Global Definability in Basic Modal Logic

Maarten de Rijke[1] and Holger Sturm[2]

[1] ILLC, University of Amsterdam, Pl. Muidergracht 24,
1018 TV Amsterdam, The Netherlands. E-mail: mdr@wins.uva.nl

[2] Institut für Philosophie, Logik und Wissenschaftstheorie, LMU München,
Ludwigstraße 31, 80598 München, Germany. E-mail: sturm@cis.uni-muenchen.de

May 26, 1999

Abstract. We present results on global definability in basic modal logic, and contrast our model-theoretic results and proof techniques with known results about local definability.

1 Introduction

In modal logic, the fundamental semantic relation is 'truth of a modal formula φ in a pair $(\mathfrak{M}, w)$,' where $\mathfrak{M}$ is a model and w is a world chosen from this model. Such a pair is called a *pointed* model [5, 18] or a model-world pair [1]. We use $(\mathfrak{M}, w) \Vdash \varphi$ to say that φ is true in $(\mathfrak{M}, w)$. Based on this relation, further semantical concepts are defined as usual. That a formula φ follows from a set of formulas Γ holds, for instance, if for every pointed model $(\mathfrak{M}, w)$ in which all $\psi \in \Gamma$ are true, we have $(\mathfrak{M}, w) \Vdash \varphi$. Or, in modal terms: it is not possible that Γ is true in some possible state without φ being true there as well.

Both the satisfaction relation and the consequence relation reflect a *local* perspective on the modal language and its models. Formulas are evaluated inside models, at some particular world w (the 'actual' world). In contrast, we can take a *global* perspective, under which models, and not pointed models, are regarded as the fundamental semantic units. The global counterparts of the satisfaction and the consequence relation are then defined as follows. A formula φ is said to be *true in a model* $\mathfrak{M}$, abbreviated by $\mathfrak{M} \models \varphi$, if φ is true in every pointed model $(\mathfrak{M}, w)$ based on $\mathfrak{M}$; and φ follows globally from Γ, if for all models $\mathfrak{M}$ with $\mathfrak{M} \models \Gamma$ it holds that $\mathfrak{M} \models \varphi$. As several authors have observed [5, 13, 22], a large number of logical properties, like completeness, canonicity, finite model property, and interpolation, come in two flavors: a local one and a global one. For instance, the notions of local and global consequence do not coincide: $\Box\psi$ follows locally from ψ, but not globally, yet the following general fact holds: a formula φ follows locally from a set Γ iff φ follows globally from the set $\{\Box^n\psi \mid \psi \in \Gamma, n \in \omega\}$. So, in a sense, the global approach can be simulated within the local setting; moreover, by adding some non-standard modal connectives to our language, the other direction holds as well; see [22, Appendix B] for details.

On the other hand, the global perspective is worth exploring for its own sake. Why? First, via a straightforward translation modal formulas φ may be regarded as terms t_φ of an algebraic language, so that φ turns out to be true in a model $\mathfrak{M}$ iff the equality $t_\varphi = 1$ holds in the corresponding algebra $\mathfrak{M}^*$ [5]. Hence, from the algebraic point of view the global setting proves the more natural one. Second, principles of supervenience, as they were applied in different areas of philosophy, are usually discussed on the level of models [12]. Third, sometimes we are interested in certain features of the whole model, independent of the fact whether they are enforceable by pointwise reasoning [1]. And fourth, global constraints on models play a key role in the closely related area of terminological reasoning [6].

The latter two points raise an important question: what is the expressive power of the given modal language? In other words: which properties of structures are definable or describable within the language? The view of properties that underlies this account is purely extensional; roughly, a property is identified with the class of structures which have this property. Therefore, a property, that is, a class K of structures, is said to be *definable* if there is a set of formulas Γ such that K equals the class of structures in which Γ holds. We can rephrase our original question as follows: is there a general characterization of the classes of structures that are definable in the above sense? What this questions asks for is a characterization of the elementary classes of the logic under investigation. Answers to such questions are usually given by stating algebraic closure conditions that are necessary and sufficient for a class K to be definable. Clearly, in the case of modal logic the characterization problem has two versions: a global and a local one. Our paper contains solutions for both versions. In Section 3 we deal with the local version, whereas Section 4 is reserved for the global setting.

Semantics-based translations form a valuable tool for investigating the meta-properties of a logic more deeply, by relating it to other logics. In the case of basic modal logic there exists a natural translation into first-order logic. Using this translation we may view modal logic as a fragment of first-order logic. Moreover, this fragment has a nice semantical characterization in terms of its preservation behavior. A famous result by van Benthem [2] tells us that a first-order formula lies in this fragment if and only if it is preserved under so-called bisimulations (see Section 3). The global counterpart of this result is proved in Section 4. Then, in Section 5 we compare our global definability results to earlier results obtained by Hansoul [11]. We mention a number of definability and preservation results concerning particular classes of models in Sections 6. We conclude with some suggestions for future research in Section 7.

2 Basic Concepts

Fix a countable set $\mathcal{P} := \{p_n \mid n \in \omega\}$ of proposition letters. The set $\mathcal{ML}$ of modal formulas (over $\mathcal{P}$) is defined as the least set X such that every proposition letter from $\mathcal{P}$ belongs to X, X contains the logical constant $\bot$ (falsum), and X is closed under the boolean connectives $\neg$, $\vee$, and $\wedge$ as well as under the modal operators $\Box$ and $\Diamond$.

A *model* for $\mathcal{ML}$ is a triple $\mathfrak{M} = (W, R, V)$, where W is a non-empty set, R a binary relation on W, and V a valuation function from $\mathcal{P}$ into the power set of W. As usual, the truth of modal formulas is defined recursively with respect to pairs $(\mathfrak{M}, w)$, consisting of a model $\mathfrak{M}$ and a distinguished element $w \in W$. The atomic and boolean cases of the definition are clear. For the modal operators we put

$$\begin{array}{lll} (\mathfrak{M}, w) \Vdash \Diamond\varphi & \text{iff} & \text{there is a } w' \in W \text{ with } Rww' \text{ and } (\mathfrak{M}, w') \Vdash \varphi, \\ (\mathfrak{M}, w) \Vdash \Box\varphi & \text{iff} & \text{for all } w' \in W \text{ with } Rww' \text{ we have } (\mathfrak{M}, w') \Vdash \varphi. \end{array}$$

If $\mathfrak{M}$ is a model and φ a modal formula, we use $\mathfrak{M} \models \varphi$ to say that for all $w \in W$ it holds that $(\mathfrak{M}, w) \Vdash \varphi$. If Φ is a set of modal formulas, $\mathfrak{M} \models \Phi$ says that $\mathfrak{M} \models \varphi$ for every $\varphi \in \Phi$. $\mathrm{Mod}(\Phi)$ denotes the class of models $\mathfrak{M}$ such that $\mathfrak{M} \models \Phi$. A class K of models is *globally definable* if there is a set Φ with $\mathsf{K} = \mathrm{Mod}(\Phi)$. Analogously, we call a class K of pointed models *locally definable* if $\mathsf{K} = \{(\mathfrak{M}, w) \mid (\mathfrak{M}, w) \Vdash \Phi\}$, for some $\Phi \subseteq \mathcal{ML}$.

A pointed model $(\mathfrak{M}, w)$ may also be viewed as a first-order model suitable for a first-order vocabulary, say τ, consisting of a countable set $\{P_n \mid n \in \omega\}$ of predicate symbols, a binary relation symbol S and an individual constant c.[1] This, together with the fact that the truth clauses for modal formulas are stated in a first-order metalanguage, suggests a standard translation ST from $\mathcal{ML}$ into the set of first-order sentences over τ:

$$\begin{array}{rcl} ST(p_n) & = & P_n c, \text{for } n \in \omega \\ ST(\bot) & = & \bot \\ ST(\neg\varphi) & = & \neg ST(\varphi) \\ ST(\varphi \vee \psi) & = & ST(\varphi) \vee ST(\psi) \\ ST(\varphi \wedge \psi) & = & ST(\varphi) \wedge ST(\psi) \\ ST(\Diamond\varphi) & = & \exists x(Scx \wedge ST(\varphi)[x/c])^{2,3} \\ ST(\Box\varphi) & = & \forall x(Scx \rightarrow ST(\varphi)[x/c]) \end{array}$$

The following lemma is then proved by an easy induction. In fact, this is the result which allows us to regard modal logic as a fragment of first-order logic.

Lemma 2.1 *For every modal formula φ, every model $\mathfrak{M} = (W, R, V)$ and every $w \in W$: $(\mathfrak{M}, w) \Vdash \varphi$ if and only if $(\mathfrak{M}, w) \Vdash ST(\varphi)$.*

As has been emphasized in the introduction, throughout this paper there are two consequence relations in use, a local one, $\models_l$, and a global one, $\models_g$. The first is defined with respect to pointed models, the second with respect to models. Below, these two relations will also be applied to (sets of) first-order sentences. For instance, let $\Delta \cup \{\alpha\}$ be a set of first-order sentences over τ, then $\Delta \models_l \alpha$ says that for every pointed model $(\mathfrak{M}, w)$, $(\mathfrak{M}, w) \Vdash \Delta$ implies $(\mathfrak{M}, w) \Vdash \alpha$.

[1] This is possible because, neglecting some harmless notational differences, $(\mathfrak{M}, w)$ may be seen as a convenient way of denoting the first-order model $(W, R, (V(p_n))_{n\in\omega}, w)$.

[2] Here, and in the following clause, the variable x is assumed to be the first variable chosen from a list of variables that do not occur in $ST(\varphi)$.

[3] In general, for a first-order formula α and individual terms t_1 and t_2, $\alpha[t_1/t_2]$ denotes the formula one gets by replacing every occurrence of t_2 in α by t_1.

3 Local Definability

In this section we introduce a well-known type of equivalence relations between models, so-called bisimulations. What makes them important is the fact that modal formulas cannot distinguish between bisimilar models, that is, if there is a bisimulation between two models $\mathfrak{M}$ and $\mathfrak{N}$ which relates worlds w and v, then w and v satisfy exactly the same modal formulas. This result is stated in Lemma 3.2. Moreover, recent work has shown that bisimulations form an important tool in modal model theory. A central result along this line is Theorem 3.3 below, which provides an algebraic characterization of the elementary classes of pointed models, that is, the classes of pointed models that are definable by (sets of) modal formulas. This result was first stated and proved in [18].

Definition 3.1 Let $\mathfrak{M} = (W, R, V)$ and $\mathfrak{N} = (W', R', V')$ be models. A non-empty relation $Z \subseteq W \times W'$ is a *bisimulation* between $\mathfrak{M}$ and $\mathfrak{N}$, if Z satisfies the following conditions:

B1 For every $w \in W$ and $v \in W'$, if Zwv then $(\mathfrak{M}, w) \Vdash p_n \Leftrightarrow (\mathfrak{N}, v) \Vdash p_n$, for every $n \in \omega$.

B2 For every $w, w' \in W$ and $v \in W'$, if Zwv and Rww', then there is some $v' \in W'$ such that $R'vv'$ and $Zw'v'$.

B3 For every $w \in W$ and $v, v' \in W'$, if Zwv and $R'vv'$ then there is some $w' \in W$ such that Rww' and $Zw'v'$.

We write $(\mathfrak{M}, w) \sim (\mathfrak{N}, v)$ to say that there is a bisimulation Z from $\mathfrak{M}$ to $\mathfrak{N}$ with Zwv.

If Z is a bisimulation such that for every $v \in W'$ there is some $w \in W$ with Zwv, then Z is called a *surjective* bisimulation from $\mathfrak{M}$ to $\mathfrak{N}$. Z is a *total* bisimulation from $\mathfrak{M}$ to $\mathfrak{N}$, if for every $w \in W$ there is some $v \in W'$ with Zwv.

Lemma 3.2 *Let Z be a bisimulation between $\mathfrak{M}$ and $\mathfrak{N}$ such that Zwv. Then for every modal formula φ, $(\mathfrak{M}, w) \Vdash \varphi$ iff $(\mathfrak{N}, v) \Vdash \varphi$.*

Theorem 3.3 *For a class K of pointed models the following equivalences hold.*

1. *K is (locally) definable by a set of modal formulas iff K is closed under bisimulations and ultraproducts, and the complement of K, abbreviated by $\overline{\mathsf{K}}$, is closed under ultrapowers.*

2. *K is (locally) definable by a single modal formula iff both K and $\overline{\mathsf{K}}$ are closed under bisimulations and ultraproducts.*

Proof. We only give a sketch. For details the reader is referred to [5].

1. The *only if* direction follows from Lemma 3.2 and a well-known result from classical model theory. For the converse suppose that K and $\overline{\mathsf{K}}$ fulfill the closure conditions stated above. Clearly, $\overline{\mathsf{K}}$ is also closed under bisimulations. Define $\Phi := \{\varphi \in \mathcal{ML} \mid \forall(\mathfrak{M}, w) \in \mathsf{K} : (\mathfrak{M}, w) \Vdash \varphi\}$. Obviously,

$\mathsf{K} \subseteq \{(\mathfrak{M}, w) \mid (\mathfrak{M}, w) \models \Phi\}$. For the other direction, assume $(\mathfrak{M}, w) \Vdash \Phi$. Put $\Delta := \{ST(\psi) \mid (\mathfrak{M}, w) \Vdash \psi\}$. It is easy to see that Δ is finitely satisfiable in K, that is, for every finite subset δ of Δ there is a pointed model $(\mathfrak{N}_\delta, v_\delta) \in \mathsf{K}$ in which δ holds. By standard first-order reasoning we find an ultraproduct $(\mathfrak{N}, v) := \prod_{\delta \subseteq_\omega \Delta} (\mathfrak{N}_\delta, v_\delta)/\mathcal{U}$ such that $(\mathfrak{N}, v) \Vdash \Delta$, and, since K is closed under ultraproducts, $(\mathfrak{N}, v) \in \mathsf{K}$. Moreover, by the choice of Δ, $(\mathfrak{N}, v) \equiv_{\mathcal{ML}} (\mathfrak{M}, w)$. Now, a modal version of the Keisler-Shelah theorem tells us that there exist ultrapowers $(\mathfrak{N}', v')$, $(\mathfrak{M}', w')$ of $(\mathfrak{N}, v)$, $(\mathfrak{M}, w)$ respectively such that $(\mathfrak{N}', v') \sim (\mathfrak{M}', w')$. Utilizing the closure conditions on K we obtain $(\mathfrak{N}', v') \in \mathsf{K}$ and $(\mathfrak{M}', w') \in \mathsf{K}$ in turns. Finally, that $\overline{\mathsf{K}}$ is closed under ultrapowers, implies $(\mathfrak{M}, w) \in \mathsf{K}$. This completes the first part of the proof.

2. Again, the *only if* direction is straightforward. The other direction follows from the first claim by an easy compactness argument. $\dashv$

From this theorem we easily obtain van Benthem's bisimulation theorem:

Corollary 3.4 *A first-order sentence α (over τ) is equivalent to the translation of a modal formula if and only if α is locally preserved under bisimulations.*[4]

In [20], Sahlqvist described a general method of transforming a pointed model $(\mathfrak{M}, w)$ into a tree-like model, called the unraveling of $(\mathfrak{M}, w)$, which is bisimilar and thus modally equivalent to the original model.

Definition 3.5 Let $\mathfrak{M} = (W, R, V)$ be a model and $w \in W$. The unraveling of $(\mathfrak{M}, w)$ is the model $(\mathfrak{M}^u, w^u)$, defined as follows. The domain W^u consists of all finite sequences $(w_0, w_1, \ldots, w_n)$ of elements from W such that $w_0 = w$ and Rw_iw_{i+1}, for every $i < n$. Let $R^u(w_0, \ldots, w_n)(v_0, \ldots, v_m)$ if $m = n + 1$, Rw_nv_m and $w_i = v_i$ for every $i \leq n$. Put $(w_0, \ldots, w_n) \in V^u(p)$ if $w_n \in V(p)$. And, finally, let $w^u := (w)$.

It is easy to see that the canonical mapping $f_{(\mathfrak{M},w)} : (w_0, \ldots, w_n) \mapsto w_n$ defines a bisimulation from $(\mathfrak{M}^u, w^u)$ to $(\mathfrak{M}, w)$. Moreover, if the original model $\mathfrak{M}$ is generated by w, then $f_{(\mathfrak{M},w)}$ is surjective.

4 Global Definability

In the previous section we took a *local* perspective on modal logic and its corresponding first-order fragment. Below we adopt what we call a *global* point of view: modal formulas will be considered on the level of *models*. The main result of this section characterizes the classes of models that are (globally) definable by modal formulas. This time, the key notions are ultraproducts and ultrapowers, as before, as well as surjective bisimulations and disjoint unions. The latter are introduced below.

[4]Here we say that a first-order sentence α is *locally preserved* under bisimulations if whenever Z is a bisimulation between models $\mathfrak{M}$ and $\mathfrak{N}$ such that Zwv and $(\mathfrak{M}, w) \Vdash \alpha$, then $(\mathfrak{N}, v) \Vdash \alpha$.

Definition 4.1 Let $\{\mathfrak{M}_i \mid i \in I\}$ be a non-empty family of models. The *disjoint union* of this family, abbreviated by $\biguplus_{i\in I} \mathfrak{M}_i$, is the model $\mathfrak{M} = (W, R, V)$, where

- $W := \bigcup_{i\in I}(W_i \times \{i\})$;
- for $(w, i), (v, j) \in W$, we put $R(w, i)(v, j)$, if $i = j$ and $R_i wv$; and
- we set $(w, i) \in V(p_n)$, for $n \in \omega$ and $(w, i) \in W$, if $w \in V_i(p_n)$.

Obviously, by a straightforward adaptation of the standard translation from Section 2 we may view $\mathcal{ML}$ as a fragment of first-order logic on the (global) level of models as well: we just have to correlate a modal formula φ with the universal closure of its standard translation, that is, with the formula $\forall x(ST(\varphi)[x/c])$. By making use of Lemma 2.1 it is then easy to see that for every model $\mathfrak{M}$, $\mathfrak{M} \models \varphi$ iff $\mathfrak{M} \Vdash \forall x(ST(\varphi)[x/c])$.

As in the local case, the modal fragment of first-order logic has a semantic characterization in terms of its preservation behavior. This time we can prove that a first-order sentence lies in this fragment iff it is preserved under surjective bisimulations and disjoint unions, where a formula φ is said to be *preserved under disjoint unions* if for every non-empty family $\{\mathfrak{M}_i \mid i \in I\}$ of models such that for every $i \in I$, $\mathfrak{M}_i \models \varphi$, we have $\biguplus_{i\in I} \mathfrak{M}_i \models \varphi$.

Lemma 4.2 *For a set Δ of first-order sentences over $\tau \setminus \{c\}$ the following are equivalent:*

1. *There is a set $\Phi \subseteq \mathcal{ML}$ with* $\mathrm{Mod}(\Phi) = \mathrm{Mod}(\Delta)$.
2. $\mathrm{Mod}(\Delta)$ *is closed under disjoint unions and surjective bisimulations.*

Proof. For the implication from 1 to 2 it is sufficient to prove that modal formulas are preserved under disjoint unions and surjective bisimulations. Here, we only deal with disjoint unions; the case for surjective bisimulations can be proved similarly. Let $\varphi \in \mathcal{ML}$ and let $\mathfrak{M} := (W, R, V)$ be the disjoint union of a family $\{\mathfrak{M}_i \mid i \in I\}$ of models, where $\mathfrak{M}_i \models \varphi$ holds for each $i \in I$. Choose $(w, i) \in W$. By definition, $w \in W_i$. Then, our assumption yields $\mathfrak{M}_i \models \varphi$, hence $(\mathfrak{M}_i, w) \Vdash \varphi$. Now it is easy to see that the following clause defines a bisimulation Z between $\mathfrak{M}_i$ and $\mathfrak{M}$: for every $w', w'' \in W_i$ put $Zw'(w'', i)$, if $w' = w''$. Thus we obtain $(\mathfrak{M}, w) \Vdash \varphi$ by Lemma 3.2. Since w was chosen arbitrary, this yields $\mathfrak{M} \models \varphi$.

For the other implication, suppose that Δ is a set of first-order sentences such that $\mathrm{Mod}(\Delta)$ is closed under disjoint unions and surjective bisimulations. Define $\Phi := \{\varphi \in \mathcal{ML} \mid \Delta \models_g \forall x(ST(\varphi)[x/c])\}$. By an application of Lemma 2.1 we obtain $\mathrm{Mod}(\Delta) \subseteq \mathrm{Mod}(\Phi)$. To prove the converse inclusion, let $\mathfrak{M}$ be a model with $\mathfrak{M} \models \Phi$. Choose an ω-saturated elementary extension $\mathfrak{M}' = (W', R', V')$ of $\mathfrak{M}$. Clearly, $\mathfrak{M}' \models \Phi$.

Next we prove the following claim

($\star$) for each $w \in W'$ there is a pointed model $(\mathfrak{N}_w, v_w)$ and a relation Z_w such that $Z_w : (\mathfrak{N}_w, v_w) \sim (\mathfrak{M}', w)$ and $\mathfrak{N}_w \models \Delta$.

Assuming for a moment that ($\star$) has already been proved, we can proceed as follows. First, let $\mathfrak{N} := \biguplus_{w \in W'} \mathfrak{N}_w$. Since Δ is preserved under disjoint unions and $\mathfrak{N}_w \models \Delta$ holds for each $w \in W'$, we infer $\mathfrak{N} \models \Delta$. Second, define a relation Z as follows: for every (v, w) from $\mathfrak{N}$ and every $w' \in W'$ put $Z(v, w)w'$ if $v = w'$. Moreover, by a routine argument, it is easy to prove that Z forms a surjective bisimulation from $\mathfrak{N}$ to $\mathfrak{M}'$. Thus, as Δ is preserved under surjective bisimulations, $\mathfrak{M}' \models \Delta$. As $\mathfrak{M}'$ is an elementary extension of $\mathfrak{M}$, we get $\mathfrak{M} \models \Delta$. This establishes $\mathrm{Mod}(\Phi) \subseteq \mathrm{Mod}(\Delta)$, as desired.

Therefore, to conclude the proof it remains to establish ($\star$). So, let $w \in W'$. Define $\Gamma := \Delta \cup \{ST(\psi) \mid \psi \in \mathcal{ML}, (\mathfrak{M}', w) \Vdash \psi\}$. Γ is finitely satisfiable. For suppose not. Then there is a modal formula ψ such that $(\mathfrak{M}', w) \Vdash \psi$ and $\Delta \models_l \neg ST(\psi)$, hence $\Delta \models_l ST(\neg\psi)$. Since the constant c does not occur in Δ, we obtain $\Delta \models_g \forall x(ST(\neg\psi)[x/c])$. Thus, by definition, $\varphi \in \Phi$. But this contradicts the assumption $\mathfrak{M}' \models \Phi$. Thus Γ is finitely satisfiable, and so, by compactness, satisfiable. Hence, there is a pointed model $(\mathfrak{N}_w, v_w)$ with $(\mathfrak{N}_w, v_w) \models \Gamma$. Without any restrictions we may assume that $(\mathfrak{N}_w, v_w)$ is ω-saturated. From $(\mathfrak{N}_w, v_w) \models \Gamma$ we easily obtain $\mathfrak{N}_w \models \Delta$ and $(\mathfrak{N}_w, v_w) \equiv_{\mathcal{ML}} (\mathfrak{M}', w)$. Now, from modal logic we know that if two ω-saturated models are modally equivalent, then they are bisimilar. Hence there exists a bisimulation Z_w between $(\mathfrak{N}_w, v_w)$ and $(\mathfrak{M}', w)$. This establishes ($\star$) and concludes the proof of the lemma. $\dashv$

Theorem 4.3 *For a class* K *of models the following equivalences hold.*

1. K *is (globally) definable by a set of modal formulas iff* K *is closed under surjective bisimulations, disjoint unions and ultraproducts, and* $\overline{\mathsf{K}}$ *is closed under ultrapowers.*

2. K *is (globally) definable by means of a single modal formula iff* K *is closed under surjective bisimulations and disjoint unions, and both* K *and* $\overline{\mathsf{K}}$ *are closed under ultraproducts.*

Proof. 1. Suppose K is globally definable by a set Φ of modal formulas, that is $\mathsf{K} = \mathrm{Mod}(\Phi)$. Obviously, K is also definable by a set of first-order sentences, namely by $\Delta := \{\forall x(ST(\varphi)[x/c]) \mid \varphi \in \Phi\}$. So, a well-known result from model theory tells us that K is closed under ultraproducts, and $\overline{\mathsf{K}}$ under ultrapowers. That K is closed under disjoint unions and surjective bisimulations is an immediate consequence of Lemma 4.2.

For the other direction assume that K fulfills the closure conditions stated in 1. It follows that K and $\overline{\mathsf{K}}$ are closed under isomorphisms. For note that if a function f is an isomorphism from a model $\mathfrak{M}$ onto a model $\mathfrak{N}$, then f is a surjective bisimulation from $\mathfrak{M}$ to $\mathfrak{N}$, and the inverse of f is a surjective bisimulation from $\mathfrak{N}$ onto $\mathfrak{M}$. From general first-order model theory we infer that K is first-order definable, that is: there is a set Δ of first-order sentences such that $\mathsf{K} = \mathrm{Mod}(\Delta)$. By assumption, $\mathrm{Mod}(\Delta)$ is closed under disjoint unions and surjective bisimulations. An application of Lemma 4.2 concludes the first part of the proof.

2. For the *only if* direction suppose that K is definable by a modal formula φ. Hence, by the first claim, K is closed under surjective bisimulations, disjoint unions and ultraproducts. Further, it is easy to see that $\overline{\mathsf{K}}$ is first-order definable by the single sentence $\neg\forall x(ST(\varphi)[x/c])$. Hence $\overline{\mathsf{K}}$ is closed under ultraproducts. This concludes the *only if* direction.

For the other direction, suppose that K and its complement satisfy the closure conditions stated above. By the first claim there is a set Φ of modal formulas such that $\mathsf{K} = \mathrm{Mod}(\Phi)$. Moreover, as in the proof of the first claim we infer that K and $\overline{\mathsf{K}}$ are closed under isomorphisms. Therefore, $\overline{\mathsf{K}}$ is definable by means of a single first-order sentence α, that is $\overline{\mathsf{K}} = \mathrm{Mod}(\alpha)$. Obviously, the set $\{\alpha\} \cup \{\forall x(ST(\varphi)[x/c]) \mid \varphi \in \Phi\}$ is not satisfiable. Hence, by compactness, there are modal formulas $\varphi_1, \ldots, \varphi_n \in \Phi$ such that $\{\alpha\} \cup \{\forall x(ST(\varphi_1)[x/c]), \ldots, \forall x(ST(\varphi_n)[x/c])\}$ is not satisfiable. It is easy to see that the conjunction $\varphi := (\varphi_1 \wedge \cdots \wedge \varphi_n)$ defines K. $\dashv$

Corollary 4.4 *For a first-order sentence α over $\tau \setminus \{c\}$ the following are equivalent:*

1. *There is a modal formula φ such that $\models \alpha \leftrightarrow \forall x(ST(\varphi)[x/c])$.*
2. *α is preserved under disjoint unions and surjective bisimulations.*

Proof. The claim follows from Lemma 4.2 by compactness. $\dashv$

Our next aim is to relativize Corollary 4.4 to the setting of finite models, which brings us into the area of finite model theory. Though it has its origins in classical model theory and complexity theory, finite model theory has developed into an independent research area with its own methods and a whole stock of fascinating results (see [7] for an excellent overview). The investigation of finite models owes its importance to the deep connections between complexity classes and the expressive power on finite models of particular logics. In addition, the study of finite models forms a natural task in computational linguistics as well as in database theory. In the context of modal logic the viewpoint of finite model theory has first been taken in a paper by Rosen [19]. The main result of this paper shows that van Benthems bisimulation result remains true over the class of finite models. Below we will adopt this result to the global setting.

We need some new notions and notation. We write $\mathfrak{M} \equiv_n^{fo} \mathfrak{N}$ to denote that $\mathfrak{M}$ and $\mathfrak{N}$ agree on all first-order sentences of quantifier rank at most n. Moreover, we need so-called *n-approximations* $\sim_n$ to bisimulations; we say that $(\mathfrak{M}, w) \sim_n (\mathfrak{N}, v)$ if there exists a sequence of relations $Z_0, \ldots, Z_n$, each on the domains of $\mathfrak{M}$ and $\mathfrak{N}$, such that

1. Z_0wv.
2. For all $m < n$, if $Z_m w'v'$, and $Rw'w''$, then there is some v'' in $\mathfrak{N}$ such that $Rv'v''$ and $Z_{m+1}w''v''$ (and vice-versa).
3. For all $m \leq n$, if $Z_m w'v'$, then $(\mathfrak{M}, w') \Vdash p_n$ iff $(\mathfrak{N}, v') \Vdash p_n$, for every $n \in \omega$.

Next, we need a technical lemma that combines a number of results due to Rosen [19] into a single statement.

Lemma 4.5 *There is a unary function f with the following property. Let $r \in \omega$, and assume that $(\mathfrak{M}, w) \sim_{f(r)} (\mathfrak{N}, v)$. Then there exist models $(\mathfrak{M}^*, w^*) \sim (\mathfrak{M}, w)$ and $(\mathfrak{N}^*, v^*) \sim (\mathfrak{M}, w)$ with $(\mathfrak{M}^*, w^*) \equiv_r^{fo} (\mathfrak{N}^*, v^*)$ as displayed in the following diagram:*

$$\begin{tikzcd}
(\mathfrak{M}, w) \arrow[r, dash, "\sim_{f(r)}"] \arrow[d, dash, "\sim"'] & (\mathfrak{N}, v) \arrow[d, dash, "\sim"] \\
(\mathfrak{M}^*, w^*) \arrow[r, dash, "\equiv_r^{fo}"] & (\mathfrak{N}^*, v^*)
\end{tikzcd}$$

Moreover, it may be assumed that there is a surjective bisimulation linking $(\mathfrak{M}^*, w^*)$ to $(\mathfrak{M}, w)$, and a surjective bisimulation linking $(\mathfrak{N}, v)$ to $(\mathfrak{N}^*, v^*)$.

Proof. The proof is a combination of Lemmas 2, 3, and 4 in Rosen [19] plus the well-known observation that every model is a p-morphic image of the disjoint union of its point-generated submodels. $\dashv$

Theorem 4.6 *For a first-order sentence α the following are equivalent:*

1. *There is a modal formula φ such that $\forall x(ST(\varphi)[x/c])$ and α are equivalent.*
2. *α is preserved under surjective bisimulations (between finite models) and finite disjoint unions.*

Proof. The direction from 1 to 2 follows from Theorem 4.3. For the other direction suppose α is a first-order sentence that satisfies the closure conditions stated in 2. Let τ be the smallest (finite) vocabulary in which α lives. For every $n \in \omega$, define

$$\Delta_n := \{\forall x(ST(\varphi)[x/c]) \mid \varphi \in \tau, dg(\varphi) \leq n, \alpha \models \forall x(ST(\varphi)[x/c])\}.$$

Observe that we may assume each Δ_n to be finite.

It suffices to show that, for some $n \in \omega$, $\Delta_n \models \alpha$, for then α will be equivalent to a finite conjunction of formulas of the required universal form (which is itself equivalent to a formula of the required form). To arrive at a contradiction, let us assume that

$$(\dagger) \qquad \Delta_n \not\models \alpha, \text{ for each } n \in \omega.$$

Then, for every $n \in \omega$, there is a model $\mathfrak{M}_n$ such that $\mathfrak{M}_n \models \Delta_n$ and $\mathfrak{M}_n \not\models \alpha$. The following observations will prove to be useful below:

1. Δ_n is true in every point-generated submodel of $\mathfrak{M}_n$.
2. There is a point-generated submodel of $\mathfrak{M}_n$ which falsifies α.

The last observation follows from $\mathfrak{M}_n \not\models \alpha$, by the closure conditions on α. As a consequence, we may assume that $\mathfrak{M}_n$ itself is point-generated, say by w_n. By a well-known result from modal logic, there is a modal formula θ of degree n such that for every pointed model $(\mathfrak{N}, v)$, $(\mathfrak{N}, v) \models \theta$ if and only if there is an n-bisimulation from $(\mathfrak{N}, v)$ to $(\mathfrak{M}_n, w_n)$, in symbols $(\mathfrak{N}, v) \sim_n (\mathfrak{M}_n, w_n)$. It is easy to see that the formula $(\alpha \wedge ST(\theta))$ is satisfiable. Otherwise we would obtain $\alpha \models \forall x(ST(\neg\theta)[x/c])$, hence $\forall x(ST(\neg\theta)[x/c]) \in \Delta_n$, which contradicts $\mathfrak{M}_n \models \Delta_n$. So there is a pointed model $(\mathfrak{N}_n, v_n)$ with $(\mathfrak{N}_n, v_n) \models \theta$, thus $(\mathfrak{N}_n, v_n) \sim_n (\mathfrak{M}_n, w_n)$, and $\mathfrak{N} \models \alpha$. Again, we can assume that $\mathfrak{N}_n$ is point-generated, namely by v_n.

Thus, we have shown the following:

($\star$) for each $n \in \omega$ there are point-generated models $(\mathfrak{N}_n, v_n)$, $(\mathfrak{M}_n, w_n)$ such that $(\mathfrak{N}_n, v_n) \sim_n (\mathfrak{M}_n, w_n)$, $\mathfrak{N}_n \models \alpha$ and $\mathfrak{M}_n \not\models \alpha$.

Now, to arrive at a contradiction, we reason as follows. Let r be the quantifier rank of α. Use ($\star$) with $n = f(r)$, where f is the function mentioned in Lemma 4.5. Then we find models $(\mathfrak{M}_{f(r)}, w_{f(r)}) \not\models \alpha$ and $(\mathfrak{N}_{f(r)}, v_{f(r)}) \models \alpha$ with

$$(\mathfrak{M}_{f(r)}, w_{f(r)}) \sim_{f(r)} (\mathfrak{N}_{f(r)}, v_{f(r)}).$$

By the conclusion of Lemma 4.5, we obtain models $(\mathfrak{M}^*, w^*)$ and $(\mathfrak{N}^*, v^*)$ with

$$(\mathfrak{M}_{f(r)}, w_{f(r)}) \sim (\mathfrak{M}^*, w^*) \equiv_r^{fo} (\mathfrak{N}^*, v^*) \sim (\mathfrak{N}_{f(r)}, v_{f(r)}).$$

As $\mathfrak{M}_{f(r)} \not\models \alpha$ and as there is a surjective bisimulation linking $(\mathfrak{M}^*, w^*)$ to $(\mathfrak{N}_{f(r)}, v_{f(r)})$, we can conclude that $\mathfrak{M}^* \not\models \alpha$. Similarly, we find that $\mathfrak{N}^* \models \alpha$ — but this contradicts $(\mathfrak{M}^*, w^*) \equiv_r^{fo} (\mathfrak{N}^*, v^*)$. Hence, we conclude that our original assumption (†) was mistaken. That is: there exists n with $\Delta_n \models \alpha$, as required. $\dashv$

5 Hansoul's Theorem

In the previous section we presented an algebraic characterization of the elementary classes of $\mathcal{ML}$. This result is not the first global definability result with respect to classes of models. In [11], Hansoul gave an alternative characterization.[5] The purpose of this section is to relate Hansoul's definability result to ours. In Theorem 5.6 we compare the conditions that were imposed by Hansoul on a class K to be definable by a set of modal formulas with the conditions that were used in our Theorem 4.3. We will show that both sets of conditions are equivalent in the sense that a class K satisfies Hansoul's conditions if and only if it satisfies ours.

Before we state Hansoul's result, it will be helpful to recall some modal notions. Let $\mathfrak{M} := (W, R, V)$ be a model. For each $\varphi \in \mathcal{ML}$ we define $V(\varphi)$ as the set $\{w \in W \mid (\mathfrak{M}, w) \Vdash \varphi\}$. By a well-known construction we can associate an algebra $\mathfrak{A}_{\mathfrak{M}} := (A_{\mathfrak{M}}, \cap, \cdot^c, W, l_R)$ with the model $\mathfrak{M}$. This algebra consists of the following ingredients: $A_{\mathfrak{M}}$ is the set of all truth-sets of the form $V(\varphi)$,

[5] In [1] another definability result was proved with respect to infinitary modal languages.

with $\varphi \in \mathcal{ML}$, $\cap$ and $\cdot^c$ are the set-theoretic operations of intersection and complement, respectively, and l_R is the algebraic counterpart of the necessity operator, defined as follows, for $X \in \mathfrak{A}_{\mathfrak{M}}$:

$$l_R(X) := \{w \in W \mid \forall v \in W(Rwv \to v \in X)\}.$$

It is easy to verify that $\mathfrak{A}_{\mathfrak{M}}$ is a *modal* algebra. Then, the canonical structure of $\mathfrak{A}_{\mathfrak{M}}$, $\mathcal{C}(\mathfrak{M})$ in symbols, is the following model $(W_{\mathfrak{M}}, R_{\mathfrak{M}}, V_{\mathfrak{M}})$:

- for $W_{\mathfrak{M}}$, we take the set of ultrafilters over $\mathfrak{A}_{\mathfrak{M}}$;
- for $w, v \in W_{\mathfrak{M}}$, we put $R_{\mathfrak{M}}wv$ if $\forall X \in A_{\mathfrak{M}}(l_R(X) \in u \Rightarrow X \in v)$;
- finally, for $u \in W_{\mathfrak{M}}$, we put $u \in V_{\mathfrak{M}}(p)$ if $V(p) \in u$.

An important feature of $\mathcal{C}(\mathfrak{M})$ is stated in Lemma 5.1. The proof is fairly standard, and we skip it here (for details, see [5]).

Lemma 5.1 *Let $\mathfrak{M}$ be a model. For every modal formula φ and every ultrafilter $u \in W_{\mathfrak{M}}$, $(\mathcal{C}(\mathfrak{M}), u) \Vdash \varphi$ if and only if $V(\varphi) \in u$.*

We are now ready to state Hansoul's result in our own words, avoiding the topological notions used by Hansoul.

Theorem 5.2 (Hansoul's Theorem) *A class K of models is globally definable if and only if K is closed under isomorphisms, generated submodels and disjoint unions, and for each model $\mathfrak{M}$, $\mathfrak{M} \in \mathsf{K}$ iff $\mathcal{C}(\mathfrak{M}) \in \mathsf{K}$.*[6]

Throughout this section we make use of a special class of models, so-called *modally saturated* models. By a modally saturated model we mean a model $\mathfrak{M} := (W, R, V)$ that satisfies the following condition: for every $w \in W$ and $\Phi \subseteq \mathcal{ML}$, if for each finite subset $\Phi' \subseteq \Phi$ there is some $v \in W$ with Rwv and $(\mathfrak{M}, v) \Vdash \Phi'$, then there is some $v \in W$ with Rwv and $(\mathfrak{M}, v) \Vdash \Phi$.

For things to come the reader should have the following two facts concerning modally saturated models in mind:

(i) Modal saturation is a weaker notion than ω-saturation; that means that every ω-saturated model is modally saturated, but the converse fails.

(ii) The class of modally saturated models forms a so-called Hennessy-Milner class, that is a class of models in which the relation of modal equivalence between pointed models from this class is a bisimulation.

Both results are well-known; for proofs the reader may consult [5].

The main result of this section, Theorem 5.6, will, essentially, be proved by combining the following three lemmas.

[6]The experienced reader may realize that Hansoul's result shows great similarity to a famous result by Goldblatt [9], which characterizes modally definable classes of generalized frames.

Lemma 5.3 *Let $\mathfrak{M} := (W, R, V)$ be a model and let $\mathfrak{N} := (W', R', V')$ be an ω-saturated elementary extension of $\mathfrak{M}$. Then there exists a total and surjective bisimulation from $\mathfrak{N}$ to $\mathcal{C}(\mathfrak{M})$.*

Proof. Suppose $\mathfrak{N}$ is an ω-saturated elementary extension of a model $\mathfrak{M}$. By (i), $\mathfrak{N}$ is modally saturated. Further, by a routine argument, it is easy to check that $\mathcal{C}(\mathfrak{M})$ is modally saturated as well. Next, define Z as follows: for $w \in W'$ and $v \in W_{\mathfrak{M}}$, put Zwv if $(\mathfrak{N}, w) \equiv_{\mathcal{ML}} (\mathcal{C}(\mathfrak{M}), v)$. According to (ii), Z is a bisimulation from $\mathfrak{N}$ to $\mathcal{C}(\mathfrak{M})$. It remains to show that Z is total and surjective.

For totality, choose $w \in W'$. Define $\Delta_w := \{\varphi \in \mathcal{ML} \mid (\mathfrak{N}, w) \Vdash \varphi\}$. Let $\varphi \in \Delta_w$. Then $(\mathfrak{N}, w) \Vdash \varphi$. Hence $(\mathfrak{N}, w) \Vdash ST(\varphi)$, by Lemma 2.1. Thus $\mathfrak{N} \Vdash \exists x(ST(\varphi)[x/c])$. Since $\mathfrak{N}$ is an elementary extension of $\mathfrak{M}$, we infer $\mathfrak{M} \Vdash \exists x(ST(\varphi)[x/c])$. So there is some $v \in W$ such that $(\mathfrak{M}, v) \Vdash \varphi$. Since Δ_w is closed under conjunction, it follows that Δ_w is finitely satisfiable in $\mathfrak{M}$. From this we easily obtain that $X := \{V(\varphi) \mid \varphi \in \Delta_w\}$ has the finite intersection property. Moreover, $X \subseteq A_{\mathfrak{M}}$. Hence, there is some ultrafilter $u \in W_{\mathfrak{M}}$ with $X \subseteq u$. Thus, we get $V(\varphi) \in u$, for every $\varphi \in \Delta_w$. By an application of Lemma 5.1 we then obtain $(\mathcal{C}(\mathfrak{M}), u) \Vdash \varphi$, for every $\varphi \in \Delta_w$. This establishes $(\mathfrak{N}, w) \equiv_{\mathcal{ML}} (\mathcal{C}(\mathfrak{M}), u)$, which yields, by definition, Zwu. So we have shown that Z is total.

To prove that Z is surjective, let $u \in W_{\mathfrak{M}}$. Put $\Gamma_u := \{\varphi \in \mathcal{ML} \mid (\mathcal{C}(\mathfrak{M}), u) \Vdash \varphi\}$. Choose $\varphi \in \Gamma_u$. By Lemma 5.1, $V(\varphi) \in u$. Thus $V(\varphi) \neq \emptyset$. Hence, there is some $w \in W$ such that $(\mathfrak{M}, w) \Vdash \varphi$. From this we deduce that Γ_u is finitely satisfiable in $\mathfrak{M}$; for note that Γ_u is closed under conjunctions. Clearly, then, Γ_u is finitely satisfiable in $\mathfrak{N}$ as well. Since $\mathfrak{N}$ is ω-saturated, Γ_u is satisfiable in $\mathfrak{N}$. Hence there is some $v \in W'$ with $(\mathfrak{N}, v) \Vdash \Gamma_u$, and therefore, Zvu. So Z is surjective. This completes the proof of the lemma. $\dashv$

Lemma 5.4 *Let $\mathfrak{M} := (W, R, V)$ and $\mathfrak{M}' := (W', R', V')$ be models, and let Z be a total and surjective bisimulation from $\mathfrak{M}$ to $\mathfrak{M}'$. Then $\mathcal{C}(\mathfrak{M})$ and $\mathcal{C}(\mathfrak{M}')$ are isomorphic.*

Proof. Suppose $\mathfrak{M}$, $\mathfrak{M}'$ and Z satisfy the conditions stated above. It suffices to show that the associated algebras $\mathfrak{A}_{\mathfrak{M}}$ and $\mathfrak{A}_{\mathfrak{M}'}$ are isomorphic. To prove this, we only need to verify that the mapping $f : V(\varphi) \mapsto V'(\varphi)$ is an isomorphism between $\mathfrak{A}_{\mathfrak{M}}$ and $\mathfrak{A}_{\mathfrak{M}'}$. Obviously, the domain of f is $A_{\mathfrak{M}}$.

To see that f is a function, suppose that $V(\varphi) = V(\psi)$. Choose $v \in V'(\varphi)$. As Z is surjective, there is some $w \in W$ with Zwv. Hence $w \in V(\varphi)$. Thus $w \in V(\psi)$, by assumption. From this we obtain $v \in V'(\psi)$. Therefore, $V'(\varphi) \subseteq V'(\psi)$. A similar argument yields the other direction. So $V'(\varphi) = V'(\psi)$. Hence f is a function.

The injectivity of f can be proved in a similar way. This time one makes use of the fact that Z is total. It remains to show that f is a homomorphism in the sense that it commutes with the algebraic operations $\cap$, $\cdot^c$ and l_R. But this is an easy task which we leave to the reader. $\dashv$

Lemma 5.5 *Let* K *be a class of models that is closed under isomorphisms, generated submodels, disjoint unions and ultrapowers. Then* K *is closed under ultraproducts.*

Proof. The claim of this lemma is a version of a well-known result from the theory of frames due to Goldblatt [9]. We only need to replace the family M by a family of frames, and we immediately get Goldblatt's Theorem 1.16.4. Since Goldblatt's proof can easily be adjusted to the context of models, we dispense with a proof. However, we will give a hint that should help the reader to find the proof: Suppose $\mathsf{M} := \{\mathfrak{M}_i \mid i \in I\}$ is a non-empty family of models and suppose $\mathfrak{N}$ is an ultraproduct of this family. Then one has to show that $\mathfrak{N}$ is isomorphic to a generated submodel of some ultrapower for the disjoint union of M. When this has been done, the claim of the lemma follows as an easy corollary. $\dashv$

Before we get to the main result of this section, a short remark is in order. In Theorem 4.3 we required that a class K of models be closed under ultraproducts for it to be modally definable. A careful examination of the proof of this theorem shows that we can do with a slightly weaker condition. It suffices to call for closure under ultraproducts modulo ultrafilters of a special sort, namely countably incomplete ultrafilters. The reason why this suffices is rather obvious: countably incomplete ultrafilters provide ω-saturated ultraproducts, and the latter are the kind of ultraproducts we are in need of.

Theorem 5.6 *For a class* K *of models the following conditions are equivalent:*

1. K *is closed under surjective bisimulations, disjoint unions and ultraproducts modulo countably incomplete ultrafilters, and* $\overline{\mathsf{K}}$ *is closed under ultrapowers modulo countably incomplete ultrafilters.*

2. K *is closed under generated submodels, isomorphisms and disjoint unions, and for each model* $\mathfrak{M}$*,* $\mathfrak{M} \in \mathsf{K}$ *iff* $\mathcal{C}(\mathfrak{M}) \in \mathsf{K}$*.*

Proof. For the *only if* direction assume that K satisfies the closure conditions stated in 1. Clearly, K is closed under generated submodels and isomorphisms. Next, suppose $\mathfrak{M} \in K$. Choose a countably incomplete ultrafilter $\mathcal{U}$ over ω and let $\mathfrak{M}' := (W', R', V')$ be the ultrapower of $\mathfrak{M}$ modulo $\mathcal{U}$. By the closure conditions on K, $\mathfrak{M}' \in \mathsf{K}$. Further, from general first-order model theory we know that $\mathfrak{M}'$ is ω-saturated. Hence, by Lemma 5.3, there is a total surjective bisimulation Z from $\mathfrak{M}'$ to $\mathcal{C}(\mathfrak{M})$. As K is closed under surjective bisimulations, we obtain $\mathcal{C}(\mathfrak{M}) \in \mathsf{K}$. For the other direction suppose $\mathcal{C}(\mathfrak{M}) \in \mathsf{K}$. Let $\mathfrak{M}'$ and Z be as above. It is easy to see that Z^{-1} is a surjective bisimulation from $\mathcal{C}(\mathfrak{M})$ to $\mathfrak{M}'$. Thus, $\mathfrak{M}' \in \mathsf{K}$. Since $\overline{\mathsf{K}}$ is closed under ultrapowers, we obtain $\mathfrak{M} \in \mathsf{K}$. This concludes the proof of the *only if* direction.

For the converse, assume that K satisfies the closure conditions stated in 2. To show that K is closed under surjective bisimulations, let $\mathfrak{M}$ and $\mathfrak{M}'$ be models such that $\mathfrak{M} \in \mathsf{K}$, and suppose that Z is a surjective bisimulation from $\mathfrak{M}$ to $\mathfrak{M}'$. Let $\mathfrak{N} := (W'', R'', V'')$ be the submodel of $\mathfrak{M}$ generated by the domain of Z. As K is closed under generated submodels, $\mathfrak{N} \in \mathsf{K}$. Moreover, it is obvious that $Z \cap (W'' \times W')$ is a bisimulation from $\mathfrak{N}$ to $\mathfrak{M}'$ which is total and surjective. Hence, by Lemma 5.4, $\mathcal{C}(\mathfrak{N}) \cong \mathcal{C}(\mathfrak{M}')$. By the closure under $\mathcal{C}$, $\mathcal{C}(\mathfrak{N}) \in \mathsf{K}$. As K is closed under isomorphisms, $\mathcal{C}(\mathfrak{M}') \in \mathsf{K}$, which yields $\mathfrak{M}' \in \mathsf{K}$. This shows that K is closed under surjective bisimulations.

Let $\mathfrak{M} \in \mathsf{K}$, and let $\mathfrak{N}$ be an ultrapower modulo a countably incomplete ultrafilter of $\mathfrak{M}$. By the closure conditions on K, $\mathcal{C}(\mathfrak{M}) \in \mathsf{K}$. From Lemma 5.3 we infer that $\mathfrak{N} \in \mathsf{K}$ if and only if $\mathcal{C}(\mathfrak{M}) \in \mathsf{K}$. Hence $\mathfrak{N} \in \mathsf{K}$. Thus K is closed under the right kind of ultrapowers. By a similar argument we can show that $\overline{\mathsf{K}}$ is closed under suitable ultrapowers. An application of Lemma 5.5 concludes the proof. $\dashv$

As a corollary to Theorem 5.6 we find that Hansoul's characterization of the globally definable classes of models (Theorem 5.2) is equivalent with ours (Theorem 4.3).

6 Universal Classes

Applying arguments and tools similar to the ones in the previous section, we can also obtain global definability and preservation results for modal formulas satisfying various syntactic constraints. Below, we restrict our attention to universal formulas. By a *universal* formula we mean a modal formula that has been built up from atomic formulas and negated atomic formulas, using $\wedge$, $\vee$ and $\Box$ only. We use Π to denote the set of universal formulas. Analogously, the set Σ of existential formulas is defined as the smallest subset of $\mathcal{ML}$ that contains all atomic formulas and their negations, and is closed under $\wedge$, $\vee$ and $\Diamond$.

When we drop clause B3 in Definition 3.1, we arrive at the notion of simulation. We write $(\mathfrak{M}, w) \rightsquigarrow (\mathfrak{N}, v)$ to denote that there is a simulation Z from $\mathfrak{M}$ to $\mathfrak{N}$ with Zwv. It is easy to prove that universal formulas are anti-preserved under simulations, that is, if $(\mathfrak{M}, w) \rightsquigarrow (\mathfrak{N}, v)$ then for every universal formula φ, if $(\mathfrak{N}, v) \Vdash \varphi$ then $(\mathfrak{M}, w) \Vdash \varphi$.

The main result of this section, Theorem 6.3, provides a precise characterization of the conditions under which classes of models are globally definable by sets of *universal* formulas. From this we easily get a preservation result for universal formulas, stated as Corollary 6.4. In the proof of Lemma 6.2, which forms the key to the proof of Theorem 6.3, we make use of the following lemma. (For a proof the reader is referred to [21]; this work also contains local counterparts of the results of this section.)

Lemma 6.1 *Let* $(\mathfrak{M}, w)$, $(\mathfrak{N}, v)$ *be unraveled pointed models with* $(\mathfrak{M}, w) \rightsquigarrow (\mathfrak{N}, v)$. *Then there is an unraveled model* $(\mathfrak{N}', v)$ *such that* $(\mathfrak{M}, w) \subseteq (\mathfrak{N}', v)$ *and a surjective bisimulation from* $(\mathfrak{N}, v)$ *to* $(\mathfrak{N}', v)$.

Lemma 6.2 *For a set* Φ *of modal formulas the following are equivalent:*

1. *There is a set* Ψ *of universal formulas such that* $\mathrm{Mod}(\Phi) = \mathrm{Mod}(\Psi)$.

2. $\mathrm{Mod}(\Phi)$ *is closed under submodels.*

Proof. For the implication from 1 to 2 we need to show that universal formulas are preserved under submodels, which can be proved by an easy induction. For the other direction, suppose Φ is a set of modal formulas such that $\mathrm{Mod}(\Phi)$ is closed under submodels. Define $\Psi := \{\psi \in \Pi \mid \Phi \models_g \psi\}$. We must show that $\Psi \models_g \Phi$. So suppose $\mathfrak{M} := (W, R, V)$ is a model with $\mathfrak{M} \models \Psi$. It suffices to prove that Φ holds in each point-generated submodel of $\mathfrak{M}$.

To see why this is enough, let $\mathfrak{M}'$ be the disjoint union of the set of all point-generated submodels of $\mathfrak{M}$. Now, suppose we have shown that Φ holds in each point-generated submodel of $\mathfrak{M}$. Since $\mathrm{Mod}(\Phi)$ is closed under disjoint unions, this yields $\mathfrak{M}' \models \Phi$. Moreover, it is easy to see that there exists a surjective bisimulation from $\mathfrak{M}'$ to $\mathfrak{M}$. By closure under surjective bisimulations we obtain the desired result.

So, let $w \in W$, and let $\mathfrak{M}_w$ be the submodel of $\mathfrak{M}$ which is generated by w. Clearly, $\mathfrak{M}_w \models \Psi$. Our next aim is to show that there is a model $\mathfrak{N}_w$, generated by a point v, such that $\mathfrak{N}_w \models \Phi$ and each existential formula true in $(\mathfrak{M}_w, w)$ also holds in $(\mathfrak{N}_w, v)$. To do this, it is enough to show that the following two sets Γ_1 and Γ_2 are simultaneously satisfiable:

$$\begin{aligned} \Gamma_1 &:= \{\forall x(ST(\varphi)[x/c]) \mid \varphi \in \Phi\} \\ \Gamma_2 &:= \{ST(\psi) \mid \psi \in \Sigma, (\mathfrak{M}_w, w) \Vdash \psi\}. \end{aligned}$$

Suppose not. Then there are $\psi_1, \ldots, \psi_n \in \Sigma$ with $(\mathfrak{M}_w, w) \Vdash \psi_i$, for $i \leq n$, such that $\Gamma_1 \models_l \neg(ST(\psi_1) \wedge \cdots \wedge ST(\psi_n))$. Hence $\Gamma_1 \models_l \neg ST(\psi_1 \wedge \cdots \wedge \psi_n)$. Let $\psi := (\psi_1 \wedge \cdots \wedge \psi_n)$. Since Σ is closed under conjunctions, $\psi \in \Sigma$. Hence $ST(\psi) \in \Gamma_2$. Now, note that c does not occur in Γ_1. So we obtain $\Gamma_1 \models_g \forall x(\neg ST(\psi)[x/c])$, hence $\Gamma_1 \models_g \forall x(ST(\neg\psi)[x/c])$. The latter implies $\Phi \models_g \psi^*$, where ψ^* is the negation normal form of ψ. Obviously, $\psi^* \in \Pi$, hence $\psi^* \in \Psi$, which contradicts $\mathfrak{M}_w \models \Psi$. From this we can conclude that $\Gamma_1 \cup \Gamma_2$ is satisfiable.

So there is a pointed model $(\mathfrak{N}_w, v)$ with $\mathfrak{N}_w \models \Phi$ and such that for each existential formula ψ, if $(\mathfrak{M}_w, w) \Vdash \psi$ then $(\mathfrak{N}_w, v) \Vdash \psi$. Choose ω-saturated elementary extensions $(\mathfrak{M}'_w, w)$, $(\mathfrak{N}'_w, v)$ of $(\mathfrak{M}_w, w)$, $(\mathfrak{N}_w, v)$ respectively. By a standard argument we obtain $(\mathfrak{M}'_w, w) \rightsquigarrow (\mathfrak{N}'_w, v)$. Further, let $(\mathfrak{M}^u_w, w)$, $(\mathfrak{N}^u_w, v)$ be the unraveling of $(\mathfrak{M}'_w, w)$, $(\mathfrak{N}'_w, v)$ respectively. It is easy to see that $(\mathfrak{M}^u_w, w) \rightsquigarrow (\mathfrak{N}^u_w, v)$ holds. By an application of Lemma 6.1 we obtain a pointed model $(\mathfrak{N}, v_w)$ with $(\mathfrak{M}^u_w, w) \subseteq (\mathfrak{N}, v)$ and a surjective bisimulation from $(\mathfrak{N}^u_w, v)$ to $(\mathfrak{N}, v)$. See Figure 1.

$$\begin{tikzcd}
(\mathfrak{M}_w, w) \arrow[r, dash, "\preceq"] \arrow[d, dash] & (\mathfrak{M}'_w, w) \arrow[r, dash] \arrow[d, dash, "\rightsquigarrow"'] & (\mathfrak{M}^u_w, w) \arrow[r, dash, "\subseteq"] \arrow[d, dash, "\rightsquigarrow"'] & (\mathfrak{N}, v_w) \\
(\mathfrak{N}_w, v) \arrow[r, dash] & (\mathfrak{N}'_w, v) \arrow[r, dash] & (\mathfrak{N}^u_w, v) \arrow[ur, dash, "\sim"'] &
\end{tikzcd}$$

Figure 1: A chain of models

To conclude the proof, we argue as follows. By construction it holds that $\mathfrak{N}_w \models \Phi$, hence, by elementary extension, $\mathfrak{N}'_w \models \Phi$. Consider the submodel $\mathfrak{N}''$ of $\mathfrak{N}'$ that is point-generated by v_w. Observe that the converse relation of $f_{(\mathfrak{N}'',v_w)}$ forms a surjective bisimulation from $\mathfrak{N}'$ to $\mathfrak{N}^u$. This, together with Theorem 4.2, implies $\mathfrak{N}^u_w \models \Phi$ and, hence, $\mathfrak{N} \models \Phi$. That Φ is preserved under submodels implies $\mathfrak{M}^u_w \models \Phi$. Again, by making use of the fact that modal formulas are preserved under surjective bisimulations and elementary extensions, we finally obtain $\mathfrak{M}_w \models \Phi$. This concludes the proof. $\dashv$

Theorem 6.3 *A class* K *of models is definable by a set of universal formulas iff* K *is closed under surjective bisimulations, disjoint unions, submodels and ultraproducts.*

Proof. Suppose K is globally definable by a set of universal formulas. Then K satisfies the above closure conditions by Theorem 4.3 and Lemma 6.2. For the other direction we reason as follows. Suppose K is closed under surjective bisimulations, disjoint unions and ultraproducts. Since every model is (isomorphic to) a submodel of its ultrapowers, $\overline{\mathsf{K}}$ is closed under ultrapowers as well. Thus by Theorem 4.3 there is a set Ψ of modal formulas such that $\mathrm{Mod}(\Psi) = \mathsf{K}$. Using the closure of K under submodels, an application of Lemma 6.2 concludes the proof. $\dashv$

Corollary 6.4 *A modal formula* φ *is globally preserved under submodels iff there is a universal formula* ψ *such that* φ *and* ψ *are true on exactly the same models.*

Proof. The claim of the corollary follows from Lemma 6.2 by compactness. Recall that the set of universal formulas is closed under conjunction. $\dashv$

We hasten to add that similar results also exist for positive, that is negation-free formulas: formulas built up from proposition letters using only $\wedge$, $\vee$, $\Diamond$ and $\Box$. In particular, in Theorem 6.5 we give necessary and sufficient conditions for a class of models K to be globally definable by a set of positive formulas.

In the following, a model $\mathfrak{N} = (W', R', V')$ is said to be a *weak extension* of a model $\mathfrak{M} = (W, R, V)$, if $W = W'$, $R = R'$, and $V(p_n) \subseteq V'(p_n)$, for every $n \in \omega$.

Theorem 6.5 *A class* K *of models is definable by a set of positive formulas iff* K *is closed under surjective bisimulations, disjoint unions, weak extensions and ultraproducts, and* $\overline{\mathsf{K}}$ *is closed under ultrapowers.*

Proof. The proof closely resembles the proof of Theorem 6.3; hence, we only give a brief sketch. For a start, one has to find a kind of bisimulations suitable for positive formulas. These are positive bisimulations, where a relation Z is said to be a *positive* bisimulation, if Z satisfies B2 and B3 from Definition 3.1 together with the following weakening of clause B1: for $w \in W$, $v \in W'$, if Zwv then $(\mathfrak{M}, w) \Vdash p_n$ implies $(\mathfrak{N}, v) \Vdash p_n$, for every $n \in \omega$. Then, one has to prove the following statement (see [21]):

> Suppose $(\mathfrak{M}, w)$, $(\mathfrak{N}, v)$ are unraveled pointed models and Z is a positive bisimulation from $(\mathfrak{M}, w)$ to $(\mathfrak{N}, v)$. Then there exist unraveled models $(\mathfrak{M}', w)$, $(\mathfrak{N}', v)$ such that $(\mathfrak{N}', v)$ is a weak extension of $(\mathfrak{M}', w)$, and there are surjective bisimulations from $(\mathfrak{M}, w)$, $(\mathfrak{N}, v)$ to $(\mathfrak{M}', w)$, $(\mathfrak{N}', v)$ respectively.

Next, using this result, one can show that a set of modal formulas Φ is preserved under weak extensions if and only if there is a set of positive formulas Ψ with $\mathrm{Mod}(\Psi) = \mathrm{Mod}(\Phi)$. The characterization for positive classes then follows by an easy argument, similar to the argument in the proof of Theorem 6.3. $\dashv$

From this theorem we easily infer a preservation result for positive formulas:

Corollary 6.6 *A modal formula φ is globally preserved under weak extensions iff there is a positive formula ψ such that φ and ψ are true on exactly the same models.*

7 Concluding Remarks

In this paper we have presented a number of definability and preservation results for (basic) modal logic. Contrary to common practice, we have emphasized the global perspective. Of course, the above results can only be considered as a modest beginning; a lot of work remains to be done. Our future research should concentrate on the following aspects and questions.

First, what other local results (and tools) can be adapted to the global setting? For example, consider existential formulas and try to prove results in the spirit of Section 6. Note that, on the global level, a modal existential formula like $\Diamond p$ does not translate into an existential first-order sentence, but into a sentence with the prefix $\forall\exists$. Another interesting task is to analyze modal Horn formulas along the same line; for local results see [21].

Second, in recent years logicians have introduced bisimulations for a great number of modal and modal-like languages and proved (local) definability results with respect to them. Examples are temporal languages with Since and Until [15], description logics [17] and negation-free languages [16]. Many of the results proved in these papers can be globalized along the lines followed in Section 4. How far does this method go? In particular, does it also apply to modal languages that live outside first-order logic, like PDL or infinitary languages?

Third, to go one step further, it would be worthwhile to look for a uniform way of connecting the local and the global setting. Indeed, we have a general result with respect to preservation, but it is not completely satisfactory. First, it looks slightly proof-generated and, second, it makes use of the notion of ω-saturated models which restricts its applicability to modal languages that lie inside first-order logic. However, the result may serve as a good starting point for further work.

In this context, Kracht and Wolter's work on transfer results [13, 14] deserves attention. They investigate the metalogical properties of modal logics, thereby considering both levels, the local and the global one. Though they take a different perspective on modal logic — their aim is to explore the lattice of normal modal logic — there are interesting connections between their research and ours. This holds, in particular, when we no longer analyze definability with respect to the class of all models only, but consider definability within certain subclasses (induced by special frame properties) as well.

Fourth, we may also change tack, and try to match up the semantic operations that were used in this paper with syntactic characterizations in the following sense. Find syntactic characterizations for the classes of models closed under surjective bisimulations, and establish a preservation result for surjective bisimulations. Tackle analogous questions for disjoint unions as well.

The case of surjective bisimulations is fairly easy. By a standard argument we are able to prove that a class K is $\bigvee$-definable iff K is closed under surjective bisimulations and ultraproducts, and $\overline{\mathsf{K}}$ is closed under ultrapowers. Here, by $\bigvee$-definable we mean definable in the set $\Theta := \{\forall x(ST(\varphi_1)[x/c]) \vee \cdots \vee \forall x(ST(\varphi_n)[x/c]) \mid \varphi_i \in \mathcal{ML}\}$. The corresponding preservation result may again be obtained as an immediate consequence. It tells us that a first-order sentence (over $\tau \setminus \{c\}$) is equivalent to some $\varphi \in \Theta$ if and only if it is preserved under surjective bisimulations.

The case of disjoint unions turns out to be much harder. Van Benthem [3] gave a characterization of the first-order sentences that are invariant under disjoint unions. This result only considers the frame language, that is the first-order vocabulary consisting of the relation symbol S and the identity symbol. It is easy to turn van Benthem's result into a result that characterizes the sentences in the modal language, that is $\tau \setminus \{c\}$, that are preserved under disjoint unions. Moreover, by using techniques developed in the present paper, we are able to give a similar characterization for the language we obtain by dropping the identity symbol. Although this is a nice result, it strikes us as not completely satisfactory. What we are after is a characterization of those first-order sentences that are *preserved* — and not invariant — under disjoint unions. The solution to this problem must be left to future research.

As a final point, we mention the connections that exist between the results in this paper and work on the universal modality [8, 10]. These links are interesting for a number of reasons. First, the global translations of $\mathcal{ML}$-formulas are expressible in the language with the universal modality. In general, the universal modality makes it possible to talk about global properties from the local point of view. Second, in [4] van Benthem formulated a preservation theorem that says that a first-order sentence is equivalent to a modal formula from the universal language if and only if it is preserved under total bisimulations. And third, the algebraic criteria that were given in [10] for a class of generalized frames to be definable in the language with universal modality may be adopted to the level of models. This should lead to a characterization result in the style of Hansoul's theorem. Moreover, the latter result and van Benthem's theorem might be linked up in the spirit of our Theorem 5.6.

Acknowledgments. Maarten de Rijke was supported by the Spinoza project 'Logic in Action' at ILLC, the University of Amsterdam.

References

[1] J. Barwise and L.S. Moss. Modal correspondence for models. *Journal of Philosophical Logic*, 27:275–294, 1998.

[2] J.F.A.K. van Benthem. *Modal Correspondence Theory*. PhD thesis, University of Amsterdam, 1976.

[3] J.F.A.K. van Benthem. *Modal Logic and Classical Logic*. Bibliopolis, Napoli, 1985.

[4] J.F.A.K. van Benthem. The range of modal logic. Technical Report, ILLC, University of Amsterdam, 1997.

[5] P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*, Manuscript, ILLC, University of Amsterdam, 1998. Available at `http://www.illc.-uva.nl/~mdr/Publications/modal-logic.html`.

[6] F.M. Donini, M. Lenzerini, D. Nardi, and A. Schaerf. Reasoning in description logics. In G. Brewka, editor, *Principles of Knowledge Representation*, Studies in Logic, Language and Information, pages 191–236. CSLI Publications, 1996.

[7] H.-D. Ebbinghaus and J. Flum. *Finite Model Theory*. Springer, Berlin 1995.

[8] G. Gargov and V. Goranko. Modal logic with names. *Journal of Philosophical Logic*, 22:607–636, 1993.

[9] R. Goldblatt. Metamathematics of modal logic I. *Reports on Mathematical Logic*, 6:41–78, 1976.

[10] V. Goranko and S. Passy. Using the universal modality. *Journal of Logic and Computation*, 1:5–31, 1992.

[11] G. Hansoul. Modal-axiomatic closure of Kripke models. In: H. Andréka, D. Monk, and I. Németi, editors, *Algebraic Logic*. North-Holland, 257–264, 1991.

[12] J. Kim. *Supervenience and Mind*. Cambridge University Press, Cambridge 1993.

[13] M. Kracht. *Tools and Techniques in Modal Logic*. Habilitationsschrift, Berlin 1996.

[14] M. Kracht and F. Wolter. Simulation and transfer results in modal logic – a survey. *Studia Logica*, 59:149–177, 1997.

[15] N. Kurtonina and M. de Rijke. Bisimulations for temporal logic. *Journal of Logic, Language and Information*, 6:403–425, 1997.

[16] N. Kurtonina and M. de Rijke. Simulating without negation. *Journal of Logic and Computation*, 7:501–522, 1997.

[17] N. Kurtonina and M. de Rijke. Expressiveness of concept expressions in first-order description logics. *Artificial Intelligence*, 107:303-333, 1999.

[18] M. de Rijke. *Extending Modal Logic*. PhD thesis, ILLC, University of Amsterdam, 1993.

[19] E. Rosen. Modal logic over finite structures. *Journal of Logic, Language and Information*, 6:427–439, 1997.

[20] H. Sahlqvist. Completeness and correspondence in the first and second order semantics for modal logic. In: S. Kanger, editor, *Proc. 3rd. Scand. Logic Symp.* North-Holland, 110–143, 1975.

[21] H. Sturm. *Modale Fragmente von* $\mathcal{L}_{\omega\omega}$ *und* $\mathcal{L}_{\omega_1\omega}$. PhD thesis, CIS, University of Munich, 1997.

[22] Y. Venema. *Many-Dimensional Modal Logic*. PhD thesis, University of Amsterdam, 1991.