Reasoning, KR'91, Cambridge, MA, Morgan Kaufmann, San Mateo, CA, 1991, pp. 230–237.

[12] G. Gottlob, The power of beliefs or translating default logic into standard autoepistemic logic, in: G. Lakemeyer, B. Nebel (Eds.), Foundations of Knowledge Representation and Reasoning, Lecture Notes in Artificial Intelligence, Vol. 810, Springer, Berlin, 1994, 133-144.

[13] K. Konolige, On the relation between default and autoepistemic logic, Artificial Intelligence 35 (1988) 343-382.

[14] S. Kraus, D. Lehmann, M. Magidor, Nonmonotonic reasoning, preferential models and cumulative logics, Artificial Intelligence 44 (1990) 167-207.

[15] V. Lifschitz, Minimal belief and negation as failure, Artificial Intelligence 70 (1994) 53-72.

[16] W. Marek, G. Schwarz, M. Truszczyński, Modal nonmonotonic logics: ranges, characterization, computation, J. ACM 40 (1993) 963-990.

[17] W. Marek, M. Truszczyński, Relating autoepistemic and default logics, in: Proceedings 1st International Conference on Principles of Knowledge Representation and Reasoning, KR'89, Toronto, Ont., Morgan Kaufmann, San Mateo, CA, 1989, pp. 276-288.

[18] W. Marek, M. Truszczyński, Modal logic for default reasoning, Ann. Math. Artif. Intell. 1 (1990) 275-302.

[19] W. Marek, M. Truszczyński, Autoepistemic logic, J. ACM 38 (1991) 588-619.

[20] W. Marek, M. Truszczyński, More on modal aspects of default logic, Fund. Inform. 17 (1992) 99-116.

[21] W. Marek, M. Truszczyński, Nonmonotonic Logic, Context-Dependent Reasoning, Springer, Berlin, 1993.

[22] D. McDermott, J. Doyle, Nonmonotonic logic I, Artificial Intelligence 13 (1980) 41-72.

[23] D. McDermott, Nonmonotonic logic II: nonmonotonic modal theories, J. ACM 29 (1982) 33-57.

[24] R.C. Moore, Semantical considerations on non-monotonic logic, Artificial Intelligence 25 (1985) 75-94.

[25] R. Reiter, A logic for default reasoning, Artificial Intelligence 13 (1980) 81-132.

[26] G. Schwarz, Autoepistemic modal logics, in: R. Parikh (Ed.), Proceedings TARK 1990, Morgan Kaufmann, San Mateo, CA, 1990, pp. 97-109.

[27] G. Schwarz, Reflexive autoepistemic logic, Fund. Inform. 17 (1992) 157-173.

[28] G. Schwarz, Bounding introspection in nonmonotonic logics, in: Proceedings 3rd International Conference on Principles of Knowledge Representation and Reasoning, KR'91, Cambridge, MA, Morgan Kaufmann, San Mateo, CA, 1991.

[29] G. Schwarz, Minimal model semantics for nonmonotonic modal logics, in: Proceedings LICS-92, 1992, pp. 34-43.

[30] G. Schwarz, M. Truszczyński, Minimal knowledge problem: a new approach, Artificial Intelligence 67 (1994) 113-141.

[31] M. Tiomkin, M. Kaminski, Nonmonotonic default modal logics, J. ACM 38 (1991) 963-984.

[32] M. Truszczyński, Modal interpretations of default logic, in: J. Myopoulos, R. Reiter (Eds.), Proceedings IJCAI-91, Sydney, Australia, Morgan Kaufmann, San Mateo, CA, 1991, pp. 393-398.

[33] M. Truszczyński, Embedding default logic into modal nonmonotonic logics, in: A. Nerode, W. Marek, V. Subrahmanian (Eds.), Logic Programming and Non-Monotonic Reasoning, MIT Press, Cambridge, MA, 1991, pp. 151-165.