A van Benthem Theorem for Atomic and Molecular Logics

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

After recalling the definitions of atomic and molecular logics, we show how notions of bisimulation can be automatically defined from the truth conditions of the connectives of any of these logics. Then, we prove a generalization of van Benthem modal characterization theorem for molecular logics. Our molecular connectives should be uniform and contain all conjunctions and disjunctions. We use modal logic, the Lambek calculus and modal intuitionistic logic as case study and compare in particular our work with Olkhovikov’s work.

1 Introduction

Modal bisimulation is the notion of invariance of modal logic: every formula of first-order logic (FOL) with a free variable whose truth value is always the same in two bisimilar models is equivalent to the translation into FOL of a formula of modal logic. This is the core of the van Benthem characterization theorem. A wide variety of non–classical logics have been introduced over the past decades: modal logics, relevant logics, Lambek calculi, to name just a few. For each of these logics, one can define a notion of invariance and prove by adapting the van Benthem’s characterization theorem that this notion of invariance characterizes the given logic as a fragment of FOL. A drawback of this logical pluralistic approach is that this has to be done by hand on a case by case basis for each non–classical logic. Each time the notion of invariance has to be found out and each time the proof of the van Benthem characterization theorem has to be adapted for that specific notion of invariance. For example, a similar characterization theorem has been proved for (modal) intuitionistic logic [19], temporal logic [15], sabotage modal logic [7], the modal \( \mu \)-calculus [14], graded modal logic [22]. This situation is obviously problematic if one shares the ideal of “universal logic” [8]. Instead, we would prefer to obtain automatically from the definitions of the connectives of a given logic a suitable definition of bisimulation and its associated characterization theorem. This is what we are going to provide in this article for a wide range of non–classical logics, those molecular logics whose connectives are uniform, a notion introduced in that paper. Atomic and molecular logics are introduced in [6]. They behave as ‘normal forms’ for logics since we show in [6] that every non-classical logic such that the truth conditions of its connectives can be expressed in terms of first-order formulas is as expressive as an atomic or molecular logic.

Organization of the article. We start in Sections 2 and 3 by recalling first–order logics, modal logic, the Lambek calculus and modal intuitionistic logic. In Section 4 we recall atomic and molecular logics. Then, in Section 5 we will show how a suitable notion of bisimulation/invariance can be defined automatically from the definition of the connectives of any atomic or molecular logic. Then, in Section 7 we will generalize van Benthem modal characterization theorem to molecular logics whose connectives are uniform. Finally, we discuss related work in Section 8 in particular the work of Olkhovikov.
2 Classical logics

The set \( \mathcal{P} \triangleq \{ \mathcal{R}_1, \ldots, \mathcal{R}_n, \ldots \} \) 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), \( \mathcal{V} \triangleq \{ v_1, \ldots, v_n, \ldots \} \) 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 \( x, x_1, x_2, \ldots, y_1, y_2, \ldots, z_1, z_2, \ldots \) to refer to arbitrary variables or constants, which can be for example \( v_4, v_5, c_{101}, c_{21}, \ldots \).

The first-order language \( \mathcal{L}_{\text{FOL}} \) is defined inductively by the following grammars in BNF:

\[
\begin{align*}
\mathcal{L}_{\text{FOL}}^C : \ t & \ ::= \ x \mid c \\
\mathcal{L}_{\text{FOL}}^P : \ \phi & \ ::= \ R_1 \ldots t \mid \bot \mid (\phi \rightarrow \psi) \mid \forall x \phi
\end{align*}
\]

where \( x \in \mathcal{V}, c \in \mathcal{C}, t \in \mathcal{L}_{\text{FOL}}^C \) and \( R \in \mathcal{P} \). Elements of \( \mathcal{L}_{\text{FOL}}^C \) are called terms and elements of \( \mathcal{L}_{\text{FOL}}^P \) are called first-order formulas. Formulas of the form \( R_1 \ldots t_k \) are called atomic formulas. If \( \phi \in \mathcal{L}_{\text{FOL}}^P \), the Boolean negation of \( \phi \), denoted \( \neg \phi \), is defined by the abbreviation \( \neg \phi \triangleq (\phi \rightarrow \bot) \). We also use the abbreviations \( \top \triangleq \neg \bot \), \( (\phi \lor \psi) \triangleq (\neg \neg \phi \lor \neg \neg \psi) \) and \( (\phi \land \psi) \triangleq (\phi \rightarrow \psi) \land (\psi \rightarrow \phi) \) as well as the abbreviations \( \exists x \phi \triangleq \forall y \neg (\forall x \neg \phi) \), \( \forall x_1 \ldots x_n \phi \triangleq \forall x_1 \ldots \forall x_n \phi \), \( \exists x_1 \ldots x_n \phi \triangleq \exists x_1 \ldots \exists x_n \phi \) and \( \forall \mathcal{T} \phi \triangleq \forall x_1 \ldots x_n \phi \) if \( \mathcal{T} \triangleq (x_1, \ldots, x_n) \) is a tuple of variables.

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

Free variables may be used to bind elements of two different subformulas. For example, the formula \( Ryx \lor R'xz \) with free variables \( x, y, z \) will be evaluated in a structure in such a way that \( x \) will be assigned the same element of the domain in the two subformulas \( Ryx \) and \( R'xz \).

We denote by \( \mathcal{L}_{\text{FOL}}(\mathcal{T}) \) the fragment of \( \mathcal{L}_{\text{FOL}} \) whose formulas all contain at least one free variable or constant.

A structure is a tuple \( M \triangleq (W, \{ R_1, \ldots, R_n, \ldots, c_1, \ldots, c_n, \ldots \}) \) 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 \( R_1, \ldots, R_n, \ldots \); respectively;
- \( c_1, \ldots, c_n, \ldots \in W \) are elements of the domain called distinguished elements.

An assignment over \( M \) is a mapping \( s : \mathcal{V} \cup \mathcal{C} \rightarrow W \) such that for all \( c_i \in \mathcal{C} \), \( s(c_i) = c_i \). 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 assignment \( (M, s) \) is called a pointed structure. The class of all pointed structures \( (M, s) \) is denoted \( \mathcal{M}_{\text{FOL}} \).

The satisfaction relation \( \models_{\text{FOL}} \subseteq \mathcal{M}_{\text{FOL}} \times \mathcal{L}_{\text{FOL}} \) is defined inductively as follows. Below, we write \( (M, s) \models \phi \) for \( ((M, s), \phi) \in \models_{\text{FOL}} \).

\[
\begin{align*}
(M, s) & \models \bot \quad \text{never}; \\
(M, s) & \models R_1 \ldots t_k \quad \text{iff} \quad (s(t_1), \ldots, s(t_k)) \in R_i; \\
(M, s) & \models (\phi \rightarrow \psi) \quad \text{iff} \quad (M, s) \not\models \psi \text{ then } (M, s) \models \psi; \\
(M, s) & \models \forall x \phi \quad \text{iff} \quad (M, s[x := w]) \models \phi \text{ for all } w \in W.
\end{align*}
\]

In the literature \( [10] \), \( (M, s) \models \phi(x_1, \ldots, x_k) \) is sometimes denoted \( M \models \phi(x_1, \ldots, x_k)[w_1, \ldots, w_k] \), \( M \models \phi[w_1/x_1, \ldots, w_k/x_k] \) or simply \( M \models \phi[w_1, \ldots, w_k] \), with \( w_1 = s(x_1), \ldots, w_k = s(x_k) \). In that case, we say
that \((M, s)\) makes \(\varphi\) true. We say that the formula \(\varphi \in \mathcal{L}_{\text{FOL}}^P\) is realized in \(M\) when there is an assignment \(s\) such that \((M, s) \models \varphi\). 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. Two (pointed) structures are \textit{elementarily equivalent} when they make true the same sentences.

A triple of the form \((\mathcal{L}_{\text{FOL}}, \mathcal{E}_{\text{FOL}}, \models_{\text{FOL}})\) is called the \textit{first-order logic} associated to \(\mathcal{L}_{\text{FOL}}\) and \(\mathcal{E}_{\text{FOL}}\). If \(\mathcal{L}_{\text{FOL}} = \mathcal{L}_{\text{FOL}}^P\), the triple \((\mathcal{L}_{\text{FOL}}^P, \mathcal{E}_{\text{FOL}}, \models_{\text{FOL}})\) is called \textit{pure predicate logic} (associated to \(\mathcal{E}_{\text{FOL}}\), if \(\mathcal{L}_{\text{FOL}} = \mathcal{L}_{\text{FOL}}(\mathcal{X})\), the triple \((\mathcal{L}_{\text{FOL}}(\mathcal{X}), \mathcal{E}_{\text{FOL}}, \models_{\text{FOL}})\) is called \textit{pure predicate logic with constants} (associated to \(\mathcal{E}_{\text{FOL}}\)). When \(\mathcal{E}_{\text{FOL}} = \mathcal{M}_{\text{FOL}}\), they are simply called respectively \textit{pure predicate logic} and \textit{pure predicate logic with constants}.

3 Non-classical logics

In this section, \(\mathbb{A}\) is a set of \textit{propositional letters} which can be finite or infinite.

3.1 Modal logic

The set \(\mathbb{I}\) is a set of indices which can be finite or infinite. The \textit{multi-modal language} \(\mathcal{L}_{\text{ML}}\) is defined inductively by the following grammar in BNF:

\[
\mathcal{L}_{\text{ML}} : \varphi ::\quad p \mid \neg p \mid (\varphi \land \varphi) \mid (\varphi \lor \varphi) \mid \Diamond_j \varphi \mid \Box_j \varphi
\]

where \(p \in \mathbb{A}\) and \(j \in \mathbb{I}\). We present the so-called \textit{possible world semantics} of modal logic. A \textit{Kripke model} \(M\) is a tuple \(M = (W, \{R_1, \ldots, R_m, P_1, \ldots, P_n, \ldots\})\) where

- \(W\) is a non-empty set whose elements are called \textit{possible worlds};
- \(R_1, \ldots, R_m, \ldots \subseteq W \times W\) are binary relations over \(W\) called \textit{accessibility relations};
- \(P_1, \ldots, P_n, \ldots \subseteq W\) are unary relations interpreting the propositional letters of \(\mathbb{P}\).

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

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

\[
\begin{align*}
(M, w) \models p_i & \quad \text{iff} \quad P_i(w) \text{ holds}; \\
(M, w) \models \neg p & \quad \text{iff} \quad P_i(w) \text{ does not hold}; \\
(M, w) \models (\varphi \land \psi) & \quad \text{iff} \quad (M, w) \models \varphi \quad \text{and} \quad (M, w) \models \psi; \\
(M, w) \models (\varphi \lor \psi) & \quad \text{iff} \quad (M, w) \models \varphi \quad \text{or} \quad (M, w) \models \psi; \\
(M, w) \models \Diamond_j \varphi & \quad \text{iff} \quad \text{there exists} \; v \in W \; \text{such that} \; R_{jvw} \; \text{and} \; (M, v) \models \varphi; \\
(M, w) \models \Box_j \varphi & \quad \text{iff} \quad \forall \; v \in W \; \text{such that} \; R_{jvw}, (M, v) \models \varphi.
\end{align*}
\]

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

3.2 Lambek calculus

The \textit{Lambek language} \(\mathcal{L}_{\text{LC}}\) is the set of formulas defined inductively by the following grammar in BNF:

\[
\mathcal{L}_{\text{LC}} : \varphi ::\quad p \mid (\varphi \otimes \varphi) \mid (\varphi \& \varphi) \mid (\varphi \bowtie \varphi)
\]

where \(p \in \mathbb{A}\). A \textit{Lambek model} is a tuple \(M = (W, \{R, P_1, \ldots, P_n, \ldots\})\) where:
• $W$ is a non-empty set;
• $R \subseteq W \times W \times W$ is a ternary relation over $W$;
• $P_1, \ldots, P_n, \ldots \subseteq W$ are unary relations over $W$.

We write $w \in M$ for $w \in W$ by abuse and $(M, w)$ is called a pointed Lambek model. The class of all pointed Lambek models is denoted $\mathcal{E}_{\text{LC}}$. We define the satisfaction relation $\models_{\text{int}} \subseteq \mathcal{E}_{\text{LC}} \times L_{\text{LC}}$ by the following truth conditions. Below, we write $(M, w) \models \varphi$ for $(M, w), \varphi \in \models_{\text{LC}}$. For all Lambek models $M = (W, \{R, P_1, \ldots, P_n, \ldots\})$, all $w \in M$, all $\varphi, \psi \in L_{\text{LC}}$ and all $p_i \in \mathbb{P}$,

\[
(M, w) \models p_i \quad \text{iff} \quad P_i(w) \text{ holds;}
(M, w) \models (\varphi \otimes \psi) \quad \text{iff} \quad \text{there are } v, u \in W \text{ such that } Rvw,
\quad (M, v) \models \varphi \text{ and } (M, u) \models \psi;
(M, w) \models (\varphi \triangleright \psi) \quad \text{iff} \quad \text{for all } v, u \in W \text{ such that } Rvwu,
\quad \text{if } (M, v) \models \varphi \text{ then } (M, u) \models \psi;
(M, w) \models (\psi \circ \varphi) \quad \text{iff} \quad \text{for all } v, u \in W \text{ such that } Rvwu,
\quad \text{if } (M, v) \models \varphi \text{ then } (M, u) \models \psi.
\]

The triple $(L_{\text{LC}}, E_{\text{LC}}, \models_{\text{LC}})$ forms a logic, that we call the Lambek calculus. Bisimulations for the Lambek calculus, called directed bisimulations, can be found in [21].

### 3.3 Modal intuitionistic logic

The modal intuitionistic language $L_{\text{int}}$ is defined inductively by the following grammar in BNF:

\[
L_{\text{int}} : \quad \varphi ::= \top | \bot | p | (\varphi \land \varphi) | (\varphi \lor \varphi) | (\varphi \rightarrow \varphi) | \Diamond \varphi | \Box \varphi
\]

where $p \in \mathbb{A}$. A modal intuitionistic model is a tuple $M = (W, \{R, R_\circ, P_1, \ldots, P_n, \ldots\})$ where:

• $W$ is a non-empty set;
• $R \subseteq W \times W$ is a binary relation over $W$ which is reflexive and transitive ($R$ is reflexive if for all $w \in W$ $Rww$ and transitive if for all $u, v, w \in W$, $Ruv$ and $Rvw$ imply $Ruw$);
• $R_\circ \subseteq W \times W$ is a binary relation over $W$;
• $P_1, \ldots, P_n, \ldots \subseteq W$ are unary relations over $W$ such that for all $v, w \in W$, if $Rvw$ and $P_n(w)$ then $P_n(v)$.

We write $w \in M$ for $w \in W$ by abuse and the pair $(M, w)$ is called a pointed modal intuitionistic model. The class of all pointed modal intuitionistic models is denoted $\mathcal{E}_{\text{int}}$. We define the satisfaction relation $\models_{\text{int}} \subseteq \mathcal{E}_{\text{int}} \times L_{\text{int}}$ by the following truth conditions. Below, we write $(M, w) \models \varphi$ for $(M, w), \varphi \in \models_{\text{int}}$. For all modal intuitionistic models $M = (W, \{R, R_\circ, P_1, \ldots, P_n, \ldots\})$, all $w \in M$, all $\varphi, \psi \in L_{\text{int}}$ and all $p_i \in \mathbb{P}$,

\[
(M, w) \models \top \quad \text{always;}
(M, w) \models \bot \quad \text{never;}
(M, w) \models p_i \quad \text{iff} \quad P_i(w) \text{ holds;}
(M, w) \models (\varphi \land \psi) \quad \text{iff} \quad (M, w) \models \varphi \text{ and } (M, w) \models \psi;
(M, w) \models (\varphi \lor \psi) \quad \text{iff} \quad (M, w) \models \varphi \text{ or } (M, w) \models \psi;
(M, w) \models (\varphi \rightarrow \psi) \quad \text{iff} \quad \text{for all } v \in W \text{ such that } Rvw, \text{if } (M, v) \models \varphi \text{ then } (M, v) \models \psi;
(M, w) \models \Diamond \varphi \quad \text{iff} \quad \text{for all } v \in W \text{ such that } R_\circ vu, (M, u) \models \varphi;
(M, w) \models \Box \varphi \quad \text{iff} \quad \text{there is } u \in W \text{ such that } R_\circ vu \text{ and } (M, u) \models \varphi.
\]
The triple \((\mathcal{L}_{\text{int}}, \mathcal{E}_{\text{int}}, \models_{\text{int}})\) forms a logic, that we call modal intuitionistic logic. Bisimulations for (modal) intuitionistic logic can be found in [18, 19].

4 Atomic and molecular logics

4.1 Atomic logics

Atomic logics are logics such that the truth conditions of their connectives are defined by first-order formulas of the form \(\forall x_1 \ldots x_n (\pm_1 P_1 x_1 \lor \ldots \lor \pm_n P_n x_n \lor \pm_1 R_1 x_1 \ldots x_n)\) or \(\exists x_1 \ldots x_n (\pm_1 P_1 x_1 \land \ldots \land \pm_n P_n x_n \land \pm_1 R_1 x_1 \ldots x_n)\) where \(\pm\) is either empty or \(\neg\). Likewise, propositional letters are defined by first-order formulas of the form \(\pm P x\). We will represent the structure of these formulas by means of so-called skeletons whose various arguments capture the different features and patterns that allow us to define them completely.

We recall that \(\mathbb{N}^+\) denotes the set of natural numbers minus 0 and that for all \(n \in \mathbb{N}^+\), \(\mathcal{S}_n\) denotes the group of permutations over the set \(\{1, \ldots, n\}\). Permutations are generally denoted \(\sigma, \tau\), the identity permutation \(\text{Id}\) is sometimes denoted \(1\) as the neutral element of every permutation group and \(\sigma^{-1}\) stands for the inverse permutation of the permutation \(\sigma\). For example, the permutation \(\sigma = (3, 1, 2)\) is the mapping that maps 1 to 3, 2 to 1 and 3 to 2 (see for instance [23] for more details).

**Definition 1** (Atomic skeletons and connectives). The sets of atomic skeletons \(\mathcal{P}\) and \(\mathcal{C}\) are defined as follows:

\[
\mathcal{P} \equiv \mathcal{S}_1 \times \{+, -\} \times \{\forall, \exists\} \times \mathbb{N}^+;
\]

\[
\mathcal{C} \equiv \mathcal{P} \cup \bigcup_{n \in \mathbb{N}^+} \left\{ \mathcal{S}_{n+1} \times \{+, -\} \times \{\forall, \exists\} \times \mathbb{N}^{n+1} \times \{+, -\}^n \right\}.
\]

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

An (atomic) connective or (atomic) propositional letter is an object to which is associated an (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. We also introduce the *Boolean connectives called conjunctions and disjunctions*:

\[
\mathcal{B} \equiv \{\land_k, \lor_k \mid k \in \mathbb{N}^+\}
\]

The type signatures of \(\land_k\) and \(\lor_k\) are \((k, k, k)\) and their arity is 2.

We say that a set of atomic connectives \(\mathcal{C}\) is complete for conjunction and disjunction when it contains all conjunctions and disjunctions \(\land_k, \lor_k\), for \(k\) ranging over all input types and output types of the atomic connectives of \(\mathcal{C}\). The set of atomic skeletons associated to \(\mathcal{C}\) is denoted \(*(\mathcal{C})\), its set of propositional letters is denoted \(\mathcal{P}(\mathcal{C})\).

Propositional letters are denoted \(p, p_1, p_2, \ldots\) and connectives are denoted \(*, *_1, *_2, \ldots\). 

**Remark 1.** The permutations \(\sigma\) mentioned in atomic skeletons play an important role in the proof theory of atomic logics, which is dealt with in [4, 5].
**Definition 2** (Atomic language). Let $C$ be a set of atomic connectives. The *(typed) atomic language* $L_C$ associated to $C$ is the smallest set that contains the propositional letters and that is closed under the atomic connectives. That is,

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

The **Boolean atomic language** $L^B_C$ is the smallest set that contains the propositional letters and that is closed under the atomic connectives of $C$ as well as the Boolean connectives $\Box$:

- for all $\varphi, \psi \in L^B_C$ of type $k$, we have that $(\varphi \land \psi), (\varphi \lor \psi) \in L^B_C$.

Elements of $L_C$ are denoted $\varphi, \psi, \alpha, \ldots$. The *type* of a formula $\varphi \in L_C$ is denoted $k(\varphi)$. For all $\varphi_1, \ldots, \varphi_n \in L$ of type $k$, $\land \{\varphi_1, \ldots, \varphi_n\}$ and $\lor \{\varphi_1, \ldots, \varphi_n\}$ stand for $((\varphi_1 \land \varphi_2) \land \ldots \land \varphi_n)$ and $((\varphi_1 \lor \varphi_2) \lor \ldots \lor \varphi_n)$ respectively. When it is clear from the context, we will omit the subscript $k$ in $\land, \lor$ and write them $\land, \lor$.

*In the sequel, we assume that all sets of connectives $C$ are such that they contain at least a propositional letter.*

**Definition 3** ($C$–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 $* \in C$ which is not a Boolean connective of type signature $(k, k_1, \ldots, k_n)$ is associated to a $k_1 + \ldots + k_n + k$–ary relation $R_* \in \mathcal{R}$.

An assignment is a tuple $(w_1, \ldots, w_k) \in W^k$ for some $k \in \mathbb{N}^+$, generally denoted $\overline{w}$. 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 $\mathcal{M}_C$.

**Definition 4** (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 $L_C$ in $M$, denoted $[\cdot]^M : L_C \rightarrow W^k$, inductively as follows: for all propositional letters $p \in C$ of skeleton $(\mathrm{Id}, \pm, \exists, \forall)$, all connectives $* \in C$ of skeleton $(\sigma, \pm, \exists, (k, k_1, \ldots, k_n), (\pm, \ldots, \pm))$ of arity $n > 0$ and all $k \in \mathbb{N}^+$, for all $\varphi, \psi, \varphi_1, \ldots, \varphi_n \in L_C$, if $k(\varphi) = k(\psi) = k$,

$$
\begin{align*}
[p]^M & \triangleq \pm R_p \\
[(\varphi \land \psi)]^M & \triangleq [\varphi]^M \cap [\psi]^M \\
[(\varphi \lor \psi)]^M & \triangleq [\varphi]^M \cup [\psi]^M \\
[*(\varphi_1, \ldots, \varphi_n)]^M & \triangleq f_*(\llbracket \varphi_1 \rrbracket^M, \ldots, \llbracket \varphi_n \rrbracket^M)
\end{align*}
$$

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

- if $\exists \mathcal{E} = \forall$: "$\forall \overline{w}_1 \in W^{k_1} \ldots \overline{w}_n \in W^{k_n} (\overline{w}_1 \upharpoonright 1 W_1 \lor \ldots \lor \overline{w}_n \upharpoonright n W_n \lor R_{\sigma}^{\sigma} \overline{w}_1 \ldots \overline{w}_n \overline{w}_{n+1})$";
- if $\exists \mathcal{E} = \exists$: "$\exists \overline{w}_1 \in W^{k_1} \ldots \overline{w}_n \in W^{k_n} (\overline{w}_1 \upharpoonright 1 W_1 \land \ldots \land \overline{w}_n \upharpoonright n W_n \land R_{\sigma}^{\sigma} \overline{w}_1 \ldots \overline{w}_n \overline{w}_{n+1})$";

where, for all $j \in [1; n]$, $\overline{w}_j \upharpoonright j W_j \triangleq \begin{cases} W_j & \text{if } \pm = + \\
W_j & \text{if } \pm = - \end{cases}$ and $R_{\sigma}^{\sigma} \overline{w}_1 \ldots \overline{w}_{n+1}$ holds iff $\pm R_\sigma \overline{w}_{\sigma^{-1}(1)} \ldots \overline{w}_{\sigma^{-1}(n+1)}$ with the notations $+R_\sigma \triangleq R_\sigma$ and $-R_\sigma \triangleq W^{k_1+k_2+\ldots+k_n} - R_\sigma$. If $\mathcal{E}_C$ is a class of pointed $C$–models, the *satisfaction relation* $\models \subseteq \mathcal{E}_C \times L_C$ is defined as follows: for all $\varphi \in L_C$ and all
((M, \overline{w}), \varphi) \in \mathcal{E}_C, ((M, \overline{w}), \varphi) \in \models \iff \overline{w} \in \llbracket \varphi \rrbracket^M. We usually write \((M, \overline{w}) \models \varphi\) instead of \(((M, \overline{w}), \varphi) \in \models\) and we say that \(\varphi\) is true in \((M, \overline{w})\).

The class of atomic logics is defined by \(\mathbb{L}_{GGGL} = \{ (\mathcal{L}_C, \mathcal{E}_C, \models) \mid C\text{ is a finite set of atomic connectives and } \mathcal{E}_C \text{ is a class of } C\text{-models}\}\). The atomic logic \((\mathcal{L}_C, \mathcal{E}_C, \models)\) is the atomic logic associated to \(\mathcal{E}_C\) and \(C\). The logics of the form \((\mathcal{L}_C, M_C, \models)\) are called basic atomic logics. We call them Boolean (basic) atomic logics when their language includes the Boolean connectives \(\neg\).

**Example 1** (Lambek calculus, modal logic). The Lambek calculus, where \(C = \{ \top, \lor, \bot, \} \) is defined in Section 3 is an example of atomic logic. Here \(\lor, \bot, \top\) are the connectives of skeletons \((\sigma_1, +, s_1)\), \((\sigma_3, -, s_3)\), \((\sigma_3, -, s_3)\). Another example of atomic logic is modal logic where \(C = \{ \top, \bot, \land, \lor, \Diamond, \Box \}\) is such that

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

- \(\land, \lor, \Diamond, \Box\) are connectives of skeletons \((\sigma_1, +, s_1)\), \((\sigma_1, -, s_1)\), \((\sigma_2, +, s_1)\) and \((\sigma_2, -, s_1)\) respectively;

- \(C\text{-models } M = (W, \mathcal{R}) \in \mathcal{E}_C\) are such that \(R_\land = R_\lor = \{ (w, w, w) \mid w \in W \}, R_\Box = R_\Box\) and \(R_\top = R_\bot = W\).

Indeed, one can easily show that, with these conditions on the \(C\text{-models}\) of \(\mathcal{E}_C\), we have that for all \((M, w) \in \mathcal{E}_C, (M, w) \models \land (\varphi, \psi) \iff (M, w) \models \varphi \text{ and } (M, w) \models \psi, \text{ and } (M, w) \models \land (\varphi, \psi) \iff (M, w) \models \land \varphi \text{ or } (M, w) \models \land \psi\). The Boolean conjunction and disjunction \(\land\) and \(\lor\) are defined using the connectives of \(C\) by means of special relations \(R_\land\) and \(R_\lor\). However, they could obviously be defined directly. Many more examples of atomic connectives are in Figure 2. They are in fact just examples of gaggale connectives since all gaggle logics [4, 5] are also atomic logics: they are all of type signature \((1, 1, \ldots, 1)\). All the possible truth conditions of unary and binary atomic connectives of this type signature are in [4, 5].

### 4.2 Molecular logics

Molecular logics are logics whose primitive connectives are compositions of atomic connectives. That is why we call them ‘molecular’, just as molecules are compositions of atoms in chemistry.

**Definition 5** (Molecular skeleton and connective). The class \(C^*\) of molecular skeletons is the smallest set such that:

- \(P \cup B \subseteq C^*\) and \(C^*\) contains for each \(k \in \mathbb{N}^*\) a symbol \(id_k\) of type signature \((k, k)\) and arity 1;
| Atomic connective | Truth condition | Non–classical con. in the literature |
|------------------|----------------|-----------------------------------|
| \( \varphi(\sigma_1,+,s_1) \psi \) | \( \exists u (v \in \varphi \land u \in \psi \land Ruwv) \) | \( \varphi \in \psi \quad 16, \varphi \odot, \psi \quad 3 \) |
| \( \varphi(\sigma_2,-,s_2) \psi \) | \( \forall u (v \in \varphi \lor u \in \psi \lor \neg Ruwv) \) | / 16, \( \varphi <_1 \psi \quad 3 \) |
| \( \varphi(\sigma_3,-,s_2) \psi \) | \( \forall u (v \in \varphi \lor u \in \psi \lor \neg Ruwv) \) | / 16, \( \varphi c_2 \psi \quad 3 \) |
| \( \varphi(\sigma_4,+,s_1) \psi \) | \( \exists u (v \in \varphi \land u \in \psi \land Ruwv) \) | / 16, \( \varphi c_2 \psi \quad 3 \) |
| \( \varphi(\sigma_5,-,s_3) \psi \) | \( \forall u (v \in \varphi \lor u \in \psi \lor \neg Ruwv) \) | / 16, \( \varphi <_1 \psi \quad 3 \) |
| \( \varphi(\sigma_6,-,s_3) \psi \) | \( \forall u (v \in \varphi \lor u \in \psi \lor \neg Ruwv) \) | / 16, \( \varphi c_2 \psi \quad 3 \) |
| \( \varphi(\sigma_7,-,s_5) \psi \) | \( \forall u (v \in \varphi \lor u \in \psi \lor \neg Ruwv) \) | / 16, \( \varphi <_1 \psi \quad 3 \) |
| \( \varphi(\sigma_8,+,s_5) \psi \) | \( \forall u (v \in \varphi \lor u \in \psi \lor \neg Ruwv) \) | / 16, \( \varphi c_2 \psi \quad 3 \) |
| \( \varphi(\sigma_9,-,s_4) \psi \) | \( \forall u (v \in \varphi \lor u \in \psi \lor \neg Ruwv) \) | / 16, \( \varphi c_2 \psi \quad 3 \) |
| \( \varphi(\sigma_{10},-,s_3) \psi \) | \( \forall u (v \in \varphi \lor u \in \psi \lor \neg Ruwv) \) | / 16, \( \varphi c_2 \psi \quad 3 \) |
| \( \varphi(\sigma_{11},-,s_2) \psi \) | \( \forall u (v \in \varphi \lor u \in \psi \lor \neg Ruwv) \) | / 16, \( \varphi c_2 \psi \quad 3 \) |
| \( \varphi(\sigma_{12},-,s_1) \psi \) | \( \forall u (v \in \varphi \lor u \in \psi \lor \neg Ruwv) \) | / 16, \( \varphi c_2 \psi \quad 3 \) |

Figure 2: Some binary connectives of atomic logics of type \((1,1,1)\)
• for all $\star \in \mathbb{C}$ of type signature $(k, k_1, \ldots, k_n)$ and all $c_1, \ldots, c_n \in \mathbb{C}^+$ of type signatures $(k_1, k_1^1, \ldots, k_{a_1}^1)$, ..., $(k_n, k_n^b, \ldots, k_n^n)$ respectively, the connective $\star(c_1, \ldots, c_n)$ belongs to $\mathbb{C}^+$, its type signature is $(k, k_1^1, \ldots, k_{a_1}^1, k_n^b, \ldots, k_n^n)$ and its arity is $a_1 + \ldots + a_n$.

We define the quantification signature $\mathcal{Q}(c) = \star(c_1, \ldots, c_n)$ by $\mathcal{Q}(c) = \mathcal{Q}(\star)$.

If $c \in \mathbb{C}^+$, we define its decomposition tree as follows. If $c = \star \in \mathbb{C}$ is of arity $n > 0$, then its decomposition tree $T_c$ is the tree of root $\star$ with $n$ children—leaves labeled by $id$. If $c = \star(c_1, \ldots, c_n) \in \mathbb{C}^+$ then its decomposition tree $T_c$ is a tree labeled with atomic connectives 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.

A **molecular connective** is an object to which is associated a molecular skeleton. Its arity, 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 $id$ of the decomposition trees of the molecular connectives of $C$.

**Example 2** (Modal intuitionistic logic). Let us consider the connectives defined by the following first-order formulas:

$$c(x) \equiv \forall y (Rx \rightarrow \exists z (R\circ yz \rightarrow P(z)))$$

$$c'(x) \equiv \forall y (Rx \rightarrow \exists z (R\circ yz \land P(z)))$$

$$\star_1(x) \equiv \forall y (Rx \rightarrow P(y))$$

$$\star_2(x) \equiv \exists z (R\circ yz \rightarrow P(z))$$

$$\star_3(x) \equiv \exists z (R\circ yz \land P(z))$$

Then, $\star_1, \star_2, \star_3$ are atomic connectives and the connectives associated to $c, c'$ are molecular connectives. Indeed, $c$ is the composition of $\star_1$ and $\star_2$, $c = \star_1(\star_2)$, and $c'$ is the composition of $\star_1$ and $\star_3$, $c' = \star_1(\star_3)$. Equivalently, $c$ and $c'$ will have the same semantics as $c = \star_1(\star_2(id))$ and $c' = \star_1(\star_3(id))$. The connective associated to $c$ corresponds to the connective $\Box$ of modal intuitionistic logic and the connective associated to $c'$ corresponds to the connective $\Diamond$ of modal intuitionistic logic [19] defined in Section 3.

**Definition 6** (Molecular language). Let $C$ be a set of molecular connectives. The **(typed) molecular language** $\mathcal{L}_C$ 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 $\mathcal{L}_C$;
- for all $\star \in \mathbb{C}$ of type signature $(k, k_1, \ldots, k_n)$ and for all $\varphi_1, \ldots, \varphi_n \in \mathcal{L}_C$ of types $k_1, \ldots, k_n$ respectively, we have that $\star(\varphi_1, \ldots, \varphi_n) \in \mathcal{L}_C$ and $\star(\varphi_1, \ldots, \varphi_n)$ is of type $k$.

The **Boolean molecular language** $\mathcal{L}_C^B$ is the smallest set that contains the propositional letters and that is closed under the molecular connectives of $C$ as well as the Boolean connectives $\mathbb{B}$:

- for all $\varphi, \psi \in \mathcal{L}_C^B$ of type $k$, we have that $(\varphi \land_k \psi), (\varphi \lor_k \psi) \in \mathcal{L}_C^B$.

We say that $C$ is complete for conjunction and disjunction when its associated set of atomic connectives is complete for conjunction and disjunction.

Elements of $\mathcal{L}_C$ are called **molecular formulas** and are denoted $\varphi, \psi, \alpha, \ldots$. The type of a formula $\varphi \in \mathcal{L}_C$ is denoted $k(\varphi)$. We use the same abbreviations as for the atomic language.
**Definition 7** (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 $C$. We also define $\overline{w}(M, C) \triangleq \overline{w}(M, C')$. The truth conditions for molecular connectives are defined naturally inductively from the truth conditions of atomic connectives of Definition 4 (we only give the new cases): for all $k \in \mathbb{N}^+$,

$$[[\ast (c_1, \ldots, c_n) (\varphi_1^1, \ldots, \varphi_1^{k_1}, \ldots, \varphi_n^1, \ldots, \varphi_n^{k_n})]]^M \triangleq \ast (\varphi_1^1, \ldots, \varphi_1^{k_1}, \ldots, \varphi_n^1, \ldots, \varphi_n^{k_n}) M$$

If $\mathcal{E}_C$ is a class of pointed $C$–models, the triple $(\mathcal{L}_C, \mathcal{E}_C, \mathfrak{I}^\mathfrak{L})$ is a logic called the molecular logic associated to $\mathcal{E}_C$ and $\mathcal{C}$. The logics of the form $(\mathcal{L}_C, \mathcal{M}_C, \mathfrak{I}^\mathfrak{L})$ are called basic molecular logics. We call them Boolean (basic) molecular logics when their language includes the Boolean connectives $\mathbb{B}$.

### 4.3 Boolean negation

Note that atomic logics do not include Boolean negation as a primitive connective. It turns out that Boolean negation can be defined systematically for each atomic connective by applying a transformation on it. 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 8** (Boolean negation). Let $\ast$ be an $n$–ary connective of skeleton $(\sigma, \pm, \mathcal{A}, \mathcal{T}, \pm_1, \ldots, \pm_n)$. The Boolean negation of $\ast$ is the connective $\neg \ast$ of skeleton $(\sigma, \neg \pm, \neg \mathcal{A}, \neg \mathcal{T}, \neg \pm_1, \ldots, \neg \pm_n)$ where $\neg \mathcal{A} \triangleq \exists$ if $\mathcal{A} = \forall$ and $\neg \mathcal{A} \triangleq \forall$ otherwise, which is associated in any $C$–model to the same relation as $\ast$. If $\varphi = \ast (\varphi_1, \ldots, \varphi_n)$ is an atomic formula, the Boolean negation of $\varphi$ is the formula $\neg \varphi \triangleq \neg \ast (\varphi_1, \ldots, \varphi_n)$.

**Proposition 1.** Let $C$ be a set of atomic connectives such that $\neg \ast \in C$ for all $\ast \in C$. Let $\varphi \in \mathcal{L}_C$ and $M$ be a $C$–model. Then, for all $\overline{w} \in \overline{w}(M, C)$, $\overline{w} \in [[\neg \varphi]]^M$ iff $\overline{w} \notin [[\varphi]]^M$.

### 5 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 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.

#### 5.1 Atomic logics

**Definition 9** (**C**–bisimulation). Let $C$ be a set of atomic connectives, let $\ast \in C$ and let $M_1 = (W_1, \mathcal{R}_1)$ and $M_2 = (W_2, \mathcal{R}_2)$ be two $C$–models. A binary relation $Z \subseteq \bigcup_{k \in \mathbb{N}}(W_1^k \times W_2^k)$ is a $C$–bisimulation between $M_1$ and $M_2$ when for all $\ast \in C$, if $\{M, M'\} = \{M_1, M_2\}$, then for all $\overline{w}_1, \ldots, \overline{w}_n, \overline{w}_1', \ldots, \overline{w}_n', \overline{w}, \overline{w}' \in \overline{w}(M, C) \cup \overline{w}(M', C)$,

1. if $\ast$ is a propositional letter $p$ then, if $\overline{w} Z \overline{w}'$ and $\overline{w} \in [p]$ then $\overline{w}' \in [p]$;
2. if $\ast$ has skeleton $(\sigma, \pm, \mathcal{K}, (\pm_1, \ldots, \pm_n))$ and we have $\overline{w} Z \overline{w}'$ and $R_{\ast}^{\pm_1} \overline{w}_1 \ldots \overline{w}_n \overline{w}$, then

$$\exists \overline{w}_1', \ldots, \overline{w}_n' (\overline{w}_1 \times \overline{w}_1' \wedge \overline{w}_2 \times \overline{w}_2' \wedge \ldots \wedge \overline{w}_n \times \overline{w}_n' \wedge R_{\ast}^{\pm_1} \overline{w}_1' \ldots \overline{w}_n' \overline{w})$$

3. if $\ast$ has skeleton $(\sigma, \pm, \mathcal{K}, (\pm_1, \ldots, \pm_n))$ and we have $\overline{w} Z \overline{w}'$ and $-R_{\ast}^{\pm_1} \overline{w}_1 \ldots \overline{w}_n' \overline{w}'$, then

$$\exists \overline{w}_1', \ldots, \overline{w}_n' (\overline{w}_1 \times \overline{w}_1' \wedge \overline{w}_2 \times \overline{w}_2' \wedge \ldots \wedge \overline{w}_n \times \overline{w}_n' \wedge -R_{\ast}^{\pm_1} \overline{w}_1' \ldots \overline{w}_n' \overline{w})$$
where, for all $j \in \{1; n\}$, we define $\overrightarrow{w_j w'_j} = \begin{cases} \overrightarrow{w_j w'_j} & \text{if } \pm_j = + \\ \overrightarrow{Z}\overrightarrow{w_j w'_j} & \text{if } \pm_j = - \end{cases}$.

When such a $\mathcal{C}$–bisimulation $Z$ exists and $\overrightarrow{w Z w'}$, we say that $(M, \overrightarrow{w})$ and $(M', \overrightarrow{w'})$ are $\mathcal{C}$–bisimilar and we write it $(M, \overrightarrow{w}) \sim_\mathcal{C} (M', \overrightarrow{w'})$.

Note that case 1. is a particular instance of cases 2. and 3. with $n = 0$. The fact that the order $M_1 - M_2$ can be possibly reversed at the level of the definition is reminiscent of the way directed bisimulations are defined for the Lambek calculus, as we will see in Example 4.

**Definition 10.** Let $\mathcal{C}$ be a set of atomic connectives. Let $(M, \overrightarrow{w})$ and $(M', \overrightarrow{w'})$ be two pointed $\mathcal{C}$–models. We write $(M, \overrightarrow{w}) \sim_\mathcal{C} (M', \overrightarrow{w'})$ when for all $\varphi \in \mathcal{L}_\mathcal{C}$, $(M, \overrightarrow{w}) \models_\varphi$ implies $(M', \overrightarrow{w'}) \models_\varphi$.

**Proposition 2.** Let $\mathcal{C}$ be a set of atomic connectives and let $M_1 = (W_1, R_1)$ and $M_2 = (W_2, R_2)$ be two $\mathcal{C}$–models. Let $Z$ be a $\mathcal{C}$–bisimulation between $M_1$ and $M_2$. Then, if $\{M, M'\} = \{M_1, M_2\}$ then for all $\overrightarrow{w} \in \overrightarrow{\mathcal{P}}(M, C)$, all $\overrightarrow{w'} \in \overrightarrow{\mathcal{P}}(M', C)$, if $\overrightarrow{w Z w'}$ then $(M, \overrightarrow{w}) \sim_\mathcal{C} (M', \overrightarrow{w'})$.

**Example 3** (Modal logic). Let us consider the connectives of modal logic: $\mathcal{C} = \{p, \neg p, \land, \lor, \Box, \Diamond\}$ where $p$ has skeleton $(\text{Id}, +, \exists, 1)$, $\neg p$ has skeleton $(\text{Id}, -, \forall, 1)$, $\Box$ has skeleton $((\tau_2, +, t_1))$ and $\Diamond$ 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 $\mathcal{C}$–models). A binary relation $Z$ between $M_1$ and $M_2$ is a $\mathcal{C}$–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 $w Z w'$ and $w \in \llbracket p \rrbracket$ then $w' \in \llbracket p \rrbracket$ (condition for $p$);
- if $w Z w'$ and $w' \in \llbracket p \rrbracket$ then $w \in \llbracket p \rrbracket$ (condition for $\neg p$);
- if $w Z w'$ and $Rwv$ then there is $v' \in W'$ such that $v Z v'$ and $R'w'v'$ (condition for $\Box = (\tau_2, +, t_1)$);
- if $w Z w'$ and $R'w'v'$ then there is $v \in W$ such that $v Z v'$ and $Rwv$ (condition for $\Diamond = (\tau_2, -, t_2)$).

Note that every $\mathcal{C}$–bisimulation can be canonically extended into a symmetric $\mathcal{C}$–bisimulation: one sets $w' Z w$ when $w Z w'$ already holds.

**Example 4** (Lambek calculus). Let us consider the connectives of the Lambek calculus: $\mathcal{C} = \{p, o, \backslash, /\}$ where $p$ has skeleton $(\text{Id}, +, \exists, 1)$, $o$ has skeleton $(\sigma_1, +, s_1)$, $\backslash$ has skeleton $(\sigma_5, -, s_2)$ and $/$ has skeleton $(\sigma_3, -, s_2)$. Let $M_1 = (W_1, \{R_1, P_1\})$ and $M_2 = (W_2, \{R_2, P_2\})$ be two Lambek models (they are also $\mathcal{C}$–models). A binary relation $Z$ between $M_1$ and $M_2$ is a $\mathcal{C}$–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 $w, v, u \in M$ and all $w', v', u' \in M'$,

- if $w Z w'$ and $w \in \llbracket p \rrbracket$ then $w' \in \llbracket p \rrbracket$ (condition for $p$);
- if $w Z w'$ and $Rwv$ then there are $v', u' \in W'$ such that $v Z v'$, $u Z u'$ and $Rv'u'w'$ (condition for $o = (\sigma_1, +, s_1))$;
- if $v Z v'$ and $Rv'u'w'$ then there are $u, w \in W$ such that $u Z u'$, $w Z w'$ and $Ruwv$ (condition for $\backslash = (\sigma_5, -, s_2)$);
- if $u Z u'$ and $Rv'u'w'$ then there are $v, w \in W$ such that $v Z v'$, $w Z w'$ and $Ruwv$ (condition for $/$ = $(\sigma_3, -, s_2)$).

The following proposition shows that the notions of $\mathcal{C}$–bisimulation for the Lambek calculus and directed bisimulation coincide (directed bisimulations are defined for example in [21] Definition 13.2) and likewise for modal logic.

**Proposition 3.** Let $\mathcal{C} = \{p, \neg p, \land, \lor, \Box, \Diamond\}$ be the connectives of Example 3 and let $M$ and $M'$ be two $\mathcal{C}$–models. Then, a $\mathcal{C}$–bisimulation between $M$ and $M'$ is a modal bisimulation between $M$ and $M'$ and vice versa.
• Let $C = \{p, o, \land, \lor, \Rightarrow\}$ be the connectives of Example 2 and let $M$ and $M'$ be two $C$–models. Then, a $C$–bisimulation between $M$ and $M'$ is a directed bisimulation between $M$ and $M'$ and vice versa.

**Example 5 (Intuitionistic logic).** Let us consider the connectives of intuitionistic logic: $C = \{p, \bot, \top, \land, \lor, \Rightarrow\}$ where $p$ has skeleton $(\text{Id}, +, \bot, \top, \land, \lor)$ and $\Rightarrow$ has skeleton $(\sigma, -, \triangleright, \leftarrow)$ (here, $\top$ and $\bot$ are represented by specific propositional letters of respective signatures $(\text{Id}, +, \bot, \top, \land, \lor)$ and $(\text{Id}, -, \lor, \triangleright, \leftarrow)$). Let $M_1 = (W_1, R_1, P)$ and $M_2 = (W_2, R_2, P)$ be two intuitionistic models. Following the results of Section 8, one can easily show that for all $u_1, v_1, w_1 \in W_1$ and all $u_2, v_2, w_2 \in W_2$,

$$R_1 \Rightarrow u_1 v_1 w_1 \text{ iff } R_1 u_1 w_1 \text{ and } R_1 v_1 w_1$$

(1)

$$R_2 \Rightarrow u_2 v_2 w_2 \text{ iff } R_2 u_2 w_2 \text{ and } R_2 v_2 w_2$$

(2)

One can show that for all $\phi \in L_C$ and all $w, v_1, w_1 \models \phi$ iff $M_1 \models \phi$ and $\neg \phi$ (and likewise for $M_2$ and $M_2'$). Now, a binary relation $Z$ between $M_1$ and $M_2'$ is a $C$–bisimulation between $M_1$ and $M_2'$ iff for all $M, M' \in \{M_1, M_2\}$, all $w, w', u' \in \mathcal{W}(M, C) \cup \mathcal{W}(M', C)$ and all $p \in \mathcal{P}$,

- if $wZw'$ and $w \in [p]$ then $w' \in [p]$ (condition for $p$);

- if $vZw'$ and $R_1 \Rightarrow v'u'w'$ then there are $u, w \in W$ such that $u'Zu$, $wZw'$ and $R_1 \Rightarrow wu w$ (*condition for $\Rightarrow$);

- conditions for $\top$ and $\bot$ trivially hold because of their semantics.

Using Expressions (1) and (2), one can easily show that condition (*) is equivalent to the following condition:

- if $vZw'$ and $R_1 \Rightarrow v'u'w'$ then there are $u, w \in W$ such that $u'Zu$, $wZw'$ and $Rw w$ (**) .

We will show in Section 8 that condition (**) is equivalent on $\sigma$–saturated models to Olkhovikov’s condition “step” of [19] Definition 1] of his “basic asimulation”.

### 5.2 Molecular logics

**Definition 11 (C–bisimulation for molecular connectives).** Let $C$ be a set of molecular connectives and let $M_1 = (W_1, R_1)$ and $M_2 = (W_2, R_2)$ be two $C$–models. For all $c_0 \in C$, let $V_{c_0}$ be the vertices of the decomposition tree $T_{c_0}$. We associate to each vertex $c \in V_{c_0}$ a binary relation $Z_c \subseteq \bigcup_{k \in \mathbb{N}^*} (W_1 \times W_2 \times W_1 \times W_2 \times W_1 \times W_2)$.

- The set of such binary relations is denoted $\{Z\} \cup \bigcup_{c \in C} \{Z_c \mid c \in V_{c_0}\}$ and is such that if $c$ is $id_k$ for some $k \in \mathbb{N}^*$ then $Z_c$ is $Z$ and we have that $Z \subseteq \bigcup \{Z_c \mid c \in C\}$. We say that this set of binary relations is a $C$–bisimulation between $M_1$ and $M_2$ when for all $c_0 \in C$, all vertices $c \in V_{c_0}$, if $\{M, M'\} = \{M_1, M_2\}$ then for all $w_1, w_2, w_1, w_2, w_1, w_2, w_1, w_2 \\
\Rightarrow \mathcal{W}(M, C) \cup \mathcal{W}(M', C)$.

1. if $c$ is a propositional letter $p$ then, $\overline{W}Z_c \overline{W}$ and $\overline{p} \in [p]$ imply $\overline{w} \in [p]$;

2. if $c$ has skeleton $*(c_1, \ldots, c_n)$ with $* = (\sigma, \pm, \top, \bot, \ldots, \pm, \top, \bot, \ldots)$ and we have $\overline{w}Z_c \overline{w}$ and $R_x \sigma \overline{w_1 \ldots w_n}$, then

$$\exists \overline{w_1}, \ldots, \overline{w_n} (\overline{w_1} \sigma c_1 \overline{w_1} \land \overline{w_2} \sigma c_2 \overline{w_2} \land \ldots \land \overline{w_n} \sigma c_n \overline{w_n} \land R_x \sigma \overline{w_1 \ldots w_n})$$

3. if $c$ has skeleton $*(c_1, \ldots, c_n)$ with $* = (\sigma, \pm, \top, \bot, \ldots, \pm, \top, \bot, \ldots)$ and we have $\overline{w}Z_c \overline{w}$ and $\neg R_x \sigma \overline{w_1 \ldots w_n}$, then

$$\exists \overline{w_1}, \ldots, \overline{w_n} (\overline{w_1} \sigma c_1 \overline{w_1} \land \overline{w_2} \sigma c_2 \overline{w_2} \land \ldots \land \overline{w_n} \sigma c_n \overline{w_n} \land \neg R_x \sigma \overline{w_1 \ldots w_n})$$
where for all \( j \in [1; n] \), we have \( \overline{w_j \wedge c_j \overline{w}_j} \equiv \begin{cases} \overline{w_j Z_{c_j} \overline{w}_j} & \text{if } \pm j = + \\ \overline{w_j Z_{c_j} \overline{w}_j} & \text{if } \pm j = - \end{cases} \)

When such a set of binary relations exists and is such that \( \overline{\pi Z \overline{w}'} \), we say that \((M, \overline{w})\) and \((M', \overline{w}')\) are \(C\text{-bisimilar}\) and we write it \((M, \overline{w}) \sim_C (M', \overline{w}')\).

Note that case 1. is a particular instance of cases 2. and 3. with \( n = 0 \).

**Definition 12.** Let \( C \) be a set of molecular connectives. For all \( c_0 \in C \) and all vertex \( c \) of the decomposition tree \( T_{c_0} \), we define the language \( L_{c,C} \) as follows:

\[
L_{c,C} \equiv \begin{cases} \{ c(\varphi_1, \ldots, \varphi_n) \mid \varphi_1, \ldots, \varphi_n \in L_C \} & \text{if } c \text{ is of arity } n > 0 \\ \{ p \} & \text{if } c = p \text{ is a propositional letter} \\ \{ \} & \text{if } c \text{ is } \text{id}_k \text{ for some } k \in \mathbb{N}^* \end{cases}
\]

Let \((M, \overline{w})\) and \((M', \overline{w}')\) be two pointed \(C\text{-models}\). We write \((M, \overline{w}) \sim_C (M', \overline{w}')\) when for all \( \varphi \in L_{c,C} \), \((M, \overline{w}) \models \varphi \) implies \((M', \overline{w}') \models \varphi \). We also write \((M, \overline{w}) \sim_C (M', \overline{w}')\) when for all \( \varphi \in L_{c,C} \), \((M, \overline{w}) \models \varphi \) implies \((M', \overline{w}') \models \varphi \).

**Proposition 4.** Let \( C \) be a set of molecular connectives and let \( M_1 = (W_1, R_1) \) and \( M_2 = (W_2, R_2) \) be two \(C\text{-models}\). Let \( C_0 \subseteq C \) and for all \( c \in C_0 \), let \( V_c \) be the vertices of the decomposition tree \( T_c \). Let \( \{ Z \} \cup \bigcup_{c_0 \in C_0} \{ Z_c \mid c \in V_{c_0} \} \) be a \( C_0\text{-bisimulation} \) between \( M_1 \) and \( M_2 \). If \( \{ M, M' \} = \{ M_1, M_2 \} \) then for all \( c_0 \in C_0 \) and all \( c \in V_{c_0} \), for all \( \overline{w} \in \overline{\pi}(M, C) \) and all \( \overline{w}' \in \overline{\pi}(M', C) \), if \( \overline{w}_Z \overline{w}' \) then \((M, \overline{w}) \sim_{C_0} (M', \overline{w}')\). In particular, if \( \overline{w} \overline{w}' \) then \((M, \overline{w}) \sim_{C_0} (M', \overline{w}')\).

**Definition 13 (Uniform connective).** A uniform connective is a molecular connective \( c \) whose skeleton is of the form \(*_1(c_1, \ldots, c_n)\) with \(*_1 = (\sigma, \pm, E, \overline{E}, (\pm_1, \ldots, \pm_m)) \in C\) such that

1. \( n \geq 1 \) and \( c_1, \ldots, c_n \) are molecular skeletons of arity 1;
2. for all \( j \in [1; n] \) such that \( c_j \neq \text{id}_k \) for all \( k \in \mathbb{N}^* \), \( E(c_j) = \begin{cases} \forall & \text{if } \pm j = + \\ \exists & \text{if } \pm j = - \end{cases} \);
3. if \( c_0 \) is a molecular skeleton appearing in the decomposition tree of \( c \) of the form \(*_0 = c_0(c_1', \ldots, c_m')\) such that the tonicity signature of \(*_0 \) is \((\pm_1, \ldots, \pm_m) \), then for all \( i \in [1; m] \), \( E(c_i') = \pm_i E(c_0) \).

According to our definition, molecular connectives of the form \(*_1(c_1'(c_1', c_2'))\) cannot be uniform connectives, unless \( c_1' \) or \( c_2' \) is a propositional letter. This is due to our first condition: in that case, \( c(c_1'(c_1', c_2')) \) should be of arity 1, which is possible only if \( c_1' \) or \( c_2' \) is a propositional letter. Hence, uniform connectives can be reduced to the composition of compound subconnectives \( c_1, \ldots, c_m' \), each of arity 1, so that molecular connectives are essentially of the form \(*_1(c_1'(c_1''(c_1''(c_1'))), \ldots, c_m'(c_m''(c_m'')))\). Basically, uniform connectives are such that the quantification patterns of their successive internal connectives are essentially of the form \( \exists \ldots \exists \ldots \lor \ldots \lor \ldots \) or \( \forall \ldots \forall \ldots \).

**Example 6 (Modal intuitionistic logic).** Let \( C = \{ p, T, \bot, \land, \lor, \Rightarrow, \ast, \ast' \} \) where \( \ast, \ast' \in C^* \) are the molecular connectives \( c, c' \) of Example 4 and where \( \{ p, T, \bot, \land, \lor, \Rightarrow \} \) are defined in Example 5. The connectives of \( C \) are all uniform connectives. Note that \( \ast' \) is not a uniform connective and that is why we consider \(-\ast'\) for the moment, which is a uniform connective. We are going to see that we can easily get the bisimilarity condition for \( \ast' \) from the bisimilarity condition for \(-\ast'\).

Let \( M_1 = (W_1, \{ R_1, R_1, P \}) \) and \( M_2 = (W_2, \{ R_2, R_2, P \}) \) be two modal intuitionistic models. The set of binary relations \( \{ Z, Z_{\ast'}, Z_{\ast''} \} \) is a \( C\text{-bisimulation} \) iff for all \( M, M' \in \{ M_1, M_2 \} \) with \( M = (W, \{ R, R, P \}) \) and \( M' = (W', \{ R', R', P \}) \), all \( w, v, u, w', v', u' \in \overline{\pi}(M, C) \cup \overline{\pi}(M', C) \) and all \( p \in \mathbb{P} \),
• if \( wZw' \) and \( w \in [p] \) then \( w' \in [p] \) (condition for \( p \), like in Example [5]);
• if \( vZv' \) and \( R'v'w' \) and \( R'w'w' \) then there are \( u,w \in W \) such that \( u'Zu, wZw' \) and \( Rvw \) (condition for \( * \), like in Example [3]);
• if \( Rwv \) and \( wZw' \) then there is \( v \in W \) such that \( vZv' \) and \( Rwv \), if \( wZ_{e_2}w' \) and \( R_e w'v' \) then there is \( v \in W \) such that \( vZv' \) and \( Rwv \) (condition for \( \star = \star_1(\star_2) \));
• if \( Rwv \) and \( wZw' \) then there is \( v \in W' \) such that \( vZw, v'w'v' \), if \( wZ_{e_3}w' \) and \( R_{e_3}wv \) then there is \( v' \in W' \) such that \( vZv' \) and \( Rwv \).

It turns out that the Conditions of \( (\star) \) are the conditions (diam–2(1)) and (diam–2(2)) of Olkhovikov [19, Definition 9], as expected.

6 Ultrafilters and ultraproducts

In that section, we recall some key notions of model theory [10], ultrafilters and ultraproducts.

Definition 14 (Filter and ultrafilter). Let \( I \) be a non–empty set. A filter \( F \) over \( I \) is a set \( F \in \mathcal{P}(I) \) 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 \( I - X \notin U \).

In the rest of this section, \( I \) is a non–empty set and \( U \) is an ultrafilter over \( I \).

Definition 15 (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 \( \prod_U W_i \triangleq \{ (v_i)_{i \in I} \in \prod_{i \in I} W_i \mid (v_i)_{i \in I} \sim_U (w_i)_{i \in I} \} \). The ultraproduct of \( (W_i)_{i \in I} \) modulo \( U \) is \( \prod_U W_i \triangleq \{ \prod_U w_i \mid (w_i)_{i \in I} \in \prod_{i \in I} W_i \} \).

Definition 16 (Ultraproduct). Let \( (M_i, s_i)_{i \in I} \) be a family of pointed structures. The ultraproduct \( \prod_U (M_i, s_i) \) is the pointed structure \( (\prod_U M_i, \prod_U s_i) \) where \( \prod_U s_i : \mathcal{V} \to \prod_U W_i \) is the assignment such that for all \( \alpha \in \mathcal{V} \), \( (\prod_U s_i)(\alpha) = \prod_U s_i(\alpha) \) and \( \prod_U M_i = (W_U, \mathcal{R}_U) \) is defined as follows:

• \( W_U = \prod_U W_i \);
• for all \( n + 1 \)-ary relations \( R_i \) of \( M_i \), the \( n + 1 \)-ary relation \( \prod_U R_i \in \mathcal{R}_U \) is defined for all \( \prod_U w_i^1, \ldots, \prod_U w_i^{n+1} \in W_U \) by \( \prod_U R_i(\prod_U w_i^1, \ldots, \prod_U w_i^{n+1}) \text{ iff } \{ i \in I \mid R_i(\prod_U w_i^1, \ldots, \prod_U w_i^{n+1}) \} \in U \).

Definition 17 (Closure under ultraproducts). Let \( K \) be a class of pointed structures. We say that \( K \) is closed under ultraproducts when for all non-empty sets \( I \), if for all \( i \in I (M_i, s_i) \in K \) then \( \prod_U (M_i, s_i) \in K \) for all ultraproducts \( U \) over \( I \).
7 A generic van Benthem characterization theorem

In this section we generalize the van Benthem characterization theorem for modal logic \cite[Theorem 2.68]{vB} to molecular logics. We first show how atomic and molecular logics can be naturally embedded into first–order logic.

**Definition 18** (Translation from atomic and molecular logics to FOL). Let \( C \) be a set of atomic connectives.

**Syntax.** For all \( k \in \mathbb{N}^* \) and all \( \pi \in (\lor \cup C)^k \), we define the mappings \( ST_{\pi} : \mathcal{L}_C^k \to \mathcal{L}_{\text{FOL}}^\pi(\pi) \), where \( \mathcal{L}_C^k \) is the set of formulas of \( \mathcal{L}_C \) of type \( k \), as follows: for all atoms \( p \in C \), all \( * \in C \) of skeleton \( (\sigma, + \in E, (k, k_1, \ldots, k_n), (\epsilon_1, \ldots, \epsilon_n)) \) and all \( \phi_1, \ldots, \phi_n \in \mathcal{L}_C 

\[
ST_{\pi}(\phi) \; \equiv \; P(\pi)
\]

\[
ST_{\pi}(\varphi \land_k \psi) \; \equiv \; ST_{\pi}(\varphi) \land ST_{\pi}(\psi)
\]

\[
ST_{\pi}(\varphi \lor_k \psi) \; \equiv \; ST_{\pi}(\varphi) \lor ST_{\pi}(\psi)
\]

\[
ST_{\pi}(\ast(\phi_1, \ldots, \phi_n)) \; \equiv \; \forall \bar{x}_1 \ldots \bar{x}_n (\ast_1 \bar{x}_1 \times \ast_n \bar{x}_n \bar{x}_1 \ldots \bar{x}_n \bar{\pi})
\]

where \( \bar{x}_1, \ldots, \bar{x}_n \) are tuples of free variables of size \( k_1, \ldots, k_n \),

\[
x = \begin{cases} 
\land & \text{if } \lor = \exists \\
\lor & \text{if } \lor = \forall 
\end{cases} \text{ and for all } j \in [1; n], \ast_j = \begin{cases} 
\neg & \text{if } \pm_j = \neg \\
\text{empty} & \text{if } \pm_j = + 
\end{cases}
\]

**Semantics.** Let \( (M, \bar{w}) \) be a pointed \( C \)-model of type \( k \) with \( \bar{w} = (w_1, \ldots, w_k) \). Let \( \bar{x} = (x_1, \ldots, x_k) \in (\lor \cup C)^k \). A pointed structure associated to \( (M, \bar{w}) \) and \( \bar{x} \) is a pointed structure \( (M, \bar{s}_{\bar{x}}) \) \( (\text{the set of predicates } P \text{ considered are a copy of the relations of } M \text{ where the assignment } s_{\bar{x}} \text{ is such that } s_{\bar{x}}(x_1) = w_1, \ldots, s_{\bar{x}}(x_k) = w_k) \).

The above translations canonically extend to molecular logics. Indeed, if \( C \) is a set of molecular connectives, every molecular formula of \( \mathcal{L}_C \) can be viewed as a formula of \( \mathcal{L}_{C'} \), where \( C' \) is the set of atomic connectives associated to \( C \). Likewise, any pointed \( C \)-model can also be viewed as a pointed \( C' \)-model. Then, we apply the above translations to obtain the translation of molecular formulas or \( C \)-models into FOL.

The following proposition follows straightforwardly from the truth conditions of Definition 4.

**Proposition 5.** Let \( C \) be a set of molecular connectives, let \( (M, \bar{w}) \) be a pointed \( C \)-model, let \( \varphi \in \mathcal{L}_C \) of type \( k \) and let \( \bar{x} \in \mathcal{V}^k \). Then, \( (M, \bar{w}) \models \varphi \) iff \( (M, \bar{s}_{\bar{x}}) \models ST_{\bar{x}}(\varphi) \).

**Theorem 1** (Characterization theorem). Let \( C \) be a set of uniform connectives complete for conjunction and disjunction. Let \( \varphi(\bar{x}) \in \mathcal{L}_{\text{FOL}}^\pi(\bar{x}) \) with \( k \) free variables \( \bar{x} = (x_1, \ldots, x_k) \) and let \( (\mathcal{L}_C, \mathcal{E}_C, \models) \) be a molecular logic such that all models of \( \mathcal{E}_C \) contain relations \( \{ R_* \mid * \in C \} \) interpreting all the predicates occurring in \( \varphi(\bar{x}) \). Let \( ST_{\pi}(\mathcal{E}_C) \models \{ (M, \bar{s}_{\bar{x}}) \mid (M, \bar{w}) \in \mathcal{E}_C \text{ of type } k \} \) and assume that \( ST_{\pi}(\mathcal{E}_C) \) is closed under ultraproducts. The two following statements are equivalent:

1. There exists a formula \( \psi \in \mathcal{L}_C \) such that \( \varphi(\bar{x}) \leftrightarrow ST_{\pi}(\psi) \) is valid on \( ST_{\pi}(\mathcal{E}_C) \);

2. \( \varphi(\bar{x}) \) is invariant for \( C \)-bisimulations on \( \mathcal{E}_C \), that is, for all pointed \( C \)-models \( (M, \bar{w}), (M', \bar{w}') \) of \( \mathcal{E}_C \) of type \( k \) such that \( (M, \bar{w}) \to C (M', \bar{w}') \), we have that \( (M, \bar{w}) \models \varphi(\bar{x}) \) implies \( (M', \bar{w}') \models \varphi(\bar{x}) \).

**Remark 2.** The assumption that \( ST_{\pi}(\mathcal{E}_C) \) is closed under ultraproducts is not really demanding since any class of structures definable by a set of first–order sentences is closed under ultraproducts by Keisler theorem \cite[Corollary 6.1.16]{Keisler}. For example, the class of (modal) intuitionistic models is closed under ultraproducts since it is definable by a set of first–order sentences (we only need to impose the reflexivity and transitivity on the binary relations by means of the validity of corresponding sentences). So, our generic theorem applies to these logics. It also applies to modal logic and to many others since the class of all Kripke models is definable by an empty set of sentences and therefore is closed under ultraproducts.
8 Related work

8.1 Comparison with Olkhovikov’s work

The closest work to ours is by Olkhovikov [18] [19] who investigates generalizations of the van Benthem characterization theorem. The publications [18] and [19] deal in particular with (modal) intuitionistic logic, following our methodology, we have rediscovered Olkhovikov’s definitions in Examples 5 [2] and 6. In particular, we have the following result.

**Fact 1.** Let \( M = (W, \{R, P\}) \) and \( M' = (W', \{R', P'\}) \) be two \( \omega \)-saturated intuitionistic models and let \( Z \) be the maximal \( C \)-bisimulation between \( M \) and \( M' \) for set inclusion (\( C \) is defined in Example 5). Then, the following two conditions are equivalent:

1. Condition (**) of Example 5 for all \( v \in W \) and all \( w', v', u' \in W' \), if \( vZv' \) and \( R'v'w' \) and \( R'u'w' \) then there are \( u, w \in W \) such that \( uZu \), \( wZw' \) and \( Rvw \) and \( Rvw' \);

2. Condition “step” of [19] Definition 1: for all \( v \in W \) and all \( w', v' \in W' \), if \( vZv' \) and \( R'v'w' \) then there is \( w \in W \) such that \( wZw' \) and \( w'Zw \) and \( Rvw \).

In [19], Olkhovikov also provides a generic van Benthem style characterization theorem for a number of logics defined by specific kinds of connectives. He introduces a normal form for connectives in terms of formulas of FOL that he calls \( k \)-ary guarded \( x \)-connectives. A \( k \)-ary guarded \( x \)-connective of degree 0 \( \mu = \psi(P_1(x), \ldots, P_k(x)) \) is a Boolean combination of the unary predicates \( P_1(x), \ldots, P_k(x) \). A \( k \)-ary \( \forall \)-guarded \( x_1 \)-connective of degree \( n + 1 \) is a formula of the form \( \forall x_2 \ldots x_{m+1} (\wedge_{i=1}^m S_i(x_i, x_{i+1}) \rightarrow \mu^-) \) where \( S_1, \ldots, S_m \) are binary predicates and \( \mu^- \) is a \( k \)-ary guarded \( x_{m+1} \)-connective of degree \( n \) (provided that formula is not equivalent to \( k \)-ary guarded \( x_1 \)-connective of a smaller degree). \( k \)-ary \( \exists \)-guarded \( x_1 \)-connectives are defined similarly. If one sets \( Rx_1 \ldots x_{m+1} \) for \( \wedge_{i=1}^m S_i(x_i, x_{i+1}) \) then \( \forall x_2 \ldots x_{m+1} (\wedge_{i=1}^m S_i(x_i, x_{i+1}) \rightarrow \mu^-) \) can be viewed as the first-order formula with free variable \( x_1 \) defining a molecular connective: \( \forall x_2 \ldots x_{m+1} (x_2 \in [\bot] \lor \ldots \lor x_m \in [\bot] \lor x_{m+1} \in [\bot] \lor \mu^- \lor -Rx_1 \ldots x_{m+1} \). Hence, guarded connectives of degree not exceeding 1 are captured by specific molecular connectives. It is unclear whether Olkhovikov’s regular connectives of degree 2 are also captured by uniform connectives.

In any case, our results strictly extend those of Olkhovikov because we are able to provide a van Benthem characterization for connectives defined by formulas of FOL with multiple free variables. It is this feature that plays a key role for first-order logic [6]. It is also made more clear and explicit than in Olkhovikov’s publication how the suitable notions of bisimulation are defined from logics given by their set of connectives. Finally, we showed in Examples 2 and 6 how his results about (modal) intuitionistic logic [18] [19] can be recovered in our setting as specific instances of our general results.

8.2 Other related work

Van Benthem theorems have been proved for many non-classical logics, such as (modal) intuitionistic logic [19], intuitionistic predicate logic [18], temporal logic [15], sabotage modal logic [7], graded modal logic [22], fuzzy modal logic [25], coalgebraic modal logics [24], neighbourhood semantics of modal logic [13], the modal mu-calculus [14], hybrid logic [2]. We showed that our generic Theorem 1 subsumes some of them [19] [18] [15]. However, some others are not in the scope of our theorem because the correspondence language to which they refer extends first-order logic. For example, the van Benthem theorem for coalgebraic modal logic [24] is w.r.t. coalgebraic predicate logic, fuzzy modal logic [25] w.r.t. first-order fuzzy predicate logic and the modal mu-calculus [14] w.r.t. monadic second-order logic.

**Acknowledgments.** I thank two anonymous reviewers for helpful comments.
References

[1] Gerard Allwein & J. Michael Dunn (1993): *Kripke Models for Linear Logic*. The Journal of Symbolic Logic 58(2), pp. 514–545, doi:10.2307/2275217.

[2] Carlos Areces, Patrick Blackburn & Maarten Marx (2001): *Hybrid Logics: Characterization, Interpolation and Complexity*. J. Symb. Log. 66(3), pp. 977–1010, doi:10.2307/2695090.

[3] Guillaume Aucher (2016): *Displaying Updates in Logic*. Journal of Logic and Computation 26(6), pp. 1865–1912, doi:10.1093/logcom/exw001.

[4] Guillaume Aucher (2020): *Towards universal logic: Gaggle logics*. Journal of Applied Logics – IfCoLog Journal of Logics and their Applications 7(6), pp. 875–945. Available at http://collegepublications.co.uk/ifcolog/?00042.

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

[6] Guillaume Aucher (2022): *On the Universality of Atomic and Molecular Logics via Protologics*. Logica Universalis (To appear).

[7] Guillaume Aucher, Johan van Benthem & Davide Grossi (2018): *Modal logics of sabotage revisited*. J. Log. Comput. 28(2), pp. 269–303, doi:10.1093/logcom/exx034.

[8] Jean-Yves Béziau (2005): *From Consequence Operator to Universal Logic: A Survey of General Abstract Logic*. In Jean-Yves Béziau, editor: *Logica Universalis*, Birkhäuser Basel, Basel, pp. 3–17, doi:10.1007/3-7643-7643-7/0_1.

[9] Patrick Blackburn, Maarten de Rijke & Yde Venema (2001): *Modal Logic*. Cambridge Tracts in Computer Science 53, Cambridge University Press, doi:10.1017/CBO9781107050884.

[10] C. C. Chang & H. J. Keisler (1998): *Model Theory*. Studies in Logic and the Foundations of Mathematics, Elsevier.

[11] Rajeev Goré (1998): *Substructural logics on display*. Logic Journal of IGPL 6(3), pp. 451–504, doi:10.1093/jigpal/6.3.451.

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

[13] Helle Hvid Hansen, Clemens Kupke & Eric Pacuit (2009): *Neighbourhood Structures: Bisimilarity and Basic Model Theory*. Logical Methods in Computer Science 5(2), doi:10.2168/LMCS-5(2:2)2009.

[14] David Janin & Igor Walukiewicz (1996): *On the Expressive Completeness of the Propositional mu-Calculus with Respect to Monadic Second Order Logic*. In Ugo Montanari & Vladimiro Sassone, editors: *CONCUR ’96, Concurrency Theory, 7th International Conference, Pisa, Italy, August 26-29, 1996, Proceedings*, Lecture Notes in Computer Science 1119, Springer, pp. 263–277, doi:10.1007/3-540-66104-7_60.

[15] Natasha Kurtonina & Maarten de Rijke (1997): *Bisimulations for Temporal Logic*. Journal of Logic, Language and Information 6(4), pp. 403–425, doi:10.1023/A:1008223921944.

[16] Joachim Lambek (1958): *The mathematics of sentence structure*. The American Mathematical Monthly 65(3), pp. 154–170, doi:10.1080/00029890.1958.11989160.
[17] Michael Moortgat (2007): *Symmetries in Natural Language Syntax and Semantics: The Lambek-Grishin Calculus*. In Daniel Leivant & Ruy de Queiroz, editors: Logic, Language, Information and Computation, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 264–284, doi:10.1007/978-3-540-73445-1_19

[18] Grigory K. Olkhovikov (2014): *Model-theoretic characterization of intuitionistic predicate formulas*. Journal of Logic and Computation 24(4), pp. 809–829, doi:10.1093/logcom/ext014

[19] Grigory K. Olkhovikov (2017): *On expressive power of basic modal intuitionistic logic as a fragment of classical FOL*. Journal of Applied Logic 21, pp. 57–90, doi:10.1016/j.jal.2016.11.036

[20] Grigory K. Olkhovikov (2017): *On generalized Van Benthem-type characterizations*. Annals of Pure and Applied Logic 168(9), pp. 1643–1691, doi:10.1016/j.apal.2017.03.002

[21] Greg Restall (2000): *An Introduction to Substructural Logics*. Routledge, doi:10.4324/9780203252642

[22] Maarten de Rijke (2000): *A Note on Graded Modal Logic*. Studia Logica 64(2), pp. 271–283, doi:10.1023/A:1005245900406

[23] Joseph J. Rotman (1995): *An Introduction to the Theory of Groups*. Graduate Texts in Mathematics 148, Springer, doi:10.1007/978-1-4612-4176-8

[24] Lutz Schröder, Dirk Pattinson & Tadeusz Litak (2017): *A van Benthem/Rosen theorem for coalgebraic predicate logic*. Journal of Logic and Computation 27(3), pp. 749–773, doi:10.1093/logcom/exv043

[25] Paul Wild, Lutz Schröder, Dirk Pattinson & Barbara König (2018): *A van Benthem Theorem for Fuzzy Modal Logic*. In Anuj Dawar & Erich Grädel, editors: Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2018, Oxford, UK, July 09-12, 2018, ACM, pp. 909–918, doi:10.1145/3209108.3209180