On Characterization, Definability and ω-Saturated Models

Facundo Carreiro*

Dto. de Computación, FCEyN, Universidad de Buenos Aires
fcarreiro@dc.uba.ar

Abstract. Two important classic results about modal expressivity are the Characterization and Definability theorems. We develop a general theory for modal logics below first order (in terms of expressivity) which exposes the following result: Characterization and Definability theorems hold for every (reasonable) modal logic whose ω-saturated models have the Hennessy-Milner property. The results are presented in a general version which is relativized to classes of models.

1 Introduction

Syntactically, *modal languages* [1] are propositional languages extended with *modal operators*. Indeed, the basic modal language is defined as the extension of the propositional language with the unary operator $\Box$. Although these languages have a very simple syntax, they are extremely useful to describe and reason about *relational structures*.

A relational structure is a nonempty set together with a family of relations. Given the generality of this definition it is not surprising that modal logics are used in a wide range of disciplines: mathematics, philosophy, computer science, computational linguistics, etc. For example, theoretical computer science uses labeled transition systems (which are nothing but relational structures) to model the execution of a program.

An important observation that might have gone unnoticed in the above paragraph is that there are *many different modal logics*. There is, nowadays, a wide variety of modal languages and an extensive menu of modal operators to choose from (e.g. *Since* and *Until* [2], universal modality [3], difference modality [4], fix-point operators [5], etc.), enabling the design of a particular logic for each specific application.

It is in this context that the notion of *bisimulation* [6] becomes fundamental. Intuitively, a bisimulation characterizes, from a structural point of view, when a state in a model is indistinguishable from a state in another model. Bisimulations are a crucial tool in the process of studying relational structures and they open the way to formally analyze the exact expressive power of modal languages.

If we comprehend in detail the notion of bisimulation, we can measure and try to balance expressiveness and complexity, and thus obtain a logic appropriate

* F. Carreiro was partially supported by a grant by CONICET Argentina.

A. Cerone and P. Pihlajasaari (Eds.): ICTAC 2011, LNCS 6916, pp. 62–76, 2011.

to the context of use with the minimum possible computational complexity. However, deciding which is the correct notion of bisimulation for a given logic is not an easy task.

In this paper we investigate Characterization and Definability, two model-theoretical results intimately related with the notion of bisimulation. We pursue a general study of these properties without referring to a particular modal logic. In general, the validity of these theorems is a good indicator that the underlying notion of bisimulation for a given logic is indeed the correct one.

Characterization results for modal logics identify them as fragments of a better known logic. This type of characterizations allows for the transfer of results and a better understanding of the modal logic. The first work in this direction was done by van Benthem [7] who used bisimulations to characterize the basic modal logic (BML) as the bisimulation invariant fragment of first order logic.

Theorem (J. van Benthem). *A first order formula α is equivalent to the translation of a* BML *formula iff α is invariant under bisimulations.*

For example, as a corollary of this theorem, we know that the first order formula $\varphi(x) = R(x, x)$ is not expressable in BML because we can construct a model with a reflexive element which is bisimilar to a non-reflexive element. Observe that in this case the notion of bisimulation is that of BML. As we have said before, every modal logic has a, potentially different, notion of bisimulation.

A theorem of this kind would identify which first order properties can be captured with each particular modal logic. But —as the syntax, semantics of the logic and the definition of bisimulation involved have changed— each such theorem needs a new ad-hoc proof.

Let us now discuss definability. Given a logic $\mathfrak{L}$, definability results identify the properties that a class of $\mathfrak{L}$-models K should satisfy in order to guarantee the existence of an $\mathfrak{L}$-formula (or a set of $\mathfrak{L}$-formulas) safisfied exactly by the models in K.[1] This question has already been addressed for first order logic [8], BML [1] and many others logics [9,10]. Whereas the answer for first order logic is presented in terms of *potential isomorphisms* [8], in the case of modal logics, the notion of bisimulation plays a fundamental role.

Theorem (M. de Rijke). *A class of Kripke models* K *is definable by means of a single* BML *formula iff both* K *and its complement are closed under bisimulations and ultraproducts.*

As a corollary of this theorem we get, for example, that the class of finite models is not definable in BML (because it is not closed under bisimulations). As with the characterization theorem, definability results similar to the one presented here also hold for a vast number of modal logics. However, every logic requires a different proof.

In this article we undertake a study of the proof techniques used for Characterization and Definability results. Our objective is to find sufficient conditions that an arbitrary logic has to fulfill to validate these theorems. Such conditions are captured in the notion of an *adequate pair* (introduced in Section 2).

[1] We consider definability of classes of models; we will not discuss frame definability.

If a given modal logic $\mathfrak{L}$ —together with its syntax, semantics and its notion of bisimulation— is compatible with our general theory, we guarantee the validity of Characterization and Definability theorems for $\mathfrak{L}$. We prove a more general version of these theorems were we allow a relativization to different classes of models.

The article is organized as follows. In Section 2 we give the set of basic definitions needed to state our main results. In Sections 3 and 4 we prove a generalization of the Characterization and Definability theorems and finally in Section 5 we draw some conclusions and propose further lines of research. Most of the proofs are presented in the main body of the article while the most technical ones can be found in the Appendix.

2 Basic Definitions

The theorems discussed in this article deal with two logics: one less expressive than the other. Different modal logics play the part of the former while we will always take first order logic (with or without equality) as the latter.

Definition 1 (Languages). *We denote the modal language as $\mathfrak{L}$ and the first order language as $\mathfrak{F}$. We consider languages $\mathfrak{L}$ extending $\mathfrak{P} = \langle (p_i)_{i\in\mathbb{N}}, \wedge, \vee, \top, \bot \rangle$ with the usual interpretations. $\mathfrak{F}$ is a countable first order language which may contain equality. For any language $\mathfrak{A}$, we call* $\mathrm{FORM}(\mathfrak{A})$ *to the set of formulas of the language $\mathfrak{A}$ and* $\mathrm{FORM}_1(\mathfrak{F})$ *to the subset of* $\mathrm{FORM}(\mathfrak{F})$ *with at most one free variable x.*

We do not impose any restrictions on the structures over which the language $\mathfrak{L}$ is interpreted. We only assume that every $\mathfrak{L}$-model $\mathcal{M}$ has a set of elements (also called worlds) which is called the domain or universe of $\mathcal{M}$ (notated $|\mathcal{M}|$).

Definition 2 (Models). *We define* $\mathrm{MODS}(\mathfrak{L})$ *to be the class of $\mathfrak{L}$-models under study (not necessarily the class of all models of the signature of $\mathfrak{L}$), and* $\mathrm{MODS}(\mathfrak{F})$ *to be the class of all $\mathfrak{F}$-models. We use $\overline{\mathsf{K}}$ to denote the complement of the class K.*

An $\mathfrak{L}$-pointed model *is pair $\langle \mathcal{M}, w \rangle$ where $\mathcal{M}$ is an $\mathfrak{L}$-model and $w \in |\mathcal{M}|$. We define the class of $\mathfrak{L}$-pointed models corresponding to* $\mathrm{MODS}(\mathfrak{L})$ *as*

$$\mathrm{PMODS}(\mathfrak{L}) = \{\langle \mathcal{M}, w \rangle : \mathcal{M} \in \mathrm{MODS}(\mathfrak{L}) \textit{ and } w \in |\mathcal{M}|\}.$$

An $\mathfrak{F}$-pointed model *is a pair $\langle \mathcal{M}, g \rangle$ where $\mathcal{M}$ is an $\mathfrak{F}$-model and $g : \{x\} \to |\mathcal{M}|$. The class of $\mathfrak{F}$-pointed models corresponding to* $\mathrm{MODS}(\mathfrak{F})$ *is defined as*

$$\mathrm{PMODS}(\mathfrak{F}) = \{\langle \mathcal{M}, g \rangle : \mathcal{M} \in \mathrm{MODS}(\mathfrak{F}) \textit{ and } g : \{x\} \to |\mathcal{M}|\}.$$

Definition 3. *We use $\mathcal{M}, w \models \varphi$ to denote that an $\mathfrak{L}$-formula φ is true in the point w of the $\mathfrak{L}$-model $\mathcal{M}$. Similarly, we use $\mathcal{M}, g \models \varphi$ to denote that a $\mathfrak{F}$-formula φ is true in the $\mathfrak{F}$-model $\mathcal{M}$ under the assignment g.*

If Γ is a set of first order formulas, we write $\Gamma \models_{\mathsf{K}} \beta$ to mean that the entailment is valid within *the class K. We say that the first order formulas α and β are K-equivalent when $\models_{\mathsf{K}} \alpha \leftrightarrow \beta$.*

Let $\langle \mathcal{M}, w\rangle, \langle \mathcal{N}, v\rangle \in \text{PMODS}(\mathfrak{L})$. *We write* $\mathcal{M}, w \rightsquigarrow_{\mathfrak{L}} \mathcal{N}, v$ *to mean that for every* $\mathfrak{L}$*-formula* φ*, if* $\mathcal{M}, w \models \varphi$ *then* $\mathcal{N}, v \models \varphi$*. We write* $\mathcal{M}, w \leftrightsquigarrow_{\mathfrak{L}} \mathcal{N}, v$ *when* $\mathcal{M}, w \rightsquigarrow_{\mathfrak{L}} \mathcal{N}, v$ *and* $\mathcal{N}, v \rightsquigarrow_{\mathfrak{L}} \mathcal{M}, w$*. This notation extends analogously to* $\text{PMODS}(\mathfrak{F})$*. We drop the subscript when the logic involved is clear from context.*

During the article, we will use and adapt some classical first order notions (such as potential isomorphism, ω-saturation and definability) to the context of our framework; see [8] for reference on these classical definitions. We start by adapting the notion of closure under potential isomorphisms to make it relative to a class of models.

Definition 4 (C-closure under potential isomorphisms). *Let* $\langle \mathcal{M}, g\rangle$ *and* $\langle \mathcal{N}, h\rangle \in \text{PMODS}(\mathfrak{F})$ *we write* $\mathcal{M}, g \cong \mathcal{N}, h$ *to mean that there exists a potential isomorphism* I *between* $\mathcal{M}$ *and* $\mathcal{N}$ *such that* $\langle g(x)\rangle I \langle h(x)\rangle$*.*

A class $\mathsf{K} \subseteq \mathsf{C} \subseteq \text{PMODS}(\mathfrak{F})$ *is* C-closed under potential isomorphisms *if for every* $\langle \mathcal{M}, g\rangle \in \mathsf{K}$ *and* $\langle \mathcal{N}, h\rangle \in \mathsf{C}$ *such that* $\mathcal{M}, g \cong \mathcal{N}, h$ *then* $\langle \mathcal{N}, h\rangle \in \mathsf{K}$*.*

We can see a model as an information repository. We need to define a way to access this information from the perspective of both the $\mathfrak{L}$ and the $\mathfrak{F}$ language.

Definition 5 (Truth preserving translations). *A* formula translation *is a function* $\mathsf{Tf}_x : \text{FORM}(\mathfrak{L}) \to \text{FORM}_1(\mathfrak{F})$ *such that* $\mathsf{Tf}_x(\varphi \wedge \psi) = \mathsf{Tf}_x(\varphi) \wedge \mathsf{Tf}_x(\psi)$ *and* $\mathsf{Tf}_x(\varphi \vee \psi) = \mathsf{Tf}_x(\varphi) \vee \mathsf{Tf}_x(\psi)$*. Given a class of models* $\mathsf{K} \subseteq \text{PMODS}(\mathfrak{F})$*, a* model translation *is a bijective function* $\mathsf{T} : \text{PMODS}(\mathfrak{L}) \to \mathsf{K}$*.*

A pair of translations $(\mathsf{Tf}_x, \mathsf{T})$ *is* truth-preserving *if for all* $\varphi \in \text{FORM}(\mathfrak{L})$ *and all* $\langle \mathcal{M}, w\rangle \in \text{PMODS}(\mathfrak{L})$ *they satisfy* $\mathcal{M}, w \models \varphi$ iff $\mathsf{T}(\mathcal{M}, w) \models \mathsf{Tf}_x(\varphi)$*. As an abuse of notation we use* $\mathsf{T}(\mathcal{M})$ *when we are not interested in the associated point of evaluation.*

For the rest of the article, we fix $(\mathsf{Tf}_x, \mathsf{T})$ to be an arbitrary pair of truth-preserving translations. We will also need to translate formulas from $\mathfrak{L}$ to $\mathfrak{F}$ and then go back to $\mathfrak{L}$-formulas. *A priori*, as we are not requiring Tf_x to be injective this could lead to a problem but it can be easily solved.

Notice that, for any α, β such that $\mathsf{Tf}_x(\alpha) = \mathsf{Tf}_x(\beta)$ we have $\alpha \models \beta$ and $\beta \models \alpha$. Hence, without loss of generality, we can work with equivalence classes of $\mathfrak{L}$-formulas (modulo $\mathfrak{L}$-equivalence) and assume that the formula translation Tf_x is injective. In the following definition we recall the classical notion of definability adapted to the context of $\mathfrak{L}$-models.

Definition 6. *A class* $\mathsf{M} \subseteq \text{PMODS}(\mathfrak{L})$ *is* defined by a set of $\mathfrak{L}$-formulas Γ *(resp. a formula* φ*) when* $\langle \mathcal{M}, w\rangle \in \mathsf{M}$ *iff* $\mathcal{M}, w \models \Gamma$ *(resp.* $\mathcal{M}, w \models \varphi$*).*

In Section 1 we talked about bisimulations. This notion is usually used in the context of logics with negation where the relation is symmetrical. In the case of negation-free logics the analogous notion is not symmetrical and it is called *simulation*. For the framework developed in this article we take a broader approach and define an abstract notion of $\mathfrak{L}$-*simulation* which generalizes it.

Definition 7 ($\mathfrak{L}$-simulation). *An* $\mathfrak{L}$-simulation *is a* non-empty *relation* $Z \subseteq \text{PMODS}(\mathfrak{L}) \times \text{PMODS}(\mathfrak{L})$ *such that if* $\langle \mathcal{M}, w\rangle Z \langle \mathcal{N}, v\rangle$ *then* $\mathcal{M}, w \rightsquigarrow \mathcal{N}, v$*.*

We write $\mathcal{M}, w \rightrightarrows_{\mathfrak{L}} \mathcal{N}, v$ *to indicate that there exists an* $\mathfrak{L}$*-simulation* Z *such that* $\langle \mathcal{M}, w \rangle Z \langle \mathcal{N}, v \rangle$*. We drop the subscript when the logic is clear from context.*

It is worth mentioning that we do not assume any particular 'structural' property for the notion of $\mathfrak{L}$-simulation (e.g., the *zig* or *zag* condition for the basic modal logic). We only require it to be defined as a non-empty binary relation between $\mathfrak{L}$-pointed models and to preserve the truth of formulas (in one direction). This is something that any reasonable observational equivalence notion should satisfy.

It is known that, in general, $\mathcal{M}, w \rightsquigarrow \mathcal{N}, v$ does not imply $\mathcal{M}, w \rightrightarrows \mathcal{N}, v$. The classes of models where this property holds are called *Hennessy-Milner classes* [1]. An important class which is closely related to this property is the class of ω-saturated models.

Definition 8 (ω-saturation). *A first order model* $\mathcal{M}$ *is* ω-saturated *if for every finite* $A \subseteq |\mathcal{M}|$ *the expansion* $\mathcal{M}_A$ *with new constants* $\overline{a}$ *for every* $a \in A$ *satisfies: Let* $\Gamma(x)$ *be a set of formulas such that every finite subset is satisfiable in* $\mathcal{M}_A$ *then* $\Gamma(x)$ *should also be satisfied in* $\mathcal{M}_A$*.*

Informally, it resembles a kind of 'intra-model' compactness. That is, given a set $\Gamma(x)$ if every finite subset is satisfied in (possibly different) elements in $\mathcal{M}$ then there is *a single element* in $\mathcal{M}$ which satisfies the whole set.

The notion of ω-saturation [8] is defined for first order models, but it also applies to $\mathfrak{L}$-models using the model translation introduced in Definition 5: we define an $\mathfrak{L}$-model $\mathcal{M}$ to be ω-saturated if $\mathsf{T}(\mathcal{M})$ is so.

Not all models are ω-saturated but it is a known result that every first order model has an ω-saturated extension with the same first order theory (i.e., an elementarily equivalent extension). The ω-saturated extension $\mathcal{M}^+$ can be constructed by taking the ultrapower of $\mathcal{M}$ with a special kind of ultrafilter.[2] Observe that, in our setting, this extension will also have the same $\mathfrak{L}$-theory (via Tf_x and T as in Definition 5).

The following definition is central in this article. It makes explicit the requirements for our main theorems of Sections 3 and 4 to hold. From now on, we fix an arbitrary modal logic $\mathfrak{L}$ and a class of models K with the following properties.

Definition 9 (Adequate pair). *A logic* $\mathfrak{L}$ *and a class* $\mathsf{K} \subseteq \mathrm{PMODS}(\mathfrak{F})$ *are said to be an* adequate pair *if they fulfill the following requirements:*

1. K *is closed under ultraproducts.*
2. *There exist truth-preserving translations* Tf_x*,* T *(c.f. Definition 5).*
3. *There exists an* $\mathfrak{L}$*-simulation notion (c.f. Definition 7).*
4. *The class of* ω*-saturated* $\mathfrak{L}$*-models has the Hennessy-Milner property with respect to* $\mathfrak{L}$*-simulations (c.f. Definition 8).*

As we will show in the rest of the paper, we will be able to establish Characterization and Definability results for any arbitrary adequate pair. The crucial condition in the definition is item 4, which will usually be the hardest property to establish when defining an adequate pair. The other conditions should be naturally satisfied by any modal logic whose expressivity is below first order.

---

[2] We assume that the reader is familiar with the definition of *ultraproducts*, *ultrapowers* and *ultrafilters*; for details, see [11].

3 Characterization

One of the central notions of the Characterization theorem stated in Section 1 was that of bisimulation invariance. In the following definition we restate this notion in terms of $\mathfrak{L}$ and K.

Definition 10 ($\mathfrak{L}$-simulation K-invariance). *Let $\langle \mathfrak{L}, \mathsf{K} \rangle$ be an adequate pair. A formula $\alpha(x) \in \mathrm{FORM}_1(\mathfrak{F})$ is* K-invariant for $\mathfrak{L}$-simulations *if for all $\mathfrak{L}$-pointed models $\mathcal{M}, w$ and $\mathcal{N}, v$, such that $\mathcal{M}, w \rightrightarrows \mathcal{N}, v$, if $\mathsf{T}(\mathcal{M}, w) \models \alpha(x)$ then $\mathsf{T}(\mathcal{N}, v) \models \alpha(x)$.*

This property is defined for first order formulas but $\mathfrak{L}$-simulations are defined between $\mathfrak{L}$-models. When trying to generalize the notion of invariance, at least two options come up naturally.

The first one is to call a first order formula $\alpha(x)$ 'invariant for $\mathfrak{L}$-simulations' if, for every two $\mathfrak{L}$-pointed models $\mathcal{M}, w$ and $\mathcal{N}, v$ such that $\mathcal{M}, w \rightrightarrows_{\mathfrak{L}} \mathcal{N}, v$ whenever $\alpha(x)$ is true in $\mathsf{T}(\mathcal{M}, w)$ then it is also true in $\mathsf{T}(\mathcal{N}, v)$. In this case we would be checking simulation in $\mathfrak{L}$ and satisfaction in $\mathfrak{F}$.

The second option is to define a simulation relation $\rightrightarrows_{\mathfrak{F}}$ for $\mathfrak{F}$-models (e.g. as done in [12]). In this case we could just say that a first order formula $\alpha(x)$ is 'invariant for $\mathfrak{L}$-simulations' if for every two $\mathfrak{F}$-models $\mathcal{M}, g$ and $\mathcal{N}, h$ such that $\mathcal{M}, g \rightrightarrows_{\mathfrak{F}} \mathcal{N}, h$ whenever $\alpha(x)$ is true in $\mathcal{M}, g$, then it is also true $\mathcal{N}, h$.

In this article we will use the first option because it is simpler and requires no new definitions.

Theorem 11 (Characterization). *Let $\langle \mathfrak{L}, \mathsf{K} \rangle$ be an adequate pair. A formula $\alpha(x) \in \mathrm{FORM}_1(\mathfrak{F})$ is K-equivalent to the translation of an $\mathfrak{L}$-formula iff $\alpha(x)$ is K-invariant for $\mathfrak{L}$-simulations.*

The proof of this theorem will be the guide in the next few pages but the reader should be aware that, in order to do it, we will prove some propositions and lemmas which will allow us to conclude the desired result.

Proof. The claim from left to right is a consequence of the invariance of $\mathfrak{L}$-formulas over $\mathfrak{L}$-simulations. The implication from the right to left, suppose that $\alpha(x)$ is K-invariant for $\mathfrak{L}$-simulations, we want to see that it is K-equivalent to the translation of an $\mathfrak{L}$-formula. Consider the set of $\mathfrak{L}$-consequences of α:

$$\mathrm{SLC}(\alpha) = \{\mathsf{Tf}_x(\varphi) : \varphi \text{ is an } \mathfrak{L}\text{-formula and } \alpha(x) \models_{\mathsf{K}} \mathsf{Tf}_x(\varphi)\}.$$

The following Proposition shows that it suffices to prove that $\mathrm{SLC}(\alpha) \models_{\mathsf{K}} \alpha(x)$.

Proposition 12. *If $\mathrm{SLC}(\alpha) \models_{\mathsf{K}} \alpha(x)$ then $\alpha(x)$ is K-equivalent to the translation of an $\mathfrak{L}$-formula.*

Proof. Suppose $\mathrm{SLC}(\alpha) \models_{\mathsf{K}} \alpha(x)$, by relative compactness (see Theorem 20 in the Appendix) there is a finite set $\Delta \subseteq \mathrm{SLC}(\alpha)$ such that $\Delta \models_{\mathsf{K}} \alpha(x)$, therefore $\models_{\mathsf{K}} \bigwedge \Delta \to \alpha(x)$. Trivially (by definition) we have that $\models_{\mathsf{K}} \alpha(x) \to \bigwedge \Delta$ so we conclude $\models_{\mathsf{K}} \alpha(x) \leftrightarrow \bigwedge \Delta$. As every $\beta \in \Delta$ is the translation of an $\mathfrak{L}$-formula and the translation preserves conjunctions then $\bigwedge \Delta$ is also the translation of some modal formula. □

Hence, the proof reduces to show that $\text{SLC}(\alpha) \models_{\mathsf{K}} \alpha(x)$. Let us suppose that $\mathsf{T}(\mathcal{M}, w) \models \text{SLC}(\alpha)$. We show that $\mathsf{T}(\mathcal{M}, w) \models \alpha(x)$. Define $\mathsf{NTh}^w(x)$ as

$$\mathsf{NTh}^w(x) = \{\neg \mathsf{Tf}_x(\varphi) : \varphi \text{ is an } \mathfrak{L}\text{-formula and } \mathcal{M}, w \not\models \varphi\}$$

Observe that if $\mathfrak{L}$ has negation then $\mathsf{NTh}^w(x)$ will be the translation of w's modal theory and every model of $\mathsf{NTh}^w(x)$ will be modally equivalent to w. If $\mathfrak{L}$ does not have negation we will only preserve formulas that are not true in w. The above definition works in both cases. Let

$$\Sigma(x) = \{\alpha(x)\} \cup \mathsf{NTh}^w(x).$$

Proposition 13. *$\Sigma(x)$ has a model in K.*

Proof. Let us suppose that there is no model in K for $\Sigma(x)$ and use the contrapositive of Theorem 20. Then there is a finite subset $\{\alpha(x), \neg\delta_1, \ldots, \neg\delta_n\} \subseteq \Sigma(x)$ with $\neg\delta_i \in \mathsf{NTh}^w(x)$ which does not have a model in K. Notice that this finite set should include $\alpha(x)$, otherwise it would have a had model, namely $\mathsf{T}(\mathcal{M}, w)$.

Observe that for every model $\mathcal{A}^f \in \mathsf{K}$, as $\mathcal{A}^f \not\models \{\alpha(x), \neg\delta_1, \ldots, \neg\delta_n\}$ then $\mathcal{A}^f \models \alpha(x) \to \neg(\neg\delta_1 \wedge \cdots \wedge \neg\delta_n)$. This means that $\alpha(x) \to (\delta_1 \vee \cdots \vee \delta_n)$ is valid in K, therefore $\alpha(x) \models_{\mathsf{K}} \delta_1 \vee \cdots \vee \delta_n$. If $\delta_1 \vee \cdots \vee \delta_n$ is a K-consequence of $\alpha(x)$ then, as the formula translation preserves disjunction, $\delta_1 \vee \cdots \vee \delta_n \in \text{SLC}(\alpha)$. But, as $\mathsf{T}(\mathcal{M}, w) \models \text{SLC}(\alpha)$ then $\mathsf{T}(\mathcal{M}, w) \models \delta_1 \vee \cdots \vee \delta_n$. This is a contradiction, since $\mathsf{T}(\mathcal{M}, w) \not\models \delta_i$ for every i. □

As $\Sigma(x)$ is satisfiable in K we have a model $\mathcal{N}$ and an element v such that $\langle \mathcal{N}, v \rangle \in \mathsf{K}$ and $\mathsf{T}(\mathcal{N}, v) \models \Sigma(x)$. We make the following proposition.

Proposition 14. *$\mathcal{N}, v \rightsquigarrow_{\mathfrak{L}} \mathcal{M}, w$.*

Proof. Take the contrapositive. Suppose that $\mathcal{M}, w \not\models \varphi$ then $\neg\mathsf{Tf}_x(\varphi) \in \mathsf{NTh}^w(x)$ and because $\mathsf{NTh}^w(x) \subseteq \Sigma(x)$ we can state that $\mathsf{T}(\mathcal{N}, v) \models \neg\mathsf{Tf}_x(\varphi)$ which implies that $\mathsf{T}(\mathcal{N}, v) \not\models \mathsf{Tf}_x(\varphi)$. By truth-preservation of the translations we get $\mathcal{N}, v \not\models \varphi$. □

Now we link $\mathsf{T}(\mathcal{N}, v)$ and $\mathsf{T}(\mathcal{M}, w)$ in a way that lets us transfer the validity of $\alpha(x)$ from the first model to the second. The next lemma, which makes a detour through the class of ω-saturated models, will be useful for this matter.

Lemma 15 (Big Detour Lemma). *Let $\alpha(x) \in \text{FORM}_1(\mathfrak{F})$ be $\mathfrak{L}$-simulation K-invariant. If $\mathcal{N}, v \rightsquigarrow_{\mathfrak{L}} \mathcal{M}, w$ and $\mathsf{T}(\mathcal{N}, v) \models \alpha(x)$ then $\mathsf{T}(\mathcal{M}, w) \models \alpha(x)$.*

Proof. We define some terminology to avoid cumbersome notation. We add a subscript f to the first order translations of $\mathfrak{L}$-models, e.g. $\mathcal{M}_f = \mathsf{T}(\mathcal{M})$. We also add a superscript $+$ to first order saturated models and a superscript $*$ to modal saturated models.

Applying Theorem 21 to $\mathcal{M}, w$ and $\mathcal{N}, v$ (with $\mathsf{M}_1 = \mathsf{M}_2 = \text{MODS}(\mathfrak{L})$) we build up new models. The theorem explicitly states the relationship among them, we will use this result to prove this lemma. Hereafter we will use the same

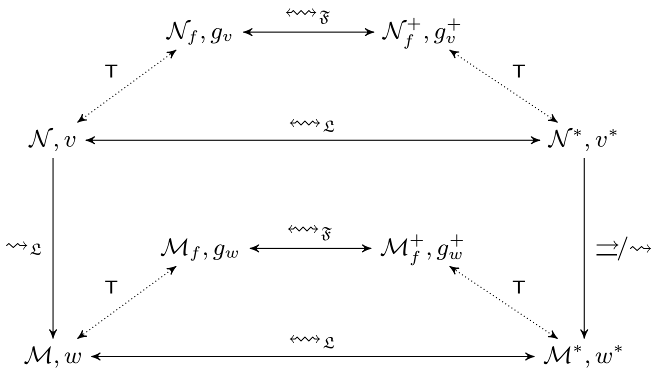

Fig. 1. Directions for the detour

notation as in Theorem 21. The diagram in Fig. 1 helps to illustrate the current situation along with the relationship between the various models. Think of it as a cube. The front face represents the $\mathfrak{L}$-models and the back face has the $\mathfrak{F}$-models.

With this new notation the lemma can be restated as follows: Let $\alpha(x)$ be an $\mathfrak{L}$-simulation K-invariant $\mathfrak{F}$-formula, if $\mathcal{N}, v \rightsquigarrow_{\mathfrak{L}} \mathcal{M}, w$ and $\mathcal{N}_f, g_v \models \alpha(x)$ then $\mathcal{M}_f, g_w \models \alpha(x)$.

We can see that, as $\mathcal{N}_f, g_v \models \alpha(x)$ and $\mathcal{N}_f^+, g_v^+$ is elementarily equivalent to $\mathcal{N}_f, g_v$, then $\mathcal{N}_f^+, g_v^+ \models \alpha(x)$. Because $\alpha(x)$ is invariant under $\mathfrak{L}$-simulations and $\mathcal{N}^*, v^* \rightrightarrows_{\mathfrak{L}} \mathcal{M}^*, w^*$ we know that $\mathcal{M}_f^+, g_w^+ \models \alpha(x)$. Again by elementary equivalence we finally conclude that $\mathcal{M}_f, g_w \models \alpha(x)$. $\square$

Since $\alpha(x) \in \Sigma(x)$ and $\mathsf{T}(\mathcal{N}, v) \models \Sigma(x)$, aplying this lemma to $\mathcal{M}, w$ and $\mathcal{N}, v$ yields $\mathsf{T}(\mathcal{M}, w) \models \alpha(x)$, and this concludes the proof of the characterization theorem.

We have proved that a $\mathfrak{F}$-formula $\alpha(x)$ is K-equivalent to the translation of an $\mathfrak{L}$-formula iff $\alpha(x)$ is K-invariant for $\mathfrak{L}$-simulations. We did it by showing that $\alpha(x)$ was equivalent to the translation of the modal consequences of $\alpha(x)$. This was accomplished by taking a detour through the class of ω-saturated *first order* models (Lemma 15).

It is worth noting that the handling of ω-saturated models has been isolated in Theorem 21. After invoking the theorem, we only used the relationships among the models named by it. Also, the requirements for the adequate pair (Definition 9) were used during the proof, e.g., closure under ultraproducts was used for compactness in Theorem 20 and the Hennessy-Milner property of the ω-saturated models was critically used in Theorem 21.

This result can already be applied to many logics, for example, ranging from sub-boolean logics to hybrid or temporal logics. It is important to notice that it allows for a relativization of the first order class of models (called K in the definition of adequate pair). This particular point will be revisited in Section 5 when we draw conclusions over the developed framework.

4 Definability

Definability theorems address the question of which properties of models are definable by means of formulas of a given logic. In this section we answer the question of Definability for $\mathfrak{L}$, our arbitrary logic under study. We present two results in this direction: one considers sets of $\mathfrak{L}$-formulas, and the other a single $\mathfrak{L}$-formula. We begin with definability by a set of $\mathfrak{L}$-formulas.

Theorem 16 (Definability by a set). *Let* $\langle \mathfrak{L}, \mathsf{K} \rangle$ *be an adequate pair and let* $\mathsf{M} \subseteq \text{PMODS}(\mathfrak{L})$*. Then* M *is definable by a set of* $\mathfrak{L}$*-formulas iff* M *is closed under* $\mathfrak{L}$*-simulations,* $\mathsf{T}(\mathsf{M})$ *is closed under ultraproducts and* $\mathsf{T}(\overline{\mathsf{M}})$ *is closed under ultrapowers.*

Proof. From left to right, suppose that M is defined by the set Γ of $\mathfrak{L}$-formulas and there is a model $\langle \mathcal{M}, w \rangle \in \mathsf{M}$ such that $\mathcal{M}, w \rightrightarrows \mathcal{N}, v$ for some pointed model $\mathcal{N}, v$. Since $\langle \mathcal{M}, w \rangle \in \mathsf{M}$, we have $\mathcal{M}, w \models \Gamma$. By simulation preservation we have $\mathcal{N}, v \models \Gamma$ and therefore $\langle \mathcal{N}, v \rangle \in \mathsf{M}$. Hence M is closed under $\mathfrak{L}$-simulations.

To see that $\mathsf{T}(\mathsf{M})$ is closed under ultraproducts we take a family of models $\langle \mathcal{M}_i^f, g_i \rangle \in \mathsf{T}(\mathsf{M})$. As $\mathcal{M}_i^f \models \mathsf{Tf}_x(\Gamma)$ for all i then, by [8, Theorem 4.1.9], the ultraproduct $\prod_D \mathcal{M}_i^f \models \mathsf{Tf}_x(\Gamma)$. We conclude that $\prod_D \mathcal{M}_i^f \in \mathsf{T}(\mathsf{M})$. We still have to check that $\mathsf{T}(\overline{\mathsf{M}})$ is closed under ultrapowers. Take $\langle \mathcal{M}^f, g \rangle \in \mathsf{T}(\overline{\mathsf{M}})$. By definition $\mathcal{M}^f, g \not\models \mathsf{Tf}_x(\Gamma)$. Let $\mathcal{M}_*^f = \prod_D \mathcal{M}^f$ be an ultrapower of $\mathcal{M}^f$. By [8, Corollary 4.1.10], the ultrapower is elementarily equivalent to the original model. Hence, for the canonical mapping $h(x) = \lambda z.g(x)$, we have $\mathcal{M}_*^f, h \not\models \mathsf{Tf}_x(\Gamma)$. This means that $\langle \mathcal{M}_*^f, h \rangle \in \mathsf{T}(\overline{\mathsf{M}})$, therefore $\mathsf{T}(\overline{\mathsf{M}})$ is closed under ultrapowers.

For the right to left direction we proceed as follows: Define the set $\Gamma = \mathsf{Th}(\mathsf{M})$, i.e. Γ is the set of all $\mathfrak{L}$-formulas which are valid in the class M. Trivially $\mathsf{M} \models \Gamma$, and it remains to show that if $\mathcal{M}, w \models \Gamma$ then $\langle \mathcal{M}, w \rangle \in \mathsf{M}$. We define the following set:

$$\mathsf{NTh}^w(x) = \{\neg \mathsf{Tf}_x(\varphi) : \varphi \text{ is an } \mathfrak{L}\text{-formula and } \mathcal{M}, w \not\models \varphi\}.$$

Using a compactness argument it can be shown that $\mathsf{NTh}^w(x)$ is satisfiable in $\mathsf{T}(\mathsf{M})$. Suppose not, then by Theorem 20, there exists a finite $\Gamma_0 \subseteq \Gamma$ such that $\Gamma_0 = \{\neg\sigma_1, \ldots, \neg\sigma_n\}$ is not satisfiable in $\mathsf{T}(\mathsf{M})$. This means that the formula $\psi = \mathsf{Tf}_x^{-1}(\sigma_1 \vee \cdots \vee \sigma_n)$ is valid in M therefore $\psi \in \Gamma$. This is absurd because $\mathcal{M}, w \not\models \sigma_i$ for any i and by hypothesis $\mathcal{M}, w \models \Gamma$. We conclude that there is a model $\langle \mathcal{N}, v \rangle \in \mathsf{M}$ such that $\mathsf{T}(\mathcal{N}, v) \models \mathsf{NTh}^w(x)$. Observe that $\mathcal{N}, v \rightsquigarrow \mathcal{M}, w$.

Let $\mathcal{M}, w \models \Gamma$ and suppose by contradiction that $\langle \mathcal{M}, w \rangle \in \overline{\mathsf{M}}$. Using Theorem 21 (with $\mathsf{M}_1 = \overline{\mathsf{M}}, \mathsf{M}_2 = \mathsf{M}$) we conclude that there exist ω-saturated extensions $\langle \mathcal{N}^*, v^* \rangle \in \mathsf{M}$ and $\langle \mathcal{M}^*, w^* \rangle \in \overline{\mathsf{M}}$ such that $\mathcal{N}^*, v^* \rightrightarrows \mathcal{M}^*, w^*$. As M is closed under $\mathfrak{L}$-simulations then $\langle \mathcal{M}, w \rangle \in \mathsf{M}$ and this is a contradiction. □

The above result gives sufficient and necessary conditions for a class of $\mathfrak{L}$-models to be definable by a set of $\mathfrak{L}$-formulas. It is worth noting most of the work is done in the first order side and therefore detached from $\mathfrak{L}$. In the last part of the theorem we make use of Theorem 21 which connects both logics through the class of ω-saturated models. This gives us another hint that this theorem isolates the very core of characterization and definability results.

Our second result considers classes of models definable by a *single* formula. To prove the result we first need the following lemmas:

Lemma 17. *Let* $\mathsf{M} \subseteq \mathrm{PMODS}(\mathfrak{L})$*. If* M *is closed under* $\mathfrak{L}$*-simulations and both* $\mathsf{T}(\mathsf{M})$ *and* $\mathsf{T}(\overline{\mathsf{M}})$ *are closed under ultrapowers then* $\mathsf{T}(\mathsf{M})$ *and* $\mathsf{T}(\overline{\mathsf{M}})$ *are* K*-closed under potential isomorphisms.*

The proof of this lemma can be found in the Appendix. Intuitively, as the notion of $\mathfrak{L}$-simulation is at most as strong as that of potential isomorphism then closure under $\mathfrak{L}$-simulations should imply closure under partial isomorphisms.

Lemma 18. *Let* $\mathsf{M} \subseteq \mathrm{PMODS}(\mathfrak{L})$*. If* M *is closed under* $\mathfrak{L}$*-simulations and both* $\mathsf{T}(\mathsf{M})$ *and* $\mathsf{T}(\overline{\mathsf{M}})$ *are closed under ultraproducts then there exists an* $\mathfrak{F}$*-formula* $\alpha(x)$ *such that for every* $\langle \mathcal{M}, g\rangle \in \mathsf{K}$ *we have* $\mathcal{M}, g \models \alpha(x)$ *iff* $\langle \mathcal{M}, g\rangle \in \mathsf{T}(\mathsf{M})$*.*

Proof. The proof is a combination of Lemma 17 and [13, Theorem A.2]. This last result uses a relativized version of the first order definability theorem (see [13] for details).

Theorem 19 (Definability by a Single Formula). *Let* $\langle \mathfrak{L}, \mathsf{K}\rangle$ *be an adequate pair and let* $\mathsf{M} \subseteq \mathrm{PMODS}(\mathfrak{L})$*. Then* M *is definable by a single* $\mathfrak{L}$*-formula iff* M *is closed under* $\mathfrak{L}$*-simulations and both* $\mathsf{T}(\mathsf{M})$ *and* $\mathsf{T}(\overline{\mathsf{M}})$ *are closed under ultraproducts.*

Proof. From left to right, suppose M is definable by a single $\mathfrak{L}$-formula φ. Using Theorem 16, as M is defined by the singleton set $T = \{\varphi\}$ we conclude that M is closed under $\mathfrak{L}$-simulations and $\mathsf{T}(\mathsf{M})$ is closed under ultraproducts.

To see that $\mathsf{T}(\overline{\mathsf{M}})$ is closed under ultraproducts proceed as follows. Observe that $\mathsf{T}(\overline{\mathsf{M}}) = \mathsf{N} \cap \mathsf{K}$. where $\mathsf{N} = \{\langle \mathcal{M}^f, g\rangle : \mathcal{M}^f, g \models \neg\mathsf{Tf}_x(\varphi)\}$. Observe that the class N is defined by the first order formula $\neg\mathsf{Tf}_x(\varphi)$, thus it is closed under ultraproducts [8]. It is easy to check that the intersection of ultraproduct-closed classes is also closed under ultraproducts. With this final observation we conclude that $\mathsf{T}(\overline{\mathsf{M}})$ is closed under ultraproducts.

For the right to left direction, because M is closed under $\mathfrak{L}$-simulations and both $\mathsf{T}(\mathsf{M})$ and $\mathsf{T}(\overline{\mathsf{M}})$ are closed under ultraproducts then, by Lemma 18, there is a first order formula $\alpha(x)$ such that for every $\langle \mathcal{M}^f, g\rangle \in \mathsf{K}$ we have $\mathcal{M}^f, g \models \alpha(x)$ iff $\langle \mathcal{M}^f, g\rangle \in \mathsf{T}(\mathsf{M})$. As M is closed under $\mathfrak{L}$-simulations then α is K-invariant for $\mathfrak{L}$-simulations. Using Theorem 11 we conclude that $\alpha(x)$ is K-equivalent to the translation of an $\mathfrak{L}$-formula φ, which defines M. □

In this result we give necessary and sufficient conditions for a class of $\mathfrak{L}$-models to be definable by a single $\mathfrak{L}$-formula. The most interesting part is the right to left direction where we use Lemma 18. For this step, standard proofs such as those found in [1,9,10] use *structural* properties of the notion of $\mathfrak{L}$-simulation. Namely, symmetry in the case of BML-bisimulation and that $\leftrightarrows \subseteq \rightrightarrows_{\mathfrak{L}}$ when $\mathfrak{L}$ is the negation free basic modal logic and $\leftrightarrows$ is the BML-bisimulation relation.

In our case we get, as a corollary of Lemma 17, that in our setting $\cong \subseteq \rightrightarrows_{\mathfrak{L}}$ for any notion of $\mathfrak{L}$-simulation regardless of its structural definition. Using this fact, the proof goes smoothly.

5 Conclusions and Further Work

We can usually think of many different notions of simulation for a given logic $\mathfrak{L}$ but, which is the correct one? At least, the following property should hold:

$$\text{If } \mathcal{M}, w \rightrightarrows_{\mathfrak{L}} \mathcal{N}, v \text{ then } \mathcal{M}, w \rightsquigarrow_{\mathfrak{L}} \mathcal{N}, v \tag{1}$$

But this is not enough. Suppose that we claim that the right notion of simulation $\rightrightarrows_{\mathfrak{L}}$ for the basic modal logic is the equivalence notion for first order (namely, partial isomorphisms). It is clear that we will be able to prove (1) but still $\rightrightarrows_{\mathfrak{L}}$ would be too strong.

In the process of finding the right simulation notion, candidates are often checked against finite models, or against image finite models. In those cases, one expects to be able to prove the converse of (1). These classes of models are special cases of ω-saturated models. The main results of this article show that proving the converse of (1) for the class of ω-saturated models is enough to develop the basic model theory for that logic, at least in what respects to Characterization and Definability. This observation stresses the crucial relationship between ω-saturated models and the suitability of a simulation notion for a given logic.

The general framework presented in this work can also be used to give new and unifying proofs of Characterization and Definability for logics where these theorems are well-known to hold, e.g. hybrid logics [13] and temporal logic with *Since* and *Until*. It is worth noting that it can even be used to prove results for non-classical modal logics such as monotonic neighbourhood logics where the models are *not* Kripke models [12]. It also establishes these theorems for logics that have not been investigated so far (e.g., *Memory Logics* [14,15]). In all cases, we only need to check that the requirements in Definition 9 are met.

In general, characterization and definability results are stated with respect to the class of *all models*. For example, BML is the fragment of first order formulas which are bisimulation invariant in the class of *all first models*. The relativization introduced in this framework (the K class in Definition 9) allows a new technique to prove this kind of results. Think of the following motivational example.

The 'Basic Temporal Logic' is a modal logic which has two modalities F and P. The classical perspective on this logic interprets it over Kripke models defined as a tuple $\langle W, R, V \rangle$ such that

$$\begin{aligned} \mathcal{M}, w \models F\varphi &\text{ iff there is a } v \text{ such that } wRv \text{ and } \mathcal{M}, v \models \varphi \\ \mathcal{M}, w \models P\varphi &\text{ iff there is a } v \text{ such that } vRw \text{ and } \mathcal{M}, v \models \varphi \end{aligned}$$

Observe that the F modality can be thought as a normal 'diamond' over the relation R but that is not possible with the P modality. An alternative is to interpret it over Kripke models with two relations R_1, R_2 where K are the models such that $R_2 = R_1^{-1}$. In this case, both modalities can be interpreted as simple 'diamonds' (which have been given fancy names F and P) over R_1 and R_2 respectively. Our framework can be used to obtain characterization and definability results for this perspective. In fact, formulas of this logic are exactly the BML-bisimulation K-invariant fragment of first order.

Although our results cover a big family of logics, they cannot be used to prove Characterization or Definability results for the class of *finite* models. This is an important class which lays beyond our framework, since it is *not* closed under ultraproducts. Many Characterization or Definability theorems are known to hold in the class of finite models [16]. Other logics outside of the scope of this paper are those without *disjuction*. For example, several description logics are known to satisfy preservation theorems but they do not have disjunction in the language (see [17]).

We think that our results can be generalized, without much difficulty, to cover the case without disjunction. On the other hand, the problem found in the class of finite models is much more difficult to avoid. Our result requires K to be closed under ultraproducts because that implies compactness and because we use ultraproducts to get ω-saturated models. Although finite models are ω-saturated, they are not closed under ultraproducts. Further study of proofs that do not use these properties [7,18] may yield a result which would be able to handle more cases uniformly.

Generalizations in this same line of thought have been pursued in the work of Hollenberg, for instance. In [19] relativized versions of the characterization and definability theorems for the so-called *normal* first order definable modalities are stated —though without proof. Those modalities are defined by Σ_1^0-formulas and, therefore, they cannot be used to define the *Since* and *Until* operators of temporal logics. Also, the results obtained in [19] do not take into account sub-boolean logics. Our proofs work for any modality with a first order translation and both boolean and sub-boolean logics. Therefore, they would subsume previous generalizations that we are aware of.

The work done by Areces and Gorín in [20] gives a uniform way to define many modalities by using the standard semantics over a restricted class of models. From our perspective, the most important point of their work is that we get a *unique* notion of model equivalence for every logic that fits in their framework. The right simulation notion turns to be the same as BML's bisimulation.

Not every modality can be expressed within their framework (e.g., the Since and Until operators [21]). Nevertheless, we think that an interesting way to continue the work in this article is to try to expand the framework developed in [20] to support more types of operators. This would allow us to give a 'canonical' simulation notion for a broader set of logics and therefore be able to easily prove the Hennessy-Milner property for them. This line of work definitely looks as a promising path to give an automatic derivation of the Characterization and Definability theorems for a greater set of modal logics.

References

1. Blackburn, P., de Rijke, M., Venema, Y.: Modal Logic. Cambridge University Press, Cambridge (2001)
2. Kamp, H.: Tense Logic and the Theory of Linear Order. PhD thesis, University of Califormia, Los Angeles (1968)
3. Goranko, V., Passy, S.: Using the universal modality: Gains and questions. Journal of Logic and Computation 2(1), 5–30 (1992)
4. de Rijke, M.: The modal logic of inequality. Journal of Symbolic Logic 57, 566–584 (1992)
5. Kozen, D.: Results on the propositional μ-calculus. Theoretical Computer Science 27(3), 333–354 (1983)
6. Sangiorgi, D.: On the origins of bisimulation and coinduction. ACM Trans. Program. Lang. Syst. 31(4), 1–41 (2009)
7. van Benthem, J.: Modal Correspondence Theory. PhD thesis, Universiteit van Amsterdam, Instituut voor Logica en Grondslagenonderzoek van Exacte Wetenschappen (1976)
8. Chang, C.C., Keisler, H.J.: Model Theory. Studies in Logic and the Foundations of Mathematics, vol. 73. Elsevier Science B.V., Amsterdam (1973)
9. Kurtonina, N., de Rijke, M.: Simulating without negation. Journal of Logic and Computation 7, 503–524 (1997)
10. Kurtonina, N., de Rijke, M.: Bisimulations for temporal logic. Journal of Logic, Language and Information 6, 403–425 (1997)
11. Keisler, H.J.: The ultraproduct construction. In: Proceedings of the Ultramath Conference, Pisa, Italy (2008)
12. Hansen, H.H.: Monotonic modal logics. Master's thesis, ILLC, University of Amsterdam (2003)
13. Carreiro, F.: Characterization and definability in modal first-order fragments. Master's thesis, Universidad de Buenos Aires (2010), arXiv:1011.4718
14. Areces, C., Figueira, D., Figueira, S., Mera, S.: Expressive power and decidability for memory logics. In: Hodges, W., de Queiroz, R. (eds.) Logic, Language, Information and Computation. LNCS (LNAI), vol. 5110, pp. 56–68. Springer, Heidelberg (2008)
15. Areces, C., Figueira, D., Figueira, S., Mera, S.: The expressive power of memory logics. Review of Symbolic Logic (to appear)
16. Rosen, E.: Modal logic over finite structures. Journal of Logic, Language and Information 6, 95–27 (1995)
17. Kurtonina, N., de Rijke, M.: Classifying description logics. In: Brachman, R.J., Donini, F.M., Franconi, E., Horrocks, I., Levy, A.Y., Rousset, M.C. (eds.) Description Logics. URA-CNRS, vol. 410 (1997)
18. Otto, M.: Elementary proof of the van Benthem-Rosen characterisation theorem. Technical Report 2342, Department of Mathematics, Technische Universität Darmstadt (2004)
19. Hollenberg, M.: Logic and Bisimulation. PhD thesis, Philosophical Institute, Utrecht University (1998)
20. Areces, C., Gorín, D.: Coinductive models and normal forms for modal logics. Journal of Applied Logic (2010) (to appear)
21. Blackburn, P., van Benthem, J., Wolter, F.: Handbook of Modal Logic. Studies in Logic and Practical Reasoning, vol. 3. Elsevier Science Inc., New York (2006)

Appendix

Theorem 20 (First order compactness relative to a class of models). *Let C be a class of first order models which is closed under ultraproducts and let Σ be a set of first order formulas. If every finite set $\Delta \subseteq \Sigma$ has a model in C then there is a model in C for Σ.*

Proof. Let $\mathcal{M}_i^f$ be a model for each finite subset $\Delta_i \subseteq \Sigma$, algebraic proofs of the compactness theorem [11, Theorem 4.3] show that the ultraproduct of the models $\mathcal{M} = \prod_U \mathcal{M}_i^f$ satisfies $\mathcal{M} \models \Sigma$ (with a suitable ultrafilter U). As each $\mathcal{M}_i^f$ is in C and C is closed under ultraproducts we conclude that $\mathcal{M} \in \mathsf{C}$.

The above argument proves the theorem for the case where Σ is a set of sentences. If Σ has some formulas with free variables $x_1, x_2, \ldots$ we proceed as follows. We extend the original first order language with constants $\overline{x}_1, \overline{x}_2, \ldots$ and use the above result with a new set Σ' where each free appearance of x_i has been replaced by $\overline{x}_i$. It is left to the reader to check that this transformation preserves satisfiability. □

Theorem 21. *Let $\langle \mathfrak{L}, \mathsf{K} \rangle$ be an adequate pair and let $\mathsf{M}_1, \mathsf{M}_2 \subseteq \mathrm{MODS}(\mathfrak{L})$ be two classes such that $\mathsf{T}(\mathsf{M}_1)$ and $\mathsf{T}(\mathsf{M}_2)$ are closed under ultrapowers. Let $\mathcal{M} \in \mathsf{M}_1$ and $\mathcal{N} \in \mathsf{M}_2$ be two $\mathfrak{L}$-models such that for some $w \in |\mathcal{M}|$, $v \in |\mathcal{N}|$ they satisfy $\mathcal{N}, v \rightsquigarrow_{\mathfrak{L}} \mathcal{M}, w$. Then there exist models $\mathcal{M}^* \in \mathsf{M}_1$ and $\mathcal{N}^* \in \mathsf{M}_2$ and elements $w^* \in |\mathcal{M}^*|$, $v^* \in |\mathcal{N}^*|$ such that*

1. $\mathsf{T}(\mathcal{M}, w) \leftrightsquigarrow_{\mathfrak{F}} \mathsf{T}(\mathcal{M}^*, w^*)$ *and* $\mathsf{T}(\mathcal{N}, v) \leftrightsquigarrow_{\mathfrak{F}} \mathsf{T}(\mathcal{N}^*, v^*)$
   Their translations are pairwise elementarily equivalent.
2. $\mathcal{M}, w \leftrightsquigarrow_{\mathfrak{L}} \mathcal{M}^*, w^*$ *and* $\mathcal{N}, v \leftrightsquigarrow_{\mathfrak{L}} \mathcal{N}^*, v^*$
   They are pairwise $\mathfrak{L}$-equivalent.
3. $\mathcal{N}^*, v^* \rightrightarrows_{\mathfrak{L}} \mathcal{M}^*, w^*$
   There is a simulation from $\mathcal{N}^*, v^*$ to $\mathcal{M}^*, w^*$.

Proof. We define some terminology for the models with which we will be working on before starting with the proof. Call $\mathcal{M}_f, g_w = \mathsf{T}(\mathcal{M}, w)$ and $\mathcal{N}_f, g_v = \mathsf{T}(\mathcal{N}, v)$. Take $\mathcal{M}_f^+, \mathcal{N}_f^+$ to be ω-saturated ultrapowers of $\mathcal{M}_f$ and $\mathcal{N}_f$. As the classes are closed under ultrapowers, the saturated models lay in the same class as the original models.

By [8, Corollary 4.1.13] we have an elementary embedding $d : |\mathcal{M}_f| \to |\mathcal{M}_f^+|$. Let g_w^+ be an assignment for $\mathcal{M}_f^+$ such that $g_w^+(x) = d(g_w(x))$. Take the modal preimage of $\mathcal{M}_f^+, g_w^+$ and call it $\mathcal{M}^*, w^* = \mathsf{T}^{-1}(\mathcal{M}_f^+, g_w^+)$. We repeat the same process and assign similar names to models and points deriving from $\mathcal{N}$.

1. As a consequence of [8, Corollary 4.1.13], since there is an elementary embedding, we have that $\mathcal{M}_f, g_w \leftrightsquigarrow_{\mathfrak{F}} \mathcal{M}_f^+, g_w^+$. The same argument works with $\mathcal{N}_f$ and $\mathcal{N}_f^+$.
2. Following the last point, we can conclude, through the truth-preserving translations, that $\mathcal{M}, w \leftrightsquigarrow_{\mathfrak{L}} \mathcal{M}^*, w^*$. The same proof works with $\mathcal{N}, v$ and $\mathcal{N}^*, v^*$. Corollary: $\mathcal{N}^*, v^* \rightsquigarrow_{\mathfrak{L}} \mathcal{M}^*, w^*$.

3. As both $\mathcal{M}_f^+$ and $\mathcal{N}_f^+$ are ω-saturated, by definition of adequate pair, that implies that they have the Hennesy-Milner property. Therefore, because we proved that $\mathcal{N}^*, v^* \rightsquigarrow_{\mathfrak{L}} \mathcal{M}^*, w^*$ we conclude that $\mathcal{N}^*, v^* \rightrightarrows_{\mathfrak{L}} \mathcal{M}^*, w^*$. □

Lemma 17. *Let* $\mathsf{M} \subseteq \mathrm{PMODS}(\mathfrak{L})$*. If* M *is closed under* $\mathfrak{L}$*-simulations and both* $\mathsf{T}(\mathsf{M})$ *and* $\mathsf{T}(\overline{\mathsf{M}})$ *are closed under ultrapowers then* $\mathsf{T}(\mathsf{M})$ *and* $\mathsf{T}(\overline{\mathsf{M}})$ *are* K*-closed under potential isomorphisms.*

Proof. Suppose that $\mathsf{T}(\mathsf{M})$ is not K-closed under potential isomorphisms. This means that there exist models $\langle \mathcal{M}^f, g\rangle \in \mathsf{T}(\mathsf{M})$ and $\langle \mathcal{N}^f, h\rangle \in \mathsf{T}(\overline{\mathsf{M}})$ such that $\mathcal{M}^f, g \cong \mathcal{N}^f, h$. Recall that $\mathsf{K} \setminus \mathsf{T}(\mathsf{M}) = \mathsf{T}(\overline{\mathsf{M}})$. For a smoother proof, call their modal counterparts $\mathcal{M}, w$ and $\mathcal{N}, v$ respectively. Therefore $\langle \mathcal{M}, w\rangle \in \mathsf{M}$ and $\langle \mathcal{N}, v\rangle \notin \mathsf{M}$.

As $\mathcal{M}^f, g \cong \mathcal{N}^f, h$ we know by [8, Proposition 2.4.4] that $\mathcal{M}^f, g \models \varphi(x)$ if and only if $\mathcal{N}^f, h \models \varphi(x)$. In particular they have the same modal theory, $\mathcal{M}, w \leftrightsquigarrow_{\mathfrak{L}} \mathcal{N}, v$. As this implies that $\mathcal{M}, w \rightsquigarrow_{\mathfrak{L}} \mathcal{N}, v$ we can use Theorem 21 (instantiating with $\mathsf{K}_1 = \mathsf{T}(\mathsf{M})$, $\mathsf{K}_2 = \mathsf{T}(\overline{\mathsf{M}})$ and $\mathcal{M}$, $\mathcal{N}$ interchanged) and get models $\langle \mathcal{M}^*, w^*\rangle \in \mathsf{M}$ and $\langle \mathcal{N}^*, v^*\rangle \in \overline{\mathsf{M}}$ such that $\mathcal{M}^*, w^* \rightrightarrows_{\mathfrak{L}} \mathcal{N}^*, v^*$.

Knowing that $\mathcal{M}^*, w^* \rightrightarrows_{\mathfrak{L}} \mathcal{N}^*, v^*$ and that M is closed under simulations we conclude that $\langle \mathcal{N}^*, v^*\rangle \in \mathsf{M}$. This is absurd because it contradicts $\langle \mathcal{N}^*, v^*\rangle \in \overline{\mathsf{M}}$. Hence $\mathsf{T}_{\mathsf{K}}(\mathsf{M})$ is K-closed under potential isomorphisms.

To see that $\mathsf{T}(\overline{\mathsf{M}})$ is K-closed under potential isomorphisms we argue by contradicction. Suppose it is not the case, then there exist $\langle \mathcal{M}^f, g\rangle \in \mathsf{T}(\overline{\mathsf{M}})$ and $\langle \mathcal{N}^f, h\rangle \in \mathsf{K} \setminus \mathsf{T}(\overline{\mathsf{M}})$ such that $\mathcal{M}^f, g \cong \mathcal{N}^f, h$. As $\langle \mathcal{N}^f, h\rangle \in \mathsf{K} \setminus \mathsf{T}(\overline{\mathsf{M}})$ this means that $\langle \mathcal{N}^f, h\rangle \in \mathsf{T}(\mathsf{M})$. We have just proved that $\mathsf{T}(\mathsf{M})$ is K-closed under potential isomorphism then, as $\mathcal{M}^f, g \cong \mathcal{N}^f, h$ (and because of the symmetry of the potential isomorphism relation), we conclude that $\langle \mathcal{M}^f, g\rangle \in \mathsf{T}(\mathsf{M})$ which contradicts our hypothesis. Absurd. □