UNIVERSITAT DE BARCELONA

MASTER'S THESIS

MASTER IN PURE AND APPLIED LOGIC

Unification in intuitionistic logic

Sebastián R. Cristancho S.

Supervised by:
Tommaso Moraschini
Amanda Vidal Wandelmer

July, 2023

Acknowledgements

First of all, I would like to thank Tommaso and Amanda for their support, patience and help during the realization of this work.

Secondly, I express my deep gratitude to the IIIA-CSIC (Instituto de Investigación en Inteligencia Artificial - Consejo Superior de Investigaciones Científicas) for financing this master's thesis through the scholarship of introduction to research JAEIntroICU-2021-IIIA-02.

Finally, I thank my family for dealing with my absence, and I thank Jesusa for dealing with my presence.

Contents

CHAPTER 1

Introduction

The concept of unification has been widely studied from a logical perspective. In the context of logic, a formula A is said to be *unifiable* in a logic $\vdash$ if there is a substitution σ that turns A into a theorem of $\vdash$. In this case, we say that σ is a unifier (in $\vdash$) of A, or that A is unifiable (in $\vdash$) by σ.

Given a logic $\vdash$ and a unifiable formula A (in $\vdash$), there is a natural way to compare its unifiers in terms of generality using the fact that, up to logical equivalence, some unifiers can be 'obtained' from others. More precisely, we say that the unifier σ_1 of A is less general than the unifier σ_2 of A if there is a substitution τ such that $\sigma_1(x)$ is logically equivalent to $\tau(\sigma_2(x))$ in $\vdash$ for all propositional variables x in the domain of σ_1 and σ_2. This gives rise to a hierarchy among the set of unifiers of A, where the unifiers in lower levels can be obtained from the unifiers in upper levels. A *basis of unifiers* of a unifiable formula A is a set of incomparable elements that 'generates' any other unifier of A. The study of the hierarchy among unifiers rises some interesting questions: Given a unifiable formula A in $\vdash$, is there a basis of unifiers of A? If so, is it finite or infinite? If it is finite, does it have one or more elements? These questions can be stated not only for formulas, but for logics in general. It prompts a classification of logics in different types:

- *Logics of unification type* 1 *(or unitary)*: logics where every unifiable formula has a basis of unifiers of size 1.

- *Logics of unification type* ω *(or finitary)*: logics where every unifiable formula has a finite basis of unifiers and there is at least one unifiable formula that does not have a basis of unifiers of size 1.

- *Logics of unification type* ∞ *(or infinitary)*: logics where every unifiable formula has an infinite basis of unifiers and there is at least one unifiable formula that does not have a finite basis of unifiers.

- *Logics of unification type* 0 *(or nullary)*: logics where there is a unifiable formula that does not have a basis of unifiers.

Solving the Unification Problem for a certain logic consists in finding the unification type of such logic. Classical propositional logic (**CPL**) has unitary unification type.

This means that for any unifiable formula A there is a unifier of A from which we can obtain all the other unifiers of A. Such unifiers are called most general unifiers. In intuitionistic propositional logic (**IPL**), however, the solution to the Unification Problem is not that nice. In example 3.34, we will show that in **IPL** there is no most general unifier for the formula $x \vee \neg x$. Nevertheless, even thought unification in **IPL** is not unitary, Ghilardi proved that the unification type of **IPL** is finitary [6], that is to say, for every unifiable formula in **IPL** there is always a finite set of incomparable unifiers that can generate any other unifier of such formula. The main goal of this work is to present a clear and detailed proof of this outstanding result. We will follow the structure of Ghilardi's proof but we will also introduce some slight modifications to adapt it to the conventions and results developed in more recent works.

The first chapter of this document includes all the preliminaries needed for the proof. We present the basic concepts and results about Kripke semantics and finish with a version of the completeness theorem that will allow us to work exclusively with finite rooted models.

The second chapter is the core part of this work. We start by presenting the basic definitions concerning unification. Then, we introduce the operator $(-)^\sigma$ as the semantical counterpart of substitutions. More precisely, while a substitution $\sigma\colon F(X) \to F(Y)$ gives us a way to transform formulas in $F(X)$ (formulas with variables in a set of variables X) into formulas in $F(Y)$, the $(-)^\sigma$ operator allows us to transform models in $\mathcal{M}_Y$ (models whose valuation takes values on $\mathcal{P}(Y)$) into models of $\mathcal{M}_X$ using the substitution σ. The operator $(-)^\sigma$ transforms a model in $\mathcal{M}_Y$ into a model in $\mathcal{M}_X$ preserving its frame and defining a new valuation on it following the rule: for all $x \in X$, x holds in a world m of the new model if and only if $\sigma(x)$ held at m in the original model. This operator will be the main tool to reason about substitutions from a semantical perspective in order to prove some of their properties. The first noticeable result obtained with this tool is a characterization of unifiable formulas as satisfiable formulas stated in Proposition 3.17.

Before moving on to the proof of the Unification Theorem for **IPL**, we prove the corresponding result for **CPL** (Section 3.4). Although the Unification Theorem for **CPL** is simpler, it will serve as a motivation for the intuitionistic case and will display the engine used later.

The proof of the Unification Theorem for **IPL** is the assembly of many concepts. Two key notions are the notion of projective formulas and the notion of implicational complexity. *Projective formulas* are the formulas with a unifier σ such that $\vdash x \leftrightarrow \sigma(x)$. We say that a formula B is projective with σ, if σ is a unifier witnessing the projectivity of B. The *implicational complexity* of a formula A, written as $c(A)$, refers to the maximum number of nested implications appearing on it. The main step in the proof of the Unification Theorem for **IPL**, and the part that will require most of our efforts, is Lemma 3.76, which states that any unifier σ of A is also unifier of some projective formula B of less or equal complexity than A such that $B \models A$. But projective formulas have the nice property that if B is projective with τ, then τ is a most general unifier of B (Remark 3.25), that is, any unifier of B can be obtained from τ. Hence, Lemma 3.76 guarantees that any unifier σ of a unifiable formula A can be obtained from the substitutions witnessing the projectivity of the projective formulas of less or equal complexity than A that imply A. If a formula has a most general unifier, moreover, it is unique up to logical equivalence. Thus, the unifiers of A can be obtained from a set of substitutions of cardinality bounded by the number of projective formulas of complexity less or equal than A. Finally, Lemma 3.53 guarantees that, picking a finite set of propositional variables including all the variables of A, this number is, up to logical equivalence, finite.

The contributions of this work are the following:

- We present a detailed proof, using Ghilardi's techniques, of the Unification Theorem for **CPL**, which in Ghilardi's article appears as a remark. This theorem was first proved by Löwenheim in [11] using different methods.
- In the proof of the Unification Theorem for **IPL** we prefer a purely semantic approach, reasoning in terms of the operator $(-)^{\sigma}$ and its properties instead of using syntactical deduction. Even though there are no new results, this semantical approach gives a more intuitive, and sometimes even graphical, idea of the contents of some lemmas. In particular, the proofs of lemmas 3.35, 3.37 and 3.41 (and its respective Corollaries 3.36, 3.38 and 3.40) are different from the proofs of the same facts displayed in Ghilardi's article.
- We prove some lemmas that are omitted in Ghilardi's article: lemmas 3.16, 3.53 and 3.62. To prove them, we also prove the auxiliary lemmas 3.52, 3.61 and 3.15. Lemmas 3.53 and 3.62 are crucial in the proof of the Unification Theorem. Lemma 3.53 guarantees the finiteness of the basis of unifiers appearing in the proof, and Lemma 3.62 is an important step in the proof of lemma 3.76, which is the hardest result before the proof of the main theorem.
- The main contribution of this work, we think, is the use of a different definition of the extension property, and the subsequent adaptation of all proofs of Ghilardi's proof to this change. Ghilardi's definition of the extension property is the following: a class of finite rooted Kripke models $\mathcal{K}$ has the extension property if for every model $\boldsymbol{M}$ whose proper generated submodels are all in $\mathcal{K}$, the class contains a variant* of $\boldsymbol{M}$. In this work, we use the definition proposed by Iemhoff in [8] and [9], where she uses the extension property to prove that Visser's rules form a basis for the admissible rules in **IPL** and gives some semantical characterizations of admissible rules for **IPL** and intermediate logics. According to Iemhoff's definition, a class of finite rooted Kripke models $\mathcal{K}$ has the extension property if it contains a variant of the disjoint sum of any finite set of models of $\mathcal{K}$. While the two properties are not equivalent †, the proof that Visser's rules form a basis for the admissible rules in **IPL** uses various facts about the extension property from Ghilardi's article. With this document we fill this small gap proving Ghilardi's results about the extension property using Iemhoff's definition. The results that were adapted are Theorem 3.49 and Lemmas 3.72 and 3.76.

*For rooted Kripke models, $\boldsymbol{M}'$ is a variant of $\boldsymbol{M}$ if and only if $\boldsymbol{M}'$ and $\boldsymbol{M}$ have the same frames and their valuations differ only at the root.

†Iemhoff's definition implies Ghilardi's under the assumption that the class of models is stable, and Ghilardi's definition implies Iemhoff's under the assumption that the class is closed under homomorphic images.

CHAPTER 2

Preliminaries

In this chapter we review some basic definitions and results of classical and intuitionistic propositional logic and we fix the notation that we will use in next chapters. We state many theorems without proof. The interested reader could find all the missing proofs in [3].

We start with the basic syntax. We will work with a fixed countable set of propositional variables Var. The basic symbols of our language will be the propositional variables in Var, the bottom symbol $\bot$, the top symbol $\top$, the binary connectives $\wedge, \vee, \rightarrow$ and parenthesis (,). We use lowercase letters $x, y, z, \ldots$ for propositional variables and capital letters $X, Y, Z, \ldots$ for sets of propositional variables.

Definition 2.1. Given a set of propositional variables $X \subseteq Var$, the set $F(X)$ of formulas in X is defined recursively as follows:

- $\top \in F(X)$;
- $\bot \in F(X)$;
- $x \in F(X)$, for all $x \in X$;
- if $A, B \in F(X)$, then $(A * B) \in F(X)$, where $* \in \{\wedge, \vee, \rightarrow\}$;
- Nothing else belongs to $F(X)$.

If there is no ambiguity, we will omit the use of some parenthesis. We use capital letters $A, B, C, A', B' \ldots$ for formulas. $\neg A$ is a shorthand for $A \rightarrow \bot$ and $A \leftrightarrow B$ is a shorthand for $(A \rightarrow B) \wedge (B \rightarrow A)$.

Definition 2.2. Let X and Y be sets of variables. A substitution σ (from X to Y) is a function from $F(X)$ to $F(Y)$ with the following properties:

- $\sigma(\top) = \top$
- $\sigma(\bot) = \bot$
- $\sigma(A * B) = \sigma(A) * \sigma(B)$ for all $A, B \in F(X)$ and each $* \in \{\wedge, \vee, \rightarrow\}$.

Remark 2.3. By the recursion theorem (see for instance [5]), every substitution $\sigma\colon F(X) \to F(Y)$ is completely characterized by its values on X, that is to say, if $\sigma_1\colon F(X) \to F(Y)$ and $\sigma_2 : F(X) \to F(Y)$ are substitutions such that $\sigma_1(x) = \sigma_2(x)$ for all $x \in X$, then $\sigma_1 = \sigma_2$. We will use this fact extensively throughout the next chapters.

The *composition* of two substitutions $\sigma : F(X) \to F(Y)$ and $\tau : F(Y) \to F(Z)$ is the function $\tau\sigma : F(X) \to F(Z)$ such that $\tau\sigma(x) = \tau(\sigma(x))$ for all $x \in X$. It is easy to check that $\tau\sigma$ is indeed a substitution.

2.1 Intuitionistic Logic

In this section we present the definition of a logic as a consequence relation and we define intuitionistic propositional logic and classical propositional logic within this context.

Definition 2.4. A *consequence relation* on a set A is a binary relation $\vdash \subseteq \mathcal{P}(A) \times A$ such that for every $X \cup Y \cup \{z\} \subseteq A$,

(i) $X \vdash x$ for all $x \in X$ and

(ii) if $X \vdash y$ for all $y \in Y$ and $Y \vdash z$, then $X \vdash z$.

To simplify the notation, we will use the following conventions:

- $\vdash y$ when $\emptyset \vdash y$
- $X, x \vdash y$ when $X \cup \{x\} \vdash y$.
- $x \dashv\vdash y$ when $x \vdash y$ and $x \vdash y$.

Given $X, Y \subseteq Var$ and a substitution $\sigma : F(X) \to F(Y)$, we define $\sigma[\Gamma] = \{\sigma(B) \in F(Y) : B \in \Gamma\}$.

Definition 2.5. A *logic* is a consequence relation $\vdash$ on the set of formulas $F(Var)$ that, moreover, is *substitution invariant*, that is to say, for every substitution $\sigma : F(X) \to F(Y)$ and every set of formulas $\Gamma \cup \{A\} \subseteq F(X)$,

$$\text{if } \Gamma \vdash A, \text{ then } \sigma[\Gamma] \vdash \sigma(A).$$

The expression $\Gamma \vdash A$ should be understood as "Γ proves A" or "A follows from Γ". The substitution invariance captures the idea that logical inferences are true only in virtue of their form. Given a logic $\vdash$, formulas such that $\vdash A$ are called *theorems* of $\vdash$.

There are many ways to determine a logic. One of them is by using a Hilbert calculus.

Definition 2.6. A *rule* is an expression of the form $\Delta \rhd B$, where $\Delta \cup \{B\} \subseteq F(Var)$. In this case, Δ is said to be the set of *premises* of the rule and B the *conclusion*. When $\Delta = \emptyset$, the rule $\Delta \rhd \boldsymbol{B}$ is called an *axiom*. A *Hilbert calculus* is a set of rules. A *finitary Hilbert calculus* is a Hilbert calculus all of whose rules have a finite set of premises.

For any finitary Hilbert calculus H, we define the notion of (finitary) proof in H. A *(finitary) proof of A from Γ in* H, with $\Gamma \subseteq F(Var)$ and $A \in F(Var)$, is a finite sequence of formulas $A_1, A_2, \ldots, A_n$ in $F(Var)$ such that $A_n = A$ and for every A_i in the sequence, either $A_i \in \Gamma$ or there is a rule $\Delta \rhd B$ in H, subsets $X, Y \subseteq Var$ and a substitution $\sigma : F(X) \to F(Y)$ such that $\sigma[\Delta] \subseteq \{A_j : j < i\}$ and $A_i = \sigma(B)$.

The logic $\vdash_{\mathsf{H}}$ induced by H is defined as follows: for every $\Gamma \cup \{A\} \subseteq F(Var)$

$$\Gamma \vdash_{\mathsf{H}} A \iff \text{there exists a (finitary) proof of } A \text{ from } \Gamma \text{ in } \mathsf{H}.$$

It is easy to check that $\vdash_{\mathsf{H}}$ is in deed a logic. Moreover, it is the least logic $\vdash$ such that $\Gamma \vdash A$, for every rule $\Gamma \rhd A$ in H.

Remark 2.7. In this work we will only deal with logics induced by finitary Hilbert calculus and finitary proofs. They are usually called *finitary logics*. Classical propositional logic and intuitionistic propositional logic are examples of them. It is easy to see that finitary logics are compact, that is, if $\Gamma \vdash A$, then $\Gamma_0 \vdash A$ for some finite $\Gamma_0 \subseteq \Gamma$. However, the notion of a logic induced by a Hilbert calculus can be extended to an arbitrary Hilbert calculus. See [13] for a generalization.

Example 2.8. Intuitionistic propositional logic **IPL** is the logic induced by the Hilbert calculus IPC with the following rules:

$$\begin{aligned}
&\emptyset \rhd \top \\
&\emptyset \rhd x \to (y \to x) \\
&\emptyset \rhd x \to (y \to (x \wedge y)) \\
&\emptyset \rhd (x \wedge y) \to x \\
&\emptyset \rhd (x \wedge y) \to y \\
&\emptyset \rhd x \to (x \vee y) \\
&\emptyset \rhd y \to (x \vee y) \\
&\emptyset \rhd (x \vee y) \to ((x \to z) \to ((y \to z) \to z)) \\
&\emptyset \rhd (x \to y) \to ((x \to (y \to z)) \to (x \to z)) \\
&\emptyset \rhd \bot \to x \\
x, x \to y &\rhd y.
\end{aligned}$$

Example 2.9. Classical propositional logic **CPL** is the logic induced by the Hilbert calculus $\mathsf{CPC} = \mathsf{IPC} \cup \{\emptyset \rhd x \vee \neg x\}$.

A well-known result of both intuitionistic propositional logic and classical propositional logic is the Deduction Theorem. For a proof, see [3].

Theorem 2.10 (Deduction Theorem). *Let $X \subseteq Var$. For all $\Gamma \subseteq F(X)$ and every $A, B \in F(X)$,*

$$\Gamma, A \vdash_{\mathsf{IPC}} B \text{ if and only if } \Gamma \vdash_{\mathsf{IPC}} A \to B$$

$$\Gamma, A \vdash_{\mathsf{CPC}} B \text{ if and only if } \Gamma \vdash_{\mathsf{CPC}} A \to B$$

2.2 Kripke Semantics

In this section we present the Kripke semantics for **IPL**. The key notion in the concept of Kripke model.

Definition 2.11. A *Kripke frame* $\mathcal{F}$ is a pair (F, R) where F is a non-empty set and R is an order relation on F (i.e., R is reflexive, antisymmetric and transitive).

Definition 2.12. Given $X \subseteq Var$, a *Kripke model* $\boldsymbol{M}$ *in* X is a triple (M, R, v) where (M, R) is a Kripke frame and $v : M \to \mathcal{P}(X)$ is a function with the following property: for all $m \in M$ and all $x \in X$, if $x \in v(m)$, then $x \in v(m')$ for each $m' \geqslant m$.

The set M in the previous definition is called the *domain* of $\boldsymbol{M}$ and the function v is called its *valuation*. The condition about v stated above is known as the *truth-preserving condition*. If $m \in M$, we will say that m is a *world* or a *point* of $\boldsymbol{M}$. With $\mathcal{M}_X$ we will denote the class of all Kripke models in X.

Given a Kripke model $\boldsymbol{M} \in \mathcal{M}_X$, we will use $\leqslant_M$ and v_M to denote, respectively, its relation and its valuation. If the model is clear form the context, we will simply use $\leqslant$ and v.

A Kripke model $\boldsymbol{M}$ is said to be *rooted* if there is an $x \in M$ such that $x \leqslant_M m$ for all $m \in \boldsymbol{M}$. The element x is called the root of $\boldsymbol{M}$. Whenever we have a rooted model $\boldsymbol{M}$, we will use r_M for its root. If the model is clear form the contexts, we will use r. Moreover, a Kripke model $\boldsymbol{M}$ is said to be *finite* if its domain is finite.

The fundamental semantic notion is the notion of truth. Given $\boldsymbol{M} \in \mathcal{M}_X$, and $m \in M$, we define recursively when a formula $A \in F(X)$ *is true (or holds) at point m in the model* $\boldsymbol{M}$, in symbols $\boldsymbol{M}, m \models \boldsymbol{A}$, as follows:.

Definition 2.13. Let X be a set of variables, $\boldsymbol{M} = (M, , \leqslant, v) \in \mathcal{M}_X$, $m \in M$, and $A, B \in F(X)$, we have:

- $\boldsymbol{M}, m \models \top$;
- $\boldsymbol{M}, m \models \bot$ does not hold;
- $\boldsymbol{M}, m \models x$ iff $x \in v(m)$;
- $\boldsymbol{M}, m \models A \wedge B$ iff $\boldsymbol{M}, m \models A$ and $\boldsymbol{M}, m \models B$;
- $\boldsymbol{M}, m \models A \vee B$ iff $\boldsymbol{M}, m \models A$ or $\boldsymbol{M}, m \models B$;
- $\boldsymbol{M}, m \models A \to B$ iff for all $n \geqslant m$, if $\boldsymbol{M}, n \models A$, then $\boldsymbol{M}, n \models B$.

We will write $\boldsymbol{M}, m \not\models \boldsymbol{A}$ when $\boldsymbol{M}, m \models A$ does not hold. A formula $A \in F(X)$ is said to be *satisfiable* if there is a model $\boldsymbol{M} \in \mathcal{M}_X$ and a world $m \in M$ such that $\boldsymbol{M}, m \models \boldsymbol{A}$. If $\boldsymbol{M}, m \models A$ for all $m \in M$, $\boldsymbol{M}$ is said to be a model of A (or A is said to be true in the model $\boldsymbol{M}$), in symbols $\boldsymbol{M} \models A$. If $\Gamma \subseteq F(X)$, $\boldsymbol{M} \models \Gamma$ stands for $\boldsymbol{M} \models A$ for all $A \in \Gamma$. Finally, we will use $Mod_X(\Gamma)$ to denote the class of models of Γ in $\mathcal{M}_X$, that is, $Mod_X(\Gamma) = \{\boldsymbol{M} \in \mathcal{M}_X : \boldsymbol{M} \models \Gamma\}$. $Mod_X(A)$ is a shorthand of $Mod_X(\{A\})$.

The following result states that the truth-preserving condition can be extended to all formulas. It can be easily proven by induction on the complexity of the formula.

Proposition 2.14. *Let $X \subseteq Var$ and $\boldsymbol{M} \in \mathcal{M}_X$. For all $A \in F(X)$ and for every $m \in M$, if $\boldsymbol{M}, m \models A$ then $\boldsymbol{M}, m' \models A$ for each $m' \geqslant m$.*

Remark 2.15. Due to the previous result, observe that for rooted models, being true in the whole model is equivalent to being true in its root. That is to say, if $\boldsymbol{M} \in \mathcal{M}_X$ is a rooted model with root r, $\boldsymbol{M} \models A$ if and only if $\boldsymbol{M}, r \models A$. We will use this simple fact often in the following chapters.

Definition 2.16. Let $A, B \in F(X)$ and $\Gamma \subseteq F(X)$. We say that Γ *implies* A (or A is a consequence of Γ), and we write $\Gamma \models A$, if and only if for all $\boldsymbol{M} \in \mathcal{M}_X$, if $\boldsymbol{M} \models \Gamma$, then $\boldsymbol{M} \models A$. Moreover, A is said to be a *valid formula*, written as $\models A$, if and only if $\boldsymbol{M} \models A$ for all $\boldsymbol{M} \in \mathcal{M}_X$ (equivalently, $\varnothing \models A$). For simplicity, we write $A \models B$ instead of $\{A\} \models B$. Finally, A and B are said to be *equivalent*, in symbols $A \equiv B$, if and only if $A \models B$ and $B \models A$ (equivalently, $\models A \leftrightarrow B$).

Remark 2.17. We have an easy characterization of logical implication in terms of classes of models. It is clear that for all $\Gamma \subseteq F(X)$ and each $A \in F(X)$, $\Gamma \models A$ if and only if $Mod_X(\Gamma) \subseteq Mod_X(A)$. In particular, for every $A, B \in F(X)$, $A \models B$ if and only if $Mod_X(A) \subseteq Mod_X(B)$, and $A \equiv B$ if and only in $Mod_X(A) = Mod_X(B)$.

2.3 Semantics for CPL

The usual semantics for **CPL** is given by classical valuations. Given $X \subseteq Var$, a *classical valuation in* X is a function $a : F(X) \to \{0, 1\}$ satisfying the following rules for constants and connectives:

- $a(\top) = 1$;
- $a(\bot) = 0$;
- $a(A \wedge B) = 1$ iff $a(A) = 1$ and $a(B) = 1$ for all $A, B \in F(X)$;
- $a(A \vee B) = 1$ iff $a(A) = 1$ or $a(B) = 1$ for all $A, B \in F(X)$;
- $a(A \to B) = 1$ iff $a(A) = 0$ or $a(B) = 1$ for all $A, B \in F(X)$.

By the recursion theorem, classical valuations in X are completely characterized by its values on the elements of X and the rules above. Given $X \subseteq Var$ and $A \in F(X)$, we say that A is *satisfiable* if and only if there is a classical valuation $a : F(X) \to \{0, 1\}$ such that $a(A) = 1$. In this case, a is said to be a model of A. The notions of logical consequence, valid formulas and equivalent formulas are analogous to the same notions in Kripke semantics, using classical valuations instead of Kripke models.

2.4 Submodels and generated submodels

Definition 2.18. Let $X \subseteq Var$ and $\boldsymbol{M}$ and $\boldsymbol{N}$ be in $\mathcal{M}_X$. The model $\boldsymbol{N}$ is a *submodel* of $\boldsymbol{M}$ if $N \subseteq M$ and $\leqslant_N$ and v_N are, respectively, the restrictions of $\leqslant_M$ and v_M to N.

Definition 2.19. Let $X \subseteq Var$ and $\boldsymbol{M} = (M, \leqslant_M, v_M)$ be a Kripke model in $\mathcal{M}_X$ and $p \in M$. *The submodel of* $\boldsymbol{M}$ *generated by* p is the model $\boldsymbol{M}_p = (M_p, \leqslant_{M_p}, (v_M)_p)$, where M_p is the upset of p, that is, $M_p = \{x \in M : p \leqslant x\}$, and $\leqslant_{M_p}$ and $(v_M)_p$ are, respectively, the restrictions of $\leqslant$ and v to M_p.

Remark 2.20. If $\boldsymbol{M} \in \mathcal{M}_X$, $p \in M$ and $q \in M_p$, it is immediate from the definition that $(\boldsymbol{M}_p)_q = \boldsymbol{M}_q$.

It is a routinary task to check that $\boldsymbol{M}_p$ is in fact a Kripke model in X. An important fact about generated submodels is that their worlds satisfy exactly the same formulas that they satisfy in the original model, as the following proposition states. It can be easily proved by induction on the complexity of the formula.

Proposition 2.21. *Let $X \subseteq Var$, $\boldsymbol{M} \in \mathcal{M}_X$ and $p \in M$. For each $q \in M_p$ and $A \in F(X)$, $\boldsymbol{M}, q \models A$ if and only if $\boldsymbol{M}_p, q \models A$.*

Remark 2.22. Obviously, given $\boldsymbol{M} \in \mathcal{M}_X$ and $p \in M$, the generated submodel $\boldsymbol{M}_p$ is a rooted model with root p. Thus, in view of Proposition 2.21 and Remark 2.15, we have $\boldsymbol{M}_p \models A$ if and only if $\boldsymbol{M}_p, p \models A$ if and only if $\boldsymbol{M}, p \models A$. We will make extensive use of this simple fact in many proofs.

Remark 2.23. Moreover, in view of the truth preserving condition for formulas and Proposition 2.21, we have that a formula is satisfiable if and only if there is a model $\boldsymbol{M} \in \mathcal{M}_X$ such that $\boldsymbol{M} \models A$.

2.5 Isomorphic models

Definition 2.24. Let $X \subseteq Var$. Two models $\boldsymbol{M} = (M, \leqslant_M, v_M)$ and $\boldsymbol{N} = (N, \leqslant_N, v_N)$ in $\mathcal{M}_X$ are said to be *isomorphic*, denoted as $\boldsymbol{M} \cong \boldsymbol{N}$, if there is a function $f : M \to N$ such that:

- f is a bijection;
- For all $p, q \in M$, $p \leqslant_M q$ iff $f(p) \leqslant_N f(q)$;
- For all $p \in M$, $v_M(p) = v_N(f(p))$.

Isomorphic models are structurally the same. It is not a surprise that they satisfy the same formulas. This fact can by proved by induction on the complexity of the formula.

Proposition 2.25. *Let $X \subseteq Vat$. If $\boldsymbol{M}, \boldsymbol{N} \in \mathcal{M}_X$ and $p \in M$, then for all formula $A \in F(X)$*

$$\boldsymbol{M}, p \models A \quad \Longleftrightarrow \quad \boldsymbol{N}, f(p) \models A$$

Remark 2.26. In view of the above result, we have that for all $X \subseteq Var$ and every $A \in F(X)$, the class of models $Mod_X(A)$ is closed under isomorphisms, in the sense that for all $\boldsymbol{M} \in Mod_X(A)$, if $\boldsymbol{N} \cong \boldsymbol{M}$, then $\boldsymbol{N} \in Mod_X(A)$.

2.6 Completeness

The most important results of Kripke semantics for **IPL** are the Completeness Theorem and the Finite Model Property theorem. Proofs of these results can be found in [3].

Theorem 2.27 (IPC - Completeness 1). *Let $X \subseteq Var$. For all $\Gamma \subseteq F(X)$ and every $A \in F(X)$, we have*

$$\Gamma \vdash_{\mathsf{IPC}} A \Longleftrightarrow \Gamma \models A$$

Theorem 2.28 (Finite model property). *Let $X \subseteq Var$. For all finite $\Gamma \subseteq F(X)$ and every $A \in F(X)$, if $\Gamma \nvdash_{\mathsf{IPC}} A$, there exists a finite model $\boldsymbol{M} \in \mathcal{M}_X$ such that $\boldsymbol{M} \models \Gamma$ and $\boldsymbol{M} \not\models A$.*

The Completeness theorem states that for *every* $\Gamma \subseteq F(X)$ and all $A \in F(X)$, Γ proves A in the logic **IPL** if and only if for *every* model $\boldsymbol{M} \in \mathcal{M}_X$, $\boldsymbol{M} \models A$ whenever $\boldsymbol{M} \models \Gamma$. If we combine the Completeness theorem with the Finite model property, we obtain a version of the Completeness theorem restricted to finite sets of formulas and finite models.

Theorem 2.29 (IPC - Completeness 2). *For all finite $\Gamma \subseteq F(X)$ and every $A \in F(X)$, we have*

$$\Gamma \vdash_{\mathsf{IPC}} A \iff \Gamma \models_{\mathit{fin}} A$$

where $\Gamma \models_{\mathit{fin}} A$ means that for every finite model $\boldsymbol{M} \in \mathcal{M}_X$, if $\boldsymbol{M} \models \Gamma$, then $\boldsymbol{M} \models A$.

Using generated submodels and Proposition 2.21 we can refine even more the previous result.

Theorem 2.30 (IPC - Completeness 3). *For all finite $\Gamma \subseteq F(X)$ and every $A \in F(X)$, we have*

$$\Gamma \vdash_{\mathsf{IPC}} A \iff \Gamma \models_{\mathit{fin\ rooted}} A$$

where $\Gamma \models_{\mathit{fin\ rooted}} A$ means that for every finite rooted model $\boldsymbol{M} \in \mathcal{M}_X$, if $\boldsymbol{M} \models \Gamma$, then $\boldsymbol{M} \models A$.

Thus, $\Gamma \models A$, $\Gamma \models_{\mathit{fin}} A$ and $\Gamma \models_{\mathit{fin\ rooted}} A$ are equivalent when Γ is finite. In view of this equivalence, when we deal with **IPL** in the following chapters we will work with finite rooted models, and whenever we use $\models$, the reader must understand $\models_{\mathit{fin\ rooted}}$.

On the other hand, we state the Completeness theorem for classical propositional logic. **CPL** turns out to be complete with respect to the semantics based on classical valuations.

Theorem 2.31 (CPC - Completeness). *Let $X \subseteq Var$. For all $\Gamma \subseteq F(X)$ and every $A \in F(X)$, we have*

$$\Gamma \vdash_{\mathsf{CPC}} A \iff \Gamma \models A$$

CHAPTER 3

Unification in Intuitionistic logic

The goal of this chapter is to prove the Unification theorem for intuitionistic propositional logic. In the first section, we give the basic definitions to state the theorem. In the second section, we define the operator $(-)^\sigma$, which will be one of the fundamental tool of this chapter. As we pointed out in the introduction, the operator $(-)^\sigma$ will allow us to reason semantically about substitutions. In the third section, we prove the unification theorem for classical propositional logic. And finally, in the fourth section, we prove the unification theorem for intuitionistic propositional logic.

3.1 Unifiable formulas

Definition 3.1. Let $\vdash$ be a logic, $X, Y \subseteq Var$ and $A \in F(X)$,

- A substitution $\sigma\colon F(X) \to F(Y)$ is a *unifier of A in $\vdash$* if and only if $\vdash \sigma(A)$.
- A is *unifiable in $\vdash$* if and only if it has a unifier in $\vdash$.

Remark 3.2. Due to the Completeness theorem, in **IPL** the first condition is equivalent to the requirement $\models \sigma(A)$, i.e., every model $\boldsymbol{M} \in \mathcal{M}_Y$ is a model of $\sigma(A)$. In **CPL**, the first condition is equivalent to say that $a(\sigma(A)) = 1$ for all classical valuations $a\colon F(X) \to \{0, 1\}$.

Definition 3.3. Let $\vdash$ be a logic, $X, Y, Z \subseteq Var$, $\sigma_1\colon F(X) \to F(Y)$ and $\sigma_2\colon F(X) \to F(Z)$ be substitutions. σ_1 is said to be *less general (in $\vdash$)* than σ_2, in symbols $\sigma_1 \preceq_\vdash \sigma_2$, if and only if there is a substitution $\tau\colon F(Z) \to F(Y)$ such that

$$\tau(\sigma_2(x)) \dashv\vdash \sigma_1(x) \quad \text{for all} \quad x \in X.$$

For simplicity, whenever the logic is clear from the context, we will omit the explicit reference to it. Thus, we will write *unifier of A* instead of *unifier of A in $\vdash$*, *unifiable* instead of *unifiable in $\vdash$* and *less general* instead of *less general (in $\vdash$)*.

Remark 3.4. The relation $\preceq_\vdash$ defines a preorder (that is, $\preceq_\vdash$ is reflexive and transitive) on the set of all substitutions with domain $F(X)$. Transitivity follows from the substitution invariance in the definition of $\vdash$.

Remark 3.5. In **IPL** and **CPL**, where $\vdash$ and $\models$ stands for the corresponding logic and semantics, we have the equivalences

$$\begin{aligned} \tau(\sigma_2(x)) \dashv\vdash \sigma_1(x) &\Longleftrightarrow \tau(\sigma_2(x)) \equiv \sigma_1(x) \\ &\Longleftrightarrow \models \tau(\sigma_2(x)) \leftrightarrow \sigma_1(x) \end{aligned}$$

for all $x \in X$. Thus, in **IPL** and **CPL**, we can use any of the previous statements to show that one substitution is less general than other.

Remark 3.6. In **IPL** and **CPL**, the condition '$\models \tau(\sigma_2(x)) \leftrightarrow \sigma_1(x)$ for all $x \in X$' implies (and therefore it is equivalent to) the condition '$\models \tau(\sigma_2(A)) \leftrightarrow \sigma_1(A)$ for all $A \in F(X)$'. This can be proved by induction. The base cases are immediate. For the inductive cases, assume $\models \tau(\sigma_2(B)) \leftrightarrow \sigma_1(B)$ and $\models \tau(\sigma_2(C)) \leftrightarrow \sigma_1(C)$. We will only prove the result for $B \to C$. We need $\models \tau(\sigma_2(B \to C)) \leftrightarrow \sigma_1(B \to C)$. But this is equivalent to $\models (\tau(\sigma_2(B)) \to \tau(\sigma_2(C))) \leftrightarrow (\sigma_1(B) \to \sigma_1(C))$, which is clearly true from the induction hypothesis.

Example 3.7. Let $X \subseteq Var$ and $z \in X$. In **CPL** and **IPL**, consider the substitution $\sigma_1 \colon F(X) \to F(\emptyset)$ such that

$$\sigma_1(x) = \begin{cases} \top & \text{if x = z} \\ \bot & \text{otherwise} \end{cases}$$

and the substitution $\sigma_2 \colon F(X) \to F(X)$ such that

$$\sigma_2(x) = \begin{cases} z \to z & \text{if x = z} \\ z \wedge x & \text{otherwise} \end{cases}$$

Clearly, both substitutions are unifiers of the formula z. Moreover, σ_1 is less general than σ_2. To show this, we need to find a substitution $\tau \colon F(X) \to F(\emptyset)$ such that $\models \tau(\sigma_2(x)) \leftrightarrow \sigma_1(x)$ for all $x \in X$. For $x = z$, this is the same as $\models (\tau(z) \to \tau(z)) \leftrightarrow \top$, and for $x \neq z$ this is the same as $\models \tau(z) \wedge \tau(x) \leftrightarrow \bot$. As the reader can check, any substitution $\tau \colon F(X) \to F(\emptyset)$ satisfying $\tau(z) \equiv \bot$ or $\tau(x) \equiv \bot$ for all $x \neq z$ satisfies what we need.

Definition 3.8. Let $\vdash$ be a logic and $A \in F(X)$ be a unifiable formula.

- Given σ_1 and σ_2 unifiers of A, σ_1 is said to be *less general* than σ_2 if σ_1 is less general than σ_2 as substitutions.
- A set S of unifiers of A is said to be a *complete set of unifiers of* A if every unifier of A is less general than some member of S.
- A complete set of unifiers of A is said to be a *basis of unifiers of* A if and only if its members are pairwise incomparable with respect to the preorder $\preceq_\vdash$.
- σ is said to be a *most general unifier (mgu) for* A if $\{\sigma\}$ is a complete set of unifiers of A.

Remark 3.9. Basis of unifiers of a formula A are minimal sets among the complete sets of unifiers of A. It is easy to prove that two basis of unifiers of a formula A have the same cardinality.

Remark 3.10. In the definitions above, we use the set of formulas $F(X)$ defined in the preliminaries. However, these definitions can be stated for other logics using their corresponding set of formulas. In particular, the definitions of this sections can be stated for modals logics and Łukasiewicz logics.

Every unifiable formula in a logic $\vdash$ has a complete set of unifiers (the set of all its unifiers). However, not every unifiable formula admits a basis of unifiers. In [10], Jeřábek shows that the $p \to \Box p$ is unifiable in **K** but it lacks a basis of unifiers. Moreover, if a unifiable formula admits a basis of unifiers, the cardinality of this basis (and therefore the cardinality of all basis) may be 1, finite (different from 1) or infinite. We use this to define the unification type of a logic.

Definition 3.11. The unification type of a logic is:

- *Unitary (or 1)* iff every unifiable formula has a basis of unifiers of size 1 (equivalently, every unifiable formula admits a mgu);
- *Finitary (or ω)* iff every unifiable formula has a finite basis of unifiers and there is at least one unifiable formula that does not have a basis of unifiers of size 1;
- *Infinitary (or ∞)* iff every unifiable formula has a finite or an infinite basis of unifiers and there is at least one unifiable formula that does not have a finite basis of unifiers;
- *Nullary (or 0)* iff there is a unifiable formula that does not have a basis of unifiers.

The Unification Theorem for intuitionistic propositional logic states that unification type of **IPL** is finitary and the Unification Theorem for **CPL** states that the unification type of **CPL** is unitary. In [6], Ghilardi also shows that any intermediate logic satisfying the De Morgan axiom $\neg(A \wedge B) \to (\neg A \vee \neg B)$ has unitary unification type. The modal logic **S5** and Łukasiewicz logics $Ł_n$ are other examples of logics with unitary unification type. Modal logics **S4** and **K4** are examples of logics with finitary unification. **K** and Łukasiewicz logic $Ł_\omega$ are examples of logics with nullary unification type. In [4], W. Dzik, S.Kost and P.Wojtylak point out that no example of a modal nor intermediate logic with an infinitary unification type is known. For a comprehensive explanation of unification in modal logics see [2] and [1], and for unification in Łukasiewicz logics see [12].

3.2 The $(-)^\sigma$ operator

Every substitution induces a transformation between classes of Kripke models and between classical valuations. This transformation will be extremely useful to obtain a characterization of unifiable formulas in purely semantical terms and will be a fundamental tool in the proof of the unification theorems for **IPL** and **CPL**.

We start with **IPL** and the transformation between Kripke models.

Definition 3.12. Let $X, Y \subseteq Var$ be sets of propositional variables, and let $\sigma\colon F(X) \to F(Y)$ be a substitution. The operator $(-)^\sigma : \mathcal{M}_Y \to \mathcal{M}_X$ sends a Kripke model $\boldsymbol{M} = (M, \leqslant_M, v_M)$ to the Kripke model $\boldsymbol{M}^\sigma = (M, \leqslant_M, v^\sigma_M)$, where $v^\sigma_M : A \to \mathcal{P}(X)$ is such that $x \in v^\sigma_M(m)$ iff $\boldsymbol{M}, m \models \sigma(x)$ for all $m \in M$.

It is easy to check that $\boldsymbol{M}^\sigma$ is a Kripke model in Y. We only need to verify the truth-preserving property for v^σ_M, but this is inherited from the truth-preserving property of v. Sometimes we will use M^σ to denote the domain of $\boldsymbol{M}^\sigma$. Even though $M^\sigma = M$, in some cases it will be clearer to say explicitly whether we think of M as the domain of $\boldsymbol{M}$ or as the domain of $\boldsymbol{M}^\sigma$.

The $(-)^\sigma$ operator has a good behavior with respect to the operation of submodel generation in the sense that both operations commute.

Proposition 3.13. *Let $X, Y \subseteq Var$, $\sigma\colon F(X) \to F(Y)$ be a substitution, $\boldsymbol{M} \in \mathcal{M}_X$ and $p \in M$. Then, $(\boldsymbol{M}_p)^\sigma = (\boldsymbol{M}^\sigma)_p$.*

Proof. The frames of $(\boldsymbol{M}_p)^\sigma$ and $(\boldsymbol{M}^\sigma)_p$ are the same. It only remains to be proved that $(v^\sigma_M)_p = ((v_M)_p)^\sigma$. But for all $q \in M_p$ and all $x \in X$, $x \in (v^\sigma_M)_p(q)$ iff $x \in (v^\sigma_M)(q)$ iff $\boldsymbol{M}, q \models \sigma(x)$ iff $\boldsymbol{M}_p, q \models \sigma(x)$ iff $x \in ((v_M)_p)^\sigma$, using the definition of $(-)^\sigma$ and Proposition 2.21. ⊠

The former proposition ensures that there is no ambiguity if we use $\boldsymbol{M}^\sigma_p$ to denote $(\boldsymbol{M}_p)^\sigma$ and $(\boldsymbol{M}^\sigma)_p$.

The following lemma describes the basic properties of the operator $(-)^\sigma$. In spite of its simplicity, this lemma will be the most used result of the chapter.

Lemma 3.14. *Let $X, Y, Z \subseteq Var$, $A \in F(X)$ and $\sigma\colon F(X) \to F(Y)$ be a substitution. Then*

(i) *for every $\boldsymbol{M} \in \mathcal{M}_Y$ and every $m \in M$, $\boldsymbol{M}^\sigma, m \models A$ iff $\boldsymbol{M}, m \models \sigma(A)$;*

(ii) *$\models \sigma(A)$ iff $\boldsymbol{M}^\sigma \models A$ for all $\boldsymbol{M} \in \mathcal{M}_Y$;*

(iii) *for every substitution $\tau\colon F(Y) \to F(Z)$ and for every model $\boldsymbol{N} \in \mathcal{M}_Z$, $\boldsymbol{N}^{\tau\sigma} = (\boldsymbol{N}^\tau)^\sigma$.*

Proof.

(i) Let $\boldsymbol{M} = (M, \leqslant, v_M)$ be in $\mathcal{M}_Y$ and $m \in M$. This part is proved by induction on A. If A is a propositional variable, the result is immediate from the definition of $(-)^\sigma$, and if A is $\bot$ or $\top$, the result is trivial. The inductive cases for $\wedge$ and $\vee$ are easy. Assume now that the result holds for formulas B and C. We will prove that $\boldsymbol{M}^\sigma, m \models B \to C$ iff $\boldsymbol{M}, m \models \sigma(B \to C)$. For the left-to-right direction assume $\boldsymbol{M}^\sigma, m \models B \to C$. We want $\boldsymbol{M}, m \models \sigma(B \to C)$. But $\sigma(B \to C) = \sigma(B) \to \sigma(C)$. So let $n \geqslant_M m$ be such that $\boldsymbol{M}, n \models \sigma(B)$. By the induction hypothesis, $\boldsymbol{M}^\sigma, n \models B$. Since $\boldsymbol{M}^\sigma, m \models B \to C$, $\boldsymbol{M}^\sigma, n \models C$. By the induction hypothesis again, we have $\boldsymbol{M}, n \models \sigma(C)$. Therefore, $\boldsymbol{M}, m \models \sigma(B) \to \sigma(C)$, as wanted. The other direction is proven similarly.

(ii) This is an immediate consequence of (i).

(iii) Let $\tau\colon F(Y) \to F(Z)$ be a substitution and $\boldsymbol{N} \in \mathcal{M}_Z$. The domains and the relations of $\boldsymbol{N}^{\tau\sigma}$ and $(\boldsymbol{N}^\tau)^\sigma$ are the same. To prove the equality we only need to check that $v_{N^{\tau\sigma}} = v_{(N^\tau)^\sigma}$. But for all $x \in X$ and all $n \in N$, we have $x \in v_{N^{\tau\sigma}}(n)$ iff $\boldsymbol{N}, n \models \tau\sigma(x)$ iff $\boldsymbol{N}, n \models \tau(\sigma(x))$ iff $\boldsymbol{N}^\tau, n \models \sigma(x)$ iff $(\boldsymbol{N}^\tau)^\sigma, n \models x$ iff $x \in v_{(N^\tau)^\sigma}(n)$. Hence, $v_{N^{\tau\sigma}} = v_{(N^\tau)^\sigma}$.

⊠

Before proving the characterization result about unifiable formulas, we need the following lemmas about models where the valuation is constant. We say that a model $\boldsymbol{M} \in \mathcal{M}_X$ has constant valuation if $v_M(m) = v_M(m')$ for all $m, m' \in M$.

Lemma 3.15. *Let $X \subseteq Var$ and $\boldsymbol{M} \in \mathcal{M}_X$ be a model with constant valuation. Then for all formula $A \in F(X)$ and all $m \in M$*

$$\boldsymbol{M} \models A \iff \boldsymbol{M}_m \models A$$

Proof. We proceed by induction on A. The case for propositional variables is immediate from the hypothesis that the valuation is constant, and the cases for $\bot$ and $\top$ are trivial. For the inductive cases, assume that for all $m \in M$, $\boldsymbol{M} \models B \iff \boldsymbol{M}_m \models B$ and $\boldsymbol{M} \models C \iff \boldsymbol{M}_m \models C$. We will only prove the result for $B \to C$, that is, we will prove that for all $m \in M$, $\boldsymbol{M} \models B \to C \iff \boldsymbol{M}_m \models B \to C$. Let $m \in M$. We want $\boldsymbol{M} \models C \to D$ iff $\boldsymbol{M}_m \models C \to D$. The left-to-right direction is immediate from the truth preserving result for formulas stated in Lemma 2.14. For the other direction, assume $\boldsymbol{M}_m \models C \to D$, and suppose towards a contradiction that $\boldsymbol{M} \not\models C \to D$. Then there is some $m' \in M$ such that $\boldsymbol{M}, m' \models C$ and $\boldsymbol{M}, m' \not\models D$, that is, $\boldsymbol{M}_{m'} \models C$ and $\boldsymbol{M}_{m'} \not\models D$. Hence, by the induction hypothesis, $\boldsymbol{M} \models C$ and $\boldsymbol{M} \not\models D$, and again by the induction hypothesis we obtain $\boldsymbol{M}_m \models C$ and $\boldsymbol{M}_m \not\models D$, contradicting $\boldsymbol{M}_m \models C \to D$. ⊠

Lemma 3.16. *Let $X \subseteq Var$ and let $\boldsymbol{M}, \boldsymbol{N} \in \mathcal{M}_X$. If for all $m \in M$ and all $n \in N$, $v_M(m) = v_N(n)$, then*

$$\boldsymbol{M} \models A \iff \boldsymbol{N} \models A \quad \textit{for all } A \in F(X)$$

.

Proof. We prove it by induction. The case when A is a propositional variable is an immediate application of the hypothesis about v_M and v_N to the roots, and the cases for $\top$ and $\bot$ are trivial. For the inductive step, we only prove the case when A is of the form $B \to C$ under the assumption that the result holds for B and C. We need $\boldsymbol{M} \models B \to C$ if and only if $\boldsymbol{N} \models B \to C$. We will prove only one direction, for the other direction is proven similarly. Assume $\boldsymbol{M} \models B \to C$, and suppose towards a contradiction that $\boldsymbol{N} \not\models B \to C$. It implies that there is an $n \in N$ such that $\boldsymbol{N}, n \models B$ and $\boldsymbol{N}, n \not\models C$. Furthermore, observe that the hypothesis about v_M and v_N in the statement implies that both $\boldsymbol{M}$ and $\boldsymbol{N}$ have constant valuation. Hence, by the previous lemma, $\boldsymbol{N} \models B$ and $\boldsymbol{N} \not\models C$, and by the induction hypothesis we get $\boldsymbol{M} \models B$ and $\boldsymbol{M} \not\models C$. It implies that there is an $m \in \boldsymbol{M}$ such that $\boldsymbol{M}, m \models B$ and $\boldsymbol{M}, m \not\models C$. Hence, $\boldsymbol{M} \not\models B \to C$, obtaining a contradiction.

⊠

As a consequence, we have the following characterization of unifiable formulas for **IPL**.

Proposition 3.17. *Let $X \subseteq Var$ and $A \in F(X)$. A is unifiable in* **IPL** *if and only if A is satisfiable.*

Proof. ($\Rightarrow$) Assume A is unifiable. Hence, there is a substitution $\sigma \colon F(X) \to F(Y)$, for some $Y \subseteq Var$, such that $\models \sigma(A)$. Let $\boldsymbol{M}$ be an arbitrary model in $\mathcal{M}_Y$ (take, for instance, the one-point model where no propositional variable of Y holds). As $\models \sigma(A)$, $\boldsymbol{M} \models \sigma(A)$. By (i) of Lemma 3.14, $\boldsymbol{M}^\sigma \models A$. Therefore, A is satisfiable.

($\Leftarrow$) Assume A is satisfiable, and let $\boldsymbol{M} \in \mathcal{M}_X$ be a model of A. Since M is finite, we can pick a maximal element w of $\boldsymbol{M}$. Define the substitution $\sigma\colon F(X) \to F(\emptyset)$ such that

$$\sigma(x) = \begin{cases} \top & \text{if } \boldsymbol{M}_w \models x \\ \bot & \text{otherwise} \end{cases}$$

We will prove that $\models \sigma(A)$. By (ii) of Lemma 3.14, it is enough to show that $\boldsymbol{N}^\sigma \models A$ for all $\boldsymbol{N} \in \mathcal{M}_\emptyset$. So let $\boldsymbol{N}$ be in $\mathcal{M}_\emptyset$. Notice that for all $n \in N^\sigma$ and all $x \in X$, $x \in v_N^\sigma(n) \Leftrightarrow \boldsymbol{N}, n \models \sigma(x) \Leftrightarrow \sigma(x) = \top \Leftrightarrow \boldsymbol{M}_w \models x$. Hence, $\boldsymbol{M}_w$ and $\boldsymbol{N}^\sigma$ satisfy the hypothesis of Lemma 3.16. As $\boldsymbol{M}_w \models A$, we conclude $\boldsymbol{N}^\sigma \models A$, as wished. ⊠

For **CPL** and classical valuations, the definitions and results are analogous to the definitions and results in **IPL**. Given a substitution $\sigma\colon F(X) \to F(Y)$ and a classical valuation $a\colon F(Y) \to \{0,1\}$, we define the classical propositional valuation $a^\sigma\colon F(X) \to \{0,1\}$ as the unique classical valuation such that $a^\sigma(x) = 1$ if and only if $a(\sigma(x)) = 1$ for all $x \in X$. We state, without proof, the result concerning the basic properties of a^σ and the semantical characteization of unifiable formulas as satisfiable formulas (with the semantics of **CPL**). The proofs are just adaptations of the proofs for Kripke models.

Lemma 3.18. *Let $A \in F(X)$ be a formula and $\sigma\colon F(X) \to F(Y)$ be a substitution. Then,*

(i) *for every classical valuation $a\colon F(Y) \to \{0,1\}$, $a^\sigma \models A$ iff $a \models \sigma(A)$.*

(ii) *$\models \sigma(A)$ iff $a^\sigma \models A$ for all classical valutaion $a\colon F(X) \to \{0,1\}$.*

(iii) *for every substitution $\tau\colon F(Y) \to F(Z)$ and avery classical valuation $b\colon F(Z) \to \{0,1\}$, $b^{\tau\sigma} = (b^\tau)^\sigma$.*

Proposition 3.19. *Let $X \subseteq Var$ and $A \in F(X)$. A is unifiable in **CPL** if and only if A is satisfiable.*

3.3 Projective formulas

Definition 3.20. Let $\vdash$ be a logic and $X \subseteq Var$. A formula $A \in F(X)$ is *projective* if and only if there is a substitution $\sigma\colon F(X) \to F(X)$ such that:

(1) $\vdash \sigma(A)$ (that is, σ is a unifier of A);

(2) $A \vdash x \leftrightarrow \sigma(x)$ for all $x \in X$.

In this case, we say that A *is projective with* σ.

Note that, by definition, every projective formula is unifiable. The definition above is meant for every logic. However, for **CPL** and **IPL** we can give equivalent definitions of projective formulas.

Remark 3.21. By completeness, in **CPL** and **IPL**, with their respective semantics, we can replace conditions (1) and (2) for

(1') $\models \sigma(A)$ (that is, σ is a unifier of A);

(2') $A \models x \leftrightarrow \sigma(x)$ for all $x \in X$.

Remark 3.22. In **CPL** and **IPL**, condition (2) is equivalent to the following (apparently stronger) condition:

(2″) $A \models B \leftrightarrow \sigma(B)$ for all $B \in F(X)$.

This is shown by proving that (2″) is equivalent to (2′). The implication from (2″) to (2′) is trivial. To show that (2′) implies (2″), we proceed by induction on A. For propositional variables, it is exactly the statement (2′), and for constants $\perp$ and $\top$, the result is trivial. Now assume that (2″) holds for formulas $C, D \in F(X)$. We want $A \models C \wedge D \leftrightarrow \sigma(C \wedge D)$. But $C \leftrightarrow \sigma(C), D \leftrightarrow \sigma(D) \models C \wedge D \leftrightarrow \sigma(C) \wedge \sigma(D)$. And by hypothesis, $A \models C \leftrightarrow \sigma(C)$ and $A \models D \leftrightarrow \sigma(D)$. Thus, $A \models C \wedge D \leftrightarrow \sigma(C) \wedge \sigma(D)$. As $\sigma(C) \wedge \sigma(D) = \sigma(C \wedge D)$, we are done. The inductive step for the other connectives is similar. Thus, we can replace (2′) by (2″) in the definition of projective formula.

Remark 3.23. In **IPL**, conditions (1) and (2) are equivalent, respectively, to

(1i) $\boldsymbol{M}^\sigma \models A$ for all $\boldsymbol{M} \in \mathcal{M}_X$ (equivalently, $(\mathcal{M}_X)^\sigma \subseteq Mod_X(A)$, where $(\mathcal{M}_X)^\sigma$ is the image of $\mathcal{M}_X$ under σ).

(2i) $\boldsymbol{M}^\sigma = \boldsymbol{M}$ for all $\boldsymbol{M} \in Mod_X(A)$.

We show it by proving that (1i) and (2i) are equivalent to (1′) and (2′). The equivalence (1′) if and only if (1i) is Lemma 3.14 (ii). The equivalence (2′) if and only if (2i) is proved as follows. For the left-to-right direction assume that σ satisfies (2′) and let $\boldsymbol{M} \in Mod_X(A)$, that is, $\boldsymbol{M} \models A$. We want $\boldsymbol{M} = \boldsymbol{M}^\sigma$, but since the frames of $\boldsymbol{M}$ and $\boldsymbol{M}^\sigma$ are the same, we only need to check that $v_M = v_M^\sigma$. But observe that for all $m \in \boldsymbol{M}$

$$\begin{aligned}
x \in v_M^\sigma(m) &\iff \boldsymbol{M}^\sigma, m \models x \\
&\iff \boldsymbol{M}, m \models \sigma(x) && \text{(by Lemma 3.14)} \\
&\iff \boldsymbol{M}, m \models x && \text{(by } \boldsymbol{M} \models A \text{ and (2′))} \\
&\iff x \in v_M(m)
\end{aligned}$$

Thus, $\boldsymbol{M}^\sigma = \boldsymbol{M}$, as we wanted. For the direction from (2i) to (2′), assume σ satisfies (2i). We want $A \models x \leftrightarrow \sigma(x)$ for all $x \in X$. So let $\boldsymbol{M} \in \mathcal{M}_X$ be a model of A. For all $m \in \boldsymbol{M}$ and all $x \in X$ we have

$$\begin{aligned}
\boldsymbol{M}, m \models x &\iff \boldsymbol{M}^\sigma, m \models x && \text{(for } \boldsymbol{M}^\sigma = \boldsymbol{M}\text{)} \\
&\iff \boldsymbol{M}, m \models \sigma(x) && \text{(by Lemma 3.14)}
\end{aligned}$$

Then $A \models x \leftrightarrow \sigma(x)$.

Notice that conditions (1i) and (2i) give us a characterization of projective formulas in terms of properties of the operator $(-)^\sigma$. Condition (1i) says that the operator $(-)^\sigma$ collapses the class $\mathcal{M}_X$ into the models of A, and condition (2i) says that the operator $(-)^\sigma$ leaves unmodified the models of A. Thus, the operator $(-)^\sigma$ projects the class of all models in X into the models of A.

Remark 3.24. In **CPL**, conditions (1) and (2) are equivalent, respectively, to

(1c) $a(A) = 1$ for all classical valuation $a \colon F(X \to \{1, 0\}$.

(2c) $a^\sigma = a$ for all classical valuation $a \colon F(X \to \{1, 0\}$ such that $a(A) = 1$.

The proof is similar to the case of **IPL**.

Remark 3.25. In **CPL** and **IPL**, if $A \in F(X)$ is projective with $\sigma\colon F(X) \to F(X)$, then σ is a most general unifier of A. To see this, simply observe that, since σ satisfies (2), for all substitution $\tau\colon F(X) \to F(Y)$ we have $\tau(A) \models \tau(x) \leftrightarrow \tau(\sigma(x))$ for each $x \in X$. Furthermore, if τ is a unifier of A, we have $\models \tau(A)$, and thus $\models \tau(x) \leftrightarrow \tau(\sigma(x))$. Therefore, $\tau \preceq \sigma$ for every τ unifier of A.

The following lemma holds both for **IPL** and **CPL**.

Lemma 3.26. *Let $A \in F(X)$. If $\sigma_1, \sigma_2\colon F(X) \to F(X)$ are substitutions satisfying (2'), then $\sigma_1\sigma_2$ also satisfies (2').*

Proof. By assumption $A \models x \leftrightarrow \sigma_1(x)$ and $A \models x \leftrightarrow \sigma_2(x)$ for all $x \in X$, and we need to prove $A \models x \leftrightarrow \sigma_1\sigma_2(x)$ for all $x \in X$. On the one hand, recall that, by Remark 3.22, (2'') also holds for σ_1, so $A \models A \leftrightarrow \sigma_1(A)$, and therefore $A \models \sigma_1(A)$. On the other hand, since $A \models x \leftrightarrow \sigma_2(x)$ for all $x \in X$, completeness and the substitution invariance, we have $\sigma_1(A) \models \sigma_1(x) \leftrightarrow \sigma_1(\sigma_2(x))$ for all $x \in X$. From this and $A \models \sigma_1(A)$, we get $A \models \sigma_1(x) \leftrightarrow \sigma_1(\sigma_2(x))$ for all $x \in X$. Since also $A \models x \leftrightarrow \sigma_1(x)$, we conclude $A \models x \leftrightarrow \sigma_1(\sigma_2(x))$ for all $x \in X$. ⊠

3.4 Boolean Unification

Before proving the unification theorem for intuitionistic propositional logic, we prove now, as an illustrative and important example, the corresponding theorem for classical propositional logic. To this end, given $X \subseteq Var$, a formula $A \in F(X)$ and a classical valuation $a\colon F(X) \to \{0,1\}$, we define $\theta_a^A(x)\colon F(X) \to F(X)$ as the substitution satisfying:

$$\theta_a^A(x) = \begin{cases} A \to x & \text{if } a(x) = 1 \\ A \wedge x & \text{otherwise} \end{cases} \tag{3.1}$$

In **CPL**, if a formula A is satisfiable with a, A turns out to be projective with $\theta_a^A(x)$. This nice property will simplify the problem of finding the unification type of **CPL**.

Proposition 3.27. . *If $A \in F(X)$ is a formula satisfiable with the classical valuation a, then A is projective with $\theta_a^A(x)$.*

Proof. We will show that A is projective with θ_a^A by showing that θ_a^A satisfies the conditions (1c) and (2c) of Remark 3.24. We start by proving that θ_a^A satisfies (2c). Let b be a classical valuation such that $b(A) = 1$. We want to check that $b^{\theta_a^A} = b$, for which it is enough to show that $b^{\theta_a^A}(x) = b(x)$ for all $x \in X$. We proceed by cases. If $a(x) = 1$, then we have

$$\begin{aligned}
b^{\theta_a^A}(x) = 1 &\iff b^{\theta_a^A} \models x \\
&\iff b \models \theta_a^A(x) && \text{(by Lemma 3.18)} \\
&\iff b \models A \to x && \text{(by the assumption } a(x) = 1) \\
&\iff b(A \to x) = 1 \\
&\iff b(x) = 1
\end{aligned}$$

If $a(x) \neq 1$, then we have

$$\begin{aligned} b^{\theta_a^A}(x) = 1 &\iff b^{\theta_a^A} \models x \\ &\iff b \models \theta_a^A(x) && \text{(by Lemma 3.18)} \\ &\iff b \models A \wedge x && \text{(by the assumption } a(x) \neq 1) \\ &\iff b(A \wedge x) = 1 \\ &\iff b(x) = 1 \end{aligned}$$

Hence $b^{\theta_a^A}(x) = b(x)$ for all $x \in X$ and thus $b^{\theta_a^A} = b$, as desired.

Now we will prove that θ_a^A satisfies (1c). So let $b\colon F(X) \to \{0,1\}$ be a classical valuation. If $b(A) = 1$, (2c) implies $b^{\theta_a^A} = b$ and we are done. It only remains to prove that if $b(A) \neq 1$, $b^{\theta_a^A}(A) = 1$ holds. Since a is already a model of A, we will do it by proving that $b^{\theta_a^A} = a$. As before, it is enough to show that $a(x) = b^{\theta_a^A}(x)$ for all $x \in X$. So let $x \in X$. We have two cases: $a(x) = 1$ or $a(x) \neq 1$. If $a(x) = 1$, then, following the same steps as for the (2c) case, we have $b^{\theta_a^A}(x) = 1 \iff b(A \to x) = 1$. Since we are assuming $b(A) \neq 1$, trivially $b(A \to x) = 1$, and then $b^{\theta_a^A}(x) = 1$. If $a(x) \neq 1$, $a(x) = 0$, and we have $b^{\theta_a^A}(x) = 1 \iff b(A \wedge x) = 1$. But $b(A) \neq 1$. Hence $b(A \wedge x) \neq 1$ and then $b^{\theta_a^A}(x) = 0$. Therefore, $a = b^{\theta_a^A}$, as wanted. ☒

Observe that in the proof above, we have showed not only that the operator $(-)^{\theta_a^A}$ collapses all classical valuations to models of A. We also showed that classical valuations that are not models of A are collapsed to one single point: the valuation a. Moreover, observe that the election of a in the proof above is arbitrary except for the requirement $a(A) = 1$. Thus, every classical valuation satisfying A induces a substitution that makes A projective.

As a consequence of Proposition 3.19 and Proposition 3.27, in **CPL** we have the following corollary.

Corollary 3.28. *Let $X \subseteq Var$ and $A \in F(X)$. The following conditions are equivalent:*

(i) *A is unifiable;*

(ii) *A is satisfiable;*

(iii) *A is projective.*

Finally, we have the unification theorem for **CPL**.

Theorem 3.29. *The unification type of* **CPL** *is unitary, that is, if $A \in F(X)$ is a unifiable formula, then there is a most general unifier of A.*

Proof. Let $A \in F(X)$ be a unifiable formula. By Proposition 3.17, A is satisfiable. Let $a\colon F(X) \to \{0,1\}$ be a classical valuation such that $a(A) = 1$. By Proposition 3.27, A is projective with θ_a^A. By Remark 3.25, θ_a^A is a most general unifier of A. Thus $\{\theta_a^A\}$ is a basis of unifiers of A, as we wanted. ☒

In the previous proof, observe that the only request for the classical valuation $a\colon F(X) \to \{0,1\}$ is that $a(A) = 1$. Hence, if $A \in F(X)$, then for all classical valuations $b\colon F(X) \to \{0,1\}$, such that $a(A) = 1$, $\{\theta_b^A\}$ is a base for the set of unifiers of A. In particular, $\theta_a^A \preceq \theta_b^A$ for every pair of classical valuations $a, b\colon F(X) \to \{0,1\}$ satisfying A.

3.5 Unification in intuitionistic logic

From now on, unless the contrary is said, we will work with **IPL**. Following remarks 3.2 and 3.5, we will use the semantical version of the definitions in the previous sections. For simplicity, since there is no ambiguity, we will omit the explicit reference to **IPL** in the notation. Thus, we will use $\vdash$ instead of $\vdash_{\mathsf{IPC}}$ and $\preceq$ instead of $\preceq_{\mathsf{IPC}}$.

We start this section by presenting a counterexample that shows that unification in intuitionistic propositional logic is not unitary. We also take this opportunity to present a very useful model construction and the disjunction property for intuitionistic propositional logic.

Definition 3.30. Let $X \subseteq Var$. A *sum of disjoint models* $\boldsymbol{M}_1, \ldots, \boldsymbol{M}_n \in \mathcal{M}_X$ (that is, $M_i \cap M_j = \varnothing$ for all $i, j \in \{1, \ldots, n\}, i \neq j$) is any model $\boldsymbol{M}$ in $\mathcal{M}_X$ satisfying the following conditions:

- $M = M_1 \cup \cdots \cup M_n \cup \{r\}$, where $r \notin M_i$ for all $1 \leqslant i \leqslant n$.
- $\leqslant_M = (\bigcup_{1 \leqslant i \leqslant n} \leqslant_{M_i}) \cup \{(r, x) : x \in M_1 \cup \cdots \cup M_n \cup \{r\}\}$.
- $v_M : M_1 \cup \cdots \cup M_n \cup \{r\} \to \mathcal{P}(X)$ is the valuation such that $v_M(r) = \varnothing$ and $v_M(m) = v_{M_i}(m)$ for $m \neq r$, where i is the only natural number such that $m \in M_i$.

A *sum* of arbitrary models $\boldsymbol{N}_1, \ldots, \boldsymbol{N}_n \in \mathcal{M}_X$ is any sum of disjoint models $\boldsymbol{N}'_1, \ldots, \boldsymbol{N}'_n \in \mathcal{M}_X$, where $\boldsymbol{N}_i \cong \boldsymbol{N}'_i$ for each i.

The intuitive idea behind this definition is to build a new model from a finite list of models putting the models of the list one side by side after the other and adding a common root where no propositional variable holds.

Remark 3.31. It is easy to check that if $\boldsymbol{N}$ and $\boldsymbol{N}'$ are sums of arbitrary models $\boldsymbol{N}_1, \ldots, \boldsymbol{N}_n \in \mathcal{M}_X$, then $\boldsymbol{N} \cong \boldsymbol{N}'$. Since isomorphic models are indistinguishable from the point of view of logic, as we showed in Proposition 2.25, we will use $\boldsymbol{N}_1 + \cdots \boldsymbol{N}_n$, or simply $\sum_{1 \leqslant i \leqslant n} \boldsymbol{N}_i$, to denote any sum of models $\boldsymbol{N}_1, \ldots, \boldsymbol{N}_n$.

Remark 3.32. Furthermore, it follows from the definitions that if $\boldsymbol{N}_1, \ldots, \boldsymbol{N}_n, \boldsymbol{N}'_1, \ldots, \boldsymbol{N}'_n \in \mathcal{M}_X$ and $\boldsymbol{N}_i \cong \boldsymbol{N}'_i$ for all $1 \leqslant i \leqslant n$, then $\sum_{1 \leqslant i \leqslant n} \boldsymbol{N}_i \cong \sum_{1 \leqslant i \leqslant n} \boldsymbol{N}'_i$.

A nice application of the sum of models is the following proof of the disjunction property for **IPL**.

Lemma 3.33 (Disjunction Property). *Let $X \subseteq Var$ and $A, B \in F(X)$. If $\models A \vee B$, then $\models A$ or $\models B$.*

Proof. Suppose $\models A \vee B$ and assume by contradiction that $\not\models A$ and $\not\models B$. Then, there are models $\boldsymbol{M}, \boldsymbol{N} \in \mathcal{M}_X$ such that $\boldsymbol{M} \not\models A$ and $\boldsymbol{N} \not\models B$. Clearly, $\boldsymbol{M} + \boldsymbol{N}, r_{M+N} \not\models A \vee B$, and then $\boldsymbol{M} + \boldsymbol{N} \not\models A \vee B$, obtaining a contradiction with $\models A \vee B$. ⊠

Example 3.34 (Unification type of IPL is not unitary). Let $X \subseteq Var$ and $y \in X$. Consider the formula $A := y \vee \neg y$. The substitutions $\sigma_1 \colon F(X) \to F(\varnothing)$ and $\sigma_2 \colon F(X) \to F(\varnothing)$ such that $\sigma_1(x) = \top$ and $\sigma_2(x) = \bot$ for all $x \in X$ are unifiers of A. However, there is no most general unifier of A. To see this, suppose by contradiction that there is a most general unifier $\sigma \colon F(X) \to F(Y)$ for A, where $Y \subseteq Var$. Then, there are substitutions $\tau_1 \colon F(Y) \to F(\varnothing)$ and $\tau_2 \colon F(Y) \to F(\varnothing)$ such that $\models \tau_1(\sigma(x)) \leftrightarrow \sigma_1(x)$ and $\models \tau_2(\sigma(x)) \leftrightarrow \sigma_2(x)$ for all $x \in X$. In particular, $\models \tau_1(\sigma(y)) \leftrightarrow \sigma_1(y)$ and

$\models \tau_2(\sigma(y)) \leftrightarrow \sigma_2(y)$, that is, $\models \tau_1(\sigma(y)) \leftrightarrow \top$ and $\models \tau_2(\sigma(y)) \leftrightarrow \bot$. But σ is a unifier of A, and consequently $\models \sigma(y \vee \neg y)$, so $\models \sigma(y) \vee \neg\sigma(y)$. By the disjunction property, $\models \sigma(y)$ or $\models \neg\sigma(y)$. Now we reason by cases. On the one hand, if $\models \sigma(y)$, then $\models \sigma(y) \leftrightarrow \top$. Thus, $\models \tau_2(\sigma(y)) \leftrightarrow \tau_2(\top)$, by the substitution invariance, and therefore $\models \tau_2(\sigma(y)) \leftrightarrow \top$, contradicting $\models \tau_2(\sigma(y)) \leftrightarrow \bot$. On the other hand, if $\models \neg\sigma(y)$, then $\models \sigma(y) \leftrightarrow \bot$. Following similar steps, we conclude $\models \tau_1(\sigma(y)) \leftrightarrow \bot$, contradicting $\models \tau_1(\sigma(y)) \leftrightarrow \top$. Therefore, A can not have a most general unifer.

Substitutions σ_Y^A and properties of $(-)^{\sigma_Y^A}$

In the proof of the unification theorem for **CPL**, we worked with the substitution θ_a^A, and the result was relatively easy due to the good behaviour of this substitution, in the sense that it makes A projective whenever A is satisfiable with a. In **IPL**, we will work with a similar tool. Given a formula $X \subseteq Var$, $A \in F(X)$ and a subset $Y \subseteq X$, we define $\sigma_Y^A \colon F(X) \to F(X)$ as the substitution such that

$$\sigma_Y^A(x) = \begin{cases} A \to x & \text{if } x \in Y \\ A \wedge x & \text{otherwise} \end{cases} \tag{3.2}$$

The following lemmas and corollaries describe the main properties of the operator σ_Y^A.

Lemma 3.35. *Let $X \subseteq Var$, $A \in F(X)$, $\boldsymbol{M} \in \mathcal{M}_X$ and $Y \subseteq X$. For all $m \in M$ such that $\boldsymbol{M}, m \models A$ and all $x \in X$, we have $\boldsymbol{M}, m \models x$ if and only if $\boldsymbol{M}^{\sigma_Y^A}, m \models x$ (equivalently, $x \in v_M(m)$ if and only if $x \in v_M^{\sigma_Y^A}(m)$).*

Proof. Let $x \in X$ and let $m \in M$ be such that $\boldsymbol{M}, m \models A$. Recall that $\boldsymbol{M}^{\sigma_Y^A}, m \models x \iff \boldsymbol{M}, m \models \sigma_Y^A(x)$. The formula $\sigma_Y^A(x)$ can be either $A \to x$ or $A \wedge x$. But since $\boldsymbol{M}, m \models A$, in any case we have $\boldsymbol{M}, m \models \sigma_Y^A(x) \iff \boldsymbol{M}, p \models x$. Therefore, $\boldsymbol{M}, m \models A \iff \boldsymbol{M}^{\sigma_Y^A}, m \models x$, as desired. ⊠

Corollary 3.36. *Let $X \subseteq Var$, $A \in F(X)$ and $Y \subseteq X$. Then $\boldsymbol{M}^{\sigma_Y^A} = \boldsymbol{M}$ for all $\boldsymbol{M} \in Mod_X(A)$ (and its equivalent statements: $A \models x \leftrightarrow \sigma_Y^A(x)$ for all $x \in X$, $A \models B \leftrightarrow \sigma_Y^A(B)$ for all $B \in F(X)$).*

Lemma 3.37. *Let $X \subseteq Var$, $A \in F(X)$, $\boldsymbol{M} \in \mathcal{M}_X$ and $Y \subseteq X$. For all $m \in M$ such that $\boldsymbol{M}, m \not\models A$ and all $x \in X$, if $\boldsymbol{M}^{\sigma_Y^A}, m \models x$, then $x \in Y$ (equivalently, $v_M^{\sigma_Y^A}(m) \subseteq Y$).*

Proof. Let $x \in X$ and let $m \in M$ be such that $\boldsymbol{M}, m \not\models A$. Assume $\boldsymbol{M}^{\sigma_Y^A}, m \models x$. Then $\boldsymbol{M}, m \models \sigma_Y^A(x)$. By contradiction, suppose $x \notin Y$. This means that $\sigma_Y^A(x) = A \wedge x$. As $\boldsymbol{M}, m \models \sigma_Y^A(x)$, $\boldsymbol{M}, m \models A$, contradiction the assumption $\boldsymbol{M}, m \not\models A$. ⊠

Corollary 3.38. *Let $X \subseteq Var$, $A \in F(X)$ and $Y \subseteq X$. For all $\boldsymbol{M} \in \mathcal{M}_X$ such that $\boldsymbol{M} \not\models A$ and for every $x \in X$, if $\boldsymbol{M}^{\sigma_Y^A} \models x$, then $x \in Y$ (equivalently, $v_M(r_M)^{\sigma_Y^A} \subseteq Y$).*

Lemma 3.39. *Let $X \subseteq Var$, $A \in F(X)$, $\boldsymbol{M} \in \mathcal{M}_X$ and $Y \subseteq X$. For all $m \in M$ and all $x \in X$, $\boldsymbol{M}^{\sigma_Y^A}, m \models x$ if and only if $\boldsymbol{M}^{\sigma_Y^A \sigma_Y^A}, m \models x$ (equivalently, $v_M^{\sigma_Y^A} = v_M^{\sigma_Y^A \sigma_Y^A}$).*

Proof. Let $m \in M$ and $x \in X$. We reason by cases. If $\boldsymbol{M}, m \models A$, a double application of Lemma 3.35 gives us $\boldsymbol{M}, m \models x$ if and only of $\boldsymbol{M}^{\sigma_Y^A}, m \models x$ if and only if $\boldsymbol{M}^{\sigma_Y^A \sigma_Y^A}, m \models x$, and we are done.

If $\boldsymbol{M}, m \not\models A$, observe that $\boldsymbol{M}^{\sigma_Y^A}, m \models x \iff \boldsymbol{M}, m \models \sigma_Y^A(x)$ and $\boldsymbol{M}^{\sigma_Y^A \sigma_Y^A}, m \models x \iff \boldsymbol{M}^{\sigma_Y^A}, m \models \sigma_Y^A(x)$. So to establish the desired conclusion, it is enough to show that $\boldsymbol{M}, m \models \sigma_Y^A(x) \iff \boldsymbol{M}^{\sigma_Y^A}, m \models \sigma_Y^A(x)$. We have two options: $x \in Y$ or $x \notin Y$.

If $x \notin Y$, $\boldsymbol{M}, m \models \sigma_Y^A(x) \iff \boldsymbol{M}, m \models A \wedge x$ and $\boldsymbol{M}^{\sigma_Y^A}, m \models \sigma_Y^A(x) \iff \boldsymbol{M}^{\sigma_Y^A}, m \models A \wedge x \iff \boldsymbol{M}, m \models \sigma_Y^A(A \wedge x) \iff \boldsymbol{M}, m \models \sigma_Y^A(A) \wedge \sigma_Y^A(x) \iff \boldsymbol{M}, m \models \sigma_Y^A(A) \wedge (A \wedge x)$. Thus, in this case the conclusion is reduced to prove $\boldsymbol{M}, m \models A \wedge x \iff \boldsymbol{M}, m \models \sigma_Y^A(A) \wedge A \wedge x$, which is true, since both statements are false, for $\boldsymbol{M}, m \not\models A$.

If $x \in Y$, $\boldsymbol{M}, m \models \sigma_Y^A(x) \iff \boldsymbol{M}, m \models A \to x$ and $\boldsymbol{M}^{\sigma_Y^A}, m \models \sigma_Y^A(x) \iff \boldsymbol{M}^{\sigma_Y^A}, m \models A \to x \iff \boldsymbol{M}, m \models \sigma_Y^A(A \to x) \iff \boldsymbol{M}, m \models \sigma_Y^A(A) \to \sigma_Y^A(x) \iff \boldsymbol{M}, m \models \sigma_Y^A(A) \to (A \to x)$. Thus, the conclusion is proved if we manage to prove $\boldsymbol{M}, m \models A \to x \iff \boldsymbol{M}, m \models \sigma_Y^A(A) \to (A \to x)$. The left-to-right direction is immediate. For the other direction, assume $\boldsymbol{M}, m \models \sigma_Y^A(A) \to (A \to x)$ and let $n \in M$ be such that $n \geqslant m$ and $\boldsymbol{M}, n \models A$. We want $\boldsymbol{M}, n \models x$. But $\boldsymbol{M}, n \models A$ is equivalent to $\boldsymbol{M}_n \models A$. By Colorally 3.36, $\boldsymbol{M}_n^{\sigma_Y^A} = \boldsymbol{M}_n$. Thus, $\boldsymbol{M}_n^{\sigma_Y^A} \models A$. By Lemma 3.14, $\boldsymbol{M}_n \models \sigma_Y^A(A)$ and then $\boldsymbol{M}, n \models \sigma_Y^A(A)$. As $\boldsymbol{M}, m \models \sigma_Y^A(A) \to (A \to x)$, $\boldsymbol{M}, n \models A \to x$. And since $\boldsymbol{M}, n \models A$, also $\boldsymbol{M}, n \models x$, as we wanted. ⊠

Corollary 3.40. *Let $X \subseteq Var$, $A \in F(X)$ and $Y \subseteq X$. For all $\boldsymbol{M} \in \mathcal{M}_X$, it holds that $\boldsymbol{M}^{\sigma_Y^A} = \boldsymbol{M}^{\sigma_Y^A \sigma_Y^A}$.*

Lemma 3.41. *Let $X \subseteq Var$, $A \in F(X)$, $\boldsymbol{M} \in \mathcal{M}_X$ and $Y, Z \subseteq X$. If for all $m \in M$*

$$\boldsymbol{M}^{\sigma_Z^A}, m \models A \quad \textit{if and only if} \quad \boldsymbol{M}, m \models A$$

then for all $m \in M$

$$\boldsymbol{M}^{\sigma_Y^A}, m \models x \quad \textit{if and only if} \quad \boldsymbol{M}^{\sigma_Z^A \sigma_Y^A}, m \models x.$$

Proof. Let $m \in M$. We may assume that $\boldsymbol{M}, m \not\models A$. Otherwise, by Lemma 3.35, the result holds. Observe that $\boldsymbol{M}^{\sigma_Y^A}, m \models x \iff \boldsymbol{M}, m \models \sigma_Y^A(x)$ and $(\boldsymbol{M}^{\sigma_Z^A})^{\sigma_Y^A}, m \models x \iff \boldsymbol{M}^{\sigma_Y^A}, m \models \sigma_Z^A(x)$. Thus, it is enough to show that $\boldsymbol{M}, m \models \sigma_Y^A(x) \iff \boldsymbol{M}^{\sigma_Z^A}, m \models \sigma_Y^A(x)$. We have two cases: $x \in Y$ or $x \notin Y$.

If $x \notin Y$, $\boldsymbol{M}, m \models \sigma_Y^A(x) \iff \boldsymbol{M}, m \models A \wedge x$ and $\boldsymbol{M}^{\sigma_Z^A}, m \models \sigma_Y^A(x) \iff \boldsymbol{M}^{\sigma_Z^A}, m \models A \wedge x$. Thus, in this case the conclusion is reduced to prove $\boldsymbol{M}, m \models A \wedge x \iff \boldsymbol{M}^{\sigma_Z^A}, m \models A \wedge x$, which is true by the hypothesis in the statement.

If $x \in Y$, $\boldsymbol{M}, m \models \sigma_Y^A(x) \iff \boldsymbol{M}, m \models A \to x$ and $\boldsymbol{M}^{\sigma_Z^A}, m \models \sigma_Y^A(x) \iff \boldsymbol{M}^{\sigma_Z^A}, m \models A \to x \iff \boldsymbol{M}, m \models \sigma_Z^A(A \to x) \iff \boldsymbol{M}, m \models \sigma_Z^A(A) \to \sigma_Z^A(x)$. Thus, the conclusion is established if we prove $\boldsymbol{M}, m \models A \to x \iff \boldsymbol{M}, m \models \sigma_Z^A(A) \to \sigma_Z^A(x)$. For the left-to-right direction, assume $\boldsymbol{M}, m \models A \to x$ and let $n \in M$ be such that

$n \geqslant m$ and $\boldsymbol{M}, n \models \sigma_Z^A(A)$. We want $\boldsymbol{M}, n \models \sigma_Z^A(x)$, which is proved as follows:

$$\begin{aligned}
\boldsymbol{M}, n \models \sigma_Z^A(A) &\iff \boldsymbol{M}^{\sigma_Z^A}, n \models A \\
&\iff \boldsymbol{M}, n \models A && \text{(by the hypothesis in the statement)} \\
&\iff \boldsymbol{M}, n \models x && \text{(since } \boldsymbol{M}, m \models A \to x\text{)} \\
&\iff \boldsymbol{M}^{\sigma_Z^A}, n \models x && \text{(by Lemma 3.35, since } \boldsymbol{M}, n \models A\text{)} \\
&\iff \boldsymbol{M}, n \models \sigma_Z^A(x)
\end{aligned}$$

For the right-to-left direction, assume $\boldsymbol{M}, m \models \sigma_Z^A(A) \to \sigma_Z^A(x)$ and let $n \in M$ be such that $n \geqslant m$ and $\boldsymbol{M}, n \models A$. We want $\boldsymbol{M}, n \models x$, which is proved as follows:

$$\begin{aligned}
\boldsymbol{M}, n \models A &\iff \boldsymbol{M}, n \models \sigma_Z^A(A) && \text{(as } A \models A \leftrightarrow \sigma_Z^A(A) \text{ by Corollary 3.36)} \\
&\implies \boldsymbol{M}, n \models \sigma_Z^A(x) && \text{(since } \boldsymbol{M}, m \models \sigma_Z^A(A) \to \sigma_Z^A(x)\text{)} \\
&\iff \boldsymbol{M}^{\sigma_Z^A}, n \models x \\
&\iff \boldsymbol{M}, n \models x && \text{(by Lemma 3.35, since } \boldsymbol{M}, n \models A\text{)}
\end{aligned}$$

$\boxtimes$

Corollary 3.42. *Let $X \subseteq Var$, $A \in F(X)$, $\boldsymbol{M} \in \mathcal{M}_X$ and $Y, Z \subseteq X$. If for all $m \in M$*

$$\boldsymbol{M}^{\sigma_Z^A}, m \models A \quad \textit{if and only if} \quad \boldsymbol{M}, m \models A$$

then $\boldsymbol{M}^{\sigma_Y^A} = \boldsymbol{M}^{\sigma_Z^A \sigma_Y^A}$.

We summarize the achivements about $(-)^{\sigma_Y^A}$ made so far:

- The operator $(-)^{\sigma_Y^A}$ does not modify the worlds/models where A holds.
- After the application of $(-)^{\sigma_Y^A}$, the propositional variables valid in any world where A originally did not hold are always in Y.
- The opertator $(-)^{\sigma_Y^A}$ is idempotent.
- If we compose two operators $(-)^{\sigma_Z^A}$ and $(-)^{\sigma_Y^A}$, and the first operator acting does not make A true in any new world, then its action is completely irrelevant in the composition.

Corollary 3.36 states that σ_Y^A satisfies the second condition in order to make A a projective formula. The reader can easily check that, in general, substitutions σ_Y^A do not satisfy the first condition, that is, σ_Y^A is not always a unifier of A, or equivalently, the operator $(-)^{\sigma_Y^A}$ does not collapse all models of $\mathcal{M}_X$ to $Mod_X(A)$. However, the application of $(-)^{\sigma_Y^A}$ to a certain model might make that A holds in more worlds of the model. By Lemma 3.35, it never occurs that A becomes false in a world after the application $(-)^{\sigma_Y^A}$ if A was already true in that world before the action of $(-)^{\sigma_Y^A}$. What if we systematically apply substitutions of the form $(-)^{\sigma_Y^A}$? Will we end up with a model of A? If so, due to Lemma 3.26 the composition of the the substitutions used would be a substitution that makes A projective. We explore this idea in detail in the rest of this section.

For the rest of this section, we will assume that $X \subseteq Var$ is finite, and $X_1, X_2, \ldots, X_n$ is a sequence of all the subsets of X satisfying the following condition: if $X_i \subseteq X_j$ then $i \leqslant j$. Examples of these kind of sequences are the lists obtained by listing first the empty set, then all the subsets of X of one element, then all the subsets of X of two elements, then all the subsets of X of three elements, and so on until listing the set X itself.

Define the substitution $\sigma^A\colon F(X) \to F(X)$ as the composition $\sigma^A_{X_n} \cdots \sigma^A_{X_2}\sigma^A_{X_1}$. The substitution σ^A has its corresponding operator $(-)^{\sigma^A} : \mathcal{M}_X \to \mathcal{M}_X$. By Lemma 3.14 (iii), we know that for all $\boldsymbol{M} \in \mathcal{M}_X$, $(((\boldsymbol{M}^{\sigma^A_{X_n}})\cdots)^{\sigma^A_{X_2}})^{\sigma^A_{X_1}} = \boldsymbol{M}^{\sigma^A}$. Therefore, the operator $(-)^{\sigma^A}$ can be characterized as the composition $(-)^{\sigma^A_{X_1}} \circ (-)^{\sigma^A_{X_2}} \circ \cdots \circ (-)^{\sigma^A_{X_n}}$. However, note the reverse order: $\sigma^A = \sigma^A_{X_n} \circ \cdots \circ \sigma^A_{X_2} \circ \sigma^A_{X_1}$ and $(-)^{\sigma^A} = (-)^{\sigma^A_{X_1}} \circ (-)^{\sigma^A_{X_2}} \circ \cdots \circ (-)^{\sigma^A_{X_n}}$.

Lemma 3.43. *Let $A \in F(X)$ and $\boldsymbol{M} \in \mathcal{M}_X$ be such that $\boldsymbol{M}^{\sigma^A} \models A$. If $X_i \subseteq v^{\sigma^A}_M(r)$ for some $1 \leqslant i \leqslant n$, then $\boldsymbol{M}^{\sigma^A_{X_n}\cdots\sigma^A_{X_i}} \models A$.*

Proof. Let $i \in \{1, \ldots, n\}$ be such that $X_i \subseteq v^{\sigma^A}_M(r)$. We may assume that $\boldsymbol{M}^{\sigma^A_{X_n}} \not\models A$. Otherwise, by Corollary 3.36, $\boldsymbol{M}^{\sigma^A_{X_n}\cdots\sigma^A_{X_j}} \models A$ for all $j \in \{1, \ldots, n\}$ and we are done. Then, since $\boldsymbol{M}^{\sigma^A} \models A$ and $\boldsymbol{M}^{\sigma^A_{X_n}} \not\models A$, there is a $k \in \{1, \ldots, n-1\}$ such that $\boldsymbol{M}^{\sigma^A_{X_n}\cdots\sigma^A_{X_k}} \models A$ but $\boldsymbol{M}^{\sigma^A_{X_n}\cdots\sigma^A_{X_{k+1}}} \not\models A$. We want $\boldsymbol{M}^{\sigma^A_{X_n}\cdots\sigma^A_{X_i}} \models A$. As $\boldsymbol{M}^{\sigma^A_{X_n}\cdots\sigma^A_{X_k}} \models A$, to obtain the result it is enough to show that $i = k$ or, by and Corollary 3.36, that $(-)^{\sigma^A_{X_i}}$ acts after $(-)^{\sigma^A_{X_k}}$, that is, $i < k$ (recall the reverse order). We will show then that $i \leqslant k$. Observe that, by Corollary 3.38, we know that $v_M^{\sigma^A_{X_n}\cdots\sigma^A_{X_k}}(r) \subseteq X_k$, and since $\boldsymbol{M}^{\sigma^A_{X_n}\cdots\sigma^A_{X_k}} \models A$, by Corollary 3.36 we have $v_M^{\sigma^A_{X_n}\cdots\sigma^A_{X_k}}(r) = v_M^{\sigma^A_{X_n}\cdots\sigma^A_{X_1}}(r) = v^{\sigma^A}_M(r)$. Thus, $v^{\sigma^A}_M(r) \subseteq X_k$. But $X_i \subseteq v^{\sigma^A}_M(r)$, by the hypothesis in the statement. Hence $X_i \subseteq X_k$. However, by the condition on the list $X_1, \ldots, X_n$, this implies $i \leqslant k$, which concludes the proof. ⊠

Definition 3.44. A model $\boldsymbol{M'} \in \mathcal{M}_X$ is said to be a *variant* of the model $\boldsymbol{M} \in \mathcal{M}_X$ if and only if $M' = M$, $\leqslant_{M'} = \leqslant_M$ and $v_{M'}(m) = v_M(m)$ for all $m \neq r$.

Remark 3.45. If $\boldsymbol{M'}$ is a variant of $\boldsymbol{M}$, the valuation $v_{M'}$ of $\boldsymbol{M'}$ is completely determined by the valuation v_M in $\boldsymbol{M}$ and $v_{M'}(r_M)$.

Lemma 3.46. *Let $A \in F(X)$ and $\boldsymbol{M} \in \mathcal{M}_X$ be such that $\boldsymbol{M} \not\models A$ but $\boldsymbol{M}, m \models A$ for all $m \neq r$. If there is a variant $\boldsymbol{M'}$ of $\boldsymbol{M}$ such that $\boldsymbol{M'} \models A$, then $\boldsymbol{M'} = \boldsymbol{M}^{\sigma^A_Y}$ where $Y = v_{M'}(r)$.*

Proof. By definition, the frames of $\boldsymbol{M'}$ and $\boldsymbol{M}^{\sigma^A_Y}$ are the same. Thus, to prove the equality $\boldsymbol{M'} = \boldsymbol{M}^{\sigma^A_Y}$ we need to show that $v_{M'} = v^{\sigma^A_Y}_M$. However, since $\boldsymbol{M}, m \models A$ for all $m \neq r$, using Lemma 3.35 and the definition of variant, we already know that $v_{M'}(m) = v^{\sigma^A_Y}_M(m)$ for every $m \neq r$. It only remains to prove that $v_{M'}(r) = v^{\sigma^A_Y}_M(r)$. We do it by showing the double inclusion.

Let $x \in v_{M'}(r) = Y$. We want $x \in v_M^{\sigma_Y^A}(r)$. But $x \in v_M^{\sigma_Y^A}(r) \iff \boldsymbol{M}^{\sigma_Y^A}, r \models x \iff \boldsymbol{M}, r \models \sigma_Y^A(x) \iff \boldsymbol{M}, r \models A \to x$. We will prove this last statement. So let $m \geqslant r$ be such that $\boldsymbol{M}, m \models A$. Observe that $m \neq r$, for $\boldsymbol{M} \not\models A$. Furthermore, since $x \in v_{M'}(r)$, $\boldsymbol{M}', m \models x$, by the truth-preservation condition. But $\boldsymbol{M}'$ is a variant of $\boldsymbol{M}$ and $m \neq r$. As a consequence, $\boldsymbol{M}, m \models x$. Hence, $\boldsymbol{M}, r \models A \to x$ and this shows the inclusion $v_{M'}(r) \subseteq v_M^{\sigma_Y^A}(r)$.

Now let $x \in v_M^{\sigma_Y^A}(r)$. Then, $\boldsymbol{M}^{\sigma_Y^A}, r \models x$ and so $\boldsymbol{M}, r \models \sigma_Y^A(x)$. Suppose, by contradiction, that $x \notin v_{M'}(r) = Y$. Reasoning as above, $x \in v_M^{\sigma_Y^A}(r) \iff \boldsymbol{M}^{\sigma_Y^A}, r \models x \iff \boldsymbol{M}, r \models \sigma_Y^A(x) \iff \boldsymbol{M}, r \models A \wedge x$. In particular $\boldsymbol{M}, r \models A$, contradicting $\boldsymbol{M} \not\models A$. Thus, $v_M^{\sigma_Y^A}(r) \subseteq v_{M'}(r)$. ☒

It is time to introduce one of the key concepts of this and next chapter.

Definition 3.47. A class $\mathcal{K}$ of Kripke models is said to have the *extension property* if and only if for every $n \in \omega$ and all $\boldsymbol{M}_1, \ldots, \boldsymbol{M}_n \in \mathcal{K}$, there is a variant of (some) $\sum_{1 \leqslant i \leqslant n} \boldsymbol{M}_i$ in $\mathcal{K}$.

Remark 3.48. If $\mathcal{K}$ is a class closed under isomorphism (that is, if for all $\boldsymbol{M} \in \mathcal{K}$ and every model $\boldsymbol{N}$ such that $\boldsymbol{M} \cong \boldsymbol{N}$, $\boldsymbol{N} \in \mathcal{K}$ holds), the definition above can be restricted to pairwise disjoint models. That is: a class $\mathcal{K}$ of Kripke models closed under isomorphism has the *extension property* if and only if for every $n \in \omega$ and all pairwise disjoint models $\boldsymbol{M}_1, \ldots, \boldsymbol{M}_n \in \mathcal{K}$, there is a variant of $\sum_{1 \leqslant i \leqslant n} \boldsymbol{M}_i$ in $\mathcal{K}$.

We end this section with a characterization of projective formulas using σ_Y^A and the extension property.

Theorem 3.49. *Let $A \in F(X)$. The following conditions are equivalent:*

(i) *σ^A is a unifier of A;*

(ii) *A is projetive;*

(iii) *$Mod_X(A)$ has the extension property.*

Proof.

(i) $\Rightarrow$ (ii). Recall that each $\sigma_{X_i}^A$ appearing in σ^A satisfies $A \models x \leftrightarrow \sigma_{X_i}^A(x)$ and substitutions satisfying such condition are closed under composition. Thus, σ^A satisfies $A \models x \leftrightarrow \sigma^A(x)$. If, moreover, σ^A is a unifier of A, then A is projective with σ^A.

(ii) $\Rightarrow$ (iii). Let $\boldsymbol{M}_1, \ldots, \boldsymbol{M}_n \in Mod_X(A)$. We need to find a variant of $\sum_{1 \leqslant i \leqslant n} \boldsymbol{M}_i$ in $Mod_X(A)$. Since A is projective, there is a substitution $\sigma\colon F(X) \to F(X)$ such that $\boldsymbol{M}^\sigma \in Mod_X(A)$ for all $\boldsymbol{M} \in \mathcal{M}_X$ and $\boldsymbol{M}^\sigma = \boldsymbol{M}$ for all $\boldsymbol{M} \in Mod_X(A)$. In particular, $(\sum_{1 \leqslant i \leqslant n} \boldsymbol{M}_i)^\sigma$ belongs to $Mod_X(A)$. We will show that $(\sum_{1 \leqslant i \leqslant n} \boldsymbol{M}_i)^\sigma$ is a variant of $\sum_{1 \leqslant i \leqslant n} \boldsymbol{M}_i$, and thus it is the model we want. Observe for all $m \in \sum_{1 \leqslant i \leqslant n} \boldsymbol{M}_i$ different from the root, $\sum_{1 \leqslant i \leqslant n} \boldsymbol{M}_i, m \models A$, as $\boldsymbol{M}_1, \ldots, \boldsymbol{M}_n \in Mod_X(A)$. Therefore, using Remark 3.23 and Proposition 3.13, we conclude that $(\sum_{1 \leqslant i \leqslant n} \boldsymbol{M}_i)^\sigma$ is a variant of $(\sum_{1 \leqslant i \leqslant n} \boldsymbol{M}_i)$.

(iii) $\Rightarrow$ (i). Assume $Mod_X(A)$ has the extension property. And suppose by contradiction that σ^A is not a unifier of A. Thus, there is a model $\boldsymbol{M} \in \mathcal{M}_X$ such that $\boldsymbol{M} \not\models \sigma^A(A)$. Without loss of generality, we may assume that:

- For all $m \in M$ with m different from the root r, $\boldsymbol{M}, m \models \sigma^A(A)$. If $\boldsymbol{M}$ does not have this property, take a submodel of $\boldsymbol{M}$ that does have that property.

- If $m_1, \ldots, m_n$ are the immediate successors of the root r, then $\boldsymbol{M}_{m_1}, \ldots, \boldsymbol{M}_{m_n}$ are mutually disjoint (that is $M_i \cap M_j = \emptyset$ for all $i \neq j$). If $\boldsymbol{M}$ does not have this property, take a model with root r and followed with isomorphic copies of $\boldsymbol{M}_1, \ldots, \boldsymbol{M}_n$ mutually disjoints.

By Lemma 3.14, $\boldsymbol{M}^{\sigma^A} \not\models A$ but $\boldsymbol{M}^{\sigma^A}, m \models A$ for all $m \neq r$. Thus, by Corollary 3.36, $\boldsymbol{M}_{m_i}^{\sigma^A} = \boldsymbol{M}_{m_i}$ for each m_i immediate successor of r, and then each $\boldsymbol{M}_{m_i}^{\sigma^A}$ is in $Mod_X(A)$. Thus, $\sum_{1 \leqslant i \leqslant n} \boldsymbol{M}_{m_i}^{\sigma^A}$ is a variant of $\boldsymbol{M}^{\sigma^A}$. Moreover, since $Mod_X(A)$ has the extension property, there is a variant $\boldsymbol{N}$ of $\sum_{1 \leqslant i \leqslant n} \boldsymbol{M}_{m_i}^{\sigma^A}$ in $Mod_X(A)$. The model $\boldsymbol{N}$ is obviously also a variant of $\boldsymbol{M}^{\sigma^A}$. By Lemma 3.46, $\boldsymbol{N} = \boldsymbol{M}^{\sigma^A \sigma_Y^A}$ where $Y = v_N(r)$. But $X_1, \ldots, X_n$ is a list of all subsets of X. Thus, $Y = X_i$ for some i, $1 \leqslant i \leqslant n$. Hence, $\boldsymbol{N} = (\boldsymbol{M}^{\sigma^A})^{\sigma_{X_i}^A}$ for some i, $1 \leqslant i \leqslant n$, that is, $\boldsymbol{N} = \boldsymbol{M}^{\sigma_{X_n}^A \cdots \sigma_{X_1}^A \sigma_{X_i}^A}$. We have two options: $i = 1$ of $i > 1$. We will see that both cases lead us to a contradiction.

If $i = 1$, then $\boldsymbol{N} = \boldsymbol{M}^{\sigma_{X_n}^A \cdots \sigma_{X_1}^A \sigma_{X_1}^A} = \boldsymbol{M}^{\sigma_{X_n}^A \cdots \sigma_{X_1}^A} = \boldsymbol{M}^{\sigma^A}$, by Corollary 3.40. But $\boldsymbol{M}^{\sigma^A} \not\models A$ and $\boldsymbol{N} \models A$, obtaining a contradiction.

If $i > 1$, it suffices to show that $\boldsymbol{M}^{\sigma_{X_n}^A \cdots \sigma_{X_1}^A \sigma_{X_i}^A} = \boldsymbol{M}^{\sigma_{X_n}^A \cdots \sigma_{X_i}^A}$. To see why this is enough, observe that if $\boldsymbol{N} = \boldsymbol{M}^{\sigma_{X_n}^A \cdots \sigma_{X_1}^A \sigma_{X_i}^A} = \boldsymbol{M}^{\sigma_{X_n}^A \cdots \sigma_{X_i}^A}$, then, as $\boldsymbol{N} \models A$, Corollary 3.36 would imply $\boldsymbol{N} = \boldsymbol{M}^{\sigma_{X_n}^A \cdots \sigma_{X_1}^A} = \boldsymbol{M}^{\sigma^A}$, contradicting $\boldsymbol{M}^{\sigma^A} \not\models A$ and $\boldsymbol{N} \models A$. To prove the equality $\boldsymbol{M}^{\sigma_{X_n}^A \cdots \sigma_{X_1}^A \sigma_{X_i}^A} = \boldsymbol{M}^{\sigma_{X_n}^A \cdots \sigma_{X_i}^A}$, notice first that $\boldsymbol{M}^{\sigma_{X_n}^A \cdots \sigma_{X_1}^A \sigma_{X_i}^A} \models X_i$ (that is, $\boldsymbol{M}^{\sigma_{X_n}^A \cdots \sigma_{X_1}^A \sigma_{X_i}^A} \models x$ for all $x \in X_i$), for $\boldsymbol{M}^{\sigma_{X_n}^A \cdots \sigma_{X_1}^A \sigma_{X_i}^A} = \boldsymbol{N}$, $\boldsymbol{N} \models v_N(r)$ and $v_N(r) = Y = X_i$. In particular, $\boldsymbol{M}_m^{\sigma_{X_n}^A \cdots \sigma_{X_1}^A \sigma_{X_i}^A} \models X_i$ for all $m \neq r$. Moreover, $\boldsymbol{M}_m^{\sigma_{X_n}^A \cdots \sigma_{X_1}^A} \models A$ for all $m \neq r$, as $\boldsymbol{M}^{\sigma^A}, m \models A$ for all $m \neq r$. But Corollary 3.36 implies $\boldsymbol{M}_m^{\sigma_{X_n}^A \cdots \sigma_{X_1}^A} = \boldsymbol{M}_m^{\sigma_{X_n}^A \cdots \sigma_{X_1}^A \sigma_{X_i}^A}$, and then, as $\boldsymbol{M}_m^{\sigma_{X_n}^A \cdots \sigma_{X_1}^A \sigma_{X_i}^A} \models X_i$, $\boldsymbol{M}_m^{\sigma_{X_n}^A \cdots \sigma_{X_1}^A} \models X_i$ for all $m \neq r$. Lemma 3.43 implies $\boldsymbol{M}_m^{\sigma_{X_n}^A \cdots \sigma_{X_i}^A} \models A$ for all $m \neq r$. Equivalently, $\boldsymbol{M}^{\sigma_{X_n}^A \cdots \sigma_{X_i}^A}, m \models A$ for all $m \neq r$. As $\boldsymbol{M}^{\sigma_{X_n}^A \cdots \sigma_{X_1}^A} \not\models A$ (for $\boldsymbol{M}^{\sigma^A} \not\models A$) and $\boldsymbol{M}^{\sigma_{X_n}^A \cdots \sigma_{X_i}^A}, m \models A$ for all $m \neq r$, Lemma 3.35 implies that the operators $(-)^{\sigma_{X_k}^A}$, for $1 \leqslant k < i$, do not make A valid in new worlds (the only world where A does not hold is r). Hence, by Lemma 3.41, we obtain $\boldsymbol{M}^{\sigma_{X_n}^A \cdots \sigma_{X_1}^A \sigma_{X_i}^A} = \boldsymbol{M}^{\sigma_{X_n}^A \cdots \sigma_{X_i}^A \sigma_{X_i}^A}$. But $\boldsymbol{M}^{\sigma_{X_n}^A \cdots \sigma_{X_i}^A} = \boldsymbol{M}^{\sigma_{X_n}^A \cdots \sigma_{X_i}^A \sigma_{X_i}^A}$, by Corollary 3.40, concluding $\boldsymbol{M}^{\sigma_{X_n}^A \cdots \sigma_{X_1}^A \sigma_{X_i}^A} = \boldsymbol{M}^{\sigma_{X_n}^A \cdots \sigma_{X_i}^A}$.

⊠

Implicational Complexity and n-equivalence

The implicational complexity of a formula $A \in F(X)$ is the number of nested implications in A. More precisely, we have the following recursive definition.

Definition 3.50. Let $X \subseteq Var$. The *implicational complexity* of a formula $A \in F(X)$, denoted as $c(A)$, is defined as:

- $c(A) = 0$ if A is a propositional variable, $\perp$ or $\top$;
- $c(A) = max\{c(B), c(C)\}$ if $A = (B * C)$, where $* \in \{\wedge, \vee\}$;
- $c(A) = max\{c(B), c(C)\} + 1$ if $A = (B \to C)$.

We will use C_n^X to denote the set of all formulas in $F(X)$ of implicational complexity less or equal than n, and $C_n^X/_\equiv$ to denote the set of all equivalence classes of C_n^X under the relation of logical equivalence.

Remark 3.51. Observe that once we have C_n^X, C_{n+1}^X can be defined inductively as:

- For all $A \in C_n^X$, $A \in C_{n+1}^X$.
- For all formulas $B, C \in C_n^X$, $B \to C \in C_{n+1}^X$.
- For all formulas $B, C \in C_{n+1}^X$, $(B * C) \in C_{n+1}^X$, where $* \in \{\wedge, \vee\}$;
- No other formula belongs to C_{n+1}^X.

This fact allows us to do inductive reasoning to prove properties valid for all the formulas in C_{n+1}. To show that every formula of implicational complexity less or equal than $n+1$ has cartain property, it will be enough to prove that the property holds for formulas in C_n, for formulas of the form $B \to C$, where $B, C \in C_n$, and that the property is preserved when applying operations $\wedge$ and $\vee$.

One important part of the the unification theorem for **IPL** rests on the fact that whenever $X \subseteq Var$ is finite, $C_n^X/_\equiv$ is also finite for all $n \in \omega$. To prove this fact, we need a preliminary result. Given $X \subseteq Var$ and $L \subseteq F(X)$, we will use $L_{\wedge\vee}$ to denote the least set of formulas containing L closed under $\wedge$ and $\vee$. That is, $L_{\wedge\vee}$ is the least set of formulas such that: (1) $A \in L_{\wedge\vee}$ for all $A \in L$, and (2) $B \wedge C \in L_{\wedge\vee}$ and $B \vee C \in L_{\wedge\vee}$ for all $C, B \in L_{\wedge\vee}$. The following result says that the truth of formulas in $L_{\wedge\vee}$ at any world (in any model) is completely determined by the truth of formulas in L at that world. The truth of formulas of L at other worlds is irrelevant.

Lemma 3.52. *Let $X \subseteq Var$ and $L \subseteq F(X)$. For all models $\boldsymbol{M}, \boldsymbol{M}' \in \mathcal{M}_X$ and every $m \in \boldsymbol{M}$ and $m' \in \boldsymbol{M}'$,*

(i) *The following equivalence holds*

$$\{[A] \in L/_\equiv : \boldsymbol{M}, m \models A\} = \{[A] \in L/_\equiv : \boldsymbol{M}', m' \models A\}$$
$$\Longleftrightarrow$$
$$\text{For all } B \in L_{\wedge\vee}, \boldsymbol{M}, m \models B \text{ if and only if } \boldsymbol{M}', m' \models B.$$

(ii) *The amount of non-equivalent formulas in $L_{\wedge\vee}$ is bounded by $2^{|L/_\equiv|}$, that is, $|L_{\wedge\vee}/_\equiv| \leqslant 2^{|L/_\equiv|}$.*

Proof.

(i) The right-to-left direction is immediate, as $L \subseteq L_{\wedge\vee}$. For the left-to-right implication, we proceed by induction on $L_{\wedge\vee}$. If $B \in L$, the result holds since $\{A \in L : \boldsymbol{M}, m \models A\} = \{A \in L : \boldsymbol{M}', m' \models A\}$ is a trivial consequence of the hypothesis $\{[A] \in L/_\equiv : \boldsymbol{M}, m \models A\} = \{[A] \in L/_\equiv : \boldsymbol{M}', m' \models A\}$. If $B = C \wedge D$, with C and D in $L_{\wedge\vee}$, and the result holds for C and D, we have

$$\begin{aligned}
\boldsymbol{M}, m \models B &\Longleftrightarrow \boldsymbol{M}, m \models C \wedge D \\
&\Longleftrightarrow \boldsymbol{M}, m \models C \text{ and } \boldsymbol{M}, m \models C \\
&\Longleftrightarrow \boldsymbol{M}, m' \models C \text{ and } \boldsymbol{M}', m' \models C && \text{(Induction Hypothesis)} \\
&\Longleftrightarrow \boldsymbol{M}, m' \models C \wedge D \\
&\Longleftrightarrow \boldsymbol{M}, m' \models B.
\end{aligned}$$

A similar argument works when $B = C \vee D$, with C and D in $L_{\wedge\vee}$, and the result holds for C and D.

(ii) Assume by contradiction that there are more than $2^{|L/_\equiv|}$ non-equivalent formulas in $L_{\wedge\vee}$. The previous result implies that there are more than $2^{|L/_\equiv|}$ subsets of $|L/_\equiv|$, which is a contradiction. ⊠

Lemma 3.53. *If $X \subseteq Var$ is finite, then $C_n^X/_\equiv$ is finite for all $n \in \omega$.*

Proof. The proof goes by induction on n. If $n = 0$, simply observe that $C_0^X = (X \cup \{\bot, \top\})_{\wedge\vee}$. The previous Lemma implies $|C_0^X/_\equiv| \leqslant 2^{|(X\cup\{\bot,\top\})/_\equiv|} = 2^{|X|+2}$, which is finite, since so is X. Now assume that $C_n^X/_\equiv$ is finite. We need to prove that $C_{n+1}^X/_\equiv$ is finite as well. Observe that $C_{n+1}^X = (C_n^X \cup I_n)_{\wedge\vee}$, where $I_n = \{A \to B : A, B \in C_n^X\}$. Notice first that $I_n/_\equiv$ is finite, since so is $C_n^X/_\equiv$. Furthermore, the previous lemma implies $|C_{n+1}^X/_\equiv| \leqslant 2^{|(C_n^X\cup I_n)/_\equiv|}$. But $|(C_n^X \cup I_n)/_\equiv| \leqslant |C_n^X/_\equiv| + |I_n/_\equiv|$, which is finite. Therefore, $|C_{n+1}/_\equiv|$ is finite too. ⊠

Definition 3.54. Let $\boldsymbol{M}, \boldsymbol{N} \in \mathcal{M}_X$.

- $\boldsymbol{M}$ and $\boldsymbol{N}$ are said to be 0-*equivalent*, denoted as $\boldsymbol{M} \sim_0 \boldsymbol{N}$, iff $v_M(r_M) = v_N(r_N)$.
- $\boldsymbol{M}$ and $\boldsymbol{N}$ are said to be $n+1$-*equivalent*, denoted as $\boldsymbol{M} \sim_{n+1} \boldsymbol{N}$, iff for all $p \in M$ there exists a $q \in N$ such that $\boldsymbol{M}_p \sim_n \boldsymbol{N}_q$ and for all $q \in N$ there exists an $p \in M$ such that $\boldsymbol{N}_q \sim_n \boldsymbol{M}_p$.
- $\boldsymbol{M}$ is said to be 0-*less than* $\boldsymbol{N}$, denoted as $\boldsymbol{M} \leqslant_0 \boldsymbol{N}$, iff $v_M(r_M) \supseteq v_N(r_N)$.
- $\boldsymbol{M}$ is said to be $n+1$-*less than* $\boldsymbol{N}$, denoted as $\boldsymbol{M} \leqslant_{n+1} \boldsymbol{N}$, iff for all $p \in M$ there exists a $q \in N$ such that $\boldsymbol{M}_p \sim_n \boldsymbol{N}_q$.

Remark 3.55. It is easy to check that for all $n \in \omega$, $\sim_n$ defines an equivalence relation on $\mathcal{M}_X$, and $\leqslant_n$ defines a preorder on $\mathcal{M}_X$ (i.e. $\leqslant_n$ is reflexive and transitive).

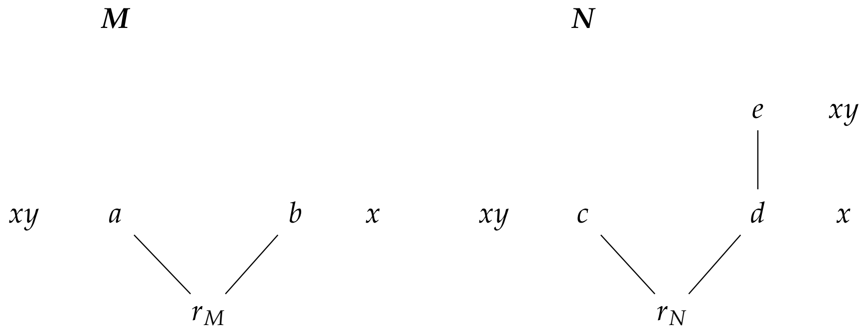
Remark 3.56. It is also immediate from the definitions that given $\boldsymbol{M}, \boldsymbol{N} \in \mathcal{M}_X$, then for all $n \in \omega$, we have $\boldsymbol{M} \leqslant_n \boldsymbol{N}$ and $\boldsymbol{N} \leqslant_n \boldsymbol{M}$ if and only if $\boldsymbol{M} \sim_n \boldsymbol{N}$. This fact, in conjunction with the transitivity of $\leqslant_n$ stated in the remark above, implies:

- If $\boldsymbol{M} \leqslant_n \boldsymbol{N}$ and $\boldsymbol{M} \sim_n \boldsymbol{S}$, then $\boldsymbol{S} \leqslant_n \boldsymbol{N}$;
- If $\boldsymbol{M} \leqslant_n \boldsymbol{N}$ and $\boldsymbol{N} \sim_n \boldsymbol{S}$, then $\boldsymbol{M} \leqslant_n \boldsymbol{S}$.

Remark 3.57. For every $n \in \omega$ and all $\boldsymbol{M}, \boldsymbol{N}, \boldsymbol{M}', \boldsymbol{N}', \boldsymbol{S} \in \mathcal{M}_X$, it can be easily checked that

- If $\boldsymbol{M} \cong \boldsymbol{N}$, then and $\boldsymbol{M} \sim_n \boldsymbol{N}$;
- If $\boldsymbol{M} \cong \boldsymbol{M}'$, $\boldsymbol{N} \cong \boldsymbol{N}'$ and $\boldsymbol{M} \leqslant_n \boldsymbol{N}$, then $\boldsymbol{M}' \leqslant_n \boldsymbol{N}'$;
- If $\boldsymbol{M} \cong \boldsymbol{M}'$, $\boldsymbol{N} \cong \boldsymbol{N}'$ and $\boldsymbol{M} \sim_n \boldsymbol{N}$, then $\boldsymbol{M}' \sim_n \boldsymbol{N}'$.

Example 3.58. Consider the models $\boldsymbol{M}$ and $\boldsymbol{N}$ on the set of variables $\{x, y\}$ depicted below:



Next to each point appear only the variables that are true at that point. For instance, neither x nor y are true in r_M, and x and y are both true in a. We have $\boldsymbol{M} \sim_0 \boldsymbol{N}$, for $v_M(r_M) = \emptyset = v_N(r_N)$, and $\boldsymbol{M} \sim_1 \boldsymbol{N}$, since for all for all $p \in M$ there exists a $q \in N$ such that $\boldsymbol{M}_p \sim_0 \boldsymbol{N}_q$ (that is, $v_M(p) = v_N(q)$) and vice versa. However, $\boldsymbol{M} \not\sim_2 \boldsymbol{N}$. To see this, notice that for $d \in \boldsymbol{N}$, there is no element $p \in \boldsymbol{M}$ such that $\boldsymbol{M}_p \sim_1 \boldsymbol{N}_d$, as the reader can check.

Remark 3.59. Both $\boldsymbol{M} \sim_n \boldsymbol{N}$ and $\boldsymbol{M} \leqslant_n \boldsymbol{N}$ have an equivalent formulation in game theoretical terms. Even though we will not use this formulation, it could shade light on the intuitive understanding of both notions. For $n = 0$, we just define $\boldsymbol{M} \sim_0 \boldsymbol{N}$ iff $v_M(r_M) = v_N(r_N)$ and $\boldsymbol{M} \leqslant_0 \boldsymbol{N}$ iff $v_M(r_M) \supseteq v_N(r_N)$, as before. For $n > 0$, consider the two-player dynamic game with the following rules. Each player has a maximum of n movements. At each movement, Player I challenges Player II by choosing either a point in $\boldsymbol{M}$ or a point in $\boldsymbol{N}$, and Player II must answer the challenge by choosing a point in the other model satisfying the same propositional variables than the point picked by Player I does. If Player II succeeds, the game continues, whenever there are more movements left, if not, Player I wins. However, once Player I and Player II have made a move, in the next move, if it is still allowed, both players must pick points grater or equal (in the order of their respective models) than the elements picked in the last move. If Player II manages to successfully answer the n challenges proposed by Player I, then Player II wins. We define $\boldsymbol{M} \sim_n \boldsymbol{N}$ if and only if Player II has a winning strategy, and $\boldsymbol{M} \leqslant_n \boldsymbol{N}$ if and only if Player II has a winning strategy in the modified game with the additional restriction that Player I must pick always an element of $\boldsymbol{M}$. Observe, for instance, that in Example 3.58, Player II can successfully answer any one-movement challenge, which means that $\boldsymbol{N} \sim_1 \boldsymbol{M}$, while it can not successfully answer the two-movement challenge when Player I plays d first and then e, which shows that $\boldsymbol{N} \not\leqslant_2 \boldsymbol{M}$ and $\boldsymbol{N} \not\sim_2 \boldsymbol{M}$. The game-theoretical definitions of $\sim_n$ and $\leqslant_n$ are studied in detail by Ghilardi in [7].

Lemma 3.60. *Let $X \subseteq Var$. For every $n \in \omega$ and all $\boldsymbol{M}, \boldsymbol{N} \in \mathcal{M}_X$,*

- *if $\boldsymbol{M} \sim_n \boldsymbol{N}$ then $\boldsymbol{M} \sim_m \boldsymbol{N}$ for all $m \leqslant n$.*
- *if $\boldsymbol{M} \leqslant_n \boldsymbol{N}$ then $\boldsymbol{M} \leqslant_m \boldsymbol{N}$ for all $m \leqslant n$.*

Proof. We prove only the first statement, since the prove of the second is similar. If $n = 0$ the result is trivial. For $n \geqslant 1$, it is enough to prove that for all $\boldsymbol{M}, \boldsymbol{N} \in \mathcal{M}_X$, if $\boldsymbol{M} \sim_n \boldsymbol{N}$, then $\boldsymbol{M} \sim_{n-1} \boldsymbol{N}$. We prove this last statement by induction on n.

For $n = 1$, assume $\boldsymbol{M} \sim_1 \boldsymbol{N}$. We want $\boldsymbol{M} \sim_0 \boldsymbol{N}$. Since $\boldsymbol{M} \sim_1 \boldsymbol{N}$, there is a $q \in N$ such that $\boldsymbol{M}_{r_M} \sim_0 \boldsymbol{N}_q$. Thus, $v_M(r_M) = v_N(q)$. But $v_N(r_N) \subseteq v_N(q)$, by the truth-preserving condition. Then, $v_N(r_N) \subseteq v_M(r_M)$. A similar argument shows that $v_M(r_M) \subseteq v_N(r_N)$. Thus, $v_N(r_N) = v_M(r_M)$, and therefore $\boldsymbol{M} \sim_0 \boldsymbol{N}$.

For the inductive step, assume that the result holds for $n \geqslant 1$ and suppose $\boldsymbol{M} \sim_{n+1} \boldsymbol{N}$. We want $\boldsymbol{M} \sim_n \boldsymbol{N}$. So let $p \in M$. Since $\boldsymbol{M} \sim_{n+1} \boldsymbol{N}$, there exist an $q \in N$ such that $\boldsymbol{M}_p \sim_n \boldsymbol{N}_q$. And by the induction hypothesis $\boldsymbol{M}_p \sim_{n-1} \boldsymbol{N}_q$. Thus, $\boldsymbol{M} \leqslant_n \boldsymbol{N}$. A similar argument shows that $\boldsymbol{N} \leqslant_n \boldsymbol{M}$. Therefore, $\boldsymbol{M} \sim_n \boldsymbol{N}$. ⊠

For the next result, we fix some basic notation. If $\boldsymbol{M} \in \mathcal{M}_X$, by $[\boldsymbol{M}]_n$ we will denote the equivalence class of $\boldsymbol{M}$ under the relation $\sim_n$, that is, $[\boldsymbol{M}]_n = \{\boldsymbol{N} \in \mathcal{M}_X : \boldsymbol{M} \sim_n \boldsymbol{N}\}$. With $\mathcal{M}_X/_{\sim_n}$ we will denote the set of all equivalence classes of $\sim_n$, that is, $\mathcal{M}_X/_{\sim_n} = \{[\boldsymbol{M}]_n : \boldsymbol{M} \in \mathcal{M}_X\}$. Moreover, we will use $\|\boldsymbol{M}\|_n$ for the set of all the equivalence classes of generated submodels of $\boldsymbol{M}$, that is, $\|\boldsymbol{M}\|_n = \{[\boldsymbol{M}_p]_n \in \mathcal{M}_X/_{\sim_n} : p \in M\}$. Obviously, we have that $\boldsymbol{M} \sim_n \boldsymbol{N}$ if and only if $[\boldsymbol{M}]_n = [\boldsymbol{N}]_n$. We have the following characterization of $\sim_{n+1}$.

Lemma 3.61. *Let $X \subseteq Var$. For all $\boldsymbol{M}, \boldsymbol{N} \in \mathcal{M}_X$, $\|\boldsymbol{M}\|_n = \|\boldsymbol{N}\|_n$ if and only if $\boldsymbol{M} \sim_{n+1} \boldsymbol{N}$.*

Proof. For the left-to-right implication, assume $\|\boldsymbol{M}\|_n = \|\boldsymbol{N}\|_n$. To prove $\boldsymbol{M} \sim_{n+1} \boldsymbol{N}$ we will show that $\boldsymbol{M} \leqslant_{n+1} \boldsymbol{N}$ and $\boldsymbol{N} \leqslant_{n+1} \boldsymbol{M}$. Let $p \in M$. Since $\|\boldsymbol{M}\|_n = \|\boldsymbol{N}\|_n$, there is $q \in N$ such that $[\boldsymbol{M}_p]_n = [\boldsymbol{N}_q]_n$. Thus $\boldsymbol{M}_p \sim_n \boldsymbol{N}_q$. This shows $\boldsymbol{M} \leqslant_{n+1} \boldsymbol{N}$. Using a similar argument we get $\boldsymbol{N} \leqslant_{n+1} \boldsymbol{M}$. Therefore, $\boldsymbol{M} \sim_{n+1} \boldsymbol{N}$. For the other implication, assume $\boldsymbol{M} \sim_{n+1} \boldsymbol{N}$. To prove $\|\boldsymbol{M}\|_n = \|\boldsymbol{N}\|_n$, consider first the case where $[\boldsymbol{M}_p]_n \in \|\boldsymbol{M}\|_n$. As $\boldsymbol{M} \sim_{n+1} \boldsymbol{N}$, there is $q \in N$ such that $\boldsymbol{M}_p \sim_n \boldsymbol{N}_q$. Thus, $[\boldsymbol{M}_p]_n = [\boldsymbol{N}_q]_n$. As $[\boldsymbol{N}_q]_n \in \|\boldsymbol{N}\|_n$, $[\boldsymbol{M}_p]_n \in \|\boldsymbol{N}\|_n$, and we conclude $\|\boldsymbol{M}\|_n \subseteq \|\boldsymbol{N}\|_n$. The other inclusion is proven in the same fashion, concluding $\|\boldsymbol{M}\|_n = \|\boldsymbol{N}\|_n$. ⊠

Lemma 3.62. *Let $X \subseteq Var$. If X is finite, then $\mathcal{M}_X/_{\sim_n}$ is finite for all $n \in \omega$.*

Proof. We proceed by induction on n. For $n = 0$, the result is clear since there are as many elements in $\mathcal{M}_X/_{\sim_0}$ as elements in $\mathcal{P}(X)$. Now assume that $\mathcal{M}_X/_{\sim_n}$ is finite. Since this is so, to prove that $\mathcal{M}_X/_{\sim_{n+1}}$ is also finite it is enough to show that $|\mathcal{M}_X/_{\sim_{n+1}}| \leqslant |\mathcal{P}(\mathcal{M}_X/_{\sim_n})|$. Suppose by contradiction that $|\mathcal{M}_X/_{\sim_{n+1}}| > |\mathcal{P}(\mathcal{M}_X/_{\sim_n})|$. Consider the application $\| - \|_n : \mathcal{M}_X/_{\sim_{n+1}} \to \mathcal{P}(\mathcal{M}_X/_{\sim_n})$ that sends a class $[\boldsymbol{M}]_{n+1}$ to $\|\boldsymbol{M}\|_n$. By lemma 3.61, $\| - \|_n$ is well defined. Furthermore, as $|\mathcal{M}_X/_{\sim_{n+1}}| > |\mathcal{P}(\mathcal{M}_X/_{\sim_n})|$, the application $\| - \|_n$ is not injective. This means that there are classes $[\boldsymbol{M}]_{n+1}, [\boldsymbol{N}]_{n+1} \in \mathcal{M}_X/_{\sim_{n+1}}$, $[\boldsymbol{M}]_{n+1} \neq [\boldsymbol{N}]_{n+1}$, such that $\|\boldsymbol{M}\|_n = \|\boldsymbol{N}\|_n$. But, by Lemma 3.61, this implies $\boldsymbol{M} \sim_{n+1} \boldsymbol{N}$, contradicting $[\boldsymbol{M}]_{n+1} \neq [\boldsymbol{N}]_{n+1}$. Therefore, $|\mathcal{M}_X/_{\sim_{n+1}}| \leqslant |\mathcal{P}(\mathcal{M}_X/_{\sim_n})|$. ⊠

The following result establishes the relationship between the implicational complexity and the n-equivalence. The relation $\leqslant_n$ can be captured by some formula of complexity n.

Lemma 3.63. *Let $X \subseteq Var$. For every $n \in \omega$ and every $\boldsymbol{N} \in \mathcal{M}_X$, there is a formula φ^n_N such that:*

(i) $c(\varphi^n_N) = n$;

(ii) *for all* $\boldsymbol{N}' \in \mathcal{M}_X$, $\boldsymbol{N}' \models \varphi^n_N$ *if and only if* $\boldsymbol{N}' \leqslant_n \boldsymbol{N}$;

(iii) *for all* $\boldsymbol{M} \in \mathcal{M}_X$, *if* $\boldsymbol{M} \sim_n \boldsymbol{N}$, *then* $\varphi^n_M \equiv \varphi^n_N$.

Proof. We prove it by induction on n. For $n = 0$, let $\boldsymbol{N} \in \mathcal{M}_X$ and take $\varphi^0_N = \bigwedge v_N(r_N)$. It is clear that φ^0_N satisfies the desired conditions. For the inductive step, assume that the result holds for n. To prove the result for $n+1$, let $\boldsymbol{N} \in \mathcal{M}_X$ and let S_N be a set with one representative of each class in $\{[\boldsymbol{M}] \in \mathcal{M}_X/_{\sim_n} : \boldsymbol{M} \not\sim_n \boldsymbol{N}_q \text{ for all } q \in N\}$, and for each $\boldsymbol{S} \in S_N$, let T_S be a set with one representative of each class in $\{[\boldsymbol{M}] \in \mathcal{M}_X/_{\sim_n} : \boldsymbol{S} \not\leqslant_n \boldsymbol{M}\}$. S_N and each T_S are finite, by Lemma 3.62. We define

$$\varphi^{n+1}_N = \bigwedge_{\boldsymbol{S} \in S_N} (\varphi^n_S \to \bigvee_{\boldsymbol{T} \in T_S} \varphi^n_T)$$

By the induction hypothesis, formulas φ^n_S and φ^n_T exist and satisfy conditions (i), (ii) and (iii). Conditions (i) and (iii) for φ^{n+1}_N are easy consequences of conditions (i) and (iii) for formulas φ^n_S and φ^n_T. It only remains to prove condition (ii).

For the left-to-right direction, let $\boldsymbol{N}' \in \mathcal{M}_X$ be such that $\boldsymbol{N}' \models \varphi^{n+1}_N$. Assume towards a contradiction that $\boldsymbol{N}' \not\leqslant_{n+1} \boldsymbol{N}$. Then, there exists a $p \in N'$ such that $\boldsymbol{N}'_p \not\sim_n \boldsymbol{N}_q$ for all $q \in N$. Thus, $[\boldsymbol{N}'_p] \in \{[\boldsymbol{M}] \in \mathcal{M}_X/_{\sim_n} : \boldsymbol{M} \not\sim_n \boldsymbol{N}_q \text{ for all } q \in N\}$. Hence, there is an $\boldsymbol{S} \in S_N$ such that $\boldsymbol{N}'_p \sim_n \boldsymbol{S}$. In particular, $\boldsymbol{N}'_p \leqslant_n \boldsymbol{S}$. By the induction hypothesis, $\boldsymbol{N}'_p \models \varphi^n_S$. Equivalently, $\boldsymbol{N}', p \models \varphi^n_S$. As $\boldsymbol{N}' \models \varphi^{n+1}_N$ and $\boldsymbol{N}', p \models \varphi^n_S$, we get $\boldsymbol{N}', p \models \varphi^n_T$ for some $\boldsymbol{T} \in T_S$. Thus, $\boldsymbol{N}'_p \models \varphi^n_T$. By the induction hypothesis, $\boldsymbol{N}'_p \leqslant_n \boldsymbol{T}$. As, moreover, $\boldsymbol{N}'_p \sim_n \boldsymbol{S}$, we have $\boldsymbol{S} \leqslant_n \boldsymbol{T}$, by Remark 3.56. But $\boldsymbol{S} \not\leqslant_n \boldsymbol{T}$, for $\boldsymbol{T} \in T_S$, obtaining a contradiction. Therefore, $\boldsymbol{N}' \leqslant_{n+1} \boldsymbol{N}$.

For the right-to-left direction, let $\boldsymbol{N}' \in \mathcal{M}_X$ be such that $\boldsymbol{N}' \leqslant_{n+1} \boldsymbol{N}$. To show that $\boldsymbol{N}' \models \varphi^{n+1}_N$, we will prove that $\boldsymbol{N}' \models \varphi^n_S \to \bigvee_{\boldsymbol{T} \in T_S} \varphi^n_T$ for all $\boldsymbol{S} \in S_N$. So let $\boldsymbol{S} \in S_N$ and $p \in N'$ be such that $\boldsymbol{N}', p \models \varphi^n_S$. By the induction hypothesis, $\boldsymbol{N}'_p \leqslant_n \boldsymbol{S}$. Moreover, since $\boldsymbol{N}' \leqslant_{n+1} \boldsymbol{N}$, there is a $q \in N$ such that $\boldsymbol{N}'_p \sim_n \boldsymbol{N}_q$. We have then $\boldsymbol{N}_q \sim_n \boldsymbol{N}'_p \leqslant_n \boldsymbol{S}$. Thus, $\boldsymbol{N}_q \leqslant_n \boldsymbol{S}$. But $\boldsymbol{N}_q \not\sim_n \boldsymbol{S}$, for $\boldsymbol{S} \in S_N$. Remark 3.56 implies $\boldsymbol{S} \not\leqslant_n \boldsymbol{N}_q$. As a consequence, there is a $\boldsymbol{T} \in T_S$ such that $\boldsymbol{N}_q \sim_n \boldsymbol{T}$. Hence, since $\boldsymbol{N}'_p \sim_n \boldsymbol{N}_q \sim_n \boldsymbol{T}$ and $\boldsymbol{T} \in T_S$, we have $\boldsymbol{N}'_p \leqslant_n \boldsymbol{T}$ for some $\boldsymbol{T} \in T_S$. By the induction hypothesis, $\boldsymbol{N}'_p \models \varphi^n_T$ for some $\boldsymbol{T} \in T_S$. Equivalently, $\boldsymbol{N}', p \models \varphi^n_T$ for some $\boldsymbol{T} \in T_S$. Thus, $\boldsymbol{N}', p \models \bigvee_{\boldsymbol{T} \in T_S} \varphi^n_T$. And therefore $\boldsymbol{N}' \models \bigwedge_{\boldsymbol{S} \in S_N} (\varphi^n_S \to \bigvee_{\boldsymbol{T} \in T_S} \varphi^n_T)$, as we wanted. ⊠

Lemma 3.64. *Let* $X \subseteq Var$. *For all* $n \in \omega$ *and all* $\boldsymbol{M}, \boldsymbol{N} \in \mathcal{M}_X$, *the following equivalence holds:* $\boldsymbol{M} \leqslant_n \boldsymbol{N}$ *if and only if for all* $A \in F(X)$ *such that* $c(A) \leqslant n$, *if* $\boldsymbol{N} \models A$, *then* $\boldsymbol{M} \models A$.

Proof. We proceed by induction on n. For $n = 0$, assume first that $\boldsymbol{M} \leqslant_0 \boldsymbol{N}$ and let $A \in F(X)$ be such that $c(A) \leqslant 0$. Then, A is a propositional variable, $\top$ or $\bot$. The cases for $\top$ are $\bot$ trivial. For propositional variables, the result holds since $v_N(r_N) \subseteq v_M(r_M)$. This proves one direction. For the other direction, assume that for all $A \in F(X)$ such that $c(A) \leqslant 0$, if $\boldsymbol{N} \models A$, then $\boldsymbol{M} \models A$. In particular, if $\boldsymbol{N} \models x$, then $\boldsymbol{M} \models x$ for all $x \in X$. Thus, $v_N(r_N) \subseteq v_M(r_M)$, that is, $\boldsymbol{M} \leqslant_0 \boldsymbol{N}$.

For the inductive case, assume that the result holds for n and let us prove it for $n+1$. For the left-to-right direction assume $\boldsymbol{M} \leqslant_{n+1} \boldsymbol{N}$. We need to show that for all $A \in F(X)$ such that $c(A) \leqslant n+1$, if $\boldsymbol{N} \models A$, then $\boldsymbol{M} \models A$. Following Remark 3.51, we will reason by induction on the set C_{n+1}. As we pointed out in such remark, to prove that the property holds for every formula in C_{n+1}, it is enough to show that the property holds for every formula in C_n, for every formula of the form $B \to C$ with $B, C \in C_n$, and that the property is preserved when applying operations $\wedge$ and $\vee$. So let $A \in C_{n+1}$ be such that $\boldsymbol{N} \models A$.

- If $A \in C_n$, then $c(A) \leqslant n$. Furthermore, Lemma 3.60 implies $\boldsymbol{M} \leqslant_n \boldsymbol{N}$, for $\boldsymbol{M} \leqslant_{n+1} \boldsymbol{N}$. Then, the induction hypothesis implies $\boldsymbol{M} \models A$.

- If $A = B \to C$ for some $B, C \in C_n$, we know that $c(B) \leqslant n$ and $c(C) \leqslant n$. To show that $\boldsymbol{M} \models B \to C$, take $p \in M$ such that $\boldsymbol{M}, p \models B$ or, equivalently, $\boldsymbol{M}_p \models B$. As $\boldsymbol{M} \leqslant_{n+1} \boldsymbol{N}$, there is a $q \in N$ such that $\boldsymbol{M}_p \sim_n \boldsymbol{N}_q$. In particular, $\boldsymbol{N}_q \leqslant_n \boldsymbol{M}_p$. As $c(B) \leqslant n$, the induction hypothesis implies $\boldsymbol{N}_q \models B$. Since $\boldsymbol{N} \models A$, $\boldsymbol{N}_q \models C$. But $\boldsymbol{M}_p \leqslant_n \boldsymbol{N}_q$ as well, for $\boldsymbol{M}_p \sim_n \boldsymbol{N}_q$, and $c(C) \leqslant n$. Hence, the induction hypothesis implies $\boldsymbol{M}_p \models C$, that is, $\boldsymbol{M}, p \models C$. Therefore, $\boldsymbol{M} \models B \to C$, as wanted.

- Now assume that $B, C \in C_{n+1}$ are such that if $\boldsymbol{N} \models B$, then $\boldsymbol{M} \models B$ and if $\boldsymbol{N} \models C$, then $\boldsymbol{M} \models C$. If $A = B \wedge C$, then, as $\boldsymbol{N} \models A$, also $\boldsymbol{N} \models B$ and $\boldsymbol{N} \models C$. By the induction hypothesis on B and C, $\boldsymbol{M} \models B$ and $\boldsymbol{M} \models C$. And therefore $\boldsymbol{M} \models A$, as we wanted. The case for $\vee$ is similar.

For the right-to-left direction, assume that for all $A \in F(X)$ such that $c(A) \leqslant n+1$, if $\boldsymbol{N} \models A$, then $\boldsymbol{M} \models A$. By Lemma 3.63, there is a formula φ_N^{n+1} of complexity $n+1$ such that for all $\boldsymbol{N}' \in \mathcal{M}_X$, $\boldsymbol{N}' \models \varphi_N^{n+1}$ if and only if $\boldsymbol{N}' \leqslant_{n+1} \boldsymbol{N}$. Obviously, $\boldsymbol{N} \models \varphi_N^{n+1}$, for $\boldsymbol{N} \leqslant_{n+1} \boldsymbol{N}$. Hence, by the hypothesis in this implication, $\boldsymbol{M} \models \varphi_N^{n+1}$, and therefore $\boldsymbol{M} \leqslant_{n+1} \boldsymbol{N}$, as desired. $\boxtimes$

Corollary 3.65. *For all $n \in \omega$ and all $\boldsymbol{M}, \boldsymbol{N} \in \mathcal{M}_X$, $\boldsymbol{M} \sim_n \boldsymbol{N}$ if and only if $\boldsymbol{M}$ and $\boldsymbol{N}$ satisfy the same formulas of complexity less or equal than n.*

Stable classes

Definition 3.66. Let $X \subset Var$ and $n \in \omega$. A class $\mathcal{K}$ of Kripke models on X is said to be $\leqslant_n$-*closed* if for each $\boldsymbol{M} \in \mathcal{K}$ and all $\boldsymbol{N} \in \mathcal{M}_X$ such that $\boldsymbol{N} \leqslant_n \boldsymbol{M}$, $\boldsymbol{N} \in \mathcal{K}$ holds.

Remark 3.67. The class $\mathcal{M}_X$ is clearly $\leqslant_n$-closed. Thus, given a class of models in X, the smallest $\leqslant_n$-closed class of models in X containing $\mathcal{K}$, denoted as $\mathcal{K}_{\downarrow_n}$, always exists. It can be easily checked that $\mathcal{K}_{\downarrow_n} = \bigcap\{\mathcal{L} \subseteq \mathcal{M}_X : \mathcal{K} \subseteq \mathcal{L}$ and $\mathcal{L}$ is $\leqslant_n$-closed$\} = \{\boldsymbol{N} \in \mathcal{M}_X : \boldsymbol{N} \leqslant_n \boldsymbol{M}$ for some $\boldsymbol{M} \in \mathcal{K}\}$. It is also clear that a class $\mathcal{K}$ is $\leqslant_n$-closed if and only if $\mathcal{K} = \mathcal{K}_{\downarrow_n}$.

Remark 3.68. In view of Remark 3.57, we know that $\mathcal{K}_{\downarrow_n}$ is closed under isomorphisms.

Definition 3.69. Let $X \subseteq Var$. A class $\mathcal{K}$ of Kripke models in X is said to be *stable* if for all $\boldsymbol{M} \in \mathcal{K}$ and every $p \in M$, $\boldsymbol{M}_p \in \mathcal{K}$ holds.

Remark 3.70. It is easy to check that if $\boldsymbol{M} \in \mathcal{K}_{\downarrow_n}$, then $\boldsymbol{M}_p \in \mathcal{K}_{\downarrow_n}$ for all $p \in M$. Thus, $\mathcal{K}_{\downarrow_n}$ is stable for all $n \in \omega$.

Lemma 3.71. *Let $X \subseteq Var$, $n \in \omega$ and $\mathcal{K}$ be a class of Kripke models on X. Then the following statements are equivalent:*

(i) *$\mathcal{K} = Mod_X(A)$ for some $A \in F(X)$ with $c(A) \leqslant n$;*

(ii) *$\mathcal{K}$ is $\leqslant_n$-closed.*

Proof. Assume first $\mathcal{K} = Mod_X(A)$ for some $A \in F(X)$ with $c(A) \leqslant n$, and let $\boldsymbol{M} \in \mathcal{K}$ and $\boldsymbol{N} \leqslant_n \boldsymbol{M}$. We want $\boldsymbol{N} \in \mathcal{K}$. As $\mathcal{K} = Mod_X(A)$, $\boldsymbol{M} \models A$. By the Lemma 3.64, $\boldsymbol{N} \models A$, and therefore $\boldsymbol{N} \in \mathcal{K}$, as wanted.

Now assume $\mathcal{K}$ is $\leqslant_n$-closed. We need to find a formula A of implicational complexity less or equal than n such that $\mathcal{K} = Mod_X(A)$. Let K be a set with one representative of each class in

$$\{[\boldsymbol{M}]_n \in \mathcal{M}_X/{\sim_n} : \boldsymbol{M} \in \mathcal{K} \text{ and there is no } \boldsymbol{N} \in \mathcal{K} \text{ such that } \boldsymbol{M} \leqslant_n \boldsymbol{N} \text{ and } \boldsymbol{N} \not\leqslant_n \boldsymbol{M}\}.$$

In other words, K contains exactly one representative for each $\leqslant_n$-maximal class of $\mathcal{K}$. Observe that $K \subseteq \mathcal{K}$, for $\mathcal{K}$ is $\leqslant_n$-closed. Moreover, notice that K is finite, for so is $\mathcal{M}_X/{\sim_n}$. Then, we can define $A = \bigvee_{\boldsymbol{M} \in K} \varphi^n_M$. The formula A has complexity n, and observe that

$$\begin{aligned}
\boldsymbol{N} \in Mod_X(A) &\iff \boldsymbol{N} \models A \\
&\iff \boldsymbol{N} \models \varphi^n_M \text{ for some } \boldsymbol{M} \in K \subseteq \mathcal{K} \\
&\iff \boldsymbol{N} \leqslant_n \boldsymbol{M} \text{ for some } \boldsymbol{M} \in \mathcal{K} && \text{(Lemma 3.63)} \\
&\iff \boldsymbol{N} \in \mathcal{K} && (\mathcal{K} \text{ is } \leqslant_n\text{-closed})
\end{aligned}$$

Hence, $\mathcal{K} = Mod_X(A)$, which concludes the proof. ⊠

The following theorem states that for stable clases the extension property is inherited from $\mathcal{K}$ to each $\mathcal{K}_{\downarrow n}$.

Lemma 3.72. *Let $X \subseteq Var$ and $\mathcal{K}$ be a stable class of Kripke models in X. If $\mathcal{K}$ has the extension property, then for all $n \in \omega$, the class $\mathcal{K}_{\downarrow n}$ has the extension property.*

Proof. First, we will prove the result for $n = 0$, and then for $n \neq 0$.

For $n = 0$, let $\boldsymbol{N}_1, \ldots, \boldsymbol{N}_l$ be in $\mathcal{K}_{\downarrow 0}$. We need to find a variant of $(\sum_{1 \leqslant i \leqslant l} \boldsymbol{N}_i)$ in $\mathcal{K}_{\downarrow 0}$. But

$$\begin{aligned}
\mathcal{K}_{\downarrow 0} &= \{\boldsymbol{N} \in \mathcal{M}_X : \boldsymbol{N} \leqslant_0 \boldsymbol{M} \text{ for some } \boldsymbol{M} \in \mathcal{K}\} \\
&= \{\boldsymbol{N} \in \mathcal{M}_X : v_M(r_M) \subseteq v_N(r_N) \text{ for some } \boldsymbol{M} \in \mathcal{K}\}
\end{aligned}$$

Thus, for each $\boldsymbol{N}_i$ there is a $\boldsymbol{M}_i \in \mathcal{K}$ such that $v_{M_i}(r_{M_i}) \subseteq v_{N_i}(r_{N_i})$. Since $\mathcal{K}$ has the extension property, there is a variant $\boldsymbol{M}$ of $(\sum_{1 \leqslant i \leqslant l} \boldsymbol{M}_i)$ in $\mathcal{K}$. By the truth-preserving condition, we know that $v_M(r_M) \subseteq v_M(r_{M_i})$ for each $1 \leqslant i \leqslant l$. Consider the variant $\boldsymbol{N}$ of $(\sum_{1 \leqslant i \leqslant l} \boldsymbol{N}_i)$ such that $v_N(r_N) = v_M(r_M)$, which is well defined since $v_N(r_N) = v_M(r_M) \subseteq v_M(r_{M_i}) = v_{M_i}(r_{M_i}) \subseteq v_{N_i}(r_{N_i}) = v_N(r_{N_i})$ for each $1 \leqslant i \leqslant l$. Since $v_N(r_N) = v_M(r_M)$ and $\boldsymbol{M} \in \mathcal{K}$, we conclude $\boldsymbol{N} \in \mathcal{K}_{\downarrow 0}$. Therefore, $\boldsymbol{N}$ is the variant we needed.

For $n \neq 0$, let $\boldsymbol{N}_1, \ldots, \boldsymbol{N}_l$ be in $\mathcal{K}_{\downarrow n}$. In view of Remarks 3.48 and 3.68, we may assume that $\boldsymbol{N}_1, \ldots, \boldsymbol{N}_l$ are pairwise disjoint. It suffices to find a variant of $(\sum_{1 \leqslant i \leqslant l} \boldsymbol{N}_i)$ in $\mathcal{K}_{\downarrow n}$. As $\mathcal{K}_{\downarrow n} = \{\boldsymbol{N} \in \mathcal{M}_X : \boldsymbol{N} \leqslant_n \boldsymbol{M} \text{ for some } \boldsymbol{M} \in \mathcal{K}\}$, we need a variant $\boldsymbol{N}$ of $(\sum_{1 \leqslant i \leqslant l} \boldsymbol{N}_i)$ and a model $\boldsymbol{M} \in \mathcal{K}$ such that $\boldsymbol{N} \leqslant_n \boldsymbol{M}$. Consider a list of disjoint models $\boldsymbol{N}'_1, \ldots, \boldsymbol{N}'_m$ consisting of $\boldsymbol{N}_1 \ldots, \boldsymbol{N}_l$ at the beginning and isomorphic copies of all their generated submodels afterwards. By Remarks 3.70 and 3.57, we know that all $N'_1, \ldots, N'_m$ are in $\mathcal{K}_{\downarrow n}$. Thus, for each $1 \leqslant j \leqslant m$, there is a model $M_j \in \mathcal{K}$ such that $N'_j \leqslant_n M_j$. This means that for all $1 \leqslant j \leqslant m$, there is a $p_j \in M_j$ such that $N'_j \sim_{n-1} (M_j)_{p_j}$. For simplicity, we will use M'_j to denote $(M_j)_{p_j}$. Since $\mathcal{K}$ is stable and each $M_j \in \mathcal{K}$, every M'_j is also in $\mathcal{K}$, for M'_j is submodel of M_j. But $\mathcal{K}$ also has the extension property. Thus, there is a variant M of $(\sum_{\leqslant j \leqslant m} M'_j)$ in $\mathcal{K}$. Let N be the variant of $(\sum_{1 \leqslant i \leqslant l} N_i)$ such that $v_N(r_N) = v_M(r_M)$. First of all, we need to show that this valuation defines a Kripke model. To do this, we need to prove that $v_N(r_N) \subseteq v_N(r_{N_i})$ for all $1 \leqslant i \leqslant l$. But $N_i \sim_{n-1} M'_i$ for all $1 \leqslant i \leqslant l$, and then, by Lemma 3.60, $N_i \sim_0 M'_i$ for all $1 \leqslant i \leqslant l$, that is, $v_{N_i}(r_{N_i}) = v_{M'_i}(r_{M'_i})$. As $v_N(r_N) = v_M(r_M) \subseteq v_M(r_{M'_i}) = v_{M'_i}(r_{M'_i}) = v_{N_i}(r_{N_i}) = v_N(r_{N_i})$ for all $1 \leqslant i \leqslant l$, we obtain what we wanted.

We will prove now that $N \leqslant_n M$. Let $p \in N$.

- If $p \neq r_N$, then p is in N_i for some $1 \leqslant i \leqslant l$. Thus, the model $(N_i)_p$ is (isomorphic to) N'_j for some $l < j \leqslant m$. As $N'_j \sim_{n-1} M'_j$, we have $(N_i)_p \sim_{n-1} M'_j$. But $M'_j \cong M''_j$, where M''_j is the isomorphic copy of M'_j in M. By Remark 3.57, $(N_i)_p \sim_{n-1} M''_j$. And since $(N_i)_p = (N_{r_{N_i}})_p = N_p$ and $M''_j = (M''_j)_{r_{M''_j}} = M_{r_{M''_j}}$, we have $N_p \sim_{n-1} M_{r_{M''_j}}$.

- For $p = r_N$, we will show that $N_{r_N} \sim_{n-1} M_{r_M}$, that is, $N \sim_{n-1} M$. We prove it by showing that for all $0 \leqslant k \leqslant n-1$, $N \sim_k M$. We do it by induction on k.

  – For $k = 0$, the result is immediate, since we defined $v_N(r_N) = v_M(r_M)$.

  – Now assume that the result holds for $k-1$, that is, $N \sim_{k-1} M$. We want $N \sim_k M$. We may assume $n \neq 1$. Otherwise, the case $k = 0$ is enough to prove the result. Thus, $n > 2$.

    We first prove $N \leqslant_k M$. For each $p \in N$ we want a $q \in M$ such that $N_p \sim_{k-1} M_q$ for some $q \in M$. If $p = r_N$, by the induction hypothesis he have $N_{r_N} \sim_{k-1} M_{r_M}$. If $p \neq r_N$, p is in N_i for some $1 \leqslant i \leqslant l$ and $N_p = (N_i)_p \sim_{n-1} M'_j \cong M''_j = M_{r_{M''_j}}$ for some $l < j \leqslant m$, where M''_j is the isomorphic copy of M'_j in M. Therefore, $N_p \sim_{n-1} M_{r_{M''_j}}$. But $1 \leqslant k \leqslant n-1$. Hence $N_p \sim_{k-1} M_{p_j}$, by Lemma 3.60. This completes the proof of $N \leqslant_k M$.

    Now we will show $M \leqslant_k N$. So let $p \in M$. If $p = r_M$, by the induction hypothesis he have $N_{r_N} \sim_{k-1} M_{r_M}$. If $p \neq r_M$, p is in the isomorphic copy M''_j (in M) of some M'_j, $1 \leqslant j \leqslant m$. But $M'_j \sim_{n-1} N'_j$, and then $M''_j \sim_{n-1} N'_j$. Thus, there is a $q_j \in N'_j$ such that $(M''_j)_p \sim_{n-2} (N'_j)_{q_j}$. Since $1 \leqslant k \leqslant n-1$, we have $0 \leqslant k-1 \leqslant n-2$. Hence, $(M''_j)_p \sim_{k-1} (N'_j)_{q_j}$. However, N'_j is either one among $N_1 \ldots, N_l$ or an isomorphic copy of a generated submodel of some N_i. In any case, there is an N_i, for some $1 \leqslant i \leqslant l$, and an element $q \in N_i$ such that $(M''_j)_p \sim_{k-1} (N_i)_q$. As $(M''_j)_p = M_p$ and $(N_i)_q = N_q$, we have $M_p \sim_{k-1} N_{q_j}$, as we wanted. This completes the proof of $M \leqslant_k N$. Therefore, $N \sim_k M$.

  We have proven that $N \sim_k M$ for all $0 \leqslant k \leqslant n-1$. In particular, we have $N \sim_{n-1} M$, that is $N_{r_N} \sim_{n-1} M_{r_M}$.

Since we have showed that for each $p \in N$ there is a $q \in M$ such that $\boldsymbol{N}_p \sim_{n-1} \boldsymbol{M}_q$, we have $\boldsymbol{N} \leqslant_n \boldsymbol{M}$, concluding the proof.

⊠

Bisimilar models

Definition 3.73. Let $X \subseteq Var$. Two models $\boldsymbol{M}, \boldsymbol{N} \in \mathcal{M}_X$ are said to be bisimilar, in symbols $\boldsymbol{M} \sim_\infty \boldsymbol{N}$, if $\boldsymbol{M} \sim_n \boldsymbol{N}$ for all $n \in \omega$.

Remark 3.74. In view of Corollary 3.65, two models $\boldsymbol{M}, \boldsymbol{N} \in \mathcal{M}_X$ are bisimilar if and only if they satisfy the same formulas. That is to say, bisimilar models are indistinguishable in **IPL**.

Lemma 3.75. *Let $X \subseteq Var$. Given $\boldsymbol{M}, \boldsymbol{N} \in \mathcal{M}_X$, we have:*

(i) $\boldsymbol{M} \sim_\infty \boldsymbol{N}$ *if and only if the following two conditions hold:*

  (a) For all $p \in M$ there is a $q \in N$ such that $\boldsymbol{M}_p \sim_\infty \boldsymbol{N}_q$;

  (b) For all $q \in N$ there is a $p \in M$ such that $\boldsymbol{M}_p \sim_\infty \boldsymbol{N}_q$.

(ii) $\boldsymbol{M} \sim_\infty \boldsymbol{N}$ *if and only the following three conditions hold:*

  (a) $\boldsymbol{M} \sim_0 \boldsymbol{N}$;

  (b) For all $p \in M$ there is a $q \in N$, $q \neq r_N$, such that $\boldsymbol{M}_p \sim_\infty \boldsymbol{N}_q$;

  (c) For all $q \in N$ there is a $p \in M$, $p \neq r_M$, such that $\boldsymbol{M}_p \sim_\infty \boldsymbol{N}_q$.

Proof.

(i) Assume $\boldsymbol{M} \sim_\infty \boldsymbol{N}$. We prove only (a), since (b) is analogous. Let $p \in M$ and assume by contradiction that there is no $q \in N$ such that $\boldsymbol{M}_p \sim_\infty \boldsymbol{N}_q$. It means that for each $q \in N$ there is an $n_q \in \omega$ such that $\boldsymbol{M}_p \not\sim_{n_q} \boldsymbol{N}_q$. Let m be the maximum of $\{n_q \in \omega : q \in N\}$, which exists because N is finite. It is clear that $\boldsymbol{M}_p \not\sim_m \boldsymbol{N}_q$ for each $q \in N$. But this implies $\boldsymbol{M} \not\sim_{m+1} \boldsymbol{N}$, contradicting $\boldsymbol{M} \not\sim_\infty \boldsymbol{N}$.

Now assume (a) and (b). Assume towards a contradiction that $\boldsymbol{M} \not\sim_\infty \boldsymbol{N}$. Hence, $\boldsymbol{M} \not\sim_{m+1} \boldsymbol{N}$ for some $m \in \omega$. This implies that there is a $p \in M$ such that $\boldsymbol{M}_p \not\sim_m \boldsymbol{N}_q$ for all $q \in N$ or that there is a $q \in N$ such that $\boldsymbol{M}_p \not\sim_m \boldsymbol{N}_q$ for all $p \in M$. In each case, we have a contradiction with (a) or (b), respectively.

(ii) The left-to-right direction is immediate from the definition of $\sim_\infty$ and (i). For the right-to-left direction assume (a), (b) and (c). We need to prove that for all $n \in \omega$, $\boldsymbol{M} \sim_n \boldsymbol{N}$. We prove it by induction on n. By (a), the result holds for $n = 0$. Now assume $\boldsymbol{M} \sim_n \boldsymbol{N}$. We want $\boldsymbol{M} \sim_{n+1} \boldsymbol{N}$. So let $p \in M$. If $p = r_M$, by the induction hypothesis we have $\boldsymbol{M}_{r_M} \sim_n \boldsymbol{N}_{r_N}$. If $p \neq r_N$, condition (b) implies that there is a $q \in N$ such that $\boldsymbol{M}_p \sim_n \boldsymbol{N}_q$. This shows $\boldsymbol{M} \leqslant_{n+1} \boldsymbol{N}$. A similar argument shows $\boldsymbol{N} \leqslant_{n+1} \boldsymbol{M}$. Hence, $\boldsymbol{M} \sim_{n+1} \boldsymbol{N}$. Therefore, $\boldsymbol{M} \sim_\infty \boldsymbol{N}$.

⊠

Main theorem

Lemma 3.76. *Let $X \subseteq Var$. If $A \in F(X)$ is a unifiable formula and $\sigma\colon F(X) \to F(Y)$ is a unifier of A, then there is formula $B \in F(X)$ such that:*

(i) $c(B) \leqslant c(A)$;

(ii) *B is projective;*

(iii) $B \models A$;

(iv) *σ is a unifier of B.*

Proof. Consider the class $\mathcal{K} = \{\boldsymbol{M} \in \mathcal{M}_X : \boldsymbol{M} \sim_\infty \boldsymbol{N}^\sigma \text{ for some } \boldsymbol{N} \in \mathcal{M}_Y\}$. The class $\mathcal{K}_{\downarrow_{c(A)}}$ is trivially $\leqslant_{c(A)}$-closed. Thus, by Lemma 3.71, we know that $\mathcal{K}_{\downarrow_{c(A)}} = Mod_X(B)$ for some formula $B \in F(X)$ such that $c(B) \leqslant c(A)$. We claim that B is the formula we need. It remains to be proven that B is projective, $B \models A$, and σ is a unifier of B.

To show that B is projective, we will use Theorem 3.49. By this theorem, it is enough to show that $Mod_X(B)$ has the extension property. To this end, we will use Lemma 3.72. This lemma implies that in order to prove that $Mod_X(B)$ (which is $\mathcal{K}_{\downarrow_{c(A)}}$) has the extension property it is enough to show that $\mathcal{K}$ is stable and has the extension property.

To show that $\mathcal{K}$ is stable, let $\boldsymbol{M} \in \mathcal{K}$ and $p \in \boldsymbol{M}$. We need to prove that $\boldsymbol{M}_p \in \mathcal{K}$, that is, we need to find a model $\boldsymbol{N}$ in $\mathcal{M}_Y$ such that $\boldsymbol{M}_p \sim_\infty \boldsymbol{N}^\sigma$. But since $\boldsymbol{M} \in \mathcal{K}$ we know, by definition of $\mathcal{K}$, that there is a model $\boldsymbol{N}' \in \mathcal{M}_Y$ such that $\boldsymbol{M} \sim_\infty (\boldsymbol{N}')^\sigma$. By Lemma 3.75 (i), there is $q \in N'$ such that $\boldsymbol{M}_p \sim_\infty ((\boldsymbol{N}')^\sigma)_q$. Taking $\boldsymbol{N} = \boldsymbol{N}'_q$ and recalling that the operator $(-)^\sigma$ commutes with the submodel generation (Proposition 3.13), we are done.

To show that $\mathcal{K}$ has the extension property, let $\boldsymbol{M}_1, \boldsymbol{M}_2 \ldots, \boldsymbol{M}_n$ be in $\mathcal{K}$. We need to find a variant of $(\sum_{1 \leqslant i \leqslant n} \boldsymbol{M}_i)$ in $\mathcal{K}$. By definition of $\mathcal{K}$, for each $\boldsymbol{M}_i$ there is a $\boldsymbol{N}_i \in \mathcal{M}_Y$ such that $\boldsymbol{M}_i \sim_\infty (\boldsymbol{N}_i)^\sigma$. Consider the model $\boldsymbol{N} = (\sum_{1 \leqslant i \leqslant n} \boldsymbol{N}_i)$ and its image $\boldsymbol{N}^\sigma$ under the operator $(-)^\sigma$. Let $\boldsymbol{M}$ be the variant of $(\sum_{1 \leqslant i \leqslant n} \boldsymbol{M}_i)$ such that $v_M(r_M) = v_N^\sigma(r_N)$. To show that this valuation indeed defines a Kripke model we need to prove that $v_M(r_M) \subseteq v_M(r_{M_i})$. But $v_{M_i}(r_{M_i}) = v_{N_i}^\sigma(r_{N_i})$, for $\boldsymbol{M}_i \sim_\infty (\boldsymbol{N}_i)^\sigma$, and $v_N^\sigma(r_N) \subseteq v_N^\sigma(r_{N_i}) = v_{N_i}^\sigma(r_{N_i})$, for the operator $(-)^\sigma$ commutes with submodel generation. We claim that $\boldsymbol{M} \sim_\infty \boldsymbol{N}^\sigma$. We will show this using Lemma 3.75 (ii). Condition (a) of 3.75 is immediately satisfied by the definition of $\boldsymbol{M}$. Conditions (b) and (c) are also easily satisfied for $\boldsymbol{M}_i \sim_\infty (\boldsymbol{N}_i)^\sigma$ for all $1 \leqslant i \leqslant n$. Hence, $\boldsymbol{M}$ is the variant of $(\sum_{1 \leqslant i \leqslant n} \boldsymbol{M}_i)$ that we wanted.

Thus, $\mathcal{K}$ is stable and has the extension property. Using Theorem 3.49 and Lemma 3.72, this implies that B is projective.

We now need to prove that $B \models A$. We will do it by showing $Mod_X(B) \subseteq Mod_X(A)$. Recall that $\mathcal{K}_{\downarrow_{c(A)}} = Mod_X(B)$ and $\mathcal{K}_{\downarrow_{c(A)}}$ is the smallest $\leqslant_{c(A)}$-closed class containing $\mathcal{K}$. Hence, to obtain the desired inclusion it is enough to show that $Mod_X(A)$ is $\leqslant_{c(A)}$-closed and contains $\mathcal{K}$.

We first prove that $Mod_X(A)$ is $\leqslant_{c(A)}$-closed. Let $\boldsymbol{N} \in Mod_X(A)$ and $\boldsymbol{M} \in \mathcal{M}_X$ be such that $\boldsymbol{M} \leqslant_{c(A)} \boldsymbol{N}$. Since $\boldsymbol{N} \models A$, Lemma 3.64 implies $\boldsymbol{M} \models A$, and we are done.

Now we proceed to prove that $Mod_X(A)$ contains $\mathcal{K}$. Let $\boldsymbol{M} \in \mathcal{K}$. We know $\boldsymbol{M} \sim_\infty \boldsymbol{N}^\sigma$ for some $\boldsymbol{N} \in \mathcal{M}_Y$. In particular, $\boldsymbol{M} \sim_{c(A)} \boldsymbol{N}^\sigma$. Corollary 3.65 implies $\boldsymbol{M} \models A$ if and only if $\boldsymbol{N}^\sigma \models A$. But σ is a unifier of A. Thus, $\boldsymbol{N}^\sigma \models A$. Therefore $\boldsymbol{M} \models A$, that is, $\boldsymbol{M} \in Mod_X(A)$, and we have the inclusion $\mathcal{K} \subseteq Mod_X(A)$.

It only remains to be proven that σ is a unifier of B, but this follows from the inclusions $(\mathcal{M}_Y)^\sigma \subseteq \mathcal{K} \subseteq \mathcal{K}_{\downarrow_{c(A)}} = Mod_X(B)$. $\boxtimes$

Finally, we can conclude this section by proving the Unification Theorem for **IPL**.

Theorem 3.77. *The unification type of* **IPL** *is finitary.*

Proof. In Example 3.34 we already showed that there is a unifiable formula not admitting a basis of unifiers of size 1. Now we will show that every unifiable formula has a finite basis of unifiers. So let A be a unifiable formula. Take $X \subseteq Var$ finite such that $A \in F(X)$. Let S_A be a set of representatives of the set of classes $C^X_{c(A)}/_\equiv$ (recall that $C^X_{c(A)}$ is the set of all formulas in $F(X)$ of implicational complexity less or equal than $c(A)$). Now take $B_A = \{B \in S_A : B \text{ is projective and } B \models A\}$. By Lemma 3.76, B_A is non-empty, as A is unifiable. For each $B \in B_A$, pick a unifier σ_B such that B is projective with σ_B, and consider the set $\Sigma_A = \{\sigma_B : B \in B_A\}$. By Lemma 3.53, S_A, B_A and Σ_A are all finite. We claim Σ_A is a basis of unifiers of A.

First, we need to show that substitutions in Σ_A are indeed unifiers of A. So let $B \in B_A$ and $\sigma_B \in \Sigma_A$. Since $B \models A$, $\sigma_B(B) \models \sigma_B(A)$. Moreover, σ_B is a unifier of B, for B is projective with σ_B. Therefore, $\models \sigma_B(A)$, and we conclude that σ_B is a unifier of A as well.

Now we need to show that every unifier of A is less general than some substitution in Σ_A. So let σ be a unifier of A. By Lemma 3.76, there is formula $B' \in F(X)$ such that B' is projective, $B' \models A$, $c(B') \leqslant c(A)$ and σ is a unifier of B'. It is easy to see that $B' \equiv B$ for some $B \in B_A$. Clearly, σ is a unifier of B as well. Furthermore, by Remark 3.25, we know that σ_B is a most general unifier of B. Hence $\sigma \preceq \sigma_B$, as we wanted.

Therefore, Σ_A is a finite basis of unifiers of A, which concludes the proof. $\boxtimes$

Bibliography

[1] P. Balbiani. Unification in modal logic. In Md. Aquil Khan and Amaldev Manuel, editors, *Logic and Its Applications*, pages 1–5, Berlin, Heidelberg, 2019. Springer Berlin Heidelberg.

[2] P. Balbiani, C. Gencer, M. Mojtahedi, M. Rostamigiv, and T. Tinchev. A gentle introduction to unification in modal logics. In *13èmes Journées d'Intelligence Artificielle Fondamentale (JIAF 2019)*, Toulouse, France, July 2019.

[3] A. V. Chagrov and M. Zakharyaschev. Modal logic. In *Oxford logic guides*, 1997.

[4] W. Dzik, S. Kost, and P. Wojtylak. Unification types and union splittings in intermediate logics, 2022.

[5] H.B. Enderton. *A Mathematical Introduction to Logic*. Elsevier Science, 2001.

[6] S. Ghilardi. Unification in intuitionistic logic. *Journal of Symbolic Logic*, 64(2):859–880, 1999.

[7] S. Ghilardi and M. Zawadowski. *Sheaves, Games, and Model Completions: A Categorical Approach to Nonclassical Propositional Logics*. Trends in Logic. Springer Netherlands, 2013.

[8] R. Iemhoff. On the admissible rules of intuitionistic propositional logic. *The Journal of Symbolic Logic*, 66(1):281–294, 2001.

[9] R. Iemhoff. Intermediate Logics and Visser's Rules. *Notre Dame Journal of Formal Logic*, 46(1):65 – 81, 2005.

[10] E. Jeřábek. Blending margins: the modal logic K has nullary unification type. *Journal of Logic and Computation*, 25(5):1231–1240, 10 2013.

[11] L. Löwenheim. Über das auflösungsproblem im logischen klassenkalkül. *Sitzungsberichte der Berliner mathematischen Gesellschaft*, (7):89–94, 1908.

[12] V. Marra and L. Spada. Duality, projectivity, and unification in Łukasiewicz logic and mv-algebras. *Annals of Pure and Applied Logic*, 164(3):192–210, 2013.

[13] T. Moraschini. The algebra of logic. *Course notes*, 2023.