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 μ–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.

A. Indrzejczak, M. Zawidzki (Eds): 10th International Conference on Non-Classical Logics. Theory and Applications (NCL 2022)
EPTCS 358, 2022, pp. 84–101, doi:10.4204/EPTCS.358.7

2 Classical logics

The set $\mathcal{P} \triangleq \{\mathsf{R}_1,\dots,\mathsf{R}_n,\dots\}$ is a set of *predicate symbols* of arity $k_1,\dots,k_n,\dots$ respectively (one of them can be the identity predicate = of arity 2), $\mathcal{V} \triangleq \{v_1,\dots,v_n,\dots\}$ is a set of *variables* and $\mathcal{C} \triangleq \{c_1,\dots,c_n,\dots\}$ is a set of *constants*. Each of these sets can be finite or infinite. $v_1,v_2,v_3,\dots$ are the names of the variables and we use the expressions $x,x_1,x_2,\dots,y,y_1,y_2,\dots,z,z_1,z_2,\dots$ to refer to arbitrary variables or constants, which can be for example $v_{42},v_5,c_{101},c_{21},\dots$

The *first-order language* $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}$ is defined inductively by the following grammars in BNF:

$$\mathcal{L}^{\mathcal{VC}}_{\mathsf{FOL}}: \; t \; ::= \; x \; | \; \mathsf{c}$$
$$\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}: \; \varphi \; ::= \; \mathsf{R}t\dots t \; | \; \bot \; | \; (\varphi\to\varphi) \; | \; \forall x\varphi$$

where $x\in\mathcal{V}$, $\mathsf{c}\in\mathcal{C}$, $t\in\mathcal{L}^{\mathcal{VC}}_{\mathsf{FOL}}$ and $\mathsf{R}\in\mathcal{P}$. Elements of $\mathcal{L}^{\mathcal{VC}}_{\mathsf{FOL}}$ are called *terms* and elements of $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}$ are called *first–order formulas*. Formulas of the form $\mathsf{R}t_1\dots t_k$ are called *atomic formulas*. If $\varphi\in\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}$, 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 $\exists x\varphi\triangleq\neg\forall x\neg\varphi$, $\forall x_1\dots x_n\varphi\triangleq\forall x_1\dots\forall x_n\varphi$, $\exists x_1\dots x_n\varphi\triangleq\exists x_1\dots\exists x_n\varphi$ and $\forall\overline{x}\varphi\triangleq\forall x_1\dots x_n\varphi$ if $\overline{x}=(x_1,\dots,x_n)$ is a tuple of variables.

Let $\varphi\in\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}$. 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. A variable which is not free (in φ) is *bound* (in φ). We say that a formula of $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}$ is a *sentence* (or is *closed*) when it contains no free variable. We denote by $\varphi(x_1,\dots,x_k)$ a formula of $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}$ whose free variables or constants coincide *exactly* with $x_1,\dots,x_k$. In doing so, we depart from the literature in which this notation means that the free variables of φ are *included* in $\{x_1,\dots,x_k\}$. Free variables may be used to bind elements of two different subformulas. For example, the formula $\mathsf{R}yx\vee\mathsf{R}'xz$ 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 $\mathsf{R}yx$ and $\mathsf{R}'xz$.

We denote by $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(\overline{x})$ the fragment of $\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}$ whose formulas all contain at least one free variable or constant.

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

- W is a non-empty set called the *domain*;
- $R_1,\dots,R_n,\dots$ are relations over W with the same arity as $\mathsf{R}_1,\dots,\mathsf{R}_n,\dots$ respectively;
- $c_1,\dots,c_n,\dots\in W$ 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 $\mathcal{M}_{\mathsf{FOL}}$.

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

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

In the literature [10], $(M,s)\models\varphi(x_1,\dots,x_k)$ is sometimes denoted $M\models\varphi(x_1,\dots,x_k)[w_1,\dots,w_k]$, $M\models\varphi[w_1/x_1,\dots,w_k/x_k]$ or simply $M\models\varphi[w_1,\dots,w_k]$, with $w_1=s(x_1),\dots,w_k=s(x_k)$. In that case, we say that (M,s) makes φ *true*. We say that the formula $\varphi \in \mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}$ 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 *elementarily equivalent* when they make true the same sentences.

A triple of the form $\left(\mathcal{L}_{\mathsf{FOL}}, \mathcal{E}_{\mathsf{FOL}}, \models_{\mathsf{FOL}}\right)$ is called the *first–order logic associated to $\mathcal{L}_{FOL}$ and $\mathcal{E}_{FOL}$*. If $\mathcal{L}_{\mathsf{FOL}} = \mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}$, the triple $\left(\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}, \mathcal{E}_{\mathsf{FOL}}, \models_{\mathsf{FOL}}\right)$ is called *pure predicate logic (associated to $\mathcal{E}_{FOL}$)*, if $\mathcal{L}_{\mathsf{FOL}} = \mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(\overline{x})$, the triple $\left(\mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(\overline{x}), \mathcal{E}_{\mathsf{FOL}}, \models_{\mathsf{FOL}}\right)$ 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 Non-classical logics

In this section, $\mathbb{A}$ is a set of *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 *multi-modal language* $\mathcal{L}_{\mathsf{ML}}$ is defined inductively by the following grammar in BNF:

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

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*;
- $R_1, \ldots, R_m, \ldots \subseteq W \times W$ are binary relations over W called *accessibility relations*;
- $P_1, \ldots, P_n, \ldots \subseteq W$ are unary relations interpreting the propositional letters of $\mathbb{P}$.

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 $p_i \in \mathbb{P}$ and all $j \in \mathbb{I}$,

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

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

3.2 Lambek calculus

The *Lambek language* $\mathcal{L}_{LC}$ is the set of formulas defined inductively by the following grammar in BNF:

$$\mathcal{L}_{\mathsf{LC}}: \ \varphi \ ::= \ p \ \mid \ (\varphi \otimes \varphi) \ \mid \ (\varphi \subset \varphi) \ \mid \ (\varphi \supset \varphi)$$

where $p \in \mathbb{A}$. A *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;
- $P_1, \dots, P_n, \dots \subseteq W$ 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 $\mathcal{E}_{\mathsf{LC}}$. We define the *satisfaction relation* $\models_{\mathsf{Int}} \subseteq \mathcal{E}_{\mathsf{LC}} \times \mathcal{L}_{\mathsf{LC}}$ by the following *truth conditions*. Below, we write $(M, w) \models \varphi$ for $((M, w), \varphi) \in \models_{\mathsf{LC}}$. For all Lambek models $M = (W, \{R, P_1, \dots, P_n, \dots\})$, all $w \in M$, all $\varphi, \psi \in \mathcal{L}_{\mathsf{LC}}$ and all $p_i \in \mathbb{P}$,

$$
\begin{array}{lll}
(M,w) \models p_i & \text{iff} & P_i(x) \text{ holds;} \\
(M,w) \models (\varphi \otimes \psi) & \text{iff} & \text{there are } v,u \in W \text{ such that } Rvuw, \\
 & & (M,v) \models \varphi \text{ and } (M,u) \models \psi; \\
(M,w) \models (\varphi \supset \psi) & \text{iff} & \text{for all } v,u \in W \text{ such that } Rwvu, \\
 & & \text{if } (M,v) \models \varphi \text{ then } (M,u) \models \psi; \\
(M,w) \models (\psi \subset \varphi) & \text{iff} & \text{for all } v,u \in W \text{ such that } Rvwu, \\
 & & \text{if } (M,v) \models \varphi \text{ then } (M,u) \models \psi.
\end{array}
$$

The triple $(\mathcal{L}_{\mathsf{LC}}, \mathcal{E}_{\mathsf{LC}}, \models_{\mathsf{LC}})$ 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* $\mathcal{L}_{\mathsf{Int}}$ is defined inductively by the following grammar in BNF:

$$\mathcal{L}_{\mathsf{Int}}: \ \varphi \ ::= \ \top \ | \ \bot \ | \ p \ | \ (\varphi \wedge \varphi) \ | \ (\varphi \vee \varphi) \ | \ (\varphi \Rightarrow \varphi) \ | \ \Diamond\varphi \ | \ \Box\varphi$$

where $p \in \mathbb{A}$. A *modal intuitionistic model* is a tuple $M = (W, \{R, R_\Diamond, P_1, \dots, P_n, \dots\})$ 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);
- $R_\Diamond \subseteq W \times W$ is a binary relation over W;
- $P_1, \dots, P_n, \dots \subseteq W$ are unary relations over W such that for all $v, w \in W$, if Rvw and $P_n(v)$ then $P_n(w)$.

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 $\mathcal{E}_{\mathsf{Int}}$. We define the *satisfaction relation* $\models_{\mathsf{Int}} \subseteq \mathcal{E}_{\mathsf{Int}} \times \mathcal{L}_{\mathsf{Int}}$ by the following *truth conditions*. Below, we write $(M, w) \models \varphi$ for $((M, w), \varphi) \in \models_{\mathsf{Int}}$. For all modal intuitionistic models $M = (W, \{R, R_\Diamond, P_1, \dots, P_n, \dots\})$, all $w \in M$, all $\varphi, \psi \in \mathcal{L}_{\mathsf{Int}}$ and all $p_i \in \mathbb{P}$,

$$
\begin{array}{lll}
(M,w) \models \top & & \text{always;} \\
(M,w) \models \bot & & \text{never;} \\
(M,w) \models p_i & \text{iff} & P_i(w) \text{ holds;} \\
(M,w) \models (\varphi \wedge \psi) & \text{iff} & (M,w) \models \varphi \text{ and } (M,w) \models \psi; \\
(M,w) \models (\varphi \vee \psi) & \text{iff} & (M,w) \models \varphi \text{ or } (M,w) \models \psi; \\
(M,w) \models (\varphi \Rightarrow \psi) & \text{iff} & \text{for all } v \in W \text{ such that } Rwv, \text{if } (M,v) \models \varphi \text{ then } (M,v) \models \psi; \\
(M,w) \models \Box\varphi & \text{iff} & \text{for all } v \in W \text{ such that } Rwv, \\
 & & \text{for all } u \in W \text{ such that } R_\Diamond vu, (M,u) \models \varphi; \\
(M,w) \models \Diamond\varphi & \text{iff} & \text{for all } v \in W \text{ such that } Rwv, \\
 & & \text{there is } u \in W \text{ such that } R_\Diamond vu \text{ and } (M,u) \models \varphi.
\end{array}
$$

The triple $(\mathcal{L}_{\mathsf{Int}}, \mathcal{E}_{\mathsf{Int}}, \models_{\mathsf{Int}})$ 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 $\forall x_1 \dots x_n(\pm_1 \mathsf{P}_1 x_1 \vee \dots \vee \pm_n \mathsf{P}_n x_n \vee \pm \mathsf{R} x_1 \dots x_n x)$ or $\exists x_1 \dots x_n(\pm_1 \mathsf{P}_1 x_1 \wedge \dots \wedge \pm_n \mathsf{P}_n x_n \wedge \pm \mathsf{R} x_1 \dots x_n x)$ where $\pm_i$ and $\pm$ are either empty or $\neg$. Likewise, propositional letters are defined by first-order formulas of the form $\pm \mathsf{P} x$. 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 $n \in \mathbb{N}^*$, $\mathfrak{S}_n$ 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 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* $\mathbb{P}$ 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 $(\sigma, \pm, Æ, k)$ 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.

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*:

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

The type signatures of $\wedge_k$ and $\vee_k$ are (k,k,k) and their arity is 2.

We say that a set of atomic connectives C *is complete for conjunction and disjunction* when it contains all conjunctions and disjunctions $\wedge_k, \vee_k$, for k ranging over all input types and output types of the atomic connectives of C. The set of atomic skeletons associated to C is denoted $\star(\mathsf{C})$, 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.* $\dashv$

Remark 1. The permutations σ 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* $\mathcal{L}_{\mathsf{C}}$ 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 $(k, k_1, \dots, k_n)$ and for all $\varphi_1, \dots, \varphi_n \in \mathcal{L}_{\mathsf{C}}$ of types $k_1, \dots, k_n$ respectively, we have that $\star(\varphi_1, \dots, \varphi_n) \in \mathcal{L}_{\mathsf{C}}$ and $\star(\varphi_1, \dots, \varphi_n)$ is of *type* k.

The *Boolean atomic language* $\mathcal{L}^{\mathbb{B}}_{\mathsf{C}}$ 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 $\mathbb{B}$:

- for all $\varphi, \psi \in \mathcal{L}^{\mathbb{B}}_{\mathsf{C}}$ of type k, we have that $(\varphi \wedge_k \psi), (\varphi \vee_k \psi) \in \mathcal{L}^{\mathbb{B}}_{\mathsf{C}}$.

Elements of $\mathcal{L}_{\mathsf{C}}$ are denoted $\varphi, \psi, \alpha, \dots$ The *type of a formula* $\varphi \in \mathcal{L}_C$ is denoted $k(\varphi)$. For all $\varphi_1, \dots, \varphi_n \in \mathcal{L}$ of type k, $\bigwedge\{\varphi_1, \dots, \varphi_n\}$ and $\bigvee\{\varphi_1, \dots, \varphi_n\}$ stand for $((\varphi_1 \wedge_k \varphi_2) \wedge_k \dots \wedge_k \varphi_n)$ and $((\varphi_1 \vee_k \varphi_2) \vee_k \dots \vee_k \varphi_n)$ respectively. When it is clear from the context, we will omit the subscript k in $\wedge_k, \vee_k$ and write them $\wedge, \vee$.

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

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

An *assignment* is a tuple $(w_1, \dots, w_k) \in W^k$ 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 $\mathcal{M}_{\mathsf{C}}$. $\dashv$

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* $\mathcal{L}_C$ *in* M, denoted $[\![\cdot]\!]^M : \mathcal{L}_{\mathsf{C}} \to \bigcup\limits_{k \in \mathbb{N}^*} W^k$, inductively as follows: for all propositional letters $p \in \mathsf{C}$ of skeleton $(\mathrm{Id}, \pm, \text{Æ}, k)$, all connectives $\star \in \mathsf{C}$ of skeleton $(\sigma, \pm, \text{Æ}, (k, k_1, \dots, k_n), (\pm_1, \dots, \pm_n))$ of arity $n > 0$ and all $k \in \mathbb{N}^*$, for all $\varphi, \psi, \varphi_1, \dots, \varphi_n \in \mathcal{L}_{\mathsf{C}}$, if $k(\varphi) = k(\psi) = k$,

$$
\begin{array}{rcl}
[\![p]\!]^M & \triangleq & \pm R_p \\
[\![(\varphi \wedge_k \psi)]\!]^M & \triangleq & [\![\varphi]\!]^M \cap [\![\psi]\!]^M \\
[\![(\varphi \vee_k \psi)]\!]^M & \triangleq & [\![\varphi]\!]^M \cup [\![\psi]\!]^M \\
[\![\star(\varphi_1, \dots, \varphi_n)]\!]^M & \triangleq & f_\star([\![\varphi_1]\!]^M, \dots, [\![\varphi_n]\!]^M)
\end{array}
$$

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

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

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 $R^{\pm\sigma}_\star \overline{w}_1 \dots \overline{w}_{n+1}$ holds iff $\pm R_\star \overline{w}_{\sigma^-(1)} \dots \overline{w}_{\sigma^-(n+1)}$ with the notations $+R_\star \triangleq R_\star$ and $-R_\star \triangleq W^{k+k_1+\dots+k_n} - R_\star$. 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

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 'orbits' of unary and binary signatures

$(M,\overline{w}) \in \mathcal{E}_{\mathsf{C}}$, $((M,\overline{w}),\varphi) \in \Vdash$ 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 class of *atomic logics* is defined by $\mathbb{L}_{\mathsf{GGL}} \triangleq \{(\mathcal{L}_{\mathsf{C}},\mathcal{E}_{\mathsf{C}},\Vdash) \mid \mathsf{C}$ is a finite set of atomic connectives and $\mathcal{E}_{\mathsf{C}}$ is a class of C–models$\}$. The atomic 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 call them *Boolean (basic) atomic logics* when their language includes the Boolean connectives $\mathbb{B}$. ⊣

Example 1 (Lambek calculus, modal logic)**.** The Lambek calculus, where $\mathsf{C} = \{p,\circ,\backslash,/\}$ is defined in Section 3, is an example of atomic logic. Here $\circ,\backslash,/$ are the connectives of skeletons $(\sigma_1,+,s_1)$, $(\sigma_5,-,s_3),(\sigma_3,-,s_2)$. Another example of atomic logic is modal logic where $\mathsf{C} = \{p,\top,\bot,\wedge,\vee,\Diamond,\Box\}$ is such that

- $\top,\bot$ are connectives of skeletons $(\mathrm{Id},+,\exists,1)$ and $(\mathrm{Id},-,\forall,1)$ respectively;
- $\wedge,\vee,\Diamond,\Box$ are connectives of skeletons $(\sigma_1,+,s_1)$, $(\sigma_1,-,s_4)$, $(\tau_2,+,t_1)$ and $(\tau_2,-,t_2)$ respectively;
- C-models $M = (W,\mathcal{R}) \in \mathcal{E}_{\mathsf{C}}$ are such that $R_\wedge = R_\vee = \{(w,w,w) \mid w \in W\}$, $R_\Diamond = R_\Box$ and $R_\top = R_\bot = W$.

Indeed, one can easily show that, with these conditions on the C–models of $\mathcal{E}_{\mathsf{C}}$, we have that for all $(M,w) \in \mathcal{E}_{\mathsf{C}}$, $(M,w) \Vdash \wedge(\varphi,\psi)$ iff $(M,w) \Vdash \varphi$ and $(M,w) \Vdash \psi$, and $(M,w) \Vdash \vee(\varphi,\psi)$ iff $(M,w) \Vdash \varphi$ or $(M,w) \Vdash \psi$. The Boolean conjunction and disjunction $\wedge$ and $\vee$ are defined using the connectives of $\mathbb{C}$ by means of special relations $R_\wedge$ and $R_\vee$. However, they could obviously be defined directly. Many more examples of atomic connectives are in Figure 2. They are in fact just examples of gaggle 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 $\mathbb{C}^*$ of *molecular skeletons* is the smallest set such that:

- $\mathbb{P} \cup \mathbb{B} \subseteq \mathbb{C}^*$ and $\mathbb{C}^*$ 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
The conjunction orbit		
$\varphi\ (\sigma_1,+,s_1)\ \psi$	$\exists vu\,(v \in \llbracket\varphi\rrbracket \wedge u \in \llbracket\psi\rrbracket \wedge Rvuw)$	$\varphi \circ \psi$ [16], $\varphi \otimes_3 \psi$ [3]
$\varphi\ (\sigma_2,-,s_2)\ \psi$	$\forall vu\,(v \in \llbracket\varphi\rrbracket \vee u \notin \llbracket\psi\rrbracket \vee -Rwuv)$	
$\varphi\ (\sigma_3,-,s_2)\ \psi$	$\forall vu\,(v \in \llbracket\varphi\rrbracket \vee u \notin \llbracket\psi\rrbracket \vee -Ruwv)$	$/$ [16], $\varphi \subset_2 \psi$ [3]
$\varphi\ (\sigma_4,+,s_1)\ \psi$ $= \psi\ (\sigma_1,+,s_1)\ \varphi$	$\exists vu\,(v \in \llbracket\varphi\rrbracket \wedge u \in \llbracket\psi\rrbracket \wedge Ruvw)$	
$\varphi\ (\sigma_5,-,s_3)\ \psi$ $= \psi\ (\sigma_2,-,s_2)\ \varphi$	$\forall vu\,(v \notin \llbracket\varphi\rrbracket \vee u \in \llbracket\psi\rrbracket \vee -Rwvu)$	$\backslash$ [16], $\varphi \supset_1 \psi$ [3]
$\varphi\ (\sigma_6,-,s_3)\ \psi$ $= \psi\ (\sigma_3,-,s_2)\ \varphi$	$\forall vu\,(v \notin \llbracket\varphi\rrbracket \vee u \in \llbracket\psi\rrbracket \vee -Rvwu)$	
The not–but orbit		
$\varphi\ (\sigma_1,+,s_6)\ \psi$	$\exists vu\,(v \notin \llbracket\varphi\rrbracket \wedge u \in \llbracket\psi\rrbracket \wedge Rvuw)$	$\varphi \succ_3 \psi$ [3]
$\varphi\ (\sigma_2,+,s_6)\ \psi$	$\exists vu\,(v \notin \llbracket\varphi\rrbracket \wedge u \in \llbracket\psi\rrbracket \wedge Rwuv)$	
$\varphi\ (\sigma_3,-,s_4)\ \psi$	$\forall vu\,(v \in \llbracket\varphi\rrbracket \vee u \in \llbracket\psi\rrbracket \vee -Ruwv)$	$\varphi \oplus_2 \psi$ [3]
$\varphi\ (\sigma_4,+,s_5)\ \psi$ $= \psi\ (\sigma_1,+,s_6)\ \varphi$	$\exists vu\,(v \in \llbracket\varphi\rrbracket \wedge u \notin \llbracket\psi\rrbracket \wedge Ruvw)$	
$\varphi\ (\sigma_5,+,s_5)\ \psi$ $= \psi\ (\sigma_2,+,s_6)\ \varphi$	$\exists vu\,(v \in \llbracket\varphi\rrbracket \wedge u \notin \llbracket\psi\rrbracket \wedge Rwvu)$	$\varphi \prec_1 \psi$ [3]
$\varphi\ (\sigma_6,-,s_4)\ \psi$ $= \psi\ (\sigma_3,-,s_4)\ \varphi$	$\forall vu\,(v \in \llbracket\varphi\rrbracket \vee u \in \llbracket\psi\rrbracket \vee -Rvwu)$	
The but–not orbit		
$\varphi\ (\sigma_1,+,s_5)\ \psi$	$\exists vu\,(v \in \llbracket\varphi\rrbracket \wedge u \notin \llbracket\psi\rrbracket \wedge Rvuw)$	$\varphi \prec_3 \psi$ [3]
$\varphi\ (\sigma_2,-,s_4)\ \psi$	$\forall vu\,(v \in \llbracket\varphi\rrbracket \vee u \in \llbracket\psi\rrbracket \vee -Rwuv)$	
$\varphi\ (\sigma_3,+,s_6)\ \psi$	$\exists vu\,(v \notin \llbracket\varphi\rrbracket \wedge u \in \llbracket\psi\rrbracket \wedge Ruwv)$	$\varphi \succ_2 \psi$ [3]
$\varphi\ (\sigma_4,+,s_6)\ \psi$ $= \psi\ (\sigma_1,+,s_5)\ \varphi$	$\exists vu\,(v \notin \llbracket\varphi\rrbracket \wedge u \in \llbracket\psi\rrbracket \wedge Ruvw)$	$\varphi \obslash \psi$ [12, 17]
$\varphi\ (\sigma_5,-,s_4)\ \psi$ $= \psi\ (\sigma_2,-,s_4)\ \varphi$	$\forall vu\,(v \in \llbracket\varphi\rrbracket \vee u \in \llbracket\psi\rrbracket \vee -Rwvu)$	$\varphi \oplus \psi$ [12, 17] $\varphi \oplus_1 \psi$ [3]
$\varphi\ (\sigma_6,+,s_5)\ \psi$ $= \psi\ (\sigma_3,+,s_6)\ \varphi$	$\exists vu\,(v \in \llbracket\varphi\rrbracket \wedge u \notin \llbracket\psi\rrbracket \wedge Rvwu)$	$\varphi \oslash \psi$ [12, 17]
The stroke orbit		
$\varphi\ (\sigma_1,+,s_7)\ \psi$	$\exists vu\,(v \notin \llbracket\varphi\rrbracket \wedge u \notin \llbracket\psi\rrbracket \wedge Rvuw)$	$\varphi\ \vert_3\ \psi$ [1, 11]
$\varphi\ (\sigma_2,+,s_7)\ \psi$	$\exists vu\,(v \notin \llbracket\varphi\rrbracket \wedge u \notin \llbracket\psi\rrbracket \wedge Rwuv)$	
$\varphi\ (\sigma_3,+,s_7)\ \psi$	$\exists vu\,(v \notin \llbracket\varphi\rrbracket \wedge u \notin \llbracket\psi\rrbracket \wedge Ruwv)$	
$\varphi\ (\sigma_4,+,s_7)\ \psi$ $= \psi\ (\sigma_1,+,s_7)\ \varphi$	$\exists vu\,(v \notin \llbracket\varphi\rrbracket \wedge u \notin \llbracket\psi\rrbracket \wedge Ruvw)$	
$\varphi\ (\sigma_5,+,s_7)\ \psi$ $= \psi\ (\sigma_2,+,s_7)\ \varphi$	$\exists vu\,(v \notin \llbracket\varphi\rrbracket \wedge u \notin \llbracket\psi\rrbracket \wedge Rwvu)$	$\varphi\ \vert_1\ \psi$ [1, 11]
$\varphi\ (\sigma_6,+,s_7)\ \psi$ $= \psi\ (\sigma_3,+,s_7)\ \varphi$	$\exists vu\,(v \notin \llbracket\varphi\rrbracket \wedge u \notin \llbracket\psi\rrbracket \wedge Rvwu)$	$\varphi\ \vert_2\ \psi$ [1, 11]

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 $(k_1, k_1^1, \ldots, k_{a_1}^1)$, $\ldots, (k_n, k_1^n, \ldots, k_{a_n}^n)$ respectively, the connective $\star(c_1, \ldots, c_n)$ belongs to $\mathbb{C}^*$, its *type signature* is $(k, k_1^1, \ldots, k_{a_1}^1, \ldots, k_1^n, \ldots, k_{a_n}^n)$ and its *arity* is $a_1 + \ldots + a_n$.

We define the *quantification signature* $Æ(c)$ of $c = \star(c_1, \ldots, c_n)$ by $Æ(c) \triangleq Æ(\star)$.

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. $\dashv$

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

$$\begin{aligned}
c(x) &\triangleq \forall y(\mathsf{R}xy \to \forall z(\mathsf{R}_\Diamond yz \to \mathsf{P}(z)))\\
c'(x) &\triangleq \forall y(\mathsf{R}xy \to \exists z(\mathsf{R}_\Diamond yz \wedge \mathsf{P}(z)))\\
\star_1(x) &\triangleq \forall y(\mathsf{R}xy \to \mathsf{P}(y))\\
\star_2(x) &\triangleq \forall z(\mathsf{R}_\Diamond yz \to \mathsf{P}(z))\\
\star_3(x) &\triangleq \exists z(\mathsf{R}_\Diamond yz \wedge \mathsf{P}(z))
\end{aligned}$$

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 $c = \star_1(\star_2(id_1))$ and $c' = \star_1(\star_3(id_1))$. 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. $\dashv$

Definition 6 (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 $\star \in \mathsf{C}$ 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.

The *Boolean molecular language* $\mathcal{L}_\mathsf{C}^\mathbb{B}$ 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 $\varphi, \psi \in \mathcal{L}_\mathsf{C}^\mathbb{B}$ of type k, we have that $(\varphi \wedge_k \psi), (\varphi \vee_k \psi) \in \mathcal{L}_\mathsf{C}^\mathbb{B}$.

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 $\mathcal{L}_\mathsf{C}$ 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. $\dashv$

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 $\overline{w}(M,\mathsf{C}) \triangleq \overline{w}(M,\mathsf{C}')$. 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 $k \in \mathbb{N}^*$,

$$\begin{array}{rcl} [\![id_k(\varphi)]\!]^M & \triangleq & [\![\varphi]\!]^M \\ [\![\star(c_1,\dots,c_n)\big(\varphi_1^1,\dots,\varphi_1^{k_1},\dots,\varphi_n^1,\dots,\varphi_n^{k_n}\big)]\!]^M & \triangleq & f_\star\big([\![c_1(\varphi_1^1,\dots,\varphi_1^{k_1})]\!]^M,\dots,[\![c_n(\varphi_n^1,\dots,\varphi_n^{k_n}]\!]^M\big) \end{array}$$

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*. The logics of the form $(\mathcal{L}_\mathsf{C},\mathcal{M}_\mathsf{C},\Vdash)$ 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 $\star$ be a n–ary connective of skeleton$(\sigma,\pm,Æ,\overline{k},\pm_1,\dots,\pm_n)$. The *Boolean negation of* $\star$ is the connective $-\star$ of skeleton $(\sigma,-\pm,-Æ,\overline{k},-\pm_1,\dots,-\pm_n)$ 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 $\varphi = \star(\varphi_1,\dots,\varphi_n)$ is an atomic formula, the *Boolean negation of* φ is the formula $-\varphi \triangleq -\star(\varphi_1,\dots,\varphi_n)$. ⊣

Proposition 1. *Let C be a set of atomic connectives such that $-\star \in \mathsf{C}$ for all $\star \in \mathsf{C}$. Let $\varphi \in \mathcal{L}_\mathsf{C}$ and M be a C–model. Then, for all $\overline{w} \in \overline{w}(M,\mathsf{C})$, $\overline{w} \in [\![-\varphi]\!]^M$ 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 $\star \in \mathsf{C}$ 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 for all $\star \in \mathsf{C}$, if $\{M,M'\} = \{M_1,M_2\}$, then for all $\overline{w}_1,\dots,\overline{w}_n,\overline{w'}_1,\dots,\overline{w'}_n,\overline{w},\overline{w'} \in \overline{w}(M,\mathsf{C}) \cup \overline{w}(M',\mathsf{C})$,

1. if $\star$ is an 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 $(\sigma,\pm,\exists,\overline{k},(\pm_1,\dots,\pm_n))$ and we have $\overline{w}Z\overline{w'}$ and $R_\star^{\pm\sigma}\overline{w}_1\dots\overline{w}_n\overline{w}$, then $\exists\overline{w'}_1,\dots,\overline{w'}_n\big(\overline{w}_1 \bowtie \overline{w'}_1 \wedge \overline{w}_2 \bowtie \overline{w'}_2 \wedge \dots \wedge \overline{w}_n \bowtie \overline{w'}_n \wedge R_\star'^{\pm\sigma}\overline{w'}_1\dots\overline{w'}_n\overline{w'}\big)$;
3. if $\star$ has skeleton $(\sigma,\pm,\forall,\overline{k},(\pm_1,\dots,\pm_n))$ and we have $\overline{w}Z\overline{w'}$ and $-R_\star'^{\pm\sigma}\overline{w'}_1\dots\overline{w'}_n\overline{w'}$, then $\exists\overline{w}_1,\dots,\overline{w}_n\big(\overline{w}_1 \bowtie \overline{w'}_1 \wedge \overline{w}_2 \bowtie \overline{w'}_2 \wedge \dots \wedge \overline{w}_n \bowtie \overline{w'}_n \wedge -R_\star^{\pm\sigma}\overline{w}_1\dots\overline{w}_n\overline{w}\big)$;

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 $n = 0$. 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 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 2. *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 $\overline{w} \in \overline{w}(M,C)$, all $\overline{w'} \in \overline{w}(M',C)$, if $\overline{w} Z \overline{w'}$ then $(M,\overline{w}) \rightsquigarrow_C (M',\overline{w'})$.*

Example 3 (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. ⊣

Example 4 (Lambek calculus)**.** Let us consider the connectives of the Lambek calculus: $\mathsf{C} = \{p,\circ,\backslash,/\}$ where p has skeleton $(\mathrm{Id},+,\exists,1)$, $\circ$ has skeleton $(\sigma_1,+,s_1)$, $\backslash$ has skeleton $(\sigma_5,-,s_3)$ 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 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 $w,v,u \in M$ and all $w',v',u' \in M'$,

- if wZw' and $w \in [\![p]\!]$ then $w' \in [\![p]\!]$ (condition for p);
- if wZw' and $Rvuw$ then there are $v',u' \in W'$ such that vZv', uZu' and $Rv'u'w'$ (condition for $\circ = (\sigma_1,+,s_1)$);
- if vZv' and $Rv'u'w'$ then there are $u,w \in W$ such that $u'Zu$, wZw' and $Rvuw$ (condition for $\backslash = (\sigma_5,-,s_3)$);
- if uZu' and $Rv'u'w'$ then there are $v,w \in W$ such that $v'Zv$, wZw' and $Rvuw$ (condition for $/ = (\sigma_3,-,s_2)$). ⊣

The following proposition shows that the notions of 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 $\mathsf{C} = \{p,\neg p,\wedge,\vee,\Diamond,\Box\}$ be the connectives of Example 3 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.*

- *Let $\mathsf{C} = \{p, \circ, \backslash, /\}$ be the connectives of Example 4 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: $\mathsf{C} = \{p, \bot, \top, \wedge, \vee, \Rightarrow\}$ where p has skeleton $(\mathrm{Id}, +, \exists, 1)$, $\top$ has skeleton $(\mathrm{Id}, +, \exists, 1)$, $\bot$ has skeleton $(\mathrm{Id}, -, \forall, 1)$, $\wedge$ and $\vee$ are Boolean connectives and $\Rightarrow$ has skeleton $(\sigma_5, -, s_3)$ (here, $\top$ and $\bot$ are represented by specific propositional letters of respective signatures $(\mathrm{Id}, +, \exists, 1)$ and $(\mathrm{Id}, -, \forall, 1)$). Let $M_1 = (W_1, R_1, P)$ and $M_2 = (W_2, R_2, P)$ be two intuitionistic models. Following the results of [3, Section 8], we represent these intuitionistic models by the C–models $M_1^{\Rightarrow} = (W_1, R_{1,\Rightarrow}, P)$ and $M_2^{\Rightarrow} = (W_2, R_{2,\Rightarrow}, P)$ respectively such that for all $u_1, v_1, w_1 \in W_1$ and all $u_2, v_2, w_2 \in W_2$,

$$R_{1,\Rightarrow} u_1 v_1 w_1 \text{ iff } R_1 u_1 w_1 \text{ and } R_1 v_1 w_1 \tag{1}$$

$$R_{2,\Rightarrow} u_2 v_2 w_2 \text{ iff } R_2 u_2 w_2 \text{ and } R_2 v_2 w_2 \tag{2}$$

One can show [3] that for all $\varphi \in \mathcal{L}_{\mathsf{C}}$ and all $w_1 \in W_1$, $M_1, w_1 \models \varphi$ iff $M_1^{\Rightarrow}, w_1 \Vdash \varphi$ (and likewise for M_2 and $M_2^{\Rightarrow}$). Now, a binary relation Z between $M_1^{\Rightarrow}$ and $M_2^{\Rightarrow}$ is a C–bisimulation between $M_1^{\Rightarrow}$ and $M_2^{\Rightarrow}$ iff for all $M, M' \in \{M_1^{\Rightarrow}, M_2^{\Rightarrow}\}$, all $w, w', v', u' \in \overline{w}(M, \mathsf{C}) \cup \overline{w}(M', \mathsf{C})$ and all $p \in \mathbb{P}$,

- if wZw' and $w \in [\![p]\!]$ then $w' \in [\![p]\!]$ (condition for p);
- if vZv' and $R'_{\Rightarrow} v'u'w'$ then there are $u, w \in W$ such that $u'Zu$, wZw' and $R_{\Rightarrow} vuw$ $(*)$ (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 vZv' and $R'v'w'$ and $R'u'w'$ then there are $u, w \in W$ such that $u'Zu$, wZw' and Rvw and Ruw $(**)$.

We will show in Section 8 that condition $(**)$ is equivalent on ω–saturated models to Olkhovikov's condition "step" of [19, Definition 1] of his "basic asimulation". $\dashv$

5.2 Molecular logics

Definition 11 (C–bisimulation for molecular connectives)**.** Let C be a set of molecular connectives and let $M_1 = (W_1, \mathcal{R}_1)$ 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 decomposition tree T_{c_0}. We associate to each vertex $c \in V_{c_0}$ a binary relation $Z_c \subseteq \bigcup_{k \in \mathbb{N}^*} (W_1^k \times W_2^k) \cup (W_2^k \times W_1^k)$. The set of such binary relations is denoted $\{Z\} \cup \bigcup_{c_0 \in \mathsf{C}} \{Z_c \mid c \in V_{c_0}\}$ and is such that if c is id_k for some $k \in \mathbb{N}^*$ then Z_c is Z and we have that $Z \subseteq \bigcap \{Z_c \mid c \in \mathsf{C}\}$. We say that this set of binary relations is a *C–bisimulation* between M_1 and M_2 when for all $c_0 \in \mathsf{C}$, all vertices $c \in V_{c_0}$, 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 an propositional letter p then, $\overline{w} Z_c \overline{w'}$ and $\overline{w} \in [\![p]\!]$ imply $\overline{w'} \in [\![p]\!]$;
2. if c has skeleton $\star(c_1, \ldots, c_n)$ with $\star = (\sigma, \pm, \exists, \overline{k}, (\pm_1, \ldots, \pm_n))$ and we have $\overline{w} Z_c \overline{w'}$ and $R_\star^{\pm\sigma} \overline{w}_1 \ldots \overline{w}_n \overline{w}$, then
$\exists \overline{w'}_1 \overline{w'}_2 \ldots \overline{w'}_n \left(\overline{w}_1 \bowtie_{c_1} \overline{w'}_1 \wedge \overline{w}_2 \bowtie_{c_2} \overline{w'}_2 \wedge \ldots \wedge \overline{w}_n \bowtie_{c_n} \overline{w'}_n \wedge R'^{\pm\sigma}_\star \overline{w'}_1 \ldots \overline{w'}_n \overline{w'} \right)$;
3. if c has skeleton $\star(c_1, \ldots, c_n)$ with $\star = (\sigma, \pm, \forall, \overline{k}, (\pm_1, \ldots, \pm_n))$ and we have $\overline{w} Z_c \overline{w'}$ and $-R'^{\pm\sigma}_\star \overline{w'}_1 \ldots \overline{w'}_n \overline{w'}$, then
$\exists \overline{w}_1 \overline{w}_2 \ldots \overline{w}_n \left(\overline{w}_1 \bowtie_{c_1} \overline{w'}_1 \wedge \overline{w}_2 \bowtie_{c_2} \overline{w'}_2 \wedge \ldots \wedge \overline{w}_n \bowtie_{c_n} \overline{w'}_n \wedge -R_\star^{\pm\sigma} \overline{w}_1 \ldots \overline{w}_n \overline{w} \right)$;

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

When such a set of binary relations exists and is such that $\overline{w} Z \overline{w'}$, 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 $n = 0$.

Definition 12. Let C be a set of molecular connectives. For all $c_0 \in \mathsf{C}$ and all vertex c of the decomposition tree T_{c_0}, we define the language $\mathcal{L}_{c\mathsf{C}}$ as follows:

$$\mathcal{L}_{c\mathsf{C}} \triangleq \begin{cases} \{c(\varphi_1,\dots,\varphi_n) \mid \varphi_1,\dots,\varphi_n \in \mathcal{L}_{\mathsf{C}}\} & \text{if } c \text{ is of arity } n > 0 \\ \{p\} & \text{if } c = p \text{ is a propositional letter} \\ \mathcal{L}_{\mathsf{C}} & \text{if } c \text{ is } id_k \text{ for some } k \in \mathbb{N}^*. \end{cases}$$

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 4. *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 V_c be the vertices of the decomposition tree T_c. Let $\{Z\} \cup \bigcup_{c_0 \in \mathsf{C}_0} \{Z_c \mid c \in V_{c_0}\}$ 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 13 (Uniform connective)**.** A *uniform connective* is a molecular connective c whose skeleton is of the form $\star(c_1,\dots,c_n)$ with $\star = (\sigma,\pm,\text{Æ},\overline{k},(\pm_1,\dots,\pm_n)) \in \mathbb{C}$ such that

1. $n \geq 1$ and $c_1,\dots,c_n$ are molecular skeletons of arity 1;
2. for all $j \in [\![1;n]\!]$ such that $c_j \neq id_k$ for all $k \in \mathbb{N}^*$, $\text{Æ}(c_j) = \begin{cases} \forall & \text{if } \pm_j = + \\ \exists & \text{if } \pm_j = - \end{cases}$;
3. if c_0 is a molecular skeleton appearing in the decomposition tree of c of the form $c_0 = \star_0(c'_1,\dots,c'_m)$ such that the tonicity signature of $\star_0$ is $(\pm_1,\dots,\pm_m)$, then for all $i \in [\![1;m]\!]$, $\text{Æ}(c'_i) = \pm_i \text{Æ}(c_0)$. ⊣

According to our definition, molecular connectives of the form $\star(c(c'(c'_1,c'_2)))$ cannot be uniform connectives, unless c'_1 or c'_2 is a propositional letter. This is due to our first condition: in that case, $c(c'(c'_1,c'_2))$ 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 $c_i^1,\dots,c_i^{m_i}$, each of arity 1, so that molecular connectives are essentially of the form $\star(c_1^1(\dots c_1^{m_1-1}(c_1^{m_1})),\dots,c_n^1(\dots c_n^{m_n-1}(c_n^{m_n})))$. Basically, uniform connectives are such that the quantification patterns of their successive internal connectives are essentially of the form $\exists\dots\exists\dots$ or $\forall\dots\forall\dots$

Example 6 (Modal intuitionistic logic)**.** Let $\mathsf{C} = \{p,\top,\bot,\wedge,\vee,\Rightarrow,\star,-\star'\}$ where $\star,\star' \in \mathbb{C}^*$ are the molecular connectives c,c' of Example 2 and where $\{p,\top,\bot,\wedge,\vee,\Rightarrow\}$ are defined in Example 5. The connectives of C are all uniform connectives. Note that $\star'$ is not a uniform connective and that is why we consider $-\star'$ for the moment, which is a uniform connective. We are going to see that we can easily get the bisimilarity condition for $\star'$ from the bisimilarity condition for $-\star'$.

Let $M_1 = (W_1,\{R_1,R_{1,\diamond},P\})$ and $M_2 = (W_2,\{R_2,R_{2,\diamond},P\})$ be two modal intuitionistic models. The set of binary relations $\{Z,Z^{\star}_{\star_2},Z^{-\star'}_{\star_3}\}$ is a C–bisimulation iff for all $M,M' \in \{M_1,M_2\}$ with $M = (W,\{R,R_\diamond,P\})$ and $M' = (W',\{R',R'_\diamond,P\})$, all $w,v,u,w',v',u' \in \overline{w}(M,\mathsf{C}) \cup \overline{w}(M',\mathsf{C})$ 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 $R'u'w'$ then there are $u, w \in W$ such that $u'Zu$, wZw' and Rvw and Ruw (condition for $\star$, like in Example 5);
- if $R'w'v'$ and wZw' then there is $v \in W$ such that $vZ^{\star}_{\star_2}v'$ and Rwv,
  if $wZ^{\star}_{\star_2}w'$ and $R'_{\diamond}w'v'$ 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 $v'Z^{-\star'}_{\star_3}v$ and $R'w'v'$,
  if $wZ^{\star'}_{\star_3}w'$ and $R_{\diamond}wv$ then there is $v' \in W'$ such that vZv' and $Rw'v'$
  (condition for $-\star' = -\star_1(\star_3)$).

To obtain the bisimilarity condition for $\star'$, it suffices to observe that for all $c \in \mathbb{C}^*$, it holds that $(M,w) \rightsquigarrow_{\{-c\}} (M',w')$ iff $(M',w') \rightsquigarrow_{\{c\}} (M,w)$. So, we just have to replace wZw' by $w'Zw$ in the condition above. We obtain:

($\ast$) if $w'Zw$ and Rwv then there is $v' \in W'$ such that $v'Z_{\star_3}v$ and $R'w'v'$,
if $wZ_{\star_3}w'$ and $R_{\diamond}wv$ then there is $v' \in W'$ such that vZv' and $Rw'v'$.

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

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 $F \subseteq \mathcal{P}(I)$ 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 $I - X \notin U$. $\dashv$

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 $\prod_U W_i \triangleq \{\prod_U w_i \mid (w_i)_{i\in I} \in \prod_{i\in I} W_i\}$. $\dashv$

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 $x \in \mathcal{V}$, $(\prod_U s_i)(x) = \prod_U s_i(x)$ 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 $R^i_{\star}$ of M_i, the $n+1$–ary relation $\prod_U R_{\star} \in \mathcal{R}_U$ is defined for all $\prod_U w^1_i, \dots, \prod_U w^{n+1}_i \in W_U$ by $\prod_U R_{\star} \prod_U w^1_i \dots \prod_U w^{n+1}_i$ iff $\{i \in I \mid R^i_{\star} w^1_i \dots w^{n+1}_i\} \in U$. $\dashv$

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. $\dashv$

7 A generic van Benthem characterization theorem

In this section we generalize the van Benthem characterization theorem for modal logic [9, Theorem 2.68] 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 $\overline{x} \in (\mathcal{V} \cup \mathcal{C})^k$, we define the mappings $ST_{\overline{x}} : \mathcal{L}^k_{\mathsf{C}} \to \mathcal{L}^{\mathcal{P}}_{\mathsf{FOL}}(\overline{x})$, where $\mathcal{L}^k_{\mathsf{C}}$ is the set of formulas of $\mathcal{L}_{\mathsf{C}}$ of type k, as follows: for all atoms $p \in \mathsf{C}$, all $\star \in \mathsf{C}$ of skeleton $(\sigma, \pm, Æ, (k, k_1, \dots, k_n), (\pm_1, \dots, \pm_n))$ and all $\varphi_1, \dots, \varphi_n \in \mathcal{L}_{\mathsf{C}}$,

$$\begin{aligned}
ST_{\overline{x}}(p) &\triangleq \mathsf{P}(\overline{x})\\
ST_{\overline{x}}(\varphi \wedge_k \psi) &\triangleq ST_{\overline{x}}(\varphi) \wedge ST_{\overline{x}}(\psi)\\
ST_{\overline{x}}(\varphi \vee_k \psi) &\triangleq ST_{\overline{x}}(\varphi) \vee ST_{\overline{x}}(\psi)\\
ST_{\overline{x}}(\star(\varphi_1, \dots, \varphi_n)) &\triangleq Æ\overline{x}_1 \dots \overline{x}_n \left(*_1 ST_{\overline{x}_1}(\varphi_1) \times \dots \times *_n ST_{\overline{x}_n}(\varphi_n) \times \mathsf{R}^{\pm\sigma}_{\star} \overline{x}_1 \dots \overline{x}_n \overline{x} \right)
\end{aligned}$$

where $\overline{x}_1, \dots, \overline{x}_n$ are tuples of free variables of size $k_1, \dots, k_n$,

$$\times = \begin{cases} \wedge & \text{if } Æ = \exists \\ \vee & \text{if } Æ = \forall \end{cases} \text{ and for all } j \in [\![1;n]\!],\ *_j = \begin{cases} \neg & \text{if } \pm_j = - \\ \text{empty} & \text{if } \pm_j = + \end{cases}.$$

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

The above translations canonically extend to molecular logics. Indeed, if C is a set of molecular connectives, every molecular formula of $\mathcal{L}_{\mathsf{C}}$ can be viewed as a formula of $\mathcal{L}_{\mathsf{C}'}$, 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 $(M, \overline{w})$ be a pointed C–model, let $\varphi \in \mathcal{L}_{\mathsf{C}}$ of type k and let $\overline{x} \in \mathcal{V}^k$. Then, $(M, \overline{w}) \Vdash \varphi$ iff $\left(M, s^{\overline{w}}_{\overline{x}}\right) \Vdash ST_{\overline{x}}(\varphi)$.*

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

1. *There exists a formula $\psi \in \mathcal{L}_{\mathsf{C}}$ such that $\varphi(\overline{x}) \leftrightarrow ST_{\overline{x}}(\psi)$ is valid on $ST_{\overline{x}}(\mathcal{E}_{\mathsf{C}})$;*
2. $\varphi(\overline{x})$ is invariant for C–bisimulations on $\mathcal{E}_{\mathsf{C}}$*, that is, for all pointed C–models $(M, \overline{w}), (M', \overline{w'})$ of $\mathcal{E}_{\mathsf{C}}$ of type k such that $(M, \overline{w}) \to_{\mathsf{C}} \left(M', \overline{w'}\right)$, we have that $\left(M, s^{\overline{w}}_{\overline{x}}\right) \models \varphi(\overline{x})$ implies $\left(M', s^{\overline{w'}}_{\overline{x}}\right) \models \varphi(\overline{x})$.*

Remark 2. The assumption that $ST_{\overline{x}}(\mathcal{E}_{\mathsf{C}})$ 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 [10, Corollary 6.1.16]. 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, 20] who investigates generalizations of the van Benthem characerization theorm. The publications [18] and [19] deal in particular with (modal) intuitionistic (predicate) 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 ω–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 $u'Zu$, wZw' and Rvw and Ruw;*
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* $\mu = \psi(\mathsf{P}_1(x), \dots, \mathsf{P}_k(x))$ is a Boolean combination of the unary predicates $\mathsf{P}_1(x), \dots, \mathsf{P}_k(x)$. A *k–ary $\forall$–guarded x_1–connective of degree $n+1$* is a formula of the form $\forall x_2 \dots x_{m+1} \left(\bigwedge_{i=1}^m S_i(x_i, x_{i+1}) \to \mu^-\right)$ where $S_1, \dots, S_m$ are binary predicates and μ^- 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 $\mathsf{R}x_1 \dots x_m x_{m+1}$ for $\bigwedge_{i=1}^m S_i(x_i, x_{i+1})$ then $\forall x_2 \dots x_{m+1}$ $\left(\bigwedge_{i=1}^m S_i(x_i, x_{i+1}) \to \mu^-\right)$ can be viewed as the first–order formula with free variable x_1 defining a molecular connective: $\forall x_2 \dots x_{m+1} (x_2 \in [\![\bot]\!] \vee \dots \vee x_m \in [\![\bot]\!] \vee x_{m+1} \in [\![\mu^-]\!] \vee -\mathsf{R}x_1 \dots x_m x_{m+1})$. 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], neighbourhoud 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 coalgebric modal logic [24] is w.r.t. coalgebric 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:10.1007/3-7643-7304-0_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:10.1007/3-540-61604-7_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.