HAL open science

Generalized Keisler Theorems for First-order Logic and Protologics

Guillaume Aucher

To cite this version:

Guillaume Aucher. Generalized Keisler Theorems for First-order Logic and Protologics. Université de Rennes 1 (UR1). 2022. hal-03884245

HAL Id: hal-03884245

https://hal.science/hal-03884245v1

Submitted on 5 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Generalized Keisler Theorems for First-order Logic and Protologics

Guillaume Aucher
Univ Rennes, IRISA, IRMAR
263, Avenue du Général Leclerc
35042 Rennes Cedex, France
guillaume.aucher@univ-rennes1.fr

December 4, 2022

Abstract

The Keisler theorems dealing with the definability in first-order logic of classes of structures are generalized and adapted to non-classical logics. On the one hand, we generalize and prove by completely different means an analogue of the Keisler-Shelah isomorphism theorem for first-order logic and countable languages, where the notion of isomorphism is replaced in that theorem by a variant of partial isomorphism. On the other hand, we adapt the Keisler theorems for first-order logics to protologics, *i.e.* logics such that the truth conditions of their connectives are expressible by first-order formulas. Our results are based on similar theorems for atomic and molecular logics proved in a companion article [5]. We indeed show in this present article that first-order logic and protologics are as expressive as some atomic and molecular logics.

1 Introduction

The model theory of non-classical logics is often developed on a case by case basis by considering each logic independently, and similar theorems are proved for each non-classical logic by frequently adapting the same proof method. For example, for modal logic, temporal logic and XPath, a number of model-theoretical results dealing with the definability of classes of models by means of a set of formulas or a single formula have been proved [9, 19, 1] by adapting the Keisler theorems of first-order logic (FOL for short) [10]. This type of theorems provides conditions of definability of classes of models in a specific logic. More precisely, they state that a class of models is definable in a given logic if this class of models and its complement are both closed under a specific construction called ultraproduct and under a specific notion of bisimulation associated to the logic. Similar adaptations to other logics have been made for the van Benthem characterization theorem of modal logic [25, 19, 7, 18, 11]. A natural question that comes up to mind is to wonder whether the Keisler theorems of FOL transfer to an arbitrary logic. We shall see in that paper that the Keisler theorems do transfer to any protologic, *i.e.* any logic such that the truth conditions of its connectives are expressible by first-order formulas, if these connectives are so-called 'normal'.

To prove our results, we will resort to the framework of atomic and molecular logics introduced in [6]. This framework is based on Dunn's gaggle theory [12, 13]. Atomic logics generalize "gaggle logics" by adding types to formulas. Gaggle logics were introduced in [4] where it is also shown

that a very large number of non-classical logics are actually gaggle logics and therefore also atomic logics (see Figures 6-11 in that article). Atomic and molecular logics are a generalization of gaggle logics which behave as 'normal form' logics. We will indeed show that every non-classical logic such that the truth conditions of its connectives are expressible in first-order logic is as expressive as an atomic or molecular logic.

Moreover, as it turns out, an appropriate notion of bisimulation can be automatically associated to any atomic or molecular logic from the truth conditions of its connectives [5]. On the one hand, our embedding of FOL into atomic logics will allow us to automatically derive our notion of invariance for FOL, that we call 'predicate bisimulation' and 'first-order bisimulation', and to obtain our generalization of the Keisler theorems for FOL. Our notion of predicate bisimulation turns out to be a slight and natural variant of the classical notion of partial isomorphism. On the other hand, our embedding of protologics into molecular logics will allow us to obtain automatically invariance notions for any protologic and our generalization of the Keisler theorems for protologics.

Structure of the article We start in Section 2 by recalling first–order logics and in Section 3 modal logic. In Section 4, we introduce a specific notion of equi-expressivity for logics which are not based on the same classes of models. In Section 5 we introduce atomic logics, in Section 6 molecular logics and in Section 7 their Boolean versions with Boolean connectives. In Section 9, we show how notions of bisimulations can be automatically defined from the truth conditions of the connectives of atomic and molecular logics, after some formal preliminaries in Section 8 where we introduce universal and existential connectives. In Section 10, we adapt the notions of ultraproducts and ultrapowers to atomic and molecular logics. In Section 11, we show that first-order logics are as expressive as atomic logics. This leads us in Section 12 to apply our general results for atomic logics to first-order logics. In doing so, we generalize and (re)discover an analogue of the Keisler theorems. Likewise, in Section, 13, we recall and show how every protologic is as expressive as some molecular logic. This leads us in Section 14 to prove similar results for protologics. Finally, we show in Section 15 that bisimulations of schematically equally expressive logics are equivalent. We conclude in Section 16. Unless otherwise stated, all the proofs of this paper are in this appendix or in the appendix of the companion article [5].

2 First-Order Logics

In this section, we recall FOL. In the sequel, all logics will always be semantically presented by following a tri-partite representation: language, class of models, satisfaction relation.

The set $\mathcal{P} \triangleq \{\mathsf{R}_1, \ldots, \mathsf{R}_n, \ldots\}$ is a set of *predicate symbols* of arity $k_1, \ldots, k_n, \ldots$ respectively (one of them can be the identity predicate $=$ of arity 2), $\mathcal{F} \triangleq \{\mathsf{f}_1, \ldots, \mathsf{f}_n, \ldots\}$ is a set of *function symbols*, $\mathcal{V} \triangleq \{v_1, \ldots, v_n, \ldots\}$ is a set of *variables* and $\mathcal{C} \triangleq \{c_1, \ldots, c_n, \ldots\}$ is a set of *constants*. Each of these sets can be finite or infinite. $v_1, v_2, v_3, \ldots$ are the names of the variables and we use the expressions $x, x_1, x_2, \ldots, y, y_1, y_2, \ldots, z, z_1, z_2, \ldots$ to refer to arbitrary variables or constants, which can be for example $v_{42}, v_5, c_{101}, c_{21}, \ldots$ $\mathrm{Arity}(\mathcal{P}, \mathcal{F})$ is the set of all arities of predicate and function symbols. The *first-order language* $\mathcal{L}^{\mathcal{P}\mathcal{F}}_{\mathsf{FOL}}$ is defined inductively by the following grammars in BNF:

$$\begin{array}{rrcl} \mathcal{T}: & t & ::= & x \mid \mathsf{c} \mid \mathsf{f}t \ldots t \\ \mathcal{L}^{\mathcal{P}\mathcal{F}}_{\mathsf{FOL}}: & \varphi & ::= & \mathsf{R}t \ldots t \mid \bot \mid (\varphi \to \varphi) \mid \forall x \varphi \end{array}$$

where $x \in \mathcal{V}$, $\mathsf{c} \in \mathcal{C}$, $\mathsf{f} \in \mathcal{F}$, $t \in \mathcal{T}$ and $\mathsf{R} \in \mathcal{P}$. Elements of $\mathcal{T}$ are called *terms* and elements of $\mathcal{L}^{\mathcal{P}\mathcal{F}}_{\mathsf{FOL}}$ are called *first–order formulas*. Formulas of the form $\mathsf{R}t_1 \ldots t_k$ are called *atomic formulas*

and first–order formulas without function symbols are called *pure predicate formulas.*

If $\varphi \in \mathcal{L}_{\mathsf{FOL}}^{\mathcal{PF}}$, the *Boolean negation* of φ, denoted $\neg\varphi$, is defined by the abbreviation $\neg\varphi \triangleq (\varphi \to \bot)$. We also use the abbreviations $\top \triangleq \neg\bot$, $(\varphi \vee \psi) \triangleq (\neg\varphi \to \psi)$, $(\varphi \wedge \psi) \triangleq \neg(\neg\varphi \vee \neg\psi)$ and $(\varphi \leftrightarrow \psi) \triangleq (\varphi \to \psi) \wedge (\psi \to \varphi)$ as well as the abbreviations $\exists x\varphi \triangleq \neg\forall x\neg\varphi$, $\forall x_1 \ldots x_n\varphi \triangleq \forall x_1 \ldots \forall x_n\varphi$, $\exists x_1 \ldots x_n\varphi \triangleq \exists x_1 \ldots \exists x_n\varphi$ and $\forall\overline{x}\varphi \triangleq \forall x_1 \ldots x_n\varphi$ if $\overline{x} = (x_1, \ldots, x_n)$ is a tuple of variables.

Let $\varphi \in \mathcal{L}_{\mathsf{FOL}}^{\mathcal{PF}}$. An occurrence of a variable x in φ is *free* (in φ) if, and only if, x is not within the scope of a quantifier of φ. We say that a formula of $\mathcal{L}_{\mathsf{FOL}}^{\mathcal{PF}}$ is a *sentence* (or is *closed*) when it contains no free variable. We denote by $\varphi(x_1, \ldots, x_k)$ a formula of $\mathcal{L}_{\mathsf{FOL}}^{\mathcal{PF}}$ whose free variables or constants coincide *exactly* with $x_1, \ldots, x_k$. We assume that these variables and constants are all distinct. In doing so, we depart from the literature in which this notation means that the free variables of φ are *included* in $\{x_1, \ldots, x_k\}$.

We denote by $\mathcal{L}_{\mathsf{FOL}}^{\mathcal{P}}$ the fragment of $\mathcal{L}_{\mathsf{FOL}}^{\mathcal{PF}}$ whose formulas do not contain function symbols. We denote by $\mathcal{L}_{\mathsf{FOL}}^{\mathcal{P}}(\overline{x})$ (resp. $\mathcal{L}_{\mathsf{FOL}}^{\mathcal{PF}}(\overline{x})$) the fragment of $\mathcal{L}_{\mathsf{FOL}}^{\mathcal{P}}$ (resp. $\mathcal{L}_{\mathsf{FOL}}^{\mathcal{PF}}$) whose formulas all contain at least one free variable or constant. For all $k \in \mathbb{N}^*$ and $\overline{x} = (x_1, \ldots, x_k) \in (\mathcal{V} \cup \mathcal{C})^k$, we denote by $\mathcal{L}_{\mathsf{FOL}}^{\mathcal{P}}(\overline{x}, k)$ the fragment of $\mathcal{L}_{\mathsf{FOL}}^{\mathcal{P}}$ whose formulas all contain *exactly* k free variables or constants and these variables or constants are $\overline{x}$. If $X \subseteq \mathcal{V} \cup \mathcal{C}$ is a non-empty set, we denote by $\mathcal{L}_{\mathsf{FOL}}^{\mathcal{P}}(X)$ (resp. $\mathcal{L}_{\mathsf{FOL}}^{\mathcal{PF}}(X)$) the fragment of $\mathcal{L}_{\mathsf{FOL}}^{\mathcal{P}}$ (resp. $\mathcal{L}_{\mathsf{FOL}}^{\mathcal{PF}}$) whose formulas are such that their free variables and constants are all contained in X. We denote by $\mathcal{L}_{\mathsf{FOL}}^{\mathcal{P}}(\emptyset)$ (resp. $\mathcal{L}_{\mathsf{FOL}}^{\mathcal{PF}}(\emptyset)$) the set of sentences of $\mathcal{L}_{\mathsf{FOL}}^{\mathcal{P}}$ (resp. $\mathcal{L}_{\mathsf{FOL}}^{\mathcal{PF}}$) without constants. A language $\mathcal{L}_{\mathsf{FOL}} \subseteq \mathcal{L}_{\mathsf{FOL}}^{\mathcal{PF}}$ is *countable* if its set of predicate symbols, function symbols, variables and constants is countable.

A *structure* is a tuple $M \triangleq (W, \{R_1, \ldots, R_n, \ldots, f_1, \ldots, f_n, \ldots, c_1, \ldots, c_n, \ldots\})$ where:

- W is a non-empty set called the *domain*;
- $R_1, \ldots, R_n, \ldots$ are relations over W with the same arity as $\mathsf{R}_1, \ldots, \mathsf{R}_n, \ldots$ respectively;
- $f_1, \ldots, f_n, \ldots$ are functions over W with the same arity as $\mathsf{f}_1, \ldots, \mathsf{f}_n, \ldots$ respectively;
- $c_1, \ldots, c_n, \ldots \in W$ are elements of the domain called *distinguished elements.*

An *assignment* over M is a mapping $s : \mathcal{V} \cup \mathcal{C} \to W$ such that for all $\mathsf{c}_i \in \mathcal{C}$, $s(\mathsf{c}_i) = c_i$. If s is an assignment, $s[x := w]$ is the same assignment as s except that the value of the variable $x \in \mathcal{V}$ is assigned to w. A pair of structure and assignement (M, s) is called a *pointed structure.* The class of all pointed structures (M, s) is denoted $\mathcal{M}_{\mathsf{FOL}}$. If K is a class of pointed structures, $\overline{K}$ is $\mathcal{M}_{\mathsf{FOL}} - K$.

If (M, s) is a pointed structure, we extend the assignment s from variables and constants to terms and define the *extended assignment* $\overline{s} : \mathcal{T} \to W$ as follows:

$$\begin{array}{rcl} \overline{s}(x) & \triangleq & s(x) \\ \overline{s}(\mathsf{c}) & \triangleq & s(\mathsf{c}) \\ \overline{s}(\mathsf{f}t_1 \ldots t_k) & \triangleq & f(\overline{s}(t_1), \ldots, \overline{s}(t_k)). \end{array}$$

The *satisfaction relation* $\models_{\mathsf{FOL}} \subseteq \mathcal{M}_{\mathsf{FOL}} \times \mathcal{L}_{\mathsf{FOL}}^{\mathcal{PF}}$ is defined inductively as follows. Below, we write $(M, s) \models \varphi$ for $((M, s), \varphi) \in \models_{\mathsf{FOL}}$.

$$\begin{array}{lll} (M, s) \models \bot & & \text{never;} \\ (M, s) \models \mathsf{R}_i t_1 \ldots t_k & \text{iff} & (\overline{s}(t_1), \ldots, \overline{s}(t_k)) \in R_i; \\ (M, s) \models (\varphi \to \psi) & \text{iff} & \text{if } (M, s) \models \varphi \text{ then } (M, s) \models \psi; \\ (M, s) \models \forall x\varphi & \text{iff} & (M, s[x := w]) \models \varphi \text{ for all } w \in W. \end{array}$$

In the literature [10], $(M, s) \models \varphi(x_1, \ldots, x_k)$ is sometimes denoted $M \models \varphi(x_1, \ldots, x_k)[w_1, \ldots, w_k]$, $M \models \varphi[w_1/x_1, \ldots, w_k/x_k]$ or simply $M \models \varphi[w_1, \ldots, w_k]$, with $w_1 = s(x_1), \ldots, w_k = s(x_k)$. Some

other times [14], it is denoted $M \models \varphi(x_1, \ldots, x_n)[s]$, $M, s \models \varphi(x_1, \ldots, x_n)$ or simply $M \models \varphi[s]$. In that case, we say that (M, s) makes φ *true*. We depart from the literature by treating constants on a par with variables: the denotation of constants is usually not dealt with by means of assignments.

We say that the formula $\varphi \in \mathcal{L}_{\mathsf{FOL}}^{\mathcal{PF}}$ is *realized in* M when there is an assignment s such that $(M, s) \models \varphi$.

A triple of the form $(\mathcal{L}_{\mathsf{FOL}}, \mathcal{E}_{\mathsf{FOL}}, \models_{\mathsf{FOL}})$ is called the *first–order logic associated to* $\mathcal{L}_{FOL}$ *and* $\mathcal{E}_{FOL}$. If $\mathcal{L}_{\mathsf{FOL}} = \mathcal{L}_{\mathsf{FOL}}^{\mathcal{P}}$, the triple $\left(\mathcal{L}_{\mathsf{FOL}}^{\mathcal{P}}, \mathcal{E}_{\mathsf{FOL}}, \models_{\mathsf{FOL}}\right)$ is called *pure predicate logic (associated to* $\mathcal{E}_{FOL}$*)*, if $\mathcal{L}_{\mathsf{FOL}} = \mathcal{L}_{\mathsf{FOL}}^{\mathcal{P}}(\overline{x})$, the triple $(\mathcal{L}_{\mathsf{FOL}}^{\mathcal{P}}(\overline{x}), \mathcal{E}_{\mathsf{FOL}}, \models_{\mathsf{FOL}})$ is called *pure predicate logic with free variables and constants (associated to* $\mathcal{E}_{FOL}$*)*. When $\mathcal{E}_{\mathsf{FOL}}$ is $\mathcal{M}_{\mathsf{FOL}}$, they are simply called respectively *pure predicate logic* and *pure predicate logic with free variables and constants*.

3 Modal Logic

In this section, $\mathbb{A}$ is a set of *propositional letters* which can be finite or infinite. The set $\mathbb{I}$ is a set of indices which can be finite or infinite. The *multi-modal language* $\mathcal{L}_{\mathsf{ML}}$ is defined inductively by the following grammar in BNF:

$$\mathcal{L}_{\mathsf{ML}}: \quad \varphi \quad ::= \quad p \quad | \quad \neg p \quad | \quad (\varphi \wedge \varphi) \quad | \quad (\varphi \vee \varphi) \quad | \quad \Diamond_j \varphi \quad | \quad \Box_j \varphi$$

where $p \in \mathbb{A}$ and $j \in \mathbb{I}$.

We present the so-called *possible world semantics* of modal logic. A *Kripke model* M is a tuple $M \triangleq (W, \{R_1, \ldots, R_m, \ldots, P_1, \ldots, P_n, \ldots\})$ where

- W is a non-empty set whose elements are called *possible worlds*;
- $R_1, \ldots, R_m, \ldots \subseteq W \times W$, $m \in \mathbb{I}$ are binary relations over W called *accessibility relations*;
- $P_1, \ldots, P_n, \ldots \subseteq W$ are unary relations interpreting the propositional letters of $\mathbb{A}$.

We write $w \in M$ for $w \in W$ by abuse and the pair (M, w) is called a *pointed Kripke model*. The class of all pointed Kripke models is denoted $\mathcal{E}_{\mathsf{ML}}$.

We define the *satisfaction relation* $\models_{\mathsf{ML}} \subseteq \mathcal{E}_{\mathsf{ML}} \times \mathcal{L}_{\mathsf{ML}}$ inductively by the following *truth conditions*. Below, we write $(M, w) \models \varphi$ for $((M, w), \varphi) \in \models_{\mathsf{ML}}$. For all $(M, w) \in \mathcal{E}_{\mathsf{ML}}$, all $\varphi, \psi \in \mathcal{L}_{\mathsf{ML}}$, all $p_i \in \mathbb{A}$ and all $j \in \mathbb{I}$,

$(M, w) \models p_i$	iff	$P_i(w)$ holds;
$(M, w) \models \neg p_i$	iff	$P_i(w)$ does not hold;
$(M, w) \models (\varphi \wedge \psi)$	iff	$(M, w) \models \varphi$ and $(M, w) \models \psi$;
$(M, w) \models (\varphi \vee \psi)$	iff	$(M, w) \models \varphi$ or $(M, w) \models \psi$;
$(M, w) \models \Diamond_j \varphi$	iff	there exists $v \in W$ such that $R_j wv$ and $(M, v) \models \varphi$;
$(M, w) \models \Box_j \varphi$	iff	for all $v \in W$ such that $R_j wv, (M, v) \models \varphi$.

The triple $(\mathcal{L}_{\mathsf{ML}}, \mathcal{E}_{\mathsf{ML}}, \models_{\mathsf{ML}})$ forms a logic, that we call *modal logic*. Bisimulations for modal logic can be found in [9].

3.1 Common Logical Notions

In the present section, we define a number of notions which are common to all logics and in particular to the logics introduced beforehand. The way we define logics is different from many proposals considered in universal logic [8] such as pairs of Suzsko's abstract logics, Tarski's

consequence operators or logical structures. Often a logic is viewed as a pair of a language together with a consequence relation on this language. Our approach to defining logics is somehow more 'semantic' in that respect than the usual proposals. It corresponds in fact to the "abstract logics" of García-Matos & Väänänen [15] or to the "rooms" of Mossakowski et al. [23].

A *logic* is a triple $\mathsf{L} \triangleq (\mathcal{L}, \mathcal{E}, \models)$ where

- $\mathcal{L}$ is a *logical language* defined as a set of well-formed expressions built from a set of *connectives* C and a set of *propositional letters* $\mathbb{A}$;
- $\mathcal{E}$ is a *class of pointed models*;
- $\models$ is a *satisfaction relation* which relates in a compositional manner elements of $\mathcal{L}$ to models of $\mathcal{E}$ by means of so-called *truth conditions*.

Let $\mathsf{L} = (\mathcal{L}, \mathcal{E}, \models)$ be a logic and let $\Gamma \subseteq \mathcal{L}$, $\varphi \in \mathcal{L}$ and $M \in \mathcal{E}$. We write $M \models \Gamma$ when for all $\psi \in \Gamma$, we have $M \models \psi$. Then, we say that

- φ is *true* (*satisfied*) at M or M is a *model* of φ when $M \models \varphi$;
- φ is a *logical consequence* of Γ, written $\Gamma \models_{\mathsf{L}} \varphi$, when for all $M \in \mathcal{E}$, if $M \models \Gamma$ then $M \models \varphi$;
- φ is *valid*, written $\models_{\mathsf{L}} \varphi$, when for all models $M \in \mathcal{E}$, we have $M \models \varphi$;
- φ is *satisfiable* when there is a model $M \in \mathcal{E}$ such that $M \models \varphi$.

If Γ is a singleton $\Gamma = \{\psi\}$, we also write by abuse $\psi \models \varphi$ for $\{\psi\} \models \varphi$.

A set of formulas of $\mathcal{L}$ is called a *theory*. A set Δ of formulas of $\mathcal{L}$ is said to be a *set of axioms* for a theory Γ iff Γ and Δ have the same logical consequences. A theory is called *finitely axiomatizable* iff it has a finite set of axioms. A logic L is *axiomatizable* if its set of validities is finitely axiomatizable.

4 On the Relative Expressivity of Logics

When two logics $\mathsf{L}_1 = (\mathcal{L}_1, \mathcal{E}_1, \models_1)$ and $\mathsf{L}_2 = (\mathcal{L}_2, \mathcal{E}_2, \models_2)$ are interpreted over different classes of models $\mathcal{E}_1$ and $\mathcal{E}_2$, there is no canonical way to compare their relative expressiveness, even if some proposals have already been made [15, 23]. We are now going to propose some new notions to deal with that issue.

Given a logic $(\mathcal{L}, \mathcal{E}, \models)$, for all $\varphi, \psi \in \mathcal{L}$, we write $\varphi \equiv \psi$ when for all $M \in \mathcal{E}$, it holds that $M \models \varphi$ iff $M \models \psi$ and for all $M, N \in \mathcal{E}$, we write $M \equiv N$ when for all $\varphi \in \mathcal{L}$, it holds $M \models \varphi$ iff $N \models \varphi$. If $\mathcal{M}, \mathcal{N} \subseteq \mathcal{E}$, we write $\mathcal{M} \equiv \mathcal{N}$ when for all $M \in \mathcal{M}$ there is $N \in \mathcal{N}$ such that $M \equiv N$, and vice versa.

Definition 1 (Equi-expressivity)**.** A logic $\mathsf{L}_1 = (\mathcal{L}_1, \mathcal{E}_1, \models_1)$ is *as expressive as* a logic $\mathsf{L}_2 = (\mathcal{L}_2, \mathcal{E}_2, \models_2)$, written $\mathsf{L}_1 \equiv \mathsf{L}_2$, when the following hold:

1. there is a mapping $T_1^{\varphi} : \mathcal{L}_1 \to \mathcal{L}_2$ and a mapping $T_1^{M} : \mathcal{E}_1 \to \mathcal{E}_2$ such that for all $\varphi_1 \in \mathcal{L}_1$ and all $M_1 \in \mathcal{E}_1$, it holds that $M_1 \models \varphi_1$ iff $T_1^{M}(M_1) \models T_1^{\varphi}(\varphi_1)$;
2. there is a mapping $T_2^{\varphi} : \mathcal{L}_2 \to \mathcal{L}_1$ and a mapping $T_2^{M} : \mathcal{E}_2 \to \mathcal{E}_1$ such that for all $\varphi_2 \in \mathcal{L}_2$ and all $M_2 \in \mathcal{E}_2$, it holds that $M_2 \models \varphi_2$ iff $T_2^{M}(M_2) \models T_2^{\varphi}(\varphi_2)$;
3. for all $M_1 \in \mathcal{E}_1$ and all $M_2 \in \mathcal{E}_2$, it holds that $T_2^{M}(T_1^{M}(M_1)) \equiv M_1$ and $T_1^{M}(T_2^{M}(M_2)) \equiv M_2$.

Let $\mathbb{L}_1$ and $\mathbb{L}_2$ be two classes of logics. We say that $\mathbb{L}_1$ is *as expressive as* $\mathbb{L}_2$, written $\mathbb{L}_1 \equiv \mathbb{L}_2$, when for all $\mathsf{L}_1 \in \mathbb{L}_1$ there is $\mathsf{L}_2 \in \mathbb{L}_2$ such that $\mathsf{L}_1 \equiv \mathsf{L}_2$, and vice versa. $\square$

Our third condition states that T_2^M and T_1^M are inverse bijections of each other (modulo some natural congruence). Hence, our definition is set in such a way that we compare the relative expressivity of each logic by comparing them over their whole class of models, taking into account the specificities of *all* the models of each logic in the comparison.

Our proposal is different from the one of García-Matos & Väänänen [15]. However, they deal with a more general notion of embedding between logics based on different classes of models, that is an embedding in only one direction, of one logic into another. If two logics L_1 and L_2 are essentially equally expressive in our sense then there exist two "model-expansive corridors" from L_1 to L_2 and from L_2 to L_1 in the sense of Mossakowski et al. [23] (with the proviso that the surjection holds modulo the congruence $\equiv$). Thus, our notion of equi-expressivity is more demanding and stronger than the one of Mossakowski et al. [23]. On the other hand, our "logics" can in fact be seen as institutions [23]. In that case, one can prove that two logics are "equally expressive" in our sense if, and only if, they are "equivalent" in the sense of institutions [24, Definition 3.5].[1]

5 Atomic Logics

Atomic logics are logics such that the truth conditions of their connectives are defined by first-order formulas of the form $\forall x_1 \dots x_n(\pm_1 \mathsf{Q}_1 x_1 \vee \dots \vee \pm_n \mathsf{Q}_n x_n \vee \pm \mathsf{R} x_1 \dots x_n x)$ or $\exists x_1 \dots x_n(\pm_1 \mathsf{Q}_1 x_1 \wedge \dots \wedge \pm_n \mathsf{Q}_n x_n \wedge \pm \mathsf{R} x_1 \dots x_n x)$ where the $\pm_i$s and $\pm$ are either empty or $\neg$. Likewise, propositional letters are defined by first-order formulas of the form $\pm \mathsf{R} x$. We will represent the structure of these formulas by means of so–called *skeletons* whose various arguments capture the different features and patterns from which they can be redefined completely. Atomic logics are also generalizations of our gaggle logics [4] with types associated to formulas.

We recall that $\mathbb{N}^*$ denotes the set of natural numbers without 0 and that for all $n \in \mathbb{N}^*$, $\mathfrak{S}_n$ denotes the group of permutations over the set $\{1, \dots, n\}$. Permutations are generally denoted σ, τ, the identity permutation Id is sometimes denoted 1 as the neutral element of every permutation group and σ^- stands for the inverse permutation of the permutation σ. For example, the permutation $\sigma = (3, 1, 2)$ is the permutation that maps 1 to 3, 2 to 1 and 3 to 2 (see for instance [26] for more details).

Definition 2 (Atomic skeletons and connectives)**.** The sets of *atomic skeletons* $\mathbb{P}$ and $\mathbb{C}$ are defined as follows:

$$\mathbb{P} \triangleq \mathfrak{S}_1 \times \{+, -\} \times \{\forall, \exists\} \times \mathbb{N}^*$$
$$\mathbb{C} \triangleq \mathbb{P} \cup \bigcup_{n \in \mathbb{N}^*} \left\{ \mathfrak{S}_{n+1} \times \{+, -\} \times \{\forall, \exists\} \times {\mathbb{N}^*}^{n+1} \times \{+, -\}^n \right\}.$$

$\mathbb{P}$ is called the set of *propositional letter skeletons* and $\mathbb{C}$ is called the set of *connective skeletons*. They can be represented by tuples $(\sigma, \pm, Æ, \overline{k}, \overline{\pm}_j)$ or $(\sigma, \pm, Æ, k)$ if it is a propositional letter skeleton, where $Æ \in \{\forall, \exists\}$ is called the *quantification signature* of the skeleton, $\overline{k} = (k, k_1, \dots, k_n) \in {\mathbb{N}^*}^{n+1}$ is called the *type signature* of the skeleton and $\overline{\pm}_j = (\pm_1, \dots, \pm_n) \in \{+, -\}^n$ is called the *tonicity signature* of the skeleton; $(Æ, \overline{k}, \overline{\pm_j})$ is called the *signature* of the skeleton. The *arity* of a propositional letter skeleton is 0 and its *type* is k. The *arity* of a skeleton $\star \in \mathbb{C}$ is n, its *input types* are $k_1, \dots, k_n$ and its *output type* is k.

[1] I thank Peter Arndt for proving that result.

A *(atomic) connective* or *propositional letter* is a symbol generally denoted $\star$ or p to which is associated a (atomic) skeleton. Its arity, signature, quantification signature, type signature, tonicity signature, input and output types are the same as its skeleton. By abuse, we sometimes identify a connective with its skeleton. If C is a set of atomic connectives, its set of propositional letters is denoted $\mathbb{P}(\mathsf{C})$.

Propositional letters are denoted p, p_1, p_2, *etc.* and connectives are denoted $\star, \star_1, \star_2$, *etc.* $\square$

We need to distinguish between connectives and skeletons because in general we need a countable number of propositional letters or connectives of the same skeleton, like in some modal logics, where we need multiple modalities of the same similarity type/skeleton.

Definition 3 (Atomic language)**.** Let C be a set of atomic connectives. The *(typed) atomic language* $\mathcal{L}_{\mathsf{C}}$ associated to C is the smallest set that contains the propositional letters and that is closed under the atomic connectives. That is,

- $\mathbb{P}(\mathsf{C}) \subseteq \mathcal{L}_{\mathsf{C}}$;
- for all $\star \in \mathsf{C}$ of arity $n > 0$ and of type signature $(k, k_1, \ldots, k_n)$ and for all $\varphi_1, \ldots, \varphi_n \in \mathcal{L}_{\mathsf{C}}$ of types $k_1, \ldots, k_n$ respectively, we have that $\star(\varphi_1, \ldots, \varphi_n) \in \mathcal{L}_{\mathsf{C}}$ and $\star(\varphi_1, \ldots, \varphi_n)$ is of *type* k.

Elements of $\mathcal{L}_{\mathsf{C}}$ are called *atomic formulas* and are denoted $\varphi, \psi, \alpha, \ldots$ The *type of a formula* $\varphi \in \mathcal{L}_{\mathsf{C}}$ is denoted $k(\varphi)$.

The *skeleton syntactic tree* of a formula $\varphi \in \mathcal{L}_{\mathsf{C}}$ is the syntactic tree of the formula φ in which the nodes labeled with subformulas of φ are replaced by the skeleton of their outermost connective.

A set of atomic connectives C is *plain* if for all $\star \in \mathsf{C}$ of skeleton $(\sigma, \pm, Æ, (k, k_1, \ldots, k_n), (\pm_1, \ldots, \pm_n))$ there are atoms $p_1, \ldots, p_n \in \mathbb{P}$ of types $k_1, \ldots, k_n$ respectively. *In the sequel, we assume that all sets of connectives* C *are plain.* $\square$

Our assumption that all sets of connectives C considered are plain makes sense. Indeed, we want all connectives of C to appear in some formula of $\mathcal{L}_{\mathsf{C}}$. If C was not plain then there would be a connective of C (with input type k) which would be necessarily composed with another connective of C (of output type k), if we want such a connective to appear in a formula of $\mathcal{L}_{\mathsf{C}}$. Yet, in that case, we should instead view C as a set of *molecular* connectives (introduced in the next section).

Definition 4 (C–models)**.** Let C be a set of atomic connectives. A C*–model* is a tuple $M = (W, \mathcal{R})$ where W is a non-empty set and $\mathcal{R}$ is a set of relations over W such that each n–ary connective $\star \in \mathsf{C}$ of type signature $(k, k_1, \ldots, k_n)$ is associated to a $k_1 + \ldots + k_n + k$–ary relation $R_\star \in \mathcal{R}$.

An *assignment* is a tuple $(w_1, \ldots, w_k) \in W^k$ for some $k \in \mathbb{N}^*$, generally denoted $\overline{w}$. The set of assignments of a C–model M is denoted $\overline{w}(M, \mathsf{C})$. A *pointed* C*–model* $(M, \overline{w})$ is a C–model M together with an assignment $\overline{w}$. In that case, we say that $(M, \overline{w})$ *is of type* k. The class of all pointed C–models is denoted $\mathcal{M}_{\mathsf{C}}$. $\square$

Note that a C–model can be canonically seen as a (first-order) structure, for some appropriate set of predicates $\mathcal{P}$ associated to the relations of $\mathcal{R}$.

Definition 5 (Atomic logics)**.** Let C be a set of atomic connectives and let $M = (W, \mathcal{R})$ be a C–model. We define the *interpretation function of* $\mathcal{L}_{\mathsf{C}}$ *in* M, denoted $[\![\cdot]\!]^M : \mathcal{L}_{\mathsf{C}} \to \bigcup_{k \in \mathbb{N}^*} W^k$,

Permutations of $\mathfrak{S}_2$	unary signatures
$\tau_1 = (1,2)$	$t_1 = (\exists, (1,1), +)$
$\tau_2 = (2,1)$	$t_2 = (\forall, (1,1), +)$
	$t_3 = (\forall, (1,1), -)$
	$t_4 = (\exists, (1,1), -)$
Permutations of $\mathfrak{S}_3$	**binary signatures**
$\sigma_1 = (1,2,3)$	$s_1 = (\exists, (1,1,1), (+,+))$
$\sigma_2 = (3,2,1)$	$s_2 = (\forall, (1,1,1), (+,-))$
$\sigma_3 = (3,1,2)$	$s_3 = (\forall, (1,1,1), (-,+))$
$\sigma_4 = (2,1,3)$	$s_4 = (\forall, (1,1,1), (+,+))$
$\sigma_5 = (2,3,1)$	$s_5 = (\exists, (1,1,1), (+,-))$
$\sigma_6 = (1,3,2)$	$s_6 = (\exists, (1,1,1), (-,+))$
	$s_7 = (\exists, (1,1,1), (-,-))$
	$s_8 = (\forall, (1,1,1), (-,-))$

Figure 1: Permutations of $\mathfrak{S}_2$ and $\mathfrak{S}_3$ and 'families' of unary and binary signatures

inductively as follows: for all propositional letters $p \in \mathsf{C}$ of type k, all connectives $\star \in \mathsf{C}$ of skeleton $(\sigma, \pm, Æ, (k, k_1, \ldots, k_n), (\pm_1, \ldots, \pm_n))$ of arity $n > 0$, for all $\varphi_1, \ldots, \varphi_n \in \mathcal{L}_{\mathsf{C}}$,

$$\begin{aligned} [\![p]\!]^M &\triangleq \begin{cases} R_p & \text{if } \pm = + \\ W^k - R_p & \text{if } \pm = - \end{cases} \\ [\![\star(\varphi_1, \ldots, \varphi_n)]\!]^M &\triangleq f_\star([\![\varphi_1]\!]^M, \ldots, [\![\varphi_n]\!]^M) \end{aligned}$$

where the function $f_\star$ is defined as follows: for all $W_1 \in \mathcal{P}(W^{k_1}), \ldots, W_n \in \mathcal{P}(W^{k_n})$, $f_\star(W_1, \ldots, W_n) \triangleq \{\overline{w}_{n+1} \in W^k \mid \mathcal{C}^\star(W_1, \ldots, W_n, \overline{w}_{n+1})\}$ where $\mathcal{C}^\star(W_1, \ldots, W_n, \overline{w}_{n+1})$ is called the *truth condition* of $\star$ and is defined as follows:

- if $Æ = \forall$: "$\forall \overline{w}_1 \in W^{k_1} \ldots \overline{w}_n \in W^{k_n} \left(\overline{w}_1 \pitchfork_1 W_1 \vee \ldots \vee \overline{w}_n \pitchfork_n W_n \vee R_\star^{\pm\sigma} \overline{w}_1 \ldots \overline{w}_n \overline{w}_{n+1}\right)$";
- if $Æ = \exists$: "$\exists \overline{w}_1 \in W^{k_1} \ldots \overline{w}_n \in W^{k_n} \left(\overline{w}_1 \pitchfork_1 W_1 \wedge \ldots \wedge \overline{w}_n \pitchfork_n W_n \wedge R_\star^{\pm\sigma} \overline{w}_1 \ldots \overline{w}_n \overline{w}_{n+1}\right)$";

where, for all $j \in [\![1; n]\!]$, $\overline{w}_j \pitchfork_j W_j \triangleq \begin{cases} \overline{w}_j \in W_j & \text{if } \pm_j = + \\ \overline{w}_j \notin W_j & \text{if } \pm_j = - \end{cases}$ and

$R_\star^{\pm\sigma} \overline{w}_1 \ldots \overline{w}_{n+1}$ holds iff $\pm R_\star \overline{w}_{\sigma^-(1)} \ldots \overline{w}_{\sigma^-(n+1)}$ holds, with the notations $+R_\star \triangleq R_\star$ and $-R_\star \triangleq W^{k+k_1+\ldots+k_n} - R_\star$. If $\mathcal{E}_{\mathsf{C}}$ is a class of pointed C–models, the *satisfaction relation* $\Vdash \subseteq \mathcal{E}_{\mathsf{C}} \times \mathcal{L}_{\mathsf{C}}$ is defined as follows: for all $\varphi \in \mathcal{L}_{\mathsf{C}}$ and all $(M, \overline{w}) \in \mathcal{E}_{\mathsf{C}}$, $((M, \overline{w}), \varphi) \in \Vdash$ iff $\overline{w} \in [\![\varphi]\!]^M$. We usually write $(M, \overline{w}) \Vdash \varphi$ instead of $((M, \overline{w}), \varphi) \in \Vdash$ and we say that φ is *true* in $(M, \overline{w})$.

The logic $(\mathcal{L}_{\mathsf{C}}, \mathcal{E}_{\mathsf{C}}, \Vdash)$ is the *atomic logic associated to $\mathcal{E}_{\mathsf{C}}$ and C*. The logics of the form $(\mathcal{L}_{\mathsf{C}}, \mathcal{M}_{\mathsf{C}}, \Vdash)$ are called *basic atomic logics*. □

We stress that the $\pm$ sign in $R_\star^{\pm\sigma}$ is the $\pm$ sign in $(\sigma, \pm, Æ, (k, k_1, \ldots, k_n), (\pm_1, \ldots, \pm_n))$.

Example 1 (Modal logic)**.** An example of atomic logic is modal logic where $\mathsf{C} = \{p, \top, \bot, \wedge, \vee, \Diamond_j, \Box_j \mid j \in \mathbb{I}\}$ is such that

- $\top, \bot$ are connectives of skeletons $(\mathrm{Id}, +, \exists, 1)$ and $(\mathrm{Id}, -, \forall, 1)$ respectively;

- $\wedge, \vee, \Diamond_j, \Box_j$ are connectives of skeletons $(\sigma_1, +, s_1)$, $(\sigma_1, -, s_4)$, $(\tau_2, +, t_1)$ and $(\tau_2, -, t_2)$ respectively;
- the C-models $M = (W, \mathcal{R}) \in \mathcal{E}_{\mathsf{C}}$ are such that $R_\wedge = R_\vee = \{(w, w, w) \mid w \in W\}$, $R_{\Diamond_j} = R_{\Box_j}$ and $R_\top = R_\perp = W$.

With these conditions on the C–models of $\mathcal{E}_{\mathsf{C}}$, for all $(M, w) \in \mathcal{E}_{\mathsf{C}}$,

$$\begin{array}{lll} w \in & [\![\Diamond_j \varphi]\!]^M & \text{iff} \quad \exists v(v \in [\![\varphi]\!]^M \wedge R_{\Diamond_j} wv) \\ w \in & [\![\Box_j \varphi]\!]^M & \text{iff} \quad \forall v(v \in [\![\varphi]\!]^M \vee -R_{\Box_j} wv) \\ w \in & [\![\wedge(\varphi, \psi)]\!]^M & \text{iff} \quad \exists vu(v \in [\![\varphi]\!]^M \wedge u \in [\![\psi]\!]^M \wedge R_\wedge vuw) \\ & & \text{iff} \quad w \in [\![\varphi]\!]^M \wedge w \in [\![\psi]\!]^M \\ w \in & [\![\vee(\varphi, \psi)]\!]^M & \text{iff} \quad \forall vu(v \in [\![\varphi]\!]^M \vee u \in [\![\psi]\!]^M \vee -R_\vee vuw) \\ & & \text{iff} \quad w \in [\![\varphi]\!]^M \vee w \in [\![\psi]\!]^M \end{array}$$

Other examples are given in Figure 2 as well as in [4, 6, 5]. □

6 Molecular Logics

Molecular logics are basically logics whose primitive connectives are compositions of atomic connectives in which it is possible to repeat the same argument at different places in the connective. That is why we call them 'molecular', just as molecules are compositions of atoms in chemistry.

Definition 6 (Molecular skeleton and connective)**.** The class $\mathbb{C}^*$ of *molecular skeletons* is the smallest set such that:

- $\mathbb{P} \subseteq \mathbb{C}^*$ and $\mathbb{C}^*$ contains as well, for each $k, l \in \mathbb{N}^*$, a symbol id_k^l of *type signature* (k, k), *output type* k and *arity* 1;
- for all atomic skeleton $\star \in \mathbb{C}$ of type signature $(k, k_1^0, \ldots, k_n^0)$ and all $c_1, \ldots, c_n \in \mathbb{C}^*$ of output types or types (if they are propositional letters) $k_1^0, \ldots, k_n^0$ respectively, $c \triangleq \star(c_1, \ldots, c_n)$ is a molecular skeleton of $\mathbb{C}^*$ of *output type* k.

If $c \in \mathbb{C}^*$, we define its *decomposition tree* as follows. If $c = p \in \mathbb{P}$ or $c = id_k^l$, then its decomposition tree T_c is the tree consisting of a single node labeled with p or id_k^l respectively. If $c = \star(c_1, \ldots, c_n) \in \mathbb{C}^*$ then its decomposition tree T_c is the tree defined inductively as follows: the root of T_c is c and it is labeled with $\star$ and one sets edges between that root and the roots $c_1, \ldots, c_n$ of the decomposition trees $T_{c_1}, \ldots, T_{c_n}$ respectively.

If $c \triangleq \star(c_1, \ldots, c_n)$ is a molecular skeleton with output type k and $k_1, \ldots, k_m$ are the ks of the different id_k^ls which appear in $c_1, \ldots, c_n$ (in an order which follows the first appearance of the id_k^ls in the inorder traversal of the decomposition trees of $c_1, \ldots, c_n$), then the *type signature* of c is $(k, k_1, \ldots, k_m)$ and its *arity* is m. We also define the *quantification signature* $Æ(c)$ of $c = \star(c_1, \ldots, c_n)$ by $Æ(c) \triangleq Æ(\star)$.

A *molecular connective* is a symbol to which is associated a molecular skeleton. Its arity, type signature, output type, quantification signature and decomposition tree are the same as its skeleton.

The set of *atomic connectives associated to a set* C *of molecular connectives* is the set of labels different from id_k^l of the decomposition trees of the molecular connectives of C. □

Atomic connective	Truth condition	Connective in the literature
	The conjunction orbit	
$\varphi\ (\sigma_1, +, s_1)\ \psi$	$\exists vu\,(v \in [\![\varphi]\!] \wedge u \in [\![\psi]\!] \wedge Rvuw)$	$\varphi \circ \psi$ [20], $\varphi \otimes_3 \psi$ [3]
$\varphi\ (\sigma_2, -, s_2)\ \psi$	$\forall vu\,(v \in [\![\varphi]\!] \vee u \notin [\![\psi]\!] \vee -Rwuv)$	
$\varphi\ (\sigma_3, -, s_2)\ \psi$	$\forall vu\,(v \in [\![\varphi]\!] \vee u \notin [\![\psi]\!] \vee -Ruwv)$	$/$ [20], $\varphi \subset_2 \psi$ [3]
$\varphi\ (\sigma_4, +, s_1)\ \psi$	$\exists vu\,(v \in [\![\varphi]\!] \wedge u \in [\![\psi]\!] \wedge Ruvw)$	
$= \psi\ (\sigma_1, +, s_1)\ \varphi$		
$\varphi\ (\sigma_5, -, s_3)\ \psi$	$\forall vu\,(v \notin [\![\varphi]\!] \vee u \in [\![\psi]\!] \vee -Rwvu)$	$\backslash$ [20], $\varphi \supset_1 \psi$ [3]
$= \psi\ (\sigma_2, -, s_2)\ \varphi$		
$\varphi\ (\sigma_6, -, s_3)\ \psi$	$\forall vu\,(v \notin [\![\varphi]\!] \vee u \in [\![\psi]\!] \vee -Rvwu)$	
$= \psi\ (\sigma_3, -, s_2)\ \varphi$		
	The but–not orbit	
$\varphi\ (\sigma_1, +, s_5)\ \psi$	$\exists vu\,(v \in [\![\varphi]\!] \wedge u \notin [\![\psi]\!] \wedge Rvuw)$	$\varphi \prec_3 \psi$ [3]
$\varphi\ (\sigma_2, -, s_4)\ \psi$	$\forall vu\,(v \in [\![\varphi]\!] \vee u \in [\![\psi]\!] \vee -Rwuv)$	
$\varphi\ (\sigma_3, +, s_6)\ \psi$	$\exists vu\,(v \notin [\![\varphi]\!] \wedge u \in [\![\psi]\!] \wedge Ruwv)$	$\varphi \succ_2 \psi$ [3]
$\varphi\ (\sigma_4, +, s_6)\ \psi$	$\exists vu\,(v \notin [\![\varphi]\!] \wedge u \in [\![\psi]\!] \wedge Ruvw)$	$\varphi \oslash \psi$ [17, 22]
$= \psi\ (\sigma_1, +, s_5)\ \varphi$		
$\varphi\ (\sigma_5, -, s_4)\ \psi$	$\forall vu\,(v \in [\![\varphi]\!] \vee u \in [\![\psi]\!] \vee -Rwvu)$	$\varphi \oplus \psi$ [17, 22]
$= \psi\ (\sigma_2, -, s_4)\ \varphi$		$\varphi \oplus_1 \psi$ [3]
$\varphi\ (\sigma_6, +, s_5)\ \psi$	$\exists vu\,(v \in [\![\varphi]\!] \wedge u \notin [\![\psi]\!] \wedge Rvwu)$	$\varphi \oslash \psi$ [17, 22]
$= \psi\ (\sigma_3, +, s_6)\ \varphi$		
	The stroke orbit	
$\varphi\ (\sigma_1, +, s_7)\ \psi$	$\exists vu\,(v \notin [\![\varphi]\!] \wedge u \notin [\![\psi]\!] \wedge Rvuw)$	$\varphi \mid_3 \psi$ [2, 16]
$\varphi\ (\sigma_2, +, s_7)\ \psi$	$\exists vu\,(v \notin [\![\varphi]\!] \wedge u \notin [\![\psi]\!] \wedge Rwuv)$	
$\varphi\ (\sigma_3, +, s_7)\ \psi$	$\exists vu\,(v \notin [\![\varphi]\!] \wedge u \notin [\![\psi]\!] \wedge Ruwv)$	
$\varphi\ (\sigma_4, +, s_7)\ \psi$	$\exists vu\,(v \notin [\![\varphi]\!] \wedge u \notin [\![\psi]\!] \wedge Ruvw)$	
$= \psi\ (\sigma_1, +, s_7)\ \varphi$		
$\varphi\ (\sigma_5, +, s_7)\ \psi$	$\exists vu\,(v \notin [\![\varphi]\!] \wedge u \notin [\![\psi]\!] \wedge Rwvu)$	$\varphi \mid_1 \psi$ [2, 16]
$= \psi\ (\sigma_2, +, s_7)\ \varphi$		
$\varphi\ (\sigma_6, +, s_7)\ \psi$	$\exists vu\,(v \notin [\![\varphi]\!] \wedge u \notin [\![\psi]\!] \wedge Rvwu)$	$\varphi \mid_2 \psi$ [2, 16]
$= \psi\ (\sigma_3, +, s_7)\ \varphi$		

Figure 2: Some binary connectives of atomic logics of type $(1, 1, 1)$

Note that the same label (atomic connective) may appear several times in a decomposition tree. Note also that the vertices of a decomposition tree are molecular connectives.

Every atomic connective $\star$ of type signature $(k, k_1, \ldots, k_n)$ can be seen as the (specific) molecular connectives $c \triangleq \star(id^1_{k_1}, \ldots, id^n_{k_n})$. One needs to introduce the connective id^l_k in order to deal with molecular connectives whose skeletons are for example of the form $\star(p, id^l_k)$ where $p \in \mathbb{P}$ or with molecular connectives in which the same argument(s) appear at different places, like for example in $\star(id^1_k, \ldots, id^1_k)$ which is of arity 1.

Definition 7 (Molecular language)**.** Let C be a set of molecular connectives. The *(typed) molecular language* $\mathcal{L}_\mathsf{C}$ associated to C is the smallest set that contains the propositional letters and that is closed under the molecular connectives while respecting the type constraints. That is,

- the propositional letters of C belong to $\mathcal{L}_\mathsf{C}$;
- for all $c \in \mathsf{C}$ of type signature $(k, k_1, \ldots, k_m)$ and for all $\varphi_1, \ldots, \varphi_m \in \mathcal{L}_\mathsf{C}$ of types $k_1, \ldots, k_m$ respectively, we have that $c(\varphi_1, \ldots, \varphi_m) \in \mathcal{L}_\mathsf{C}$ and $c(\varphi_1, \ldots, \varphi_m)$ is of *type* k.

Elements of $\mathcal{L}_\mathsf{C}$ are called *molecular formulas* and are denoted $\varphi, \psi, \alpha, \ldots$ The *type of a formula* $\varphi \in \mathcal{L}_\mathsf{C}$ is denoted $k(\varphi)$. We use the same abbreviations as for the atomic language. □

Definition 8 (Molecular logic)**.** If C is a set of molecular connectives, then a *C–model* M is a C'–model M where C' is the set of atomic connectives associated to C. The truth conditions for molecular connectives are defined naturally from the truth conditions of atomic connectives. We define the *interpretation function of $\mathcal{L}_\mathsf{C}$ in M*, denoted $[\![\cdot]\!]^M : \mathcal{L}_\mathsf{C} \to \bigcup_{k \in \mathbb{N}^*} W^k$, inductively as follows: for all propositional letters $p \in \mathsf{C}$ of skeleton $(\sigma, \pm, Æ, k)$, all molecular connectives $\star(c_1, \ldots, c_n) \in \mathsf{C}$ of arity $m > 0$ and all $k, l \in \mathbb{N}^*$, for all $\varphi, \varphi_1, \ldots, \varphi_m \in \mathcal{L}_\mathsf{C}$,

$$
\begin{aligned}
[\![p]\!]^M &\triangleq \pm R_p \\
[\![id^l_k(\varphi)]\!]^M &\triangleq [\![\varphi]\!]^M \\
[\![\star(c_1, \ldots, c_n)\,(\varphi_1, \ldots, \varphi_m)]\!]^M &\triangleq f_\star\left([\![c_1(\varphi^1_1, \ldots, \varphi^1_{i_1})]\!]^M, \ldots, [\![c_n(\varphi^n_1, \ldots, \varphi^n_{i_n})]\!]^M\right)
\end{aligned}
$$

where for all $j \in \{1, \ldots, n\}$, the formulas $\varphi^j_1, \ldots, \varphi^j_{i_j}$ are those $\varphi_1, \ldots, \varphi_m$ for which there is a corresponding id^l_k in c_j (the φ^j_is appear in the same order as their corresponding id^l_ks in c_j).

If $\mathcal{E}_\mathsf{C}$ is a class of pointed C–models, the triple $(\mathcal{L}_\mathsf{C}, \mathcal{E}_\mathsf{C}, \Vdash)$ is a logic called the *molecular logic associated to $\mathcal{E}_\mathsf{C}$ and C.* □

As one can easily notice, every atomic logic can be canonically mapped to an equi-expressive molecular logic: each atomic connective $\star$ of type signature $(k, k_1, \ldots, k_n)$ of the given atomic logic has to be transformed into the molecular connective of skeleton $\star(id^1_{k_1}, \ldots, id^n_{k_n})$. Note that the id^l_k are in fact specific atomic connectives whose associated relations are the identity relations.

Examples of specific molecular logics can be found in [6, 5].

7 Boolean Atomic and Molecular Logics

Atomic and molecular logics do not include Boolean connectives as primitive connectives. In fact, they can be defined in terms of specific atomic connectives, as follows.

Definition 9 (Boolean connectives)**.** The *Boolean connectives* called *conjunctions*, *disjunctions*, *negations* and *Boolean constants* (of type k) are the atomic connectives denoted, respectively:

$$\mathbb{B} \triangleq \{\wedge_k, \vee_k, \neg_k, \top_k, \bot_k \mid k \in \mathbb{N}^*\}$$

The skeleton of $\wedge_k$ is $(\mathrm{Id}, +, \exists, (k, k, k), (+, +))$, the skeleton of $\vee_k$ is $(\mathrm{Id}, -, \forall, (k, k, k), (+, +))$, the skeleton of $\neg_k$ is $(\mathrm{Id}, +, \exists, (k, k), -)$, the skeleton of $\top_k$ is $(\mathrm{Id}, +, \exists, k)$ and the skeleton of $\bot_k$ is $(\mathrm{Id}, -, \forall, k)$.

In any C-model $M = (W, \mathcal{R})$ containing Boolean connectives, the associated relation of any $\vee_k$ or $\wedge_k$ is $R_{\wedge_k} = R_{\vee_k} \triangleq \{(\overline{w}, \overline{w}, \overline{w}) \mid \overline{w} \in W^k\}$, the associated relation of any $\neg_k$ is $R_{\neg_k} \triangleq \{(\overline{w}, \overline{w}) \mid \overline{w} \in W^k\}$ and the associated relation of any $\top_k$ or $\bot_k$ is $R_{\bot_k} = R_{\top_k} \triangleq W^k$.

Atomic or molecular logics containing Boolean connectives are called *Boolean atomic or molecular logics*. We say that a set of atomic connectives C *is complete for conjunctions and disjunctions* when it contains all conjunctions, disjunctions and constants $\wedge_k, \vee_k, \top_k, \bot_k$, for k ranging over all input types and output types of the atomic connectives of C. We say that a set of atomic connectives C *is complete for Boolean connectives* when it contains all conjunctions, disjunctions, constants as well as negations $\wedge_k, \vee_k, \top_k, \bot_k, \neg_k$, for k ranging over all input types and output types of the atomic connectives of C. □

Proposition 1. *Let C be a set of atomic connectives containing Boolean connectives. and let $M = (W, \mathcal{R})$ be a C-model. Then, for all $k \in \mathbb{N}^*$, all $\varphi, \psi \in \mathcal{L}_{\mathsf{C}}$, if $k(\varphi) = k(\psi) = k$, then*

$$\begin{array}{rcl} [\![\top_k]\!]^M & \triangleq & W^k \\ [\![\bot_k]\!]^M & \triangleq & \emptyset \\ [\![\neg_k \varphi]\!]^M & \triangleq & W^k - [\![\varphi]\!]^M \\ [\![(\varphi \wedge_k \psi)]\!]^M & \triangleq & [\![\varphi]\!]^M \cap [\![\psi]\!]^M \\ [\![(\varphi \vee_k \psi)]\!]^M & \triangleq & [\![\varphi]\!]^M \cup [\![\psi]\!]^M. \end{array}$$

It turns out that Boolean negation can also be simulated systematically at the level of atomic connectives by applying a transformation on them. The Boolean negation of a formula then boils down to taking the Boolean negation of the outermost connective of the formula. This transformation is defined as follows.

Definition 10 (Boolean negation)**.** Let $\star$ be a n–ary connective of skeleton $(\sigma, \pm, Æ, \overline{k}, \pm_1, \ldots, \pm_n)$. The *Boolean negation of* $\star$ is the connective $-\star$ of skeleton $(\sigma, -\pm, -Æ, \overline{k}, -\pm_1, \ldots, -\pm_n)$ where $-Æ \triangleq \exists$ if $Æ = \forall$ and $-Æ \triangleq \forall$ otherwise, which is associated in any C–model to the same relation as $\star$. If $\varphi = \star(\varphi_1, \ldots, \varphi_n)$ is an atomic formula, the *Boolean negation of* φ is the formula $-\varphi \triangleq - \star (\varphi_1, \ldots, \varphi_n)$. □

Proposition 2 ([6])**.** *Let C be a set of atomic connectives such that $-\star \in \mathsf{C}$ for all $\star \in \mathsf{C}$. Let $\varphi \in \mathcal{L}_{\mathsf{C}}$ of type k and let $M = (W, \mathcal{R})$ be a C–model. Then, for all $\overline{w} \in W^k$, $\overline{w} \in [\![-\varphi]\!]^M$ iff $\overline{w} \notin [\![\varphi]\!]^M$.*

8 Universal and Existential Molecular Connectives

Universal and existential molecular connectives are essentially molecular connectives such that the quantification patterns of the quantification signatures of their successive atomic connectives are of the form $\forall \ldots \forall$ or $\exists \ldots \exists$ respectively. So, they essentially behave as 'macroscopic' atomic connectives of quantification signatures $\forall$ or $\exists$.

Definition 11 (Universal and existential molecular connective)**.** A *universal (resp. existential) molecular skeleton* is a molecular skeleton c different from any id^l_k for any $k, l \in \mathbb{N}^*$ such that $Æ(c) = \forall$ (resp. $Æ(c) = \exists$) and such that for each node of its decomposition tree labeled with $\star = (\sigma, \pm, Æ, \overline{k}, (\pm_1, \ldots, \pm_n))$ and each of its jth children labeled with some $\star_j \in \mathbb{C}$ such that the subtree generated by this j^{th} children contains at least one id^l_k, we have that $Æ(\star_j) = \pm_j Æ$. A *universal (resp. existential) molecular connective* is a molecular connective with a universal (resp. existential) skeleton. □

Example 2. On the one hand, the molecular connective $\star(p, id^l_k)$ is a universal (resp. existential) molecular connective if $Æ(\star) = \forall$ (resp. $Æ(\star) = \exists$). Likewise, $\supset (id^1_1, \Box id^2_1)$ and $\otimes(\Diamond id^1_1, p)$ are universal and existential molecular connectives respectively. On the other hand, the molecular connectives $\Box\Diamond^- id^1_1$ and $\supset (\Box id^1_1, \Box id^2_1)$ are neither universal nor existential molecular connectives. □

Just as we have tonicity signatures for atomic connectives, we can also define an adaptation of this notion for universal and existential molecular connectives, which, we repeat, are some sort of 'macroscopic' atomic connectives.

Definition 12 (Tonicity signature of a molecular connective)**.** Let c be a molecular connective and let c' be a molecular subconnective of c. We define the *tonicity of c' w.r.t. c*, denoted $tn(c', c)$ inductively as follows. If $c = c'$ then $tn(c', c) = +$. Otherwise, if $c = \star(c_1, \ldots, c_n)$ with $\star = (\sigma, \pm, Æ, \overline{k}, (\pm_1, \ldots, \pm_n))$ and c' appears in c_j then $tn(c', c) = \pm_j tn(c', c_j)$. The *tonicity signature* of a molecular connective is the tuple $(\pm_1, \ldots, \pm_l)$ of the tonicities $tn(id^i_k, c)$ of the connectives labeling the leafs of the decomposition tree of c of the form id^i_k (possibly with repetition). □

A molecular connective c is generally represented as $\star(c_1, \ldots, c_n)$ where $\star$ is an *atomic* connective, but it can also be represented as $c_0(c_1, \ldots, c_n)$ where c_0 is a *molecular* connective such that the id^i_ks which appear in c_0 are all distincts. This representation and decomposition in terms of molecular connectives is equivalent to the original one. For example the modal connective of weakly agregative modal logic $\star(id^1_1, \ldots, id^1_1)$ can be represented in terms of molecular connectives with distinct id^i_ks as follows: $\star(id^1_1, \ldots, id^n_1)(id^{n+1}_1, \ldots, id^{n+1}_1)$. We now formalize this idea of decomposition into molecular connectives.

Definition 13 (Decomposition of a molecular connective)**.** Let c be a molecular connective and let T_c be its decomposition tree. A *decomposition* of c is an expression of the form $c_0(c_1, \ldots, c_n)$ where c_0 is a molecular connective whose decomposition tree is a subtree of T_c with root c and whose leafs, corresponding to the nodes $c_1, \ldots, c_n$ of T_c, have been replaced by distinct id^i_ks and where $c_1, \ldots, c_n$ are molecular connectives whose decomposition trees are the subtrees of T_c generated by the nodes $c_1, \ldots, c_n$ of T_c. □

Example 3. For any molecular connective $\star(id^1_1, \ldots, id^n_1)$ corresponding to an atomic connective $\star$ of arity n, a decomposition is $\star(id^{n+1}_1, \ldots, id^{2n}_1)(id^1_1, \ldots, id^n_1)$. Likewise, a decomposition of $\star(p, id^l_k)$ is $\star(id^1_1, id^2_k)(p, id^l_k)$ (if p is of type 1). □

Based on this observation, we can decompose molecular connectives into an alternation of universal and existential molecular subconnectives. That is what the following definition captures, at least at the first level of alternation depth. In that definition, if c_0 is universal and the tonicity of the jth molecular connective c_j is positive for example, then c_j will have an existential quantification signature (and its 'head' will behave as an existential molecular connective).

Definition 14 (Maximal decomposition)**.** A decomposition $c = c_0(c_1, \ldots, c_n)$ of a molecular connective c is *maximal* when c_0 is a universal or existential molecular connective of tonicity

signature $(\pm_1, \ldots, \pm_n)$ with the id_k^is all distinct, and we have that $Æ(c_j) = - \pm_j Æ(c_0)$ for all $j \in [\![1; n]\!]$ such that c_j is not of the form id_k^l for some $k, l \in \mathbb{N}^*$. □

Example 4. The decomposition of $\star(p, id_k^l)$ in $\star(id_1^1, id_k^2)(p, id_k^l)$ (if p is of type 1) is not maximal because $\star(p, id_k^l)$ is already a universal or existential molecular connective. On the other hand, the decomposition of $\star(p, id_k^l)$ in $\star(p, id_k^l)(id_k^{l'})$ is maximal. □

The alternation of universal and existential molecular subconnectives inherent to any molecular connective is fully captured by the following notion of *quantified decomposition tree*. It is an abstraction of the notion of decomposition tree of Definition 6 which considers as first-class citizens universal and existential molecular (sub)connectives.

Definition 15 (Quantified decomposition tree)**.** If $c \in \mathbb{C}^*$ is a molecular skeleton, we define its *quantified decomposition tree* T'_c inductively as follows. If c is a propositional letter or id_k^l for some $k, l \in \mathbb{N}^*$ then its quantified decomposition tree T'_c is the tree consisting of a single node labeled with c. Otherwise, c admits a maximal decomposition $c = c_0(c_1, \ldots, c_n)$. Then, its quantified decomposition tree T'_c is the tree defined inductively as follows: the root of T'_c is c and it is labeled with c_0 and one sets edges between that root and the roots $c_1, \ldots, c_n$ of the quantified decomposition trees $T'_{c_1}, \ldots, T'_{c_n}$ respectively. The quantified decomposition tree of a molecular connective is the quantified decomposition tree of its skeleton. □

Note that a propositional letter p can occur as the label of a node only if the quantified decomposition tree in which it appears consists of this single node only (Example 4 illustrates this phenomenon).

Definition 16 (Relation associated to a molecular connective)**.** Let C be a set of molecular connectives and let $c = \star(c_1, \ldots, c_n) \in \mathsf{C}$ be a molecular connective with $\star$ of skeleton $(\sigma, \pm, Æ, (k^\star, k_1^\star, \ldots, k_n^\star), (\pm_1^\star, \ldots, \pm_n^\star))$. Assume that the decomposition tree of c has l different leaves labeled with id_k^is of output types $k_1, \ldots, k_l$, in that order. Then, the arity of c is l.

Let $M = (W, \mathcal{R})$ be a C–model. We define the relation R_c inductively as follows. First, we set $R_{-\star} = R_\star$ if $\star$ is an atomic connective and if $\pm \in \{+, -\}$,

$$\pm p \triangleq \begin{cases} p & \text{if } \pm = + \\ -p & \text{if } \pm = - \end{cases} \text{ and } \pm \star (c_1, \ldots, c_n) = \begin{cases} \star(c_1, \ldots, c_n) & \text{if } \pm = + \\ - \star (c_1, \ldots, c_n) & \text{if } \pm = - \end{cases}.$$

- If the arity l of c is 0 then $R_c = [\![c]\!]^M$;
- If the arity l of c is greater than 0 and $Æ = \exists$ then for all $\overline{w} \in W^k, \overline{w}_1 \in W^{k_1}, \ldots, \overline{w}_l \in W^{k_l}$, we set

  $$R_c \overline{w}_1 \ldots \overline{w}_l \overline{w} \text{ iff } \exists \overline{v}_1 \in W^{k_1^\star}, \ldots, \overline{v}_n \in W^{k_n^\star}$$
  $$\big(R_\star^{\pm\sigma} \overline{v}_1 \ldots \overline{v}_n \overline{w} \wedge R_{\pm_1^\star c_1} \overline{w}_1^1 \ldots \overline{w}_{i_1}^1 \overline{v}_1 \wedge \ldots \wedge R_{\pm_n^\star c_n} \overline{w}_1^n \ldots \overline{w}_{i_n}^n \overline{v}_n\big)$$

  where, for all $j \in [\![1; n]\!]$, the tuples $\overline{w}_1^j, \ldots, \overline{w}_{i_j}^j$ are those associated to the leafs of the decomposition tree of c which are also leafs in the decomposition (sub)tree of c_j (the $\overline{w}_i^j$s appear in the same order as their corresponding leafs in the decomposition tree of c_j). Moreover, we set $R_{id_k^i} \overline{v}\overline{w}$ iff $R_{-id_k^i} \overline{v}\overline{w}$ iff $\overline{v} = \overline{w}$.
- If the arity l of c is greater than 0 and $Æ = \forall$ then for all $\overline{w} \in W^k, \overline{w}_1 \in W^{k_1}, \ldots, \overline{w}_l \in W^{k_l}$, we set

$R_c\overline{w}_1 \dots \overline{w}_l\overline{w}$ iff $\forall \overline{v}_1 \in W^{k^\star_1}, \dots, \overline{v}_n \in W^{k^\star_n}$
$\big(R_\star^{\pm\sigma}\overline{v}_1 \dots \overline{v}_n\overline{w} \vee R_{\pm^\star_1 c_1}\overline{w}^1_1 \dots \overline{w}^1_{i_1}\overline{v}_1 \vee \dots \vee R_{\pm^\star_n c_n}\overline{w}^n_1 \dots \overline{w}^n_{i_n}\overline{v}_n\big)$

where, for all $j \in [\![1;n]\!]$, the tuples $\overline{w}^j_1, \dots, \overline{w}^j_{i_j}$ are defined as above. However, for that universal case, we set $R_{id^i_k}\overline{vw}$ iff $R_{-id^i_k}\overline{vw}$ iff $\overline{v} \neq \overline{w}$. □

Unsurprisingly, the semantics of universal and existential molecular connectives is similar to the semantics of atomic connectives of quantification signature $\forall$ and $\exists$ respectively. That is what the following proposition shows. In this proposition and the definition above, the molecular connective $c = \star(c_1, \dots, c_n)$ yields three different values: n, m and l. They can be all different in general. The value m is the arity of c, n is the arity of $\star$ and l is the number of leaves of the decomposition tree T_c of c. The value m is also the number of different labels id^i_k of the leafs of the decomposition tree T_c. Hence, the size l of the tuple of the tonicity signature of a molecular connective c is larger than the arity m of c because in the former case we take into account repetitions of the same id^i_ks.

Proposition 3. *Let C be a set of molecular connectives and let $c = \star(c_1, \dots, c_n) \in \mathsf{C}$ be a molecular connective of type signature $(k, k_1, \dots, k_m)$ and tonicity signature $(\pm_1, \dots, \pm_l)$ with $\star = (\sigma, \pm, Æ, \overline{k}, (\pm^\star_1, \dots, \pm^\star_n))$. Assume that the decomposition tree of c has $l > 0$ different leaves labeled by some id^i_k. Let $M = (W, \mathcal{R})$ be a C–model and let $\overline{w} \in W^k$.*

- *If c is an existential molecular connective then*

$$\overline{w} \in [\![c(\varphi_1, \dots, \varphi_m)]\!]^M \text{ iff } \exists \overline{w}_1 \dots \overline{w}_l \big(\overline{w}_1 \pitchfork_1 [\![\varphi_{i_1}]\!]^M \wedge \dots \wedge \overline{w}_l \pitchfork_l [\![\varphi_{i_l}]\!]^M \wedge R_c\overline{w}_1 \dots \overline{w}_l\overline{w}\big)$$

- *If c is a universal molecular connective then*

$$\overline{w} \in [\![c(\varphi_1, \dots, \varphi_m)]\!]^M \text{ iff } \forall \overline{w}_1 \dots \overline{w}_l \big(\overline{w}_1 \pitchfork_1 [\![\varphi_{i_1}]\!]^M \vee \dots \vee \overline{w}_l \pitchfork_l [\![\varphi_{i_l}]\!]^M \vee R_c\overline{w}_1 \dots \overline{w}_l\overline{w}\big)$$

where the R_cs are defined in Definition 16 and for all $j \in [\![1;l]\!]$,
$\overline{w}_j \pitchfork_j [\![\varphi_{i_j}]\!]^M \triangleq \begin{cases} \overline{w}_j \in [\![\varphi_{i_j}]\!]^M & \text{if } \pm_j = + \\ \overline{w}_j \notin [\![\varphi_{i_j}]\!]^M & \text{if } \pm_j = - \end{cases}$ *and $i_1, \dots, i_l \in [\![1;m]\!]$ are those indices corresponding to the m different id^i_ks appearing in c (we basically map the l leafs of the decomposition tree of c to their labels id^i_ks in this tree).*

9 Automatic Bisimulations for Atomic and Molecular Logics

In this section, we are going to see that notions of bisimulations can be automatically defined for atomic and molecular logics on the basis of the definition of the truth conditions of their connectives, not only for plain atomic logics but also for molecular logics. These notions are such that they preserve the truth of the formulas of the atomic logic considered between models.

9.1 Atomic Logics

Definition 17 (C–bisimulation for atomic connectives)**.** Let C be a set of atomic connectives, let $\star \in \mathsf{C}$ and let $M_1 = (W_1, \mathcal{R}_1)$ and $M_2 = (W_2, \mathcal{R}_2)$ be two C–models. A binary relation

$Z \subseteq \bigcup_{k \in \mathbb{N}^*} (W_1^k \times W_2^k) \cup (W_2^k \times W_1^k)$ is a C*–bisimulation* between M_1 and M_2 when it is non-empty and for all $\star \in \mathsf{C}$, if $\{M, M'\} = \{M_1, M_2\}$, then for all $\overline{w}_1, \ldots, \overline{w}_n, \overline{w'}_1, \ldots, \overline{w'}_n, \overline{w}, \overline{w'} \in \overline{w}(M, \mathsf{C}) \cup \overline{w}(M', \mathsf{C})$,

1. if $\star$ is a propositional letter p then, if $\overline{w} Z \overline{w'}$ and $\overline{w} \in [\![p]\!]$ then $\overline{w'} \in [\![p]\!]$;
2. if $\star$ has skeleton $(\sigma, \pm, \exists, \overline{k}, (\pm_1, \ldots, \pm_n))$ and we have $\overline{w} Z \overline{w'}$ and $R_\star^{\pm\sigma} \overline{w}_1 \ldots \overline{w}_n \overline{w}$, then $\exists \overline{w'}_1, \ldots, \overline{w'}_n (\overline{w}_1 \bowtie \overline{w'}_1 \wedge \overline{w}_2 \bowtie \overline{w'}_2 \wedge \ldots \wedge \overline{w}_n \bowtie \overline{w'}_n \wedge R_\star^{'\pm\sigma} \overline{w'}_1 \ldots \overline{w'}_n \overline{w'})$;
3. if $\star$ has skeleton $(\sigma, \pm, \forall, \overline{k}, (\pm_1, \ldots, \pm_n))$ and we have $\overline{w} Z \overline{w'}$ and $-R_\star^{'\pm\sigma} \overline{w'}_1 \ldots \overline{w'}_n \overline{w'}$, then $\exists \overline{w}_1, \ldots, \overline{w}_n (\overline{w}_1 \bowtie \overline{w'}_1 \wedge \overline{w}_2 \bowtie \overline{w'}_2 \wedge \ldots \wedge \overline{w}_n \bowtie \overline{w'}_n \wedge -R_\star^{\pm\sigma} \overline{w}_1 \ldots \overline{w}_n \overline{w})$;

where, for all $j \in [\![1; n]\!]$, we define $\overline{w}_j \bowtie \overline{w'}_j \triangleq \begin{cases} \overline{w}_j Z \overline{w'}_j & \text{if } \pm_j = + \\ \overline{w'}_j Z \overline{w}_j & \text{if } \pm_j = - \end{cases}$.

When such a C–bisimulation Z exists and $\overline{w} Z \overline{w'}$, we say that $(M, \overline{w})$ and $(M', \overline{w'})$ are C*–bisimilar* and we write it $(M, \overline{w}) \rightarrow_{\mathsf{C}} (M', \overline{w'})$. □

Note that case 1. is a particular instance of cases 2. and 3. with $n = 0$. Importantly, note also that the clause(s) defining a C-bisimulation allow us to define back the atomic connective(s) that led to their definition(s): we can recover the skeleton of the connectives considered (tonicity signature, quantification signature, type signature, *etc*) from the mere expression of the clauses of a given C-bisimulation. Hence, in that sense, a C–bisimulation completely characterizes a given (atomic) logic.

Example 5 (Modal logic)**.** Let us consider the connectives of modal logic: $\mathsf{C} = \{p, \neg p, \wedge, \vee, \Diamond, \Box\}$ where p has skeleton $(\mathrm{Id}, +, \exists, 1)$, $\neg p$ has skeleton $(\mathrm{Id}, -, \forall, 1)$, $\Diamond$ has skeleton $(\tau_2, +, t_1)$ and $\Box$ has skeleton $(\tau_2, -, t_2)$. Let $M_1 = (W_1, \{R_1, P_1\})$ and $M_2 = (W_2, \{R_2, P_2\})$ be two Kripke models (they are also C-models). A binary relation Z between M_1 and M_2 is a C–bisimulation between M_1 and M_2 when for all $M, M' \in \{M_1, M_2\}$ with $M = (W, \{R, P\})$ and $M' = (W', \{R', P'\})$, all $w, v \in M$ and all $w', v' \in M'$,

- if wZw' and $w \in [\![p]\!]$ then $w' \in [\![p]\!]$ (condition for p);
- if wZw' and $w' \in [\![p]\!]$ then $w \in [\![p]\!]$ (condition for $\neg p$);
- if wZw' and Rwv then there is $v' \in W'$ such that vZv' and $R'w'v'$ (condition for $\Diamond = (\tau_2, +, t_1)$);
- if wZw' and $R'w'v'$ then there is $v \in W$ such that vZv' and Rwv (condition for $\Box = (\tau_2, -, t_2)$).

Note that every C–bisimulation can be canonically extended into a *symmetric* C–bisimulation: one sets $w'Zw$ when wZw' already holds. □

Proposition 4. *Let $\mathsf{C} = \{p, \neg p, \wedge, \vee, \Diamond, \Box\}$ be the connectives of Example 5 and let M and M' be two C–models. Then, a C–bisimulation between M and M' is a modal bisimulation between M and M' and vice versa.*

Definition 18. Let C be a set of atomic connectives. Let $(M, \overline{w})$ and $(M', \overline{w'})$ be two pointed C–models. We write $(M, \overline{w}) \rightsquigarrow_{\mathsf{C}} (M', \overline{w'})$ when for all $\varphi \in \mathcal{L}_{\mathsf{C}}$, $(M, \overline{w}) \Vdash \varphi$ implies $(M', \overline{w'}) \Vdash \varphi$. □

Proposition 5 (Truth preservation of C-bisimilar pointed C-models)**.** *Let C be a set of atomic connectives and let $M_1 = (W_1, \mathcal{R}_1)$ and $M_2 = (W_2, \mathcal{R}_2)$ be two C–models. Let Z be a C–bisimulation between M_1 and M_2. Then, if $\{M, M'\} = \{M_1, M_2\}$ then for all $\overline{w} \in \overline{w}(M, \mathsf{C})$, all $\overline{w'} \in \overline{w}(M', \mathsf{C})$, if $\overline{w} Z \overline{w'}$ then $(M, \overline{w}) \rightsquigarrow_{\mathsf{C}} (M', \overline{w'})$.*

9.2 Molecular Logics

Definition 19 (C–bisimulation for molecular connectives)**.** Let C be a set of molecular connectives and let $M_1 = (W_1, \mathcal{R}_1)$ and $M_2 = (W_2, \mathcal{R}_2)$ be two C–models. For all $c_0 \in \mathsf{C}$, let V'_{c_0} be the vertices of the quantified decomposition tree of c_0. We associate to each vertex $c \in V'_{c_0}$ of output type or type (if it is a propositional letter) k a relation $Z_c \subseteq (W_1^k \times W_2^k) \cup (W_2^k \times W_1^k)$ such that

i. $Z \triangleq \bigcup \left\{ Z_{id_k^i} \mid id_k^i \text{ appears in some } c \in \mathsf{C} \right\}$ is non-empty;

ii. for all $\overline{w}, \overline{w'}$ of size k such that $\overline{w} Z \overline{w'}$, we also have that for all $c \in \mathsf{C}$ of output type or type (if it is a propositional letter) k that $\overline{w} Z_c \overline{w'}$.

We say that this set of relations $\left\{ Z_c \mid c \in V'_{c_0}, c_0 \in \mathsf{C} \right\}$ is a *C–bisimulation* between M_1 and M_2 when for all $c_0 \in \mathsf{C}$, all vertice $c \in V'_{c_0}$ whose label is different from any id_k^l, if $\{M, M'\} = \{M_1, M_2\}$ then for all $\overline{w}_1, \ldots, \overline{w}_n, \overline{w'}_1, \ldots, \overline{w'}_n, \overline{w}, \overline{w'} \in \overline{w}(M, \mathsf{C}) \cup \overline{w}(M', \mathsf{C})$,

1. if c is of arity 0 then, $\overline{w} Z_c \overline{w'}$ and $\overline{w} \in [\![c]\!]$ imply $\overline{w'} \in [\![c]\!]$;
2. if c is of arity greater than 0 and is maximally decomposed into $c = c_0(c_1, \ldots, c_n)$, with c_0 an existential molecular connective of tonicity signature $(\pm_1, \ldots, \pm_n)$, and we have that $\overline{w} Z_c \overline{w'}$ and $R_{c_0} \overline{w}_1 \ldots \overline{w}_n \overline{w}$, then there are $\overline{w'}_1 \overline{w'}_2 \ldots \overline{w'}_n$ such that $R_{c_0} \overline{w'}_1 \ldots \overline{w'}_n \overline{w'}$ and
   i. for all $j \in [\![1; n]\!]$ such that $\pm_j = +$, there is $i \in [\![1; n]\!]$ such that $c_j = c_i$ and $\overline{w}_i Z_{c_j} \overline{w'}_j$;
   ii. for all $j \in [\![1; n]\!]$ such that $\pm_j = -$, there is $i \in [\![1; n]\!]$ such that $c_j = c_i$ and $\overline{w'}_j Z_{c_j} \overline{w}_i$.
3. if c is of arity greater than 0 and is maximally decomposed into $c = c_0(c_1, \ldots, c_n)$, with c_0 a universal molecular connective of tonicity signature $(\pm_1, \ldots, \pm_n)$, and we have that $\overline{w} Z_c \overline{w'}$ and $-R_{c_0} \overline{w'}_1 \ldots \overline{w'}_n \overline{w'}$, then there are $\overline{w}_1 \overline{w}_2 \ldots \overline{w}_n$ such that $-R_{c_0} \overline{w}_1 \ldots \overline{w}_n \overline{w}$ and
   i. for all $j \in [\![1; n]\!]$ such that $\pm_j = +$, there is $i \in [\![1; n]\!]$ such that $c_j = c_i$ and $\overline{w}_j Z_{c_j} \overline{w'}_i$;
   ii. for all $j \in [\![1; n]\!]$ such that $\pm_j = -$, there is $i \in [\![1; n]\!]$ such that $c_j = c_i$ and $\overline{w'}_i Z_{c_j} \overline{w}_j$.

When we have that $\overline{w} Z \overline{w'}$, we say that $(M, \overline{w})$ and $(M', \overline{w'})$ are *C–bisimilar* and we write it $(M, \overline{w}) \to_{\mathsf{C}} (M', \overline{w'})$. □

Note that if we consider molecular connectives of the form $c \triangleq \star(id_{k_1}^1, \ldots, id_{k_n}^n)$, representing the atomic connective $\star$ of type signature $(k, k_1, \ldots, k_n)$, then the definition of a C–bisimulation for these kind of molecular connectives is identical to the definition of a C–bisimulation for the associated atomic connectives. Hence, our definition of C–bisimulation for molecular connectives is a genuine generalization of our definition of C–bisimulation for atomic connectives. The bisimulation relation Z in the atomic case corresponds in the molecular case to $\bigcup \left\{ Z_{id_k^i} \mid id_k^i \text{ appears in some } c \in \mathsf{C} \right\}$.

Definition 20. Let C be a set of molecular connectives. For all $c_0 \in \mathsf{C}$ and all vertex c of the quantified decomposition tree T_{c_0}, we define the language $\mathcal{L}_{c\mathsf{C}}$ as follows:

$$\mathcal{L}_{c\mathsf{C}} \triangleq \begin{cases} \{c(\varphi_1, \ldots, \varphi_n) \mid \varphi_1, \ldots, \varphi_n \in \mathcal{L}_{\mathsf{C}}\} & \text{if } c \text{ is of arity } n > 0 \\ \{c\} & \text{if } c \text{ is of arity } 0 \end{cases}$$

Let $(M, \overline{w})$ and $(M', \overline{w'})$ be two pointed C–models. We write $(M, \overline{w}) \rightsquigarrow_{c\mathsf{C}} (M', \overline{w'})$ when for all $\varphi \in \mathcal{L}_{c\mathsf{C}}$, $(M, \overline{w}) \Vdash \varphi$ implies $(M', \overline{w'}) \Vdash \varphi$. We also write $(M, \overline{w}) \rightsquigarrow_{\mathsf{C}} (M', \overline{w'})$ when for all $\varphi \in \mathcal{L}_{\mathsf{C}}$, $(M, \overline{w}) \Vdash \varphi$ implies $(M', \overline{w'}) \Vdash \varphi$. □

Proposition 6 (Truth preservation of C-bisimilar pointed C-models)**.** *Let C be a set of molecular connectives and let $M_1 = (W_1, \mathcal{R}_1)$ and $M_2 = (W_2, \mathcal{R}_2)$ be two C–models. Let $\mathsf{C}_0 \subseteq \mathsf{C}$ and for all $c \in \mathsf{C}_0$, let D_c be the vertices of the quantified decomposition tree T_c. Let $\left\{Z_c \mid c \in V'_{c_0}, c_0 \in \mathsf{C}_0\right\}$ be a C_0–bisimulation between M_1 and M_2. If $\{M, M'\} = \{M_1, M_2\}$ then for all $c_0 \in \mathsf{C}_0$ and all $c \in V'_{c_0}$, for all $\overline{w} \in \overline{w}(M, \mathsf{C})$ and all $\overline{w'} \in \overline{w}(M', \mathsf{C})$, if $\overline{w} Z_c \overline{w'}$ then $(M, \overline{w}) \rightsquigarrow_{c\mathsf{C}_0} (M', \overline{w'})$. In particular, if $\overline{w} Z \overline{w'}$ then $(M, \overline{w}) \rightsquigarrow_{\mathsf{C}_0} (M', \overline{w'})$.*

Definition 21 (Normal connectives)**.** A molecular connective is *normal* when its molecular skeleton can be decomposed maximally into a molecular skeleton $c_0(c_1, \ldots, c_n)$ such that for all $i \in [\![1; n]\!]$,

1. c_i is either id^i_k or $c'_i(id^i_k, \ldots, id^i_k)$ with c'_i either universal or existential,
   and in that case we note $id(c_i)$ the unique id^i_k appearing in c_i;
2. if $id(c_i) = id(c_j)$ then $c_i = c_j$. □

In the above definition, since the decomposition $c_0(c_1, \ldots, c_n)$ is maximal, we have in particular, by definition, that if c_0 is of tonicity signature $(\pm_1, \ldots, \pm_n)$, then $Æ(c_i) = - \pm_i Æ(c_0)$.

Example 6. The molecular connectives of modal intuitionistic logic, weakly aggregative modal logic and temporal logic [6, 5] are normal. □

10 Ultrafilters, Ultraproducts and Ultrapowers

In this section, we are going to recall and generalize to molecular logics a number of key notions and results of model theory, such as ultrafilters and ultraproducts. Our definitions are basically the same as those of FOL [10].

Definition 22 (Filter and ultrafilter)**.** Let I be a non–empty set. A *filter F over I* is a set $F \subseteq \mathcal{P}(I)$ such that $I \in F$; if $X, Y \in F$ then $X \cap Y \in F$; if $X \in F$ and $X \subseteq Z \subseteq I$ then $Z \in F$. A filter is called *proper* if it is distinct from $\mathcal{P}(I)$. An *ultrafilter over I* is a proper filter U such that for all $X \in \mathcal{P}(I)$, $X \in U$ iff $I - X \notin U$. A *countably incomplete ultrafilter* is an ultrafilter which is not closed under countable intersections. □

In the rest of this section, I is a non-empty set and U is an ultrafilter over I.

Definition 23 (Ultraproduct of sets)**.** For each $i \in I$, let W_i be a non-empty set. For all $(w_i)_{i\in I}, (v_i)_{i\in I} \in \prod_{i\in I} W_i$, we say that $(w_i)_{i\in I}$ and $(v_i)_{i\in I}$ are *U-equivalent*, written $(w_i)_{i\in I} \sim_U (v_i)_{i\in I}$, if $\{i \in I \mid w_i = v_i\} \in U$. Note that $\sim_U$ is an equivalence relation on $\prod_{i\in I} W_i$. The equivalence class of $(w_i)_{i\in I}$ under $\sim_U$ is denoted $\prod_U w_i \triangleq \left\{(v_i)_{i\in I} \in \prod_{i\in I} W_i \mid (v_i)_{i\in I} \sim_U (w_i)_{i\in I}\right\}$.

The *ultraproduct of $(W_i)_{i\in I}$ modulo U* is $\prod_U W_i \triangleq \left\{\prod_U w_i \mid (w_i)_{i\in I} \in \prod_{i\in I} W_i\right\}$. When $W_i = W$ for all $i \in I$, the ultraproduct is called the *ultrapower* of W modulo U, written $\prod_U W$. □

Definition 24 (Ultraproduct and ultrapower)**.** Let C be a set of molecular connectives and let $(M_i, \overline{w}_i)_{i\in I}$ be a family of pointed C–models. The *ultraproduct* $\prod_U (M_i, \overline{w}_i)$ of $(M_i, \overline{w}_i)$ modulo U is the pointed C–model $\left(\prod_U M_i, \prod_U \overline{w}_i\right)$ where $\prod_U M_i = (W_U, \mathcal{R}_U)$ and $\prod_U \overline{w}_i$ are defined by:

- $W_U = \prod_U W_i$;
- for all $n+1$–ary relations $R^i_\star$ of M_i, the $n+1$–ary relation $\prod_U R_\star \in \mathcal{R}_U$ is defined for all $\prod_U w^1_i, \ldots, \prod_U w^{n+1}_i \in W_U$ by $\prod_U R_\star \prod_U w^1_i \ldots \prod_U w^{n+1}_i$ iff $\left\{i \in I \mid R^i_\star w^1_i \ldots w^{n+1}_i\right\} \in U$;
- $\prod_U \overline{w}_i \triangleq \left(\prod_U w^1_i, \ldots, \prod_U w^k_i\right)$ if $(\overline{w}_i)_{i\in I} = (w^1_i, \ldots, w^k_i)_{i\in I}$.

If $(M_i, s_i)_{i\in I}$ is a family of pointed structures, the *ultraproduct* $\prod_U (M_i, s_i)$ is the pointed structure $\left(\prod_U M_i, \prod_U s_i\right)$ where $\prod_U M_i$ is defined as above (the M_i are viewed as C–models) and $\prod_U s_i : \mathcal{V} \to \prod_U W_i$ is the assignment such that for all $x \in \mathcal{V}$, $\left(\prod_U s_i\right)(x) = \prod_U s_i(x)$.

If for all $i \in I$, $(M_i, \overline{w}_i) = (M, \overline{w})$ (and $(M_i, s_i) = (M, s)$) then $\prod_U (M_i, \overline{w}_i)$ is also called an *ultrapower* of $(M, \overline{w})$ (resp. (M, s)) modulo U, also denoted $\prod_U (M, \overline{w})$ (resp. $\prod_U (M, s)$). □

11 First-Order Logics in Atomic Logics

In this section, we introduce predicate atomic logics, which are, as we will show, as expressive as the predicate atomic logics introduced in [6]. That is why we call both of them with the same name.

Definition 25 (Predicate atomic connectives)**.** Let $\mathcal{P}$ be a set of predicate symbols. The set of *predicate atomic connectives associated to* $\mathcal{P}$ is $\mathsf{C}^{\mathcal{P}} \triangleq \mathcal{P} \cup \{\bot, U\} \cup \{\Box_k, \supset_k \mid k \in \mathbb{N}^*\} \cup \{[f^l_k] \mid k, l \in \mathbb{N}^*$ and $f^l_k : [\![1; l]\!] \to [\![1; k]\!]$ is a mapping$\}$ where, for all $k, l \in \mathbb{N}^*$,

- R has skeleton $(\mathrm{Id}, +, \forall, k)$ if $\mathsf{R} \in \mathcal{P}$ is of arity k;
- $\bot$ has skeleton $(\mathrm{Id}, -, \forall, 1)$;
- U has skeleton $(\mathrm{Id}, -, \forall, (1, 1), +)$;
- $\Box_k$ has skeleton $(\mathrm{Id}, -, \forall, (k, k+1), +)$;
- $\supset_k$ has skeleton $(\mathrm{Id}, -, \forall, (k, k, k), (-, +))$;
- $[f^l_k]$ has skeleton $(\mathrm{Id}, -, \forall, (k, l), +)$. □

Definition 26 (Predicate atomic logic)**.** Let $\mathcal{P}$ be a set of predicate symbols. A *predicate* $\mathsf{C}^{\mathcal{P}}$*–model* is a $\mathsf{C}^{\mathcal{P}}$–model $M = (W, \mathcal{R})$ such that:

- for all $k \in \mathbb{N}^*$, the connectives R of type k are associated to k–ary relations R over W;
- the connective $\perp$ is associated to the 1–ary relation $R_\perp \triangleq W$;
- for all $k, l \in \mathbb{N}^*$ and all mappings $f_k^l : [\![1; l]\!] \to [\![1; k]\!]$, the connective $[f_k^l]$ is associated to the $k + l$–ary relation $R_{[f_k^l]}$ such that for all $w_1, \ldots, w_k, v_1, \ldots, v_l \in W$, we have that $R_{[f_k^l]} v_1 \ldots v_l w_1 \ldots w_k$ iff for all $i \in [\![1; l]\!]$, $v_i = w_{f_k^l(i)}$;
- for all $k \in \mathbb{N}^*$, the connectives $\supset_k$ are associated to the $3k$–ary relation $R_k \subseteq W^{3k}$ such that for all $\overline{w}_1 \in W^k$, all $\overline{w}_2 \in W^k$ and all $\overline{w}_3 \in W^k$, we have that $R_k \overline{w}_1 \overline{w}_2 \overline{w}_3$ iff $\overline{w}_1 = \overline{w}_2 = \overline{w}_3$;
- for all $k \in \mathbb{N}^*$, the connectives $\Box_k$ are associated to the $2k + 1$–ary relation $R_{\Box_k} \subseteq W^{2k+1}$ such that for all $\overline{v} \in W^{k+1}$ and all $\overline{w} \in W^k$, we have that $R_{\Box_k} \overline{v}\overline{w}$ iff $\overline{v} = (\overline{w}, u)$ for some $u \in W$;
- U is associated to the 2–ary relation $R_U \triangleq W^2$.

The class of all pointed predicate $\mathsf{C}^{\mathcal{P}}$–models is denoted $\mathcal{M}_{\mathsf{C}^{\mathcal{P}}}$. The satisfaction relation $\Vdash \subseteq \mathcal{M}_{\mathsf{C}^{\mathcal{P}}} \times \mathcal{L}_{\mathsf{C}^{\mathcal{P}}}$ is then defined following Definition 5. If $\mathcal{E}_{\mathsf{C}^{\mathcal{P}}}$ is a specific class of pointed $\mathsf{C}^{\mathcal{P}}$–model, the triple $(\mathcal{L}_{\mathsf{C}^{\mathcal{P}}}, \mathcal{E}_{\mathsf{C}^{\mathcal{P}}}, \Vdash)$ is called the *predicate atomic logic associated to* $\mathcal{E}_{\mathsf{C}^{\mathcal{P}}}$ *and* $\mathcal{P}$. We also define $\mathcal{L}_{\mathsf{C}^{\mathcal{P}}}(k) \triangleq \{\varphi \in \mathcal{L}_{\mathsf{C}^{\mathcal{P}}} \mid k(\varphi) = k\}$. A *predicate atomic logic of type* k is a triple $(\mathcal{L}_{\mathsf{C}^{\mathcal{P}}}(k), \mathcal{E}_{\mathsf{C}^{\mathcal{P}}}(k), \Vdash)$ such that all pointed $\mathsf{C}^{\mathcal{P}}$–models of $\mathcal{E}_{\mathsf{C}^{\mathcal{P}}}(k)$ are of type k. $\Box$

Our connectives $[f_k^l]$ are a generalization of the connectives $\odot_\sigma$ of multi-dimensional modal logics [21]. In fact, if $k = l$ then $[f_k^k]$ is exactly $\odot_{f_k^k}$.

Definition 27 (Translation from FOL to predicate atomic logics)**.** Let $\mathcal{P}$ be a set of predicate symbols.
Syntax. For all $k \in \mathbb{N}$ and all $\overline{x} = (x_1, \ldots, x_k) \in (\mathcal{V} \cup \mathcal{C})^k$, we define the mappings $T_{\overline{x}} : \mathcal{L}_{\mathsf{FOL}}^{\mathcal{P}}(\overline{x}, k) \to \mathcal{L}_{\mathsf{C}^{\mathcal{P}}}(k)$ and $T_\emptyset : \mathcal{L}_{\mathsf{FOL}}^{\mathcal{P}}(\emptyset) \to \mathcal{L}_{\mathsf{C}^{\mathcal{P}}}(1)$ inductively on the formula $\varphi(x_1, \ldots, x_k) \in \mathcal{L}_{\mathsf{FOL}}^{\mathcal{P}}$ (with or without free variables) as follows.

- if $\varphi = \perp$ then we define $T_\emptyset(\perp) \triangleq \perp$;
- if $\varphi = \mathsf{R} x_{i_1} \ldots x_{i_l}$ with $k \leq l$ (some variables or constants can be the same and variables and constants may be in a different order than $(x_1, \ldots, x_k)$) then there is a unique (surjective) mapping $f_k^l : [\![1; l]\!] \to [\![1; k]\!]$ such that for all $j \in [\![1; l]\!]$, we have that $f_k^l(j) = i_j$. Finally, we define

  $T_{\overline{x}}(\mathsf{R} y_1 \ldots y_l) \triangleq \left[f_k^l\right] \mathsf{R}$;
- if $\varphi = (\varphi_1 \to \varphi_2)$ where φ_1 and φ_2 are both sentences, then we define

  $T_\emptyset(\varphi) \triangleq \supset_1 (T_\emptyset(\varphi_1), T_\emptyset(\varphi_2))$

  if $\varphi = (\varphi_1 \to \varphi_2(x))$ where φ_1 is a sentence, then we define

  $T_{(x)}(\varphi) \triangleq \supset_1 \left(T_\emptyset(\varphi_1), T_{(x)}(\varphi_2(x))\right)$

  if $\varphi = (\varphi_1(x) \to \varphi_2)$ where φ_2 is a sentence, then we define

  $T_{(x)}(\varphi) \triangleq \supset_1 \left(T_{(x)}(\varphi_1(x)), T_\emptyset(\varphi_2)\right)$

  if $\varphi = (\varphi_1 \to \varphi_2(x_1, \ldots, x_k))$ where φ_1 is a sentence and $k > 1$, then we define

$T_{\overline{x}}(\varphi) \triangleq \supset_k \left([f_k^1]T_\emptyset(\varphi_1), T_{(x_1,\ldots,x_k)}(\varphi_2(x_1,\ldots,x_k))\right)$
where $f_k^1 : \{1\} \to [\![1;k]\!]$ is defined by $f_k^1(1) = 1$;
if $\varphi = (\varphi_1(x_1,\ldots,x_k) \to \varphi_2)$ where φ_2 is a sentence and $k > 1$, then we define
$T_{\overline{x}}(\varphi) \triangleq \supset_k \left(T_{(x_1,\ldots,x_k)}(\varphi_1(x_1,\ldots,x_k)), [f_k^1]T_\emptyset(\varphi_2)\right)$
where f_k^1 is defined as above;
if $\varphi = (\varphi_1(x_{i_1},\ldots,x_{i_{k_1}}) \to \varphi_2(y_{j_1},\ldots,y_{j_{k_2}}))$ then we define
$T_{\overline{x}}(\varphi) \triangleq \supset_k \left(\mathrm{Exp}_{\overline{x}}(\varphi_1(x_{i_1},\ldots,x_{i_{k_1}})), \mathrm{Exp}_{\overline{x}}(\varphi_2(x_{j_1},\ldots,x_{j_{k_2}}))\right)$
where

$$\mathrm{Exp}_{\overline{x}}(\varphi_1(x_{i_1},\ldots,x_{i_{k_1}})) \triangleq \begin{cases} T_{(x_1,\ldots,x_k)}(\varphi_1(x_{i_1},\ldots,x_{i_{k_1}})) & \text{if } k_1 = k \\ \left[f_k^{k_1}\right] T_{(x_{i_1},\ldots,x_{i_{k_1}})}(\varphi_1(x_{i_1},\ldots,x_{i_{k_1}})) & \text{if } k_1 < k \end{cases}$$

where for all $j \in [\![1;k_1]\!]$, we have that $f_k^{k_1}(j) = i_j$. The definition is similar for $\mathrm{Exp}_{\overline{x}}(\varphi_2(x_{j_1},\ldots,x_{j_{k_2}}))$, one only needs to replace i by j and k_1 by k_2.

- if $\varphi = \forall x\psi(x)$ where φ is a sentence, then we define
  $T_\emptyset(\varphi) \triangleq UT_{(x)}(\psi(x))$;
  if $\varphi = \forall x\psi(x_1,\ldots,x_k,x)$ with $k \geq 1$, then we define
  $T_{\overline{x}}(\varphi) \triangleq \Box_k T_{(x_1,\ldots,x_k,x)}(\psi(x_1,\ldots,x_k,x))$.

Semantics. Let $M = (W, \{R_1,\ldots,R_n,\ldots,c_1,\ldots,c_n,\ldots\})$ be a structure without functions. We define the $\mathsf{C}^{\mathcal{P}}$–model $T(M) = (W,\mathcal{R})$ by adding to the relations of M the relations defined in Definition 26. If $\overline{x} = (x_1,\ldots,x_k)$ is a tuple of k variables then we define $T_{\overline{x}}(M,s) \triangleq (T(M), (s(w_1),\ldots,s(w_k)))$ and $T_\emptyset(M,s) \triangleq (T(M),(s(x_0)))$ for an arbitrary $x_0 \in \mathcal{V}$. □

Lemma 1. *Let $\varphi \in \mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}$, let $\overline{x}$ be the tuple of free variables and constants of φ (possibly empty) and let (M,s) be a pointed structure. Then, we have that*

$$(M,s) \models \varphi \text{ iff } T_{\overline{x}}(M,s) \Vdash T_{\overline{x}}(\varphi).$$

Definition 28 (Translation from predicate atomic logics to FOL)**.**
Syntax. For all $k \in \mathbb{N}^*$ and all tuples $\overline{x} = (x_1,\ldots,x_k)$ of variables or constants, we define the mappings $ST_{\overline{x}} : \mathcal{L}^k_{\mathsf{C}^{\mathcal{P}}} \to \mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}$, where $\mathcal{L}^k_{\mathsf{C}^{\mathcal{P}}}$ is the set of formulas of $\mathcal{L}_{\mathsf{C}^{\mathcal{P}}}$ of type k, inductively as follows:

$$\begin{array}{rcl} ST_{\overline{x}}(\mathsf{R}) & \triangleq & \mathsf{R}\overline{x} \\ ST_{\overline{x}}(\bot) & \triangleq & \bot \\ ST_{\overline{x}}([f_k^l]\varphi) & \triangleq & ST_{(x_{f_k^l(1)},\ldots,x_{f_k^l(l)})}(\varphi) \\ ST_{\overline{x}}(\supset_k(\varphi_1,\varphi_2)) & \triangleq & (ST_{\overline{x}}(\varphi_1) \to ST_{\overline{x}}(\varphi_2)) \\ ST_{\overline{x}}(\Box_k\varphi) & \triangleq & \forall x ST_{(\overline{x},x)}(\varphi) \\ ST_{(x)}(U\varphi) & \triangleq & \forall x ST_{(x)}(\varphi) \end{array}$$

where t is an arbitrary term of $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}$ (we recall that $\vee$ is here an abbreviation).
Semantics. Let $(M,(w_1,\ldots,w_k))$ be a pointed $\mathsf{C}^{\mathcal{P}}$–model of type k and let $\overline{x} = (x_1,\ldots,x_k)$ be a tuple of free variables or constants of size k. The *(pointed) structure associated to* $(M,\overline{w})$, denoted $ST_{\overline{x}}(M,\overline{w}) \triangleq (ST(M), s^{\overline{w}}_{\overline{x}})$, is defined as follows. The assignment $s^{\overline{w}}_{\overline{x}}$ is such that for all $i \in \{1,\ldots,k\}$, $s(x_i) = w_i$ and for all $x \in \mathcal{V} - \{x_1,\ldots,x_k\}$, $s(x) = w_1$ and $ST(M)$ is the structure $ST(M) = (W, ST(\mathcal{R}))$ where $ST(\mathcal{R})$ is the set $\mathcal{R}$ to which we remove the relations stemming from Definition 26. □

Lemma 2. *Let $(M, \overline{w})$ be a pointed predicate $\mathcal{C}^{\mathcal{P}}$–model of type k, let $\varphi \in \mathcal{L}_{\mathcal{C}^{\mathcal{P}}}$ of type k and let $\overline{x} \in (\mathcal{V} \cup \mathcal{C})^k$. Then,*

$$(M, \overline{w}) \Vdash \varphi \text{ iff } ST_{\overline{x}}(M, \overline{w}) \models ST_{\overline{x}}(\varphi).$$

Moreover, for all pointed structures (M, s) without distinguished elements, we have that $ST_{\overline{x}}(T_{\overline{x}}(M, s)) \equiv_{\overline{x}} (M, s)$ and for all pointed predicate $\mathcal{C}^{\mathcal{P}}$–model $(M, \overline{w})$, we have that $T_{\overline{x}}(ST_{\overline{x}}(M, \overline{w})) \equiv_k (M, \overline{w})$.

Proposition 7. *Let $\mathcal{P}$ be a set of predicate symbols and $k \in \mathbb{N}^*$. If $\mathcal{E}_{\mathcal{C}^{\mathcal{P}}}$ is a class of pointed $\mathcal{C}^{\mathcal{P}}$–model, the predicate atomic logic (of type k) associated to $\mathcal{E}_{\mathcal{C}^{\mathcal{P}}}$ and $\mathcal{P}$ is as expressive as a predicate atomic logic (of type k) as defined in [6, Definition 18], and vice versa.*

Therefore, [6, Theorem 1] and [6, Corollary 1] hold as well for the versions of predicate atomic logics introduced in the present article.

12 Generalized Keisler Theorems for First-Order Logics

In that section, we assume that the first–order languages are countable.

12.1 Bisimulations for Pure Predicate Logic

Applying our Definition 17 to our predicate atomic logics of Section 11 we obtain the following conditions: for all $M, M' \in \{M_1, M_2\}$ with $M = (W, \mathcal{R})$ and $M' = (W', \mathcal{R}')$, all $\overline{w}, \overline{v}, \overline{u} \in \overline{w}(M, \mathsf{C})$, all $w, v \in W$ and all $\overline{w'}, \overline{v'}, \overline{u'} \in \overline{w}(M', \mathsf{C})$, all $w', v' \in W'$, all $\mathsf{R} \in \mathcal{P}$,

- condition for R: if $\overline{w} Z \overline{w'}$ and $R\overline{w}$ then $R'\overline{w'}$, for all relations R and R' both associated to R (in the associated structure);
- condition for $\bot$: if wZw' and $R'_{\bot}(w')$ then $R_{\bot}(w)$, this condition always holds;
- condition for $[f^l_k]$: for all $k, l \in \mathbb{N}^*$ and all mappings f^l_k, if $\overline{w} Z \overline{w'}$ and $R'_{[f^l_k]}\overline{w'v'}$ then there is $\overline{v} \in \overline{w}(M, \mathsf{C})$ such that $\overline{v} Z \overline{v'}$ and $R_{[f^l_k]}\overline{wv}$,

  that is, if $\overline{w} Z \overline{w'}$ then $f^l_k(\overline{w}) Z f^l_k(\overline{w'})$, where for all $\overline{w} = (w_1, \ldots, w_k)$, $f^l_k(\overline{w}) \triangleq (w_{f^l_k(1)}, \ldots, w_{f^l_k(l)})$;
- condition for $\supset_k$: for all $k \in \mathbb{N}^*$, if $\overline{w} Z \overline{w'}$ and $R_k\overline{v'u'w'}$ then there are $\overline{v}, \overline{u} \in \overline{w}(M, \mathsf{C})$ such that $\overline{v'} Z \overline{v}$, $\overline{u} Z \overline{u'}$ and $R_k\overline{vuw}$,

  that is, if $\overline{w} Z \overline{w'}$ then $\overline{w'} Z \overline{w}$;
- condition for $\Box_k$: for all $k \in \mathbb{N}^*$, if $\overline{w} Z \overline{w'}$ and $R'_{\Box_k}\overline{v'w'}$ then there is $\overline{v} \in \overline{w}(M, \mathsf{C})$ such that $\overline{v} Z \overline{v'}$ and $R_{\Box_k}\overline{vw}$,

  that is, if $\overline{w} Z \overline{w'}$ then for all $v' \in W'$ there is $v \in W$ such that $(\overline{w}, v) Z (\overline{w'}, v')$;
- condition for U: if wZw' and $v' \in W'$ then there is $v \in W$ such that vZv',

  that is, for all $w' \in W'$ there is $w \in W$ such that wZw'.

If we combine the conditions above, we obtain the following definition.

Definition 29 (Predicate bisimulation)**.** A *predicate bisimulation* between two structures M_1 and M_2 is a non–empty relation Z between finite sequences of the same length of M_1 and M_2 or M_2 and M_1 such that for all $M, M' \in \{M_1, M_2\}$ with $M = (W, \mathcal{R})$ and $M' = (W', \mathcal{R}')$, all $m, n \in \mathbb{N}^*$, all $w_1, \ldots, w_n \in W$, all $w'_1, \ldots, w'_n \in W'$,

1. if $(w_1, \ldots, w_n) Z (w'_1, \ldots, w'_n)$ then for all n–ary relations $R \in \mathcal{R}$ and $R' \in \mathcal{R}'$ associated to the same predicate R, if $R w_1 \ldots w_n$ then $R' w'_1 \ldots w'_n$;
2. if $(w_1, \ldots, w_n) Z (w'_1, \ldots, w'_n)$ then for all $w \in M$ there is $w' \in M'$ such that $(w_1, \ldots, w_n, w) Z (w'_1, \ldots, w'_n, w')$;
3. if $(w_1, \ldots, w_n) Z (w'_1, \ldots, w'_n)$ then for all mappings $f : [\![1; m]\!] \to [\![1; n]\!]$, we have that $(w_{f(1)}, \ldots, w_{f(m)}) Z (w'_{f(1)}, \ldots, w'_{f(m)})$;
4. if $(w_1, \ldots, w_n) Z (w'_1, \ldots, w'_n)$ then $(w'_1, \ldots, w'_n) Z (w_1, \ldots, w_n)$. □

Let $X \subseteq \mathcal{V} \cup \mathcal{C}$ be a non-empty set. An *X–compatible predicate bisimulation* Z between two pointed structures (M, s) and (M', s') is a predicate bisimulation between M and M' such that

5. for all $x_1, \ldots, x_n \in X$, it holds that $(s(x_1), \ldots, s(x_n))\, Z\, (s'(x_1), \ldots, s'(x_n))$.

A $\mathcal{C}$-compatible predicate bisimulation between two pointed structures is often abusively called a predicate bisimulation. □

Now, we recall the usual notion of partial isomorphism.

Definition 30 (Partial isomorphism)**.** A *partial isomorphism* between two structures M_1 and M_2 is a relation Z on the set of pairs of finite sequences $(w_1, \ldots, w_n), (v_1, \ldots, v_n)$ of elements of M_1 and M_2 of the same length such that:

1. if $(w_1, \ldots, w_n) Z (v_1, \ldots, v_n)$ then (M_1, s_1) and (M_2, s_2) make true the same atomic formulas $\varphi(x_1, \ldots, x_n) \in \mathcal{L}^{\mathcal{PF}}_{\mathsf{FOL}}$, where s_1 and s_2 are such that for all $i \in \{1, \ldots, n\}$, $s_1(x_i) = w_i$ and $s_2(x_i) = v_i$;
2. if $(w_1, \ldots, w_n) Z (v_1, \ldots, v_n)$ then for all $w \in M_1$ there is $v \in M_2$ such that $(w_1, \ldots, w_n, w) Z (v_1, \ldots, v_n, v)$, and vice versa;
3. for all $w \in M_1$ there is $v \in M_2$ such that $(w) Z (v)$, and vice versa. □

Condition 3 could be equivalently replaced by the condition $\emptyset Z \emptyset$, like in [10]. Predicate bisimulations are basically partial isomorphisms as one can easily notice:

Proposition 8. *A predicate bisimulation between pointed structures without functions is a partial isomorphism.*

Proposition 9. *Let $X \subseteq \mathcal{V} \cup \mathcal{C}$ be a non-empty set. If there is an X–compatible predicate bisimulation between two pointed structures then they make true the same formulas of $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$.*

The analogue of the Keisler-Shelah theorem below generalizes it in the sense that it applies to larger classes of first-order formulas than the class of sentences only.

Theorem 1. *Let $X \subseteq \mathcal{V} \cup \mathcal{C}$ be non-empty and assume that $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$ is countable. Let (M, s) and (M', s') be two pointed structures. Then the following are equivalent:*

1. *(M, s) and (M', s') make true the same formulas of $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$;*

2. *there exists an ultrafilter U and a X–compatible predicate bisimulation between* $\prod_U (M, s)$ *and* $\prod_U (M', s')$.

The following corollary with sentences only is an analogue of the Keisler-Shelah isomorphism theorem [10, Theorem 6.1.15] in which isomorphisms are replaced by predicate bisimulations.

Corollary 1. *Assume that $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}$ is countable and let (M, s) and (M', s') be two pointed structures. Then the following are equivalent:*

1. *(M, s) and (M', s') make true the same sentences of $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}$;*
2. *there exists an ultrafilter U and a predicate bisimulation between* $\prod_U (M, s)$ *and* $\prod_U (M', s')$.

Theorem 2. *Let $X \subseteq \mathcal{V} \cup \mathcal{C}$ be non-empty and assume that $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$ is countable. Let K and K' be classes of pointed structures such that K' is closed under ultraproduct. Then, the following are equivalent:*

1. *K is definable in K' by a set of formulas of $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$;*
2. *K is closed under X–compatible predicate bisimulations in K' and closed under ultraproducts, and $K' - K$ is closed under ultrapowers.*

Corollary 2. *Let K be a class of pointed structures. Then, the following are equivalent:*

1. *K is definable by a set of sentences of $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}$;*
2. *K is closed under predicate bisimulations and ultraproducts, and $\overline{K}$ is closed under ultrapowers.*

A result similar to Corollary 2 was already proved by van Benthem & Doets [27, p. 285]. In their result, isomorphisms are replaced by partial isomorphisms.

Theorem 3. *Let $X \subseteq \mathcal{V} \cup \mathcal{C}$ be non-empty and assume that $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$ is countable. Let K and K' be classes of pointed structures such that K' is closed under ultraproducts. Then, the following are equivalent:*

1. *K is definable in K' by means of a single formula $\varphi \in \mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$;*
2. *Both K and $K' - K$ are closed under X–compatible predicate bisimulations in K' and closed under ultraproducts.*

Corollary 3. *Let K be a class of pointed structures. Then, the following are equivalent:*

1. *K is definable by means of a sentence of $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}$;*
2. *Both K and $\overline{K}$ are closed under predicate bisimulations and ultraproducts.*

12.2 Bisimulations for First–order Logic

The results of the previous section hold for pure predicate logic, that is first–order logic without functions. In this section, we are going to extend them to first–order logic (with functions).

Definition 31 (First–order bisimulation)**.** A *first–order bisimulation* between two structures M_1 and M_2 is a pair of non–empty relations (Z, Z_0) between finite sequences of M_1 and M_2 of the same length such that for all $M, M' \in \{M_1, M_2\}$ with $M = (W, \mathcal{R})$ and $M' = (W', \mathcal{R}')$, all $m, n \in \mathbb{N}^*$, all $w_1, \ldots, w_n \in W$, all $w'_1, \ldots, w'_n \in W'$, all functions $f_1, \ldots, f_n$ of M and corresponding functions $f'_1, \ldots, f'_n$ of M', all tuples $\overline{v}_1, \ldots, \overline{v}_n$ of W and $\overline{v'}_1, \ldots, \overline{v'}_n$ of W',

1. if $(w_1, \ldots, w_n) Z_0 (w'_1, \ldots, w'_n)$ then for all n–ary relations $R \in \mathcal{R}$ and $R' \in \mathcal{R}'$ associated to the same predicate R, if $R w_1 \ldots w_n$ then $R' w'_1 \ldots w'_n$;
2. if $(w_1, \ldots, w_n) Z_0 (w'_1, \ldots, w'_n)$ and $f_1(\overline{v'}_1) = w'_1, \ldots, f_n(\overline{v'}_n) = w'_n$ then there are finite sequences $\overline{v}_1, \ldots, \overline{v}_n$ of M such that $\overline{v}_1 Z_0 \overline{v'}_1, \ldots, \overline{v}_n Z_0 \overline{v'}_n$ and $f_1(\overline{v}_1) = w_1, \ldots, f_n(\overline{v}_n) = w_n$;
3. if $(w_1, \ldots, w_n) Z_0 (w'_1, \ldots, w'_n)$ then for all mappings $f : [\![1; m]\!] \to [\![1; n]\!]$, we have that $(w_{f(1)}, \ldots, w_{f(m)}) Z_0 (w'_{f(1)}, \ldots, w'_{f(m)})$;
4. if $(w_1, \ldots, w_n) Z (w'_1, \ldots, w'_n)$ then $(w_1, \ldots, w_n) Z_0 (w'_1, \ldots, w'_n)$;
5. if $(w_1, \ldots, w_n) Z (w'_1, \ldots, w'_n)$ then for all $w \in M$ there is $w' \in M'$ such that $(w_1, \ldots, w_n, w) Z (w'_1, \ldots, w'_n, w')$;
6. if $(w_1, \ldots, w_n) Z (w'_1, \ldots, w'_n)$ then for all mappings $f : [\![1; m]\!] \to [\![1; n]\!]$, we have that $(w_{f(1)}, \ldots, w_{f(m)}) Z (w'_{f(1)}, \ldots, w'_{f(m)})$;
7. if $(w_1, \ldots, w_n) Z (w'_1, \ldots, w'_n)$ then $(w'_1, \ldots, w'_n) Z (w_1, \ldots, w_n)$.

Let $X \subseteq \mathcal{V} \cup \mathcal{C}$ be a non-empty set. An X*–compatible first-order bisimulation* Z between two pointed structures (M, s) and (M', s') is a first-order bisimulation between M and M' such that

8. for all $x_1, \ldots, x_n \in X$, it holds that $(s(x_1), \ldots, s(x_n))\, Z\, (s'(x_1), \ldots, s'(x_n))$.

A $\mathcal{C}$-compatible first-order bisimulation between two pointed structures is often abusively called a first-order bisimulation. □

The bisimulation notion that we obtain in this extended setting with functions is again different from the usual notion of (partial) isomorphism, but the difference, which concerns the atomic base, is more visible in that case than for pure predicate logics. Basically, condition 1 of a (partial) isomorphism is replaced by a set of back and forth semantic conditions dealing with the functions, but without any reference to, or evaluation of, first-order (atomic) formulas, like for our predicate bisimulations.

Theorem 4. *Theorems 1, 2, 3 and Corollaries 1, 2, 3 hold if we replace at the same time the term "predicate bisimulation" with "first–order bisimulation" and the languages $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}$ and $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$ with the languages with function symbols $\mathcal{L}^{\mathcal{PF}}_{\mathsf{FOL}}$ and $\mathcal{L}^{\mathcal{PF}}_{\mathsf{FOL}}(X)$ respectively.*

The invariance notions of pure predicate logic and first–order logic *with equality* are also predicate bisimulation and first–order bisimulation because the equality predicate is considered on a par with the other predicates, it is just like any other predicate. In that case, the

corresponding condition 1. of Definitions 29 and 31 for the equality predicate is the following: if $(w, w)Z_0(w'_1, w'_2)$ then $w'_1 = w'_2$. With the other conditions, this condition is equivalent to: $(w_1, \ldots, w_i, \ldots, w_i, \ldots, w_n)Z_0(w'_1 \ldots, w'_i, \ldots, w'_j, \ldots, w'_n)$ implies $w'_i = w'_j$. This condition does not follow from the other conditions of a first-order or a predicate bisimulation. On the other hand, it holds for these notions that if $(w_1, \ldots, w_i, \ldots, w_n)Z(w'_1, \ldots, w'_i, \ldots, w'_n)$ then $(w_1, \ldots, w_i, \ldots, w_i, \ldots, w_n)Z(w'_1, \ldots, w'_i, \ldots, w'_i, \ldots, w'_n)$.

Proposition 10. *If there is a first–order bisimulation between two pointed structures then there is a predicate bisimulation and a partial isomorphism between them. Any two finite or countable structures such that there is a predicate or (if they contain functions) a first–order bisimulation between them are isomorphic.*

13 Protologics in Molecular Logics

Protologics, introduced in [6], are logics such that the truth conditions of their connectives can be expressed by means of first-order formulas. Hence, protologics represent and correspond to a very large fragment of non-classical logics. We show that every protologic is as expressive as a molecular logic. The proofs are essentially the same as in [6]. In this section, $\mathcal{Q}$ and $\mathcal{P}$ are sets of predicates such that $\mathcal{Q} \subseteq \mathcal{P}$.

Definition 32 (Translation from protologics to molecular logics)**.** Let C^a be a set of abstract connectives.
Syntax. We define the mapping t from the abstract connectives of C^a to molecular connectives as follows:

- $t(\mathsf{Q}) = \mathsf{Q}$ for all predicate $\mathsf{Q} \in \mathcal{Q}$;
- For abstract connectives of the form $(\chi(\overline{x}), (\mathsf{Q}_1, \ldots, \mathsf{Q}_n))$, we proceed as follows. We first translate $\chi(\overline{x})$ into predicate atomic logic using the translation $T_{\overline{x}}$ of Definition 27 We obtain a formula $\chi' = T_{\overline{x}}(\chi(\overline{x}))$ of predicate atomic logic. That transformation is injective. That formula χ' is then transformed into a molecular connective c. The skeleton decomposition tree of that molecular connective is the skeleton syntactic tree of χ' where the leafs labeled with the skeleton of a predicate Q_i are all replaced by the same skeleton symbol $id^l_{k_i}$ where k_i is the arity of each Q_i and l is a natural number which allows to distinguish $id^l_{k_i}$ from the other symbols of the form $id^{l'}_{k_i}$ appearing in the skeleton. The resulting molecular connective c of arity n and type signature $(k, k_1, \ldots, k_n)$ is denoted $t(\chi(\overline{x}), (\mathsf{Q}_1, \ldots, \mathsf{Q}_n))$.

The resulting set of molecular connectives is denoted $t(\mathsf{C}^a)$. Then, this translation t is extended to the whole language as follows: for all $\star \in \mathsf{C}^a$ and all $\varphi_1, \ldots, \varphi_n \in \mathcal{L}_{\mathsf{C}^a}$ of appropriate types, $t(\star(\varphi_1, \ldots, \varphi_n)) = t(\star)(t(\varphi_1), \ldots, t(\varphi_n))$. By construction, this transformation t is injective.
Semantics. Let $(M, \overline{w})$ be a pointed C^a–model with $M = (W, \mathcal{R})$. The pointed $t(\mathsf{C}^a)$–model $t(M, \overline{w})$ is the $t(\mathsf{C}^a)$–model $((W, \mathcal{R}'), \overline{w})$ where $\mathcal{R}'$ is $\mathcal{R}$ together with the relations of Definition 26. □

Lemma 3. *Let C^a be a set of abstract connectives, let $\varphi \in \mathcal{L}_{\mathsf{C}^a}$ and let $(M, \overline{w})$ be a pointed C^a–model where $\overline{w}$ is of size the type of φ. Then, we have that*

$$(M, \overline{w}) \Vdash \varphi \text{ iff } t(M, \overline{w}) \Vdash t(\varphi)$$

Proof: It is the same as [6, Lemma 3]. □

Definition 33 (Translation from molecular logics to protologics)**.** Let C be a set of molecular connectives.
Syntax. We define the mapping st from molecular connectives to abstract connectives inductively as follows:

- $st(p) \triangleq p$ for all propositional letters p of C.
- For all molecular connectives of the form $c = \star(c_1, \ldots, c_m)$ we proceed as follows. First, we replace all symbols id_k^l appearing in c by fresh and distinct propositional letters $(p_1, \ldots, p_n)$. This yields a formula $\varphi \in \mathcal{L}_{\mathsf{C}}$ of some type k. Then, we pick a tuple of free variables $\overline{x}$ of size k and we define the first–order formula $st_{\overline{x}}(\varphi)$ inductively as follows. If φ is a propositional letter p then $st_{\overline{x}}(p) \triangleq \mathsf{Q}\overline{x}$, where Q is a predicate symbol of $\mathcal{Q}$. If φ is of the form $\star(\varphi_1, \ldots, \varphi_m)$ then
  - if $\star = (\sigma, \pm, \forall, \overline{k}, (\pm_1, \ldots, \pm_m))$ then
    $st_{\overline{x}}(\star(\varphi_1, \ldots, \varphi_m)) \triangleq \forall \overline{y}_1 \ldots \overline{y}_m \left(\pm_1 st_{\overline{y}_1}(\varphi_1) \vee \ldots \vee \pm_n st_{\overline{y}_m}(\varphi_m) \vee \pm \mathsf{R}^\sigma \overline{y}_1 \ldots \overline{y}_m \overline{x}\right)$;
  - if $\star = (\sigma, \pm, \exists, \overline{k}, (\pm_1, \ldots, \pm_m))$ then
    $st_{\overline{x}}(\star(\varphi_1, \ldots, \varphi_m)) \triangleq \exists \overline{y}_1 \ldots \overline{y}_m \left(\pm_1 st_{\overline{y}_1}(\varphi_1) \wedge \ldots \wedge \pm_n st_{\overline{y}_n}(\varphi_m) \wedge \pm \mathsf{R}^\sigma \overline{y}_1 \ldots \overline{y}_m \overline{x}\right)$;

  where $\overline{y}_1, \ldots, \overline{y}_n$ are fresh tuples of free variables and R is a predicate symbol of $\mathcal{P} - \mathcal{Q}$. We recall that for all formulas ψ, $\pm_i \psi$ stands for ψ if $\pm_i = +$ and for $\neg\psi$ if $\pm_i = -$, and that $\mathsf{R}^\sigma \overline{y}_1 \ldots \overline{y}_m \overline{y}_{m+1} \triangleq \mathsf{R}\overline{y}_{\sigma^-(1)} \cdots \overline{y}_{\sigma^-(m)} \overline{y}_{\sigma^-(m+1)}$.

  Finally, we define the abstract connective $st(c) \triangleq (st_{\overline{x}}(\varphi), (\mathsf{Q}_1, \ldots, \mathsf{Q}_n))$ where for all $i \in [\![1; n]\!]$, $\mathsf{Q}_i = st(p_i)$.

The resulting set of abstract connectives is denoted $st(\mathsf{C})$. Then, this translation st is extended to the whole language as follows: for all $\star \in \mathsf{C}$ and all $\varphi_1, \ldots, \varphi_n \in \mathcal{L}_{\mathsf{C}}$ of appropriate types, $st(\star(\varphi_1, \ldots, \varphi_n)) = st(\star)(st(\varphi_1), \ldots, st(\varphi_n))$. By construction, this transformation st is injective. □

Lemma 4. *Let C be a set of molecular connectives, let $\varphi \in \mathcal{L}_{\mathsf{C}}$ and let $(M, \overline{w})$ be a pointed C–model where $\overline{w}$ is of size the type of φ. Then, we have that*

$$(M, \overline{w}) \Vdash \varphi \text{ iff } (M, \overline{w}) \Vdash st(\varphi)$$

Proof: It is the same as [6, Lemma 4]. □

Theorem 5. *The class of protologics is as expressive as the class of molecular logics.*

Proof: It is the same as [6, Theorem 2]. □

Definition 34 (Predicate molecular logic)**.** A *predicate molecular logic* is a molecular logic whose molecular connectives are built up from the predicate atomic connectives $\{\bot, U\} \cup \{\Box_k, \supset_k | \ k \in \mathbb{N}^*\} \cup \{[f_k^l] \mid k, l \in \mathbb{N}^*$ and $f_k^l : [\![1; l]\!] \to [\![1; k]\!]$ is a mapping$\}$ of Definitions 25 and 26, together with some set of propositional letters (possibly empty). □

Theorem 6. *The class of predicate molecular logics is as expressive as the class of molecular logics and therefore also as expressive as the class of protologics.*

Proof: It is the same as [6, Theorem 3]. □

14 Generalized Keisler Theorems for Protologics

Our ultimate objective in this article is to provide an appropriate notion of bisimulation for any protologic, together with Keisler type theorems, when this is possible. This is what we are going to achieve in this section thanks to the translation from protologics to (predicate) molecular logics which was defined in the previous sections. Our method is the following:

1. Translate the given protologic L into a predicate molecular logic L' (thanks to the translation t of Definition 32);
2. Define the notion of C^a–bisimulation associated to L via the notion of C–bisimulation associated to L' (thanks to Definition 35);
3. State Keisler type characterization theorems for L if its connectives are normal (our Theorem 7).

Every protologic $(\mathcal{L}_{\mathsf{C}^a}, \mathcal{E}_{\mathsf{C}^a}, \Vdash)$ is as expressive as the predicate molecular logic $(t(\mathcal{L}_{\mathsf{C}^a}), t(\mathcal{E}_{\mathsf{C}^a}), \Vdash)$ (see Definition 32) which is itself based on a set of molecular connectives $t(\mathsf{C}^a)$ defined by the translation t of Definition 32. The notion of $t(\mathsf{C}^a)$–bisimulation canonically induces a notion of bisimulation for the protologic $(\mathcal{L}_{\mathsf{C}^a}, \mathcal{E}_{\mathsf{C}^a}, \Vdash)$, as follows.

Definition 35 (C^a–bisimulation for abstract connectives)**.** Let C^a be a set of abstract connectives and let $(\mathcal{L}_{\mathsf{C}^a}, \mathcal{E}_{\mathsf{C}^a}, \Vdash)$ be a protologic (defined in [6, Sect. 4.1]). Let $(M, \overline{w}), (M', \overline{w'}) \in \mathcal{E}_{\mathsf{C}^a}$ be two pointed C^a–models. We say that $(M, \overline{w})$ and $(M', \overline{w'})$ are *C^a–bisimilar*, written $(M, \overline{w}) \rightarrow_{\mathsf{C}^a} (M', \overline{w'})$, when $t(M, \overline{w})$ and $t(M', \overline{w'})$ (defined in Definition 32) are $t(\mathsf{C}^a)$–bisimilar, that is $t(M, \overline{w}) \rightarrow_{t(\mathsf{C}^a)} t(M', \overline{w'})$. We also say that the set of abstract connectives C^a is *normal* when the set of molecular connectives $t(\mathsf{C}^a)$ is itself normal. □

This is well-defined because for any pointed C^a–model $(M, \overline{w})$, we have by definition of t that $(M, \overline{w})$ and $t(M, \overline{w})$ are based on the same set of worlds W.

Theorem 7 (Keisler theorems for protologics)**.** *Let C^a be a set of normal abstract connectives complete for conjunction and disjunction. Then, theorems 1, 2, 3 hold if we replace "pointed structures" by "pointed C^a–models", the language "$\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$" by the language "$\mathcal{L}_{\mathsf{C}^a}$" and "$X$–compatible predicate bisimulation" by "C^a–bisimulation".*

Atomic and molecular connectives C can be equivalently seen as abstract connectives C^a, in the sense that one can find abstract connectives which have the same truth conditions as any given atomic or molecular connective.[2] For example, the connective of tense logic of skeleton $(\tau_1, -, t_2)$ has the same truth conditions as the abstract connective $(\chi(x), \mathsf{Q})$, where $\chi(x) \triangleq \forall y(\mathsf{R}yx \rightarrow \mathsf{Q}y)$. Theorems 7 suggests that the notion of C–bisimulation for atomic and molecular connectives C and the notion of C^a–bisimulation for the corresponding abstract connectives should be the same. We can indeed check on the example of tense logic that they do coincide.

Example 7 (Tense logic)**.** Let us consider the translation by t of the abstract connective $(\chi(x), \mathsf{Q})$, where $\chi(x) \triangleq \forall y(\mathsf{R}yx \rightarrow \mathsf{Q}y)$. This connective corresponds to the box modality $\Box^-$ of tense logic. We obtain that $c \triangleq t(\chi(x), \mathsf{Q}) = \Box_1 \supset_2 ([f^2_2]\mathsf{R}, [f^1_2]id^1_1)$ is a universal molecular connective, where R is a propositional letter of skeleton $(\mathrm{Id}, +, \exists, 2)$, $f^2_2 : [\![1;2]\!] \rightarrow [\![1;2]\!]$ is defined by $f^2_2(1) \triangleq 2$ and $f^2_2(2) = 1$ and $f^1_2 : \{1\} \rightarrow [\![1;2]\!]$ is defined by $f^1_2(1) = 2$.

According to the definition of C–bisimulation, the clause for this molecular connective is the following:

[2] See [6, Definition 3] for a precise definition of abstract connectives.

if wZ_cw' and $-R'_cv'w'$ then there is v such that $-R_cvw$ and vZv'.

(We recall that $Z_{id^1_1} = Z$ in that case.)

The C–bisimulation clause for the atomic connective $\Box^-$ of skeleton $(\tau_1, -, t_2)$ is the following:

if wZ_cw' and $-R'^{-\tau_1}_{\Box}v'w'$ then there is v such that $-R'^{-\tau_1}_{\Box}v'w'$ and vZv', that is,

if wZ_cw' and $R'_{\Box}v'w'$ then there is v such that $R'_{\Box}v'w'$ and vZv'.

So, if we prove that for all w, v we have that R_cvw iff $-R_{\Box}vw$, then we will have proved that the two definitions of bisimulation are equivalent (at least for this bisimulation clause, but the other clause for the $\Diamond$ modality is proved similarly). This is what the following proposition proves.

Proposition 11. *Let $\mathcal{P} = \{\mathsf{R}\}$ and let $M = (W, \mathcal{R})$ be a predicate $\mathsf{C}^{\mathcal{P}}$–model such that the predicate symbol R is associated to the relation $R_{\Box}$. Let c be the universal molecular connective $\Box_1 \supset_2 ([f^2_2]\mathsf{R}, [f^1_2]id^1_1)$. Then, for all $w, v \in W$, we have that R_cvw iff $-R_{\Box}vw$. (We recall that R_c is defined in Definition 16.)* $\Box$

This leads us to a more general problem, to determine whether two equally expressive logics have the same notion of bisimulation.

15 Bisimulations of Equally Expressive Logics are Equivalent

Our definitions of C–bisimulation are defined from the truth conditions of atomic or molecular conectives. However, these truth conditions may take many equivalent forms, in particular if we take two logics which are equally expressive but are not defined on the basis of the same connectives. In that case, we would nevertheless want to obtain the same notion of bisimulation or at least equivalent definitions. That is, we would want that when we consider two equally expressive molecular logics, their notions of bisimulations coincide. To be more precise, we would want that, in that case, if we take two bisimilar models in one logic then the translations of these two models in the second equally expressive logic are also bisimilar. That is what we are going to prove in that section. We first spell out what we mean by "schematically equally expressive" logics.

Definition 36 (Schematic equi–expressivity)**.** Let $\mathsf{L}_1 = (\mathcal{L}_1, \mathcal{E}_1, \Vdash_1)$ and $\mathsf{L}_2 = (\mathcal{L}_2, \mathcal{E}_2, \Vdash_2)$ be two molecular logics such that $\mathcal{L}_1$ and $\mathcal{L}_2$ are built up from the sets of molecular connectives C_1 and C_2. We say that L_1 is *schematically as expressive as* L_2 when they are equally expressive and, moreover, if $\{i, j\} = \{1, 2\}$, the mappings T^{φ}_i and T^M_i are such that:

i. The mapping T^{φ}_i is induced from a translation of the connectives of C_i which respects the quantified decomposition of the connectives. That is, for all molecular connectives $c \in \mathsf{C}_i$ and all vertices c_0 of the quantified decomposition tree of c, the mapping T^{φ}_i defines a molecular connective $T^{\varphi}_i(c_0)$ built up from the molecular connectives of C_j which is existential (resp. universal) if c_0 is existential (resp. universal) and which has the same arity, tonicity signature and id^i_ks as c_0 (which can be equal to some id^i_k). We extend this definition inductively to molecular connectives of C_i as follows: if $c \in \mathsf{C}_i$ is maximally decomposed into $c = c_0(c_1, \ldots, c_n)$ then

$$T^{\varphi}_i(c) \triangleq T^{\varphi}_i(c_0)(T^{\varphi}_i(c_1), \ldots, T^{\varphi}_i(c_n)).$$

Then, we extend inductively the translation to the whole language $\mathcal{L}_i$ as follows. For all $c \in \mathsf{C}_i$ and all $\varphi_1, \ldots, \varphi_m \in \mathcal{L}_i$, we define

$$T_i^{\varphi}(c(\varphi_1, \ldots, \varphi_m)) \triangleq T_i^{\varphi}(c)(T_i^{\varphi}(\varphi_1), \ldots, T_i^{\varphi}(\varphi_m))$$

ii. The mapping T_i^M is induced from a translation of the relations of the C_i–models. That is, for all $M_i, M_i' \in \mathcal{E}_i$ such that $M_i = (M, \overline{w}_i)$ and $M_i' = (M, \overline{w}_i')$, we have that $T_i^M(M_i) = (N, \overline{w}_2)$ and $T_i^M(M_i') = (N, \overline{w}_2')$, *i.e.* they are also based on the same C_j–model. In that case, we write them $T_i^M(\overline{w}_i)$ and $T_i^M(\overline{w}_i')$. Then for all $M_i \in \mathcal{E}_i$, all $\overline{w}_1, \ldots, \overline{w}_n, \overline{w} \in \overline{w}(M_i, \mathsf{C}_i)$, all $c_0 \in \mathsf{C}_i$ and all vertices $c \in V_{c_0}'$ of the quantified decomposition tree of c_0,

$$R_c T_j^M(\overline{w}_1) \ldots T_j^M(\overline{w}_n) T_j^M(\overline{w}) \text{ iff } R_{T_i^{\varphi}(c)} \overline{w}_1 \ldots \overline{w}_n \overline{w} \quad \square$$

Our notion of *schematic* equi–expressivity is a refinement of the notion of equi–expressivity of Definition 1.

Proposition 12. *Let $\mathsf{L}_1 = (\mathcal{L}_1, \mathcal{E}_1, \Vdash_1)$ and $\mathsf{L}_2 = (\mathcal{L}_2, \mathcal{E}_2, \Vdash_2)$ be two molecular logics such that $\mathcal{L}_1$ and $\mathcal{L}_2$ are built up from the sets of molecular connectives C_1 and C_2. If L_1 and L_2 are schematically equally expressive (via some mappings $T_1^M, T_1^{\varphi}, T_2^M, T_2^{\varphi}$) then for all $M_1, M_1' \in \mathcal{E}_1$ which are C_1–bisimilar, $T_1^M(M_1)$ and $T_1^M(M_1')$ are also C_2–bisimilar.*

16 Conclusion

We have introduced a generic method which allows us to find out an appropriate notion of bisimulation for an arbitrary logic whose truth conditions are defined by first–order formulas. This bisimulation notion comes as well with a number of associated model–theoretical results of the logic considered. In doing so, we have discovered new invariance notions for first–order logic. Compared to the original Keisler theorems, our results replace isomorphisms by predicate bisimulations and first–order bisimulations and we generalize existing results and consider arbitrary sets of first–order formulas, and not only the set of all sentences. On countable structures, the notions of isomorphism, partial isomorphism, predicate bisimulation and first–order bisimulation coincide. We expect that our notions of predicate and first–order bisimulation differ from isomorphisms on uncountable structures. By the (upward) Löwenheim–Skolem theorem and because our results hold for countable languages, they open new perspectives for the study of uncountable structures such as the non–standard models of arithmetic.

These generalizations and new versions of existing theorems confirm, together with the rediscovery of numerous existing results [5], the soundness and generic character of our overall approach. We do not claim to have introduced brand new notions of invariance for first-order logic, they are in fact natural and intuitive variants of the usual notions of (partial) isomorphism. Instead, we claim to have introduced a generic notion of C-bisimulation which is somehow 'deeper' and more basic than the usual notion of bisimulation for modal logic or even the usual notion of (partial) isomorphism for first-order logic. Indeed, all invariance notions introduced in the literature (including those) can all be seen as instances of our general notions of C-bisimulation of Definitions 17 and 19.

References

[1] Sergio Abriola, María Emilia Descotte, and Santiago Figueira. Model theory of XPath on data trees. part II: binary bisimulation and definability. *Inf. Comput.*, 255:195–223, 2017.

[2] Gerard Allwein and J. Michael Dunn. Kripke models for linear logic. *The Journal of Symbolic Logic*, 58(2):514–545, June 1993.

[3] Guillaume Aucher. Displaying Updates in Logic. *Journal of Logic and Computation*, 26(6):1865–1912, March 2016.

[4] Guillaume Aucher. *Selected Topics from Contemporary Logics*, chapter Towards Universal Logic: Gaggle Logics, pages 5–73. Landscapes in Logic. College Publications, October 2021.

[5] Guillaume Aucher. Generalized Keisler Theorems for Atomic and Molecular Logics. Research report, Université de Rennes 1, December 2022.

[6] Guillaume Aucher. On the universality of atomic and molecular logics via protologics. *Logica Universalis*, 16(1):285–322, 2022.

[7] Guillaume Aucher, Johan van Benthem, and Davide Grossi. Modal logics of sabotage revisited. *J. Log. Comput.*, 28(2):269–303, 2018.

[8] Jean-Yves Béziau. *Logica Universalis*, chapter From Consequence Operator to Universal Logic: A Survey of General Abstract Logic. Birkhäuser Basel, 2007.

[9] Patrick Blackburn, Maarten de Rijke, and Yde Venema. *Modal Logic*, volume 53 of *Cambridge Tracts in Computer Science*. Cambridge University Press, 2001.

[10] C. C. Chang and H. J. Keisler. *Model Theory*. Studies in Logic and the Foundations of Mathematics. Elsevier, 3rd edition, 1990.

[11] Maarten de Rijke. A note on graded modal logic. *Studia Logica*, 64(2):271–283, 2000.

[12] J Michael Dunn. Gaggle theory: an abstraction of galois connections and residuation, with applications to negation, implication, and various logical operators. In *European Workshop on Logics in Artificial Intelligence*, pages 31–51. Springer Berlin Heidelberg, 1990.

[13] J Michael Dunn. Partial-gaggles applied to logics with restricted structural rules. In Peter Schroeder-Heister and Kosta Dosen, editors, *Substructural Logics*, pages 63–108. Clarendon Press: Oxford, 1993.

[14] Herbert B. Enderton. *An Introduction to Mathematical Logic*. Academic Press, 2001.

[15] Marta García-Matos and Jouko Väänänen. Abstract model theory as a framework for universal logic. In Jean-Yves Béziau, editor, *Logica Universalis*, pages 19–33, Basel, 2007. Birkhäuser Basel.

[16] Rajeev Goré. Substructural logics on display. *Logic Journal of IGPL*, 6(3):451–504, 1998.

[17] Viktor Grishin. On a generalization of the Ajdukiewicz-Lambek system. In A. I. Mikhailov, editor, *Studies in Nonclassical Logics and Formal Systems*, pages 315–334. Nauka, Moscow, 1983.

[18] David Janin and Igor Walukiewicz. On the expressive completeness of the propositional mu-calculus with respect to monadic second order logic. In Ugo Montanari and Vladimiro Sassone, editors, *CONCUR '96, Concurrency Theory, 7th International Conference, Pisa, Italy, August 26-29, 1996, Proceedings*, volume 1119 of *Lecture Notes in Computer Science*, pages 263–277. Springer, 1996.

[19] Natasha Kurtonina and Maarten de Rijke. Bisimulations for temporal logic. *Journal of Logic, Language and Information*, 6(4):403–425, 1997.

[20] Joachim Lambek. The mathematics of sentence structure. *American mathematical monthly*, 65:154–170, 1958.

[21] Maarten Marx and Yde Venema. *Multi-dimensional modal logic*, volume 4 of *Applied logic series*. Kluwer, 1997.

[22] Michael Moortgat. Symmetries in natural language syntax and semantics: the Lambek-Grishin calculus. In *Logic, Language, Information and Computation*, pages 264–284. Springer, 2007.

[23] Till Mossakowski, Razvan Diaconescu, and Andrzej Tarlecki. What is a logic translation? *Logica Universalis*, 3(1):95–124, 2009.

[24] Till Mossakowski, Joseph Goguen, Răzvan Diaconescu, and Andrzej Tarlecki. What is a logic? In Jean-Yves Beziau, editor, *Logica Universalis*, pages 111–133, Basel, 2007. Birkhäuser Basel.

[25] Grigory K. Olkhovikov. On expressive power of basic modal intuitionistic logic as a fragment of classical FOL. *J. Appl. Log.*, 21:57–90, 2017.

[26] Joseph J. Rotman. *An Introduction to the Theory of Groups*, volume 148 of *Graduate texts in mathematics*. Springer, 1995.

[27] Johan van Benthem and Kees Doets. Higher-order logic. In Dov Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume 1, pages 275–329. Reidel Publishing Company, 1983.

A Proofs of Lemmas 1, 2, Theorems 1, 2, 3 and Corollaries 1, 2, 3

Lemma 1. *Let $\varphi \in \mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}$, let $\overline{x}$ be the tuple of free variables and constants of φ (possibly empty) and let (M, s) be a pointed structure. Then, we have that*

$$(M, s) \models \varphi \text{ iff } T_{\overline{x}}(M, s) \Vdash T_{\overline{x}}(\varphi).$$

Proof: (sketch) Similar to the proof of [6, Lemma 1], by induction on φ. We only prove some of the cases which differ significantly. Let $\varphi(x_1, \ldots, x_k) \in \mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(\overline{x}, k)$.

- If $\varphi = \mathsf{R}x_{i_1} \ldots x_{i_l}$ with $k \leq l$ then there is a unique (surjective) mapping $f^l_k : [\![1; l]\!] \to [\![1; k]\!]$ such that for all $j \in [\![1; l]\!]$, we have that $f^l_k(j) = i_j$. Then, $T_{\overline{x}}(M, s) \Vdash T_{\overline{x}}(\varphi)$

  iff $(T(M), (s(x_1), \ldots, s(x_k)) \Vdash [f^l_k]\mathsf{R}$

  iff $(T(M), (s(x_{f^l_k(1)}), \ldots, s(x_{f^l_k(l)}))) \Vdash \mathsf{R}$ by definition of $[f^l_k]$

  iff $(T(M), (s(x_{i_1}), \ldots, s(x_{i_l}))) \Vdash \mathsf{R}$

  iff $Rs(x_{i_1}) \ldots s(x_{i_l})$ holds

  iff $(M, s) \models \mathsf{R}x_{i_1} \ldots x_{i_l}$

  iff $(M, s) \models \varphi$

- If $\varphi = \forall x \psi(x_1, \ldots, x_k, x)$ with $k \geq 1$, then $T_{\overline{x}}(M, s) \Vdash T_{\overline{x}}(\varphi)$
  iff $(T(M), (s(x_1), \ldots, s(x_k))) \Vdash \Box_k T_{(x_1,\ldots,x_k,x)}(\psi(x_1, \ldots, x_k, x))$
  iff for all $w \in W$, $(T(M), (s(x_1), \ldots, s(x_k), w)) \Vdash T_{(x_1,\ldots,x_k,x)}(\psi(x_1, \ldots, x_k, x))$
  iff for all $w \in W$, $T_{(x_1,\ldots,x_k,x)}(M, s[x := w]) \Vdash T_{(x_1,\ldots,x_k,x)}(\psi(x_1, \ldots, x_k, x))$ by definition of $T_{(x_1,\ldots,x_k,x)}$
  iff for all $w \in W$, $(M, s[x := w]) \models \psi(x_1, \ldots, x_k, x)$ by Induction Hypothesis
  iff $(M, s) \models \forall x \psi(x_1, \ldots, x_k, x)$ by definition of $\forall x$
  iff $(M, s) \models \varphi$.

$\Box$

Lemma 2. *Let $(M, \overline{w})$ be a pointed predicate $\mathcal{C}^{\mathcal{P}}$–model of type k, let $\varphi \in \mathcal{L}_{\mathcal{C}^{\mathcal{P}}}$ of type k and let $\overline{x} \in (\mathcal{V} \cup \mathcal{C})^k$. Then,*

$$(M, \overline{w}) \Vdash \varphi \text{ iff } ST_{\overline{x}}(M, \overline{w}) \models ST_{\overline{x}}(\varphi).$$

Moreover, for all pointed structures (M, s) without distinguished elements, we have that $ST_{\overline{x}}(T_{\overline{x}}(M, s)) \equiv_{\overline{x}} (M, s)$ and for all pointed predicate $\mathcal{C}^{\mathcal{P}}$–model $(M, \overline{w})$, we have that $T_{\overline{x}}(ST_{\overline{x}}(M, \overline{w})) \equiv_k (M, \overline{w})$.

Proof: (sketch) Similar to the proof of [6, Lemma 2], by induction on φ. We only prove two cases which differ significantly. Let $\varphi \in \mathcal{L}_{\mathcal{C}^{\mathcal{P}}}$ of type k and let $\overline{x} = (x_1, \ldots, x_k) \in (\mathcal{V} \cup \mathcal{C})^k$.

- If $\varphi = [f_k^l]\psi$ then $ST_{\overline{x}}(M, \overline{w}) \models ST_{\overline{x}}(\varphi)$
  iff $ST_{\overline{x}}(M, \overline{w}) \models ST_{\overline{x}}([f_k^l]\psi)$
  iff $(ST(M), s_{\overline{x}}^{\overline{w}}) \models ST_{(x_{f_k^l(1)},\ldots,x_{f_k^l(l)})}(\psi)$ by definition of $ST_{\overline{x}}$
  iff $(ST(M), s_{\overline{y}}^{\overline{v}}) \models ST_{\overline{y}}(\psi)$ where $\overline{y} = (x_{f_k^l(1)}, \ldots, x_{f_k^l(l)})$ and $\overline{v} = (w_{f_k^l(1)}, \ldots, w_{f_k^l(l)})$
  iff $(M, \overline{v}) \Vdash \psi$ by Induction Hypothesis
  iff $(M, \overline{w}) \Vdash [f_k^l]\psi$ because we have that $R_{[f_k^l]}\overline{v}\overline{w}$
  iff $(M, \overline{w}) \Vdash \varphi$.
- If $\varphi = \Box_k \psi$ then $ST_{\overline{x}}(M, \overline{w}) \models ST_{\overline{x}}(\varphi)$
  iff $ST_{\overline{x}}(M, \overline{w}) \models ST_{\overline{x}}(\Box_k \psi)$
  iff $(ST(M), s_{\overline{x}}^{\overline{w}}) \models \forall x ST_{(\overline{x},x)}(\psi)$ by definition
  iff for all $w \in W$, $(ST(M), s_{\overline{x}}^{\overline{w}}[x := w]) \models ST_{(\overline{x},x)}(\psi)$
  iff for all $w \in W$, $(ST(M), s_{(\overline{x},x)}^{(\overline{w},w)}) \models ST_{(\overline{x},x)}(\psi)$ by definition
  iff for all $w \in W$, $(M, (\overline{w}, w)) \Vdash \psi$ by Induction Hypothesis
  iff $(M, \overline{w}) \Vdash \Box_k \psi$
  iff $(M, \overline{w}) \Vdash \varphi$.

$\Box$

Lemma 5. *Let $\mathcal{C}^{\mathcal{P}}$ be the set of predicate atomic connectives (introduced in Definition 25). A $\mathcal{C}^{\mathcal{P}}$–bisimulation between two $\mathcal{C}^{\mathcal{P}}$–models is a predicate bisimulation between their associated structures (defined in Definition 28), but which does not necessarily fulfill condition 5 of Definition 29.*

Proof: Let us consider the predicate atomic connectives $\mathsf{C}^{\mathcal{P}} = \mathcal{P} \cup \{\bot, U\} \cup \{\|_k, \supset_k, \Box_k \mid k \in \mathbb{N}^*\} \cup \{[f_k^l] \mid k, l \in \mathbb{N}^* \text{ and } f_k^l : [\![1; l]\!] \to [\![1; k]\!] \text{ is surjective}\}$ (defined in Definition 25). Let M_1 and M_2 be two $\mathsf{C}^{\mathcal{P}}$–models and let Z be a non-empty binary relation between finite sequences of M_1 and M_2 of the same length. Then, by Definition 17, Z is a $\mathsf{C}^{\mathcal{P}}$–bisimulation between M_1 and M_2 when for all $M, M' \in \{M_1, M_2\}$ with $M = (W, \mathcal{R})$ and $M' = (W', \mathcal{R}')$, all $\overline{w}, \overline{v}, \overline{u} \in \overline{w}(M, \mathsf{C})$, all $w, v \in W$ and all $\overline{w'}, \overline{v'}, \overline{u'} \in \overline{w}(M', \mathsf{C})$, all $w', v' \in W'$, all $\mathsf{R} \in \mathcal{P}$, it holds that

- condition for R: if $\overline{w} Z \overline{w'}$ and $R\overline{w}$ then $R'\overline{w'}$, for all relations R and R' both associated to R (in the associated structure);
- condition for $\bot$: if wZw' and $R'_{\bot}(w')$ then $R_{\bot}(w)$, this condition always holds;
- condition for $[f_k^l]$: for all $k, l \in \mathbb{N}^*$ and all surjections f_k^l, if $\overline{w} Z \overline{w'}$ and $R'_{[f_k^l]}\overline{w'v'}$ then there is $\overline{v} \in \overline{w}(M, \mathsf{C})$ such that $\overline{v} Z \overline{v'}$ and $R_{[f_k^l]}\overline{wv}$,

  that is, if $\overline{w} Z \overline{w'}$ then $f_k^l(\overline{w}) Z f_k^l(\overline{w'})$, where for all $\overline{w} = (w_1, \ldots, w_k)$, $f_k^l(\overline{w}) \triangleq (w_{f_k^l(1)}, \ldots, w_{f_k^l(l)})$;
- condition for $\supset_k$: for all $k \in \mathbb{N}^*$, if $\overline{w} Z \overline{w'}$ and $R_k\overline{v'u'w'}$ then there are $\overline{v}, \overline{u} \in \overline{w}(M, \mathsf{C})$ such that $\overline{v'} Z \overline{v}$, $\overline{u} Z \overline{u'}$ and $R_k\overline{vuw}$,

  that is, if $\overline{w} Z \overline{w'}$ then $\overline{w'} Z \overline{w}$;
- condition for $\Box_k$: for all $k \in \mathbb{N}^*$, if $\overline{w} Z \overline{w'}$ and $R'_{k,1}\overline{w'}v'\overline{u'}$ then there are $v, \overline{u} \in \overline{w}(M, \mathsf{C})$ such that vZv' and $\overline{u} Z \overline{u'}$ and $R_{k,1}\overline{w}v\overline{u}$,

  that is, if $\overline{w} Z \overline{w'}$ then for all $v' \in W'$ there is $v \in W$ such that vZv' and $(\overline{w}, v) Z (\overline{w'}, v')$;
- condition for U: if wZw' and $v' \in W'$ then there is $v \in W$ such that vZv',

  that is, for all $w' \in W'$ there is $w \in W$ such that wZw'.

Now, if we combine the conditions above, we obtain the definition of a predicate bisimulation on the structures associated to M_1 and M_2 without condition 5. $\Box$

Theorem 1. *Let $X \subseteq \mathcal{V} \cup \mathcal{C}$ be non-empty and assume that $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$ is countable. Let (M, s) and (M', s') be two pointed structures. Then the following are equivalent:*

1. *(M, s) and (M', s') make true the same formulas of $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$;*
2. *there exists an ultrafilter U and a X–compatible predicate bisimulation between $\prod_U (M, s)$ and $\prod_U (M', s')$.*

Proof: The proof follows the same reasoning as the proof of [5, Theorem 1], using [5, Lemma 1] and [5, Proposition 9, Appendix] and [5, Proposition 11] as well as the 'translation' Lemmas 1 and 2.

First, we prove that 2. implies 1. By [5, Proposition 9, Appendix], for all formulas $\varphi \in \mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$, we have that $(M, s) \Vdash \varphi$ iff $\prod_U (M, s) \Vdash \varphi$. By assumption and Proposition 9, this implies that $\prod_U (M', s') \Vdash \varphi$ and, again by [5, Proposition 9, Appendix], the latter is equivalent to $(M', s') \Vdash \varphi$.

Now, we prove that 1. implies 2. Let U be a countably incomplete ultrafilter over $\mathbb{N}$ (it exists by [9, Example 2.72]). By [5, Proposition 11], the ultrapowers $\prod_U M$ and $\prod_U M'$ are ω–saturated and therefore $T\left(\prod_U M\right)$ and $T\left(\prod_U M'\right)$ are ω–saturated as well (for the first-order language $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}$). Now, for all formula $\varphi(\overline{x}) \in \mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$, $\prod_U (M,s) \Vdash \varphi(\overline{x})$ implies $\prod_U (M',s') \Vdash \varphi(\overline{x})$ because of 1. and [5, Proposition 9, Appendix]. That is, for all formula $\varphi(\overline{x}) \in \mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$, $\left(\prod_U M, \prod_U s\right) \Vdash \varphi(\overline{x})$ implies $\left(\prod_U M', \prod_U s'\right) \Vdash \varphi(\overline{x})$. Therefore, by Lemma 1, for all formula $\varphi(\overline{x}) \in \mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$, we have that $T_{\overline{x}}\left(\prod_U M, \prod_U s\right) \Vdash T_{\overline{x}}(\varphi)$ implies $T_{\overline{x}}\left(\prod_U M', \prod_U s'\right) \Vdash T_{\overline{x}}(\varphi)$ (if φ is a sentence without constant, we take in Definition 27 of $T_\emptyset$ a variable x_0 such that $x_0 \in X$). That is, for all formula $\varphi \in \mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$, we have that $\left(T\left(\prod_U M\right), \left(\prod_U s(x_1), \ldots, \prod_U s(x_k)\right)\right) \Vdash T_{\overline{x}}(\varphi)$ implies $\left(T\left(\prod_U M'\right), \left(\prod_U s'(x_1), \ldots, \prod_U s'(x_k)\right)\right) \Vdash T_{\overline{x}}(\varphi)$ $(*)$. Then, using Lemmas 1 and 2, we can prove the following:

Claim 1. *For all* $x_1, \ldots, x_k \in X$, $\left(T\left(\prod_U M\right), \left(\prod_U s(x_1), \ldots, \prod_U s(x_k)\right)\right) \rightsquigarrow_{\mathsf{C}^{\mathcal{P}}} \left(T\left(\prod_U M'\right), \left(\prod_U s'(x_1), \ldots, \prod_U s'(x_k)\right)\right)$.

Proof: (of the Claim) Let $\varphi \in \mathcal{L}_{\mathsf{C}^{\mathcal{P}}}$ of type $k \in \mathbb{N}^*$ and let $x_1, \ldots, x_k \in X$. Assume that

$$\left(T\left(\prod_U M\right), \left(\prod_U s(x_1), \ldots, \prod_U s(x_k)\right)\right) \Vdash \varphi.$$

Then, by Lemma 2,

$$ST_{\overline{x}}\left(T\left(\prod_U M\right), \left(\prod_U s(x_1), \ldots, \prod_U s(x_k)\right)\right) \Vdash ST_{\overline{x}}(\varphi).$$

That is,

$$\left(ST\left(T\left(\prod_U M\right)\right), s_{\overline{x}}^{(\prod_U s(x_1), \ldots, \prod_U s(x_k))}\right) \Vdash ST_{\overline{x}}(\varphi).$$

So, by Lemma 1,

$$\left(T\left(ST\left(T\left(\prod_U M\right)\right)\right), \left(\prod_U s(x_1), \ldots, \prod_U s(x_k)\right)\right) \Vdash T_{\overline{x}}\left(ST_{\overline{x}}(\varphi)\right).$$

That is, again by Lemma 2,

$$\left(T\left(\prod_U M\right), \left(\prod_U s(x_1), \ldots, \prod_U s(x_k)\right)\right) \Vdash T_{\overline{x}}\left(ST_{\overline{x}}(\varphi)\right).$$

So, by $(*)$, we have that

$$\left(T\left(\prod_U M'\right), \left(\prod_U s'(x_1), \ldots, \prod_U s'(x_k)\right)\right) \Vdash T_{\overline{x}}\left(ST_{\overline{x}}(\varphi)\right).$$

That is, again by Lemma 2,

$$\left(T\left(ST\left(T\left(\prod_U M'\right)\right)\right), \left(\prod_U s'(x_1), \ldots, \prod_U s'(x_k)\right)\right) \Vdash T_{\overline{x}}\left(ST_{\overline{x}}(\varphi)\right).$$

Then, by Lemma 1,

$$\left(ST\left(T\left(\prod_U M'\right)\right), s_{\overline{x}}^{(\prod_U s'(x_1), \ldots, \prod_U s'(x_k))}\right) \Vdash ST_{\overline{x}}(\varphi).$$

Thus, by Lemma 2,

$$\left(T\left(\prod_U M'\right), \left(\prod_U s'(x_1), \ldots, \prod_U s'(x_k)\right)\right) \Vdash \varphi.$$

and we have proved the claim. □

Hence, by [5, Lemma 1] and Lemma 5, there is a X–compatible predicate bisimulation between the structures associated to $\left(T\left(\prod_U M\right), \prod_U s\right)$ and $\left(T\left(\prod_U M'\right), \prod_U s\right)$ and therefore between $\left(\prod_U M, \prod_U s\right)$ and $\left(\prod_U M', \prod_U s\right)$, that is, between $\prod_U (M, s)$ and $\prod_U (M', s')$. □

Corollary 1. *Assume that $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}$ is countable and let (M, s) and (M', s') be two pointed structures. Then the following are equivalent:*

1. *(M, s) and (M', s') make true the same sentences of $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}$;*
2. *there exists an ultrafilter U and a predicate bisimulation between $\prod_U (M, s)$ and $\prod_U (M', s')$.*

Proof: The set of sentences is $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(\mathcal{C})$. The result follows by a direct application of Theorem 1. □

Theorem 2. *Let $X \subseteq \mathcal{V} \cup \mathcal{C}$ be non-empty and assume that $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$ is countable. Let K and K' be classes of pointed structures such that K' is closed under ultraproduct. Then, the following are equivalent:*

1. *K is definable in K' by a set of formulas of $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$;*
2. *K is closed under X–compatible predicate bisimulations in K' and closed under ultraproducts, and $K' - K$ is closed under ultrapowers.*

Proof: The proof follows the same reasoning as the proof of [5, Theorem 2]. The notion of X–compatibility is used in the direction from 1. to 2. As for the direction from 2. to 1., we consider the following set of pure predicate formulas:

$$T \triangleq \left\{\varphi \in \mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X) \mid \text{ for all } (M,s) \in K, (M,s) \models \varphi\right\}$$

Let $(M_0, s_0) \in K'$ be a pointed structure of K' such that $(M_0, s_0) \Vdash T$. We are going to prove that $(M_0, s_0) \in K$. Let us consider the following set of formulas:

$$\Sigma \triangleq \{\varphi \in \mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X) \mid (M_0, s_0) \models \varphi\}$$

Σ is finitely satisfiable in K. Indeed, assume that the finite set $\{\psi_1, \ldots, \psi_n\} \subseteq \Sigma$ is not satisfiable in K. Then, $\neg\psi_1 \vee \ldots \vee \neg\psi_n$ is true on all pointed structures of K. So, $\neg\psi_1 \vee \ldots \vee \neg\psi_n \in T$. However, it would be false on (M_0, s_0), which is impossible. But then, [5, Proposition 10] shows that Σ is satisfiable in an ultraproduct $\prod_U (N_i, s_i)$ of pointed structures $(N_i, s_i) \in K$. Let us take $(N, s) = \prod_U (N_i, s_i)$. Then, $(N, s) \in K$ by closure of K under ultraproduct. Moreover, $(N, s) \models \Sigma$. So, for all $\varphi \in \mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$, $(M_0, s_0) \models \varphi$ implies $(N, s) \models \varphi$. Thus, by closure of $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$ under Boolean negation, for all $\varphi \in \mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$, $(M_0, s_0) \models \varphi$ iff $(N, s) \models \varphi$. Thus, by Theorem 1, there exists an ultrafilter U over a non-empty set I and a X–compatible predicate bisimulation between $\prod_U (M_0, s_0)$ and $\prod_U (N, s)$. Now, $\prod_U (N, s) \in K$ because $(N, s) \in K$ and closure of K under ultraproduct. But $\prod_U (M_0, s_0) \in K'$ by closure of K' under ultraproducts. So, $\prod_U (M_0, s_0) \in K$ by closure of K under X–compatible bisimulation in K'. Finally, since $K' - K$ is closed under ultrapower, (M_0, s_0) must belong to K, since otherwise $\prod_U (M_0, s_0)$ would not be in K. This completes the proof of the theorem. □

Corollary 2. *Let K be a class of pointed structures. Then, the following are equivalent:*

1. *K is definable by a set of sentences of $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}$;*
2. *K is closed under predicate bisimulations and ultraproducts, and $\overline{K}$ is closed under ultrapowers.*

Proof: The set of sentences is $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(\mathcal{C})$. The result follows from a direct application of Theorem 2 with $X = \mathcal{C}$ and $K' = \mathcal{M}_{\mathsf{FOL}}$. □

Theorem 3. *Let $X \subseteq \mathcal{V} \cup \mathcal{C}$ be non-empty and assume that $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$ is countable. Let K and K' be classes of pointed structures such that K' is closed under ultraproducts. Then, the following are equivalent:*

1. *K is definable in K' by means of a single formula $\varphi \in \mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$;*
2. *Both K and $K' - K$ are closed under X–compatible predicate bisimulations in K' and closed under ultraproducts.*

Proof: It is similar to the proof of [5, Theorem 3]. The direction from 1. to 2. is easy. For the converse, we assume that K and $K' - K$ satisfy the stated closure conditions. Hence by Theorem 2, there are sets of formulas of $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$, T_1 and T_2, defining K and $K' - K$ respectively.

Obviously, their union is inconsistent in the sense that there is no pointed structure (M, s) of K' such that $(M, s) \Vdash T_1 \cup T_2$. So then, by [5, Proposition 10], there exist $\varphi_1, \ldots \varphi_n \in T_1$ and $\psi_1, \ldots, \psi_m \in T_2$ such that for all pointed structures (M, s) of K', it is not the case that $(M, s) \Vdash \varphi_1 \wedge \ldots \wedge \varphi_n \wedge \psi_1 \wedge \ldots \wedge \psi_m$ $(*)$. To complete the proof, we show that K is in fact defined in K' by the conjunction $\varphi_1 \wedge \ldots \wedge \varphi_n$. By definition, for any (M, s) in K we have $(M, s) \Vdash \varphi_1 \wedge \ldots \wedge \varphi_n$. Conversely, for all (M, s) of K', if $(M, s) \Vdash \varphi_1 \wedge \ldots \wedge \varphi_n$ then there must be $i \in [\![1; m]\!]$ such that $(M, s) \Vdash \psi_i$ does not hold. Indeed, otherwise, we would have that $(M, s) \Vdash \varphi_1 \wedge \ldots \wedge \varphi_n \wedge \psi_1 \wedge \ldots \wedge \psi_m$, which is impossible by $(*)$. Hence, it is not the case that $(M, s) \Vdash T_2$. Therefore, (M, s) does not belong to $K' - K$, whence (M, s) belongs to K. □

Corollary 3. *Let K be a class of pointed structures. Then, the following are equivalent:*

1. *K is definable by means of a sentence of $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}$;*
2. *Both K and $\overline{K}$ are closed (in $\mathcal{M}_{\mathsf{FOL}}$) under predicate bisimulations and ultraproducts.*

Proof: The set of sentences is $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(\mathcal{C})$. The result follows from a direct application of Theorem 3 with $X = \mathcal{C}$ and $K' = \mathcal{M}_{\mathsf{FOL}}$. □

B Proof of Theorem 4

We recall that $\mathcal{P}$ is a set of predicate symbols, $\mathcal{F}$ is a set of function symbols, $\mathcal{V}$ is a set of variables and $\mathcal{C}$ is a set of constants (see Section 2). We also define the *head* $h(t)$ and the *body* $b(t)$ of a term $t \in \mathcal{T}$ inductively as follows: if $t = x$ or $t = \mathsf{c}$ then $h(t) = b(t) = t$; if $t = f(t_1, \ldots, t_m)$ then $h(t) = f$ and $b(t) = (t_1, \ldots, t_m)$.

Definition 37 (First–order atomic connectives)**.** The set of *propositional letter atomic connectives* is $\mathsf{C}^{\mathcal{F}} \triangleq \{p_{\mathsf{R}} \mid \mathsf{R} \in \mathcal{P}\} \cup \{[h_1, \ldots, h_m] \mid k \in \mathrm{Arity}(\mathcal{P}, \mathcal{F}) \text{ and } h_1, \ldots, h_m \in \mathcal{V} \cup \mathcal{C} \cup \mathcal{F}\} \cup \{[f^l_k] \mid k, l \in \mathbb{N}^* \text{ and } f^l_k : [\![1; l]\!] \to [\![1; k]\!] \text{ is a surjection}\}$ where:

- p_{R} has skeleton $(\mathrm{Id}, +, \forall, k)$ with k the arity of $\mathsf{R} \in \mathcal{P}$;
- $[h_1, \ldots, h_m]$ has skeleton $(\mathrm{Id}, -, \forall, (n_1 + \ldots + n_m, m), +)$ in which n_i is the arity of the function h_i (it is 1 if $h_i \in \mathcal{V} \cup \mathcal{C}$);
- $[f^l_k]$ has skeleton $(\mathrm{Id}, -, \forall, (k, l), +)$, for all $k, l \in \mathbb{N}^*$ (such that $l \geq k$).

The set of *first–order atomic connectives* $\mathsf{C}^{\mathcal{PF}}$ is $\mathsf{C}^{\mathcal{F}}$ to which we add $\mathsf{C}^{\mathcal{P}}$ and the propositional letters $\{p_\chi \mid \chi \in \mathcal{L}_{\mathsf{C}^{\mathcal{F}}}\}$ of skeletons $(1, +, \forall, k)$, where k is the type of $\chi \in \mathcal{L}_{\mathsf{C}^{\mathcal{F}}}$.

Definition 38 (First–order atomic logic)**.** A *first–order $\mathsf{C}^{\mathcal{PF}}$–model* is a $\mathsf{C}^{\mathcal{PF}}$–model $M = (W, \mathcal{R})$ such that:

- the relations associated to the connectives of $\mathsf{C}^{\mathcal{P}}$ satisfy the conditions of Definition 26;
- for all $\mathsf{R} \in \mathcal{P}$ of arity n, the connectives p_{R} are associated to n–ary relations $R_{p_{\mathsf{R}}}$;
- each connective $[h_1, \ldots, h_m]$ is associated to a relation $R_{[h_1, \ldots, h_m]}$ over W^l where $l = m + n_1 + \ldots + n_m$ with n_i the arity of the function h_i (it is 1 if $n_i \in \mathcal{V} \cup \mathcal{C}$);
- for all $k, l \in \mathbb{N}^*$, the connective $[f^l_k]$ is associated to a $k + l$–ary relation $R_{[f^l_k]}$ such that for all $w_1, \ldots, w_k, v_1, \ldots, v_l \in W$, it holds that $R_{[f^l_k]} v_1 \ldots v_l w_1 \ldots w_k$ iff for all $j \in [\![1; l]\!]$, we have that $v_j = w_{\sigma^l_k(j)}$;

- for all $\chi \in \mathcal{L}_{\mathsf{C}^{\mathcal{F}}}$, the connectives p_χ are associated to k–ary relations R_{p_χ}, where k is the type of χ.

The class of all pointed first–order $\mathsf{C}^{\mathcal{PF}}$–models is denoted $\mathcal{M}_{\mathsf{C}^{\mathcal{PF}}}$. We define the two–tiered language $\mathcal{L}_{\mathsf{C}^{\mathcal{PF}}}$ as follows:

$$\begin{array}{rcccccc} \mathcal{L}_{\mathsf{C}^{\mathcal{F}}}: & \chi & ::= & p_{\mathsf{R}} & \mid & [h_1,\ldots,h_m]\,\chi & \mid & [f_k^l]\chi \\ \mathcal{L}_{\mathsf{C}^{\mathcal{PF}}}: & \varphi & ::= & p_\chi & \mid & \star\,(\varphi,\ldots,\varphi) & & \end{array}$$

where $\mathsf{R} \in \mathcal{P}$, $[h_1,\ldots,h_m], [f_k^l] \in \mathsf{C}^{\mathcal{F}}$, $\star \in \mathsf{C}^{\mathcal{PF}}$ (different from propositional letters) and χ is of type m in the expression $[h_1,\ldots,h_m]\,\chi$. The satisfaction relation $\Vdash\ \subseteq \mathcal{M}_{\mathsf{C}^{\mathcal{PF}}} \times \mathcal{L}_{\mathsf{C}^{\mathcal{PF}}}$ is then defined following Definition 5. If $\mathcal{E}_{\mathsf{C}^{\mathcal{PF}}}$ is a specific class of abstract $\mathsf{C}^{\mathcal{PF}}$–models, the triple $(\mathcal{L}_{\mathsf{C}^{\mathcal{PF}}}, \mathcal{E}_{\mathsf{C}^{\mathcal{PF}}}, \Vdash)$ is called the *first–order atomic logic associated to* $\mathcal{E}_{\mathsf{C}^{\mathcal{PF}}}$. The triple $(\mathcal{L}_{\mathsf{C}^{\mathcal{PF}}}, \mathcal{M}_{\mathsf{C}^{\mathcal{PF}}}, \Vdash)$ is called *first–order atomic logic.*

We also define $\mathcal{L}_{\mathsf{C}^{\mathcal{PF}}}(k) \triangleq \{\varphi \in \mathcal{L}_{\mathsf{C}^{\mathcal{PF}}} \mid k(\varphi) = k\}$ and for all $(M,\overline{w}), (N,\overline{v}) \in \mathcal{M}_{\mathsf{C}^{\mathcal{PF}}}$, we write $(M,\overline{w}) \equiv_{\mathcal{F},k} (N,\overline{v})$ when for all $\varphi \in \mathcal{L}_{\mathsf{C}^{\mathcal{PF}}}(k)$ it holds that $(M,\overline{w}) \Vdash \varphi$ iff $(N,\overline{v}) \Vdash \varphi$.

Definition 39 (Translation from FOL to first–order atomic logic)**.**
Syntax. For all $\overline{x} = (x_1,\ldots,x_k) \in (\mathcal{V} \cup \mathcal{C})^k$, we define the mapping $T^+_{\overline{x}} : \mathcal{L}^{\mathcal{PF}}_{\mathsf{FOL}}(\overline{x},k) \to \mathcal{L}_{\mathsf{C}^{\mathcal{PF}}}$ inductively on the formula $\varphi \in \mathcal{L}^{\mathcal{PF}}_{\mathsf{FOL}}(\overline{x},k)$ as follows.

- if $\varphi = \mathsf{R}(t_1,\ldots,t_m)$ and φ contains k different variables or constants $\overline{x}$ which occur at l different places in φ then we define $T^+_{\overline{x}}(\varphi) \triangleq [f_k^l]T^+_{\overline{x}}(t_1,\ldots,t_m)p_{\mathsf{R}}$ where $[f_k^l]$ is defined as follows:
  - if $(x_{i_1},\ldots,x_{i_l})$ is the tuple of size l made up of the k variables or constants $\overline{x} = (x_1,\ldots,x_k)$ which occur in $(t_1,\ldots,t_m)$ in that order (and possibly with repetition) then $f_k^l : [\![1;l]\!] \to [\![1;k]\!]$ is defined for all $j \in [\![1;l]\!]$ by $f_k^l(j) = i_j$.

  and where $T^+_{\overline{x}}(t_1,\ldots,t_m)$ is defined inductively as follows:
  - if $(t_1,\ldots,t_m)$ is a tuple of only variables and constants then $T^+_{\overline{x}}(t_1,\ldots,t_m) \triangleq [t_1,\ldots,t_m]$;
  - otherwise, if one of the terms t_i contains a function symbol f_i, then $T^+_{\overline{x}}(t_1,\ldots,t_m) \triangleq T^+_{\overline{x}}\,(b(t_1),\ldots,b(t_m))\,[h(t_1),\ldots,h(t_m)]$.

  Note that $T^+_{\overline{x}}(\varphi)$ belongs to $\mathcal{L}_{\mathsf{C}^{\mathcal{F}}}$.
- if φ is of one of the other forms then the translation $T^+_{\overline{x}}$ is the same as $T_{\overline{x}}$.

Semantics. Let $M = ((W, \{R_1,\ldots,R_n,\ldots,f_1,\ldots,f_n,\ldots\}), s)$ be a pointed structure. We define the $\mathsf{C}^{\mathcal{PF}}$–model $T^+(M) \triangleq (W,\mathcal{R})$ as follows:

- the n–ary relation $R_{p_{\mathsf{R}}}$ is the n–ary relation R associated to $\mathsf{R} \in \mathcal{P}$;
- for all $h_1,\ldots,h_m \in \mathcal{V} \cup \mathcal{C} \cup \mathcal{F}$, the connectives $[h_1,\ldots,h_m] \in \mathsf{C}^{\mathcal{F}}$ are associated to the relation $R_{[h_1,\ldots,h_m]} \triangleq \{(v_1,\ldots,v_m,\overline{w}_1,\ldots,\overline{w}_m) \mid$ for all $i \in [\![1;k]\!]$, if $h_i = f_i$ then $f_i(\overline{w}_i) = v_i$, if $h_i \in \mathcal{V} \cup \mathcal{C}$ then $\overline{w}_i = v_i\}$;
- the other relations of $\mathcal{R}$ for the other connectives are defined like in Definition 38.

If $\overline{x} = (x_1,\ldots,x_k) \in (\mathcal{V} \cup \mathcal{C})^k$ then we define $T^+_{\overline{x}}(M,s) \triangleq \big(T^+(M), (s(x_1),\ldots,s(x_k))\big)$.

Lemma 6. *Let $k \in \mathbb{N}^*$, let $\varphi = \varphi(x_1, \ldots, x_k) \in \mathcal{L}^{\mathcal{PF}}_{\mathsf{FOL}}$ and let (M, s) be a pointed structure. Then we have that*

$$(M, s) \models \varphi \text{ iff } T^+_{\overline{x}}(M, s) \Vdash T^+_{\overline{x}}(\varphi).$$

Proof: The proof follows the same reasoning as the proof of Lemma 1. The only real new case is for $\varphi = \mathsf{R}(t_1, \ldots, t_n)$. It follows from the truth condition for that formula: we have that $(M, s) \models \mathsf{R}(t_1, \ldots, t_n)$ iff $R(w_1, \ldots, w_n)$ and $w_1 = \overline{s}(t_1), \ldots, w_n = \overline{s}(t_n)$. The condition $R(w_1, \ldots, w_n)$ is captured by p_{R} and the condition $w_1 = \overline{s}(t_1), \ldots, w_n = \overline{s}(t_n)$ is captured by the composition of appropriate connectives $[h_1, \ldots, h_m]$. □

Definition 40 (Translation from first–order atomic logic to FOL)**.**
Syntax. For all $n \in \mathbb{N}^*$ and all tuples $\overline{t} = (\overline{t}_1, \ldots, \overline{t}_m)$ of tuples of terms such that each tuple of terms $\overline{t}_i$ is of size n_i, we define the mappings $ST^+_{\overline{t}} : \mathcal{L}^k_{\mathsf{C}^{\mathcal{F}}} \to \mathcal{L}^{\mathcal{PF}}_{\mathsf{FOL}}$, where $\mathcal{L}^k_{\mathsf{C}^{\mathcal{F}}}$ is the set of formulas of $\mathcal{L}_{\mathsf{C}^{\mathcal{F}}}$ of type k, inductively as follows. For all $[h_1, \ldots, h_m]$ of type $(n_1 + \ldots + n_m, m)$, where $n_1, \ldots, n_m$ are the arities of $h_1, \ldots, h_m$, we define

$$\begin{array}{rcl} ST^+_{\overline{t}}(p_{\mathsf{R}}) & \triangleq & \mathsf{R}\overline{t} \\ ST^+_{\overline{t}}([h_1, \ldots, h_m]\,\chi) & \triangleq & ST^+_{(h_1(\overline{t}_1), \ldots, h_m(\overline{t}_m))}(\chi). \end{array}$$

Then, we define the mapping $ST^+_{\overline{t}} : \mathcal{L}_{\mathsf{C}^{\mathcal{PF}}} \to \mathcal{L}^{\mathcal{PF}}_{\mathsf{FOL}}$ inductively as follows.

$$ST^+_{\overline{t}}(p_\chi) \quad \triangleq \quad ST^+_{\overline{t}}(\chi).$$

where $\overline{t}$ is an appropriate tuple of tuples of terms. The clauses for the other connectives of $\mathsf{C}^{\mathcal{PF}}$ are defined like in [6, Definition 20], tuples of variables $\overline{x}$ only have to be replaced by tuples of terms $\overline{t}$.

Semantics. Let $(M, \overline{w})$ be a pointed $\mathsf{C}^{\mathcal{PF}}$–model of type k and let $\overline{x}$ be a tuple of free variables or constants of size k. The *pointed structure associated to* $(M, \overline{w})$, denoted $ST^+_{\overline{x}}(M, \overline{w}) \triangleq (ST^+(M), s^{\overline{w}}_{\overline{x}})$, is defined as follows. The assignment $s^{\overline{w}}_{\overline{x}}$ is defined like in [6, Definition 20] and $ST^+(M)$ is M to which we remove the relations of the form $R_{[f^l_k]}, R_{k_1}, R_{\Box_k}$ and $R_{\forall_0}$ and we replace relations of the form $R_{[h_1, \ldots, h_n]}$ with functions f defined as follows: $f(w_1, \ldots, w_k) = w$ iff $R_{[f]}ww_1, \ldots w_k$.

Lemma 7. *Let $(M, \overline{w})$ be a pointed first–order $\mathsf{C}^{\mathcal{PF}}$–model, let $\varphi \in \mathcal{L}_{\mathsf{C}^{\mathcal{PF}}}$ of type k and let $\overline{x}$ be a tuple of k variables. Then, we have that*

$$(M, \overline{w}) \Vdash \varphi \text{ iff } ST^+_{\overline{x}}(M, \overline{w}) \Vdash ST^+_{\overline{x}}(\varphi)$$

Moreover, for all pointed structures (M, s), we have that $ST^+_{\overline{x}}\left(T^+_{\overline{x}}(M, s)\right) \equiv_{\mathcal{F}, \overline{x}} (M, s)$ and for all pointed $\mathsf{C}^{\mathcal{PF}}$–model $(M, \overline{w})$ of type k, we have that $T^+_{\overline{x}}\left(ST^+_{\overline{x}}(M, \overline{w})\right) \equiv_{\mathcal{F}, k} (M, \overline{w})$.

Proof: (Sketch) By induction on φ. The tricky case is when φ is of the form $[h_1, \ldots, h_n]\,\chi$. This is where we need to index the translation by a tuple of terms. These terms keep track of the terms or subterms of the formula which is being translated. □

Lemma 8. *Let $\mathsf{C}^{\mathcal{PF}}$ be the set of first–order atomic connectives and let $M_1 = (W_1, \mathcal{R}_1)$ and $M_2 = (W_2, \mathcal{R}_2)$ be two ω–saturated $\mathsf{C}^{\mathcal{PF}}$–models. We define the binary relation $Z_0 \subseteq (W_1 \times W_2) \cup (W_2 \times W_1)$ as follows: if $\{M, M'\} = \{M_1, M_2\}$ then for all $\overline{w} \in \overline{w}(M, \mathsf{C})$, all $\overline{w'} \in \overline{w}(M', \mathsf{C})$, $\overline{w} Z_0 \overline{w'}$ iff $(M, \overline{w}) \rightsquigarrow_{\mathsf{C}^{\mathcal{F}}} (M', \overline{w'})$. We also define the binary relation Z between finite sequences of M_1 and M_2 of the same length by $\overline{w} Z \overline{w'}$ iff $(M, \overline{w}) \rightsquigarrow_{\mathsf{C}^{\mathcal{PF}}} (M', \overline{w'})$. Then, the pair of binary relations (Z, Z_0) is a first–order bisimulation between M_1 and M_2.*

Proof: Applying [5, Lemma 1] to $\mathsf{C}^{\mathcal{F}}$, from the definition of a $\mathsf{C}^{\mathcal{F}}$–bisimulation, we obtain conditions 1. – 2. of a first–order bisimulation. Likewise, by [5, Lemma 1] and Lemma 5, Z is a predicate bisimulation. Hence, conditions 4. – 7. of a first–order bisimulation are fulfilled. Moreover, we also have the condition that if $(w_1, \ldots, w_n)Z(w'_1, \ldots, w'_n)$ then for all n–ary relations $R_{p_\chi} \in \mathcal{R}$ and $R'_{p_\chi} \in \mathcal{R}'$ associated to the same propositional letter p_χ, if $R_{p_\chi} w_1 \ldots w_n$ then $R'_{p_\chi} w'_1 \ldots w'_n$. This last condition reformulates as if $\overline{w} Z \overline{w'}$ then $(M, \overline{w}) \rightsquigarrow_{\mathsf{C}^{\mathcal{F}}} (M', \overline{w'})$. That is, if $\overline{w} Z \overline{w'}$ then $\overline{w} Z_0 \overline{w'}$, which is condition 3. Hence, (Z, Z_0) is a first–order bisimulation between M_1 and M_2. □

Theorem 4. *Theorems 1, 2, 3 and Corollaries 1, 2, 3 hold if we replace at the same time the term "predicate bisimulation" with "first–order bisimulation" and the languages $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}$ and $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$ with the languages with function symbols $\mathcal{L}^{\mathcal{PF}}_{\mathsf{FOL}}$ and $\mathcal{L}^{\mathcal{PF}}_{\mathsf{FOL}}(X)$ respectively.*

Proof: It is the same as the proofs of Theorems 1, 2 and 3 and Corollaries 1 and 2. Only the [6, Lemmas 1 & 2] have to be replaced by the 'translation' Lemmas 6 and 7 and [5, Lemma 1] and Lemma 5 have to be replaced by Lemma 8. □

C Proof of Theorem 7

Lemma 9. *Let C^a be a set of abstract connectives, let $(M, \overline{w})$ be a pointed C^a–model and let U be an ultrafilter over a non-empty set I. Then, $\prod_U t(M, \overline{w})$ is isomorphic to $t\left(\prod_U (M, \overline{w})\right)$ (where t is defined in Definition 32).*

Proof: The translation t of $(M, \overline{w})$ adds relations $R_{[f^l_k]}, R_k, R_{\Box_k}, R_U, R_\perp$ to $(M, \overline{w})$. We are going to prove that the (ultraproduct) relations $\prod_U R_{[f^l_k]}, \prod_U R_k, \prod_U R_{\Box_k}, \prod_U R_U, \prod_U R_\perp$ in $\prod_U t(M, \overline{w})$ are the same as the relations $R_{[f^l_k]}, R_k, R_{\Box_k}, R_U, R_\perp$ in $t\left(\prod_U (M, \overline{w})\right)$.

- Let $k, l \in \mathbb{N}^*$ and let $f^l_k : [\![1; l]\!] \to [\![1; k]\!]$ be a mapping.
  $\prod_U R_{[f^l_k]} \prod_U v^i_1 \ldots \prod_U v^i_l \prod_U w^i_1 \ldots \prod_U w^i_k$
  iff $\left\{i \in I \mid R^i_{[f^l_k]} v^i_1 \ldots v^i_l w^i_1 \ldots w^i_k\right\} \in U$
  iff $\left\{i \in I \mid \forall j \in [\![1; l]\!], v^i_j = w^i_{f^l_k(j)}\right\} \in U$
  iff for all $j \in [\![1; l]\!]$, $\left\{i \in I \mid v^i_j = w^i_{f^l_k(j)}\right\} \in U$ by properties of ultrafilters
  iff for all $j \in [\![1; l]\!]$, $\prod_U v^i_j = \prod_U w^i_{f^l_k(j)}$. So, the relation $\prod_U R_{[f^l_k]}$ in $\prod_U t(M, \overline{w})$ is the same as the relation $R_{[f^l_k]}$ in $t\left(\prod_U (M, \overline{w})\right)$.

- $\prod_U R_k \prod_U \overline{w}^i_1 \prod_U \overline{w}^i_2 \prod_U \overline{w}^i_3$ iff $E \triangleq \left\{i \in I \mid \overline{w}^i_1 = \overline{w}^i_2 = \overline{w}^i_3\right\} \in U$. We have to show that $\prod_U \overline{w}^i_1 = \prod_U \overline{w}^i_2 = \prod_U \overline{w}^i_3$, *i.e.* $A \triangleq \left\{i \in I \mid \overline{w}^i_1 = \overline{w}^i_2\right\} \in U$ and $B \triangleq \left\{i \in I \mid \overline{w}^i_2 = \overline{w}^i_3\right\} \in U$.

However, $E \subseteq A$ and $E \subseteq B$, so $A \in U$ and $B \in U$. So, the relation $\prod_U R_k$ in $\prod_U t(M, \overline{w})$ is the same as the relation R_k in $t\left(\prod_U (M, \overline{w})\right)$.

- $\prod_U R_{\Box_k} \prod_U \overline{v}^i \prod_U \overline{w}^i$
  iff $\left\{i \in I \mid R_{\Box_k} \overline{v}^i \overline{w}^i\right\} \in U$
  iff $\left\{i \in I \mid \text{there is } u^i \in W \text{ such that } \overline{v}^i = (\overline{w}^i, u^i)\right\} \in U$
  iff there is $(u^i)_{i \in I}$ such that $\prod_U \overline{v}^i = \prod_U (\overline{w}^i, u^i)$
  iff there is $\prod_U u^i$ such that $\prod_U \overline{v}^i = \left(\prod_U \overline{w}^i, \prod_U u^i\right)$ by definition of $\prod_U (\overline{w}^i, u^i)$. So, the relation $\prod_U R_{\Box_k}$ in $\prod_U t(M, \overline{w})$ is the same as the relation $R_{\Box_k}$ in $t\left(\prod_U (M, \overline{w})\right)$.
- $\prod_U R_U \prod_U \overline{w}_1^i \prod_U \overline{w}_2^i$ iff $E = \left\{i \in I \mid R_U \overline{w}_1^i \overline{w}_2^i\right\} \in U$. However, $R_U \overline{w}_1^i \overline{w}_2^i$ always holds, for any $\overline{w}_1^i, \overline{w}_2^i$. So, $E = I$ which does belong to U.
- The same reasoning as in the previous case applies to $\prod_U R_\perp$.

$\Box$

Theorem 7 (Keisler theorems for protologics)**.** *Let C^a be a set of normal abstract connectives complete for conjunction and disjunction. Then, theorems 1, 2, 3 hold if we replace "pointed structures" by "pointed C^a–models", the language "$\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$" by the language "$\mathcal{L}_{\mathsf{C}^a}$" and "X–compatible predicate bisimulation" by "C^a–bisimulation".*

Proof: Let C^a be a set of normal abstract connectives complete for conjunction and disjunction. Then, the set $t(\mathsf{C}^a)$ of molecular connectives are also normal and complete for conjunction and disjunction. Then, applying Theorem 1 to $t(\mathsf{C}^a)$, we obtain that for all pointed C^a–models $(M, \overline{w})$ and $(M', \overline{w'})$, the following are equivalent:

1. $t(M, \overline{w}) \rightsquigarrow_{t(\mathsf{C}^a)} t(M', \overline{w'})$;
2. there exists a countably incomplete ultrafilter U over $\mathbb{N}$ such that $\prod_U t\left(M, \overline{w}\right) \rightarrow_{t(\mathsf{C}^a)} \prod_U t\left(M', \overline{w'}\right)$.

So, if we prove that (a) $t(M, \overline{w}) \rightsquigarrow_{t(\mathsf{C}^a)} t(M', \overline{w'})$ iff $(M, \overline{w}) \rightsquigarrow_{\mathsf{C}^a} (M', \overline{w'})$ and (b) $\prod_U t\left(M, \overline{w}\right) \rightarrow_{t(\mathsf{C}^a)} \prod_U t\left(M', \overline{w'}\right)$ iff $\prod_U (M, \overline{w}) \rightarrow_{\mathsf{C}^a} \prod_U \left(M', \overline{w'}\right)$ then we will have proved Theorem 1 for protologics. Expression (a) follows directly from Lemma 3. Expression (b) is proved if we prove that $t\left(\prod_U (M, \overline{w})\right)$ is isomorphic to $\prod_U t\left(M, \overline{w}\right)$, for all pointed C^a–models $(M, \overline{w})$. This is Lemma 9.

The proofs of Theorems 2 and 3 for protologics are the same as the proofs of [5, Theorem 2] and [5, Theorem 3] for molecular logics. The only difference is that in the proof we use the version of Theorem 1 for protologics that we have just proved. □

D Proofs of Propositions 7, 8, 9, 10, 11, 12

The proofs of Propositions 1, 2, 3, 4, 5, 6 are in the appendix of the companion article [5].

Proposition 7. *Let $\mathcal{P}$ be a set of predicate symbols and $k \in \mathbb{N}^*$. If $\mathcal{E}_{\mathsf{C}^{\mathcal{P}}}$ is a specific class of pointed $\mathsf{C}^{\mathcal{P}}$–model, the predicate atomic logic (of type k) associated to $\mathcal{E}_{\mathsf{C}^{\mathcal{P}}}$ and $\mathcal{P}$ is as expressive as a predicate atomic logic (of type k) in the sense of [6, Definition 18].*

Proof: We fix a tuple of k variables and constants $\overline{x} \in (\mathcal{V} \cup \mathcal{C})^k$. We denote by $T'_{\overline{x}}$ and $ST'_{\overline{x}}$ the translations $T_{\overline{x}}$ and $ST_{\overline{x}}$ of [6, Definition 19] and [6, Definition 20] respectively, both for formulas and models. Then, we define the predicate atomic logic (of type k) $\left(\mathcal{L}'_{\mathsf{C}^{\mathcal{P}}}, \mathcal{E}'_{\mathsf{C}^{\mathcal{P}}}, \Vdash'\right)$ in the sense of [6, Definition 18] by $\mathcal{L}'_{\mathsf{C}^{\mathcal{P}}} \triangleq \{ST'_{\overline{x}}T_{\overline{x}}(\varphi) \mid \varphi \in \mathcal{L}_{\mathsf{C}^{\mathcal{P}}}\}$ and $\mathcal{E}'_{\mathsf{C}^{\mathcal{P}}} \triangleq \{ST'_{\overline{x}}T_{\overline{x}}(M, \overline{w}) \mid (M, \overline{w}) \in \mathcal{E}_{\mathsf{C}^{\mathcal{P}}}\}$. Then, using Lemmas 1 and 2 as well as [6, Lemma 1] and [6, Lemma 2], we can easily prove that $(\mathcal{L}_{\mathsf{C}^{\mathcal{P}}}, \mathcal{E}_{\mathsf{C}^{\mathcal{P}}}, \Vdash)$ and $\left(\mathcal{L}'_{\mathsf{C}^{\mathcal{P}}}, \mathcal{E}'_{\mathsf{C}^{\mathcal{P}}}, \Vdash'\right)$ are equally expressive. The translations are $ST'_{\overline{x}}T_{\overline{x}}$ from the former to the latter and $ST_{\overline{x}}T'_{\overline{x}}$ from the latter to the former. Reciprocally, the proof that every predicate atomic logic (of type k) in the sense of [6, Definition 18] is as expressive as a predicate atomic logic (of type k) in the sense of this article is similar, one only needs to swap the $'$s. □

Proposition 8. *A predicate bisimulation between structures without functions is a partial isomorphism.*

Proof: Condition 1. of a partial isomorphism follows from conditions 1. and 5. of a predicate bisimulation. Condition 2. is the same. The "back" condition of a partial isomorphism is obtained by condition 2 and 4 of a predicate bisimulation. We have to show that condition 3 of Definition 30 of a partial isomorphism is deducible from the definition of a predicate bisimulation. Let Z be a predicate bisimulation between the pointed structures (M, s) and (M', s'). Let $\overline{w} \in M^k$ and $\overline{w'} \in M'^k$ with $k > 0$ and assume that $\overline{w}Z\overline{w'}$ (such a pair exists since a predicate bisimulation is non empty). For all $v \in M$, there is $v' \in M'$ such that $(\overline{w}, v)Z(\overline{w'}, v')$. Then, we have that $(v)Z(v')$ by condition 3 of a predicate bisimulation. Hence, we have proved condition 3 of a partial isomorphism. □

Proposition 9. *Let $X \subseteq \mathcal{V} \cup \mathcal{C}$ be a non-empty set. If there is an X–compatible predicate bisimulation between two pointed structures then they make true the same formulas of $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(X)$.*

Proof: This follows from Proposition 8 and the corresponding proposition for partial isomorphisms [10, Proposition 2.4.4] together with Condition 5 of Definition 29. □

Proposition 10. *If there is a first–order bisimulation between two pointed structures then there is a predicate bisimulation and a partial isomorphism between them. Any two finite or countable structures such that there is a predicate or (if they contain functions) a first–order bisimulation between them are isomorphic.*

Proof: The first part follows from the fact that conditions 1.-3. of first-order bisimulations imply conditions 1. of predicate bisimulations and partial isomorphisms (see Definition 30): one can easily prove by induction on formulas φ that if $(w_1, \ldots, w_n)Z_0(v_1, \ldots, v_n)$ then (M_1, s_1) and (M_2, s_2) make true the same formulas $\varphi(x_1, \ldots, x_n) \in \mathcal{L}^{\mathcal{PF}}_{\mathsf{FOL}}$, where (M_1, s_1) and (M_2, s_2) are

defined like in Definition 30. The other conditions of a partial isomorphism are fulfilled by the remaining conditions of a predicate or first–order bisimulation. Because we proved that a predicate bisimulation is a partial isomorphism (without function), the second part of the proposition follows from [10, Proposition 2.4.4]. $\square$

Proposition 11. *Let $\mathcal{P} = \{\mathsf{R}\}$ and let $M = (W, \mathcal{R})$ be a predicate $\mathsf{C}^{\mathcal{P}}$–model such that the predicate symbol R is associated to the relation $R_\square$. Let c be the universal molecular connective $\square_1 \supset_2 ([f_2^2]\mathsf{R}, [f_2^1]id_1^1)$. Then, for all $w, v \in W$, we have that $R_c vw$ iff $-R_\square vw$. (We recall that R_c is defined in Definition 16.)*

Proof: We define the following molecular subconnectives of c:

$$c \triangleq \square_1(c_1) \qquad c_1 \triangleq \supset_2 (c_2, c_3) \qquad c_2 \triangleq [f_2^2]\mathsf{R} \qquad c_3 \triangleq [f_2^1](id_1^1)$$

Let $w, v \in W$. Then, we have that $R_c vw$
iff $\forall \overline{u}_0(R^-_{\square_1}\overline{u}_0 w \vee R_{c_1} v\overline{u}_0)$
iff $\forall \overline{u}_0\overline{u}_1\overline{u}_2(R^-_{\square_1}\overline{u}_0 w \vee (R^-_{\supset_2}\overline{u}_1\overline{u}_2\overline{u}_0 \vee R_{-c_2}\overline{u}_1 \vee R_{c_3} v\overline{u}_2))$
iff $\forall \overline{u}_0\overline{u}_1\overline{u}_2 u_3(R^-_{\square_1}\overline{u}_0 w \vee (R^-_{\supset_2}\overline{u}_1\overline{u}_2\overline{u}_0 \vee R_{-[f_2^2]\mathsf{R}}\overline{u}_1 \vee (R^-_{[f_2^1]}u_3\overline{u}_2 \vee R_{id_1^1} vu_3)))$
iff $\forall \overline{u}_0\overline{u}_1\overline{u}_2 u_3(R_{\square_1}\overline{u}_0 w \rightarrow (R^-_{\supset_2}\overline{u}_1\overline{u}_2\overline{u}_0 \vee R_{-[f_2^2]\mathsf{R}}\overline{u}_1 \vee (R^-_{[f_2^1]}u_3\overline{u}_2 \vee R_{id_1^1} vu_3)))$
iff $\forall u_0\overline{u}_1\overline{u}_2 u_3(\neg(\overline{u}_1 = \overline{u}_2 = (w, u_0)) \vee R_{-[f_2^2]\mathsf{R}}\overline{u}_1 \vee \neg R_{[f_2^1]}u_3\overline{u}_2 \vee \neg(v = u_3))$
iff $\forall u_0\overline{u}_1\overline{u}_2 u_3(\overline{u}_1 = \overline{u}_2 = (w, u_0) \wedge v = u_3 \rightarrow (R_{[f_2^1]}u_3\overline{u}_2 \rightarrow R_{-[f_2^2]\mathsf{R}}\overline{u}_1))$
iff $\forall u_0 u_3(v = u_3 \rightarrow (u_3 = u_0 \rightarrow R_{-[f_2^2]\mathsf{R}}(w, u_0)))$
iff $R_{-[f_2^2]\mathsf{R}}(w, v)$
iff $(w, v) \in [\![-[f_2^2]\mathsf{R}]\!]$ because $-[f_2^2]\mathsf{R}$ is of arity 0
iff $\exists \overline{u}(R_{[f_2^2]}\overline{u}(w, v) \wedge \neg R_\square \overline{u})$
iff $\neg R_\square vw$ $\square$

Proposition 12. *Let $\mathsf{L}_1 = (\mathcal{L}_1, \mathcal{E}_1, \Vdash_1)$ and $\mathsf{L}_2 = (\mathcal{L}_2, \mathcal{E}_2, \Vdash_2)$ be two molecular logics such that $\mathcal{L}_1$ and $\mathcal{L}_2$ are built up from the sets of molecular connectives C_1 and C_2. If L_1 and L_2 are schematically equally expressive (via some mappings $T_1^M, T_1^\varphi, T_2^M, T_2^\varphi$) then for all $M_1, M_1' \in \mathcal{E}_1$ which are C_1–bisimilar, $T_1^M(M_1)$ and $T_1^M(M_1')$ are also C_2–bisimilar.*

Proof: Let $\{Z_c \mid c \in V'_{c_0}, c_0 \in \mathsf{C}_1\}$ be a C_1–bisimulation between M_1 and M_1'. We define the C_2–bisimulation $\{Z_c \mid c \in V'_{c_0}, c_0 \in \mathsf{C}_2\}$ between $T_1^M(M_1)$ and $T_1^M(M_1')$ as follows. Let $c_2 \in \mathsf{C}_2$ and $c \in V'_{c_2}$. For all $\overline{w}_1 \in \overline{w}(M_1, \mathsf{C}_1)$ and all $\overline{w'}_1 \in \overline{w}(M_1', \mathsf{C}_1)$, we set $T_1^M(\overline{w}_1)Z_c T_1^M(\overline{w'}_1)$ iff $\overline{w}_1 Z \overline{w'}_1$ (we recall that $Z \triangleq \bigcup\left\{Z_{id_k^i} \mid id_k^i \text{ appears in some } c_1 \in \mathsf{C}_1\right\}$). Now, we prove the conditions i., ii. and 1.-3. of Definition 19. Conditions i., ii. trivially hold by definition of Z. Likewise, condition 1. trivially holds by definition of equi-expressivity. Now, we prove condition 2. Assume that c is maximally decomposed into $c = c_0(c_1, \ldots, c_n)$ where c_0 is an existential molecular connective with n distinct id_k^is. Assume moreover that we have that $T_1^M(\overline{w})Z_c T_1^M(\overline{w})$ and $R_{c_0}T_1^M(\overline{w}_1)\ldots T_1^M(\overline{w}_n)T_1^M(\overline{w})$. Then, by definition of Z_c, we have $\overline{w}Z\overline{w'}$ and by application of condition iii. of Definition 36, we have that $R_{T_2^\varphi(c_0)}\overline{w}_1\ldots\overline{w}_n\overline{w}$. Now, by condition i. of Definition 36, $T_2^\varphi(c_0)$ is an existential molecular connective which is of the same arity and tonicity signature as c_0 and with n distinct id_k^is, like c_0. So, by condition 2 of Definition 19, there are $\overline{w'}_1, \ldots, \overline{w'}_n$ such that $R_{T_2^\varphi(c_0)}\overline{w'}_1\ldots\overline{w'}_n\overline{w'}$ and for all $j \in [\![1; n]\!]$ such that $\pm_j = +$, we have $\overline{w}_j Z \overline{w'}_j$ and for all $j \in [\![1; n]\!]$ such that $\pm_j = -$, we have $\overline{w'}_j Z \overline{w}_j$. This implies that there are $\overline{w'}_1, \ldots, \overline{w'}_n$ such that $R_{c_0}T_1^M(\overline{w'}_1)\ldots T_1^M(\overline{w'}_n)T_1^M(\overline{w'})$ and for all $j \in [\![1; n]\!]$ such that $\pm_j = +$, we have $T_1^M(\overline{w}_j)Z_{c_j}T_1^M(\overline{w'}_j)$ and for all $j \in [\![1; n]\!]$ such that $\pm_j = -$, we have $T_1^M(\overline{w'}_j)Z_{c_j}T_1^M(\overline{w}_j)$. This implies in turn that there are $\overline{w'}_1, \ldots, \overline{w'}_n$ such that $R_{c_0}T_1^M(\overline{w'}_1)\ldots T_1^M(\overline{w'}_n)T_1^M(\overline{w'})$ and

for all $j \in [\![1;n]\!]$ such that $\pm_j = +$, there is $i \in [\![1;n]\!]$ (take $i = j$) such that $c_j = c_i$ and $T_1^M(\overline{w}_j)Z_{c_j}T_1^M(\overline{w'}_j)$ and for all $j \in [\![1;n]\!]$ such that $\pm_j = -$, there is $i \in [\![1;n]\!]$ (take $i = j$) such that $c_j = c_i$ and $T_1^M(\overline{w'}_j)Z_{c_j}T_1^M(\overline{w}_j)$. This proves condition 2 in that case. The case where c is itself an existential molecular connective is proved similarly. Likewise, the proof of condition 3 is similar, with universal molecular connectives instead of existential molecular connectives. $\square$