UNIVERSITY OF LIVERPOOL

On The Relative Succinctness of Some Modal Logics

Thesis submitted in accordance with the requirements of
the University of Liverpool for the degree of Doctor in Philosophy by

Petar Vanev Iliev

November 2013

Contents

Illustrations

List of Figures

List of Tables

Preface

This thesis is primarily my own work. The sources of other materials are identified. The main technical contributions can be found in Chapter 4 and Chapter 5 which are based on the following publications: [21] presented at the Twenty Second International Joint Conference on Artificial Intelligence (IJCAI 2011), [22] from Artificial Intelligence Journal, [61] presented at the Turing Centenary Conference and Eighth Conference on Computability in Europe (CiE 2012). Some of the results in Chapter 4 are from an unpublished joint paper with Wiebe van der Hoek entitled *Public Announcements are Exponentially More Succinct than Everybody Knows and Somebody Knows* which was presented at the Fifth Workshop on Logical Aspects of Multi-Agent Systems (LAMAS 2012).

Abstract

The aim of this thesis is to compare several extensions of multimodal logic in terms of their representational succinctness on different classes of models. Succinctness is a natural refinement on the notion of expressivity. Intuitively, given two logics L_1 and L_2, we say that L_1 expresses more succinctly than L_2 some properties of a class of models if the L_1-formulae expressing the properties in question are significantly shorter than all the equivalent L_2-formulae. The precise technical interpretation of "significantly shorter" depends on the case at hand and may mean "exponentially shorter", "non-elementary shorter", etc. This work was motivated by the question of whether public announcement logic (PAL) is exponentially more succinct than multimodal logic (ML) on the class $\mathbf{S}_5$ of Kripke models with underlying structures in which all relations are reflexive, symmetric, and transitive.

Using techniques based on a generalisation of Ehrenfeucht-Fraïssé games called Adler-Immerman games, we show that extending ML in two different ways: by allowing formulae $[\cup_\Gamma]\varphi$, where $[\cup_\Gamma]\varphi$ is equivalent to $\bigwedge_{i\in\Gamma}[i]\varphi$, and by introducing formulae $[\exists_\Gamma]\varphi$ such that $[\exists_\Gamma]\varphi$ is equivalent to $\bigvee_{i\in\Gamma}[i]\varphi$, leads to two well-known equally expressive extensions of ML, called $[\cup]$ML and $[\exists]$ML, respectively, such that each one of them is exponentially more succinct than the other on the class of all Kripke models. We also compare $[\cup]$ML and $[\exists]$ML with PAL and prove that there are properties of Kripke models that can be expressed exponentially more succinctly in PAL than in both $[\cup]$ML and $[\exists]$ML. In order to show that even logics with different expressive power can be meaningfully compared in terms of their representational succinctness, we consider specific modal formulae containing modalities of the form $[a \cap b]$ and study the lengths of equivalent formulae from the logics $[\cup]$ML and $[\exists]$ML on some suitably chosen Kripke models. We obtain meaningful succinctness results even in this case.

As far as $\mathbf{S}_5$-models are concerned, we are able to answer in the affirmative our initial motivating question; moreover, we compare the logics $[\cup]$ML and $[\exists]$ML with ML and show that there is an exponential succinctness gap in this case, too. Again, we are able to exhibit formulae containing modalities $[a \cap b]$ that have ML-equivalents on a very restricted subset of $\mathbf{S}_5$-models but the length of these equivalent formulae is exponentially bigger.

Acknowledgements

First, and foremost, I would like to thank my supervisors Wiebe van der Hoek and Michael Wooldridge for their scientific guidance, constant moral support, and seemingly unlimited patience.

I owe a huge debt of gratitude to my examiners Jerome Lang and Frank Wolter. Thanks to them, my viva turned out to be one of the most helpful and enlightening scientific meetings at which I have ever been. Their feedback and insights improved this thesis considerably.

I am extremely grateful to Tim French whose ideas and suggestions played a major role in the initial stages of the work presented here.

I am eternally grateful to my early mentors and teachers Dimiter Vakarelov and Ivan Soskov. Dimiter introduced me to modal logic and guided me through my first research attempts with patience and kindness. Ivan taught me recursion theory and offered both moral and administrative help on many occasions. His untimely death was a huge blow to all of us who knew him.

I am convinced that my days in Liverpool would not have been so enjoyable if I had not met Dirk Walther, Michel Ludwig, Charlotte Kappel, Chris Everritt, John Fearnley, Thomas Varghese, Savas Konur, Yoshihito Tanaka, and Sven Schewe. Thank you guys for all the laughs.

Finally, to my family and Laska, thank you for everything.

Chapter 1

Introduction

The study of the expressive power of logics is one of the major topics in mathematical logic and theoretical computer science. The general framework for such investigations can be described as follows. We begin with the question of whether a particular formalism can express some property on some class of models or not. The intuitive notion of property is given a formal expression through the concept of query and, therefore, the formal version of our initial question is whether a particular query is definable in some logic under investigation. Such questions are of great theoretical interest but work on problems in knowledge representation, database theory, automated verification and model checking has shown that the study of expressivity should be more refined as suggested by the next examples.

It has been argued in [25] that, at least as far as knowledge representation formalisms are concerned, the comparison of two such formalisms, L_1 and L_2, cannot be meaningfully accomplished just in terms of expressive power or the computational complexity of their inference problems. This is due to the fact that often we have the following situation:

1. L_1 and L_2 are equally expressive, and/or

2. L_1 and L_2 have the same complexity of the satisfiability problem, or

3. the complexities of L_1 and L_2 are different but both are so high that it cannot be honestly claimed to be of any practical relevance.

Therefore, the authors of [25] suggest that a better comparison criterion is the representational succinctness of such formalisms. Intuitively, if we are interested in some particular query Q that is expressible with formulae φ_1 and φ_2 from L_1 and L_2 respectively, we should ask if there is a significant difference in the lengths of φ_1 and φ_2. Hence, the notion of succinctness is a refinement of the notion of expressivity.

Similarly, it is pointed out in [26] that the expressive power of node-selecting query languages for XML-documents studied in database theory is not the main issue. However, the interplay between the time in which queries can be evaluated and their succinctness becomes very important.

An analogous situation is found in the study of the relationship between different fragments of first order logic and versions of temporal logic or between different temporal logics where the issue of representational succinctness becomes very important (see, for example, [66], [17], [45])

Aims and Objectives Our starting point is multimodal propositional logic (ML), which extends classical propositional logic with an indexed set of unary modal operators $[i]$, where each such operator is interpreted on a Kripke accessibility relations R_i. We aim to study the relative succinctness of four well-known extensions of ML on different classes of models. These extensions are:

- $[\cup]$ML which adds to ML formulae $[\cup_\Gamma]\varphi$ for every finite set of relation indices Γ, where $[\cup_\Gamma]\varphi$ is equivalent to $\bigwedge_{i\in\Gamma}[i]\varphi$;
- $[\exists]$ML which extends ML with formulae $[\exists_\Gamma]\varphi$, where $[\exists_\Gamma]\varphi$ is equivalent to $\bigvee_{i\in\Gamma}[i]\varphi$;
- $[\cap]$ML obtained by adding to ML formulae of the form $[\cap_\Gamma]\varphi$ that are interpreted in Kripke models as usual

$$(M, w) \models [\cap_\Gamma]\varphi \text{ if and only if } (M, v) \models \varphi \text{ for all } v \text{ such that } \bigwedge_{i\in\Gamma} wR_i v;$$

- $[\varphi]$ML, or more popularly known as public announcement logic (PAL) (see, for example, [50]), is an extension of ML with formulae $[\varphi]\psi$ interpreted in Kripke models as follows

$$(M, w) \models [\varphi]\psi \text{ if and only if } (M, w) \models \varphi \text{ implies } (M|_\varphi, w) \models \psi,$$

  where $M|_\varphi$ is the restriction of the model M to the points that satisfy φ.

$[\cup_\Gamma]$ is the familiar from Boolean modal logic ([24]) and propositional dynamic logic ([29]) modality $[i \cup j \cup \ldots \cup m]$ where $\Gamma = \{i, j, \ldots m\}$. In epistemic logic ([62],[18]), formulae of the form $[\cup_\Gamma]\varphi$ are used to express the notion "everybody in the group of agents Γ knows that φ is true". In the same way, a formula $[\exists_\Gamma]\varphi$ captures the intuitive notion "somebody in the group of agents Γ knows that φ" while formulae $[\varphi]\psi$ are a natural way of formalizing the idea "ψ is true after the public announcement of φ". Similarly, allowing modalities of the form $[\cup_\Gamma]$ in ML corresponds to adding role disjunctions to the description logic $\mathcal{ALC}$ ([4]). Introducing these modal operators does not lead to an increase of expressive power but we will show that they increase the succinctness of the resulting system exponentially.

Although modalities $[\cap_\Gamma]$ also occur naturally in Boolean modal logic, propositional dynamic logic, and epistemic logic where they are used as a formalisation of the notion of distributed knowledge among the agents in the group Γ, it is well-known that they cannot be expressed in ML. However, we are going to show that there are formulae containing

$[\cap_\Gamma]$ modalities that have equivalent, on some simple classes of Kripke models, ML-formulae whose length is exponentially bigger and this cannot be improved even if we allow the use of $[\cup_\Gamma]$ or $[\exists_\Gamma]$ operators.

The initial motivation for the study of these problems came from the open question in [43] whether $[\varphi]$ML is exponentially more succinct than ML on $\mathbf{S}_5$-models. We give a positive answer in Chapter 5.

Thesis Structure

Chapter 2 serves as a practical introduction to the type of questions and mathematical techniques with which we are dealing in the thesis. The first sub-section contains a slow-paced proof of the well-known fact that the shortest formula in disjunctive normal form that is equivalent to the formula $l_n \stackrel{\text{def}}{=} p_1 \oplus p_2 \oplus \ldots \oplus p_n$, where $\oplus$ denotes the "*exclusive or*" operator, is its canonical disjunctive normal form. The second sub-section is dedicated to a very short overview of a small number of well-known results and methods for proving lower bounds on the size of formulae of Boolean logic, first-order logic, modal logic and some more specialised logics for knowledge representation.

Chapter 3 consists of 3 sub-sections. In the first, we consider carefully the notion of succinctness and give a sufficient condition for proving that one logic is exponentially more succinct than another on a given class of models. The modal logics we study in the thesis are defined in the second sub-section. The third is dedicated to the definition of extended-syntax trees for modal logic formulae and related notions.

Chapter 4 contains our succinctness results on classes of models with no restrictions on the relations in the underlying Kripke structures.

Chapter 5 is dedicated to succinctness results on $\mathbf{S}_5$-models.

In Chapter 6, we summarise all the results we obtain in Chapter 4 and Chapter 5 and list some open problems.

The notational conventions adopted in this thesis are more or less standard. We use the usual abbreviation "iff" for "if and only if". The operator $\stackrel{\text{def}}{=}$ is used to designate a definition of the object on the left side; $\in$, $\subset$ and $\subseteq$ are used in the usual way, i.e. $x \in S$ means that x is an element of the set S, $A \subset B$ means that A is a strict subset of B, whereas $A \subseteq B$ means that $A \subset B$ or $A = B$. The sets of the non-negative integers $\{0, 1, 2, \ldots\}$ and the non-negative real numbers are denoted $\mathbb{N}$ and $\mathbb{R}^+$, respectively.

The formal precision of our exposition increases from one chapter to the next. Sometimes we repeat or make the same point using a slightly different perspective in order to achieve clarity, intuitive understanding and help the reader grasp easily the main points of an argument.

Chapter 2

Technical Introduction

The purpose of this chapter is to introduce the reader to the type of problems and mathematical techniques with which we are dealing in this thesis. In the first sub-section we give a slow-paced proof of the well-known result [42] that the shortest formula in disjunctive normal form that is equivalent to the formula $l_n \stackrel{\text{def}}{=} p_1 \oplus p_2 \oplus \ldots \oplus p_n$, where $\oplus$ denotes the "*exclusive or*" operator, is its canonical disjunctive normal form. Although, to the best of our knowledge, the argument we offer here is new, our main interest lies not so much in finding yet another proof of a classic theorem but in the introduction of the Adler-Immerman game [1] and extended syntax trees [27] on which the proof is based and which are our main technical tools for obtaining all the results in the next chapters. Our exposition is somewhat non-standard and not very precise because we aim at developing the reader's intuition in the simplest possible setting. The second sub-section is dedicated to a very short overview of a small number of well-known results and methods for proving lower bounds on the size of formulae of Boolean logic, first-order logic, modal logic and some more specialised logics for knowledge representation. Inevitably, the selection of problems and methods we offer is extremely limited and somewhat arbitrary.

2.1 An Example From Propositional Logic

One of the notoriously difficult open problems in Boolean function complexity ([15], [38], [47], [65]) is the following.

> Given a certain Boolean function of n variables $f(x_1 \ldots x_n)$, what is the shortest (smallest) Boolean formula computing $f(x_1 \ldots x_n)$?

The usual strategy when dealing with a seemingly intractable question is to try and solve a restricted version of it. Here, we will reprove one of the classic such results as a means of introducing some of the notions and types of arguments found in the next chapters. Since this thesis is dedicated to lower bounds on formula size in the setting of modal logic as an extension of propositional logic, we will not use the language of

Boolean function theory but that of propositional logic in order to ensure uniformity in notation and presentation.

We fix a countably infinite set of propositional symbols $P = \{\mathbf{1}, p_1, p_2, \ldots\}$ that contains the constant symbol $\mathbf{1}$. The formulae of Boolean logic over P are defined relative to a non-empty countable set $\mathbf{B} = \{op_1, \ldots, op_i, \ldots\}$ of Boolean operators called **basis** and a function $\alpha : \mathbf{B} \to \mathbb{N}$ assigning to each operator from $\mathbf{B}$ a natural number called **arity**. Traditionally, the most popular basis is the De Morgan basis $\mathbf{DM} = \{\neg, \vee, \wedge\}$ where $\neg$ has arity 1 whereas $\vee$ and $\wedge$ have arity 2. Another basis which will play a role in this section is $\mathbf{Mod2} = \{\wedge, \oplus\}$, where beside the operator $\wedge$, we have the operator $\oplus$ which has arity 2.

Definition 2.1 (Boolean formulae)**.** The set of Boolean formulae $\Phi(P, \mathbf{B})$ over the propositional symbols in P and a basis $\mathbf{B}$ is the smallest set of strings containing all propositional symbols and closed under the operators in $\mathbf{B}$, i.e.,

- $P \subset \Phi(P, \mathbf{B})$;
- for any $op \in \mathbf{B}$, if $\alpha(op) = n$ and $\varphi_1, \varphi_2, \ldots, \varphi_n$ belong to $\Phi(P, \mathbf{B})$, then so does $op(\varphi_1, \varphi_2, \ldots, \varphi_n)$.

Example 2.1. *In the case of the De Morgan basis* $\mathbf{DM}$*, we have the well-known definition below where the formulae are defined using the infix notation.*

- $P \subset \Phi(P, \mathbf{DM})$*;*
- *if* $\varphi \in \Phi(P, \mathbf{DM})$*, then* $\neg\varphi \in \Phi(P, \mathbf{DM})$*;*
- *if* φ *and* ψ *belong to* $\Phi(P, \mathbf{DM})$*, then so do* $(\varphi \vee \psi)$ *and* $(\varphi \wedge \psi)$*.*

The set $\Phi(P, \mathbf{Mod2})$ *is defined similarly.*

- $P \subset \Phi(P, \mathbf{Mod2})$*;*
- *if* φ *and* ψ *belong to* $\Phi(P, \mathbf{Mod2})$*, then so do* $(\varphi \wedge \psi)$ *and* $(\varphi \oplus \psi)$*.*

The length of a formula $\varphi \in \Phi(P, \mathbf{B})$, denoted $|\varphi|$ is defined recursively on the structure of φ.

$$|\mathbf{1}| = |p_i| = 1 \text{ for any } p_i \in P;$$
$$|op(\varphi_1, \ldots, \varphi_n)| = 1 + |\varphi_1| + |\varphi_2| + \ldots + |\varphi_n|.$$

Therefore, when defining the length of a formula φ, we do not take into consideration the number of parentheses and the length of indices of the propositional symbols occurring in φ. Another way of defining the length of φ would be to encode all propositional symbols and Boolean operators as binary strings and use this encoding to represent φ as a word over the alphabet $\{0, 1\}$. Then the length of φ is the length of its binary representation. The reader can find more on this in [23] (p. 8) and [26] but the main point that is made by the authors there is that the relations between the lower bounds we prove using our length measure remain valid for any reasonable algorithmic encoding of formulae as binary strings.

The semantics of the formulae $\Phi(P, \mathbf{B})$ is given, as usual, via valuation functions, or valuations for short, $V : P \to \{0, 1\}$ that map the set of propositional symbols P to the set $\{0, 1\}$ with the provision that $\mathbf{1}$ is always mapped to 1. It is a standard result of propositional logic that every such mapping $V : P \to \{0, 1\}$ has a unique extension V_{ext} that maps $\Phi(P, \mathbf{B})$ to $\{0, 1\}$ and respects the operators from $\mathbf{B}$. We will be particularly interested in the following function

$$\begin{array}{lll} V_{ext}(\mathbf{1}) = 1 & & \text{always;} \\ V_{ext}(p_i) = 1 & \text{iff} & V(p_i) = 1; \\ V_{ext}(\neg\psi) = 1 & \text{iff} & V_{ext}(\psi) = 0; \\ V_{ext}(\psi \vee \theta) = 1 & \text{iff} & V_{ext}(\psi) = 1 \text{ or } V_{ext}(\theta) = 1; \\ V_{ext}(\psi \wedge \theta) = 1 & \text{iff} & V_{ext}(\psi) = 1 \text{ and } V_{ext}(\theta) = 1; \\ V_{ext}(\psi \oplus \theta) = 1 & \text{iff} & V_{ext}(\psi) = 1 \text{ or } V_{ext}(\theta) = 1 \text{ but not both.} \end{array}$$

For our purposes, it is convenient to identify V_{ext} with V and write $V \models \psi$ instead of $V_{ext}(\psi) = 1$. Hence, we can rewrite the above as follows.

$$\begin{array}{lll} V \models \mathbf{1} & & \text{always;} \\ V \models p_i & \text{iff} & V(p_i) = 1; \\ V \models \neg\psi & \text{iff} & V \not\models \psi; \\ V \models \psi \vee \theta & \text{iff} & V \models \psi \text{ or } V \models \theta; \\ V \models \psi \wedge \theta & \text{iff} & V \models \psi \text{ and } V \models \theta; \\ V \models \psi \oplus \theta & \text{iff} & V \models \psi \text{ or } V \models \theta \text{ but not both.} \end{array}$$

A formula φ such that $V \not\models \varphi$ for any V (equivalently, $V(\varphi) = 0$ for any V) is called a **contradiction**. Two formulae φ_1 and φ_2 are called **equivalent** iff for every valuation V, we have $V \models \varphi_1$ iff $V \models \varphi_2$. It is easy to see that $V \models \varphi \oplus \psi$ iff $V \models (\neg\varphi \wedge \psi) \vee (\varphi \wedge \neg\psi)$ and $V \models \neg\varphi$ iff $V \models \mathbf{1} \oplus \varphi$. These equivalences are the essential facts behind the standard result of propositional logic that for any formula $\varphi \in \Phi(P, \mathbf{DM})$ there is an equivalent formula $\psi \in \Phi(P, \mathbf{Mod2})$ and vice versa.

Let σ denote an element of the set $\{0, 1\}$. As usual, for any $p \in P$, we call p^σ a **literal** and define it as follows

$$p^\sigma = \begin{cases} p, & \text{if } \sigma = 1; \\ \neg p, & \text{if } \sigma = 0. \end{cases}$$

Until the end of this subsection, we are going to write $\varphi\psi$ instead of $\varphi \wedge \psi$. The length of $\varphi\psi$ is defined as $|\varphi\psi| = |\varphi| + |\psi|$. It is obvious that for any formula $\varphi \in \Phi(P, \mathbf{DM})$, there is an equivalent formula φ' obtained from φ by replacing every sub-formula $\psi_1 \wedge \psi_2$ with $\psi_1\psi_2$; moreover, $|\varphi'| \leq |\varphi|$. Let us denote the set of all such formulae by $\Phi_{\setminus\wedge}(P, \mathbf{DM})$. It is clear that lower bounds proven with respect to the length of formulae in $\Phi_{\setminus\wedge}(P, \mathbf{DM})$ will remain valid for the formulae in $\varphi \in \Phi(P, \mathbf{DM})$, too, i.e., if we prove that the shortest formula in $\Phi_{\backslash\wedge}(P, \mathbf{DM})$ that has a property P cannot have length smaller than n, then the same would be true for the formulae in $\Phi(P, \mathbf{DM})$. A formula of the form $p_k^{\sigma_k} p_l^{\sigma_l} \dots p_n^{\sigma_n}$ such that there are no two different occurrences of the same propositional symbol p_j, where $j \in \{k, l, \dots, n\}$, is called an **elementary conjunction** (EC). To simplify our exposition, we assume that the propositional symbols occurring in an EC are ordered in increasing order of their indices. The set of all elementary conjunctions built from propositional symbols from P is denoted $EC(P)$.

Example 2.2. *Given the above definition, we see that* $p_1^0 p_4^1 p_{17}^0$ *is an elementary conjunction but* $p_4^1 p_{17}^1 p_{17}^0$ *is not because there are two occurrences of the same propositional symbol* p_{17} *in the latter.*

It is easy to see that for any valuation $V : P \to \{0, 1\}$,

$$V \models p_k^{\sigma_k} p_l^{\sigma_l} \dots p_n^{\sigma_n} \text{ iff } V(p_j) = \sigma_j \text{ for any } j \in \{k, l, \dots n\}.$$

We say that the elementary conjunction EC_1 is contained in the elementary conjunction EC_2, and write $EC_1 \preceq EC_2$, if all the literals of EC_1 occur in EC_2.

Example 2.3. *If* $EC_1 \stackrel{def}{=} p_1^1 p_3^1$ *and* $EC_2 \stackrel{def}{=} p_1^1 p_2^0 p_3^1$*, then* $EC_1 \preceq EC_2$.

As is well-known from elementary propositional logic, for any basis $\mathbf{B}$ and any formula $\varphi \in \Phi(P, \mathbf{B})$, whether $V \models \varphi$ is true or not does not depend on the value of $V(p)$ for any propositional symbol p that does not occur in φ. This fact, of course, is the reason why the semantics of Boolean formulae is often given via truth tables. Hence, when appropriate, we will sacrifice precision and write,

$$\sigma_i \dots \sigma_m \models \varphi$$

to mean that only the propositional symbols $p_i, \dots, p_m$ occur in φ and when the propositional symbol p_i is assigned the value σ_i, ..., and p_m is assigned the value σ_m, then φ evaluates to 1. The reader can think about the sequence $\sigma_i \dots \sigma_m$ as a row of the truth table for φ that results in φ's being true.

Example 2.4. *We have*

$$0_1 1_2 0_3 \models p_1^0 p_2^1 p_3^0 \quad \textit{and} \quad 0_2 1_5 0_6 \models p_2^0 p_5^1 p_6^0$$

but

$$0_2 1_5 0_6 \not\models p_2^1 p_5^1 p_6^0$$

because the valuation $0_2 1_5 0_6$ *assigns* 0 *to* p_2.

The following simple fact will be important later.

Proposition 2.2. $\sigma_1 \sigma_2 \dots \sigma_n \models p_1 \oplus p_2 \oplus \dots \oplus p_n$ *iff the number of* 1*'s in the sequence* $\sigma_1 \sigma_2 \dots \sigma_n$ *is odd.*

In this section we are particularly interested in the set of formulae defined below.

Definition 2.3 (Disjunctive Forms)**.** The set $DF(P)$ of formulae in disjunctive form over P is the smallest subset of $\Phi(P, \mathbf{DM})$ such that

- $EC(P) \subset DF(P)$;
- if φ and ψ belong to $DF(P)$, then so does $(\varphi \vee \psi)$.

Hence, φ is a formula in disjunctive form iff $\varphi \stackrel{\text{def}}{=} EC_1 \vee EC_2 \vee \ldots \vee EC_n$, where $n \geq 1$. A formula φ is in **disjunctive normal form** if $\varphi \stackrel{\text{def}}{=} EC_1 \vee EC_2 \vee \ldots \vee EC_n$, $n \geq 1$, and all the elementary conjunctions are pairwise different. Note that an elementary conjunction is both in a disjunctive form and in a disjunctive normal form.

The next standard result is a simple corollary of Shannon's decomposition of Boolean functions [56].

Theorem 2.4 (Canonical Disjunctive Normal Form)**.** *For any basis* $\mathbf{B}$ *and any formula* $\varphi \in \Phi(P, \mathbf{B})$ *that is not a contradiction, there is an equivalent formula in disjunctive normal form.*

Proof. Without loss of generality, let $p_1, p_2, \ldots, p_n$ be all the different propositional symbols occurring in φ. Let $\mathbb{1} = \{\sigma_1\sigma_2 \ldots \sigma_n \mid \sigma_1\sigma_2 \ldots \sigma_n \models \varphi\}$ be the set of all rows of the truth table of φ for which φ evaluates to 1. Note that $\mathbb{1}$ is non-empty because φ is not a contradiction[1]. The formula

$$CDNF(\varphi) \stackrel{\text{def}}{=} \bigvee_{\{\sigma_1\sigma_2\ldots\sigma_n \in \mathbb{1}\}} p_1^{\sigma_1} p_2^{\sigma_2} \ldots p_n^{\sigma_n}$$

is equivalent to φ and is called its **canonical disjunctive normal form**.

□

This theorem shows that there is no difference between $\Phi(P, \mathbf{DM})$, $\Phi(P, \mathbf{Mod2})$ and $DF(P)$ in terms of the properties of valuations they can express, i.e., their expressive power is the same. However, there is a big difference between these sets of formulae in terms of the "efficiency" or "succinctness" with which they express certain properties, as we will now demonstrate.

Example 2.5. *Note that, in general, there is more than one formula in disjunctive normal form that is equivalent to a formula* φ*. For example, both* $p_1^1 p_3^1 p_4^1 \vee p_1^0 p_3^1 p_4^1$ *and* $p_3^1 p_4^1$ *are formulae in disjunctive normal form that are equivalent to the formula* $(p_1 \vee \neg p_1) p_3 p_4$*.*

It is a classic result that goes at least as far back as [42] that the shortest formula in disjunctive normal form that is equivalent to the formula $l_n \stackrel{\text{def}}{=} p_1 \oplus p_2 \oplus \ldots \oplus p_n$ is its canonical disjunctive normal form $CDNF(l_n)$ which contains 2^{n-1} elementary conjunctions each one of them containing n pairwise different propositional symbols. Therefore, $CDNF(l_n)$ contains $n \times 2^{n-1}$ different occurrences of propositional symbols.

[1]It is easy to see that contradictions do not have equivalent disjunctive normal forms.

Example 2.6. *If $n = 3$, then the shortest formula in disjunctive normal form that is equivalent to $l_3 \stackrel{def}{=} p_1 \oplus p_2 \oplus p_3$ is $CDNF(l_3) \stackrel{def}{=} p_1^0 p_2^0 p_3^1 \vee p_1^0 p_2^1 p_3^0 \vee p_1^1 p_2^0 p_3^0 \vee p_1^1 p_2^1 p_3^1$.*

As a gentle introduction to the type of reasoning employed in the later chapters, we are going to reprove this result and introduce some important concepts and techniques along the way.

The general problem we are facing can be intuitively described as follows. We want to prove lower bounds on the size of certain type of formulae φ (in our case, formulae in disjunctive form) that express certain property P with respect to a given semantics (in our case the semantics is given via valuation function and the property P is given by Proposition 2.2, namely, "the valuation V assigns the value 1 to an odd number of propositional symbols among the symbols $\{p_1, \ldots, p_n\}$"). Therefore, we need a tool that allows us to reason about the length of such formulae. Intuitively, it will be helpful if this tool is tailored simultaneously to the definition of the length of the formulae φ and the fact that they can differentiate between objects that have the property P and those that do not. We can achieve this by associating, as usual, with every formula φ its syntax tree[2] T_φ. Then we can use the number of nodes of T_φ as an estimate of the length of φ. In addition, we add some new features to these syntax trees in order to be able to reason about formulae that separate objects that have a given property from the ones that do not. Extended syntax trees were introduced in [27] in the setting of first order logic and can be used as a formalisation of the above intuition. We define the corresponding notion for Boolean formulae in disjunctive form below in Definition 2.5.

Intuitively, as its name suggests, an extended syntax tree of a Boolean formula in disjunctive form φ is a syntax tree of φ where, apart from a syntax label that can be either an elementary conjunction occurring in φ or the symbol $\vee$, each node has a semantic label $\langle \mathbb{A}, \mathbb{B} \rangle$ where $\mathbb{A}$ and $\mathbb{B}$ are two sets of valuations. From now on, a tree node with a semantic label $\langle \mathbb{A}, \mathbb{B} \rangle$ will be denoted $\mathbb{A} \circ \mathbb{B}$. The valuations in $\mathbb{A}$ will be called **the valuations on the left**. Similarly, the valuations in $\mathbb{B}$ will be called **the valuations on the right**. Additionally, if $\mathbb{V}$ is a set of valuations, we write $\mathbb{V} \models \varphi$ to mean that for all $V \in \mathbb{V}$, it is true that $V \models \varphi$. We would like to stress that this definition applies also to the case where $\mathbb{V} = \emptyset$. In particular, it is trivially true that for all valuations $V \in \emptyset$, we have $V \models \varphi$ and therefore, $\emptyset \models \varphi$.

Definition 2.5 (Extended Syntax Trees for Disjunctive Forms)**.** Let $\varphi \in DF(P)$ and let $\mathbb{A}$ and $\mathbb{B}$ be two (not necessarily non-empty) sets of valuations such that $\mathbb{A} \models \varphi$ and $\mathbb{B} \models \neg\varphi$. The extended syntax tree $T_\varphi^{\langle \mathbb{A}, \mathbb{B} \rangle}$ of φ is defined inductively on the structure of φ as follows.

(φ is an elementary conjunction $p_l^{\sigma_l} p_m^{\sigma_m} \ldots p_n^{\sigma_n}$): $T_\varphi^{\langle \mathbb{A}, \mathbb{B} \rangle}$ consists of a single node that has a **syntax label** $p_l^{\sigma_l} p_m^{\sigma_m} \ldots p_n^{\sigma_n}$ and a **semantic label** $\langle \mathbb{A}, \mathbb{B} \rangle$. Note that $\mathbb{A} \models \varphi$ while $\mathbb{B} \models \neg\varphi$.

[2] We assume that the reader is familiar with the graph-theoretic notion of tree and do not give formal definitions.

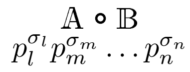


FIGURE 2.1: The extended syntax tree $T^{\langle \mathbb{A},\mathbb{B}\rangle}_{p_l^{\sigma_l} p_m^{\sigma_m} \dots p_n^{\sigma_n}}$.

(φ is $\psi_1 \vee \psi_2$): $T^{\langle \mathbb{A},\mathbb{B}\rangle}_{\psi_1 \vee \psi_2}$ has a root r with syntax label $\vee$ and semantic label $\langle \mathbb{A}, \mathbb{B}\rangle$. The first child of r is the root of $T^{\langle \mathbb{A}_1,\mathbb{B}\rangle}_{\psi_1}$. The second child of r is the root of $T^{\langle \mathbb{A}_2,\mathbb{B}\rangle}_{\psi_2}$ for some (possibly empty) $\mathbb{A}_1$ and $\mathbb{A}_2$ such that $\mathbb{A}_1 \models \psi_1$, $\mathbb{A}_2 \models \psi_2$, and $\mathbb{A} = \mathbb{A}_1 \cup \mathbb{A}_2$. Note that there is at least one subset $\mathbb{A}_1$ such that $\mathbb{A}_1 \models \psi_1$ and at least one $\mathbb{A}_2$ such that $\mathbb{A}_2 \models \psi_2$ because the opposite means that $\mathbb{A} \models \psi_1 \vee \psi_2$ is not true. Additionally, since $\mathbb{B} \models \neg(\psi_1 \vee \psi_2)$, we have $\mathbb{B} \models \neg\psi_1$ and $\mathbb{B} \models \neg\psi_2$.

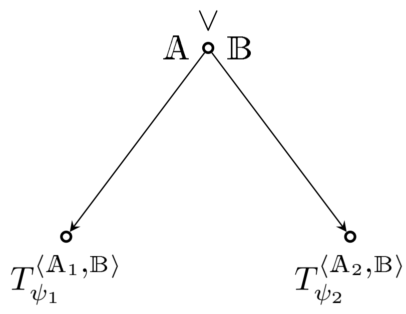


FIGURE 2.2: The extended syntax tree $T^{\langle \mathbb{A},\mathbb{B}\rangle}_{\psi_1 \vee \psi_2}$.

The number of nodes n of an extended syntax tree is called its **size**.

Example 2.7. *Let the set* $\mathbb{A}$ *consist of the valuations* $0_1 0_2 1_3$, $0_1 1_2 0_3$, $1_1 0_2 0_3$. *Let* $\mathbb{B}$ *contain just one valuation, namely* $0_1 1_2 1_3$. *The extended syntax tree* $T^{\langle \mathbb{A},\mathbb{B}\rangle}_{p_1^0 p_2^0 p_3^1 \vee p_1^0 p_2^1 p_3^0 \vee p_1^1 p_2^0 p_3^0}$ *of the formula* $p_1^0 p_2^0 p_3^1 \vee p_1^0 p_2^1 p_3^0 \vee p_1^1 p_2^0 p_3^0$ *is shown in Figure 2.3 below.*

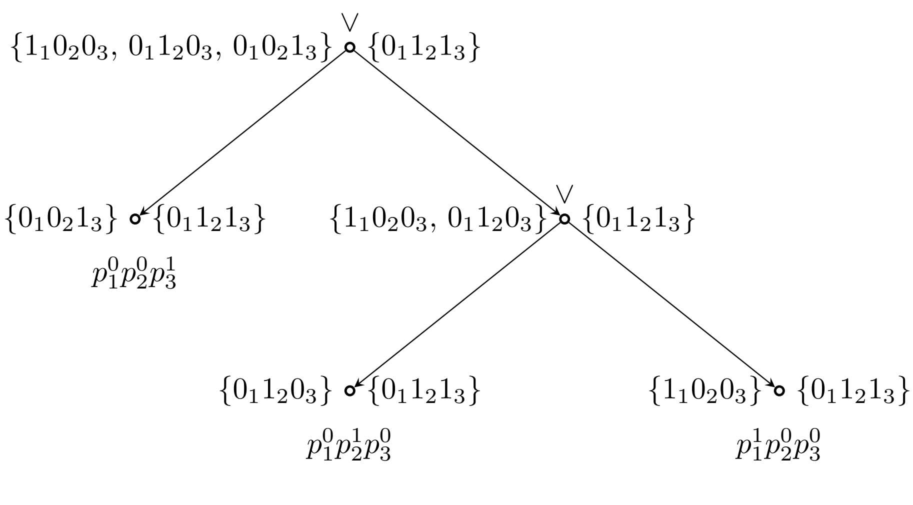


FIGURE 2.3: The extended syntax tree $T^{\langle \mathbb{A},\mathbb{B}\rangle}_{p_1^0 p_2^0 p_3^1 \vee p_1^0 p_2^1 p_3^0 \vee p_1^1 p_2^0 p_3^0}$.

Note that $\mathbb{A} \models p_1^0 p_2^0 p_3^1 \vee p_1^0 p_2^1 p_3^0 \vee p_1^1 p_2^0 p_3^0$ *and* $\mathbb{B} \models \neg(p_1^0 p_2^0 p_3^1 \vee p_1^0 p_2^1 p_3^0 \vee p_1^1 p_2^0 p_3^0)$.

The next proposition follows immediately from Definition 2.5

Proposition 2.6. *For any two formulae in disjunctive form* φ *and* ψ *and any sets of valuations* $\mathbb{A}$, $\mathbb{B}$, $\mathbb{C}$, *and* $\mathbb{D}$, *if the extended syntax tree* $T_\varphi^{\langle \mathbb{A}, \mathbb{B} \rangle}$ *contains a node* $\mathbb{C} \circ \mathbb{D}$ *that is the root of the syntax tree* $T_\psi^{\langle \mathbb{C}, \mathbb{D} \rangle}$, *then* $\mathbb{C} \models \psi$ *and* $\mathbb{D} \models \neg\psi$.

The initial motivation for the introduction of extended syntax tree came from the seminal [1] where the main idea, as in the case of Ehrenfeucht-Fraïssé (EF) games [16, 20] which are used to capture the combinatorial content of first-order quantification, is to capture the combinatorial content of the truth definition of first order formulae in the form of a game. The resulting games are called Adler-Immerman (AI) games, a versatile generalisation of EF games, that provide us with a technical means for proving lower bounds on formula size in different logics. Although the proofs of our main results in the next chapters are formulated in terms of extended syntax trees, we think that the reader will find it useful to be given two slightly different perspectives on essentially the same object and that is why we define here a version of AI games that is suitable for Boolean formulae in disjunctive form. After that, we make the connection between these games and extended syntax trees formally precise.

Definition 2.7 (AI games for disjunctive forms)**.** The one-person (called Spoiler) AI game on two sets of valuations $\mathbb{A}$ and $\mathbb{B}$ is played as follows. During the course of the game, a game tree is constructed in such a way that each node is labelled with a pair $\langle \mathbb{C}, \mathbb{D} \rangle$ of (not necessarily non-empty) sets of valuations and one element from the set $EC(P) \cup \{\vee\}$. A node labelled with the pair $\langle \mathbb{C}, \mathbb{D} \rangle$ is denoted $\mathbb{C} \circ \mathbb{D}$. As before, the valuations in $\mathbb{C}$ are called the valuations on the left while the valuations in $\mathbb{D}$ are called the valuations on the right. A node can be declared either **open** or **closed**. Once a node has been declared "closed", no further game-moves are played at it. Moves can be played only at open nodes. The game begins with the root $\mathbb{A} \circ \mathbb{B}$ of the game tree that is declared "open".

Let an open node $\mathbb{C} \circ \mathbb{D}$ be given. Spoiler can make one of the following moves at this node:

EC-move: Spoiler chooses an elementary conjunction $p_l^{\sigma_l} p_m^{\sigma_m} \dots p_n^{\sigma_n} \in EC(P)$ such that $\mathbb{C} \models p_l^{\sigma_l} p_n^{\sigma_n} \dots p_n^{\sigma_n}$ and $\mathbb{D} \models \neg(p_l^{\sigma_l} p_m^{\sigma_m} \dots p_n^{\sigma_n})$. The node is declared closed and labelled with the elementary conjunction $p_l^{\sigma_l} p_m^{\sigma_m} \dots p_n^{\sigma_n}$.

$\vee$-move: Spoiler labels the node with the symbol $\vee$ and chooses two (not necessarily non-empty) subsets $\mathbb{C}_1 \subseteq \mathbb{C}$ and $\mathbb{C}_2 \subseteq \mathbb{C}$ such that $\mathbb{C} = \mathbb{C}_1 \cup \mathbb{C}_2$. Two new open nodes are added to the tree as successors to the node $\mathbb{C} \circ \mathbb{D}$ (which is declared closed), namely $\mathbb{C}_1 \circ \mathbb{D}$ and $\mathbb{C}_2 \circ \mathbb{D}$.

A game tree T is called **closed** iff T is finite and all its leaves are closed. The **size** of a closed game tree T is the number of its nodes.

Definition 2.8 (Winning Condition for AI games)**.** We say that Spoiler wins the AI game on a pair of sets of valuations $\langle \mathbb{A}, \mathbb{B} \rangle$ in n moves if and only if there is a closed game tree T of size n with root $\mathbb{A} \circ \mathbb{B}$.

Example 2.8. *Let the set of valuations $\mathbb{A}$ consist of the three valuation functions V_1, V_2, and V_3 defined as follows:*

- $V_1(p_1) = 1$ *and* $V_1(p_n) = 0$ *for all* $n > 1$*;*
- $V_2(p_2) = 1$ *while* $V_2(p_n) = 0$ *for all* $n \neq 2$*;*
- $V_3(p_3) = 1$ *whereas* $V_2(p_n) = 0$ *for all* $n \neq 3$*;*

Let the set $\mathbb{B}$ contain only one valuation V such that

- $V(p_2) = V(p_3) = 1$ *and* $V(p_n) = 0$ *for all* $n \notin \{2, 3\}$*.*

Then, it is easy to see that the extended syntax tree from Figure 2.3 is actually a closed game tree for the AI game on the pair $\langle \mathbb{A}, \mathbb{B} \rangle$.

The above example makes it obvious that there is a strong connection between extended syntax trees and AI games which we make explicit in the next two theorems

Theorem 2.9. *For every pair $\langle \mathbb{A}, \mathbb{B} \rangle$ of sets of valuations, every closed game tree T with root $\mathbb{A} \circ \mathbb{B}$ is an extended syntax tree $T_{\varphi}^{\langle \mathbb{A}, \mathbb{B} \rangle}$ of a formula φ in disjunctive form such that $\mathbb{A} \models \varphi$ and $\mathbb{B} \models \neg\varphi$.*

Proof. We have to prove that T satisfies Definition 2.5. This is done by induction on the size of T.

Base case:

If T has just one node, then according to Definition 2.7 it is labelled with an elementary conjunction EC such that $\mathbb{A} \models EC$ and $\mathbb{B} \models \neg EC$. Therefore, T is an extended syntax tree of EC.

Induction hypothesis: We assume that the statement is true for all closed game trees of size at most n.

Induction step:

Let T contain $n + 1$ nodes where $n \geq 1$. Using Definition 2.7, we see that the root $\mathbb{A} \circ \mathbb{B}$ of T is labelled with a syntax label $\vee$ and it has two successors $\mathbb{A}_1 \circ \mathbb{B}$ and $\mathbb{A}_2 \circ \mathbb{B}$, where $\mathbb{A}_1 \cup \mathbb{A}_2 = \mathbb{A}$, that are roots of two closed sub-trees T_1 and T_2, respectively. Applying the induction hypothesis, we see that there is a formula in disjunctive form ψ_1 such that T_1 is an extended syntax tree of ψ_1; analogously, there is a formula ψ_2 for which T_2 is an extended syntax tree; moreover, we have that $\mathbb{A}_1 \models \psi_1$ and $\mathbb{A}_2 \models \psi_2$ while $\mathbb{B} \models \neg\psi_1$ and $\mathbb{B} \models \neg\psi_2$. Hence, T is an extended syntax tree of $\psi_1 \vee \psi_2$. □

Theorem 2.10. *For every formula φ in disjunctive form and any pair of sets of valuations $\mathbb{A}$ and $\mathbb{B}$ such that $\mathbb{A} \models \varphi$ and $\mathbb{B} \models \neg\varphi$, the extended syntax tree $T_{\varphi}^{\langle \mathbb{A}, \mathbb{B} \rangle}$ is a closed game tree T with root $\mathbb{A} \circ \mathbb{B}$.*

Proof. Let $\mathbb{A} \models \varphi$ and $\mathbb{B} \models \neg\varphi$. We prove the statement by induction on the structure of φ. The intuition behind the proof is that the extended syntax tree represents a game starting at $\mathbb{A} \circ \mathbb{B}$ that Spoiler has won by "playing according to" φ, i.e., φ encodes a winning strategy for Spoiler.

Base case:

If φ is an elementary conjunction $p_l^{\sigma_l} p_m^{\sigma_m} \dots p_n^{\sigma_n}$, then, according to the first item from Definition 2.5, the extended syntax tree $T_\varphi^{\langle \mathbb{A}, \mathbb{B} \rangle}$ contains one node with a semantic label $\langle \mathbb{A}, \mathbb{B} \rangle$ and a syntax label $p_l^{\sigma_l} p_m^{\sigma_m} \dots p_n^{\sigma_n}$. However, using the first item from Definition 2.7, we see that this is a closed game tree for a game in which Spoiler used the elementary conjunction $p_l^{\sigma_l} p_m^{\sigma_m} \dots p_n^{\sigma_n}$ to close the only node $\mathbb{A} \circ \mathbb{B}$.

Induction hypothesis: Let the statement be true for ψ_1 and ψ_2.

Induction step: If φ is $\psi_1 \vee \psi_2$, where both ψ_1 and ψ_2 are in disjunctive form, using the second item from Definition 2.5, we see that the extended syntax tree $T_\varphi^{\langle \mathbb{A}, \mathbb{B} \rangle}$ has a root r with syntax label $\vee$ and semantic label $\langle \mathbb{A}, \mathbb{B} \rangle$; moreover, the first child of r is the root of $T_\varphi^{\langle \mathbb{A}_1, \mathbb{B} \rangle}$ while the second child of r is the root of $T_\varphi^{\langle \mathbb{A}_2, \mathbb{B} \rangle}$ for some sets $\mathbb{A}_1$ and $\mathbb{A}_2$ such that $\mathbb{A} = \mathbb{A}_1 \cup \mathbb{A}_2$ and $\mathbb{A}_1 \models \psi_1$, $\mathbb{A}_2 \models \psi_2$. Applying the induction hypothesis, we see that the extended syntax trees $T_{\psi_1}^{\langle \mathbb{A}_1, \mathbb{B} \rangle}$ and $T_{\psi_2}^{\langle \mathbb{A}_2, \mathbb{B} \rangle}$ are closed game trees with roots $\mathbb{A}_1 \circ \mathbb{B}$ and $\mathbb{A}_2 \circ \mathbb{B}$, respectively. Hence, according to the second item from Definition 2.7, $T_\varphi^{\langle \mathbb{A}, \mathbb{B} \rangle}$ is a closed game tree with root $\mathbb{A} \circ \mathbb{B}$.

□

Having shown the connection between extended syntax trees and AI games, we are going to base the proofs of our results on the former because we want to highlight the fact that, when proving lower bounds on the size of all formulae φ such that $\mathbb{A} \models \varphi$ and $\mathbb{B} \models \neg\varphi$, we are actually proving lower bounds on the number of nodes in the syntax tree of φ. We would like to stress however that thinking in terms of AI games can be intuitively very helpful especially if we are interested in upper bounds on the size of formulae that are true in $\mathbb{A}$ and false in $\mathbb{B}$. In this case, Theorem 2.9 tells us that we simply have to find one closed game tree with root $\mathbb{A} \circ \mathbb{B}$.

Of course, lower-bound arguments based on AI games are also possible. Indeed, if we are interested in finding a formula φ of minimal length such that $\mathbb{A} \models \varphi$ and $\mathbb{B} \models \neg\varphi$, we have to reason about the shortest possible winning game for Spoiler that starts at $\mathbb{A} \circ \mathbb{B}$. The presentation of such arguments however is often mathematically less clean than a presentation based on extended syntax trees.

Now we go back to our argument that for every $n \geq 2$, the shortest formula in disjunctive normal form that is equivalent to the formula l_n is actually $CDNF(l_n)$. How can we prove such a statement? It is worth pointing out that we claim that there is only one formula in disjunctive normal form that is the shortest equivalent to l_n. A priori, it is not clear why there must be only one such formula. It is perfectly conceivable that there might be at least two of the same length containing different elementary conjunctions. However, we are going to show that there are two sets of valuations $\mathbb{O}_n$ and $\mathbb{E}_n$ such that:

(∗) $\mathbb{O}_n$ contains only valuations that assign the value 1 to an odd number of propositional symbols among $\{p_1, \ldots, p_n\}$ while $\mathbb{E}_n$ consists only of valuations that assign 1 to an even number of propositional symbols in the set $\{p_1, \ldots, p_n\}$ and, therefore, applying Proposition 2.2, we have $\mathbb{O}_n \models \mathit{CDNF}(l_n)$ and $\mathbb{E}_n \models \neg \mathit{CDNF}(l_n)$;

(∗∗) for any two different elementary conjunctions EC_1 and EC_2 occurring in $\mathit{CDNF}(l_n)$ and any formula in disjunctive form φ such that $\mathbb{O}_n \models \varphi$ and $\mathbb{E}_n \models \neg\varphi$, the extended syntax tree of φ with root $\mathbb{O}_n \circ \mathbb{E}_n$ has two different leaves labelled with elementary conjunctions EC' and EC'' such that $EC_1 \preceq EC'$ and $EC_2 \preceq EC''$.

It follows immediately from these two items and the obvious fact that the shortest elementary conjunctions EC' and EC'' such that $EC_1 \preceq EC'$ and $EC_2 \preceq EC''$ are actually EC_1 and EC_2, that the shortest formula in disjunctive form that is equivalent to l_n is $\mathit{CDNF}(l_n)$.

Let us show that there are indeed such sets of valuations $\mathbb{O}_n$ and $\mathbb{E}_n$. The following notation will be helpful. As before, we write $\sigma_1\sigma_2\ldots\sigma_n$ to represent the valuation V that assigns the value σ_1 to the propositional symbol p_1, the value σ_2 to p_2, …, and the value σ_n to p_n but this time we assume that the rest of the variables p_i, where $i > n$, are assigned the value 0. Let $\mathbb{O}_n$ be the set of valuations $\sigma_1\sigma_2\ldots\sigma_n$ such that the number of 1's among $\sigma_1\sigma_2\ldots\sigma_n$ is odd. Similarly, $\mathbb{E}_n$ is the set of valuations $\sigma_1\sigma_2\ldots\sigma_n$ such that the number of 1's among $\sigma_1, \sigma_2, \ldots, \sigma_n$ is even. Therefore,

$$\mathbb{O}_n \models l_n \text{ and } \mathbb{E}_n \models \neg l_n$$

which implies

$$\mathbb{O}_n \models \mathit{CDNF}(l_n) \text{ and } \mathbb{E}_n \models \neg\mathit{CDNF}(l_n).$$

It is easy to see that $\mathbb{E}_n$ and $\mathbb{O}_n$ each contain 2^{n-1} different valuations.

Example 2.9. *In the case of* $l_3 \overset{def}{=} p_1 \oplus p_2 \oplus p_3$*, we have that the set* $\mathbb{O}_3$ *consists of* $0_1 0_2 1_3$*,* $0_1 1_2 0_3$*,* $1_1 0_2 0_3$*, and* $1_1 1_2 1_3$ *while* $\mathbb{E}_3$ *contains only the valuation* $0_1 0_2 0_3$*,* $0_1 1_2 1_3$*,* $1_1 1_2 0_3$*, and* $1_1 0_2 1_3$*.*

Let us establish now that item (∗∗) is true. The proof is best understood via our running example for $l_3 \overset{\text{def}}{=} p_1 \oplus p_2 \oplus p_3$.

Let φ be a formula in disjunctive form that is equivalent to $\mathit{CDNF}(l_3)$. Using the sets $\mathbb{O}_3$ and $\mathbb{E}_3$, we can try and prove a lower bound on the number of nodes in the extended syntax tree $T_\varphi^{\langle \mathbb{O}_3, \mathbb{E}_3 \rangle}$ of φ which will give us a lower bound on the length of φ. However, there are significant combinatorial difficulties involved in working with the sets $\mathbb{O}_3$ and $\mathbb{E}_3$ as described in Definition 2.5. Therefore, we can try to reason with simpler subsets of $\mathbb{O}_3$ and $\mathbb{E}_3$. One way of doing this is the following. Let us suppose that we can prove that the extended syntax tree with root $\{0_1 0_2 1_3\} \circ \mathbb{E}_3$ of any formula

ψ in disjunctive form such that

$$\{0_1 0_2 1_3\} \models \psi \text{ and } \mathbb{E}_3 \models \neg\psi$$

contains a leaf labelled with an elementary conjunction EC_1 where $p_1^0 p_2^0 p_3^1 \preceq EC_1$. Then this will imply that $T_\varphi^{\langle \mathbb{O}_3, \mathbb{E}_3 \rangle}$ also contains a leaf labelled with such an elementary conjunction EC_1 because $\{0_1 0_2 1_3\} \models \varphi$ and $\mathbb{E}_3 \models \neg\varphi$ follows from our assumption that $\mathbb{O}_3 \models \varphi$ and $\mathbb{E}_3 \models \neg\varphi$, and $\{0_1 0_2 1_3\} \subset \mathbb{O}_3$.

Similarly, if we prove that the extended syntax tree with root $\{0_1 1_2 0_3\} \circ \mathbb{E}_3$ of any formula in disjunctive form ψ such that

$$\{0_1 1_2 0_3\} \models \psi \text{ and } \mathbb{E}_3 \models \neg\psi$$

contains a leaf labelled with an elementary conjunction EC_2 where $p_1^0 p_2^1 p_3^0 \preceq EC_2$, then, using the same reasoning as above, we see that $T_\varphi^{\langle \mathbb{O}_3, \mathbb{E}_3 \rangle}$ also contains a leaf labelled with such an elementary conjunction EC_2. We can follow this strategy with the other two elements of $\mathbb{O}_3$, namely $1_1 0_2 0_3$ and $1_1 1_2 1_3$. Therefore, we conclude that $T_\varphi^{\langle \mathbb{O}_3, \mathbb{E}_3 \rangle}$ has leaves labelled with elementary conjunctions EC_1, EC_2, EC_3 and EC_4 where

$$p_1^0 p_2^0 p_3^1 \preceq EC_1,\ p_1^0 p_2^1 p_3^0 \preceq EC_2,\ p_1^1 p_2^0 p_3^0 \preceq EC_3,\ p_1^1 p_2^1 p_3^1 \preceq EC_4.$$

The next proposition implies that EC_1, EC_2, EC_3, and EC_4 are pairwise different.

Proposition 2.11. *If EC^1 and EC^2 are two different elementary conjunctions occurring in the canonical disjunctive normal form $CDNF(l_n)$ of l_n, then there is no elementary conjunction EC such that $EC^1 \preceq EC$ and $EC^2 \preceq EC$.*

Proof. Let us suppose that there is such an elementary conjunction EC. Then every propositional symbol p_i, where $1 \leq i \leq n$ occurs just once in EC. However, since $EC^1 \stackrel{\text{def}}{=} p_1^{\sigma_1^1} p_2^{\sigma_2^1} \dots p_n^{\sigma_n^1}$ is different from $EC^2 \stackrel{\text{def}}{=} p_1^{\sigma_1^2} p_2^{\sigma_2^2} \dots p_n^{\sigma_n^2}$, there is at least one p_i, where $1 \leq i \leq n$, such that $p_i^{\sigma_i^1}$ occurs in EC^1 and $p_i^{\sigma_i^2}$ occurs in EC^2, and $\sigma_i^1 \neq \sigma_i^2$. Therefore, $EC^1 \preceq EC$ and $EC^2 \preceq EC$ imply that both $p_i^{\sigma_i^1}$ and $p_i^{\sigma_i^2}$ occur in EC and hence, EC is not an elementary conjunction.

□

In this way, we see that $T_\varphi^{\langle \mathbb{O}_3, \mathbb{E}_3 \rangle}$ contains at least 4 leaves labelled with elementary conjunctions EC_1, EC_2, EC_3, and EC_4 respectively, such that

$$\begin{array}{ll} p_1^0 p_2^0 p_3^1 \preceq EC_1 & p_1^0 p_2^1 p_3^0 \preceq EC_2 \\ p_1^1 p_2^0 p_3^0 \preceq EC_3 & p_1^1 p_2^1 p_3^1 \preceq EC_4. \end{array}$$

Hence, given the fact that $p_1^0 p_2^0 p_3^1$ is the shortest elementary conjunction such that $p_1^0 p_2^0 p_3^1 \preceq p_1^0 p_2^0 p_3^1$ and, similarly, $p_1^0 p_2^1 p_3^0$ is the shortest elementary conjunction such that $p_1^0 p_2^1 p_3^0 \preceq p_1^0 p_2^1 p_3^0$, etc., we see that the shortest formula in disjunctive normal form that is equivalent to l_3 is $CDNF(l_3)$.

It is obvious that the reasoning above can be applied to the case of l_n, where $n \neq 3$. Therefore, to complete the proof, we need only prove the following lemma.

Lemma 2.12. *For any $\sigma_1\sigma_2\ldots\sigma_n \in \mathbb{O}_n$ and any formula ψ such that $\sigma_1\sigma_2\ldots\sigma_n \models \psi$ and $\mathbb{E}_n \models \neg\psi$, the extended syntax tree of ψ with root $\{\sigma_1\sigma_2\ldots\sigma_n\}\circ\mathbb{E}_n$, has a leaf with syntax label EC where $p_1^{\sigma_1}\ldots p_n^{\sigma_n} \preceq EC$.*

Proof. It is easy to see that any extended syntax tree of ψ with root $\{\sigma_1\sigma_2\ldots\sigma_n\}\circ\mathbb{E}_n$ has a leaf $\{\sigma_1\sigma_2\ldots\sigma_n\}\circ\mathbb{E}_n$. The easy proof is by induction on the structure of ψ. If ψ is an elementary conjunction, the statement is obvious. If $\psi \overset{\text{def}}{=} \psi_1 \vee \psi_2$, then using the second item from Definition 2.5, we see that at least one of the two children of the root $\{\sigma_1\sigma_2\ldots\sigma_n\}\circ\mathbb{E}_n$ of the tree $T_\psi^{\langle\{\sigma_1\sigma_2\ldots\sigma_n\},\mathbb{E}_n\rangle}$ has a semantic label $\langle\{\sigma_1\sigma_2\ldots\sigma_n\},\mathbb{E}_n\rangle$. Hence, applying the induction hypothesis, we see that $T_\psi^{\langle\{\sigma_1\sigma_2\ldots\sigma_n\},\mathbb{E}_n\rangle}$ has at least one leaf with semantic label $\langle\{\sigma_1\sigma_2\ldots\sigma_n\},\mathbb{E}_n\rangle$. Next, we consider the possible syntax labels of a leaf $\{\sigma_1\sigma_2\ldots\sigma_n\}\circ\mathbb{E}_n$.

Since, for any $i > n$, the valuation $\sigma_1\sigma_2\ldots\sigma_n$ assigns the value 0 to the propositional symbol p_i, it is clear that an elementary conjunction EC containing a literal p_i^1 cannot be used as a syntax label for a leaf $\{\sigma_1\sigma_2\ldots\sigma_n\}\circ\mathbb{E}_n$ because $\sigma_1\sigma_2\ldots\sigma_n \not\models EC$ which contradicts Proposition 2.6. Similarly, an elementary conjunction EC that contains a literal p_i^σ, where $1 \leq i \leq n$ and $\sigma \neq \sigma_i$, cannot be used as a syntax label for a leaf $\{\sigma_1\sigma_2\ldots\sigma_n\}\circ\mathbb{E}_n$ because $\sigma_1\sigma_2\ldots\sigma_i\ldots\sigma_n \not\models EC$.

Example 2.10. *Consider a node $\{0_1 1_2 0_3\}\circ\mathbb{E}_3$. It is obvious that this node cannot have a syntax label $p_1^1 p_4^0$ or $p_2^1 p_3^0 p_4^1$ because $0_1 1_2 0_3 \not\models p_1^1 p_4^0$ and $0_1 1_2 0_3 \not\models p_2^1 p_3^0 p_4^1$.*

Hence, there are two possibilities left for any elementary conjunction EC that is a syntax label of a leaf $\{\sigma_1\sigma_2\ldots\sigma_n\}\circ\mathbb{E}_n$:

1. $p_1^{\sigma_1}\ldots p_n^{\sigma_n} \preceq EC$ and all the literals $p_i^{\sigma_i}$ occurring in EC such that $n < i$ are of the form p_i^0;

2. $p_1^{\sigma_1}\ldots p_n^{\sigma_n} \not\preceq EC$ and for all the literals $p_i^{\sigma_i}$ occurring in EC such that $n < i$, we have $\sigma_i = 0$.

The first case is obvious. Indeed, we have $\sigma_1\ldots\sigma_n \models EC$, whereas $\mathbb{E}_n \models \neg EC$ and the first item of Definition 2.5 is satisfied.

We claim that an elementary conjunction of the type described in item 2 cannot be used as a syntax label for $\{\sigma_1\sigma_2\ldots\sigma_n\}\circ\mathbb{E}_3$. Without loss of generality, let us suppose that the first m literals $p_1^{\sigma_1},\ldots p_m^{\sigma_m}$ among $p_1^{\sigma_1}\ldots p_n^{\sigma_n}$ do not appear in EC.

- If $m \geq n$, then all the literals p_j^σ occurring in EC are such that $\sigma = 0$ and $n \leq j$. However, in this case, it is true that $\sigma_1\sigma_2\ldots\sigma_n \models EC$ and $\mathbb{E}_n \models EC$. Therefore EC cannot be used as a syntax label of such a leaf.

- Let $m < n$.

– If the number of 1's among $\sigma_{m+1}, \sigma_{m+2}, \ldots \sigma_n$ is odd, then the valuation $1_1 0_2 0_3 \ldots 0_m \sigma_{m+1} \ldots \sigma_n$ belongs to $\mathbb{E}_n$ because it has an even number of 1's; what is more, we have that $1_1 0_2 0_3 \ldots 0_m \sigma_{m+1} \ldots \sigma_n \models EC$. Therefore, EC cannot be used as a syntax label for the node $\{\sigma_1 \sigma_2 \ldots \sigma_{m+1} \ldots \sigma_n\} \circ \mathbb{E}_n$.

– If the number of 1's among $\sigma_{m+1}, \sigma_{m+2}, \ldots \sigma_n$ is even, then the valuation $0_1 0_2 \ldots 0_m \sigma_{m+1} \ldots \sigma_n$ belongs to $\mathbb{E}_n$. It is obvious that in this case we have $0_1 0_2 \ldots 0_m \sigma_{m+1} \ldots \sigma_n \ldots \models EC$ and, hence, EC cannot be used as a syntax label for a node $\{\sigma_1 \sigma_2 \ldots \sigma_{m+1} \ldots \sigma_n\} \circ \mathbb{E}_n$.

Perhaps the argument above is best understood via an example.

Example 2.11. *Let us consider the node $\langle\{1_1 0_2 1_3 1_4\} \circ \mathbb{E}_4\rangle$ and the elementary conjunction $p_1^1 p_2^0 p_3^1 p_4^1$ that can be used as a syntax label for this node. Let us assume that $p_1^1 p_2^0 p_3^1 p_4^1 \not\preceq EC$. The considerations preceding Example 2.10 show that if EC contains a literal p_i^σ where $i > 4$, then $\sigma = 0$. Similarly, if EC contains a literal p_i^σ where $i \leq 4$, then p_i^σ coincides with the literal p_i^σ from $p_1^1 p_2^0 p_3^1 p_4^1$. For example, if p_2^σ occurs in EC, then $\sigma = 0$. Therefore, if $p_1^1 p_2^0 p_3^1 p_4^1 \not\preceq EC$, then for at least one of the propositional symbols $p_i \in \{p_1, p_2, p_3, p_4\}$, neither p_i^0 nor p_i^1 occurs in EC. Let us suppose that $EC \stackrel{\text{def}}{=} p_2^0 p_4^1 p_6^0 p_9^0$. In this case, we have $1_1 0_2 1_3 1_4 \models p_2^0 p_4^1 p_6^0 p_9^0$. However, the same is true about the valuation $1_1 0_2 0_3 1_4 \in \mathbb{E}_4$, i.e, $1_1 0_2 0_3 1_4 \models p_2^0 p_4^1 p_6^0 p_9^0$. In short, if some of the literals $p_1^1, p_2^0, p_3^1, p_4^1$ are missing from EC (and we know that none of the literals $p_1^0, p_2^1, p_3^0, p_4^0$ can appear in EC), then we can find a valuation in $\mathbb{E}_4$ that makes EC true. Thus, EC cannot be used as a syntax label for the node $\{1_1 0_2 1_3 1_4\} \circ \mathbb{E}_4$.*

This completes the proof that the shortest formula in disjunctive normal form that is equivalent to l_n is $CDNF(l_n)^3$. □

We established that the next items hold true.

There is an infinite sequence $Seq \stackrel{\text{def}}{=} l_2, l_3, \ldots$ of $\Phi(P, \mathbf{Mod2})$-formulae such that

(*a*) the length of each formula l_n from *Seq* is $2 \times n - 1$ or, in other words, the length of the formulae in *Seq* is bounded from above by a linear function of their indices;

(*b*) the length of the shortest formula $\varphi_n \in DF(P)$ that is equivalent to l_n is at least 2^{n-1}, i.e., there is no sequence $Seq_1 \stackrel{\text{def}}{=} \varphi_2, \varphi_3, \ldots$ of formulae in disjunctive form such that for every n, l_n is equivalent to φ_n and at the same time the length of the formulae in Seq_1 can be bounded from above by a sub-exponential function of their indices.

In the next chapters we will obtain a number of results that have the form above. Namely, for two logics L_1 and L_2 we will exhibit an infinite sequence of L_1-formulae

[3] Actually, it is not difficult to see that $CDNF(l_n)$ is the only formula in disjunctive normal form that is equivalent to l_n.

$\varphi_1, \varphi_2, \ldots$ such that every L_2-formula ψ_i that is equivalent to φ_i has size that is exponential in the size of φ_i. Using the terminology that we are going to introduce in Chapter 3, we say that if this is true then L_1 is exponentially more succinct than L_2. Thus items (a) and (b) above say that Boolean logic in the basis **Mod2** is exponentially more succinct than Boolean logic in the De Morgan basis in which we allow only formulae in disjunctive form. In contrast to this, a result from [67] says that if we drop the restriction to formulae in disjunctive form, then for every $n \geq 2$, there is a formula in the De Morgan basis of size less or equal to $\frac{9}{8}n^2$ that is equivalent to l_n; a much more difficult and deeper proof is required in order to show that, this formula is optimal [57]. Additionally, note that the formula $1 \oplus l_n$ is equivalent to $\overline{l_n} \stackrel{\text{def}}{=} p_1 \leftrightarrow p_2 \leftrightarrow \ldots \leftrightarrow p_n$ and analogous results as the ones above hold true for $\overline{l_n}$, i.e., the shortest formula in disjunctive normal form that is equivalent to $\overline{l_n}$ is $\mathit{CDNF}(\overline{l_n})$ but there is a formula in the De Morgan basis of size at most $\frac{9}{8}n^2$ that is equivalent to l_n; moreover, the lower bound $\frac{9}{8}n^2$ cannot be improved. The interested reader can consult [47] for a very clear exposition of this material.

2.2 A Short Overview of Some Results

Although, strictly speaking, this section is not necessary for the technical developments in the next chapters, we think that the reader would find it helpful to be given some historical perspective on the problem of proving lower bounds on formula size relative to given semantics. It goes without saying that the rough sketches below are best thought of as just pointers to a small fraction of the relevant literature.

Before going any further, it is worth pointing out that the term "formula size" is somewhat ambiguous. For example, in Boolean function complexity, a lower bound n on the size of a Boolean formula φ can mean that φ contains at least n occurrences of not necessarily distinct propositional variables or n Boolean operators. Of course, in both cases n gives a useful estimate of the "actual" length of φ. Similarly, in the case of first order logic, the size of a formula ψ is often defined as its quantifier rank or as the number of variables occurring in ψ not counting different occurrences of the same variable. Probably, a neutral term like "complexity" would be a better choice than "size", however, since these different meanings of "size" do not cause any big problems, we will continue using it in a way that makes explicit the resource we are talking about.

2.2.1 Lower Bounds On Formula Size for Boolean Functions

It seems fair to say that the great theoretical and practical importance of proving lower bounds on Boolean circuits and Boolean formulae expressing given Boolean functions was first realised by the Russian school in Boolean function complexity. In short, the most difficult open problem in this area can be loosely formulated as follows. John Riordan and Claude Shannon [53] used a counting argument to prove that for almost any Boolean function f, the syntax tree of any Boolean formula expressing f contains at least $\frac{2^n}{\log_2 n}$

leaves. A natural question then is whether we can exhibit an infinite sequence of "concrete" or "explicitly defined" Boolean functions $f_1(x_1), f_2(x_1, x_2), \ldots, f_n(x_1, x_2, \ldots x_n), \ldots$ having the property

$P \stackrel{\text{def}}{=}$ For every n, the smallest Boolean formula that computes the function f_n has size $\frac{2^n}{\log_2 n}$.

The reader can find more about the connections between this problem and some of the open problems in Computational complexity in [3] and [49]. Already in 1959, Sergey Yablonskii formulated his conjecture [68] which can be very informally explained as follows.

> If we want to "construct" a sequence of Boolean functions $S \stackrel{\text{def}}{=} f_1(x_1)$, $f_2(x_1, x_2)$, ..., $f_n(x_1, x_2, \ldots x_n)$, ... satisfying the property P above, for every $n \geq 1$, in order to find the n-th member of S, we have to do a brute force search among all the 2^{2^n} functions of n variables.

Now the reader can verify that what we actually did in the previous section was to present an algorithm that produces a sequence of Boolean functions, namely $l_1 \stackrel{\text{def}}{=} p_1 \oplus p_2$, $l_2 \stackrel{\text{def}}{=} p_1 \oplus p_2 \oplus p_3$, ..., $l_n \stackrel{\text{def}}{=} p_1 \oplus p_2 \oplus \ldots \oplus p_{n+1}$, ... such that for every n, the smallest formula in disjunctive normal form that computes the function l_n is $CDNF(l_n)$, i.e, has size at least 2^{n-1}. This was easy because of the structural properties of the formulae in disjunctive form.

The first big breakthrough in the search of high lower bounds on the formula size of "explicitly" given Boolean functions was Bella Subbotovskaya's result [57] that the size of any formula in the De Morgan basis equivalent to l_n is at least $n^{\frac{3}{2}}$. In order to obtain this bound, she invented the method or random restrictions. We state below the main theorem proven by Subbotovskaya on which this result was based.

Let $L(f)$ denote the minimal number of leaves in the syntax tree of any formula in the De Morgan basis, i.e., the number of occurrences of propositional symbols, that expresses or computes a Boolean function f and let $f(x_1, \ldots, x_n)$ be a function of n variables. Then

> for any k, where $1 \leq k \leq n$, it is possible to set $n - k$ variables among $(x_1, \ldots, x_n)$ to the constants 0 or 1 in such a way that the resulting new Boolean function f_1 of k variables satisfies the equation
>
> $$L(f_1) \leq \left(\frac{k}{n}\right)^{\frac{3}{2}} \times L(f).$$

Hence, if we fix $n - 1$ variables of l_n, we are left with a function of one variable l_1 such that any formula in the De Morgan basis computing l_1 has at least one propositional symbol, i.e., $L(l_1) \geq 1$. Using the statement above, we find

$$1 \leq L(l_1) \leq \left(\frac{1}{n}\right)^{\frac{3}{2}} \times L(l_n),$$

which, of course, implies $L(l_n) \geq n^{\frac{3}{2}}$.

Generalising Subbotovskaya's methods, Alexandr Andreev [2] constructed for any n, a function f_n such that the shortest formula in the De Morgan basis expressing f_n has size at least

$$\frac{n^{\frac{5}{2}}}{(\log_2 n)^{\frac{3}{2}} \log_2 \log_2 n}$$

Andreev's bound was improved to n^3 in 1998 by Johan Håstad [32]. Since then, this has been the largest known lower bound on the size of Boolean formulae computing an explicitly given sequence of Boolean functions

2.2.2 Lower Bounds on the Size of First Order Formulae

The quest for proving lower bounds on the size of first order formulae seems to be a relatively recent endeavour when compared to the Boolean logic case. It has been motivated by the advent of the field of descriptive complexity [35] where model theoretic methods are applied to the study of computational complexity problems.

In a series of papers, Neil Immerman and co-authors showed the deep connections between proving lower bounds on the quantifier rank and/or the number of variables of first order formulae and the unanswered questions about the relationship between SPACE and TIME in computational complexity theory [33], [34], [36], the problem of testing if two graphs are isomorphic [11], and shed a new light on some theorems in temporal logic [1] and [37]. Currently, the most exciting result in this area is Benjamin Rossman's proof [54] that any first order formula in a language containing $+, \times, \leq$ that says "a graph has a clique of size n" must have at least $\frac{n}{4}$ variables. Using this, Rossman and Immerman solved in the affirmative the long standing open problem whether the $n + 1$ variable fragment of first order logic is strictly more expressive on finite ordered graphs than the n variable fragment.

In all of the above papers, the main technical tool for obtaining the respective lower bounds on formula size is a suitable form of the EF game. Paper [1] is of special interest to us because it contains the original exposition of Adler-Immerman games which, as we said above, are one of our main technical tools. Readers familiar with EF games can think about AI games as EF games played on two **sets** of structures. We need sets of structures, instead of just two structures as in the case of EF games, in order to have game-moves corresponding to the Boolean connectives $\neg$, $\wedge$, and $\vee$. The versatility of AI games stems from the fact that we can formulate a suitable version for practically any logic. However, it was noted already in [1] that unlike EF games, AI games are not truly two-player games because the second player, called Delilah or Duplicator, has in a sense an "optimal" answer to every move by the first player called Samson or Spoiler. This optimal answer can be incorporated in the definition of the game and, therefore, there is no need for a second player. What is more, as we saw in the previous section, AI games can be replaced completely by the extended syntax trees defined by Martin Grohe and Nicole Schweikardt in [27]. Theorem 2.9 and Theorem 2.10 in the previous section give us the precise expression of this fact for Boolean formulae in disjunctive form. The main technical hurdle in this approach, of course, are the combinatorial difficulties involved in counting the nodes of extended syntax trees or equivalently, the number of moves in AI games. Currently, there are just two known techniques for doing this. Both of them can be best understood only after we have developed an appropriate technical tool-kit. That is why we postpone the discussion until the next chapter but just for the sake of completeness we add that the first approach is based on trying to count the number of different branches (equivalently, the number of leaves) in an extended syntax tree which seems easier to do than counting nodes in general. This is exactly what we did in the previous section. The second one is based on the notion of "complexity measure" first introduced in the setting of Boolean function complexity.

Of course, EF and AI games and extended syntax trees are not the only methods available for proving lower bounds on the size of first order formulae. Generally speaking, beside these games there are two other main types of proofs in this area. The first one relies on unproven but widely believed computational complexity conjectures. An example of this line of reasoning can be found in [26]. The second method is the so-called method of encoding large natural numbers by small trees. For a very clear exposition of this technique, the reader can consult [14].

2.2.3 Lower Bounds on Formulae Size in Modal Logic

It seems that the first to investigate the general problem of proving lower bounds on formula size in modal logic was Johan van Benthem who proved [60] that any modal formula that locally corresponds to the first order condition

$$\forall y \forall z((xRy \wedge xRz) \rightarrow (yRz \vee zRy \vee y = z))$$

must contain at least two different propositional variables. After that, the interest in proving such bounds was focussed mainly on temporal logic where, using results from automata theory, it was shown, among other things, that there is a sequence of first order formulae with three variables $\varphi_1, \varphi_2, \ldots$ such that the lengths of these formulae are bounded from above by a polynomial in their indices or, more formally, there is a polynomial p such that for any natural number $n \geq 1$, the length of φ_n is less or equal to $p(n)$. On the other hand, there is no sequence of temporal formulae $\psi_1, \psi_2, \ldots$ such that φ_i is equivalent on ω-words to ψ_i and the lengths of the formulae $\psi_1, \psi_2, \ldots$ can be bounded from above by an elementary function of their indices. Two papers dedicated to lower bounds on the size of temporal formulae deserve special mention. The first one is Thomas Wilke's automata-based proof [66] that every computation tree logic (or μ-calculus) formula that expresses the property

> there is a path along which there are n positions v_1, v_2, ..., v_n (not necessary in this order) satisfying the propositions p_1, p_2, ..., p_n respectively

must have size at least $\frac{2^n}{\sqrt{n}}$. This estimate was later improved in [1] to $n!$ with the help of AI games. The second paper we would like to mention is [17], in which the authors proved, again using automata theoretic techniques, that every formula of linear-time temporal logic (LTL) expressing the property

> for any two positions v and w on a path π, if $\pi, v \models p_i$ iff $\pi, w \models p_i$ for any $1 \leq i \leq n$, then $\pi, v \models p_{n+1}$ if and only if $\pi, w \models p_{n+1}$, i.e., if v and w agree on the first n propositions, then they agree on p_{n+1} too

has size at least 2^n. Using a slightly different property, namely

> any position v on a path π that agrees with the initial position v_0 on p_1, p_2, ..., p_n must also agree on p_{n+1}

it was proven in [40] that any LTL formula that

- expresses this property;
- contains only future temporal operators and is evaluated at the initial position of the path

must have size at least 2^n.

In contrast to temporal logic, results on lower bounds on formula size in the general setting of modal logic seem to be scarce. Besides van Benthem's work discussed above, we would like to mention the following articles. In [44], Carsten Lutz, Ulrike Sattler and Frank Wolter studied a modal language that has modal operators of 5 possible forms $[R]$, $[id]$, $[\neg S]$, $[S_1 \cap S_2]$, and $[S^-]$. The intuition is that, as usual, R is an atomic binary relation on the underlying Kripke structure while the more complex modalities are obtained from the identity relation id, the complement $\neg S$ of S, the intersection $S_1 \cap S_2$ of S_1 and S_2, and the converse S^- of the relation S. They proved, using model theoretic methods, that for any n, any modal formula in this language that defines the property

> the underlying set of the Kripke model has cardinality at least 2^n

must have size at least 2^{n-1}.

Using EF games, Martin Otto proved in [48] that every formula in the basic modal language with one modality $\Box$ that "says"

> there is a point that satisfies the proposition p reachable from the current point in at most 2^n steps

must have modal depth of at least 2^n.

In [59], Dimiter Vakarelov showed by algebraic means the following result. Let $n \geq 1$ and let us suppose that $2^n < k \leq 2^{n+1}$. Any formula in the basic modal language that modally defines the property

the current point has less than k successors

must contain at least $n + 1$ different propositional symbols.

Two other papers that establish lower bounds on the size of modal formulae are [19] and [43]. In the first one, Santiago Figuiera and Daniel Gorín used techniques based on EF and AI games to prove lower bounds on formulae in the basic modal language and suggest a connection between such problems and Kolmogorov complexity. In the second one, Carsten Lutz proved lower bounds on formulae expressing certain reachability properties in a multimodal language containing at least two modalities $[a]$ and $[b]$. The initial motivation for this thesis came from an open problem in [43].

2.2.4 Lower Bounds on Formulae Size in Knowledge Representation

Readers familiar with the field of knowledge representation [64] and reasoning about knowledge, and epistemic logic [18], [63] know that there is no general agreement on the “right” formalism for modelling and reasoning purposes in this area. Therefore, depending on the task at hand, various formal systems (or fragments of them) like propositional logic, first-order logic, description logics, modal and temporal logics, non-monotonic logics, etc., are used. Hence, in a sense, the discussion in the previous subsections is relevant to knowledge representation, too. Here we will mention a small number of papers containing lower bound results that were explicitly motivated by knowledge representation problems although, mathematical speaking, some of them are best understood as results about Boolean and modal logic.

The most thorough study of lower bounds on formula size in knowledge representation formalisms of which we are aware is [25] that builds upon results and techniques from [9] and [10]. It is argued there that comparing two such formalisms L_1 and L_2 in terms of their representational succinctness (i.e., establishing if there are formulae in L_1 such that all equivalent formulae from L_2 are significantly longer) is sometimes much more meaningful than comparing L_1 and L_2 in terms of expressivity or computational complexity. It is worth pointing out that many of the results in [25] are conditional, i.e., their main ingredients are unproven computational complexity conjectures.

An extensive study on lower bounds on formula-size in the setting of belief revision is presented in [7]. The authors consider the size of a propositional theory T_1 that is the result of revising a propositional theory T with a propositional formula φ. They prove that some of the formalisation of belief revision necessarily lead to the conclusion that the size of T_1 is exponential in the size of T and φ. As in the case of [25], the proofs in [7] rely on unproven computational conjectures. However, it is shown that this is justified to a certain extend because providing unconditional proofs of their results is equivalent to showing, among other things, that the complexity class NP is different from coNP. Another interesting observation made in the paper is that, if our present notion of belief revision leads to an exponential blow up of the revised theory, then,

> ...from the cognitive point of view it is questionable to assume belief revision as the evolutionary model of an agent's mind: An agent would either need an unreasonable amount of storing space, or change the format it uses to represent knowledge. ([7], p.4)

This quote shows that proving lower bounds on formula size can be used not only for obtaining new mathematical results but for supporting certain philosophical views, too.

The study of planning formalisms is another area that is related to Knowledge representation. Unlike the traditional definition that one logic L_2 is at least as expressive as another L_1 if and only if for every L_1 formula φ there is an equivalent formula in L_2, it seems that there is no generally accepted notion of relative expressive power in this context. We recommend [46] as an introduction to this type of problems. It is argued there that, intuitively, a planning formalism P_2 is at least as expressive as another one P_1 if and only if planning domains and plans formulated in P_1 have polynomially sized equivalents in P_2. Therefore, the size of the equivalent plans is an important part of the definition of expressive power. Several planning formalisms, all of them based on propositional logic, are compared in [46] with respect to this definition of expressivity. Some of the obtained lower bound results on formula-size again rely on computational complexity conjectures while the rest are unconditional.

Another paper dedicated to planning problems that is explicitly concerned with lower bounds on the size of the formal representation of plans is [39]. The authors connect knowledge-based programs as defined in [18] to knowledge based programs that have a special form called standard policies and prove, among other things, that both formalisms are equally expressive but there are knowledge based programs such that their shortest equivalent standard policies are exponentially bigger.

A relatively recent area of knowledge representation is the field of knowledge compilation (see, for example, [6]). The guiding idea on which it is based is that, since propositional reasoning is, in general, computationally intractable, then it is better to translate or compile a propositional theory off-line into a space-efficient language, which can be used then on-line for answering a large number of queries in polynomial time. One such very efficient representation is obtained by associating with every propositional formula not a syntax tree but a directed acyclic graph. A good source of lower-bound results on the size of such graphs in this setting can be found in [13].

We already discussed [43] in the previous section in connection with lower bounds on formulae in modal logic. Given the relevance of modal logic to reasoning about knowledge, we once again encourage the reader interested in epistemic logic to consult [43] for a clear exposition of a model theoretic proof of lower bounds on the size of modal formulae.

Description logics [4] form another family of very popular formalisms for knowledge representation. Here we recommend [41] for a nice result on representational succinctness in this setting and a thorough overview of the relevant literature.

Readers interested in lower bounds on formula size in default logic, circumscription, Horn formulae, conjunctive and disjunctive normal forms as used in knowledge representation will find a lot of information and pointers to the relevant literature in [8], [9], [10], and [25].

Chapter 3

Preliminaries

In this chapter we collect some technical notions and results that are used later. We try to reach a reasonable compromise between precision and intuitive clarity and that is why we do not always give the most general definitions, statements, and proofs possible.

Unless otherwise stated, all functions we consider are mappings from $\mathbb{R}^+$ to $\mathbb{R}^+$. We use the following standard definition (see for example [12]).

Definition 3.1. For any function of one variable $g(x)$, the classes of functions $O(g(x))$ and $o(g(x))$ are

$$O(g(x)) = \{f(x) \mid \text{there exist } c, z_0 \in \mathbb{R}^+ \text{ such that } f(z) \leq cg(z) \text{ for all } z \geq z_0\};$$

$$o(g(x)) = \{f(x) \mid \text{for all } c \in \mathbb{R}^+ \text{ there is a } z_0 \in \mathbb{R}^+ \text{ so that } f(z) < cg(z) \text{ for all } z \geq z_0\}.$$

Intuitively, $O(g(x))$ is the class of functions $f(x)$ that do not grow asymptotically faster than $g(x)$ while $o(g(x))$ consists of all the functions $f(x)$ that grow strictly more slowly than $g(x)$.

The class $poly(x)$ of single-variable polynomial functions is defined as

$$poly(x) = \bigcup_{c \in \mathbb{N} \text{ and } c \geq 1} O(x^c).$$

The class of exponential functions is denoted *EXP* and is identified with the class $2^{poly(x)}$. The class of sub-exponential functions, denoted *SUBEXP*, is identified with the class $2^{o(x)}$.

Example 3.1. *Intuitively, we call functions like* $2^{\frac{1}{3}x}$, $2^{3x^4+x^2+1}$, *etc., exponential while functions like* $\log_k(x)$, $3x^4 + x^2 + 1$, $2^{\sqrt[3]{x}}$ *are called sub-exponential. Another useful way of thinking about a sub-exponential function* $f(x)$ *is that it becomes insignificantly small when compared to* 2^x *as* x *grows to infinity or*

$$\lim_{x \to \infty} \frac{f(x)}{2^x} = 0.$$

3.1 On the Definition of Succinctness

In the next chapter we are going to compare a number of different modal logics in terms of the length of their formulae and we will see that some of these logics offer much more succinct way of expressing certain semantic properties than others. However, in this sub-section we give definitions and prove general results that apply to basically any type of logic. That is why we postpone talking about modal logics and related notions and begin with a general discussion of the concept of succinctness as applied to the study of formal systems.

We begin by giving an abstract definition of the notion of "logic" that is of utmost generality and may seem too broad to the reader accustomed to the more precise versions given in the literature, cf., [31] Chapter 7, Section 6.

Intuitively, we treat logics as triples consisting of a set of formulae a class of models and a binary truth relation between formulae and models.

Definition 3.2 (Logic)**.** A logic $\mathrm{L} = (\Phi, \models, \mathbb{M})$ is a triple where Φ is a non-empty set of objects called **formulae**, $\mathbb{M}$ is a non-empty class of objects called **models**, and $\models \subseteq \mathbb{M} \times \Phi$ is a binary relation called **truth relation**. If $(\mathscr{M}, \varphi) \in \models$, we write $\mathscr{M} \models \varphi$ and say that the formula φ is true in the model $\mathscr{M}$.

Note that we have not defined the set of formulae Φ and the truth relation $\models$. They are treated as parameters that will be given specific interpretations later. Our only assumption at this point is:

> Formulae are finite words over a countable alphabet Σ and the length of any formula φ, denoted $|\varphi|$, is the length of the word φ over Σ;

Of course, later, when we study specific logics, we will formally define formulae, the truth relation, the class of models, and formula length.

Definition 3.3 (Expressivity)**.** Let $\mathrm{L}_1 = (\Phi_1, \models_1, \mathbb{M}_1)$ and $\mathrm{L}_2 = (\Phi_2, \models_2, \mathbb{M}_2)$ be two logics. For any formulae $\varphi_1 \in \Phi_1$ and $\varphi_2 \in \Phi_2$ and a non-empty class $\mathbb{M}$ of models such that $\mathbb{M} \subseteq \mathbb{M}_1 \cap \mathbb{M}_2$,

> we say that the formula φ_1 is equivalent to φ_2 on $\mathbb{M}$ iff for every $\mathscr{M} \in \mathbb{M}$, it is true that $\mathscr{M} \models_1 \varphi_1$ iff $\mathscr{M} \models_2 \varphi_2$;
>
> L_2 is said to be at least as expressive as L_1 on $\mathbb{M}$ iff for every formula $\varphi_1 \in \Phi_1$, there is an equivalent on $\mathbb{M}$ formula $\varphi_2 \in \Phi_2$;
>
> we say that L_1 and L_2 are equally expressive on $\mathbb{M}$ iff L_1 is at least as expressive as L_2 on $\mathbb{M}$ and vice versa.

We begin to work our way to the main notion in this sub-section, namely the concept of **succinctness**. Although results saying that one logic is exponentially (doubly-exponentially, non-elementarily, etc.) more succinct than another are well-known, as far as we are aware, the first general definition of these notions was given in [26] and [27]. Roughly speaking succinctness is defined there as a refinement of expressivity in the following way.

Definition 3.4 (Succinctness [26, 27])**.** Let L_1 and L_2 be two logics and let F be a class of functions. We say that L_1 is F-succinct in L_2 on the class of models $\mathbb{M}$ if and only if there is a function $f \in F$ such that for every φ from L_1, there is an equivalent on $\mathbb{M}$ formula ψ from L_2 such that $|\psi| \leq f(|\varphi|)$.

Intuitively, when we say that L_1 is F-succinct in L_2 on $\mathbb{M}$, not only do we mean that L_2 is at least as expressive as L_1 on $\mathbb{M}$, but, in addition, we can give an F-upper bound on the size of L_2-formulae needed to express all of L_1 on $\mathbb{M}$. Hence, if, for example, the length of the L_2-formulae expressing all of L_1 on $\mathbb{M}$ cannot be bounded from above by a sub-exponential function, i.e., L_1 is not *SUBEXP*-succinct in L_2 on $\mathbb{M}$, we say that L_1 is exponentially more succinct than L_2 on $\mathbb{M}$.

This is a very nice general definition but we feel that some of its features can be improved. In order to make our reasoning explicit, we restate it in a semi-formal language. The notion "L_1 is F-succinct in L_2 on $\mathbb{M}$" is defined as

$$\exists f \in F\ \forall \varphi \in L_1\ \exists \psi \in L_2((\varphi \text{ is equivalent to } \psi \text{ on } \mathbb{M}) \text{ and } |\psi| \leq f(|\varphi|)).$$

Hence, if following [26, 27], we define "L_1 is exponentially more succinct than L_2 on $\mathbb{M}$" as L_1 is not sub-exponentially succinct in L_2 on $\mathbb{M}$, we obtain from the above

$$\forall f \in \mathit{SUBEXP}\ \exists \varphi \in L_1\ \forall \psi \in L_2((\varphi \text{ is equivalent to } \psi \text{ on } \mathbb{M}) \text{ implies } |\psi| > f(|\varphi|)).$$

Note that this definition is trivially true in the case when there is, e.g., one L_1 formula φ for which there is no equivalent formula in L_2. Of course, we can avoid this situation by imposing the requirement that L_2 be at least as expressive as L_1 on $\mathbb{M}$. This, however, would seem restrictive for the following reasons. It is possible for L_1 to be exponentially more succinct than L_2 on $\mathbb{M}$ and vice versa[1], and hence, in order to have such a result, we have to stipulate that L_1 and L_2 must be equally expressive on $\mathbb{M}$. On the other hand, to quote from [27]

> Succinctness as a measure for comparing the strength of logics is not restricted to logics of the same expressive power. Even if a logic L_1 is more expressive than a logic L_2, it is interesting to know whether those properties that can be expressed in both L_1 and L_2 can be expressed more succinctly in one of the logics. Sometimes, this may even be more important than the fact that some esoteric property is expressible in L_1, but not in L_2.

Using this quotation we suggest below a new version of Definition 3.4.

From now on, unless otherwise stated, we consider only non-empty sets of formulae, classes of models, and classes of functions. We begin with the following useful definition.

[1]We will exhibit such reasults in the next chapter.

Definition 3.5. Let $L_1 = (\Phi_1, \models_1, \mathbb{M}_1)$ and $L_2 = (\Phi_2, \models_2, \mathbb{M}_2)$ be two logics, $\Phi' \subseteq \Phi_1$ and $\Phi'' \subseteq \Phi_2$ be sets of formulae and let $\mathbb{M}$ be a class of models such that $\mathbb{M} \subseteq \mathbb{M}_1 \cap \mathbb{M}_2$.

> We write $\Phi' \leq_{\mathbb{M}} \Phi''$ iff for every L_1-formula $\varphi_1 \in \Phi'$, there is an equivalent on $\mathbb{M}$ L_2-formula $\varphi_2 \in \Phi''$.
>
> We write $\Phi' \equiv_{\mathbb{M}} \Phi''$ iff both $\Phi' \leq_{\mathbb{M}} \Phi''$ and $\Phi'' \leq_{\mathbb{M}} \Phi'$ are true.

Intuitively, we use Φ' and Φ'' as witnesses to the fact that some properties P of a class of models $\mathbb{M}$ in which we are interested are expressible in both L_1 and L_2[2].

We are ready to state our version of Definition 3.4.

Definition 3.6 (Succinctness)**.** Let $L_1 = (\Phi_1, \models_1, \mathbb{M}_1)$ and $L_2 = (\Phi_2, \models_2, \mathbb{M}_2)$ be two logics. For any sets of formulae $\Phi' \subseteq \Phi_1$ and $\Phi'' \subseteq \Phi_2$, any set of models $\mathbb{M}$ such that $\mathbb{M} \subseteq \mathbb{M}_1 \cap M_2$, and $\Phi' \leq_{\mathbb{M}} \Phi''$, and any class of functions F, we say that the set of L_1-formulae Φ' is F-succinct in the set of L_2-formulae Φ'' on the class of models $\mathbb{M}$ iff there is a function $f \in F$ such that for every $\varphi' \in \Phi'$ there is an equivalent on $\mathbb{M}$ formula $\varphi'' \in \Phi''$ such that $|\varphi''| \leq f(|\varphi'|)$.

Example 3.2. *In the previous chapter, when we considered Boolean logic in the De Morgan basis* **DM** *and Boolean logic in the basis* **Mod2***, we saw that for any formula in the set* $L_n = \{l_2, l_3, \dots,\} \subset \Phi(P, \mathbf{Mod2})$*, there is an equivalent formula in* $DF(P) \subset \Phi(P, \mathbf{DM})$ *of exponential size. Therefore, the set* L_n *is EXP-succinct in the set* $DF(P)$*.*

Using Definition 3.6, it is easy to see now what we mean when we say, for two sets of formulae Φ' and Φ'' where $\Phi' \leq_{\mathbb{M}} \Phi''$, that Φ' is **not** F-succinct in Φ'' on $\mathbb{M}$. Intuitively, we claim that there is no translation function $\tau : \Phi' \to \Phi''$ for which there is a function $f \in F$ such that, for all $\varphi' \in \Phi'$, we have that $\tau(\varphi'') \in \Phi''$ is equivalent to φ' on $\mathbb{M}$ and, at the same time, $|\tau(\varphi'')|$ is bounded from above by $f(|\varphi'|)$. Of course, in the case when Φ' is *EXP*-succinct in Φ'' but it is impossible to find a translation from Φ' to Φ'' that always produces equivalent formulae of sub-exponential length, i.e., when Φ' is not *SUBEXP*-succinct in Φ'' on $\mathbb{M}$, we are going to use the popular terminology and simply say that Φ' is exponentially more succinct than Φ'' on $\mathbb{M}$.

Definition 3.7. Let the sets Φ', Φ'' and $\mathbb{M}$ be as in Definition 3.6. We say that

- Φ' is at least exponentially more succinct than Φ'' on $\mathbb{M}$ iff Φ' is not *SUBEXP*-succinct in Φ'' on $\mathbb{M}$
- Φ' is exponentially more succinct than Φ'' on $\mathbb{M}$ iff Φ' is *EXP*-succinct but not *SUBEXP*-succinct in Φ'' on $\mathbb{M}$.

[2] We avoid giving an abstract definition of the notion of property because this will unnecessarily complicate our exposition; moreover, we are not strictly interested in the question whether a given property is expressible by a formula from a certain logic or not (of course, such investigation cannot be carried out without a definition of "property") but in the comparative lengths of formulae from two logics L_1 and L_2 for which we know that they express the same property on a given class of models $\mathbb{M}$. In the latter situation we can avoid talking about properties by saying that we are interested in the comparative lengths of the formulae from L_1 and L_2 that are equivalent on $\mathbb{M}$.

Let us list some easy consequences of Definition 3.6 and Definition 3.7.

The next proposition follows immediately from Definition 3.6. One of its obvious corollaries is that if a set of formulae Φ' is exponentially more succinct than another set of formulae Φ'' on a class of models $\mathbb{M}$, then Φ' remains exponentially more succinct than Φ'' on any class $\mathbb{O}$ such that $\mathbb{M} \subseteq \mathbb{O}$ and Φ'' is at least as expressive as Φ' on $\mathbb{O}$.

Proposition 3.8. *For any logics* L_1 *and* L_2 *and any class of functions* F*, if the set of* L_1*-formulae* Φ' *is not* F*-succinct in the set of* L_2*-formulae* Φ'' *on the class of models* $\mathbb{M}$ *and, if* Φ'_1 *is a set of* L_1*-formulae,* Φ''_2 *is a set of* L_2*-formulae, and* $\mathbb{O}$ *is a class of models such that*

- $\Phi' \subseteq \Phi'_1$,
- $\Phi''_2 \subseteq \Phi''$,
- $\mathbb{M} \subseteq \mathbb{O}$,
- $\Phi'_1 \leq_{\mathbb{O}} \Phi''_2$,

then Φ'_1 *is not* F*-succinct in* Φ''_2 *on* $\mathbb{O}$.

Since we are interested in the asymptotic behaviour of the gap between the lengths of formulae from Φ' and their equivalents from Φ'', if we want to prove interesting[3] results like, e.g., that this gap grows exponentially, we need both sets Φ' and Φ'' to be infinite. If Φ' were a finite set $\{\varphi'_1, \ldots, \varphi'_j, \ldots, \varphi'_n\}$, then, assuming without loss of generality that $\Phi''_n = \{\varphi''_1, \ldots, \varphi''_j, \ldots, \varphi''_n\} \subseteq \Phi''$ were the set of equivalent formulae such that φ'_i is equivalent to φ''_i for $1 \leq i \leq n$, and φ''_j is the longest formula in Φ''_n, it is obvious that Φ' is $CONST$-succinct in Φ'' on $\mathbb{M}$ where $CONST$ is the class of constant functions. Indeed, let us consider the function $f(x) = k$, where $k = |\varphi''_j|$, i.e., k is the length of the longest formula in Φ''_n, namely $|\varphi''_j|$. Then for every formula $\varphi \in \Phi'$, there is an equivalent on $\mathbb{M}$ formula $\psi \in \Phi''$ such that $|\psi| \leq f(|\varphi|)$. The proposition below shows that we need not only infinitely many formulae but infinitely many **pairwise non-equivalent** formulae in Φ' if we want to show an exponential succinctness gap because, otherwise, we are essentially back in the case for a finite Φ'.

Proposition 3.9. *Let* $\mathrm{L}_1 = (\Phi_1, \models_1, \mathbb{M}_1)$ *and* $\mathrm{L}_2 = (\Phi_2, \models_2, \mathbb{M}_2)$ *be two logics. For any sets of formulae* $\Phi' \subseteq \Phi_1$ *and* $\Phi'' \subseteq \Phi_2$*, any set of models* $\mathbb{M}$ *such that* $\mathbb{M} \subseteq \mathbb{M}_1 \cap M_2$ *and* $\Phi' \leq_{\mathbb{M}} \Phi''$*, if* Φ' *is at least exponentially more succinct than* Φ''*, then there are infinitely many pairwise non-equivalent on* $\mathbb{M}$ *formulae in* Φ'.

Proof. Suppose that there are only finitely many non-equivalent on $\mathbb{M}$ formulae $\varphi'_i, \ldots, \varphi'_k$ among the formulae in Φ'. Let $\varphi''_i, \ldots, \varphi''_k$ be the respective equivalent on $\mathbb{M}$ formulae in Φ''. It follows that any formula $\varphi' \in \Phi'$ is equivalent on $\mathbb{M}$ to one of the formulae $\varphi''_i, \ldots, \varphi''_k$. Therefore, for any natural number $n \in \mathbb{N}$ that is equal or greater to the length of the longest formula among the formulae $\varphi''_i, \ldots, \varphi''_k$, we have that Φ' is f-succinct in Φ'', where $f : \mathbb{N} \to \mathbb{N}$ is the constant function $f(x) = n$. □

[3]Here, "interesting" means "bigger than a constant".

We can use Proposition 3.9 to define the notion "exponentially more succinct" in a slightly different (but equivalent) way than the one found in Definition 3.7. Namely, we can proceed as follows.

> The set of formulae Φ is at least exponentially more succinct than the set of formulae Ψ on the class of models $\mathbb{M}$ iff Φ is *EXP*-succinct in Ψ on $\mathbb{M}$ and there is an infinite sequence of pairwise non-equivalent on $\mathbb{M}$ formulae in Φ
>
> $$\varphi_1,\ \varphi_2,\ \varphi_3,\ \ldots$$
>
> and an infinite sequence of formulae in Ψ
>
> $$\psi_1,\ \psi_2,\ \psi_3,\ \ldots$$
>
> for which the following are true.
>
> 1. There is a function $f(x)$ such that $|\varphi_n| \leq f(n)$ for every $n \in \mathbb{N}$.
> 2. There is a function $g(x)$ such that $|\psi_n| \geq g(n)$ for every $n \in \mathbb{N}$.
> 3. For every $n \in \mathbb{N}$, we have that ψ_n is the shortest formula in Ψ that is equivalent to φ_n on $\mathbb{M}$.
> 4. For every sub-exponential function $\tau(x) \in SUBEXP$, there is an $n \in \mathbb{N}$ such that $\tau(f(n)) < g(n)$.

All the succinctness results of which we are aware proceed by specifying two infinite sequence as above and showing that

- the length of formulae are indeed bounded by a suitable pair of functions $f(x)$ and $g(x)$ as specified in items 1, 2, and 4, where, e.g, $f(x)$ may be linear in x while $g(x)$ is at least exponential, or $f(x)$ is logarithmic whereas $g(x)$ is at least linear (the proof that $f(x)$ and $g(x)$ satisfy item 4 is almost always omitted because it usually follows from elementary calculus considerations);
- item 3 is satisfied which is the most difficult part.

For the sake of completeness and clarity however, we capture the general shape of a large number of such arguments in Lemma 3.10 below.

Lemma 3.10. *For any logics* $L_1 = (\Phi_1, \models_1, \mathbb{M}_1)$ *and* $L_2 = (\Phi_2, \models_2, \mathbb{M}_2)$*, any sets of formulae* $\Phi' \subseteq \Phi_1$ *and* $\Phi'' \subseteq \Phi_2$*, and any class of models* $\mathbb{M} \subseteq \mathbb{M}_1 \cap \mathbb{M}_2$ *such that* $\Phi' \leq_{\mathbb{M}} \Phi''$*, the set* Φ' *is at least exponentially more succinct than* Φ'' *on* $\mathbb{M}$ *if the following conditions are fulfilled.*

1. *There are functions* $f(x) : \mathbb{N} \to \mathbb{R}^+$ *and* $g(x) : \mathbb{R} \to \mathbb{R}$ *such that* $f(x) \in O(g(x))$*.*
2. *There is an infinite sequence of formulae* $\varphi'_1, \varphi'_2, \varphi'_3 \ldots$ *of strictly increasing length (i.e.,* $|\varphi'_i| < |\varphi'_{i+1}|$*) in* Φ' *and an infinite sequence of formulae* $\varphi''_1, \varphi''_2, \varphi''_3 \ldots$ *in* Φ'' *satisfying the property:*

For every $n \in \mathbb{N}$,

(a) $|\varphi'_n| \leq f(n)$*;*

(b) $|\varphi''_n| \geq 2^{g(n)}$*;*

(c) φ''_n *is the shortest formula in* Φ'' *that is equivalent to* φ'_n *on* $\mathbb{M}$*.*

Proof. We have to prove that for every function $i(x) \in SUBEXP$ there is a formula $\varphi'_{\mathbf{i}} \in \Phi'$ such that $|\varphi''| > i(|\varphi'_{\mathbf{i}}|)$ for any formula $\varphi'' \in \Phi''$ that is equivalent to $\varphi'_{\mathbf{i}}$ on $\mathbb{M}$.

Since $f(x) \in O(g(x))$, there are natural numbers n_0 and c such that $f(n) \leq cg(n)$ for all $n \geq n_0$. Hence $\frac{1}{c}f(n) \leq g(n)$ for all $n \geq n_0$. Suppose now that $i(x) \in SUBEXP$. Then there is a natural number m_0 such that $i(m) < 2^{\frac{1}{c}m}$ for all $m \geq m_0$. Since, for every $n \in \mathbb{N}$, it is true that $|\varphi'_n| < |\varphi'_{n+1}|$, there is a natural number $\mathbf{i} \geq n_0$, such that $f(\mathbf{i}) \geq |\varphi'_{\mathbf{i}}| \geq m_0$. Therefore, $i(|\varphi'_{\mathbf{i}}|) < 2^{\frac{1}{c}|\varphi'_{\mathbf{i}}|}$. Now, the statement follows from the condition that $\varphi''_{\mathbf{i}}$ is the shortest formula that is equivalent to $\varphi'_{\mathbf{i}}$, the fact that $|\varphi''_{\mathbf{i}}| \geq 2^{g(\mathbf{i})}$, and $g(\mathbf{i}) \geq \frac{1}{c}f(\mathbf{i}) \geq \frac{1}{c}|\varphi'_{\mathbf{i}}|$. □

Note that the condition $f(x) \in O(g(x))$ in the statement of the lemma is essential and cannot be replaced by $g(x) \in O(f(x))$. To see this, let us consider the following example:

Example 3.3. *Let us suppose that for every $n \in \mathbb{N}$, there are two formulae $\varphi'_n \in \Phi'$ and $\varphi''_n \in \Phi''$ satisfying the properties:*

$1'$. $|\varphi'_n| = n^3$*;*

$2'$. $|\varphi''_n| = 2^n$*;*

$3'$. φ''_n *is the shortest formula in* Φ'' *that is equivalent to* φ'_n *on* $\mathbb{M}$*.*

It is obvious that the linear function $f(x) = x$ is in the class $O(x^3)$. In this case however, we cannot use these two sequences of formulae to prove that there is an exponential succinctness gap between the sets of formulae Φ' and Φ'' because the sub-exponential function $2^{\sqrt[3]{x}}$ bounds the length of each φ''_n in the length of φ'_n, i.e., $\varphi''_n = 2^{\sqrt[3]{|\varphi'_n|}}$. Hence, we must impose a stronger condition on the length of $|\varphi''_n|$, namely, $|\varphi''_n| = 2^{f(n)}$ where $x^3 \in O(f(x))$.

Some articles (e.g. [40], [43], [44], [45], [51]) use implicitly instances of Lemma 3.10, while others define explicitly “exponentially more succinct” as some particular instance of it. For example, the following definition is used in [58].

If two languages, L_1 and L_2, are equally expressive, we say that L_1 is exponentially more succinct than L_2 if there is an infinite sequence of L_1 expressions $R_1, R_2, \ldots$ where the length of R_k is polynomial in k, such that for every sequence of equivalent L_2-expressions $P_1, P_2, \ldots$, the length of P_k is exponential in k.

Undoubtedly, such a definition provides a sufficient condition[4] for L_1 to be exponentially more succinct than L_2, but, unfortunately, it does not fit results like the one in,

[4]Provided that we are careful what “exponential in k” means in this case as explained in Example 3.3.

e.g., [51], where, in the wording above, the length of the expression R_k is logarithmic in k while the length of P_k is linear in k. Of course, this case is covered by Lemma 3.10.

We finish this subsection with some comments on Definition 3.6. We think that defining succinctness in terms of sets of formulae instead of in terms of logics is closer to our intuitive understanding of the meaning of this notion and helps us avoid, e.g., saying that L_1 is exponentially more succinct than L_2 on $\mathbb{M}$ and L_2 is exponentially more succinct than L_1 on $\mathbb{M}$ when we have such results. Although this a perfectly legitimate statement, we still think it is somewhat unpleasant and that the situation is better described by saying that there is a set of L_1-formulae Φ' such that the lengths of the equivalent L_2-formulae from the set Φ'' cannot be bounded from above by a sub-exponential function and vice versa; moreover, note that this formalisation makes it possible, by imposing relevant restrictions on Φ' and Φ'', to study refined versions of succinctness problems, e.g., when one of Φ' or Φ'' consists of formulae that have some specific syntactic shapes. For example, we can differentiate between the case where Φ'' is the set of all L_2-formulae (e.g., Boolean formulae in the De Morgan basis), and the case where Φ'' is a subset of L_2-formulae that have some specific property (e.g., the set of formulae in disjunctive normal form is a subset of the set of Boolean formulae in the De Morgan basis).

3.2 Four Extensions of Multimodal Logic

In this subsection, we define the modal logics that we are going to compare in terms of their succinctness in the rest of this thesis. The interest in the succinctness of different modal formalisms stems from their wide applicability. Many important intuitive concepts like “knowledge”, “belief”, “run of a computer program”, “temporal relations between different events”, etc., can be modelled as edge and vertex coloured graphs plus a suitably chosen formal language that is used to talk about such graphs. Multimodal logic [5], denoted ML, provides us with one such popular language that extends Boolean logic with formulae of the form $[i]\varphi$. The formulae of ML are evaluated locally at a designated node of a given edge and vertex coloured graph which in this setting is called a Kripke model. Intuitively, $[i]\varphi$ is true in a vertex s if every vertex reachable from s in one step along an i-coloured edge satisfies the formula φ. As we said above, given the applicability of such graphs to the modelling of interesting natural phenomena, two vertices v_1 and v_2 connected with an i-edge can represent, among other things, two epistemically indistinguishable for an agent i situations (this is the road taken in epistemic logic [18, 62]), a transition from a state v_1 to a new state v_2 caused by program i (as used in dynamic logic of programs [29]), or, in description logic [4], a role like v_2 is a brother (sister) of v_1, etc. In its abstract form, the set of pairs of vertices that are connected by an i-coloured edge is just a relation R_i on the underlying set of the graph. Therefore, a natural way of generalising this line of reasoning (as this is done in Boolean modal logic [24]) is to introduce formulae capturing various operations on relations, for example,

$[\cup_\Gamma]\varphi$ for every subset Γ of the colours used as labels for the edges of the graph. Then, a formula $[\cup_\Gamma]\varphi$ can be interpreted as true in vertex s if making one step along an edge coloured with an element from Γ always leads to a vertex t where φ is true. One can easily see that this is the same as requiring that, for all $i \in \Gamma$, the formula $[i]\varphi$ is true in s. In description logic, $[\cup_\Gamma]\varphi$ can be used to say that φ is true about all the siblings of an individual s where we define siblings as the union of brothers and sisters of s. In dynamic logic, $[\cup_\Gamma]\varphi$ can be used to express the fact that no matter which program from the set of programs Γ is executed at s, the resulting state t will satisfy φ. In epistemic logic, $[\cup_\Gamma]\varphi$ is interpreted as "everybody in the group of agents Γ knows that φ is true" (see, for example, [55] or[18]). We use $[\cup]$ML to denote ML extended with formulae $[\cup_\Gamma]\varphi$.

Analogously, we can introduce formulae $[\exists_\Gamma]\varphi$. Such a formula is true in a vertex s if there is an $i \in \Gamma$ such that making one step from s along an i-coloured vertex always leads to a vertex t where φ is true. In other words, there is an i in Γ that makes the formula $[i]\varphi$ true at s. In epistemic logic, such formulae are used to express that "somebody in the group of agents Γ knows φ" ([55]). In description logic, we can use $[\exists_\Gamma]\varphi$ to say, e.g., some of s' friends are males. ML extended with formulae $[\exists_\Gamma]\varphi$ will be denoted $[\exists]$ML.

Another natural operator on graphs is the intersection modality $[\cap_\Gamma]$. Intuitively, a formula $[\cap_\Gamma]\varphi$ is true in a vertex s when any vertex t that is R_i related to s for all the possible $i \in \Gamma$ satisfy φ. In epistemic logic, this yields the notion of distributed knowledge (see, for example, [18] Chapter 2) while, in dynamic logic, $[\cap_\Gamma]$ can be used to formalise the notion of parallel execution of the programs in Γ. Let $[\cap]$ML denote ML extended with formulae $[\cap_\Gamma]\varphi$.

Finally, we are interested in formulae of the form $[\varphi]\psi$. Intuitively, $[\varphi]\psi$ is true in a vertex s if, after removing all vertices that do not satisfy the formula φ, the formula ψ is true at s in the resulting new graph. Formulae $[\varphi]\psi$ are used in dynamic epistemic logic [62] to formalise the intuitive notion "after the public announcement of the fact expressed by φ, the fact expressed by ψ will be true". We denote ML extended with formulae $[\varphi]\psi$ by $[\varphi]$ML.

Now, we proceed by formally defining the logics described above.

Definition 3.11 (Formulae)**.** A signature is a pair $S = (P, I)$, where $P = \{p_1, p_2 \ldots\}$ is a countable set of propositional symbols and $I = \{i_1, i_2, \ldots, i_n, \ldots\}$ is a countable set of relation indices. Let p, i, and Γ vary over the elements of P, I, and the non-empty finite subsets of I, respectively. The formulae of the Multimodal logic ML and its extensions $[\cup]$ML, $[\cap]$ML, $[\exists]$ML, and $[\varphi]$ML in the signature $S = (P, I)$ are built as follows.

- The set Φ_{ML} of formulae of Multimodal Logic ML, is the smallest set such that
  - $p \in \Phi_{\mathrm{ML}}$ for any $p \in P$;
  - if $\psi \in \Phi_{\mathrm{ML}}$, then $\neg\psi \in \Phi_{\mathrm{ML}}$;
  - for any $\psi_1 \in \Phi_{\mathrm{ML}}$ and $\psi_2 \in \Phi_{\mathrm{ML}}$, we have $(\psi_1 \vee \psi_2) \in \Phi_{\mathrm{ML}}$;

  - for all $i \in I$, if $\psi \in \Phi_{\mathrm{ML}}$, then $[i]\psi \in \Phi_{\mathrm{ML}}$.

- The set of $[\cup]$ML-formulae is the smallest set $\Phi_{[\cup]\mathrm{ML}}$ such that
  - $p \in \Phi_{[\cup]\mathrm{ML}}$ for any $p \in P$;
  - if $\psi \in \Phi_{[\cup]\mathrm{ML}}$, then $\neg\psi \in \Phi_{[\cup]\mathrm{ML}}$;
  - for any $\psi_1 \in \Phi_{[\cup]\mathrm{ML}}$ and $\psi_2 \in \Phi_{[\cup]\mathrm{ML}}$, we have $(\psi_1 \vee \psi_2) \in \Phi_{[\cup]\mathrm{ML}}$;
  - for all $i \in I$, if $\psi \in \Phi_{[\cup]\mathrm{ML}}$, then $[i]\psi \in \Phi_{[\cup]\mathrm{ML}}$;
  - for all Γ, if $\psi \in \Phi_{[\cup]\mathrm{ML}}$, then $[\cup_\Gamma]\psi \in \Phi_{[\cup]\mathrm{ML}}$.

- The set of formulae $\Phi_{[\cap]\mathrm{ML}}$ of the logic $[\cap]$ML is the smallest set
  - $p \in \Phi_{[\cap]\mathrm{ML}}$ for any $p \in P$;
  - if $\psi \in \Phi_{[\cap]\mathrm{ML}}$, then $\neg\psi \in \Phi_{[\cap]\mathrm{ML}}$;
  - for any $\psi_1 \in \Phi_{[\cap]\mathrm{ML}}$ and $\psi_2 \in \Phi_{[\cap]\mathrm{ML}}$, we have $(\psi_1 \vee \psi_2) \in \Phi_{[\cap]\mathrm{ML}}$;
  - for all $i \in I$, if $\psi \in \Phi_{[\cap]\mathrm{ML}}$, then $[i]\psi \in \Phi_{[\cap]\mathrm{ML}}$;
  - for all Γ, if $\psi \in \Phi_{[\cap]\mathrm{ML}}$, then $[\cap_\Gamma]\psi \in \Phi_{[\cap]\mathrm{ML}}$.

- The set $\Phi_{[\exists]\mathrm{ML}}$ of formulae of the logic $[\exists]$ML is the smallest set with the properties
  - $p \in \Phi_{[\exists]\mathrm{ML}}$;
  - if $\psi \in \Phi_{[\exists]\mathrm{ML}}$, then $\neg\psi \in \Phi_{[\exists]\mathrm{ML}}$;
  - for any $\psi_1 \in \Phi_{[\exists]\mathrm{ML}}$ and $\psi_2 \in \Phi_{[\exists]\mathrm{ML}}$, we have $(\psi_1 \vee \psi_2) \in \Phi_{[\exists]\mathrm{ML}}$;
  - for all $i \in I$, if $\psi \in \Phi_{[\exists]\mathrm{ML}}$, then $[i]\psi \in \Phi_{[\exists]\mathrm{ML}}$;
  - for all Γ, if $\psi \in \Phi_{[\exists]\mathrm{ML}}$, then $[\exists_\Gamma]\psi \in \Phi_{[\exists]\mathrm{ML}}$.

- The set of formulae of the logic $[\varphi]$ML is the smallest set $\Phi_{[\varphi]\mathrm{ML}}$ such that
  - $p \in \Phi_{[\varphi]\mathrm{ML}}$;
  - if $\psi \in \Phi_{[\varphi]\mathrm{ML}}$, then $\neg\psi \in \Phi_{[\varphi]\mathrm{ML}}$;
  - $(\psi_1 \vee \psi_2) \in \Phi_{[\varphi]\mathrm{ML}}$ for any $\psi_1 \in \Phi_{[\varphi]\mathrm{ML}}$ and $\psi_2 \in \Phi_{[\varphi]\mathrm{ML}}$;
  - for all $i \in I$, if $\psi \in \Phi_{[\varphi]\mathrm{ML}}$, then $[i]\psi \in \Phi_{[\varphi]\mathrm{ML}}$;
  - if $\psi_1 \in \Phi_{[\varphi]\mathrm{ML}}$ and $\psi_2 \in \Phi_{[\varphi]\mathrm{ML}}$, then $[\psi_1]\psi_2 \in \Phi_{[\varphi]\mathrm{ML}}$.

As usual, we have the following abbreviations $\varphi \wedge \psi \stackrel{\text{def}}{=} \neg(\neg\varphi \vee \neg\psi)$, $\langle i\rangle\varphi \stackrel{\text{def}}{=} \neg[i]\neg\varphi$, $\langle\cup_\Gamma\rangle\varphi \stackrel{\text{def}}{=} \neg[\cup_\Gamma]\neg\varphi$, $\langle\cap_\Gamma\rangle\varphi \stackrel{\text{def}}{=} \neg[\cap_\Gamma]\neg\varphi$, $\langle\exists_\Gamma\rangle\varphi \stackrel{\text{def}}{=} \neg[\exists_\Gamma]\neg\varphi$, and $\langle\varphi\rangle\psi \stackrel{\text{def}}{=} \neg[\varphi]\neg\psi$.

Definition 3.12 (Length of Formulae)**.** The length of a formula φ is denoted $|\varphi|$. It is defined as follows: $|p| = 1$, $|[\varphi_1]\varphi_2| = |\varphi_1 \vee \varphi_2| = |\varphi_1| + |\varphi_2| + 1$, $|\neg\varphi| = |[i]\varphi| = |[\exists_\Gamma]\varphi| = |[\cup_\Gamma]\varphi| = |[\cap_\Gamma]\varphi| = 1 + |\varphi|$.

The classes of models for these formulae and the respective truth relations are defined in the usual way (e.g., see [5]).

Definition 3.13 (Kripke Model)**.** A Kripke model for the signature $S = (P, I)$ is a triple $\mathscr{M} = (M, R, V)$, where M is a non-empty set, $R : I \to 2^{M\times M}$ is a mapping that assigns a binary relation R_i on M to every $i \in I$, and $V : P \to 2^M$ is a function that assigns a subset of M, i.e., a unari relation, to every $p \in P$.

Kripke models are denoted $\mathscr{A}$, $\mathscr{B}$, ..., etc. We write sR_it for $(s,t) \in R(i)$ and say that t is an i-successor of s or that s and t are i-connected.

The pair $(\mathscr{M}, w)$, where $w \in M$, is called pointed model. Sets of pointed models are denoted $\mathbb{A}$, $\mathbb{B}$, ... The class of all pointed models is denoted **K**. The class of all models where, for any i, the relation $R(i)$ is reflexive, symmetric, and transitive is denoted **S**5.

Next, we specify the respective truth relations corresponding to the logics from Definition 3.11. However, we are not going to be completely precise when defining these relations, because, technically speaking, we need five truth relations corresponding to the five logics ML, $[\cup]$ML, $[\exists]$ML, $[\cap]$ML, and $[\varphi]$ML. Such precision will complicate unnecessarily our exposition; moreover, it will always be clear from the context which one of the truth relations we mean.

Definition 3.14 (Truth)**.** Let φ be a formula in the signature S and let $(\mathscr{M}, w) \in \mathbf{K}$. The relation φ is true in the pointed model $(\mathscr{M}, w)$, written $(\mathscr{M}, w) \models \varphi$, is defined recursively on the structure of φ as follows.

$$
\begin{array}{lll}
(\mathscr{M}, w) \models p & \text{iff} & w \in V(p); \\
(\mathscr{M}, w) \models \neg\psi & \text{iff} & (\mathscr{M}, w) \not\models \psi; \\
(\mathscr{M}, w) \models \psi_1 \vee \psi_2 & \text{iff} & (\mathscr{M}, w) \models \psi_1 \text{ or } (\mathscr{M}, w) \models \psi_2; \\
(\mathscr{M}, w) \models [i]\psi & \text{iff} & (\mathscr{M}, v) \models \psi \text{ for all } v\text{, such that } wR_iv; \\
(\mathscr{M}, w) \models [\cup_\Gamma]\psi & \text{iff} & \text{for all } i \in \Gamma,\ (\mathscr{M}, w) \models [i]\psi; \\
(\mathscr{M}, w) \models [\cap_\Gamma]\psi & \text{iff} & (\mathscr{M}, v) \models \psi \text{ for all } v \text{ such that } wR_iv \text{ for all } i \in \Gamma; \\
(\mathscr{M}, w) \models [\exists_\Gamma]\psi & \text{iff} & \text{there is an } i \in \Gamma \text{ such that } (\mathscr{M}, w) \models [i]\psi; \\
(\mathscr{M}, w) \models [\psi_1]\psi_2 & \text{iff} & (\mathscr{M}, w) \models \psi_1 \text{ implies } (\mathscr{M}|_{\psi_1}, w) \models \psi_2, \text{ where} \\
 & & \mathscr{M}|_{\psi_1} = (M|_{\psi_1}, R|_{\psi_1}, V|_{\psi_1}) \text{ is such that} \\
 & & M|_{\psi_1} = \{v \in M | (\mathscr{M}, v) \models \psi_1\} \text{ and} \\
 & & \text{for all } i \in I \text{ and all } p \in P, \\
 & & R|_{\psi_1}(i) = R(i) \cap (M|_{\psi_1} \times M|_{\psi_1}), \\
 & & V|_{\psi_1}(p) = V(p) \cap M|_{\psi_1}.
\end{array}
$$

Intuitively, the model $\mathscr{M}|_{\psi_1}$ used to define the $\models$ relation for the formula $[\psi_1]\psi_2$ is the restriction of the model $\mathscr{M}$ to the points in which ψ_1 is true.

It is obvious that for any pointed model $(\mathscr{M}, v)$ and any formula $[\cup_\Gamma]\varphi$, we have

$$(\mathscr{M}, v) \models [\cup_\Gamma]\varphi \text{ iff } (\mathscr{M}, v) \models \bigwedge_{i\in\Gamma}[i]\varphi.$$

Similarly,

$$(\mathscr{M}, v) \models [\exists_\Gamma]\varphi \text{ iff } (\mathscr{M}, v) \models \bigvee_{i\in\Gamma}[i]\varphi.$$

The above equivalences mean that if Γ consists of a single relation index i, then the formulae $[\cup_\Gamma]\varphi$ and $[\exists_\Gamma]\varphi$ are equivalent to $[i]\varphi$. Therefore, we always assume that Γ contains at least two indices and this applies to formulae $[\cap_\Gamma]\varphi$, too. More importantly, these equivalences mean that $\Phi_{\mathrm{ML}} \equiv_{\mathbf{K}} \Phi_{[\cup]\mathrm{ML}} \equiv_{\mathbf{K}} \Phi_{[\exists]\mathrm{ML}}$.

It is perhaps a little surprising that $\Phi_{\mathrm{ML}} \equiv_{\mathbf{K}} \Phi_{[\varphi]\mathrm{ML}}$ is true, too. This follows from the equivalences below first given in [50].

$$\begin{array}{lll}
(\mathscr{M}, v) \models [\varphi]p & \text{iff} & (\mathscr{M}, v) \models \varphi \to p; \\
(\mathscr{M}, v) \models [\varphi](\psi_1 \wedge \psi_2) & \text{iff} & (\mathscr{M}, v) \models [\varphi]\psi_1 \wedge [\varphi]\psi_2; \\
(\mathscr{M}, v) \models [\varphi]\neg\psi & \text{iff} & (\mathscr{M}, v) \models \varphi \to \neg[\varphi]\psi; \\
(\mathscr{M}, v) \models [\varphi][i]\psi & \text{iff} & (\mathscr{M}, v) \models \varphi \to [i][\varphi]\psi; \\
(\mathscr{M}, v) \models [\varphi_1][\varphi_2]\psi & \text{iff} & (\mathscr{M}, v) \models [\varphi_1 \wedge [\varphi_1]\varphi_2]\psi.
\end{array} \tag{3.1}$$

If $\mathbb{A}$ is a (not necessarily non-empty) set of pointed models and φ is a formula of one of the logics above, we write $\mathbb{A} \models \varphi$ to mean $(\mathscr{M}, w) \models \varphi$ for all $(\mathscr{M}, w) \in \mathbb{A}$. This implies that if $\mathbb{A} = \emptyset$, then $\mathbb{A} \models \varphi$ is trivially true for any formula φ.

We are going to use the following well-known notion (see section 2.2 in [5]).

Definition 3.15 (Bisimulation)**.** For any two pointed models $(\mathscr{M}_1, v)$ and $(\mathscr{M}_2, w)$ for the signature $S = (I, P)$, we say that $(\mathscr{M}_1, v)$ and $(\mathscr{M}_2, w)$ are bisimilar iff there is relation B between $(\mathscr{M}_1, v)$ and $(\mathscr{M}_2, w)$ such that if $(\mathscr{M}_1, v)B(\mathscr{M}_2, w)$, then

1. $(\mathscr{M}_1, v) \models p$ iff $(\mathscr{M}_2, w) \models p$ for any propositional symbol $p \in P$;
2. for any relation index $i \in I$,
    - for any i-successor $v_1 \in \mathscr{M}_1$ of v, there is an i-successor $w_1 \in \mathscr{M}_2$ of w and $(\mathscr{M}, v_1)B(\mathscr{M}_2, w_1)$;
    - for any i-successor $w_1 \in \mathscr{M}_2$ of w, there is an i-successor $v_1 \in \mathscr{M}_1$ of v and $(\mathscr{M}_2, w_1)B(\mathscr{M}_1, v_1)$.

Of course, we have the following fundamental result with respect to the notion of bisimulation (see, for example, [5] sec. 2.2 or [63] sec. 8.3).

Theorem 3.16. *For any signature $S = (P, I)$ and any pair of pointed models $(\mathscr{M}_1, v)$ and $(\mathscr{M}_2, w)$, if $(\mathscr{M}_1, v)$ and $(\mathscr{M}_2, w)$ are bisimilar then*

$$(\mathscr{M}_1, v) \models \varphi \textit{ iff } (\mathscr{M}_2, w) \models \varphi$$

for any formula φ that belongs to at least one of the sets $\Phi_{[\cup]\mathrm{ML}}$, $\Phi_{[\exists]\mathrm{ML}}$, or $\Phi_{[\varphi]\mathrm{ML}}$.

Because of this theorem, we say that the logics ML, $[\cup]$ML, $[\exists]$ML, and $[\varphi]$ML are invariant under bisimulation. Although $[\cap]$ML is not invariant under bisimulation as defined above, a suitably extended version of Definition 3.15 that takes care of intersection of relations can be given and $[\cap]$ML becomes invariant under this extended notion of bisimulation. Later, we are going to exhibit models that are bisimilar in this respect. Next, we give a definition that will be useful in what follows. Intuitively, we describe a situation in which two models bisimulate relative to one relation but may not be bisimilar relative to another. We call this relativised bisimulation.

Example 3.4. *Let us suppose that we have a signature $S = \{P, I\}$, where P contains a single propositional symbol* **b** *that intuitively means black; I contains two indices a and d where a intuitively means solid arrow and d means dashed arrow. The pointed models $(\mathscr{O}, o)$ and $(\mathscr{P}, \rho)$ in Figure 3.1, where black points satisfy the propositional symbol* **b** *and white points do not, are bisimilar with respect to a or, in other words, with respect to the relation represented by the solid arrow but they are not bisimilar with respect to d or the relation represented by the dashed arrow. We have that $(\mathscr{O}, \rho_1)$ and $(\mathscr{P}, \rho_1)$ are bisimilar; what is more they are bisimilar with respect to $a \cap d$. The same applies to $(\mathscr{O}, o_1)$ and $(\mathscr{P}, o_1)$.*

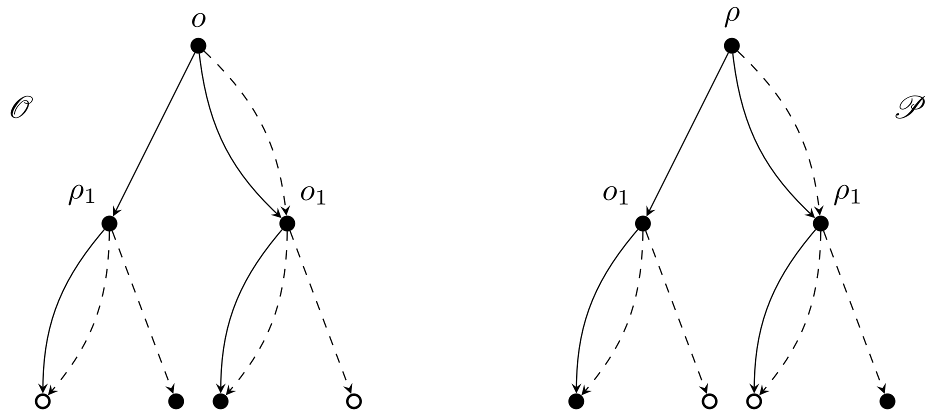


FIGURE 3.1: The pointed models $(\mathscr{O}, o)$ and $(\mathscr{P}, \rho)$.

The formal definition is as follows.

Definition 3.17 (Relativised Bisimulation)**.** For any pointed models $(\mathscr{M}_1, v)$ and $(\mathscr{M}_2, w)$ for the signature $S = \{I, P\}$ and any relation index $i \in I$, we say that $(\mathscr{M}_1, v)$ and $(\mathscr{M}_2, w)$ are bisimilar with respect to i iff there is a relation B between $(\mathscr{M}_1, v)$ and $(\mathscr{M}_2, w)$ such that if $(\mathscr{M}_1, v)B(\mathscr{M}_2, w)$, then

1. $(\mathscr{M}_1, v) \models p$ iff $(\mathscr{M}_2, w) \models p$ for any propositional symbol $p \in P$;
2. for any i-successor $v_1 \in \mathscr{M}_1$ of v, there is an i-successor $w_1 \in \mathscr{M}_2$ of w such that $(\mathscr{M}, v_1)$ and $(\mathscr{M}_2, w_1)$ are bisimilar in the sense of Definition 3.15;
3. for any i-successor $w_1 \in \mathscr{M}_2$ of w, there is an i-successor $v_1 \in \mathscr{M}_1$ of v such that $(\mathscr{M}_2, w_1)$ and $(\mathscr{M}_1, v_1)$ are bisimilar in the sense of Definition 3.15.

We have the following easy result.

Theorem 3.18. *For any signature $S = (I, P)$, any pair of pointed models $(\mathscr{M}_1, v)$ and $(\mathscr{M}_2, w)$ for S and any relation index $i \in I$, if $(\mathscr{M}_1, v)$ and $(\mathscr{M}_2, w)$ are bisimilar with respect to i, then*

$$(\mathscr{M}_1, v) \models [i]\varphi \text{ iff } (\mathscr{M}_2, w) \models [i]\varphi$$

for any formula φ in the signature S.

3.3 Extended Syntax Trees

As in the previous chapter, we proceed to defining extended syntax trees for modal formulae[5]. The main idea underlying the construction of these trees is the same as in the Boolean case but, naturally, this time, the sets of models labelling each node will be more complicated. For the sake of conceptual clarity, we are more careful in this section and begin by defining the usual syntax trees followed by their extended counterparts.

Definition 3.19 (Syntax trees)**.** The syntax tree T_φ of a modal formula φ is built recursively on the structure of φ as follows.

(φ is a propositional symbol $p \in P$): T_p consists of a single node t that has a **syntax label** $synl(t) = p$.

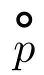


FIGURE 3.2: The syntax tree T_p.

(φ is $\neg\psi$): $T_{\neg\psi}$ has a root node t with $synl(t) = \neg$. The unique child of t is the root of the syntax tree T_ψ.

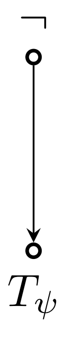


FIGURE 3.3: The syntax tree $T_{\neg\psi}$.

(φ is $\psi_1 \vee \psi_2$): $T_{\psi_1 \vee \psi_2}$ has a root node t with $synl(t) = \vee$. The left child of t is the root of the syntax tree T_{ψ_1}. The right child of t is the root of the syntax tree T_{ψ_2}.

[5]When no confusion arises, we call any formula "modal formula" if it belongs to at least one of the sets Φ_{ML}, $\Phi_{[\cup]\mathrm{ML}}$, etc, defined in Definition 3.11.

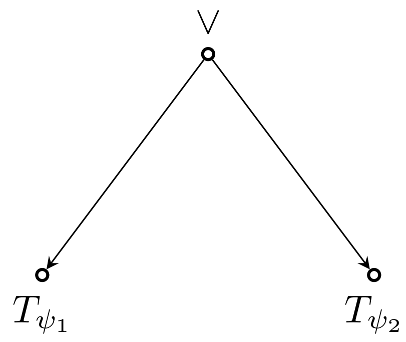


FIGURE 3.4: The syntax tree $T_{\psi_1 \vee \psi_2}$.

(φ is $[i]\psi$): $T_{[i]\psi}$ has a root node t with $synl(t) = [i]$. The unique child of t is the root of the syntax tree T_ψ.

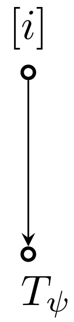


FIGURE 3.5: The syntax tree $T_{[i]\psi}$.

(φ is $[\cup_\Gamma]\psi$): $T_{[\cup_\Gamma]\psi}$ has a root node t with $synl(t) = [\cup_\Gamma]$. The unique child of t is the root of the syntax tree T_ψ.

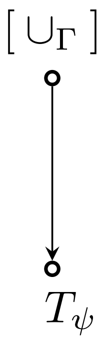


FIGURE 3.6: The syntax tree $T_{[\cup_\Gamma]\psi}$.

(φ is $[\cap_\Gamma]\psi$): $T_{[\cap_\Gamma]\psi}$ has a root node t with $synl(t) = [\cap_\Gamma]$. The unique child of t is the root of the syntax tree T_ψ.

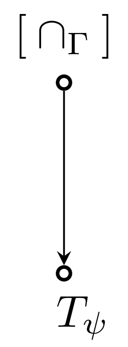


FIGURE 3.7: The syntax tree $T_{[\cap_\Gamma]\psi}$.

(φ is $[\exists_\Gamma]\psi$): $T_{[\exists_\Gamma]\psi}$ has a root node t with $synl(t) = [\exists_\Gamma]$. The unique child of t is the root of the syntax tree T_ψ.

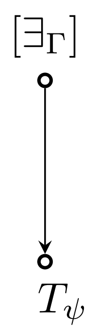


FIGURE 3.8: The syntax tree $T_{[\exists_\Gamma]\psi}$.

(φ is $[\varphi]\psi$): $T_{[\varphi]\psi}$ has a root node t with $synl(t) = [\varphi]$. The left child of t is the root of the extended syntax tree T_φ. The right child of t is the root of the syntax tree T_ψ.

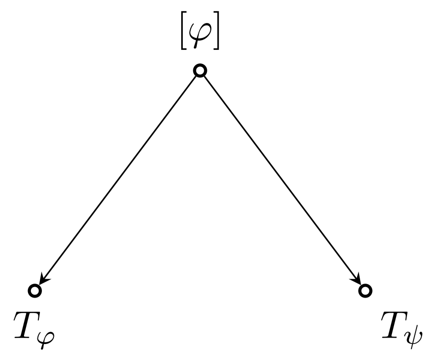


FIGURE 3.9: The syntax tree $T_{[\varphi]\psi}$.

Since every syntax tree is a rooted tree in which every edge is naturally oriented away from the root as shown in Definition 3.19 and the accompanying figures, the next definition, although formally not very precise, should be unambiguous and intuitively clear.

Definition 3.20 (Branches)**.** A branch B in a syntax tree is any path starting at the root of the tree and ending in a leaf.

1. For any branch B, the word $l \dots m$, formed by the relation indices of all the nodes with syntactic labels of the form $[i]$ occurring along B when traversing the branch from the root to its leaf is denoted $I(B)$.

2. Let two not necessarily distinct syntax trees T_1 and T_2 be given and let the branch B_1, consisting of the nodes $\eta_0, \eta_1, \dots, \eta_k$, and the branch B_2, consisting of the nodes $\eta'_0, \eta'_2, \dots, \eta'_k$ (where the nodes in each branch have been numbered in increasing order starting from the root of the tree), belong to T_1 and T_2, respectively. We call B_1 and B_2 isomorphic, and write $B_1 \cong B_2$, iff $k = l$ and the symbols from the set $\Sigma = \{p, \neg, \vee, [i], [\exists_\Gamma], [\cup_\Gamma], [\cap_\Gamma], [\varphi]\}$ labelling the nodes η_j, η'_j, where $1 \leq j \leq k$, are the same.

It is obvious that if a syntax tree T_φ contains branches B_1 and B_2 such that $I(B_1) \neq I(B_2)$, then $B_1 \ncong B_2$. Of course if $B_1 \ncong B_2$, then these branches are different, i.e, they do not coincide.

Example 3.5. *The syntax tree* $T_{[\cup_\Gamma]\mathbf{b} \vee [a]\neg[d]\neg\mathbf{b}}$ *of the formula* $[\cup_\Gamma]\mathbf{b} \vee [a]\neg[d]\neg\mathbf{b}$*, where* $\Gamma = \{a, d\}$*, is shown in Figure 3.10. We have* $I(B_r) = ad$ *for the right branch* B_r *of*

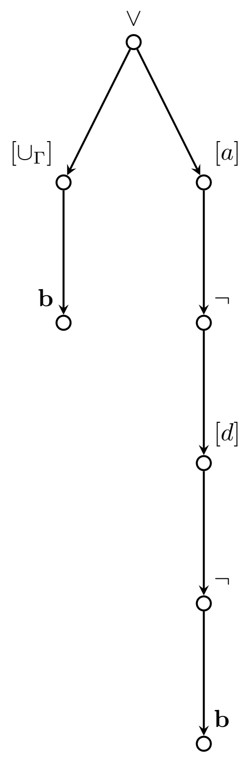


FIGURE 3.10: The syntax tree $T_{[\cup_\Gamma]\mathbf{b} \vee [a]\neg[d]\neg\mathbf{b}}$ of the formula $[\cup_\Gamma]\mathbf{b} \vee [a]\neg[d]\neg\mathbf{b}$.

$T_{[\cup\Gamma]\mathbf{b}\vee[a]\neg[d]\neg\mathbf{b}}$ *that consists of the root of the tree and the nodes with syntax labels* $[a]$*,* $\neg$*,* $[d]$*,* $\neg$*, and* $\mathbf{b}$ *in that order. It is obvious that* $B_l \not\cong B_r$ *where* B_l *is the left branch that consists of the root of* $T_{[\cup\Gamma]\mathbf{b}\vee[a]\neg[d]\neg\mathbf{b}}$ *and the nodes with syntax labels* $[\cup\Gamma]$ *and* $\mathbf{b}$*.*

As in the previous chapter, in order to prove a lower bound on the size of a formula $|\varphi|$ that expresses certain property P, the idea is to connect the number of nodes n in the syntax tree of φ with the fact that φ differentiates between models that have the property P and, therefore, satisfy φ, and those that do not. To this end we introduce extended syntax trees for modal formulae.

We begin by defining a number of operations on pointed models and sets of pointed models that will be useful later.

Definition 3.21. Let $\mathbb{M} = \{(\mathscr{M}_1, w_1), \ldots, (\mathscr{M}_k, w_k), \ldots\}$ be a set of pointed models and let $(\mathscr{M}, w) \in \mathbb{M}$.

- For any $i \in I$, let

$$[i](\mathscr{M}, w) = \{(\mathscr{M}, v) \mid v \in \mathscr{M} \text{ and } wR_i v\}.$$

  Intuitively, $[i](\mathscr{M}, w)$ is the set of all pointed models that can be reached from w by making one R_i-step. Note that if there is no point $v \in \mathscr{M}$ such that $wR_i v$, then $[i](\mathscr{M}, w) = \emptyset$.

- Using the first item, $[i]\mathbb{M}$ is defined as

$$[i]\mathbb{M} = \bigcup_{(\mathscr{M},w)\in\mathbb{M}} [i](\mathscr{M}, w).$$

  It is obvious that $[i](\mathscr{M}, w)$ can be empty because, for each $(\mathscr{M}, w) \in \mathbb{M}$ it is possible to have $[i](\mathscr{M}, w) = \emptyset$.

- If $(\mathscr{M}, w) \models \langle i\rangle\psi$, there is at least one $v \in \mathscr{M}$ such that $wR_i v$ and $(\mathscr{M}, v) \models \psi$. We construct the non-empty set of all such pointed models, i.e.,

$$\langle i(\psi)\rangle(\mathscr{M}, w) = \{(\mathscr{M}, v) \mid v \in \mathscr{M} \text{ such that } wR_i v \text{ and } (\mathscr{M}, v) \models \psi\}.$$

- If $\mathbb{M} \models \langle i\rangle\psi$, then, using the item above, we form the non-empty set of pointed models

$$\langle i(\psi)\rangle\mathbb{M} = \bigcup_{(\mathscr{M},w)\in\mathbb{M}} \langle i(\psi)\rangle(\mathscr{M}, w).$$

- For a pointed model $(\mathscr{M}, v)$, and a modal formula φ, we define $\varphi((\mathscr{M}, v))$ as the set of all pointed models with underlying structure $\mathscr{M}$ that satisfy φ, i.e.,

$$\varphi((\mathscr{M}, v)) = \{(\mathscr{M}, w) \mid (\mathscr{M}, w) \models \varphi \text{ and } w \in \mathscr{M}\}.$$

  It is clear that this set can be empty.

- Using the previous item, we define

$$\mathbb{M}|_\varphi = \{(\mathscr{M}_i|_\varphi, v_i) \mid (\mathscr{M}_i, v_i) \in \mathbb{M} \text{ and } (\mathscr{M}_i, v_i) \in \varphi((\mathscr{M}_i, v_i))\}.$$

  Note that $\mathbb{M}|_\varphi$ can be empty if for all $(\mathscr{M}_i, v_i) \in \mathbb{M}$, it is true that $(\mathscr{M}_i, v_i) \notin \varphi((\mathscr{M}_i, v_i))$, i.e., $(\mathscr{M}_i, v_i) \not\models \varphi$.

For any finite non-empty set $\Gamma \subseteq I$, we define the sets below.

- $[\cup_\Gamma]\mathbb{M}$ is the union of $[i]\mathbb{M}$ for all $i \in \Gamma$, i.e.,

$$[\cup_\Gamma]\mathbb{M} = \bigcup_{i\in\Gamma} [i]\mathbb{M}.$$

  It follows that if for all $i \in \Gamma$, we have $[i]\mathbb{M} = \emptyset$, then $[\cup_\Gamma]\mathbb{M} = \emptyset$.

- Let us suppose that $\mathbb{M} \models \langle\cup_\Gamma\rangle\psi$. Therefore, for every $(\mathscr{M}_j, w_j) \in \mathbb{M}$, there is a non-empty subset $\Gamma_j \subseteq \Gamma$ such that for every $i \in \Gamma_j$ there is at least one $v \in \mathscr{M}_j$ for which $w_j R_i v$ and $(\mathscr{M}_j, v) \models \psi$. Hence, we can construct the non-empty set

$$\langle\cup_\Gamma(\psi)\rangle\mathbb{M} = \bigcup_{i\in\Gamma_1} \langle i(\psi)\rangle(\mathscr{M}_1, w_1) \cup \ldots \cup \bigcup_{i\in\Gamma_k} \langle i(\psi)\rangle(\mathscr{M}_k, w_k) \ldots.$$

- We form the possible empty set $[\cap_\Gamma](\mathscr{M}, w)$ as follows.

$$[\cap_\Gamma](\mathscr{M}, w) = \{(\mathscr{M}, v) \mid v \in \mathscr{M} \text{ and } \bigwedge_{r_i\in\Gamma} wR_i v\}.$$

- Using the previous item, we define

$$[\cap_\Gamma]\mathbb{M} = \bigcup_{(\mathscr{M},w)\in\mathbb{M}} [\cap_\Gamma](\mathscr{M}, w).$$

  It is clear that this set could be empty.

- If $(\mathscr{M}, w) \models \langle\cap_\Gamma\rangle\psi$, there is at least one $v \in \mathscr{M}$ such that $wR_i v$ for any $i \in \Gamma$ and $(\mathscr{M}, v) \models \psi$. We construct the non-empty set of all such pointed models, i.e.,

$$\langle\cap_\Gamma(\psi)\rangle(\mathscr{M}, w) = \{(\mathscr{M}, v) \mid v \in \mathscr{M} \text{ such that } \bigwedge_{i\in\Gamma} wR_i v \text{ and } (\mathscr{M}, v) \models \psi\}.$$

- Let $\mathbb{M} \models \langle\cap_\Gamma\rangle\psi$. Then we can define the non-empty set

$$\langle\cap_\Gamma(\psi)\rangle\mathbb{M} = \bigcup_{(\mathscr{M},w)\in\mathbb{M}} \langle\cap_\Gamma(\psi)\rangle(\mathscr{M}, w).$$

- If $\mathbb{M} \models [\exists_\Gamma]\psi$, then for every $(\mathscr{M}_j, w_j) \in \mathbb{M}$, there is a subset $\Gamma_j \subseteq \Gamma$ such that for every $i \in \Gamma_j$, $(\mathscr{M}_j, w_j) \models [i]\psi$. Therefore, we can construct the (possibly empty) set
$$[\exists_\Gamma(\psi)]\mathbb{M} = \bigcup_{i\in\Gamma_1}[i](\mathscr{M}_1, w_1) \cup \ldots \cup \bigcup_{i\in\Gamma_k}[i](\mathscr{M}_k, w_k)\ldots.$$

- Let $\mathbb{M} \models \langle\exists_\Gamma\rangle\psi$. Therefore, for every $(\mathscr{M}, w) \in \mathbb{M}$ and every $i \in \Gamma$, there is at least one $v \in \mathscr{M}$ such that wR_iv and $(\mathscr{M}, v) \models \psi$. We form the non-empty set
$$\langle\exists_\Gamma(\psi)\rangle\mathbb{M} = \bigcup_{i\in\Gamma}\langle i(\psi)\rangle(\mathscr{M}_1, w_1) \cup \ldots \cup \bigcup_{i\in\Gamma}\langle i(\psi)\rangle(\mathscr{M}_k, w_k)\ldots.$$

We are ready to define extended syntax trees for modal formulae. The reader can easily see that these trees can be thought of as closed game trees for suitably defined versions of the Adler-Immerman games. In fact, the rules of such games corresponding to the modal logics we study can be extracted from the definition below and relevant versions of Theorem 2.9 and Theorem 2.10 can be proven. More about this line of reasoning can be found in [22].

As before, we use $\mathbb{A} \circ \mathbb{B}$ to denote a tree-node t that has a semantic label $\langle\mathbb{A}, \mathbb{B}\rangle$.

Definition 3.22 (Extended Syntax Trees)**.** For any modal formula φ and any sets of pointed models $\mathbb{A}$ and $\mathbb{B}$ such that $\mathbb{A} \models \varphi$ and $\mathbb{B} \models \neg\varphi$, the extended syntax tree $T_\varphi^{\langle\mathbb{A},\mathbb{B}\rangle}$ is defined inductively on the structure of φ as follows:

(φ **is a propositional symbol** $p \in P$)**:** $T_p^{\langle\mathbb{A},\mathbb{B}\rangle}$ consists of a single node t that has a syntax label $synl(t) = p$ and a **semantic label** $seml(t) = \langle\mathbb{A}, \mathbb{B}\rangle$.

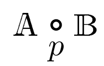


Figure 3.11: The extended syntax tree $T_p^{\langle\mathbb{A},\mathbb{B}\rangle}$.

(φ **is** $\neg\psi$)**:** $T_{\neg\psi}^{\langle\mathbb{A},\mathbb{B}\rangle}$ has a root node t with $synl(t) = \neg$ and $seml(t) = \langle\mathbb{A}, \mathbb{B}\rangle$. The unique child of t is the root $\mathbb{B} \circ \mathbb{A}$ of the extended syntax tree $T_\psi^{\langle\mathbb{B},\mathbb{A}\rangle}$. Note that since $\mathbb{A} \models \varphi$ and $\mathbb{B} \models \neg\varphi$, we have $\mathbb{B} \models \psi$ and $\mathbb{A} \models \neg\psi$.

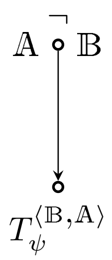


Figure 3.12: The extended syntax tree $T_{\neg\psi}^{\langle\mathbb{A},\mathbb{B}\rangle}$.

(φ is $\psi_1 \vee \psi_2$): $T^{\langle \mathbb{A},\mathbb{B}\rangle}_{\psi_1\vee\psi_2}$ has a root node t with $synl(t) = \vee$ and $seml(t) = \langle \mathbb{A}, \mathbb{B}\rangle$.

The left child of t is the root $\mathbb{A}_1 \circ \mathbb{B}$ of $T^{\langle \mathbb{A}_1,\mathbb{B}\rangle}_{\psi_1}$. The right child of t is the root $\mathbb{A}_2 \circ \mathbb{B}$ of $T^{\langle \mathbb{A}_2,\mathbb{B}\rangle}_{\psi_2}$ where the sets $\mathbb{A}_1$ and $\mathbb{A}_2$ are defined as follows:

$$\mathbb{A}_1 = \{(\mathscr{M}, v) \in \mathbb{A} \mid (\mathscr{M}, v) \models \psi_1\} \text{ and } \mathbb{A}_2 = \{(\mathscr{M}, v) \in \mathbb{A} \mid (\mathscr{M}, v) \models \psi_2\}.$$

Therefore, $\mathbb{A}_1 \models \psi_1$ and $\mathbb{A}_2 \models \psi_2$ while $\mathbb{B} \models \neg(\psi_1 \vee \psi_2)$. Although it may seem obvious, we would like to stress that $\mathbb{A} = \mathbb{A}_1 \cup \mathbb{A}_2$ does **not** imply that $\mathbb{A}_1$ and $\mathbb{A}_2$ are non-intersecting, i.e., that $\mathbb{A}_1 \cap \mathbb{A}_2 = \emptyset$.

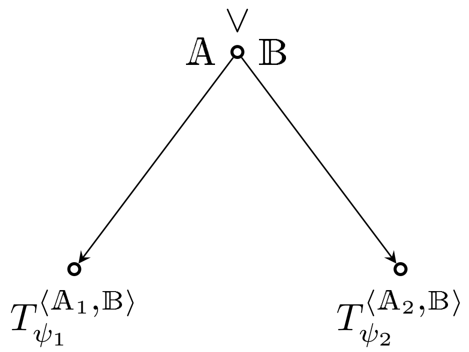


FIGURE 3.13: The extended syntax tree $T^{\langle \mathbb{A},\mathbb{B}\rangle}_{\psi_1\vee\psi_2}$.

(φ is $[i]\psi$): $T^{\langle \mathbb{A}\circ\mathbb{B}\rangle}_{[i]\psi}$ has a root node t with $synl(t) = [i]$ and $seml(t) = \langle \mathbb{A}, \mathbb{B}\rangle$. The unique child of t is the root $[i]\mathbb{A} \circ \langle i(\neg\psi)\rangle\mathbb{B}$ of $T^{\langle [i]\mathbb{A},\langle i(\neg\psi)\rangle\mathbb{B}\rangle}_{\psi}$. Again, we would like to emphasise that $[i]\mathbb{A} \models \psi$ and $\langle i(\neg\psi)\rangle\mathbb{B} \models \neg\psi$.

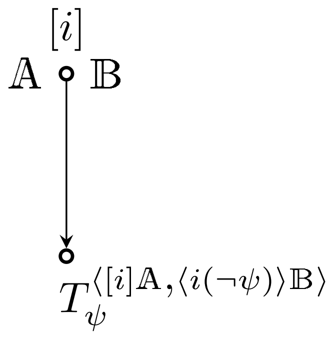


FIGURE 3.14: The extended syntax tree $T^{\langle \mathbb{A},\mathbb{B}\rangle}_{[i]\psi}$.

(φ is $[\cup_\Gamma]\psi$): $T^{\langle \mathbb{A},\mathbb{B}\rangle}_{[\cup_\Gamma]\psi}$ has a root node t with $synl(t) = [\cup_\Gamma]$ and $seml(t) = \langle \mathbb{A}, \mathbb{B}\rangle$. The unique child of t is the root $[\cup_\Gamma]\mathbb{A} \circ \langle \cup_\Gamma(\neg\psi)\rangle\mathbb{B}$ of $T^{\langle [\cup_\Gamma]\mathbb{A},\langle \cup_\Gamma(\neg\psi)\rangle\mathbb{B}\rangle}_{\psi}$. Note that $[\cup_\Gamma]\mathbb{A} \models \psi$ and $\langle \cup_\Gamma(\neg\psi)\rangle\mathbb{B} \models \neg\psi$.

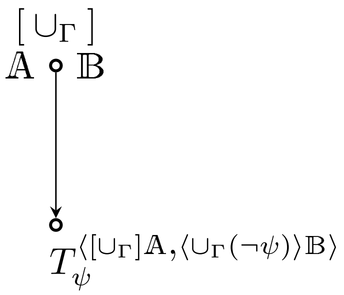


FIGURE 3.15: The extended syntax tree $T_{[\cup_\Gamma]\psi}^{\langle \mathbb{A}, \mathbb{B}\rangle}$.

(φ is $[\cap_\Gamma]\psi$): The extended syntax tree $T_{[\cap_\Gamma]\psi}^{\langle \mathbb{A}, \mathbb{B}\rangle}$ has a root node t with $synl(t) = [\cap_\Gamma]$ and $seml(t) = \langle \mathbb{A}, \mathbb{B}\rangle$. The unique child of t is the root $[\cap_\Gamma]\mathbb{A} \circ \langle \cap_\Gamma(\neg\psi)\rangle \mathbb{B}$ of the extended syntax tree $T_\psi^{\langle [\cap_\Gamma]\mathbb{A}, \langle \cap_\Gamma(\neg\psi)\rangle \mathbb{B}\rangle}$. Again, we have $[\cap_\Gamma]\mathbb{A} \models \psi$ and $\langle \cap_\Gamma(\neg\psi)\rangle \mathbb{B} \models \neg\psi$.

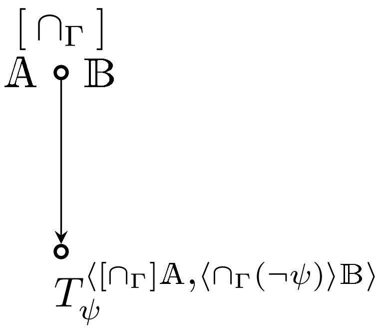


FIGURE 3.16: The extended syntax tree $T_{[\cap_\Gamma]\psi}^{\langle \mathbb{A}, \mathbb{B}\rangle}$.

(φ is $[\exists_\Gamma]\psi$): The extended syntax tree $T_{[\exists_\Gamma]\psi}^{\langle \mathbb{A}, \mathbb{B}\rangle}$ has a root t with $synl(t) = [\exists_\Gamma]$ and $seml(t) = \langle \mathbb{A}, \mathbb{B}\rangle$. The unique child of t is the root $[\exists_\Gamma]\mathbb{A} \circ \langle \exists_\Gamma(\neg\psi)\rangle \mathbb{B}$ of the extended syntax tree $T_\psi^{\langle [\exists_\Gamma]\mathbb{A}, \langle \exists_\Gamma(\neg\psi)\rangle \mathbb{B}\rangle}$. Note that $[\exists_\Gamma]\mathbb{A} \models \psi$ and $\langle \exists_\Gamma(\neg\psi)\rangle \mathbb{B} \models \neg\psi$.

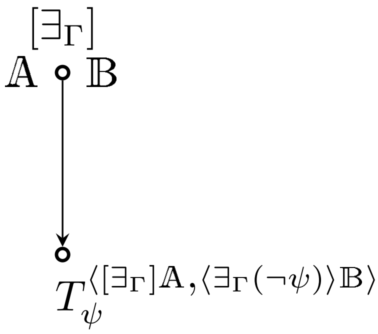


FIGURE 3.17: The extended syntax tree $T_{[\exists_\Gamma]\psi}^{\langle \mathbb{A}, \mathbb{B}\rangle}$.

(φ is $[\varphi]\psi$): The extended syntax tree $T^{\langle \mathbb{A},\mathbb{B}\rangle}_{[\varphi]\psi}$ has a root node t with $synl(t) = [\varphi]$ and $seml(t) = \langle \mathbb{A}, \mathbb{B}\rangle$. The left child of t is the root $\varphi(\mathbb{A}) \cup \varphi(\mathbb{B}) \circ \neg\varphi(\mathbb{A}) \cup \neg\varphi(\mathbb{B})$ of the extended syntax tree $T_{\varphi}^{\langle \varphi(\mathbb{A})\cup\varphi(\mathbb{B}),\neg\varphi(\mathbb{A})\cup\neg\varphi(\mathbb{B})\rangle}$. The right child of t is is the root $\mathbb{A}|_{\varphi} \circ (\mathbb{B})|_{\varphi}$ of the extended syntax tree $T_{\psi}^{\langle \mathbb{A}|_{\varphi},(\mathbb{B})|_{\varphi}\rangle}$. Using the relevant items from Definition 3.21, the reader can easily verify that the following items are true.

- $\varphi(\mathbb{A}) \cup \varphi(\mathbb{B}) \models \varphi$ and $\neg\varphi(\mathbb{A}) \cup \neg\varphi(\mathbb{B}) \models \neg\varphi$;
- $\mathbb{A}|_{\varphi} \models \psi$ and $\mathbb{B}|_{\varphi} \models \neg\psi$.

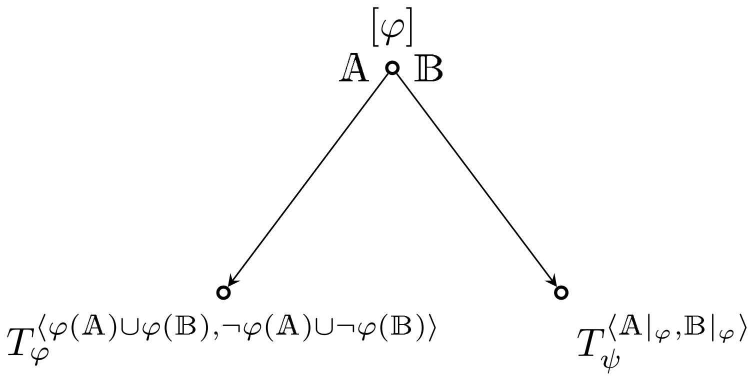


FIGURE 3.18: The extended syntax tree $T^{\langle \mathbb{A},\mathbb{B}\rangle}_{[\varphi]\psi}$.

Example 3.6. *The extended syntax tree* $T^{\langle \mathbb{A},\mathbb{D}\rangle}_{[\cup\Gamma]\mathbf{b}\vee[a]\neg[d]\neg\mathbf{b}}$ *of* $[\cup\Gamma]\mathbf{b}\vee[a]\neg[d]\neg\mathbf{b}$ *is shown in Figure 3.19. We have that* $\Gamma = \{a, d\}$. *Pointed models occurring in the semantic labels of the nodes of the syntax tree are the pairs consisting of the relevant Kripke model* $\mathscr{A}_i$ *or* $\mathscr{D}_1$ *and the nodes marked by* $\triangleright$ *and* $\triangleleft$, *respectively. Hence,* $\mathbb{A}$ *consists of the pointed models on the left of the root of the tree which has syntax label* $\vee$ *while* $\mathbb{D}$ *is on the right and contains only one pointed model. Black circles in the Kripke models denote the points where the atom* **b** *is true; white circles denote points that do not satisfy any proposition. The solid arrows in the models denote relation steps indexed with* a, *the dashed arrows are indexed with* d. *Again, we have* $I(B_r) = ad$ *for the right branch* B_r *of* $T^{\langle \mathbb{A}\circ\mathbb{D}\rangle}_{[\cup\Gamma]\mathbf{b}\vee[a]\neg[d]\neg\mathbf{b}}$ *that consists of the root and the nodes with syntax labels* $[a]$, $\neg$, $[d]$, $\neg$, *and* **b**.

It is worth pointing out that given an extended syntax tree $T_{\varphi}^{\langle \mathbb{A},\mathbb{D}\rangle}$, the "shape" of the tree depends solely on φ. In other words, if, for example, we disregard the semantic labels of the nodes, the three extended syntax trees $T^{\langle \mathbb{A},\mathbb{D}\rangle}_{[\cup\Gamma]\mathbf{b}\vee[a]\neg[d]\neg\mathbf{b}}$, $T^{\langle \emptyset,\mathbb{D}\rangle}_{[\cup\Gamma]\mathbf{b}\vee[a]\neg[d]\neg\mathbf{b}}$, and $T^{\langle \emptyset,\emptyset\rangle}_{[\cup\Gamma]\mathbf{b}\vee[a]\neg[d]\neg\mathbf{b}}$ are actually the syntax tree of the formula $[\cup\Gamma]\mathbf{b}\vee[a]\neg[d]\neg\mathbf{b}$ from Figure 3.10. Therefore, for any modal formula φ, if the syntax tree T_{φ} of φ contains a branch B for which $I(B) = i_1 \ldots i_n$, then for any pair of sets of pointed models $\mathbb{E}$ and $\mathbb{F}$ such that

$$\mathbb{E} \models \varphi \text{ and } \mathbb{F} \models \neg\varphi,$$

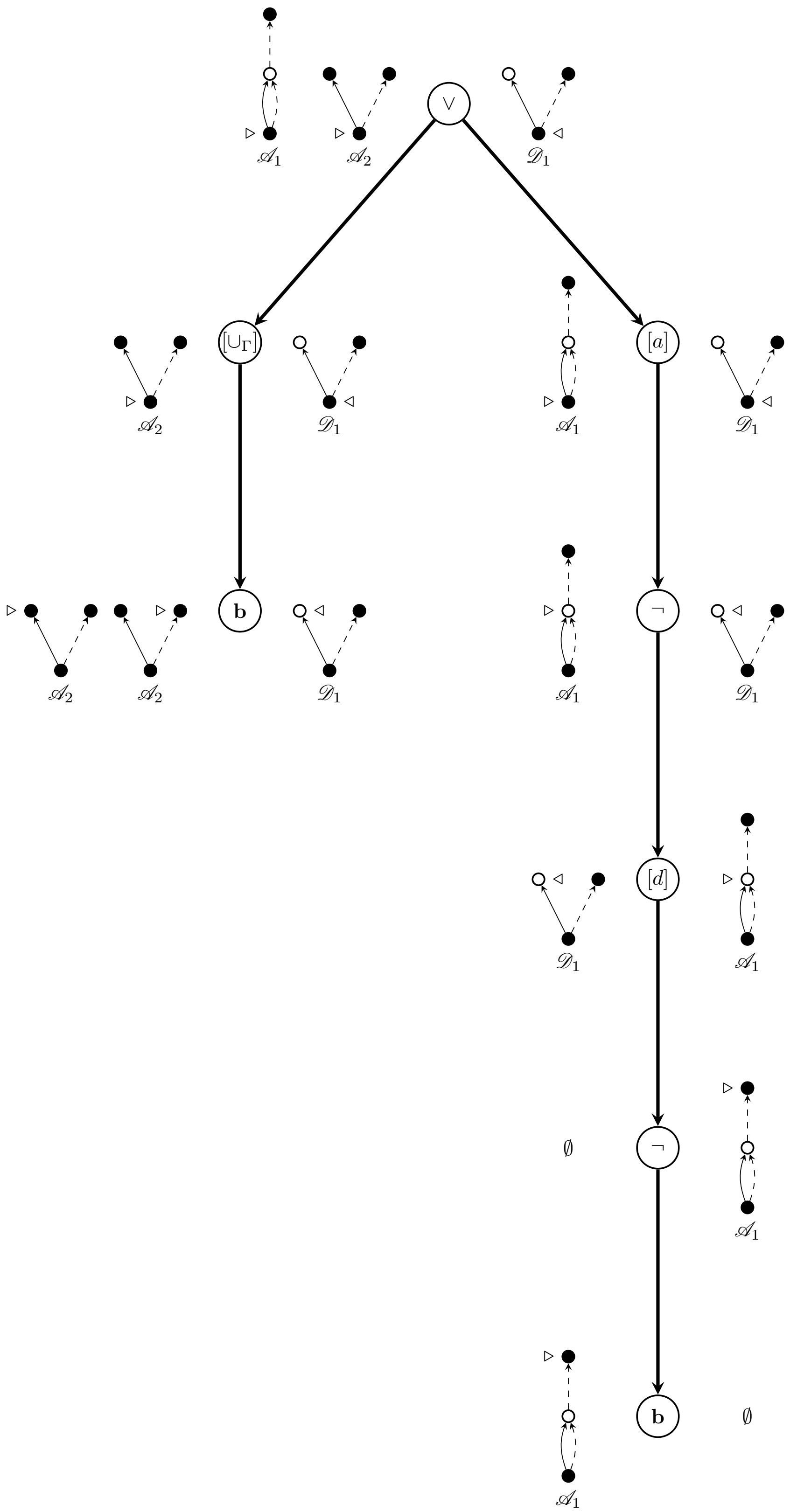


FIGURE 3.19: The extended syntax tree $T^{\langle \mathbb{A}, \mathbb{D} \rangle}_{[\cup_\Gamma]\mathbf{b} \vee [a]\neg[d]\neg\mathbf{b}}$.

the extended syntax tree of φ with root $\mathbb{E} \circ \mathbb{F}$ contains a branch B for which $I(B) = i_1 \dots i_n$. Thus, we will continue using the notions $I(B)$ and $B_1 \cong B_2$ given in Definition 3.20 for extended syntax trees, too.

Example 3.7. *The extended syntax tree $T^{\langle \mathbb{A},\mathbb{B} \rangle}_{[\mathbf{b}][a]p}$ is shown on Figure 3.20. Again, we use Kripke structures plus $\lhd$ and $\rhd$ to denote the pointed models occurring in the semantic labels of the nodes of the extended syntax tree. As before, we have that black circles denote points satisfying the proposition* **b** *and arrows are labelled with the relation index a. It is obvious that the formula $[\mathbf{b}][a]p$ is true in the set $\mathbb{A}$ that contains the only pointed model on the left of the root of $T^{\langle \mathbb{A},\mathbb{D} \rangle}_{[\mathbf{b}][a]p}$ and false in the set $\mathbb{B}$ that has just one pointed model, namely, the pointed model on the right of the root of the tree. Note how we have constructed the sets $\mathbf{b}(\mathbb{A}) \cup \mathbf{b}(\mathbb{B})$ on the left of the tree node with syntax label* **b** *and the set $\neg\mathbf{b}(\mathbb{A}) \cup \neg\mathbf{b}(\mathbb{B})$ on the right. The sets $\mathbb{A}|_\mathbf{b}$ and $\mathbb{B}|_\mathbf{b}$ form the semantic label of the node with syntax label $[a]$.*

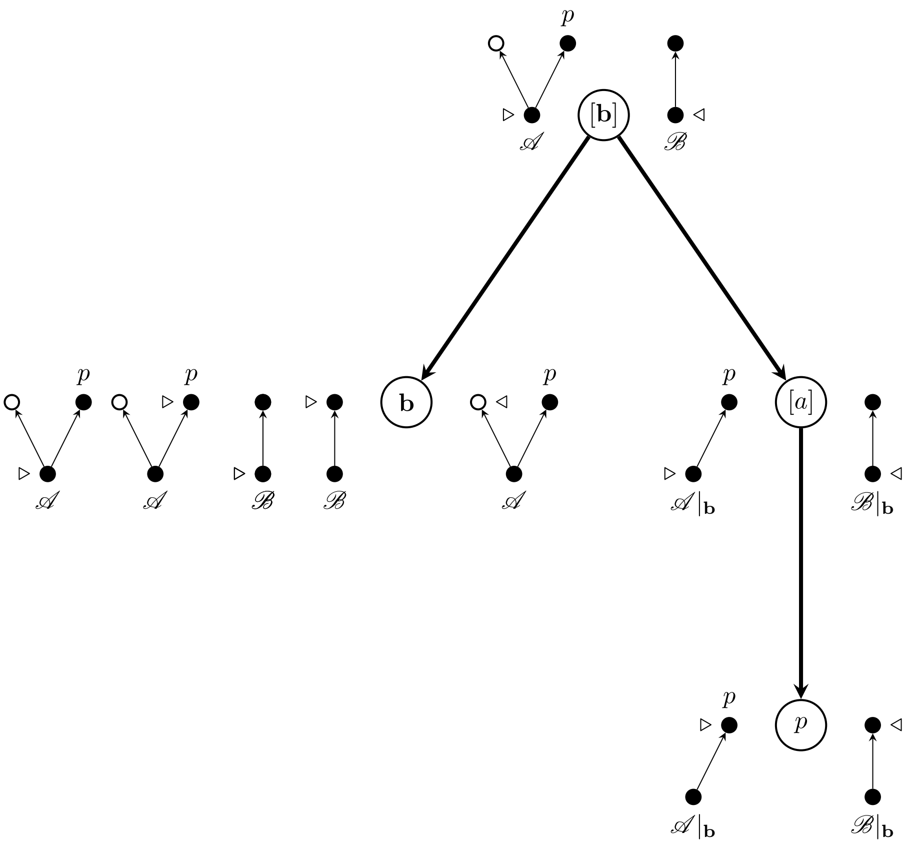


FIGURE 3.20: The extended syntax tree $T^{\langle \mathbb{A},\mathbb{D} \rangle}_{[\mathbf{b}][a]p}$.

Having defined extended syntax trees for modal formulae, it is easy to see that an analogue of Proposition 2.6 holds in the present case, too.

Proposition 3.23. *For any modal formulae φ and ψ and any sets of pointed models $\mathbb{A}$, $\mathbb{B}$, $\mathbb{C}$, and $\mathbb{D}$, if the extended syntax tree $T^{\langle \mathbb{A},\mathbb{B} \rangle}_{\varphi}$ contains a node $\mathbb{C} \circ \mathbb{D}$ that is the root of the extended syntax tree $T^{\langle \mathbb{C},\mathbb{D} \rangle}_{\psi}$, then $\mathbb{C} \models \psi$ and $\mathbb{D} \models \neg\psi$.*

Proof. Immediate from Definition 3.21 and Definition 3.22. □

Corollary 3.24. *For any modal formula $[i]\varphi$, if the two sets of pointed models $\mathbb{C}$ and $\mathbb{D}$ contain a pair of models $(\mathscr{C}, c) \in \mathbb{C}$ and $(\mathscr{D}, d) \in \mathbb{D}$ such that $(\mathscr{C}, c)$ and $(\mathscr{D}, d)$ are bisimilar with respect to i, then there is no modal formula ψ and a pair of pointed models $\mathbb{A}$ and $\mathbb{B}$ such that the extended syntax tree $T_\psi^{\langle \mathbb{A}, \mathbb{B} \rangle}$ contains a node $\mathbb{C} \circ \mathbb{D}$ that is the root of the extended syntax tree $T_{[i]\varphi}^{\langle \mathbb{C}, \mathbb{D} \rangle}$.*

Proof. If we assume otherwise, then Proposition 3.23 implies that $(\mathscr{C}, c) \models [i]\varphi$ and $(\mathscr{D}, d) \models \neg[i]\varphi$ which contradicts Theorem 3.18. □

Extended syntax trees will be the main tool for obtaining our exponential succinctness results formulated in the next sections. There, we apply the general recipe based on Lemma 3.10 that can be informally described as follows.

Suppose that we want to prove that a set Φ_1 of L_1-formulae is exponentially more succinct than a set of L_2-formulae Φ_2 on a set of pointed models $\mathbb{M}$.

> *For every $n \in \mathbb{N}$, find a formula $\varphi_n \in \Phi_1$ such that $|\varphi_n| = f(n)$ where $f(x)$ is some strictly increasing function. After that, find two sets of pointed models $\mathbb{A}_n \subseteq \mathbb{M}$ and $\mathbb{B}_n \subseteq \mathbb{M}$, such that $\mathbb{A}_n \models_1 \varphi_n$ and $\mathbb{B}_n \models_1 \neg\varphi_n$ and prove that the extended syntax tree with root $\mathbb{A}_n \circ \mathbb{B}_n$ of any L_2-formula $\psi_n \in \Phi_2$ such that $\mathbb{A}_n \models_2 \psi_n$ and $\mathbb{B}_n \models_2 \neg\psi_n$ contains at least $2^{g(n)}$ nodes where $g(x)$ is a function that grows asymptotically at least as fast as $f(x)$.*

Then, all the conditions from Lemma 3.10 are fulfilled and hence, Φ_1 is exponentially more succinct than Φ_2 on $\mathbb{M}$.

Obviously, there are two main difficulties with this strategy. The first one is finding the right models which requires an intuitive understanding of the type of semantic properties that are expressed more efficiently by formulae in Φ_1 than by Φ_2. For example, it is clear that a formula of the form $\neg[\cup_{\{a,d\}}]\neg\mathbf{b}$ is equivalent to the formula $\langle a \rangle \mathbf{b} \vee \langle d \rangle \mathbf{b}$. Therefore, intuitively, the first formula expresses more efficiently the fact that we can make either an a or a d-step from the current point and reach a point that satisfies the proposition $\mathbf{b}$. Hence, by stacking n boxes, as in the formula

$$\neg \underbrace{[\cup_{\{a,d\}}] \dots [\cup_{\{a,d\}}]}_{n \text{ times}} \neg\mathbf{b},$$

we describe an exponential in n number of different paths consisting of a and/or d-steps starting at the current point such that at least one of them leads to a point that satisfies $\mathbf{b}$. Indeed, this is one of the properties we are going to exploit later. The second difficulty is proving lower bounds on the number of nodes of the extended syntax trees $T_\psi^{\langle \mathbb{A}_n, \mathbb{B}_n \rangle}$. This is an underdeveloped area and there are currently just two known techniques which we explain below by using Φ_{ML}-formulae. We would like to stress however that the main ideas are applicable to any other logic.

Diverging Pairs: This technique was used in [1]. We formulate a specific instance of it in Theorem 3.25 below. The general idea is as follows. Let us suppose that we want to prove a lower bound on the size of Φ_{ML}-formulae that are true in some set of pointed models $\mathbb{A}$ and false in another set $\mathbb{B}$. This can be achieved by proving a lower bound on the size of all extended syntax trees with root $\mathbb{A} \circ \mathbb{B}$. Intuitively, Definition 3.22 gives us a means of estimating the size of an extended syntax tree $T_\varphi^{\langle \mathbb{A}, \mathbb{B} \rangle}$ by applying combinatorial reasoning involving the sets $\mathbb{A}$ and $\mathbb{B}$. Depending on the case at hand, these sets can be very big and may contain models that make such combinatorial considerations extremely difficult. To avoid such difficulties, we choose some "simpler", according to a useful criterion, subsets $\mathbb{A}^1, \ldots, \mathbb{A}^m$ of $\mathbb{A}$ and, similarly, $\mathbb{B}^1, \ldots, \mathbb{B}^m$ of $\mathbb{B}$ and then we "reduce" the problem of reasoning about $T_\varphi^{\langle \mathbb{A}, \mathbb{B} \rangle}$ to reasoning about extended syntax trees with roots $\mathbb{A}^i \circ \mathbb{B}^i$ as shown in the next theorem.

Theorem 3.25 (Principle of Diverging Pairs for Modal Formulae)**.** *Let* $\mathbb{A}$ *and* $\mathbb{B}$ *be two sets of pointed models and let* $\mathbb{A}^1, \ldots, \mathbb{A}^m$ *be subsets of* $\mathbb{A}$ *and* $\mathbb{B}^1, \ldots, \mathbb{B}^m$ *be subsets of* $\mathbb{B}$. *If any extended syntax tree with root* $\mathbb{A}^1 \circ \mathbb{B}^1$ *contains a branch* B_1 *and any extended syntax tree with root* $\mathbb{A}^2 \circ \mathbb{B}^2$ *contains a branch* B_2*, and …, and any extended syntax tree with root* $\mathbb{A}^m \circ \mathbb{B}^m$ *contains a branch* B_m*, where* $B_i \not\cong B_j$ *for all* i *and* j *such that* $1 \leq i \neq j \leq m$*, then the syntax tree* T_φ *of any modal formula* φ *such that* $\mathbb{A} \models \varphi$ *and* $\mathbb{B} \models \neg\varphi$ *contains* m *different branches* B'_i*, where* $1 \leq i \leq m$ *and* $B'_l \not\cong B'_j$ *for all* l *and* j *such that* $1 \leq l \neq j \leq m$*. Hence,* φ *has size at least* $(2 \times m) - 1$.

Proof. Since $\mathbb{A} \models \varphi$ and $\mathbb{B} \models \neg\varphi$, it follows that $\mathbb{A}^i \models \varphi$ and $\mathbb{B}^i \models \neg\varphi$ for any i such that $1 \leq i \leq m$. According to our assumption, the sequence of the m extended syntax trees of φ

$$T_\varphi^{\langle \mathbb{A}^1, \mathbb{B}^1 \rangle}, T_\varphi^{\langle \mathbb{A}^2, \mathbb{B}^2 \rangle}, \ldots, T_\varphi^{\langle \mathbb{A}^m, \mathbb{B}^m \rangle}$$

gives us a sequence of m branches $B_1, \ldots, B_m$, where B_1 is a branch in $T_\varphi^{\langle \mathbb{A}^1, \mathbb{B}^1 \rangle}$, B_2 is a branch in $T_\varphi^{\langle \mathbb{A}^2, \mathbb{B}^2 \rangle}$, …, B_m is a branch in $T_\varphi^{\langle \mathbb{A}^m, \mathbb{B}^m \rangle}$; moreover, $B_i \not\cong B_j$ for any i and j such that $1 \leq i \neq j \leq m$. It follows immediately that the syntax tree T_φ of φ contains m different branches. This means that φ contains at least $m - 1$ disjunctions that give us the m different branches in T_φ and, since each branch ends in a leaf labelled with a propositional symbol, we see that $|\varphi|$ is at least $(2 \times m) - 1$. □

The pairs $(\mathbb{A}^i, \mathbb{B}^i)$ are called diverging because, intuitively, they cannot be kept in the same branch in the extended syntax tree $T_\varphi^{\langle \mathbb{A}, \mathbb{B} \rangle}$. All our proofs in the next chapters rely on Theorem 3.25. The reader can easily verify that we used a similar technique in Chapter 2 as explained in Example 2.9.

Weight Function: The weight-function technique was used in [27] for proving lower bounds on the size of first-order formulae on linear orders. It can be thought of as an application of the idea of defining formal complexity measures in order to prove lower bounds on Boolean formulae (cf. Definition 8.1 on p. 259 in [65] and [52]). Here we give again an intuitive description of this technique by using Φ_{ML}-formulae. Let us suppose that we want to prove a lower bound on the size of the shortest formula $\varphi \in \Phi_{\mathrm{ML}}$ such that $\mathbb{A} \models \varphi$ and $\mathbb{B} \models \neg\varphi$. One way to achieve this is to define a weight function $w : 2^{\mathbf{K}} \times 2^{\mathbf{K}} \to \mathbb{N}$ such that $w = (\langle \mathbb{A}, \mathbb{B} \rangle) = n$, where $n > 0$, and for any pair of sets of pointed models $\langle \mathbb{C}, \mathbb{D} \rangle$, the following properties are true.

1. $w(\langle \mathbb{C}, \mathbb{D} \rangle) = w(\langle \mathbb{D}, \mathbb{C} \rangle)$;
2. if $\mathbb{C}_1 \subseteq \mathbb{C}$ and $\mathbb{C}_2 \subseteq \mathbb{C}$, then $w(\langle \mathbb{C}, \mathbb{D} \rangle) = w(\langle \mathbb{C}_1, \mathbb{D} \rangle) + w(\langle \mathbb{C}_2, \mathbb{D} \rangle)$;
3. if for some formula $[i]\psi \in \Phi_{\mathrm{ML}}$, we have $\mathbb{C} \models [i]\psi$ and $\mathbb{D} \models \langle i \rangle \neg\psi$, then $w(\langle [i]\mathbb{C}, \langle i(\neg\psi) \rangle \mathbb{D} \rangle) = w(\langle \mathbb{C}, \mathbb{D} \rangle) - 1$;
4. If $\mathbb{C} \models p$ and $\mathbb{D} \models \neg p$ for some propositional symbol p, then $w(\langle \mathbb{C}, \mathbb{D} \rangle) = 0$.

Let us consider the extended syntax tree $T_\varphi^{\langle \mathbb{A}, \mathbb{B} \rangle}$. The weight of the root $\mathbb{A} \circ \mathbb{B}$ of T is n, the weight of each leaf is 0. Items 1, 2, and 3 specify how the weight increases from the leaves to the root for nodes with syntax label $\neg$, $\vee$, and $[i]$, respectively. Then, it is obvious that $T_\varphi^{\langle \mathbb{A}, \mathbb{B} \rangle}$, and therefore T_φ, will have at least n nodes with syntax labels of the form $[i]$. Hence, φ contains at least n different occurrences of "$[i]$", that gives us $|\varphi| > n$.

We would like to stress that this artificial and very simple example was chosen so that we can showcase the main idea behind the weight function technique. We invite the reader to consult [27] for a technically sophisticated application of this method to proving lower bounds on the size of first-order formulae on the class of linear orders. For the original exposition of this method in the setting of Boolean function complexity, we recommend [65], where the idea of introducing such complexity measures or weight functions is attributed to Michael Paterson. Alexander Razborov [52] simplified the definition given in [65] and showed that practically all known complexity measures cannot help us prove even super-linear lower bounds on the size of Boolean circuits in sharp contrast to the case of Boolean formulae where they are used to prove quadratic lower bounds. A modern treatment of this material and its connections to communication complexity and graph theory can be found in [38].

We finish this chapter with some comments on the way we defined syntax trees and extended syntax trees of modal formulae. It is not difficult to introduce such trees for formulae containing symbols like $\top$, $\bot$, $\to$, or $\langle i \rangle$. We have chosen, however, to use only disjunctions and negations as Boolean operators and boxes as modal ones. The main motivation for our choice is that this is a popular way of defining modal formulae which helps us reduce the number of cases we have to consider in our proofs. On the other hand, since we are dealing with lower bounds on formula size, it is important to say that introducing $\top$, $\bot$, $\wedge$, $\rightarrow$, $\langle i \rangle$ or duals of $[\cup_\Gamma]$, $[\exists_\Gamma]$, etc. as official operators will lead only to a linear decrease of formula length and will not affect any of our results. In particular, all our succinctness theorems in the next chapter are proven as follows. We show that one modal logic L_1 is exponentially more succinct on a class of models $\mathbb{M}$ than another modal logic L_2 by exhibiting an infinite sequence of L_1-formulae $\varphi_1, \varphi_2, \ldots$ and a linear function $f(x)$ for which $|\varphi_n| = f(n)$; moreover, we prove that the shortest L_2-formula ψ_n that is equivalent to φ_n on $\mathbb{M}$ has size at least 2^n. Let us assume now that we are allowed to use the operators $\top$, $\bot$, $\wedge$, $\rightarrow$, $\langle i \rangle$, $\langle \cup_\Gamma \rangle$, $\langle \exists_\Gamma \rangle$, $\langle \cap_\Gamma \rangle$, and $\langle \varphi \rangle$ in L_2. Let us call the resulting logic L_2'. It is obvious that there is a natural number k such that for every L_2'-formula θ' there is an equivalent on $\mathbb{M}$ formula θ from L_2 for which $|\theta| \leq k \times |\theta'|$, i.e., there is a linear translation from L_2' to L_2 which works by expressing the above operators with $\neg$, $\vee$, $[i]$, $[\cup_\Gamma]$, $[\exists_\Gamma]$, $[\cap_\Gamma]$, and $[\varphi]$. Let us suppose that there is a sub-exponential function $g(x)$ and a sequence of L_2'-formulae $\vartheta_1', \vartheta_2' \ldots$ such that for all n, ϑ_n' is equivalent to φ_n and $|\vartheta_n'| = g(n)$. We know that for every n, there is an equivalent to ϑ_n' formula ϑ_n from L_2 such that $|\vartheta_n| \leq k \times |\vartheta_n'|$, which means that $|\vartheta_n| \leq k \times g(n)$. But this means that there is a sequence of L_2-formulae $\vartheta_1, \vartheta_2, \ldots$ such that ϑ_n is equivalent to φ_n and at the same time the length of the formulae $\vartheta_1, \vartheta_2, \ldots$ is bounded from above by the sub-exponential function $k \times g(x)$. This, however, contradicts our assumption that the shortest L_2-formula ψ_n that is equivalent to φ_n has size 2^n. Thus, we arrive at a contradiction.

Chapter 4

Succinctness Results on Unrestricted Models

In this and the next chapter we are going to compare in terms of their succinctness on different classes of Kripke models the logics ML, $[\cup]$ML, $[\exists]$ML, $[\cap]$ML, and $[\varphi]$ML. The initial motivation for this study came from Carten Lutz' paper [43] in which he showed that $[\varphi]$ML is exponentially more succinct than ML on a suitably chosen class of models that, unfortunately, was different from the class $\mathbf{S}_5$ of models in which all relations are relations of equivalence, i.e., the models used in epistemic logic. He conjectured that $[\varphi]$ML is exponentially more succinct than ML on $\mathbf{S}_5$, too. In the present chapter, we will give, among other things, another proof of his result while his conjecture will be confirmed in Chapter 5.

We are going to work with signatures that contain at least two relational indices and at least one propositional symbol. That is why we fix one such signature $S = (I, P)$, where $I = \{a, d\}$ and $P = \{\mathbf{b}\}$, and all formulae we consider are formulae in the signature S. The reader can find the following mnemonic useful later:

a stands for "solid arrow";
d stands for "dashed arrow";
$\mathbf{b}$ stands for "black node".

The lower bounds we obtain transfer in an obvious way to signatures with more indices and/or propositional symbols.

Our main results here can be summarised as follows.

A. There are sets of formulae

- $\Delta_{[\varphi]} \subset \Phi_{[\varphi]\text{ML}}$,
- $\Xi_{[\cup]} \subset \Phi_{[\cup]\text{ML}}$,

and a set of models $\mathbb{C} \subset \mathbf{K}$ such that $\Delta_{[\varphi]}$ and $\Xi_{[\cup]}$ are exponentially more succinct than $\Phi_{[\exists]\text{ML}}$ and $\Phi_{[\cap]\text{ML}}$ on $\mathbb{C}$.

B. There are sets of formulae

- $\Theta_{[\exists]} \subset \Phi_{[\exists]\mathrm{ML}}$,
- $\Psi_{[\cap]} \subset \Phi_{[\cap]\mathrm{ML}}$,
- $\Omega_{[\varphi]} \subset \Phi_{[\varphi]\mathrm{ML}}$,

and a set of models $\mathbb{D} \subset \mathbf{K}$ such that $\Theta_{[\exists]}$, $\Psi_{[\cap]}$, and $\Omega_{[\varphi]}$ are exponentially more succinct than $\Phi_{[\cup]\mathrm{ML}}$ on $\mathbb{D}$.

Of course, it follows immediately from **A** and **B** that the sets $\Delta_{[\varphi]}$ and $\Xi_{[\cup]}$ are exponentially more succinct than Φ_{ML} on $\mathbb{C}$ while $\Theta_{[\exists]}$, $\Psi_{[\cap]}$, and $\Omega_{[\varphi]}$ are exponentially more succinct than Φ_{ML} on $\mathbb{D}$. Furthermore, the results concerning the sets $\Delta_{[\varphi]} \subset \Phi_{[\varphi]\mathrm{ML}}$ and $\Omega_{[\varphi]} \subset \Phi_{[\varphi]\mathrm{ML}}$ can be considered as one particular way of strengthening Lutz' Theorem 2 from [43]. Moreover, it is clear from Proposition 3.8 and the following items

- $\mathbb{C} \subset \mathbf{K}$ and $\mathbb{D} \subset \mathbf{K}$,
- $\Xi_{[\cup]} \subset \Phi_{[\cup]\mathrm{ML}}$,
- $\Theta_{[\exists]} \subset \Phi_{[\exists]\mathrm{ML}}$,
- $\Delta_{[\varphi]} \subset \Phi_{[\varphi]\mathrm{ML}}$ and $\Omega_{[\varphi]} \subset \Phi_{[\varphi]\mathrm{ML}}$,
- $\Phi_{[\cup]\mathrm{ML}} \equiv_{\mathbf{K}} \Phi_{[\exists]\mathrm{ML}} \equiv_{\mathbf{K}} \Phi_{[\varphi]\mathrm{ML}}$

that

1. $\Phi_{[\cup]\mathrm{ML}}$ is exponentially more succinct than $\Phi_{[\exists]\mathrm{ML}}$ on $\mathbf{K}$;
2. $\Phi_{[\exists]\mathrm{ML}}$ is exponentially more succinct than $\Phi_{[\cup]\mathrm{ML}}$ on $\mathbf{K}$;
3. $\Phi_{[\varphi]\mathrm{ML}}$ is exponentially more succinct than both $\Phi_{[\cup]\mathrm{ML}}$ and $\Phi_{[\exists]\mathrm{ML}}$ on $\mathbf{K}$.

In fact, the work presented in this chapter was initially motivated by our wish to find proofs of items 1, 2, and 3 above and we feel that giving an informal explanation of the way in which we attacked these problems might help the reader's intuition.

One way of proving item 1 by using Lemma 3.10 is to find a sequence of pairwise non-equivalent formulae $\xi_1, \xi_2, \ldots$ in $\Phi_{[\cup]\mathrm{ML}}$ and a linear function $f(x)$ such that $|\xi_n| = f(n)$ for all $n \geq 1$. After that, we show that for any formula $\lambda_n \in \Phi_{[\exists]\mathrm{ML}}$ that is equivalent to ξ_n on $\mathbf{K}$, we have $|\lambda_n| \geq 2^n$. Having established this, Lemma 3.10 gives us the desired result. Let us elaborate on the strategy we just described. Suppose that we feel that we have found suitable formulae $\xi_1, \xi_2, \ldots$. In order to prove that they have the required property, for every $n \geq 1$, we can try to define a set of pointed models $\mathbb{A}^n$ such that $\mathbb{A}^n \models \xi_n$. Then we have to find a set of pointed models $\mathbb{B}^n$ such that $\mathbb{B}^n \models \neg\xi_n$ and prove that for every formula $\lambda \in \Phi_{[\exists]\mathrm{ML}}$ such that $\mathbb{A}^n \models \lambda$ whereas $\mathbb{B}^n \models \neg\lambda$, the extended syntax tree of λ with root $\mathbb{A}^n \circ \mathbb{B}^n$ has at least 2^n nodes. Intuitively, it is clear that the main difficulty in such a proof stems from the power of the $[\exists_\Gamma]$ operator.

While for any pointed model $(\mathscr{M}, w)$, the statement $(\mathscr{M}, w) \models [i]\varphi$ means that all points reachable from w in one i-step satisfy the formula φ, $(\mathscr{M}, w) \models [\exists_\Gamma]\varphi$ means that *there is at least one index* $i \in \Gamma$ such that *all* points reachable from w in one i-step satisfy φ. Therefore, if we manage to define the models in $\mathbb{A}^n$ and $\mathbb{B}^n$ in such a way so that to make the $[\exists_\Gamma]$ operator "useless" or "powerless", i.e., the possibility offered by $[\exists_\Gamma]$ to non-deterministically choose relation steps is eliminated, our task will be easier.

We can try to prove item 2 guided by the same intuition. Namely, we exhibit a sequence of pairwise non-equivalent formulae $\theta_1, \theta_2, \ldots$ from $\Phi_{[\exists]\mathrm{ML}}$ and a strictly increasing linear function $g(x)$ such that $|\theta_n| = g(n)$. Then we define two sets of pointed models $\mathbb{O}^n$ and $\mathbb{P}^n$, where $\mathbb{O}^n \models \theta_n$ and $\mathbb{P}^n \models \neg\theta_n$, and prove that the syntax tree with root $\mathbb{O}^n \circ \mathbb{P}^n$ of any formula $\mu \in \Phi_{[\cup]\mathrm{ML}}$ such that $\mathbb{O}^n \models \mu$ and $\mathbb{P}^n \models \neg\mu$ has at least 2^n nodes. As before, we define the models in $\mathbb{O}^n$ and $\mathbb{P}^n$ so that the operator $[\cup_\Gamma]$ is of no "use".

Finally, item 3 will come for free if we find two sequences of $\Phi_{[\varphi]\mathrm{ML}}$-formulae $\delta_1, \delta_2, \ldots$ and $\omega_1, \omega_2, \ldots$ and two strictly increasing linear functions $h(x)$ and $l(x)$ such that for all n,

- $|\delta_n| = h(n)$;
- $|\omega_n| = l(n)$;
- $\mathbb{A}^n \models \delta_n$ and $\mathbb{B}^n \models \neg\delta_n$;
- $\mathbb{O}^n \models \omega_n$ and $\mathbb{B}^n \models \neg\omega_n$.

After this brief introduction, we begin our work by first defining the sets of formulae and the sets of models mentioned in items **A** and **B** above followed by complete proofs of the stated succinctness results.

Let us begin by defining the sets of formulae and the sets of models from item **A**. The sets of formulae $\Delta_{[\varphi]} \subset \Phi_{[\varphi]\mathrm{ML}}$ and $\Xi_{[\cup]} \subset \Phi_{[\cup]\mathrm{ML}}$are defined as follows.

$\Delta_{[\varphi]}$		$\Xi_{[\cup]}$	
$\delta_1 \stackrel{\text{def}}{=}$	$\langle a\rangle\mathbf{b} \vee \langle d\rangle\mathbf{b}$	$\xi_1 \stackrel{\text{def}}{=}$	$\neg[\cup_{\{a,d\}}]\neg\mathbf{b}$
$\vdots$		$\vdots$	
$\delta_n \stackrel{\text{def}}{=}$	$\langle\delta_{n-1}\rangle\delta_1$	$\xi_n \stackrel{\text{def}}{=}$	$\neg\underbrace{[\cup_{\{a,d\}}]\ldots[\cup_{\{a,d\}}]}_{n \text{ times}}\neg\mathbf{b}$
$\vdots$		$\vdots$	

TABLE 4.1: The sets of formulae $\Delta_{[\varphi]}$ and $\Xi_{[\cup]}$.

Note that, technically speaking, the set $\Delta_{[\varphi]}$ consists of formulae that are obtained from the relevant formulae in Table 4.1 by replacing the defined operators $\langle\delta_i\rangle$ with $\neg[\delta_i]\neg$ and $\langle a\rangle$ and $\langle d\rangle$ with $\neg[a]\neg$ and $\neg[d]\neg$, respectively. It is obvious that the lengths of the formulae in $\Delta_{[\varphi]}$ and $\Xi_{[\cup]}$ are bounded from above by a linear function.

Definition 4.1 (The models $\mathbb{A}^n$ and $\mathbb{B}^n$)**.** For every natural number $n \geq 1$, the sets of pointed models $\mathbb{A}^n$ (containing 2^n different models) and $\mathbb{B}^n$ (containing a single model) are defined as follows.

1. The set $\mathbb{A}^1$ consists of the two pointed models $(\mathscr{A}_a^1, \alpha_a^1)$ and $(\mathscr{A}_d^1, \alpha_d^1)$ shown on the left of the dotted line in Figure 4.1. The set $\mathbb{B}^1$ contains only one pointed model namely, $(\mathscr{B}^1, \beta^1)$ that is shown on the right of the dotted line.

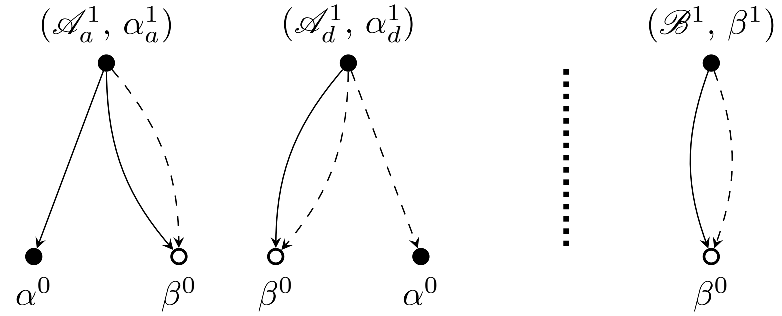


FIGURE 4.1: The sets of pointed models $\mathbb{A}^1$ and $\mathbb{B}^1$.

The black nodes satisfy the proposition **b** whereas the white nodes do not. The subscripts in the names of the pointed Kripke models encode the way a black node can be reached from the uppermost node which is denoted by α with the relevant subscripts and superscripts in the case of the models in $\mathbb{A}^1$ and by β with a superscript in the case of the model in $\mathbb{B}^1$. For example, in the model $(\mathscr{A}_a^1, \alpha_a^1)$, a black node (namely, α^0) can be reached from α_a^1 by making one step along the relation R_a represented by the arrow connecting these two nodes. In the model $(\mathscr{A}_d^1, \alpha_d^1)$, a black node can be reached from the node α_d^1 by making one step along the relation R_d represented by the dashed arrow connecting the two nodes.

$n+1$. The set $\mathbb{A}^{n+1}$ consists of all the models built from the models in $\mathbb{A}^n \cup \mathbb{B}^n$ as shown in the Figure 4.2 below on the left of the dotted vertical line.

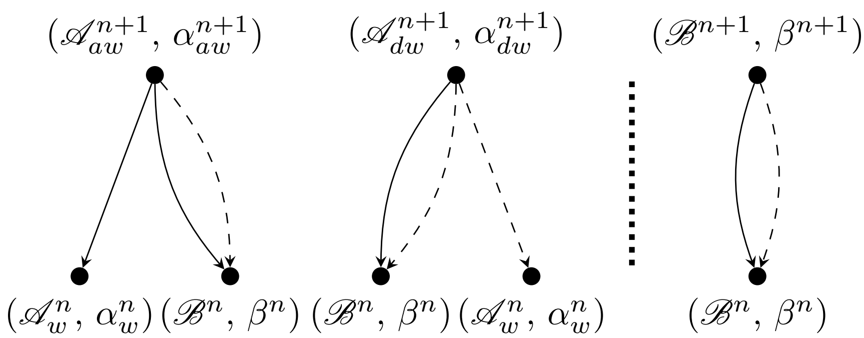


FIGURE 4.2: The sets of models $\mathbb{A}^{n+1}$ and $\mathbb{B}^{n+1}$.

For any pointed model $(\mathscr{A}^n_w, \alpha^n_w) \in \mathbb{A}$, the pointed model $(\mathscr{A}^{n+1}_{aw}, \alpha^{n+1}_{aw})$ is obtained by taking a black node, denoted by α^{n+1}_{aw}, and connecting it to the point α^n_w in $\mathscr{A}^n_w$ and the point β^n in the model $\mathscr{B}^n$ as shown. The pointed model $(\mathscr{A}^{n+1}_{dw}, \alpha^{n+1}_{dw})$ is constructed in a similar fashion. The set $\mathbb{B}^{n+1}$ contains only the model $(\mathscr{B}^{n+1}, \beta^{n+1})$ shown on the right of the dotted line. Again, black nodes satisfy the proposition **b** whereas white nodes do not.

Intuitively, the subscript w and the superscript n in the pointed model $(\mathscr{A}^n_w, \alpha^n_w) \in \mathbb{A}^n$ say that there is a sequence of relation steps of length n encoded by w, leading from the uppermost point of $\mathscr{A}^n_w$, i.e., α^n_w, to the only lowermost node satisfying the proposition **b**. Since there are 2^n different words w of length n over the alphabet $\{a, d\}$ and for every such w, there is a corresponding pointed model in $\mathbb{A}^n$, this means that $\mathbb{A}^n$ contains 2^n different pointed models.

The most important property of these models is the following. Let w_1 be a (possibly empty) word over the alphabet $\{a, d\}$. For every $(\mathscr{A}^n_w, \alpha^n_w) \in \mathbb{A}^n$, where $w = aw_1$, it is true that $(\mathscr{A}^n_w, \alpha^n_w)$ and $(\mathscr{B}^n, \beta^n)$ are bisimilar with respect to d and $a \cap d$. Similarly, if $w = dw_1$, then $(\mathscr{A}^n_w, \alpha^n_w)$ and $(\mathscr{B}^n, \beta^n)$ are bisimilar with respect to a and $a \cap d$. We hope that the next example makes our definition completely transparent and intuitively clear.

Example 4.1. *The models $(\mathscr{A}^2_{aa}, \alpha^2_{aa})$ and $(\mathscr{B}^2, \beta^2)$ that are bisimilar with respect to d and $a \cap d$ are shown in Figure 4.3 below. Note how the pointed models $\mathscr{A}^1_a$ and $\mathscr{B}^1$ are used in the construction of $\mathscr{A}^2_{aa}$.*

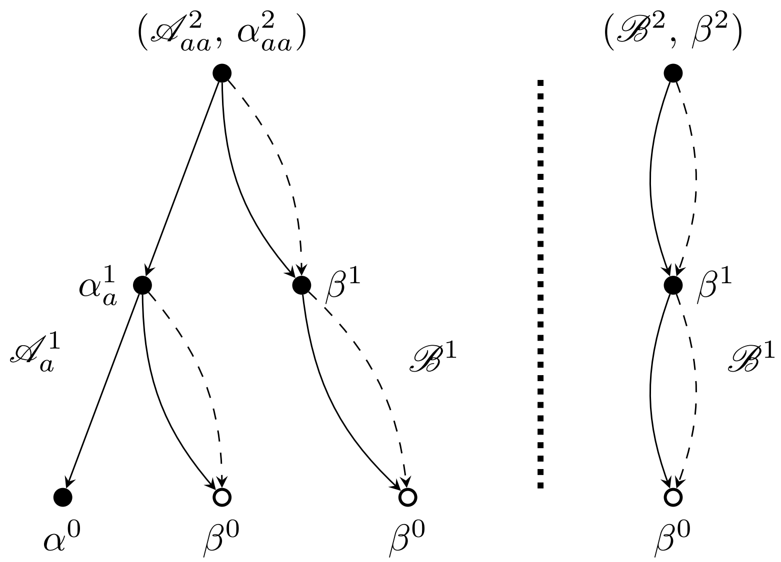


FIGURE 4.3: The pointed models $(\mathscr{A}^2_{aa}, \alpha^2_{aa})$ and $(\mathscr{B}^2, \beta^2)$.

Items 2 and 3 from the next proposition are a formal expression of our idea of making the operators $[\exists_\Gamma]$ and $[\cap_\Gamma]$ "powerless". Intuitively, no formula $[\exists_{\{a,d\}}]\varphi \in \Phi_{[\cup]\mathrm{ML}}$ or $[\cap_{\{a,d\}}]\varphi \in \Phi_{[\cap]\mathrm{ML}}$ can differentiate between $(\mathscr{A}^n_w, \alpha^n_w)$ and $(\mathscr{B}^n, \beta^n)$.

Proposition 4.2. *The following are true.*

1. $\mathbb{A}^n \models \delta_n$ *and* $\mathbb{A}^n \models \xi_n$ *whereas* $\mathbb{B}^n \models \neg\delta_n$ *and* $\mathbb{B}^n \models \neg\xi_n$.

2. *For any model* $(\mathscr{A}^{n+1}_{jw}, \alpha^{n+1}_{jw}) \in \mathbb{A}^{n+1}$, *where* $n \geq 0$ *and* $j \in \{a, d\}$, *and any formula* $[\exists_{\{a,d\}}]\varphi \in \Phi_{[\exists]\mathrm{ML}}$,

$$(\mathscr{A}^{n+1}_{jw}, \alpha^{n+1}_{jw}) \models [\exists_{\{a,d\}}]\varphi \textit{ iff } (\mathscr{B}^{n+1}, \beta^{n+1}) \models [\exists_{\{a,d\}}]\varphi.$$

3. *For any model* $(\mathscr{A}^{n+1}_{jw}, \alpha^{n+1}_{jw}) \in \mathbb{A}^{n+1}$ *and any formula* $[\cap_{\{a,d\}}]\varphi \in \Phi_{[\cap]\mathrm{ML}}$,

$$(\mathscr{A}^{n+1}_{jw}, \alpha^{n+1}_{jw}) \models [\cap_{\{a,d\}}]\varphi \textit{ iff } (\mathscr{B}^{n+1}, \beta^{n+1}) \models [\cap_{\{a,d\}}]\varphi.$$

Proof.

1. The proof of this item is by an easy induction on n. The case for δ_n can be proven by noting that δ_{n+1} is equivalent to $\delta_n \wedge (\langle a \rangle(\mathbf{b} \wedge \delta_n) \vee \langle d \rangle(\mathbf{b} \wedge \delta_n))$ which shows us how to construct the inductive argument. The reader can grasp the main idea immediately from verifying that the model $(\mathscr{A}^2_{aa}, \alpha^2_{aa})$ from Figure 4.3 satisfies the formula $\delta_1 \wedge (\langle a \rangle(\mathbf{b} \wedge \delta_1) \vee \langle d \rangle(\mathbf{b} \wedge \delta_1))$ that is equivalent to δ_2 whereas $(\mathscr{B}^2, \beta^2)$ does not.

   The case for ξ_n is also easy. It is obvious that the models $(\mathscr{A}^1_a, \alpha^1_a)$ and $(\mathscr{A}^1_d, \alpha^1_d)$ (see Figure 4.1) both satisfy ξ_1 which is equivalent to $\langle a \rangle\mathbf{b} \vee \langle d \rangle\mathbf{b}$ whereas $(\mathscr{B}^1, \beta^1)$ does not. Furthermore, we have that ξ_{n+1} is equivalent to $\langle a \rangle\xi_n \vee \langle d \rangle\xi_n$. Hence, the inductive hypothesis that $\mathbb{A}^n \models \xi_n$ whereas $\mathbb{B}^n \models \neg\xi_n$ and the construction of the models in $\mathbb{A}^{n+1}$ and $\mathbb{B}^{n+1}$ (see Figure 4.2) give us the desired conclusion.

2. Let $n \geq 0$, $j = a$, and w be (a possibly empty) word over the alphabet $\{a, d\}$. The case $j = d$ is analogous.

   (If) Let us suppose that there is a formula $[\exists_{\{a,d\}}]\varphi \in \Phi_{[\exists]\mathrm{ML}}$ such that

   $$(\mathscr{B}^{n+1}, \beta^{n+1}) \models [\exists_{\{a,d\}}]\varphi \text{ and } (\mathscr{A}^{n+1}_{aw}, \alpha^{n+1}_{aw}) \models \neg[\exists_{\{a,d\}}]\varphi.$$

   This means that

   $$(\mathscr{B}^{n+1}, \beta^{n+1}) \models [a]\varphi \vee [d]\varphi \text{ and } (\mathscr{A}^{n+1}_{aw}, \alpha^{n+1}_{aw}) \models \langle a \rangle\neg\varphi \wedge \langle d \rangle\neg\varphi.$$

   However, $(\mathscr{B}^{n+1}, \beta^{n+1})$ and $(\mathscr{A}^{n+1}_{aw}, \alpha^{n+1}_{aw})$ are bisimilar with respect to d, i.e., the relation R_d represented by the dashed arrow (see Figures 4.1 and 4.2) and therefore, using Theorem 3.18, we obtain $(\mathscr{B}^{n+1}, \beta^{n+1}) \models \langle d \rangle\neg\varphi$. This, together with our assumption that $(\mathscr{B}^{n+1}, \beta^{n+1}) \models [a]\varphi \vee [d]\varphi$, implies that $(\mathscr{B}^{n+1}, \beta^{n+1}) \models [a]\varphi \wedge \langle d \rangle\neg\varphi$. However, the only a-successor of β^{n+1} is β^n, for which we have $(\mathscr{B}^{n+1}, \beta^n) \models \neg\varphi$ because it is the only d-successor of β^{n+1} and $(\mathscr{B}^{n+1}, \beta^{n+1}) \models \langle d \rangle\neg\varphi$. Thus, we arrive at a contradiction.

(Only if) Again, we assume the contrary. As before, this means that for some formula $[\exists_{\{a,d\}}]\varphi \in \Phi_{[\exists]\mathrm{ML}}$, we have

$$(\mathscr{A}_{aw}^{n+1}, \alpha_{aw}^{n+1}) \models [a]\varphi \vee [d]\varphi \text{ and } (\mathscr{B}^{n+1}, \beta^{n+1}) \models \langle a\rangle\neg\varphi \wedge \langle d\rangle\neg\varphi.$$

Similar reasoning as the one above shows that $(\mathscr{A}_{aw}^{n+1}, \alpha_{aw}^{n+1}) \models \langle d\rangle\neg\varphi$. Therefore, $(\mathscr{A}_{aw}^{n+1}, \alpha_{aw}^{n+1}) \models [a]\varphi$ and thus, $(\mathscr{A}_{aw}^{n+1}, \beta^{n}) \models \varphi$. However, it follows that $(\mathscr{A}_{aw}^{n+1}, \beta^{n}) \models \neg\varphi$ because $(\mathscr{A}_{aw}^{n+1}, \alpha_{aw}^{n+1}) \models \langle d\rangle\neg\varphi$. Again, we arrive at a contradiction.

3. The proof of this item follows from Theorem 3.18 and the fact that the models in $\mathbb{A}^n$ and the model in $\mathbb{B}^n$ are bisimilar with respect to $a \cap d$, i.e, the relation $R_{a\cap d}$.

□

Thus, for every $n \geq 1$, we can differentiate between any pointed model in $\mathbb{A}^n$ and the only pointed model in $\mathbb{B}^n$ by using the formulae ξ_n and δ_n.

We continue by defining the sets of formulae and the sets of pointed models from item **B**. Let the sets of formulae $\Theta_{[\exists]}$, $\Psi_{[\cap]}$ and $\Omega_{[\varphi]}$ be defined as in Table 4.2.

$\Theta_{[\exists]}$		$\Psi_{[\cap]}$		$\Omega_{[\varphi]}$	
$\theta_1 \stackrel{\text{def}}{=}$	$[\exists_{\{a,d\}}]\mathbf{b}$	$\psi_1 \stackrel{\text{def}}{=}$	$[\cap_{\{a,d\}}]\mathbf{b}$	$\omega_1 \stackrel{\text{def}}{=}$	$[a]\mathbf{b} \vee [d]\mathbf{b}$
$\vdots$		$\vdots$		$\vdots$	
$\theta_n \stackrel{\text{def}}{=}$	$\underbrace{[\exists_{\{a,d\}}]\ldots[\exists_{\{a,d\}}]}_{n \text{ times}}\mathbf{b}$	$\psi_n \stackrel{\text{def}}{=}$	$\underbrace{[\cap_{\{a,d\}}]\ldots[\cap_{\{a,d\}}]}_{n \text{ times}}\mathbf{b}$	$\omega_n \stackrel{\text{def}}{=}$	$\langle\omega_{n-1}\rangle(\langle a\rangle\mathbf{b} \wedge \langle d\rangle\mathbf{b})$
$\vdots$		$\vdots$		$\vdots$	

TABLE 4.2: The sets of formulae $\Theta_{[\exists]}$, $\Psi_{[\cap]}$, and $\Omega_{[\varphi]}$.

Again, we have that the actual formulae in $\Omega_{[\varphi]}$ are obtained by replacing the defined symbols $\langle\omega_n\rangle$, $\langle a\rangle$, and $\langle d\rangle$ with $\neg[\omega_n]\neg$, $\neg[a]\neg$, and $\neg[d]\neg$, respectively. It is obvious that the length of all the formulae in $\Theta_{[\exists]}$, $\Psi_{[\cap]}$, and $\Omega_{[\varphi]}$ is given by a linear function in their indices.

Following the conventions we used in Definition 4.1, for every natural number n, we construct two sets of pointed models $\mathbb{O}^n$ and $\mathbb{P}^n$.

Definition 4.3 (The models $\mathbb{O}^n$ and $\mathbb{P}^n$)**.** The two sets of pointed models $\mathbb{O}^n$ and $\mathbb{P}^n$, containing 2^n different models each, are built recursively as follows.

1. The set $\mathbb{O}^1$ consists of the two pointed models $(\mathscr{O}_a^1, o_a^1)$ and $(\mathscr{O}_d^1, o_d^1)$ shown on the left of the dotted line in Figure 4.4 below. The set $\mathbb{P}^1$ contains the models $(\mathscr{P}_a^1, \rho_a^1)$ and $(\mathscr{P}_d^1, \rho_d^1)$ shown on the right of the dotted line. As in the case of Definition 4.1, black points denote nodes where the proposition **b** is true while white points denote nodes that do not satisfy **b**; dashed arrows represent the relation R_d whereas non-dashed arrows represent the relation R_a.

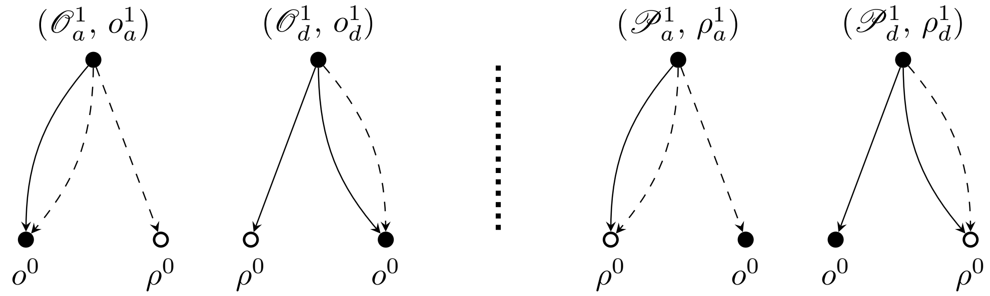


FIGURE 4.4: The models $\mathbb{O}^1$ and $\mathbb{P}^1$.

$n+1$. The sets $\mathbb{O}^{n+1}$ and $\mathbb{P}^{n+1}$ are built as shown in Figure 4.5. We follow an algorithmic pattern analogous to the one we used in the construction of the models in $\mathbb{A}^{n+1}$ and $\mathbb{B}^{n+1}$ from Definition 4.1. It should be clear that each of the sets $\mathbb{O}^n$ and $\mathbb{P}^n$ contains 2^n different pointed models - one for each subscript w of length n.

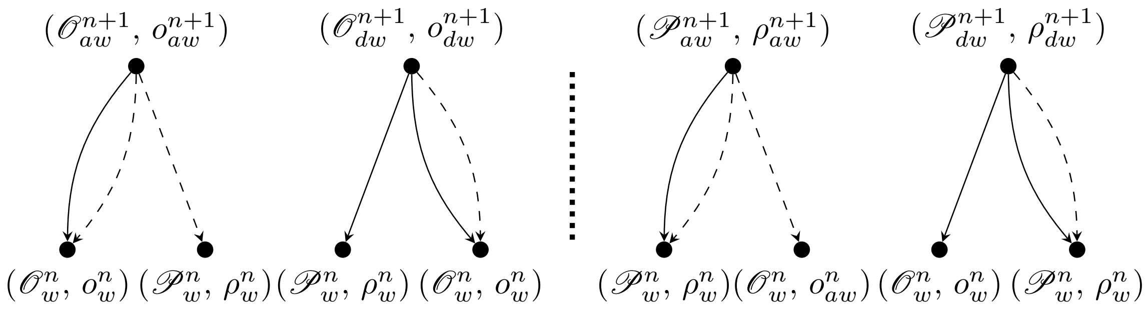


FIGURE 4.5: The sets of models $\mathbb{O}^n$ and $\mathbb{P}^n$ where $n > 1$.

Intuitively, as before, for any pair of pointed models $(\mathscr{O}^n_w, o^n_w)$ and $(\mathscr{P}^n_w, \rho^n_w)$, we have that the subscript w encodes a sequence of n relation steps that lead from o^n_w to a black point o^0; the same sequence w leads from ρ^n_w to a white point ρ^0.

Again, the most important property of the models in $\mathbb{O}^n$ and $\mathbb{P}^n$ is the following. For any (possibly empty) word w_1 over the alphabet $\{a, d\}$, the pair of models $(\mathscr{O}^n_w, o^n_w) \in \mathbb{O}^n$, $(\mathscr{P}^n_w, o^n_w) \in \mathbb{P}^n$, where $w = aw_1$, are bisimilar with respect to d. Similarly, if $w = dw_1$, then the models $(\mathscr{O}^n_w, o^n_w)$ and $(\mathscr{P}^n, \rho^n_w)$ are bisimilar with respect to a.

Example 4.2. *The pair of models $(\mathscr{O}^2_{da}, o^2_{da})$ and $(\mathscr{P}^2_{da}, \rho^2_{da})$ are shown in Figure 4.6 below. The subscript "da" and the superscript "2" in $(\mathscr{O}^2_{da}, o^2_{da})$ mean that starting at o^2_{da} and making one step along the relation R_d (represented by the dashed arrow) followed by a step along R_a (represented by the solid arrow), we arrive at a black point, i.e., a point satisfying the proposition* **b**. *The same sequence of relation steps "da" leads to a white point, i.e., a point that does not satisfy the proposition* **b** *from ρ^2_{da} in the Kripke model $\mathscr{P}^2_{da}$. Note that $(\mathscr{O}^2_{da}, o^2_{da})$ and $(\mathscr{P}^2_{da}, \rho^2_{da})$ are bisimilar with respect to a, i.e., the relation R_a represented by the solid arrow.*

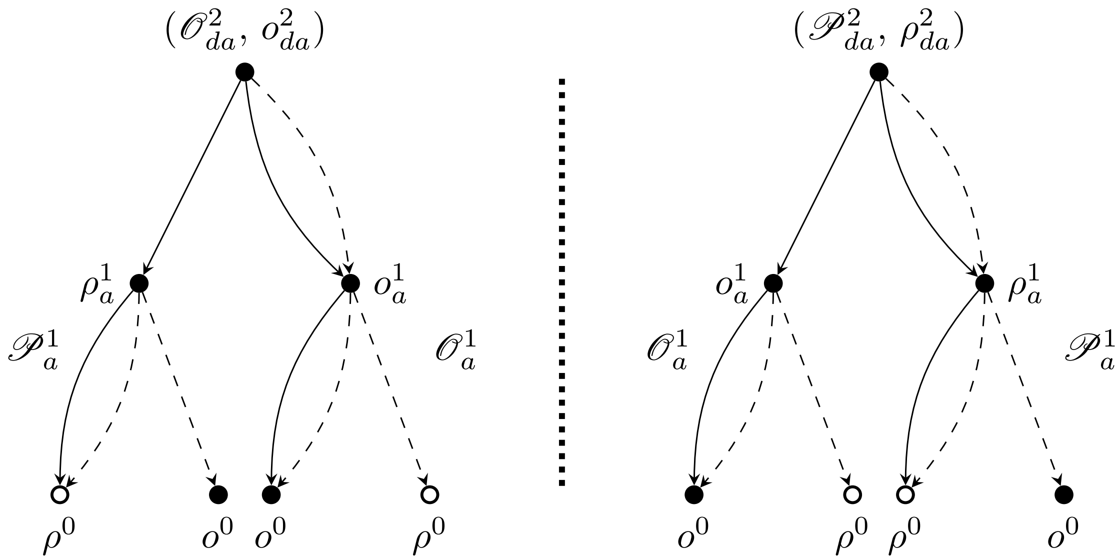


FIGURE 4.6: The pointed models $(\mathscr{O}^2_{da}, o^2_{da})$ and $(\mathscr{P}^2_{da}, \rho^2_{da})$.

Analogously to Proposition 4.2, we have the next proposition where the second item is an expression of our idea to "neutralise" the $[\cup_\Gamma]$ operator.

Proposition 4.4. *The following are true.*

1. $\mathbb{O}^n \models \theta_n$ *and* $\mathbb{O}^n \models \psi_n$*, and* $\mathbb{O}^n \models \omega_n$ *whereas* $\mathbb{P}^n \models \neg\theta_n$ *and* $\mathbb{P}^n \models \neg\psi_n$*, and* $\mathbb{P}^n \models \neg\omega_n$*.*
2. *For any pair of models* $(\mathscr{O}^{n+1}_{jw}, o^{n+1}_{jw}) \in \mathbb{O}^{n+1}$ *and* $(\mathscr{P}^{n+1}_{jw}, \rho^{n+1}_{jw}) \in \mathbb{P}^{n+1}$*, where* $n \geq 0$ *and* $j \in \{a, d\}$*, and any formula* $[\cup_{\{a,d\}}]\varphi \in \Phi_{[\cup]\mathrm{ML}}$*,*

$$(\mathscr{O}^{n+1}_{jw}, o^{n+1}_{jw}) \models [\cup_{\{a,d\}}]\varphi \textit{ iff } (\mathscr{P}^{n+1}_{jw}, \rho^{n+1}_{jw}) \models [\cup_{\{a,d\}}]\varphi.$$

Proof.

1. The proof of this item is a very easy but tedious inductive argument. It is based on the shape of the models in $\mathbb{O}^n$ and $\mathbb{P}^n$, and the following facts.
   - θ_1 is equivalent to $[a]\mathbf{b} \vee [d]\mathbf{b}$ and θ_{n+1} is equivalent to $[a]\theta_n \vee [d]\theta_n$.
   - ω_{n+1} is equivalent to $\omega_n \wedge (\langle a\rangle(\mathbf{b} \wedge \omega_n) \wedge \langle d\rangle(\mathbf{b} \wedge \omega_n))$
2. As in the proof of the second item of Proposition 4.2, let $n \geq 0$ and w be (a possibly empty) word over the alphabet $\{a, d\}$. We consider the case $j = d$. When $j = a$, the proof is analogous.

   (If) Let us suppose that for some formula $[\cup_{\{a,d\}}]\varphi \in \Phi_{[\cup]\mathrm{ML}}$, it is true that

$$(\mathscr{P}^{n+1}_{dw}, \rho^{n+1}_{dw}) \models [\cup_{\{a,d\}}]\varphi \text{ and } (\mathscr{O}^{n+1}_{dw}, o^{n+1}_{dw}) \models \neg[\cup_{\{a,d\}}]\varphi.$$

This implies

$$(\mathscr{P}^{n+1}_{dw}, \rho^{n+1}_{dw}) \models [a]\varphi \wedge [d]\varphi \text{ and } (\mathscr{O}^{n+1}_{dw}, o^{n+1}_{dw}) \models \langle a\rangle\neg\varphi \vee \langle d\rangle\neg\varphi.$$

It is obvious that $(\mathscr{P}^{n+1}_{dw}, \rho^{n+1}_{dw})$ and $(\mathscr{O}^{n+1}_{dw}, o^{n+1}_{dw})$ are bisimilar with respect to a, i.e., the relation R_a represented by the solid arrow. Hence, our assumption that $(\mathscr{P}^{n+1}_{dw}, \rho^{n+1}_{dw}) \models [a]\varphi$ and Theorem 3.18 imply that $(\mathscr{O}^{n+1}_{dw}, o^{n+1}_{dw}) \models [a]\varphi$ and therefore, $(\mathscr{O}^{n+1}_{dw}, o^{n+1}_{dw}) \models [a]\varphi \wedge \langle d\rangle\neg\varphi$. The only d-successor of o^{n+1}_{dw} is the point o^n_w which means that $(\mathscr{O}^{n+1}_{dw}, o^n_w) \models \neg\varphi$. On the other hand, since $(\mathscr{O}^{n+1}_{dw}, o^{n+1}_{dw}) \models [a]\varphi$ and o^n_w is an a-successor of o^{n+1}_{dw}, we have $(\mathscr{O}^{n+1}_{dw}, o^n_w) \models \varphi$ which is a contradiction.

(Only if) Let us assume that there is a formula $[\cup_{\{a,d\}}]\varphi \in \Phi_{[\cup]\mathrm{ML}}$ such that

$$(\mathscr{O}^{n+1}_{dw}, o^{n+1}_{dw}) \models [\cup_{\{a,d\}}]\varphi \text{ and } (\mathscr{P}^{n+1}_{dw}, \rho^{n+1}_{dw}) \models \neg[\cup_{\{a,d\}}]\varphi.$$

Therefore,

$$(\mathscr{O}^{n+1}_{dw}, o^{n+1}_{dw}) \models [a]\varphi \wedge [d]\varphi \text{ and } (\mathscr{P}^{n+1}_{dw}, \rho^{n+1}_{dw}) \models \langle a\rangle\neg\varphi \vee \langle d\rangle\neg\varphi.$$

Similar reasoning as above shows that $(\mathscr{P}^{n+1}_{dw}, \rho^{n+1}_{dw}) \models [a]\varphi \wedge \langle d\rangle\neg\varphi$. Note however, that ρ^n_w is both an a and a d-successor of ρ^{n+1}_{dw} and thus, we arrive at a contradiction.

□

Now, we are ready to state and prove the main theorems in this chapter. To this end, we are going to use the following notation. Let $\mathbb{A}$ and $\mathbb{B}$ denote the union of all $\mathbb{A}^n$ and $\mathbb{B}^n$, respectively, i.e.,

$$\mathbb{A} = \bigcup_{n\geq 1} \mathbb{A}^n \text{ and } \mathbb{B} = \bigcup_{n\geq 1} \mathbb{B}^n.$$

Similarly,

$$\mathbb{O} = \bigcup_{n\geq 1} \mathbb{O}^n \text{ and } \mathbb{P} = \bigcup_{n\geq 1} \mathbb{P}^n.$$

Theorem 4.5. *Let the sets of formulae $\Delta_{[\varphi]}$ and $\Xi_{[\cup]}$ be defined as in Table 4.1. Then*

1. *$\Delta_{[\varphi]}$ is exponentially more succinct than $\Phi_{[\exists]\mathrm{ML}}$ and $\Phi_{[\cap]\mathrm{ML}}$ on $\mathbb{A} \cup \mathbb{B}$;*
2. *$\Xi_{[\cup]}$ is exponentially more succinct than $\Phi_{[\exists]\mathrm{ML}}$ and $\Phi_{[\cap]\mathrm{ML}}$ on $\mathbb{A} \cup \mathbb{B}$.*

Theorem 4.6. *Let the sets of formulae $\Theta_{[\exists]}$, $\Psi_{[\cap]}$, and $\Omega_{[\varphi]}$ be defined as in Table 4.2. Then $\Theta_{[\exists]}$, $\Psi_{[\cap]}$, and $\Omega_{[\varphi]}$ are exponentially more succinct than $\Phi_{[\cup]\mathrm{ML}}$ on $\mathbb{O} \cup \mathbb{P}$.*

Note that, since it is claimed in Theorem 4.6 that $\Psi_{[\cap]}$ is exponentially more succinct than $\Phi_{[\cup]\mathrm{ML}}$, we must make sure that, for every formula $\psi_n \in \Psi_{[\cap]}$, there is an equivalent on $\mathbb{O} \cup \mathbb{P}$ formula from $\Phi_{[\cup]\mathrm{ML}}$. Indeed, the reader can easily verify that each $\psi_n \in \Psi_{[\cap]}$ is equivalent on $\mathbb{O} \cup \mathbb{P}$ to the formula $\theta_n \in \Theta_{[\exists]}$.

The main argument in the proofs of both theorems is supplied by Theorem 3.25 and Lemma 4.7 below, i.e., our reasoning is an application of the Diverging Pairs technique to the models in $\mathbb{A}^n$, $\mathbb{B}^n$ and $\mathbb{O}^n$, $\mathbb{P}^n$. An informal explanation based on the models in $\mathbb{O}^n$ and $\mathbb{P}^n$ is as follows.

We know that there are two pointed models in the set $\mathbb{O}^1$, namely $(\mathscr{O}^1_a, o^1_a)$ and $(\mathscr{O}^1_d, o^1_d)$. Similarly, there are two corresponding models $(\mathscr{P}^1_a, \rho^1_a)$ and $(\mathscr{P}^1_d, \rho^1_d)$ in $\mathbb{P}^1$.

Let $\mu_a \in \Phi_{[\cup]\mathrm{ML}}$ be a formula such that $(\mathscr{O}^1_a, o^1_a) \models \mu_a$ and $(\mathscr{P}^1_a, \rho^1_a) \models \neg\mu_a$. Similarly, let $\mu_d \in \Phi_{[\cup]\mathrm{ML}}$ be a formula for which $(\mathscr{O}^1_b, o^1_d) \models \mu_d$ and $(\mathscr{P}^1_d, \rho^1_d) \models \neg\mu_d$.

If we prove that

1. there is a branch B_a such that $I(B_a) = a$ in the extended syntax tree of μ_a with root $\{(\mathscr{O}^1_a, o^1_a)\} \circ \{(\mathscr{P}^1_a, \rho^1_a)\}$,
2. the extended syntax tree of μ_d with root $\{(\mathscr{O}^1_d, o^1_d)\} \circ \{(\mathscr{P}^1_d, \rho^1_d)\}$ has a branch B_d for which $I(B_d) = d$,

then it follows from Theorem 3.25 that the syntax tree of any formula $\mu \in \Phi_{[\cup]\mathrm{ML}}$ such that $\mathbb{O}^1 \models \mu$ and $\mathbb{P}^1 \models \neg\mu$ must contain the two branches described in the items above, and, therefore $|\mu| \geq 3$. Generalising this way of reasoning, we see that the extended syntax tree of any formula $\mu_n \in \Phi_{[\cup]\mathrm{ML}}$ for which $\mathbb{O}^n \models \mu_n$ and $\mathbb{P}^n \models \neg\mu_n$ contains 2^n different branches B_w, each one corresponding to the pair $(\mathscr{O}^n_w, o^n_w)$ and $(\mathscr{P}^n_w, \rho^n_w)$. The formalisation of this intuition is given in Lemma 4.7 below.

Lemma 4.7. *For any $j \in \{a, d\}$, if w is a (possibly empty) word of length $n \geq 0$ over the alphabet $\{a, d\}$, then the following are true.*

1. *For any pointed model $(\mathscr{A}^{n+1}_{jw}, \alpha^{n+1}_{jw}) \in \mathbb{A}^n$,*
   (a) *the syntax tree of any formula $\lambda \in \Phi_{[\exists]\mathrm{ML}}$ such that $(\mathscr{A}^{n+1}_{jw}, \alpha^{n+1}_{jw}) \models \lambda$ and $(\mathscr{B}^{n+1}, \beta^{n+1}) \models \neg\lambda$ contains a branch B for which $I(B) = jw$;*
   (b) *the syntax tree of any formula $\nu \in \Phi_{[\cap]\mathrm{ML}}$ such that $(\mathscr{A}^{n+1}_{jw}, \alpha^{n+1}_{jw}) \models \nu$ and $(\mathscr{B}^{n+1}, \beta^{n+1}) \models \neg\nu$ contains a branch B for which $I(B) = jw$.*
2. *For any pair of pointed models $(\mathscr{O}^{n+1}_{jw}, o^{n+1}_{jw}) \in \mathbb{O}^n$ and $(\mathscr{P}^{n+1}_{jw}, \rho^{n+1}_{jw}) \in \mathbb{P}^n$, the syntax tree of any formula $\mu \in \Phi_{[\cup]\mathrm{ML}}$ such that $(\mathscr{O}^{n+1}_{jw}, o^{n+1}_{jw}) \models \mu$ and $(\mathscr{P}^{n+1}_{jw}, \rho^{n+1}_{jw}) \models \neg\mu$, contains a branch B, such that $I(B) = jw$.*

Proof. Since the proofs of both items 1 and 2 are completely analogous, we deal with the latter. The argument for the former goes along the same lines but with the use of the second and third items from Proposition 4.2. Furthermore, the considerations in the cases $j = a$ and $j = d$ are the same modulo replacing a with d and that is why we reason only about $j = d$.

Before presenting the proof, we make explicit the main intuition behind it with the help of the models $(\mathscr{O}^2_{da}, o^2_{da})$ and $(\mathscr{P}^2_{da}, \rho^2_{da})$ from Figure 4.6.

Example 4.3. *Let* $T_\mu^{\langle\{(\mathscr{O}^2_{da},o^2_{da})\},\{(\mathscr{P}^2_{da},\rho^2_{da})\}\rangle}$ *be the extended syntax tree of a formula* μ *such that* $(\mathscr{O}^2_{da}, o^2_{da}) \models \mu$ *and* $(\mathscr{P}^2_{da}, \rho^2_{da}) \models \neg\mu$*. Then the first node of the branch* B *we want to construct is the root* $\{(\mathscr{O}^2_{da}, o^2_{da})\} \circ \{(\mathscr{P}^2_{da}, \rho^2_{da})\}$ *of the tree. Using the fact that both pointed models* $(\mathscr{O}^2_{da}, o^2_{da})$ *and* $(\mathscr{P}^2_{da}, \rho^2_{da})$ *satisfy the same propositional symbols and the second item from Proposition 4.4, we see that the root may have a syntax label* $\neg$, $\vee$, *or* $[i]$ *for some* $i \in \{a, d\}$*. If the syntax label is* $\neg$*, we add its successor* $\{(\mathscr{P}^2_{da}, \rho^2_{da})\} \circ \{(\mathscr{O}^2_{da}, o^2_{da})\}$ *to* B*. If the syntax label of the root is* $\vee$*, then it has at least one successor* $\{(\mathscr{O}^2_{da}, o^2_{da})\}\circ\{(\mathscr{P}^2_{da}, \rho^2_{da})\}$ *and we add it to* B*. Continuing in this way, we follow up the resulting nodes* $\{(\mathscr{P}^2_{da}, \rho^2_{da})\} \circ \{(\mathscr{O}^2_{da}, o^2_{da})\}$ *or* $\{(\mathscr{O}^2_{da}, o^2_{da})\} \circ \{(\mathscr{P}^2_{da}, \rho^2_{da})\}$*, adding them to* B*, until a node of this form that has a syntax label* $[i]$ *appears. We know that* i *must be* d*. If the node is* $\{(\mathscr{P}^2_{da}, \rho^2_{da})\} \circ \{(\mathscr{O}^2_{da}, o^2_{da})\}$*, then we add its successor* $\{(\mathscr{P}^2_{da}, \rho^1_{a})\}\circ\{(\mathscr{O}^2_{da}, o^1_{a})\}$ *to* B*. Similarly, if the node is* $\{(\mathscr{O}^2_{da}, o^2_{da})\}\circ\{(\mathscr{P}^2_{da}, \rho^2_{da})\}$*, then we add its successor* $\{(\mathscr{O}^2_{da}, o^1_{a})\}\circ\{(\mathscr{P}^2_{da}, \rho^1_{a})\}$ *to* B*. Again the possible syntax labels for a node of one of these forms is* $\neg$, $\vee$*, and* $[i]$*. We follow the nodes through a possible number of syntax labels of the form* $\neg$ *or* $\vee$ *until a node* η *that has a syntax label* $[i]$ *appears. We know that* i *must be* a *because* $(\mathscr{O}^2_{da}, o^1_{a})$ *and* $(\mathscr{P}^2_{da}, \rho^1_{a})$ *are bisimilar with respect to* d*. Therefore, if* $\eta \stackrel{\text{def}}{=} \{(\mathscr{O}^2_{da}, o^1_{a})\} \circ \{(\mathscr{P}^2_{da}, \rho^1_{a})\}$*, then we add its successor*$\{(\mathscr{O}^2_{da}, o^0)\} \circ \{(\mathscr{P}^2_{da}, \rho^0)\}$ *to* B*. On the other hand, if* $\eta \stackrel{\text{def}}{=} \{(\mathscr{P}^2_{da}, \rho^1_{a})\} \circ \{(\mathscr{O}^2_{da}, o^1_{a})\}$*, we add its successor* $\{(\mathscr{P}^2_{da}, \rho^0)\} \circ \{(\mathscr{O}^2_{da}, o^0)\}$ *to* B*. Note that none of the nodes* $\{(\mathscr{O}^2_{da}, o^0)\} \circ \{(\mathscr{P}^2_{da}, \rho^0)\}$ *or* $\{(\mathscr{P}^2_{da}, \rho^0)\} \circ \{(\mathscr{O}^2_{da}, o^0)\}$ *can have a syntax label* $[i]$ *or* $[\cup_{\{a,d\}}]$ *because they are not* R_a *or* R_d *related to any node. Therefore, we may follow nodes of these two forms through a number of* $\neg$ *and* $\vee$ *syntax labels, but we will eventually reach a node* $\{(\mathscr{O}^2_{da}, o^0)\} \circ \{(\mathscr{P}^2_{da}, \rho^0)\}$ *that has a syntax label* $\mathbf{b}$ *because* $(\mathscr{O}^2_{da}, o^0) \models \mathbf{b}$ *and* $(\mathscr{P}^2_{da}, \rho^0) \models \neg\mathbf{b}$*. Thus, for the branch* B *we constructed, it is true that* $I(B) = da$*.*

We proceed by induction on n to prove the stronger statement that the syntax tree of any formula μ such that

$$(\mathscr{O}^1_d, o^1_d) \models \mu \text{ and } (\mathscr{P}^1_d, \rho^1_d) \models \neg\mu$$

or

$$(\mathscr{P}^1_d, \rho^1_d) \models \mu \text{ and } (\mathscr{O}^1_d, o^1_d) \models \neg\mu$$

contains a branch B such that $I(B) = jw$.

Base case $n = 1$**.** We claim that the extended syntax tree of μ with root $\{(\mathscr{O}^1_d, o^1_d)\} \circ \{(\mathscr{P}^1_d, \rho^1_d)\}$ or $\{(\mathscr{P}^1_d, \rho^1_d)\} \circ \{(\mathscr{O}^1_d, o^1_d)\}$ (see Figure 4.4) contains a node η with syntax label $[d]$ where η has one of the following two forms.

$$\eta \stackrel{\text{def}}{=} \{(\mathscr{O}^1_d, o^1_d)\} \circ \{(\mathscr{P}^1_d, \rho^1_d)\} \text{ or } \eta \stackrel{\text{def}}{=} \{(\mathscr{P}^1_d, \rho^1_d)\} \circ \{(\mathscr{O}^1_d, o^1_d)\}.$$

Let us assume otherwise. It follows from the second item of Proposition 4.4, that the root $\{(\mathscr{O}^1_d, o^1_d)\} \circ \{(\mathscr{P}^1_d, \rho^1_d)\}$ or $\{(\mathscr{P}^1_d, \rho^1_d)\} \circ \{(\mathscr{O}^1_d, o^1_d)\}$ of the tree cannot have a syntax label $[\cup_{\{a,d\}}]$; moreover, the construction of the models $(\mathscr{O}_d^1, o_d^1)$ and $(\mathscr{P}_d^1, \rho_d^1)$ and Corollary 3.24 imply that, because $(\mathscr{O}_d^1, \rho_d^1)$ and $(\mathscr{P}_d^1, \rho_d^1)$ are bisimilar with respect to a, the root cannot have a syntax label $[a]$.

The above considerations and our assumption imply that the root can have a syntax label that is either $\vee$ or $\neg$. In either case, the relevant items from Definition 3.22 imply that at least one of the successor nodes is either $\{(\mathscr{O}_d^1, o_d^1)\} \circ \{(\mathscr{P}_d^1, \rho_d^1)\}$ or $\{(\mathscr{P}_d^1, \rho_d^1)\} \circ \{(\mathscr{O}_d^1, o_d^1)\}$. Again this node can have only a syntax label that is either $\vee$ or $\neg$, etc. This means that the extended syntax tree of μ with root $\{(\mathscr{O}_d^1, o_d^1)\} \circ \{(\mathscr{P}_d^1, \rho_d^1)\}$ or $\{(\mathscr{P}_d^1, \rho_d^1)\} \circ \{(\mathscr{O}_d^1, o_d^1)\}$ contains an infinite branch which is absurd because syntax trees and extended syntax trees are finite. Therefore, there must be a node η with a syntax label $[d]$ that has one of the forms $\{(\mathscr{O}_d^1, o_d^1)\} \circ \{(\mathscr{P}_d^1, \rho_d^1)\}$ or $\{(\mathscr{P}_d^1, \rho_d^1)\} \circ \{(\mathscr{O}_d^1, o_d^1)\}$. The successor of η is either $\eta_1 \stackrel{\text{def}}{=} \{(\mathscr{O}_d^1, o^0)\} \circ \{(\mathscr{P}_d^1, \rho^0)\}\rangle$ or $\eta_1 \stackrel{\text{def}}{=} \langle\{(\mathscr{P}_d^1, \rho^0)\} \circ \{(\mathscr{O}_d^1, o^0)\}$, respectively. Note that such a node cannot have a syntax label $[j]$ for any $j \in \{a, d\}$ nor can it have a syntax label $[\cup_{\{a,d\}}]$. It is obvious that η_1 can have a syntax label $\mathbf{b}$, $\vee$ or $\neg$. In either case we can find the desired branch B such that $I(B) = d$ by following the unique path leading from the root $\{(\mathscr{O}_d^1, o_d^1)\} \circ \{(\mathscr{P}_d^1, \rho_d^1)\}$ or $\{(\mathscr{P}_d^1, \rho_d^1)\} \circ \{(\mathscr{O}_d^1, o_d^1)\}$ of the extended syntax tree of μ to the node η and then randomly choosing a path leading from η to a leaf $\{(\mathscr{O}_d^1, o^0)\} \circ \{(\mathscr{P}_d^1, \rho^0)\}$ with syntax label $\mathbf{b}$.

Induction hypothesis. We assume that the syntax tree of any formula μ such that

$$(\mathscr{O}_w^n, o_w^n) \models \mu \text{ and } (\mathscr{P}_w^n, \rho_w^n) \models \neg\mu$$

or

$$(\mathscr{P}_w^n, \rho_w^n) \models \mu \text{ and } (\mathscr{O}_w^n, o_w^n) \models \neg\mu$$

contains a branch B for which $I(B) = w$.

Inductive step. Using the same reasoning as in the Base case, we see that the extended syntax tree of μ with root $\{(\mathscr{O}_{dw}^{n+1}, o_{dw}^{n+1})\} \circ \{(\mathscr{P}_{dw}^{n+1}, \rho_{dw}^{n+1})\}$ or $\{(\mathscr{P}_{dw}^{n+1}, \rho_{dw}^{n+1})\} \circ \{(\mathscr{O}_{dw}^{n+1}, o_{dw}^{n+1})\}$ contains a node η with syntax label $[d]$ where η has one of the following two forms.

$$\eta \stackrel{\text{def}}{=} \{(\mathscr{O}_{dw}^{n+1}, o_{dw}^{n+1})\} \circ \{(\mathscr{P}_{dw}^{n+1}, \rho_{dw}^{n+1})\} \text{ or } \eta \stackrel{\text{def}}{=} \{(\mathscr{P}_{dw}^{n+1}, \rho_{dw}^{n+1})\} \circ \{(\mathscr{O}_{dw}^{n+1}, o_{dw}^{n+1})\}.$$

According to Definition 3.22, the successor of η is
$\eta_1 \stackrel{\text{def}}{=} \{(\mathscr{O}_{dw}^{n+1}, o_w^n)\} \circ \{(\mathscr{P}_{dw}^{n+1}, \rho_w^n)\}$ or $\eta_1 \stackrel{\text{def}}{=} \{(\mathscr{P}_{dw}^{n+1}, \rho_w^n)\} \circ \{(\mathscr{O}_{dw}^{n+1}, o_w^n)\}$, respectively. Using Definition 3.22 and Proposition 3.23, we see that η_1 is a root of a sub-tree T that is an extended syntax tree of a formula ν such that either

$$(\mathscr{O}_{dw}^{n+1}, o_w^n) \models \nu \text{ and } (\mathscr{P}_{dw}^{n+1}, \rho_w^n) \models \neg\nu$$

or

$$(\mathscr{P}^{n+1}_{dw}, \rho^n_w) \models \nu \text{ and } (\mathscr{O}^{n+1}_{dw}, o^n_w) \models \neg\nu.$$

Since $(\mathscr{P}^{n+1}_{dw}, \rho^n_w)$ and $(\mathscr{P}^n_w, \rho^n_w)$ are bisimilar and the same applies to $(\mathscr{O}^{n+1}_{dw}, o^n_w)$ and $(\mathscr{O}^n_w, o^n_w)$ we see that

$$(\mathscr{O}^n_w, o^n_w) \models \nu \text{ and } (\mathscr{P}^n_w, \rho^n_w) \models \neg\nu$$

or

$$(\mathscr{P}^n_w, \rho^n_w) \models \nu \text{ and } (\mathscr{O}^n_w, o^n_w) \models \neg\nu.$$

Therefore, applying the induction hypothesis, we see that T contains a branch Br for which $I(Br) = w$. Hence, the desired branch B such that $I(B) = dw$ in the syntax tree of μ is constructed by starting at the root, going through the nodes η, η_1, and then following the branch Br.

□

Before finishing this chapter, we present one last result that follows almost immediately from the proof of Lemma 4.7. Namely, that there are sets of models, which, abusing notation, we will call again $\mathbb{O}$ and $\mathbb{P}$ such that the set of formulae $\Theta_{[\exists]}$ is exponentially more succinct than $\Phi_{[\cap]\mathrm{ML}}$ on $\mathbb{O} \cup \mathbb{P}$. Indeed, let, for $n = 1$, the set of pointed models $\mathbb{O}^1$ consists of the models on the left of the doted line in Figure 4.7 while the set $\mathbb{P}^1$ consists of the models on the right on the dotted line.

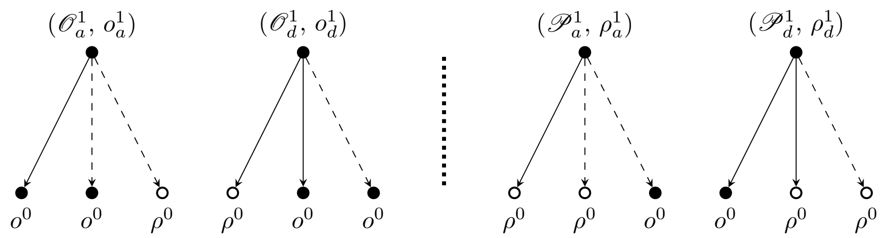


FIGURE 4.7: The pointed models $(\mathscr{O}^2_{da}, o^2_{da})$ and $(\mathscr{P}^2_{da}, \rho^2_{da})$.

Similarly, for $n > 1$ let the set $\mathbb{O}^n$ consists of the models on the left of the dotted line in Figure 4.8 and the set $\mathbb{P}^n$ consists of the pointed models on the right. The reader can easily verify that each of these new models is bisimilar to the respective models from Definition 4.3. Unlike the models from Definition 4.3 however, the new models are (trivially) bisimilar with respect to $a \cap d$, too. Therefore, applying Theorem 3.18, we see that for any formula $[a \cap d]\varphi \in \Phi_{[\cap]\mathrm{ML}}$, and any pair $(\mathscr{O}^n_w, o^n_w)$ and $(\mathscr{P}^n_w, \rho^n_w)$ of the newly defined models, we have

$$(\mathscr{O}^n_w, o^n_w) \models [\cap_{\{a,d\}}]\varphi \text{ iff } (\mathscr{P}^n_w, \rho^n_w) \models [\cap_{\{a,d\}}]\varphi. \quad (4.1)$$

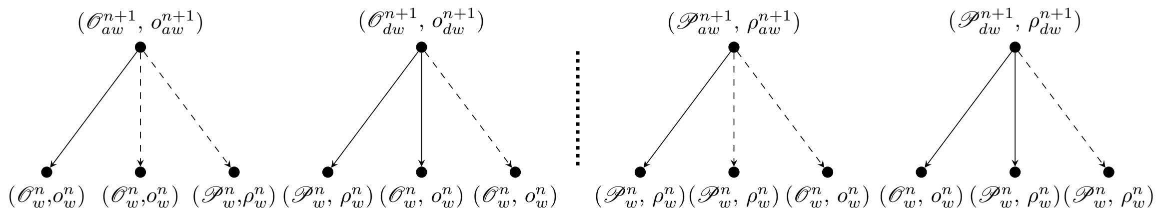


FIGURE 4.8: The pointed models $(\mathscr{O}^2_{da}, o^2_{da})$ and $(\mathscr{P}^2_{da}, \rho^2_{da})$.

Let

$$\mathbb{O} = \bigcup_{n\geq 1} \mathbb{O}^n \text{ and } \mathbb{P} = \bigcup_{n\geq 1} \mathbb{P}^n.$$

We can apply the proof of the second item of Lemma 4.7 almost verbatim with the only change being that every reference to the second item of Proposition 4.4 should be replaced with a reference to 4.1 above.

Chapter 5

Succinctness Results on S5-Models

The succinctness results in this chapter presuppose signatures that contain at least 4 relation indices and at least 4 propositional symbols. For the sake of concreteness, we fix one such signature S where $I = \{a, b, c, d\}$ and $P = \{\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}\}$. As in the case for unrestricted models, our results transfer in an obvious way to signatures with more relation indices and/or more propositional symbols.

We are going to prove the following theorem.

Theorem 5.1. *There are sets of* $\mathbf{S}5$*-models* $\mathbb{A}$*,* $\mathbb{C}$*,* $\mathbb{E}$*, and* $\mathbb{G}$ *and sets of formulae*

- $\Delta_{[\cup]} \subset \Phi_{[\cup]\mathrm{ML}}$,
- $\Theta_{[\cap]} \subset \Phi_{[\cap]\mathrm{ML}}$
- $\Sigma_{[\exists]} \subset \Phi_{[\exists]\mathrm{ML}}$,
- $\Omega_{[\varphi]} \subset \Phi_{[\varphi]\mathrm{ML}}$,

such that

1. $\Delta_{[\cup]}$ *is exponentially more succinct than* Φ_{ML} *on* $\mathbb{A}$*;*
2. $\Theta_{[\cap]}$ *is exponentially more succinct than* Φ_{ML} *on* $\mathbb{G}$*.*
3. $\Sigma_{[\exists]}$ *is exponentially more succinct than* Φ_{ML} *on* $\mathbb{C}$*;*
4. $\Omega_{[\varphi]}$ *is exponentially more succinct than* Φ_{ML} *on* $\mathbb{E}$*;*

The proof of this theorem follows the strategy employed in the previous chapter. We begin by defining the sets of formulae and the sets of models, and then proceed to proving the four items above. The main ingredient of our argument is again the diverging pairs technique. This time, however, our reasoning will be more complicated and that is why we will give each successive step of the proof in a separate section.

5.1 The sets of formulae

Definition 5.2 (Formulae)**.**

- The sets of formulae $\Delta_{[\cup]} \subset \Phi_{[\cup]\mathrm{ML}}$ and $\Sigma_{[\exists]} \subset \Phi_{[\exists]\mathrm{ML}}$ are defined recursively as shown in Table 5.1.

	$\Delta_{[\cup]}$		$\Sigma_{[\exists]}$
δ_1	$\neg[\cup_{\{a,b\}}]\neg\mathbf{c}$	σ_1	$[\exists_{\{a,b\}}]\mathbf{c}$
$\vdots$		$\vdots$	
δ_n	$\neg[\cup_{\{a,b\}}][c][d]\neg\delta_{n-1}$	σ_n	$[\exists_{\{a,b\}}](\mathbf{c} \wedge [c][d]\sigma_{n-1})$
$\vdots$		$\vdots$	

Table 5.1: The sets of formulae $\Delta_{[\cup]}$ and $\Sigma_{[\exists]}$.

- The set of formulae $\Omega_{[\varphi]} \subset \Phi_{[\varphi]\mathrm{ML}}$, and $\Theta_{[\cap]} \subset \Phi_{[\cap]\mathrm{ML}}$ are defined in Table 5.2 below.

	$\Theta_{[\cap]}$		$\Omega_{[\varphi]}$
θ_1	$\langle\cap_{\{a,b\}}\rangle\mathbf{c}$	ω_1	$\langle c\rangle(\mathbf{c} \wedge \langle d\rangle(\mathbf{d} \wedge (\langle a\rangle\mathbf{a} \vee \langle b\rangle\mathbf{b})))$
$\vdots$		$\vdots$	
θ_n	$\langle\cap_{\{a,b\}}\rangle\langle c\rangle\langle d\rangle\theta_{n-1}$	ω_n	$\langle\omega_{n-1}\rangle\omega_1$
$\vdots$		$\vdots$	

Table 5.2: The sets of formulae $\Theta_{[\cap]}$ and $\Omega_{[\varphi]}$.

It is easy to see that for any of the sets of formulae specified in Tables 5.1 and 5.2 there is a linear function that bounds from above the length of the respective formulae in terms of their indices. This, of course, applies to the actual formulae in $\Theta_{[\cap]}$ and $\Omega_{[\varphi]}$ that are obtained by replacing the defined symbols $\langle a\rangle$, $\langle c\rangle$, $\langle d\rangle$, $\langle\varphi\rangle$, and $\langle\cap_{\{a,b\}}\rangle$ with $\neg[a]\neg$, $\neg[c]\neg$, $\neg[d]\neg$, $\neg[\varphi]\neg$, and $\neg[\cap_{\{a,b\}}]\neg$, respectively.

5.2 The sets of pointed models

Next, we define the sets of $\mathbf{S}_5$-models $\mathbb{A}$, $\mathbb{C}$, $\mathbb{E}$, and $\mathbb{G}$. In our figures, we will not include reflexive edges, and, since all relations are symmetric, we will not use arrows when denoting edges. Intuitively, our models have the shape of a ladder. The nodes that form the left support of the ladder are taken from the set $\{\grave{x}_i, \grave{y}_i, \grave{z}_i\}$, while points that form the right support are taken from $\{\acute{x}_i, \acute{y}_i, \acute{z}_i\}$. A point $\grave{x}_i$ is horizontally connected, and thus forms a rung, with $\acute{x}_i$. Likewise for $\grave{y}_i$ and $\acute{y}_i$, and for $\grave{z}_i$ and $\acute{z}_i$. The class of models $\mathbb{C}$ consists of ladders with a third, middle support: the points on this support are denoted $\overline{x}_i$, $\overline{y}_i$ and $\overline{z}_i$. On such models, we will still say that $\grave{x}_i$,$\acute{x}_i$ and $\overline{x}_i$ are horizontally connected, so are $\grave{y}_i$, $\acute{y}_i$ and $\overline{y}_i$, and another horizontal rung is formed by $\grave{z}_i$,$\acute{z}_i$ and $\overline{z}_i$. The general idea governing the definition of the relation steps is that vertical relation steps are different from the horizontal ones. The nodes $\grave{x}_n$ and $\acute{x}_n$ are i-connected exactly for those $i \in I$ for which $\grave{x}_n$ is not i-connected to any $\grave{z}$ or $\grave{y}$ node. Similarly, $\grave{y}_n$ and $\acute{y}_n$ are i-connected exactly for those $i \in I$ for which $\grave{y}_n$ is not i-connected to any $\grave{x}$ or $\grave{z}$ node. Finally, $\grave{z}_n$ and $\acute{z}_n$ are i-connected exactly for those $i \in I$ for which $\grave{z}_n$ is not i-connected to any $\grave{x}$ or $\grave{y}$ node.

The following notation will be employed in what follows. We let p vary over the set $\{x, y, z\}$. Hence, if for example, we use $\grave{p}_i$ to denote $\grave{x}_i$, then, by $\acute{p}_i$, we mean $\acute{x}_i$ and $\overline{p}_i$ is $\overline{x}_i$. If w is a word of length n over the alphabet $\{a, b, c, d\}$, we write w^i for the i-th symbol of w. If $j \in \{a, b\}$, then

$$\overline{j} = \begin{cases} b, & \text{if } j = a; \\ a, & \text{if } j = b. \end{cases}$$

We begin by defining carefully the set of models $\mathbb{A}$. Along the way, we establish some conventions that are used later in the definitions of the models $\mathbb{C}$, $\mathbb{E}$ and $\mathbb{G}$.

Definition 5.3 (The set of models $\mathbb{A}$)**.** For every natural number $n \geq 1$, the sets $\mathbb{A}^n$ and $\mathbb{B}^n$, containing 2^n different pointed models each, are defined recursively as follows.

1. The set $\mathbb{A}^1$ consists of the two pointed models $(\mathscr{A}_a^1, \grave{x}_1)$ and $(\mathscr{A}_b^1, \grave{x}_1)$ shown on the left of the dotted line in Figure 5.1. The set $\mathbb{B}^1$ contains the pointed models $(\mathscr{A}_a^1, \acute{x}_1)$ and $(\mathscr{A}_b^1, \acute{x}_1)$ that are shown on the left of the dotted line. The nodes $\grave{y}_1$

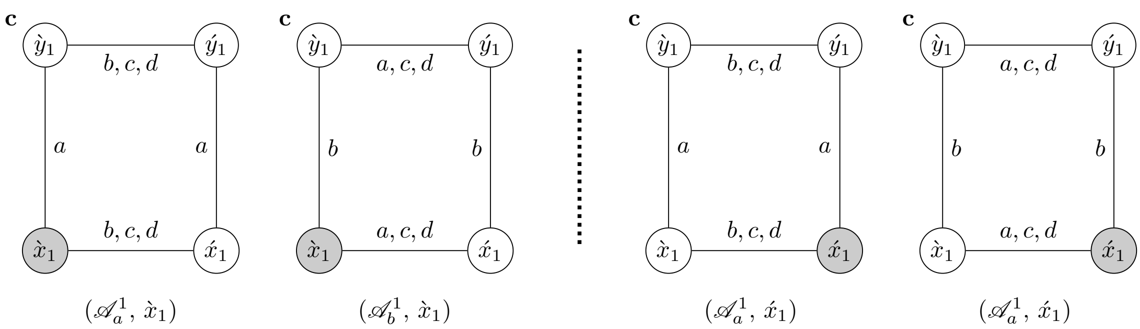


FIGURE 5.1: The sets of pointed models $\mathbb{A}^1$ and $\mathbb{B}^1$.

satisfy the proposition **c** whereas the other nodes do not satisfy any propositions. The subscripts in the names of the Kripke models encode the shortest path from the node $\grave{x}_1$ to the node $\grave{y}_1$. For example, the node $\grave{y}_1$ can be reached from the node $\grave{x}_1$ in the model $\mathscr{A}_a^1$ by making one vertical step along the relation R_a. Note that in the same model the point $\grave{y}_1$ cannot be reached from $\acute{x}_1$ by making only one step along the relation R_a; furthermore, the nodes $\grave{x}_1$ and $\acute{x}_1$ are connected via the relations R_b, R_c, and R_d and the same applies to the nodes $\grave{y}_1$ and $\acute{y}_1$.

$n+1$. The set $\mathbb{A}^{n+1}$ consists of all the pointed models models built from the models in $\mathbb{A}^n$ as shown in the Figure 5.2. We construct a pointed model $(\mathscr{A}_{acdw}^{n+1}, \grave{x}_{n+1})$ by

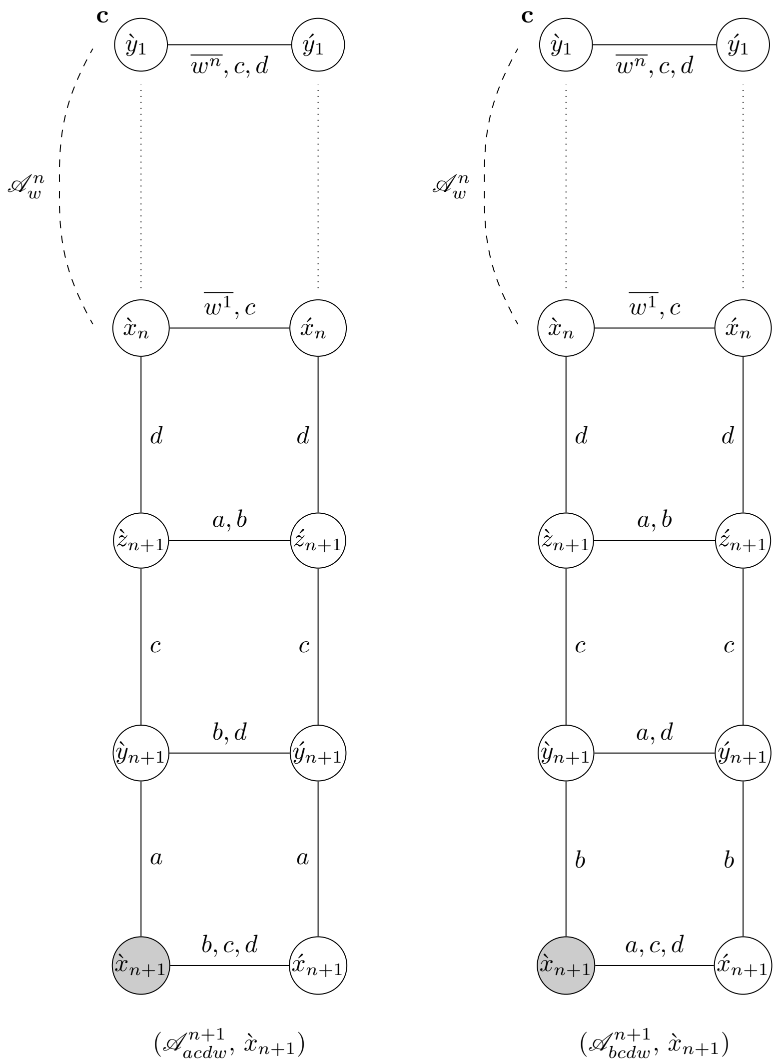
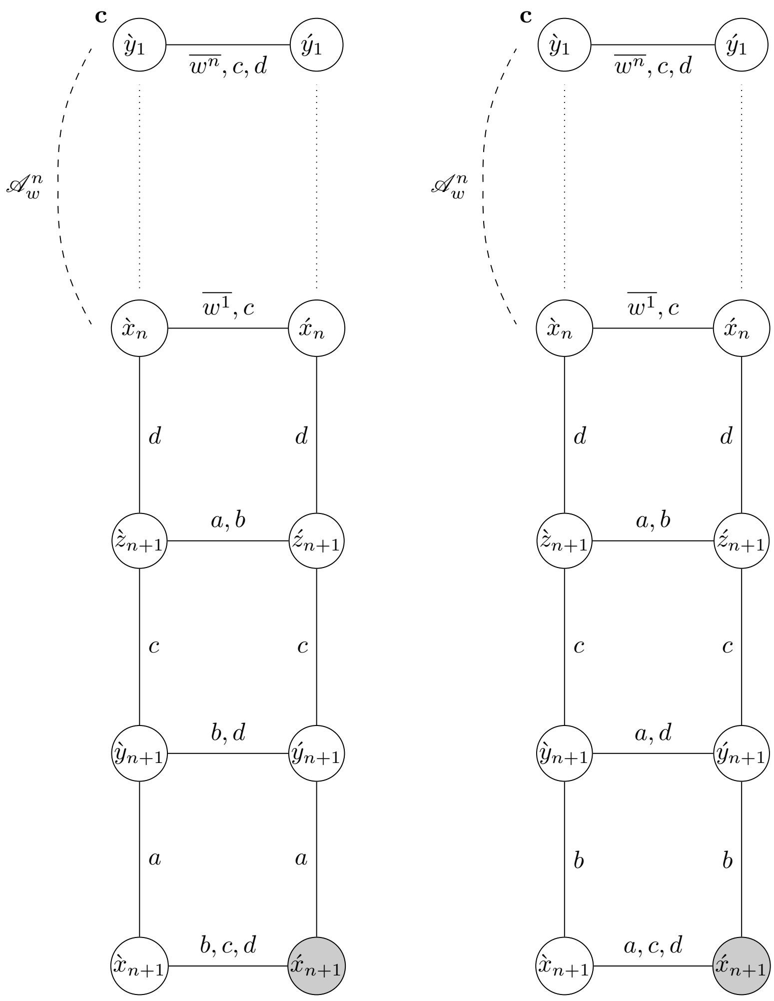


FIGURE 5.2: The set of pointed models $\mathbb{A}^{n+1}$.

taking a model $\mathscr{A}^n_w \in \mathbb{A}^n$ and erasing the d-step connecting the nodes $\grave{x}_n$ and $\acute{x}_n$. Although, strictly speaking, the model obtained in this way is different from the model $\mathscr{A}^n_w$, we will denote both models by $\mathscr{A}^n_w$ in order to show the simple idea on which our construction is based. After that, we add the new points $\grave{x}_{n+1}$, $\grave{y}_{n+1}$, $\grave{z}_{n+1}$, $\acute{x}_{n+1}$, $\acute{y}_{n+1}$, $\acute{z}_{n+1}$, and define the relation steps as shown. The pointed models $(\mathscr{A}^{n+1}_{bcdw}, \grave{x}_{n+1})$ are built in a similar way. The set of pointed models $\mathbb{B}^{n+1}$ is constructed analogously as shown in Figure 5.3. We would like to remind the reader that the relation step $\overline{w^1}$ connecting the nodes $\grave{x}_n$ and $\acute{x}_n$ is obtained by taking the first symbol of the word w in the subscript of $\mathscr{A}^n_w$ and if this symbol is a, then $\overline{w^1} = b$ and vice versa.

Let

$$\mathbb{A} = \bigcup_{n \geq 1} (\mathbb{A}^n \cup \mathbb{B}^n).$$



$(\mathscr{A}^{n+1}_{acdw}, \acute{x}_{n+1})$ $(\mathscr{A}^{n+1}_{bcdw}, \acute{x}_{n+1})$

FIGURE 5.3: The set of pointed models $\mathbb{B}^n$.

Intuitively, the subscript w in the name of the Kripke model $\mathscr{A}^n_w$ encodes the shortest sequence of relation steps leading from $\grave{x}_n$ to the point $\grave{y}_1$ that satisfies the proposition **c**. Since there are 2^n different words $w \stackrel{\text{def}}{=} j_1 cdj_2 cd \dots cdj_n$, where $j_i \in \{a, b\}$ and for every such w, there is a corresponding pointed model $(\mathbb{A}^n_w, \grave{x}_n) \in \mathbb{A}^n$ and $(\mathbb{A}^n_w, \acute{x}_n) \in \mathbb{B}^n$, we see that both $\mathbb{A}^n$ and $\mathbb{B}^n$ contain 2^n different pointed models.

The most important property of the models $\mathscr{A}^n_w$ is that $(\mathscr{A}^n_w, \grave{p}_i)$ and $(\mathscr{A}^n_w, \acute{p}_i)$ are bisimilar with respect to all horizontal relation steps that connect them, i.e., for any j such that $\grave{p}_i R_j \acute{p}_i$. We hope that the next example clarifies Definition 5.3.

Example 5.1. *The pointed models* $(\mathscr{A}^2_{acda}, \grave{x}_2)$ *and* $(\mathscr{A}^2_{acda}, \acute{x}_2)$ *are shown in Figure 5.4. Note how the model* $\mathscr{A}^1_a$ *that is used in the construction of* $\mathscr{A}^2_{acda}$ *differs from the model* $\mathscr{A}^1_a$ *from Figure 5.1; namely, for the model* $\mathscr{A}^1_a$ *from Figure 5.1, we have* $\grave{x}_1 R_d \acute{x}_1$ *whereas this is not true for the nodes* $\grave{x}_1$ *and* $\acute{x}_1$ *from the model* $\mathscr{A}^2_{acda}$ *in Figure 5.4.*

It is easy to see that

- $(\mathscr{A}^2_{acda}, \grave{x}_2)$ *and* $(\mathscr{A}^2_{acda}, \acute{x}_2)$ *are bisimilar with respect to b,c, and d;*

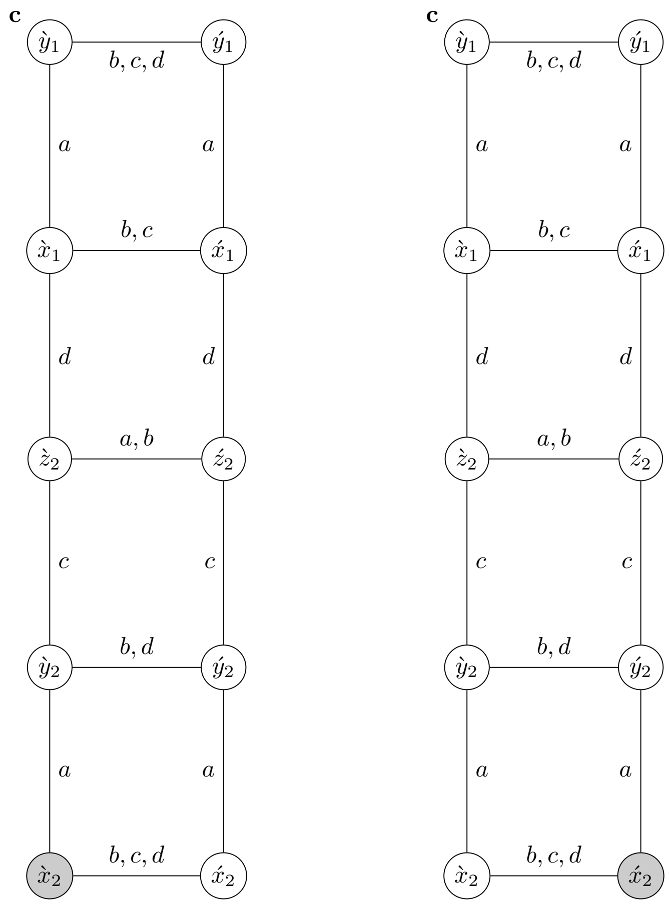


FIGURE 5.4: The models $(\mathscr{A}^2_{acda}, \grave{x}_2)$, (left) and $(\mathscr{A}^2_{acda}, \acute{x}_2)$ (right).

- $(\mathscr{A}^2_{acda}, \grave{y}_2)$ *and* $(\mathscr{A}^2_{acda}, \acute{y}_2)$ *are bisimilar with respect to* b *and* d*;*
- $(\mathscr{A}^2_{acda}, \grave{z}_2)$ *and* $(\mathscr{A}^2_{acda}, \acute{z}_2)$ *are bisimilar with respect to* a *and* b*;*
- $(\mathscr{A}^2_{acda}, \grave{x}_1)$ *and* $(\mathscr{A}^2_{acda}, \acute{x}_1)$ *are bisimilar with respect to* b *and* c*;*
- $(\mathscr{A}^2_{acda}, \grave{y}_1)$ *and* $(\mathscr{A}^2_{acda}, \acute{y}_1)$ *are bisimilar with respect to* b*,* c*, and* d*.*

We have the easy proposition below.

Proposition 5.4. *The following items are true.*

- *For all* $n \geq 1$ *and all formulae* $\delta_n \in \Delta_{[\cup]}$*,*

$$\mathbb{A}^n \models \delta_n \text{ and } \mathbb{B}^n \models \neg\delta_n.$$

- *For any pair of pointed models* $(\mathscr{A}^n_w, \grave{p}_i)$ *and* $(\mathscr{A}^n_w, \acute{p}_i)$*, where* $1 \leq i \leq n$*, and any formula* $[j]\varphi \in \Phi_{\mathrm{ML}}$*, if* $\grave{p}_i$ *and* $\acute{p}_i$ *are* j*-connected, then*

$$(\mathscr{A}^n_w, \grave{p}_i) \models [j]\varphi \text{ iff } (\mathscr{A}^n_w, \acute{p}_i) \models [j]\varphi.$$

Proof. We outline the intuition on which the proof of the first item is based. It is easy to see that $\delta_1 \stackrel{\text{def}}{=} \neg[\cup_{\{a,b\}}]\neg\mathbf{c}$ is equivalent to $\langle a\rangle\mathbf{c} \vee \langle b\rangle\mathbf{c}$; similarly, for $n > 1$, $\delta_n \stackrel{\text{def}}{=} \neg[\cup_{\{a,b\}}][c][d]\neg\delta_{n-1}$ is equivalent to $\langle a\rangle\langle c\rangle\langle d\rangle\delta_{n-1} \vee \langle b\rangle\langle c\rangle\langle d\rangle\delta_{n-1}$. Therefore, δ_n is equivalent to a formula $\kappa \in \Phi_{\mathrm{ML}}$ of modal depth $3n-2$ in which the modal operators are nested in such a way that κ "describes" all the different paths $w \stackrel{\text{def}}{=} j_1 cdj_2 cd \ldots cdj_n$, where $j_i \in \{a, b\}$, of length $3n-2$ that lead from the current node to a point satisfying the proposition $\mathbf{c}$. This means that for any pointed model $(\mathscr{A}^n_w, \grave{x}_n) \in \mathbb{A}^n$, it is true that $(\mathscr{A}^n_w, \grave{x}_n) \models \delta_n$ and, thus, $\mathbb{A}^n \models \delta_n$. On the other hand, the construction of the models $\mathscr{A}^n_w$ is such that the shortest path leading from $\acute{x}_n$ to $\grave{y}_1$ is of length $3n-1$. Therefore, for any $(\mathscr{A}^n_w, \acute{x}) \in \mathbb{B}^n$, it is true that $(\mathscr{A}^n_w, \acute{x}) \models \neg\delta_n$. Thus, $\mathbb{A}^n \models \delta_n$ and $\mathbb{B}^n \models \neg\delta_n$. The second item follows immediately from Theorem 3.18 and the fact that

- $(\mathscr{A}^n_w, \grave{x}_n)$ and $(\mathscr{A}^n_w, \acute{x}_n)$ are bisimilar with respect to any j such that $\grave{x}_n R_j \acute{x}_n$;
- $(\mathscr{A}^n_w, \grave{y}_n)$ and $(\mathscr{A}^n_w, \acute{y}_n)$ are bisimilar with respect to any j such that $\grave{y}_n R_j \acute{y}_n$;
- $(\mathscr{A}^n_w, \grave{z}_n)$ and $(\mathscr{A}^n_w, \acute{z}_n)$ are bisimilar with respect to any j such that $\grave{z}_n R_j \acute{z}_n$;
- $(\mathscr{A}^n_w, \grave{x}_{n-1})$ and $(\mathscr{A}^n_w, \acute{x}_{n-1})$ are bisimilar with respect to all j for which $\grave{x}_{n-1} R_j \acute{x}_{n-1}$;

$\vdots$

- $(\mathscr{A}^n_w, \grave{y}_1)$ and $(\mathscr{A}^n_w, \acute{y}_1)$ are bisimilar with respect to any j such that $\grave{y}_1 R_j \acute{y}_1$.

□

Using the conventions established above, we proceed by constructing a pair of sets of pointed models $\mathbb{C}^n$ and $\mathbb{D}^n$ for each $n \geq 1$.

Definition 5.5 (The set of models $\mathbb{C}$)**.** For every $n \geq 1$, the sets $\mathbb{C}^n$ and $\mathbb{D}^n$, each containing 2^n different pointed models, are constructed recursively as follows.

1. The set $\mathbb{C}^1$ consists of the pointed models $(\mathscr{C}^1_a, \grave{x}_1)$ and $(\mathscr{C}^1_b, \grave{x}_1)$ shown on the left of the dotted line in Figure 5.5. Note that the nodes $\grave{x}_1$, $\grave{y}_1$, and $\acute{x}_1$ satisfy the

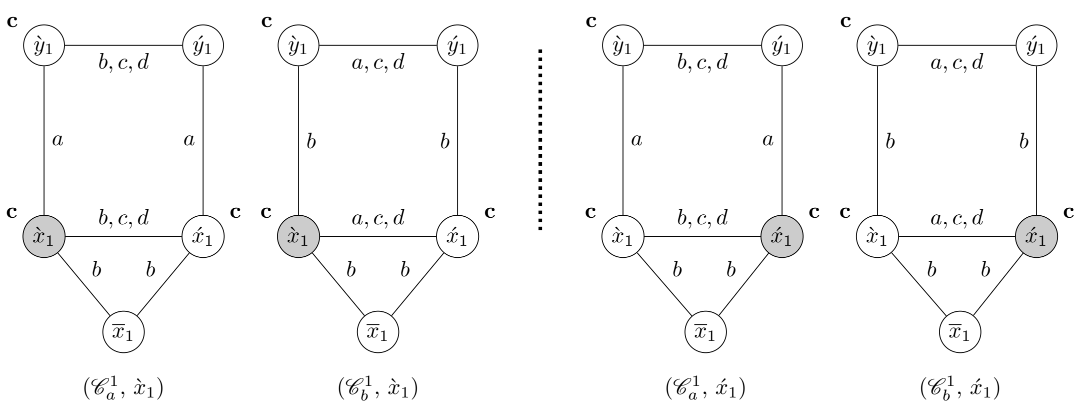


FIGURE 5.5: The pointed models $\mathbb{C}^1$ and $\mathbb{D}^1$.

proposition **c** whereas $\overline{x}_1$ and $\acute{y}_1$ do not. The set $\mathbb{D}^1$ consists of the pointed models $(\mathscr{C}_a^1, \acute{x}_1)$ and $(\mathscr{C}_b^1, \acute{x}_1)$ shown on the right of the dotted line.

n+1. The set $\mathbb{C}^{n+1}$ consists of all pointed models built as shown in Figure 5.6. Intuitively,

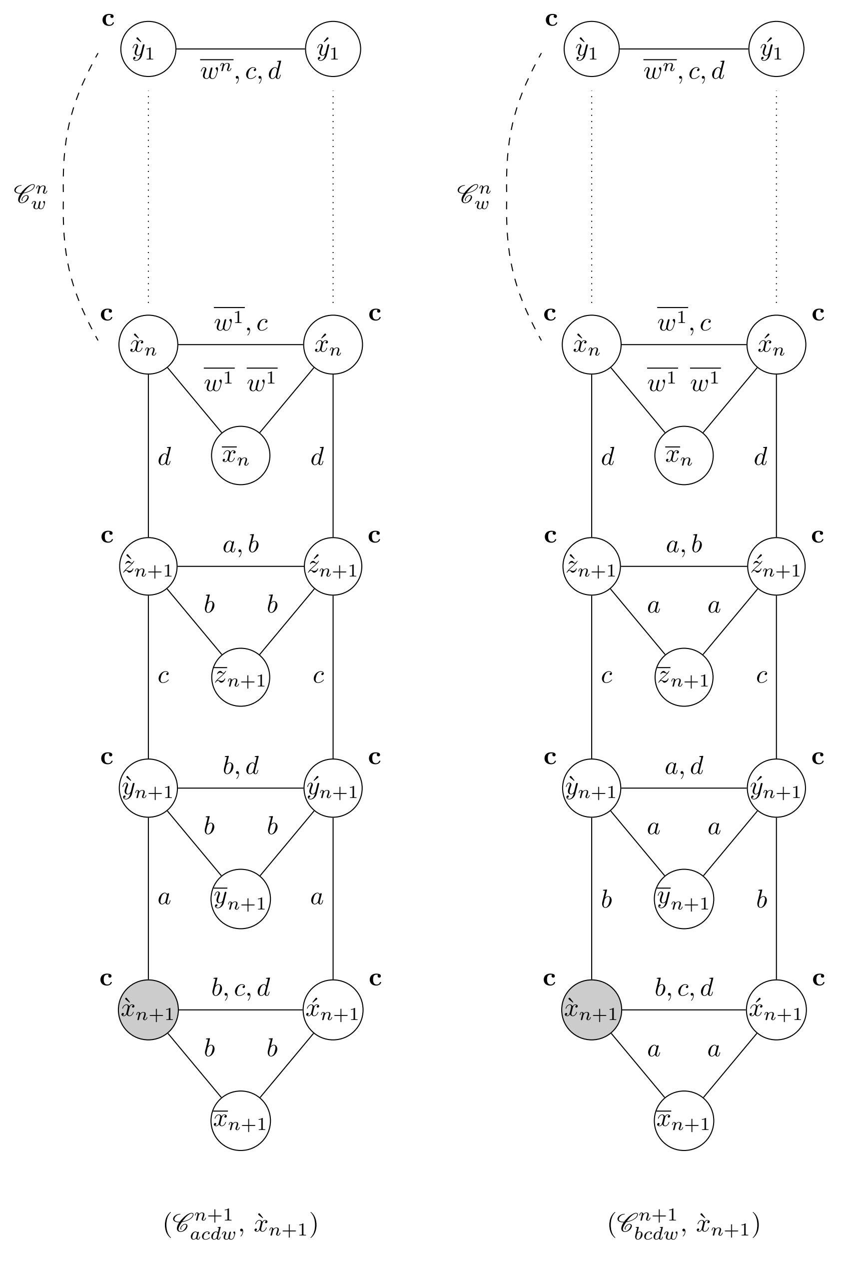


FIGURE 5.6: The set of pointed models $\mathbb{C}^{n+1}$.

we take a model $\mathscr{C}_w^n$ and erase the relation R_d connecting the points $\grave{x}_n$ and $\acute{x}_n$. Then we add 9 new points with subscript $n+1$ and define the relations as shown. The set $\mathbb{D}^{n+1}$ contains all models constructed as in Figure 5.8.

Note that all points of the form $\grave{p}_i$ in the model $\mathscr{C}_w^n$ satisfy the proposition **c**; no point of the form $\overline{p}_i$ satisfies **c** and all the points of the form $\acute{p}_i$ with the exception of $\acute{y}_1$ satisfy **c**.

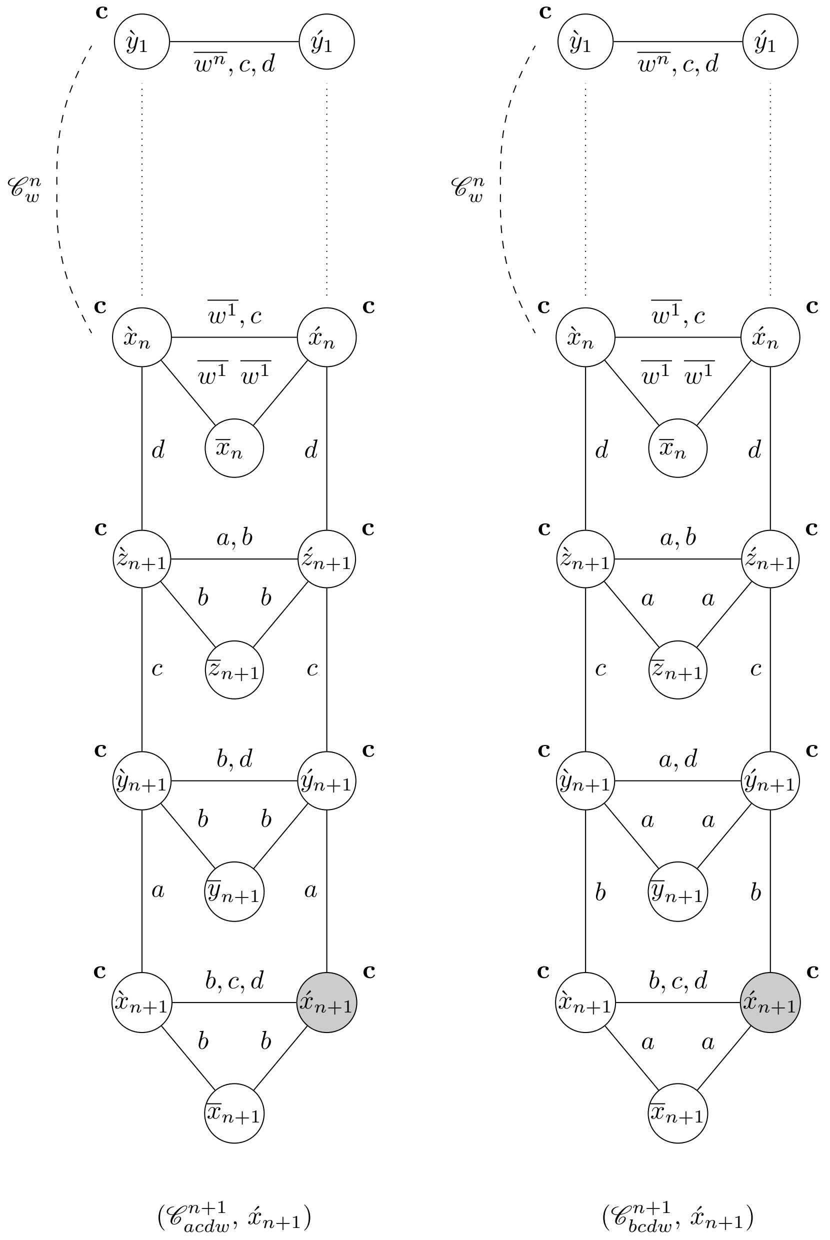


FIGURE 5.7: The set of pointed models $\mathbb{D}^{n+1}$.

The set $\mathbb{C}$ is defined as the union of all $\mathbb{C}^n$ and $\mathbb{D}^n$, i.e.,

$$\mathbb{C} = \bigcup_{n \geq 1} (\mathbb{C}^n \cup \mathbb{D}^n).$$

Again, the guiding intuition is that the word w used as a subscript in $\mathscr{C}^n_w$ encodes the shortest path of relations steps that leads from $\grave{x}_n$ to $\grave{y}_1$. Since w has the form $j_1 cdj_2 cdj_3 \ldots cdj_n$, where $j_i \in \{a, b\}$ and there are 2^n different such words, we see that both $\mathbb{C}^n$ and $\mathbb{D}^n$ contain 2^n different pointed models corresponding to the different words w. As in the case of the models in $\mathbb{A}^n$ and $\mathbb{B}^n$, it is true is that $(\mathscr{C}^n_w, \grave{p}_i)$ and $(\mathscr{C}^n_w, \acute{p}_i)$ are bisimilar with respect to all horizontal relation steps that connect them.

Example 5.2. *Figure 5.8 shows the pointed models* $(\mathscr{C}^2_{acda}, \grave{x}_2) \in \mathbb{C}^2$ *and* $(\mathscr{C}^2_{acda}, \acute{x}_2) \in \mathbb{D}^2$*. We see that*

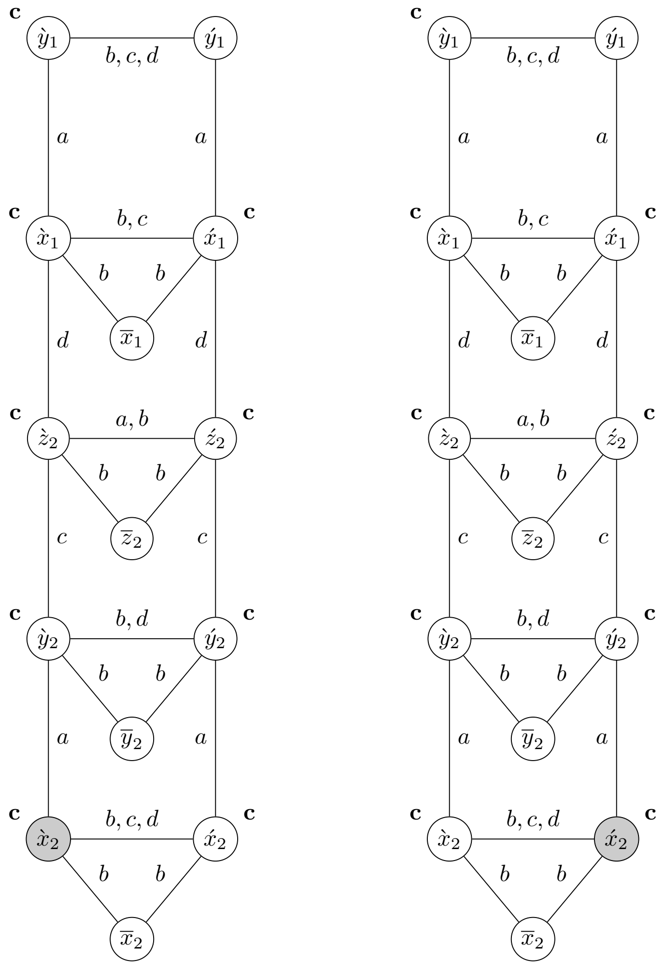


FIGURE 5.8: The pointed models $(\mathscr{C}^2_{acda}, \grave{x}_2)$ (left) and $(\mathscr{C}^2_{acda}, \acute{x}_2)$ (right).

- $(\mathscr{C}^2_{acda}, \grave{x}_2)$ *and* $(\mathscr{C}^2_{acda}, \acute{x}_2)$ *are bisimilar with respect to b,c, and d;*
- $(\mathscr{C}^2_{acda}, \grave{y}_2)$ *and* $(\mathscr{C}^2_{acda}, \acute{y}_2)$ *are bisimilar with respect to b and d;*
- $(\mathscr{C}^2_{acda}, \grave{z}_2)$ *and* $(\mathscr{C}^2_{acda}, \acute{z}_2)$ *are bisimilar with respect to a and b;*
- $(\mathscr{C}^2_{acda}, \grave{x}_1)$ *and* $(\mathscr{C}^2_{acda}, \acute{x}_1)$ *are bisimilar with respect to b and c;*
- $(\mathscr{A}^2_{acda}, \grave{y}_1)$ *and* $(\mathscr{A}^2_{acda}, \acute{y}_1)$ *are bisimilar with respect to b, c, and d.*

We have an analogue of Proposition 5.4.

Proposition 5.6. *The following are true.*

- *For all* $n \geq 1$ *and all formulae* $\sigma_n \in \Sigma_{[\exists]}$,

$$\mathbb{C}^n \models \sigma_n \text{ and } \mathbb{D}^n \models \neg\sigma_n$$

- *For all pairs of pointed models* $(\mathscr{C}^n_w, \grave{p}_i)$ *and* $(\mathscr{C}^n_w, \acute{p}_n)$*, where* $1 \leq i \leq n$*, and any formula* $[j]\varphi \in \Phi_{\mathrm{ML}}$*, if* $\grave{p}_i$ *and* $\acute{p}_i$ *are* j*-connected, then*

$$(\mathscr{C}^n_w, \grave{p}_i) \models [j]\varphi \textit{ iff } (\mathscr{C}^n_w, \acute{p}_i) \models [j]\varphi.$$

Proof. We deal only with the first item. The proof of the second is the same as the proof of the second item from Proposition 5.4.

The proof of the first item is as follows. It is easy to see that $\neg\sigma_{n+1}$ is equivalent to $\langle a\rangle(\neg\mathbf{c} \vee \langle c\rangle\langle d\rangle\neg\sigma_n) \wedge \langle b\rangle(\neg\mathbf{c} \vee \langle c\rangle\langle d\rangle\neg\sigma_n)$. Since all relations in the models $\mathscr{C}^w_n$ are reflexive and for every middle point $m \in \{\overline{x}_1\} \cup \{\overline{x}_i, \overline{y}_i, \overline{z}_i \mid 2 \leq i \leq n\}$, it is true that $(\mathscr{C}^n_w, m) \models \neg\mathbf{c}$, it follows immediately that for all $j \geq 1$, we have $(\mathscr{C}^n_w, m) \models \neg\sigma_j$.

We show next that

A: for every $1 \leq j \leq n$, it is true that

> **A1:** if $j < i \leq n$ and $p_i \in \{\grave{x}_i, \acute{x}_i, \grave{y}_i, \acute{y}_i, \grave{z}_i, \acute{z}_i\}$, then $(\mathscr{C}^n_w, p_i) \models \sigma_j$, and
>
> **A2:** if $j = i$, then $(\mathscr{C}^n_w, \grave{x}_i) \models \sigma_j$ whereas $(\mathscr{C}^n_w, \acute{x}_i) \models \neg\sigma_j$.

The proof of **A** is by induction on j.

Base case. Let $j = 1$. It is obvious that σ_1 is equivalent to $[a]\mathbf{c} \vee [b]\mathbf{c}$. Let us suppose that $1 < i$. It follows from the construction of the models $\mathscr{C}^n_w$ that all points $\grave{x}_i, \acute{x}_i, \grave{y}_i, \acute{y}_i, \grave{z}_i, \acute{z}_i$ satisfy the proposition $\mathbf{c}$; moreover, there is a relation index $l \in \{a, b\}$ such that $\grave{x}_i$ and $\grave{y}_i$ are l-connected and the same is true about $\acute{x}_i$ and $\acute{y}_i$. Furthermore, there is no point p such that p is an l-successor of one of the points $\grave{x}_i$, $\acute{x}_i$, $\grave{y}_i$, $\acute{y}_i$ and $(\mathscr{C}^n_w, p) \models \neg\mathbf{c}$. Hence

- $(\mathscr{C}^n_w, \grave{x}_i) \models \sigma_1$;
- $(\mathscr{C}^n_w, \acute{x}_i) \models \sigma_1$;
- $(\mathscr{C}^n_w, \grave{y}_i) \models \sigma_1$;
- $(\mathscr{C}^n_w, \acute{y}_i) \models \sigma_1$.

Let us consider now the points $\grave{z}_i$ and $\acute{z}_i$. Again, the construction of the model $\mathscr{C}^n_w$ is such that there is an $l \in \{a, b\}$, such that $\grave{z}_i$ and $\acute{z}_i$ are l-connected but there is no point p that is an l-successor of one of these points and at the same time $(\mathscr{C}^n_w, p) \models \neg\mathbf{c}$. Therefore, $(\mathscr{C}^n_w \grave{z}_i) \models \sigma_1$ and $(\mathscr{C}^n_w \acute{z}_i) \models \sigma_1$. To complete the base case, we have to prove A2 which says that $(\mathscr{C}^n_w, \grave{x}_1) \models \sigma_1$ and $(\mathscr{C}^n_w, \acute{x}_1) \models \neg\sigma_1$. The construction of the models is such that $\grave{x}_1$ and $\grave{y}_1$ satisfy the proposition $\mathbf{c}$ whereas $\overline{x}_1$ and $\acute{y}_1$ do not; moreover, $\acute{x}_1 R_l \acute{y}_1$ and $\acute{x}_1 R_k \overline{x}_1$ for some $l \neq k$ and $l, k \in \{a, b\}$. Hence $(\mathscr{C}^n_w, \acute{x}_1) \models \langle a\rangle\neg\mathbf{c} \wedge \langle b\rangle\neg\mathbf{c}$, i.e., $(\mathscr{C}^n_w, \acute{x}_1) \models \neg\sigma_1$. At the same time $\grave{x}_1 R_l \grave{y}_1$ and there is no point p such that $\grave{x}_1 R_l p$ and $(\mathscr{C}^n_w, p) \models \neg\mathbf{c}$. Therefore $(\mathscr{C}^n_w, \grave{x}_1) \models [a]\mathbf{c} \vee [b]\mathbf{c}$, i.e., $(\mathscr{C}^n_w, \grave{x}_1) \models \sigma_1$.

Induction step. Let us suppose that $j + 1 < i \leq n$ and let us assume that **A** is true for j. It is obvious that σ_{j+1} is equivalent to the formula $[a](\mathbf{c} \wedge [c][d]\sigma_j) \vee [b](\mathbf{c} \wedge [c][d]\sigma_j)$. We first prove that **A1** is true for the points $\grave{x}_n$ and $\acute{x}_n$. All the other cases are analogous. According to the induction hypothesis, it is true that

- $(\mathscr{C}^n_w, \grave{x}_n) \models \sigma_j$ and $(\mathscr{C}^n_w, \acute{x}_n) \models \sigma_j$;
- $(\mathscr{C}^n_w, \grave{y}_n) \models \sigma_j$ and $(\mathscr{C}^n_w, \acute{y}_n) \models \sigma_j$;
- $(\mathscr{C}^n_w, \grave{z}_n) \models \sigma_j$ and $(\mathscr{C}^n_w, \acute{z}_n) \models \sigma_j$;
- $(\mathscr{C}^n_w, \grave{x}_{n-1}) \models \sigma_j$ and $(\mathscr{C}^n_w, \acute{x}_{n-1}) \models \sigma_j$.

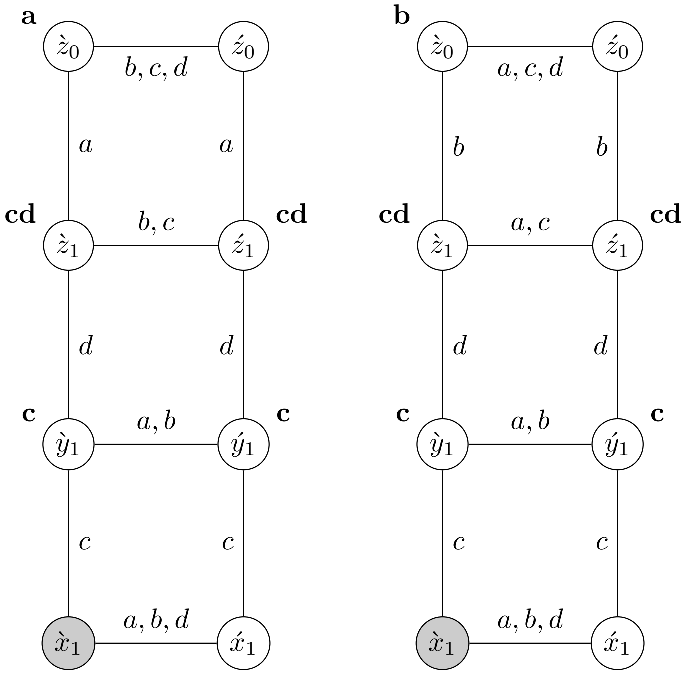
Since $\grave{x}_n$ and $\acute{x}_n$ are both c and d-connected and there is no other point p that is either a d or a c-successor of $\grave{x}_n$ or $\acute{x}_n$ and at the same time $(\mathscr{C}^n_w, p) \models \neg\sigma_j$, we see that $(\mathscr{C}^n_w, \grave{x}_n) \models \mathbf{c} \wedge [c][d]\sigma_j$ and $(\mathscr{C}^n_w, \acute{x}_n) \models \mathbf{c} \wedge [c][d]\sigma_j$. On the other hand, the induction hypothesis and the fact that $\grave{y}_n R_d \acute{y}_n$ and there is no other point p that is a d-successor of either $\grave{y}_n$ or $\acute{y}_n$ and $(\mathscr{C}^n_w, p) \models \neg\sigma_j$ imply that $(\mathscr{C}^n_w, \grave{y}_n) \models [d]\sigma_j$ and $(\mathscr{C}^n_w, \acute{y}_n) \models [d]\sigma_j$. In a similar fashion, the induction hypothesis and the fact that $\grave{z} R_d \grave{x}_{n-1}$, $\acute{z}_n R_d \acute{x}_{n-1}$, and there is no point p that is a d-successor of either $\grave{z}_n$ or $\acute{z}_n$ and $(\mathscr{C}^n_w, p) \models \neg\sigma_j$ imply that $(\mathscr{C}^n_w, \grave{z}_n) \models [d]\sigma_j$ and $(\mathscr{C}^n_w, \acute{z}_n) \models [d]\sigma_j$. Given the shape of the model $\mathscr{C}^n_w$, we obtain that $(\mathscr{C}^n_w, \grave{y}_n) \models \mathbf{c} \wedge [c][d]\sigma_j$ and $(\mathscr{C}^n_w, \acute{y}_n) \models \mathbf{c} \wedge [c][d]\sigma_j$. We know that for some $l \in \{a, b\}$, $\grave{y}_n$ is an l-successor of $\grave{x}_n$, $\acute{y}_n$ is an l-successor of $\grave{x}_n$ and there is no other point p that is an l-successor of either $\grave{x}_n$ or $\acute{x}_n$ and $(\mathscr{C}^n_w, p) \models \neg(\mathbf{c} \wedge [c][d]\sigma_j)$. Therefore, $(\mathscr{C}^n_w, \grave{x}_n) \models [l](\mathbf{c} \wedge [c][d]\sigma_j)$ and $(\mathscr{C}^n_w, \acute{x}_n) \models [l](\mathbf{c} \wedge [c][d]\sigma_j)$. Hence, $(\mathscr{C}^n_w, \grave{x}_n) \models \sigma_{j+1}$ and $(\mathscr{C}^n_w, \acute{x}_n) \models \sigma_{j+1}$. The fact that $(\mathscr{C}^n_w, \grave{x}_n) \models \sigma_n$ can be established in the same way. Next, we show that **A2** is true, i.e., that $(\mathscr{C}^n_w, \acute{x}_n) \models \neg\sigma_n$. According to the induction hypothesis, $(\mathscr{C}^n_w, \acute{x}_{n-1}) \models \neg\sigma_{n-1}$. Additionally, we have that $\acute{x}_n$ and $\acute{y}_n$ are l-connected for some $l \in \{a, b\}$; moreover $\acute{y}_n$ and $\acute{z}_n$ are c-connected while $\acute{z}_n$ and $\acute{x}_{n-1}$ are d-connected. Therefore, $(\mathscr{C}^n_w, \acute{x}_{n-1}) \models \langle l \rangle(\neg\mathbf{c} \vee \langle c \rangle\langle d \rangle\neg\sigma_{n-1})$. On the other hand, we know already that $(\mathscr{C}^n_w, \overline{x}_n) \models \neg\sigma_{n-1}$. Given the shape of the model $\mathscr{C}^n_w$, there is a $k \in \{a, b\}$ such that $k \neq l$ and $\overline{x}_n$ is a k-successor or $\acute{x}_n$. Hence, $(\mathscr{C}^n_w, \acute{x}_{n-1}) \models \langle k \rangle(\neg\mathbf{c} \vee \langle c \rangle\langle d \rangle\neg\sigma_{n-1})$, and, therefore, $(\mathscr{C}^n_w, \acute{x}_n) \models \neg\sigma_n$.

□

Next, following the conventions from Definitions 5.3 and 5.5, we define suitable sets of models $\mathbb{E}^n$ and $\mathbb{F}^n$ for every formula $\omega_n \in \Omega_{[\varphi]}$.

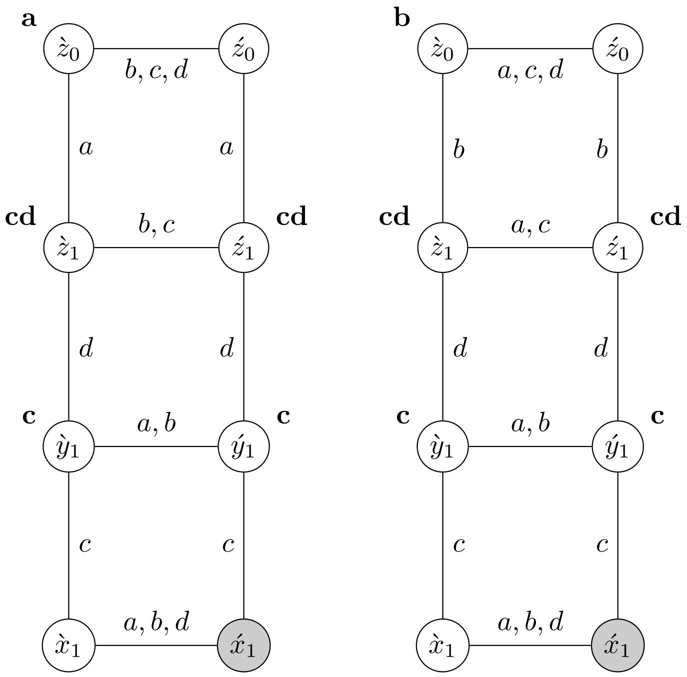
Definition 5.7 (The models $\mathbb{E}$)**.** For every natural number $n \geq 1$, the sets $\mathbb{E}^n$ and $\mathbb{F}^n$, each containing 2^n different pointed models, are defined recursively as follows.

1. The set $\mathbb{E}^1$ consists of the pointed models $(\mathscr{E}^1_{cda}, \grave{x}_1)$ and $(\mathscr{E}^1_{cdb}, \grave{x}_1)$ (see Figure 5.9). The set $\mathbb{F}^1$ contains the pointed models $(\mathscr{E}^1_{cda}, \acute{x}_1)$ and $(\mathscr{E}^1_{cdb}, \acute{x}_1)$ shown in Figure 5.10. Note that this time the nodes $\grave{z}_1$ and $\acute{z}_1$ satisfy both propositions **c** and **d**.



$(\mathscr{E}^1_{cda}, \grave{x}_1)$ $(\mathscr{E}^1_{cdb}, \grave{x}_1)$

FIGURE 5.9: The set of pointed models $\mathbb{E}^1$.



$(\mathscr{E}^1_{cda}, \acute{x}_1)$ $(\mathscr{E}^1_{cdb}, \acute{x}_1)$

FIGURE 5.10: The set of pointed models $\mathbb{F}^1$.

n+1. The set $\mathbb{E}^{n+1}$ consists of all pointed models shown in Figure 5.11. The models comprising the set $\mathbb{F}^{n+1}$ are shown in Figure 5.12.

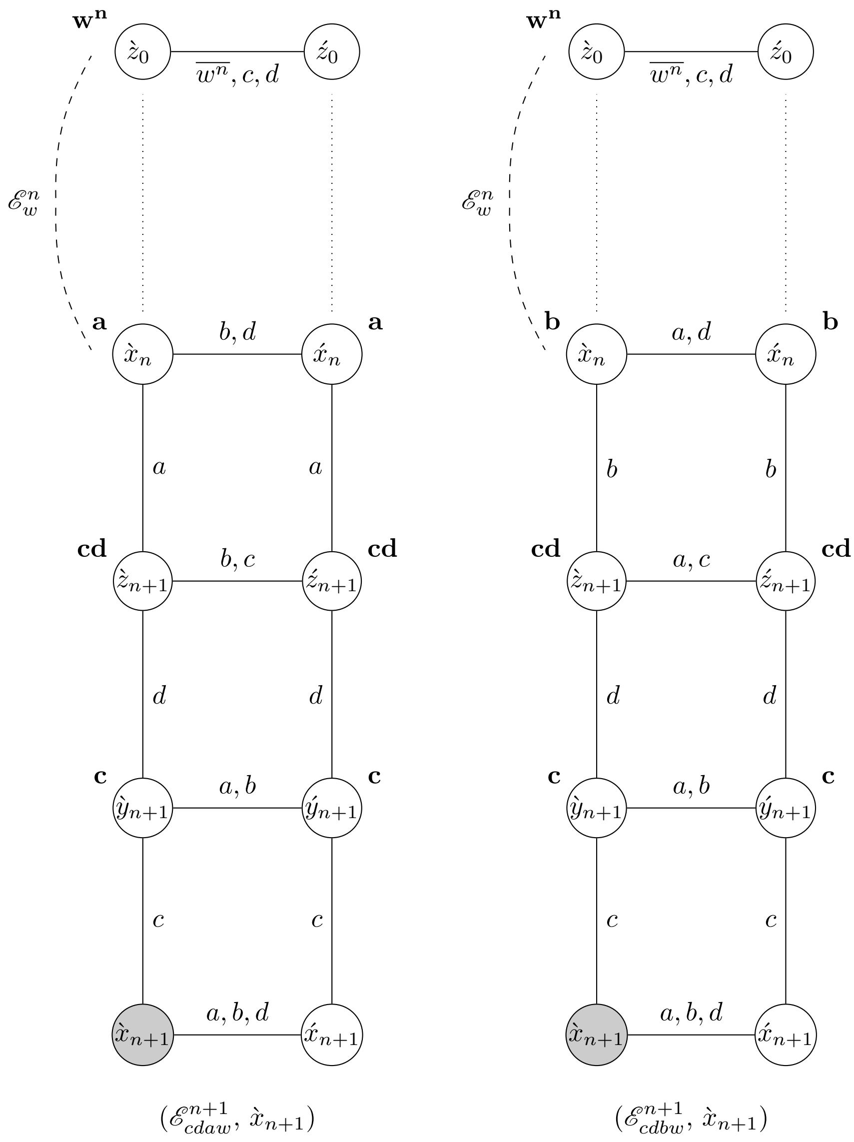


FIGURE 5.11: The pointed models in $\mathbb{E}^{n+1}$

Let $\mathbb{E}$ be the union of all $\mathbb{E}^n$ and $\mathbb{F}^n$.

As before, the subscript w in the name of the model $\mathscr{E}^n_w$ encodes the shortest path of relations steps that leads from $\grave{x}_n$ to the upper leftmost node which in this case is $\grave{z}_0$. Since w has the form $cdj_1 cdj_2 \ldots cdj_n$, where $j_i \in \{a, b\}$ and there are 2^n different such words, we see again that both $\mathbb{E}^n$ and $\mathbb{F}^n$ contain 2^n different pointed models corresponding to the different words w. We use $\mathbf{w^n}$ to denote the propositional symbol that corresponds to the last letter in the word w that is used as a subscript in the name of the model $\mathscr{E}^n_w$. For example, if $w = cdj_1 cdj_2 \ldots cdj_n$, where $j_n = a$, then $\mathbf{w^n}$ is the propositional symbol **a**. Similarly, if $j_n = b$, then $\mathbf{w^n}$ stands for **b**.

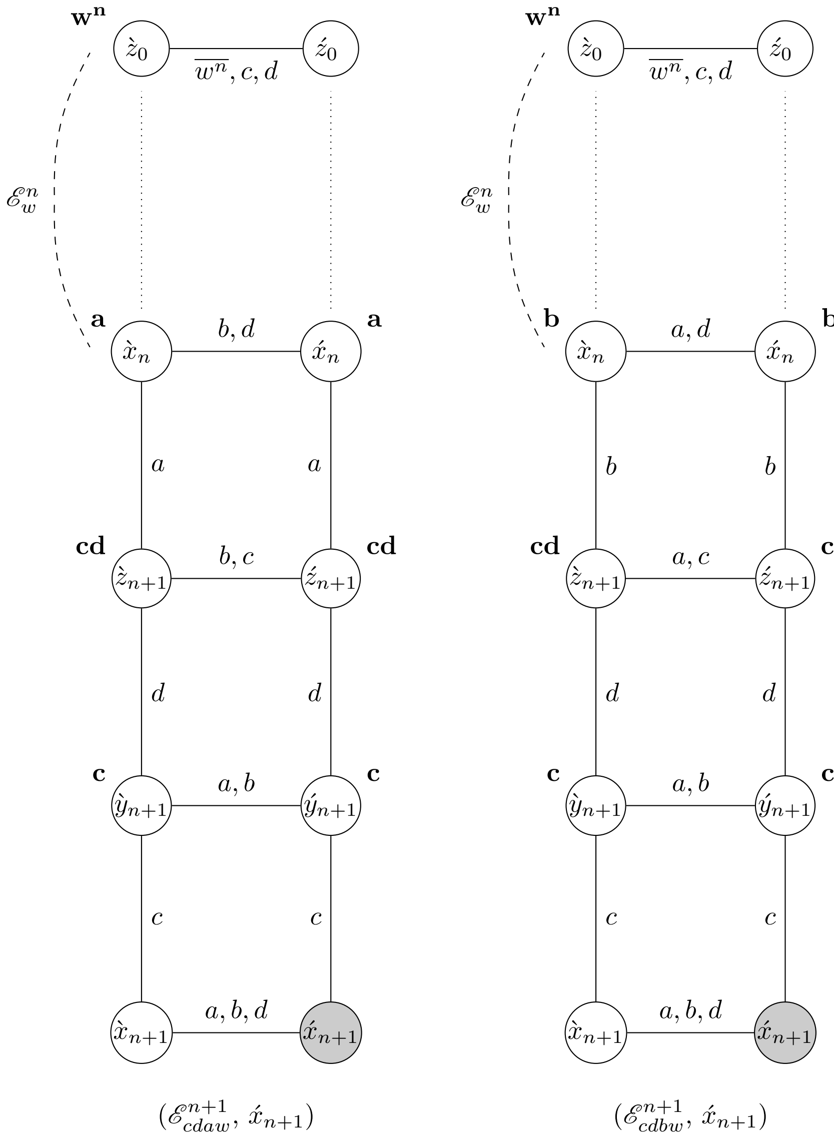


FIGURE 5.12: The pointed models in $\mathbb{F}^{n+1}$.

Example 5.3. *Figure 5.13 shows the pair of pointed models* $(\mathscr{E}^2_{cdacda}, \grave{x}_2)$ *and* $(\mathscr{E}^2_{cdacda}, \acute{x}_2)$*. Note how the model* $\mathscr{E}^1_{cda}$ *that is used in the construction of* $\mathscr{E}^2_{cdacda}$ *differs from the model* $\mathscr{E}^1_{cda}$ *from Figure 5.9. We have that* $\grave{x}_1 R_a \acute{x}_1$ *for the model* $\mathscr{E}^1_{cda}$ *from Figure 5.9. This is not true for the nodes* $\grave{x}_1$ *and* $\acute{x}_1$ *from the model* $\mathscr{E}^2_{cdacda}$ *depicted in Figure 5.13.*

The analogue of Propositions 5.4 and 5.6 in the case of the sets of pointed models $\mathbb{E}^n$ and $\mathbb{F}^n$ is Proposition 5.8 below. We have again that the models in $\mathbb{E}^n$ can be differentiated from the models in $\mathbb{F}^n$ by using the relevant formulae from Table 5.2 which, in this case, are the formulae from the set $\Omega_{[\varphi]}$; moreover, no two pointed models $(\mathscr{E}^n_w, \grave{p})$ and $(\mathscr{E}^n_w, \acute{p})$ can be differentiated by a formula $[j]\varphi \in \Phi_{\mathrm{ML}}$, where j is such that $\grave{p} R_j \acute{p}$.

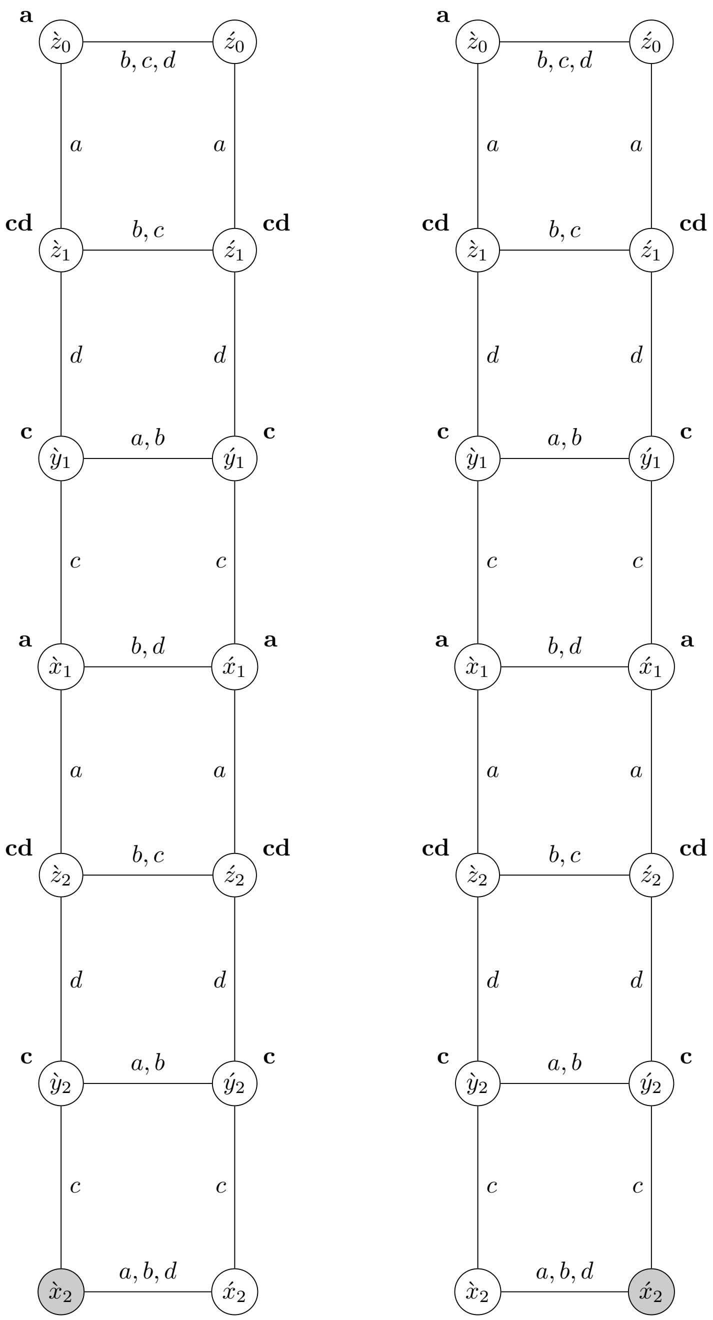


FIGURE 5.13: The pointed models $(\mathscr{E}^2_{cdacda}, \grave{x}_2)$ (left) and $(\mathscr{E}^2_{cdacda}, \acute{x}_2)$ (right).

Proposition 5.8. *The following are true.*

- *For all $n \geq 1$ and all $\omega_n \in \Omega_{[\varphi]}$,*

$$\mathbb{E}^n \models \omega_n \text{ and } \mathbb{F}^n \models \neg\omega_n$$

- *For all pairs of pointed models* $(\mathscr{E}^n_w, \grave{p}_i)$ *and* $(\mathscr{E}^n_w, \acute{p}_n)$*, where* $1 \leq i \leq n$*, and any formula* $[j]\varphi \in \Phi_{\mathrm{ML}}$*, if* $\grave{p}_i$ *and* $\acute{p}_i$ *are* j*-connected, then*

$$(\mathscr{E}^n_w, \grave{p}_i) \models [j]\varphi \textit{ iff } (\mathscr{E}^n_w, \acute{p}_i) \models [j]\varphi.$$

Proof. Using the equivalences 3.1 and the fact that $\langle\varphi\rangle\psi$ is defined as $\neg[\varphi]\neg\psi$, it is easy to see that the formula $\omega_n \stackrel{\text{def}}{=} \langle\omega_{n-1}\rangle\omega_1$ is equivalent to

$$\omega_{n-1} \wedge \langle c\rangle(\omega_{n-1} \wedge \mathbf{c} \wedge \langle d\rangle(\omega_{n-1} \wedge \mathbf{d} \wedge (\langle a\rangle(\omega_{n-1} \wedge \mathbf{a}) \vee \langle b\rangle(\omega_{n-1} \wedge \mathbf{b}))))$$

whereas $\neg\langle\omega_{n-1}\rangle\omega_1$ is equivalent to

$$\neg\omega_{n-1} \vee [c](\neg\omega_{n-1} \vee \neg\mathbf{c} \vee [d](\neg\omega_{n-1} \vee \neg\mathbf{d} \vee ([a](\neg\omega_{n-1} \vee \neg\mathbf{a}) \wedge [b](\neg\omega_{n-1} \vee \neg\mathbf{b}))))$$

We prove by induction on j that for every model $\mathscr{E}^n_w$ and every $1 \leq j \leq n$,

1. If $j < i \leq n$, then
   - $(\mathscr{E}^n_w, \grave{x}_i) \models \omega_j$;
   - $(\mathscr{E}^n_w, \acute{x}_i) \models \omega_j$;
   - $(\mathscr{E}^n_w, \grave{y}_i) \models \omega_j$;
   - $(\mathscr{E}^n_w, \acute{y}_i) \models \omega_j$;
   - $(\mathscr{E}^n_w, \grave{z}_i) \models \omega_j$;
   - $(\mathscr{E}^n_w, \acute{z}_i) \models \omega_j$.
2. If $j = i$, then $(\mathscr{E}^n_w, \grave{x}_i) \models \omega_j$, whereas $(\mathscr{E}^n_w, \acute{x}_i) \models \neg\omega_j$.

We assume that $1 < n$. The case $n = 1$ is covered in the base case below which we prove in details. The induction step is left to the reader.

Base case. Let $j = 1$ and $j < i$.

1. For every $1 < k \leq n$, we have:
   (a) $\grave{x}_k R_c \grave{y}_k$ and $\acute{x}_k R_c \acute{y}_k$, and both $\grave{y}_k$, $\acute{y}_k$ satisfy $\mathbf{c}$;
   (b) $\grave{y}_k R_d \grave{z}_k$ and $\acute{y}_k R_d \acute{z}_k$, and $\grave{z}_k$, $\acute{z}_k$ satisfy $\mathbf{d}$;
   (c) There are $l \in \{a, b\}$ and $\mathbf{l} \in \{\mathbf{a}, \mathbf{b}\}$, such that $\grave{z}_k R_l \grave{x}_{k-1}$ and $\acute{z}_k R_l \acute{x}_{k-1}$, and both $\grave{x}_{k-1}$, and $\acute{x}_{k-1}$ satisfy $\mathbf{l}$.

   Therefore, $(\mathscr{E}^n_w, \grave{x}_k) \models \langle c\rangle(\mathbf{c} \wedge \langle d\rangle(\mathbf{d} \wedge (\langle a\rangle\mathbf{a} \vee \langle b\rangle\mathbf{b})))$ and the same is true about $(\mathscr{E}^n_w, \acute{x}_k)$, i.e., $(\mathscr{E}^n_w, \acute{x}_k) \models \langle c\rangle(\mathbf{c} \wedge \langle d\rangle(\mathbf{d} \wedge (\langle a\rangle\mathbf{a} \vee \langle b\rangle\mathbf{b})))$.

   Let us prove the statement for x_k and y_k. Given the shape of $\mathscr{E}^n_w$, the following are true:

   (a) $\grave{y}_k R_c \grave{y}_k$ and $\acute{y}_k R_c \acute{y}_k$, and $\grave{y}_k$ and $\acute{y}_k$ satisfy $\mathbf{c}$;

(b) $\grave{y}_k R_d \grave{z}_k$ and $\acute{y}_k R_d \acute{z}_k$, and both $\grave{z}_k$ and $\acute{z}_k$ satisfy $\mathbf{d}$;

(c) There are $l \in \{a, b\}$ and $\mathbf{l} \in \{\mathbf{a}, \mathbf{b}\}$, such that $\grave{z}_k R_l \grave{x}_{k-1}$ and $\acute{z}_k R_l \acute{x}_{k-1}$, and $\grave{x}_{k-1}$, $\acute{x}_{k-1}$ satisfy $\mathbf{l}$.

Therefore, $(\mathscr{E}^n_w, \grave{y}_k) \models \varphi_1$ and $(\mathscr{E}^n_w, \acute{y}_k) \models \varphi_1$.

Similarly,

(a) $\grave{z}_k R_c \grave{z}_k$, $\grave{z}_k R_d \grave{z}_k$ and $\acute{z}_k R_c \acute{z}_k$, $\acute{z}_k R_d \acute{z}_k$; moreover, $\grave{z}_k$, $\acute{z}_k$ satisfy both $\mathbf{c}$ and $\mathbf{d}$;

(b) There are $l \in \{a, b\}$ and $\mathbf{l} \in \{\mathbf{a}, \mathbf{b}\}$, such that $\grave{z}_k R_l \grave{x}_{k-1}$ and $\acute{z}_k R_l \acute{x}_{k-1}$, and $\grave{x}_{k-1}$, $\acute{x}_{k-1}$ satisfy $\mathbf{l}$;

Hence, both $(\mathscr{E}^n_w, \grave{z}_k)$ and $(\mathscr{E}^n_w, \acute{z}_k)$ satisfy $\langle c \rangle(\mathbf{c} \wedge \langle d \rangle(\mathbf{d} \wedge (\langle a \rangle \mathbf{a} \vee \langle b \rangle \mathbf{b})))$, i.e., $(\mathscr{E}^n_w, \grave{z}_k) \models \varphi_1$ and $(\mathscr{E}^n_w, \acute{z}_k) \models \varphi_1$.

2. To complete the base case, we have to prove that $(\mathscr{E}^n_w, \grave{x}_1) \models \omega_1$ whereas $(\mathscr{E}^n_w, \acute{x}_1) \models \neg\omega_1$. Indeed, we have

   (a) $\grave{x}_1 R_c \grave{y}_1$ and $\grave{y}_1$ satisfies $\mathbf{c}$;

   (b) $\grave{y}_1 R_d \grave{z}_1$ and $\grave{z}_1$ satisfies $\mathbf{d}$;

   (c) There are $l \in \{a, b\}$ and $\mathbf{l} \in \{\mathbf{a}, \mathbf{b}\}$, such that $\grave{z}_1 R_l \grave{z}_0$ and $\grave{z}_0$ satisfies $\mathbf{l}$;

   Therefore, $(\mathscr{E}^n_w, \grave{x}_1) \models \langle c \rangle(\mathbf{c} \wedge \langle d \rangle(\mathbf{d} \wedge (\langle a \rangle \mathbf{a} \vee \langle b \rangle \mathbf{b})))$ and, thus, $(\mathscr{E}^n_w, \grave{x}_1) \models \omega_1$.

   It is easy to see that $\neg\omega_1$ is equivalent to $[c](\neg\mathbf{c} \vee [d](\neg\mathbf{d} \vee ([a]\neg\mathbf{a} \wedge [b]\neg\mathbf{b})))$.

   Again, given the shape of the model $\mathscr{E}^n_w$, we see that the following are true

   (a) $(\mathscr{E}^n_w, \acute{x}_1) \models \neg\mathbf{c}$;

   (b) $(\mathscr{E}^n_w, \acute{y}_1) \models \neg\mathbf{d}$;

   (c) $(\mathscr{E}^n_w, \acute{z}_1) \models ([a]\neg\mathbf{a} \wedge [b]\neg\mathbf{b})$;

   (d) $\acute{y}_1 R_d \acute{z}_1$, and there is no point p that is a d-successor of either $\acute{y}_1$ or $\acute{z}_1$ and such that $(\mathscr{E}^n_w, p) \models (\mathbf{d} \wedge (\langle a \rangle \mathbf{a} \vee \langle b \rangle \mathbf{b}))$.

   Therefore, we have $(\mathscr{E}^n_w, \acute{y}_1) \models [d](\neg\mathbf{d} \vee ([a]\neg\mathbf{a} \wedge [b]\neg\mathbf{b}))$.

   Given all the items above, the fact that $\acute{x}_1 R_c \acute{y}_1$ and there is no point p such that p is a c-successor of either $\acute{x}_1$ or $\acute{y}_1$, and $(\mathscr{E}^n_w, p) \models \langle c \rangle(\mathbf{c} \wedge \langle d \rangle(\mathbf{d} \wedge (\langle a \rangle \mathbf{a} \vee \langle b \rangle \mathbf{b})))$, we obtain $(\mathscr{E}^n_w, \acute{x}_1) \models [c](\neg\mathbf{c} \vee [d](\neg\mathbf{d} \vee ([a]\neg\mathbf{a} \wedge [b]\neg\mathbf{b})))$ and, thus, $(\mathscr{E}^n_w, \acute{x}_1) \models \neg\omega_1$.

□

Finally, following the pattern established above, for every $n \geq 1$, we construct two sets of pointed models $\mathbb{G}^n$ and $\mathbb{H}^n$ associated with each formula $\theta_n \in \Theta_{[\cap]}$.

Definition 5.9 (The models $\mathbb{G}$)**.** The sets $\mathbb{G}^n$ and $\mathbb{H}^n$, each containing 2^n different pointed models, are defined recursively as shown.

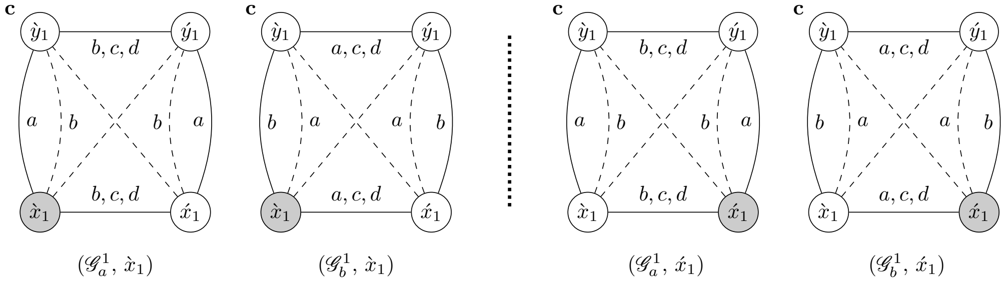


FIGURE 5.14: The sets $\mathbb{G}^1$ and $\mathbb{H}^1$.

1. The set $\mathbb{G}^1$ consists of the pointed models $(\mathscr{G}_a^1, \grave{x}_1)$ and $(\mathscr{G}_b^1, \grave{x}_1)$ shown on the left of the dotted line in Figure 5.14. The set $\mathbb{H}^1$ consists of the models $(\mathscr{G}_a^1, \acute{x}_1)$ and $(\mathscr{G}_b^1, \acute{x}_1)$ shown on the right. Note that the dashed line represents the relation R_b in the model $\mathscr{G}_a^1$ which means that all four nodes in $\mathscr{G}_a^1$ are b-connected; similarly, the dashed line in $\mathscr{G}_b^1$ represents the relation R_a connecting all nodes in the model.

$n+1$. The set $\mathbb{G}^{n+1}$ consists of all pointed models $(\mathscr{G}_{acdw}^{n+1}, \grave{x}_{n+1})$ and $(\mathscr{G}_{bcdw}^{n+1}, \grave{x}_{n+1})$ shown in Figure 5.15. Similarly, the set $\mathbb{H}^{n+1}$ consists of all the pointed mod-

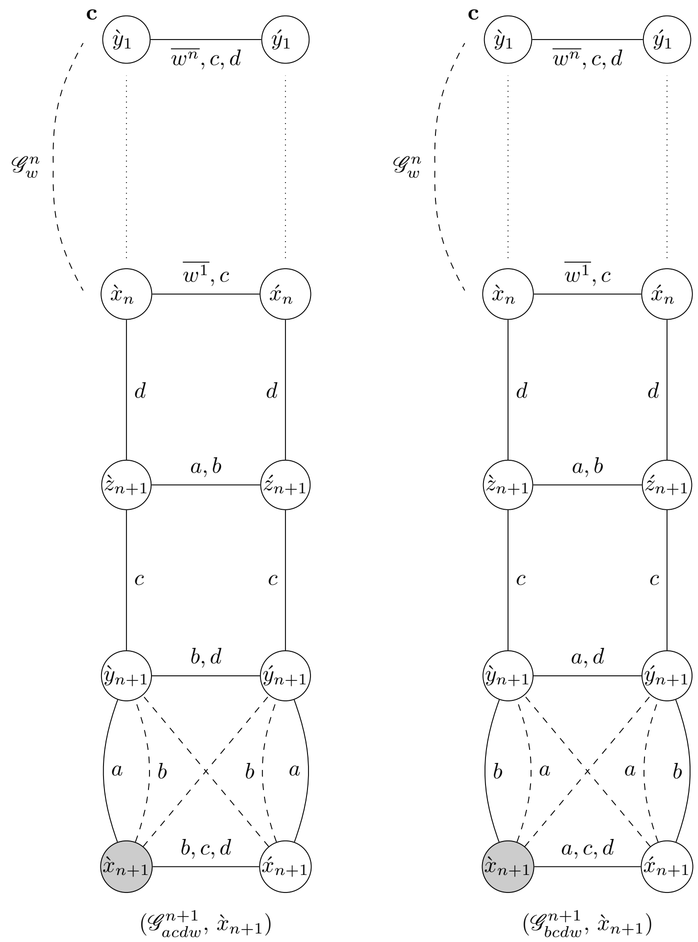


FIGURE 5.15: The set of pointed models pointed models $\mathbb{G}^{n+1}$.

els $(\mathscr{G}^{n+1}_{acdw}, \acute{x}_{n+1})$ and $(\mathscr{G}^{n+1}_{bcdw}, \acute{x}_{n+1})$ shown in Figure 5.16.

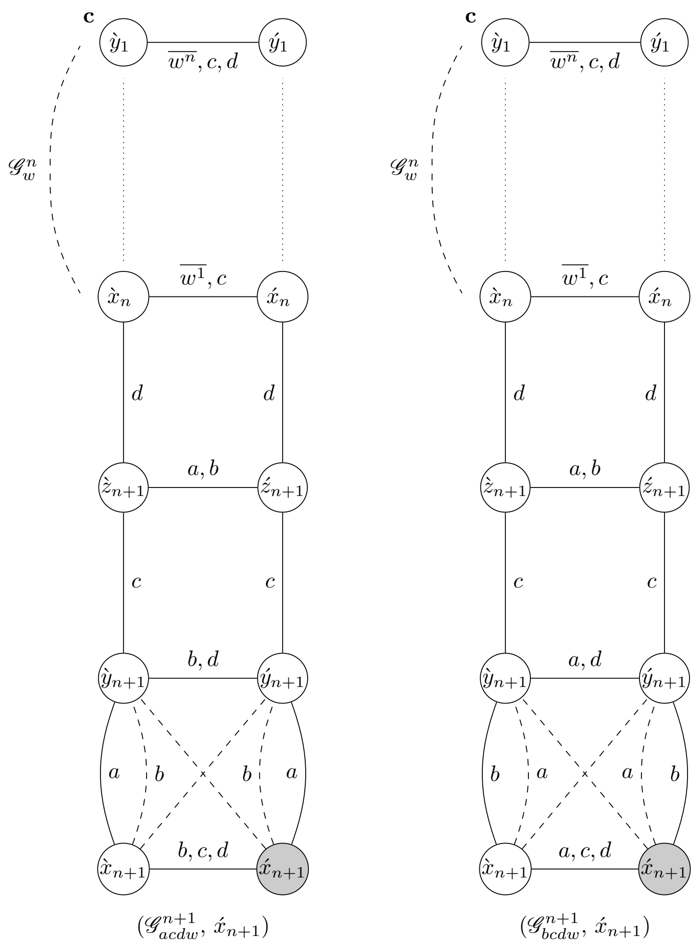


FIGURE 5.16: The set of pointed models pointed models $\mathbb{H}^{n+1}$.

Again, we have an additional relation step, either a or b, that is represented by the dashed line connecting the nodes $\grave{x}_{n+1}$, $\grave{y}_{n+1}$, $\acute{x}_{n+1}$, and $\acute{y}_{n+1}$. We would like to remind the reader that the nodes $\grave{x}_n$ and $\acute{x}_n$ are d-connected in the model $\mathscr{G}^n_w$ but they are **not** d-connected in the model $\mathscr{G}^{n+1}_{jw}$, where $j \in \{a, b\}$. For any model $\mathscr{G}^n_w$, the only point that satisfies the proposition **c** is $\grave{y}_1$. The rest of the points do not satisfy any proposition.

The set $\mathbb{G}$ is defined as the union of all sets $\mathbb{G}^n$ and $\mathbb{H}^n$.

Example 5.4. *Figure 5.17 shows the pointed models* $(\mathscr{G}^2_{acda}, \grave{x}_2)$ *and* $(\mathscr{G}^2_{acda}, \acute{x}_2)$*. Note that the dashed line represents the relation* R_b*. Furthermore, although* $\grave{x}_1$ *and* $\acute{x}_1$ *are* d*-connected in the model* G^1_a *used to build the model* $\mathscr{G}^2_{acda}$*, they are not* d*-connected in the latter.*

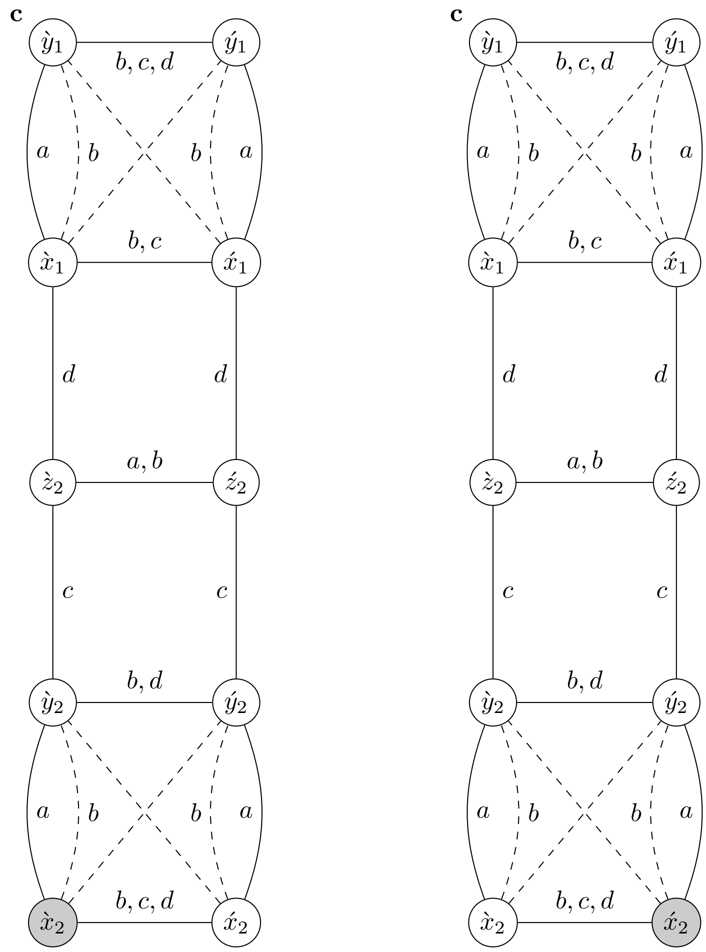


FIGURE 5.17: The pointed models $(\mathscr{G}^2_{acda}, \grave{x}_2)$ (left) and $(\mathscr{G}^n_{acda}, \acute{x}_2)$ (right).

Yet again, we have an analogue to Propositions 5.8, 5.6, and 5.4 .

Proposition 5.10. *The following are true.*

- *For all* $n \geq 1$ *and any formula* $\theta_n \in \Theta_{[\cap]}$,

$$\mathbb{G}^n \models \theta_n \text{ and } \mathbb{H}^n \models \neg\theta_n.$$

- *For any pair of pointed models* $(\mathscr{G}^n_w, \grave{p}_i)$ *and* $(\mathscr{G}^n_w, \acute{p}_i)$*, where* $1 \leq i \leq n$*, and any formula* $[j]\varphi \in \Phi_{\mathrm{ML}}$*, if* $\grave{p}_i$ *and* $\acute{p}_i$ *are* j*-connected, then*

$$(\mathscr{G}^n_w, \grave{p}_i) \models [j]\varphi \text{ iff } (\mathscr{G}^n_w, \acute{p}_i) \models [j]\varphi.$$

Proof. The truth of the first item is easily seen. However, we would like to point out that, for every formula θ_n, there is an equivalent on $\mathbb{G}$ formula $\theta'_n \in \Phi_{\mathrm{ML}}$ defined recursively as follows. The formula $\theta_1 \stackrel{\text{def}}{=} \langle \cap_{\{a,b\}} \rangle \mathbf{c}$ is equivalent to $\theta'_1 \stackrel{\text{def}}{=} \langle a \rangle \mathbf{c} \wedge \langle b \rangle \mathbf{c}$ on $\mathbb{G}$ and for $n > 1$, θ_n is equivalent to $\langle a \rangle \langle c \rangle \langle d \rangle \theta'_{n-1} \wedge \langle b \rangle \langle c \rangle \langle d \rangle \theta'_{n-1}$.

The proof of the second item is the same as in Proposition 5.4. □

5.3 The proof

The main idea behind the proofs of the succinctness results in Chapter 4 was that the two sets of pointed models $\mathbb{A}^n$ and $\mathbb{B}^n$ or $\mathbb{O}^n$ and $\mathbb{P}^n$ contained 2^n diverging pairs. In the case of $\mathbb{A}^n$ and $\mathbb{B}^m$, these were the pairs $\langle(\mathscr{A}^n_w, \alpha^n_w), (\mathscr{B}^n, \beta^n)\rangle$ corresponding to the different words w; similarly, in the case of $\mathbb{O}^n$ and $\mathbb{P}^n$, we had the pairs $\langle(\mathscr{O}^n_w, o^n_w), (\mathscr{P}^n_w, \rho^n_w)\rangle$. The argument that we used to show that, for example, for any two different words w and $\underline{w}$, the pairs $\langle(\mathscr{O}^n_w, o^n_w), (\mathscr{P}^n_w, \rho^n_w)\rangle$ and $\langle(\mathscr{O}^n_{\underline{w}}, o^n_{\underline{w}}), (\mathscr{P}^n_{\underline{w}}, \rho^n_{\underline{w}})\rangle$ were diverging was based on the fact that any extended syntax tree with root $\{(\mathscr{O}^n_w, o^n_w)\} \circ \{(\mathscr{P}^n_w, \rho^n_w)\}$ contains a branch B such that $I(B) = w$ whereas any extended syntax tree with root $\{(\mathscr{O}^n_{\underline{w}}, o^n_{\underline{w}})\} \circ \{(\mathscr{P}^n_{\underline{w}}, \rho^n_{\underline{w}})\}$ contains a branch $\underline{B}$ such that $I(\underline{B}) = \underline{w}$. Since the words w and $\underline{w}$ are different, we conclude that the pairs are diverging. We will follow a similar strategy in the present chapter, too. However, this time, we cannot prove, for example, that any extended syntax tree with root $\{(\mathscr{G}^n_w, \grave{x}_n)\} \circ \{(\mathscr{G}^n_w, \acute{x}_n)\}$ contains a branch B such that $I(B) = w$ because the relations in our models are reflexive, symmetric and transitive. Nevertheless, we can prove a suitable version of Lemma 4.7 which says that any extended syntax tree with root $\{(\mathscr{G}^n_w, \grave{x}_n)\} \circ \{(\mathscr{G}^n_w, \acute{x}_n)\}$ contains a branch B such that $I(B)$ is a path over $\mathscr{G}^n_w$; moreover, different w give rise to different paths.

Intuitively, a path over a model $\mathscr{N}^n_w \in \{\mathscr{A}^n_w, \mathscr{C}^n_w, \mathscr{E}^n_w, \mathscr{G}^n_w\}$ is a sequence of relation steps leading from the left bottom point $\grave{x}_n$ of $\mathscr{N}^n_w$ to the left top point such that only points on the left support of $\mathscr{N}^n_w$ are visited. For example, let us consider the model $\mathscr{G}^2_{acda}$ from Figure 5.17. One path over this model is the sequence of relation steps $acdcdda$ that leads from $\grave{x}_2$ to $\grave{y}_1$ in the following way $\grave{x}_2 R_a \grave{y}_2 R_c \grave{z}_2 R_d \grave{z}_2 R_c \grave{z}_2 R_d \grave{x}_1 R_d \grave{x}_1 R_a \grave{y}_1$". In other words, the individual steps in the sequence can only be "up", "down", or "make a reflexive step and stay at the current point". However, in the last case, a reflexive step at some point $\grave{p}_j$ must not use an index i that also connects $\grave{p}_j$ with $\acute{p}_j$. For example, a path over $\mathscr{G}^2_{acda}$ that starts with b is not allowed because in this model, we have $\grave{x}_2 R_b \acute{x}_2$. We formalise this intuition in the next definition.

Definition 5.11 (Paths Of Relation Steps)**.**

For any model $\mathscr{M}^n_w \in \{\mathscr{A}^n_w, \mathscr{C}^n_w, \mathscr{G}^n_w\}$, let the collection of left hand nodes $LHN(\mathscr{M}^n_w)$ be the set $\{\grave{x}_n, \grave{y}_n, \grave{z}_n, \ldots, \grave{x}_2, \grave{y}_2, \grave{z}_2, \grave{x}_1, \grave{y}_1\}$ and the top left node $tln(\mathscr{M}^n_w)$ be $\grave{y}_1$. Likewise, $LHN(\mathscr{E}^n_w)$ be the set $\{\grave{x}_n, \grave{y}_n, \grave{z}_n, \ldots, \grave{x}_2, \grave{y}_2, \grave{z}_2, \grave{x}_1, \grave{y}_1, \grave{z}_1, \grave{z}_0\}$ and $tln(\mathscr{E}^n_w) = \grave{z}_0$.

For any $\mathscr{N}^n_w \in \{\mathscr{A}^n_w, \mathscr{C}^n_w, \mathscr{E}^n_w, \mathscr{G}^n_w\}$[1], a path π over $\mathscr{N}^n_w$ is a sequence of indices $\pi_1, \pi_2 \ldots, \pi_k$ such that there is a sequence of points $u_1, \ldots, u_{k+1}$ for which the following are true.

1. $u_i \in LHN(\mathscr{N}^n_w)$ for all $1 \leq i \leq k+1$;

[1]Note, that technically speaking, using $\mathscr{N}^n_w$ as a variable over the set $\{\mathscr{A}^n_w, \mathscr{C}^n_w, \mathscr{E}^n_w, \mathscr{G}^n_w\}$ may lead to the confusion that the word w is the same for the models in the set $\{\mathscr{A}^n_w, \mathscr{C}^n_w, \mathscr{G}^n_w\}$ and $\mathscr{E}^n_w$. This is true for the former but we must keep in mind that a word w that is used as a subscript in the names of the models from $\{\mathscr{A}^n_w, \mathscr{C}^n_w, \mathscr{G}^n_w\}$ has the form $i_1 cd i_2 cd \ldots cd i_n$, where $i_k \in \{a, b\}$ whereas a word w in $\mathscr{E}^n_w$ has the form $cd i_1 cd i_2 \ldots cd i_n$ where $i_k \in \{a, b\}$. Although this is an abuse of notation, we prefer to formulate our definition in this way rather than giving a separate explanation for the models $\mathscr{E}^n_w$. We hope that this does not lead to any misunderstanding.

2. $u_1 = \grave{x}_n$ and $u_{k+1} = tln(\mathscr{N}_w^n)$

3. u_{i+1} is a π_i-successor of u_i in $\mathscr{N}_w^n$ for all $i \leq k$;

4. for all $i \leq k$, if $u_i = \grave{p}_j$ for some $\grave{p}_j \in LHM(\mathscr{N}_w^n)$, then $\grave{p}_j$ and $\acute{p}_j$ are not π_i-connected in $\mathscr{N}_w^n$.

We have the following very simple proposition.

Proposition 5.12.

- *If* $\mathscr{N}_w^n \in \{\mathscr{A}_w^n, \mathscr{C}_w^n, \mathscr{G}_w^n\}$ *and* $w = i_1 cdi_2 cd \dots cdi_n$, *where* $i_k \in \{a, b\}$, *then the sequence of indices* $\pi_w = i_1 cdi_2 cd \dots cdi_n$ *is a path over* $\mathscr{N}_w^n$.

- *Let us consider a model* $\mathscr{E}_w^n$. *If* $w = cdi_1 cdi_2 \dots cdi_n$, *where* $i_k \in \{a, b\}$, *then the sequence of indices* $\pi_w = cdi_1 cdi_2 \dots cdi_n$ *is a path over* $\mathscr{E}_w^n$.

Proof. The proofs of both items are analogous. Let us consider the first. It is obvious that there is a sequence of points, namely $\grave{x}_n, \grave{y}_n, \grave{z}_n \dots, \grave{x}_2, \grave{y}_2, \grave{z}_2, \grave{x}_1, \grave{y}_1$ such that the requirements of Definition 5.11 are fulfilled. In the case of $\mathscr{E}_w^n$, we see that $\grave{x}_n, \grave{y}_n, \grave{z}_n \dots, \grave{x}_1, \grave{y}_1, \grave{z}_1, \grave{z}_0$ fulfil the requirements of Definition 5.11. □

Definition 5.13 (Canonical paths)**.** For any $\mathscr{N}_w^n \in \{\mathscr{A}_w^n, \mathscr{C}_w^n, \mathscr{E}_w^n, \mathscr{G}_w^n\}$, the word w is called the canonical path over $\mathscr{N}_w^n$. Let $\pi = \pi_1 \dots \pi_k$ be a path over $\mathscr{N}_w^n$ and let $u_1, u_2 \dots, u_{k+1}$ be the points satisfying Definition 5.11. The word $u_1 \pi_1 u_2 \dots, u_k \pi_k u_{k+1}$ is called an extended path. If $\pi = \pi_w$, then this word is called the extended canonical path.

Note that the extended canonical path is unique, i.e., the point p_1 is always $\grave{x}_n$, the point p_2 is always $\grave{y}_n$, the point p_3 is always $\grave{z}_n$ etc.

As we said earlier, intuitively, for any $\mathscr{N}_w^n \in \{\mathscr{A}_w^n, \mathscr{C}_w^n, \mathscr{E}_w^n, \mathscr{G}_w^n\}$ the word w encodes the shortest path from $\grave{x}_n$ to $tln(\mathscr{N}_w^n)$. Using Definitions 5.11 and 5.13, we can express this more formally as follows.

Proposition 5.14. *For any* $\mathscr{N}_w^n \in \{\mathscr{A}_w^n, \mathscr{C}_w^n, \mathscr{E}_w^n, \mathscr{G}_w^n\}$, *the canonical path* π_w *is the shortest path over the model* $\mathscr{N}_w^n$.

Proof. Using Definition 5.11, it is easy to see that each one of the points $\grave{x}_n$, $\grave{y}_n$, $\grave{z}_n$, $\dots$, $\grave{x}_2$, $\grave{y}_2$, $\grave{z}_2$, $\grave{x}_1$, $tln(\mathscr{N}_w^n)$ must appear at least once in any extended path over $\mathscr{N}_w^n$. Each one of these points appears exactly once in the extended canonical path π_w. Hence, there is no shorter path than π_w.

□

We are ready now to formulate a lemma analogous to Lemma 4.7.

Lemma 5.15. *For any* $\mathscr{N}_w^n \in \{\mathscr{A}_w^n, \mathscr{C}_w^n, \mathscr{E}_w^n, \mathscr{G}_w^n\}$ *and any formula* $\varphi \in \Phi_{\mathrm{ML}}$ *such that* $(\mathscr{N}_w^n, \grave{x}_n) \models \varphi$ *and* $(\mathscr{N}_w^n, \acute{x}_n) \models \neg\varphi$, *the extended syntax tree of* φ *with root* $\{(\mathscr{N}_w^n, \grave{x}_n)\} \circ \{(\mathscr{N}_w^n, \acute{x}_n)\}$ *has a branch* B *such that* $I(B)$ *is a path over* $\mathscr{N}_w^n$.

Proof. The required branch B is constructed inductively. To present the main idea with sufficient precision, we assume that $n > 1$, and $\mathscr{N}_w^n \in \{\mathscr{A}_w^n, \mathscr{C}_w^n, \mathscr{G}_w^n\}$. The case $n = 1$ or $\mathscr{N}_w^n = \mathscr{E}_w^n$ is completely analogous modulo the fact that if $\mathscr{N}_w^n \in \{\mathscr{A}_w^n, \mathscr{C}_w^n, \mathscr{G}_w^n\}$, then for $i_k \in \{a, b\}$, the word w has the form $i_1 cdi_2 cd \dots cdi_n$, whereas if $\mathscr{N}_w^n = \mathscr{E}_w^n$ then $w = cdi_1 cdi_2 \dots cdi_n$.

During the construction, certain pointed models $(\mathscr{N}_w^n, p)$ occurring in the semantic label $seml(\eta)$ of a given node η will be declared **marked**. Let us suppose that we have a node η such that $seml(\eta)$ contains two marked pointed models, $(\mathscr{N}_w^n, p)$ on the left and $(\mathscr{N}_w^n, q)$ on the right.

- If η has a syntax label $\neg$, then, according to second item from Definition 3.22, its successor η_1 has a semantic label $seml(\eta_1)$ that contains $(\mathscr{N}_w^n, q)$ on the left and $(\mathscr{N}_w^n, p)$ on the right. We again declare $(\mathscr{N}_w^n, q)$ and $(\mathscr{N}_w^n, p)$ marked in $seml(\eta_1)$.

- If η has a syntax label $\vee$, then, according to third item from Definition 3.22, η has at least one successor η_1 with a semantic label $seml(\eta_1)$ that contains $(\mathscr{N}_w^n, p)$ on the left and $(\mathscr{N}_w^n, q)$ on the right. In this case, too, $(\mathscr{N}_w^n, p)$ and $(\mathscr{N}_w^n, q)$ remain marked in $seml(\eta_1)$.

Let us construct the desired branch B. Suppose that

$$w = i_1 cdi_2 cd \dots cdi_n, \text{ where } i_k \in \{a, b\}.$$

The first point η_0 of B is the root $\{(\mathscr{N}_w^n, \grave{x}_n)\} \circ \{(\mathscr{N}_w^n, \acute{x}_n)\}$ of T and both models $(\mathscr{N}_w^n, \grave{x}_n)$, and $(\mathscr{N}_w^n, \acute{x}_n)$ are marked.

If η_0 has a syntax label $\vee$ or $\neg$, the above considerations show that η_0 has at least one successor η such that $seml(\eta)$ contains $(\mathscr{N}_w^n, \acute{x}_n)$ and $(\mathscr{N}_w^n, \acute{x}_n)$ on opposite sides. Let η be the second node of B. Since these models satisfy the same propositional symbols, η cannot have a syntax label that is a propositional symbol. This means that it can have a syntax label $\vee$ or $\neg$. Generalising these considerations, the intuitive idea behind the construction of B is best explain as follows. We simply "follow" the marked models $(\mathscr{N}_w^n, \grave{x}_n)$, and $(\mathscr{N}_w^n, \acute{x}_n)$ starting from the root of the syntax tree through a possible number of nodes $\eta_0, \dots \eta_k$, where $k \geq 0$, that contain them. We add $\eta_0, \dots \eta_k$ to B. Since these models satisfy the same propositional symbols, η_k cannot have a syntax label that is a propositional symbol. Therefore, continuing in this way, we must encounter a node η_l such that $seml(\eta_l)$ contains $(\mathscr{N}_w^n, \grave{x}_n)$, and $(\mathscr{N}_w^n, \acute{x}_n)$ on opposite sides and has a syntax label $[i]$. Since $\mathscr{N}_w^n \in \{\mathscr{A}_w^n, \mathscr{C}_w^n, \mathscr{G}_w^n\}$, and $w = i_1 cdi_2 cd \dots cdi_n$ the second item from Proposition 5.4 or Proposition 5.6 or Proposition 5.10 respectively implies that $i = i_1$. We have to consider two cases.

- $(\mathscr{N}_w^n, \grave{x}_n)$ is on the left and $(\mathscr{N}_w^n, \acute{x}_n)$ is on the right;

- $(\mathscr{N}_w^n, \grave{x}_n)$ is on the right and $(\mathscr{N}_w^n, \acute{x}_n)$ is on the left.

For each one of these cases, we have two possibilities for the semantic label of the successor η_{l+1} of η_l according to the forth item from Definition 3.22. We consider these two possibilities for the first case. The second case follows by symmetry.

1. $seml(\eta_{l+1})$ contains $(\mathscr{N}_w^n, \acute{x}_n)$ on the right. This is possible because all the relations in $\mathscr{N}_w^n$, including i_1, are reflexive. It follows immediately that one of the models on the left is $(\mathscr{N}_w^n, \grave{x}_n)$. We declare these two models marked and add the node η_{l+1} to B;

2. $seml(\eta_{l+1})$ contains $(\mathscr{N}_w^n, \acute{y}_n)$ on the right. This is possible because $\acute{y}_n$ is an i_1-successor of $\acute{x}_n$. It follows that one of the chosen models on the left is $(\mathscr{N}_w^n, \grave{y}_n)$. These models are declared marked and the node η_{l+1} is added to B.

Since the marked models in $seml(\eta_{l+1})$ satisfy the same propositional symbols, η_{l+1} cannot have a syntax label that is a propositional symbol. Again, we "follow" the newly marked models to a node where a $[i]$-move was played.

In general,

- If $seml(\eta_l)$ contains marked models of the form $(\mathscr{N}_w^n, \grave{x}_j)$ on the left and $(\mathscr{N}_w^n, \acute{x}_j)$ on the right (or vice versa) and its syntax label is $[r]$, then, using the second item from one of Proposition 5.4 or Proposition 5.6 or Proposition 5.10, we see that $\grave{x}_j$ and $\acute{x}_j$ are not r-connected. Hence r is either d or $r \in \{a, b\}$ and $\grave{x}_j$ and $\grave{y}_j$ are r-connected. Therefore, there are not more than three possibilities for the semantic label of the successor node η_{l+1}.

    1. $seml(\eta_{l+1})$ contains $(\mathscr{N}_w^n, \grave{x}_j)$ on the left and $(\mathscr{N}_w^n, \acute{x}_j)$ on the right (or vice versa) and these models are marked;
    2. $seml(\eta_{l+1})$ contains $(\mathscr{N}_w^n, \grave{z}_{j+1})$ on the left and $(\mathscr{N}_w^n, \acute{z}_{j+1})$ on the right (or vice versa) and these models are marked;
    3. $seml(\eta_{l+1})$ contains $(\mathscr{N}_w^n, \grave{y}_j)$ on the left and $(\mathscr{N}_w^n, \acute{y}_j)$ on the right (or vice versa) and these models are marked.

- If the semantic label of a node η_l contains marked models of the form $(\mathscr{N}_w^n, \grave{y}_j)$ on the left and $(\mathscr{N}_w^n, \acute{y}_j)$ on the right (or vice versa) and its syntax label is $[r]$, then again the second item of one of Proposition 5.4 or Proposition 5.6 or Proposition 5.10 implies that $\grave{y}_j$ and $\acute{y}_j$ are not r-connected. Hence, r is either c or $r \in \{a, b\}$ and $\grave{y}_j$ is an r-successor of $\grave{x}_j$. As before, there are not more than three possibilities for the semantic label of the successor node η_{l+1}.

    1. $seml(\eta_{l+1})$ contains $(\mathscr{N}_w^n, \grave{y}_j)$ on the left and $(\mathscr{N}_w^n, \acute{y}_j)$ on the right (or vice versa) and these models are marked;
    2. $seml(\eta_{l+1})$ contains $(\mathscr{N}_w^n, \grave{z}_j)$ on the left and $(\mathscr{N}_w^n, \acute{z}_j)$ on the right (or vice versa) and these models are marked;

3. $seml(\eta_{l+1})$ contains $(\mathscr{N}_w^n, \grave{x}_j)$ on the left and $(\mathscr{N}_w^n, \acute{x}_j)$ on the right (or vice versa) and these models are marked.

- If the semantic label of a node η_l contains marked models of the form $(\mathscr{N}_w^n, \grave{z}_j)$ on the left and $(\mathscr{N}_w^n, \acute{z}_j)$ on the right (or vice versa) and its syntax label is $[r]$ then, yet again, the second item of one of Proposition 5.4 or Proposition 5.6 or Proposition 5.10 tells us that $\grave{z}_j$ and $\acute{z}_j$ are not r-connected. Therefore, $r \in \{c, d\}$ and there are not more than three possibilities for the semantic label of the successor node η_{l+1}.

1. $seml(\eta_{l+1})$ contains $(\mathscr{N}_w^n, \grave{z}_j)$ on the left and $(\mathscr{N}_w^n, \acute{z}_j)$ on the right (or vice versa) and these models are marked;
2. $seml(\eta_{l+1})$ contains $(\mathscr{N}_w^n, \grave{y}_j)$ on the left and $(\mathscr{N}_w^n, \acute{y}_j)$ on the right (or vice versa) and these models are marked;
3. $seml(\eta_{l+1})$ contains $(\mathscr{N}_w^n, \grave{x}_{j-1})$ on the left and $(\mathscr{N}_w^n, \acute{x}_{j-1})$ on the right (or vice versa) and these models are marked.

It is obvious that the semantic label of every node of B contains a marked pair of one of the following forms

- $(\mathscr{N}_w^n, \grave{x}_j)$ on the left and $(\mathscr{N}_w^n, \acute{x}_j)$ on the right or vice versa;
- $(\mathscr{N}_w^n, \grave{y}_j)$ on the left and $(\mathscr{N}_w^n, \acute{y}_j)$ on the right or vice versa;
- $(\mathscr{N}_w^n, \grave{z}_j)$ on the left and $(\mathscr{N}_w^n, \acute{z}_j)$ on the right or vice versa.

Given the construction of the models $\mathscr{N}_w^n$, we see that no node of the extended syntax tree can have a syntax label that is a propositional symbol if the semantic label of the node contains a marked pair of one of the forms above that is different from the pair $(\mathscr{N}_w^n, \grave{y}_1)$ on the left and $(\mathscr{N}_w^n, \acute{y}_1)$ on the right. Therefore, B ends with a node that has a semantic label of this form. The reader can easily check that $I(B)$ is a path over the model $\mathscr{N}_w^n$. This follows from the fact that every node with syntax label $[r]$ has a semantic label containing a marked model of one of the forms $(\mathscr{N}_w^n, \grave{x}_j)$ or $(\mathscr{N}_w^n, \grave{y}_j)$ or $(\mathscr{N}_w^n, \grave{z}_j)$. □

Since for any n, any of the sets $\mathbb{A}^n$, $\mathbb{C}^n$, $\mathbb{E}^n$, $\mathbb{G}^n$ contains 2^n different pointed models, one for each word w, the next lemma is crucial for the proof of Theorem 5.1.

Lemma 5.16. *For any two pointed models $(\mathscr{N}_w^n, \grave{x}_n)$ and $(\mathscr{N}_{\underline{w}}^n, \acute{x}_n)$, if $w \neq \underline{w}$, then for any two paths π over $\mathscr{N}_w^n$ and $\underline{\pi}$ over $\mathscr{N}_{\underline{w}}^n$, it is true that $\pi \neq \underline{\pi}$.*

The proof of Lemma 5.16 follows the steps below.

Step 1. We begin with the observation that for any model $\mathscr{N}_w^n$, the word w does not contain any sub-word of the form xx or xyx, where $x, y \in \{a, b, c, d\}$. Therefore, the canonical path π_w does not contain two successive indices π_k, π_{k+1} such that $\pi_k = \pi_{k+1}$ or three successive indices $\pi_{k-2}, \pi_{k-1}, \pi_k$, such that $\pi_{k-2} = \pi_k$. We show that π_w is the only path over $\mathscr{N}_w^n$ with this property.

Step 2. We formulate a rewriting rule as follows.

Let a path $\pi = \pi_1\pi_2\ldots\pi_n$ over $\mathscr{N}_w^n$ be given. Reading π from left to right, if a sub-string of the form xx or xyx is encountered, we replace it with x or xx respectively, and continue with the symbols following xx or xyx (if any). Having reached the end of π, we go back to the leftmost symbol of the newly obtained word and repeat the procedure. This algorithm terminates if no sub-string of the form xx or xyx is encountered.

It is obvious that this algorithm always terminates; moreover it has the following important properties.

1. If π is a path over $\mathscr{N}_w^n$, then replacing a sub-string of the form xx with x or xyx with xx in π results in a new path π_1 over $\mathscr{N}_w^n$;
2. The procedure terminates with a path π^* over $\mathscr{N}_w^n$ that does not contain any sub-string of the form xx or xyx. Therefore, using Step 1, we see that $\pi^* = \pi_w$.

Step 3. Suppose that there are two pointed models $(\mathscr{N}_w^n, \dot{x}_n)$ and $(\mathscr{N}_{\underline{w}}^n, \dot{x}_n)$ such that there are two paths π over $\mathscr{N}_w^n$ and $\underline{\pi}$ over $\mathscr{N}_w^n$ respectively, for which $\pi = \underline{\pi}$. We apply the rewriting rule to π and obtain π_w. Since $\underline{\pi}$ is equal to π, we see that $\pi_w = \pi_{\underline{w}}$. Hence $w = \underline{w}$.

Let us begin by formalising **Step 1.**

<u>Step 1.</u>

Proposition 5.17. *For any pointed model model $(\mathscr{N}_w^n, x_n)$, if $\pi_w = \pi_1\pi_2\ldots\pi_m$ and the extended canonical path is $u_1\pi_1u_2\ldots u_m\pi_m u_{m+1}$, then for any path $\underline{\pi} = \underline{\pi_1}\ldots\underline{\pi_p}$ over $\mathscr{N}_w^n$ the following is true. If $\pi_1 = \underline{\pi_1}$, $\pi_2 = \underline{\pi_2}$, $\ldots$, $\pi_i = \underline{\pi_i}$, where $1 \leq i \leq m$, then for any extended path $\underline{u_1}\ \underline{\pi_1}\ \underline{u_2}$, $\ldots$, $\underline{u_i}\ \underline{\pi_i}\ \underline{u_{i+1}}$, $\ldots$, $\underline{\pi_p}\ \underline{u_{p+1}}$ over $\underline{\pi}$, it is true that*

1. $u_1 = \underline{u_1}, \ldots, u_i = \underline{u_i}$;
2. $\underline{u_{i+1}} = u_i$ *or* $\underline{u_{i+1}} = u_{i+1}$.

In other words, if $\underline{\pi} = \underline{\pi_1}\ldots\underline{\pi_i}\ldots\underline{\pi_p}$ coincides with π_w on the first i relation steps, then any extended path $\underline{u_1}\ \underline{\pi_1}\ \underline{u_2}$, $\ldots$, $\underline{u_i}\ \underline{\pi_i}\ \underline{u_{i+1}}$, $\ldots$, $\underline{\pi_p}\ \underline{u_{p+1}}$ over $\underline{\pi}$ coincides with the extended canonical path $u_1\pi_1u_2\ldots u_i\pi_i u_{i+1}\ldots u_m\pi_m u_{m+1}$ on the points $u_1, \ldots u_i$ whereas $\underline{u_{i+1}}$ is either u_i or u_{i+1}.

Proof. The proof is by induction on $1 \leq i \leq m$. Again, we assume that $(\mathscr{N}_w^n, \dot{x}_n)$ is one of $(\mathscr{A}_w^n, \dot{x}_n)$, $(\mathscr{C}_w^n, \dot{x}_n)$ or $(\mathscr{G}_w^n, \dot{x}_n)$. The case $(\mathscr{E}_w^n, u_n)$ is analogous.

Base case. Suppose that $i = 1$, i.e., $\pi_1 = \underline{\pi_1}$. It follows from Definition 5.11 that

- $u_1 = \dot{x}_n = \underline{u_1}$;
- $\underline{u_1}$ is $\underline{\pi_1}$ related to $\underline{u_2}$ in $(\mathscr{N}_w^n, \dot{x}_n)$;

- $\underline{\pi_1}$ is such that $\grave{x}_n$ is $\underline{\pi_1}$ related to $\grave{y}_n$.

Given the shape of $(\mathscr{N}_w^n, \grave{x}_n)$, these items imply that $\underline{u_2} = \grave{x}_n$ or $\underline{u_2} = \grave{y}_n$. Hence, $\underline{u_2} = \grave{x}_n = u_1$ or $\underline{u_2} = \grave{y}_n = u_2$.

Induction step. Let us assume that the statement is true for some $1 \leq i < m$. The proof for $i+1$ is as follows.

Since $\pi_1 = \underline{\pi_1}$, $\pi_2 = \underline{\pi_2}$, …, $\pi_{i+1} = \underline{\pi_{i+1}}$, it is obvious that $\pi_1 = \underline{\pi_1}$, $\pi_2 = \underline{\pi_2}$, …, $\pi_i = \underline{\pi_i}$. The induction hypothesis implies that

- $u_1 = \underline{u_1}, \ldots, u_i = \underline{u_i}$;
- $\underline{u_{i+1}} = u_i$ or $\underline{u_{i+1}} = u_{i+1}$.

First, we show that the case $\underline{u_{i+1}} = u_i$ is impossible. Indeed, let us suppose otherwise. Then we have the following situation.

$$\begin{array}{ccccccccccc} u_1 & \pi_1 & u_2 & \ldots & u_i & \pi_i & u_{i+1} & \pi_{i+1} & u_{i+2} & \ldots \\ \shortparallel & \shortparallel & \shortparallel & \ldots & \shortparallel & \shortparallel & & \shortparallel & & \\ \underline{u_1} & \underline{\pi_1} & \underline{u_2} & \ldots & u_i & \underline{\pi_i} & u_i & \underline{\pi_{i+1}} & \underline{u_{i+2}} & \ldots \end{array}$$

We have the following possibilities for u_i.

- $u_i = \grave{x}_j$:
  - If $j = n$, then $i = 1$, i.e., $u_1 = u_i = \underline{u_i} = \grave{x}_n$. Definition 5.11 and the shape of our models imply that $\underline{\pi_i} = \pi_{i+1}$. However, $\underline{\pi_i} = \pi_i \neq \pi_{i+1} = \underline{\pi_{i+1}}$, and we arrive at a contradiction.
  - If $j < n$, then $\underline{\pi_{i-1}} = d$. Since $\underline{\pi_i} \neq \underline{\pi_{i-1}}$, Definition 5.11 implies that $\underline{\pi_i} \in \{a, b\}$ and u_i is $\underline{\pi_i}$-connected to $\grave{y}_j$. In the same way, $d = \underline{\pi_{i-1}} \neq \underline{\pi_{i+1}}$ implies that u_i is $\underline{\pi_{i+1}}$-connected to $\grave{y}_j$ Given the shape of $(\mathscr{N}_w^n, \grave{x}_n)$ and the fact that $\underline{\pi_i} \neq \underline{\pi_{i+1}}$, we arrive at a contradiction.
- $u_i = \grave{y}_j$, where $1 < j$:
  - In this case, Definition 5.11 implies that $\pi_i = c = \underline{\pi_i}$ and $\pi_{i-1} = \underline{\pi_{i-1}} \in \{a, b\}$ and u_i is π_{i-1}-connected to $\grave{x}_j$. Since $c = \pi_i \neq \pi_{i+1} = \underline{\pi_{i+1}}$, again using Definition 5.11, we obtain that $\underline{\pi_{i+1}} \in \{a, b\}$ and u_i is π_{i+1}-connected to $\grave{x}_j$. Given the shape of $(\mathscr{N}_w^n, \grave{x}_n)$ and the fact that $\underline{\pi_{i-1}} \neq \underline{\pi_{i+1}}$, we arrive at a contradiction.
- $u_i = \grave{z}_j$: In this case, Definition 5.11 implies that $\underline{\pi_{i-1}}, \underline{\pi_i}, \underline{\pi_{i+1}} \in \{c, d\}$ and, at the same time, we have that $\underline{\pi_{i-1}} \neq \underline{\pi_i}$, $\underline{\pi_{i-1}} \neq \underline{\pi_{i+1}}$, and $\underline{\pi_i} \neq \underline{\pi_{i+1}}$, which leads to a contradiction.

Therefore, the induction hypothesis implies

- $u_1 = \underline{u_1}, \ldots, u_i = \underline{u_i}$;

- $\underline{u_{i+1}} = u_{i+1}$.

Hence, we have that

$$\begin{array}{cccccccccc} u_1 & \pi_1 & u_2 & \dots & u_i & \pi_i & u_{i+1} & \pi_{i+1} & u_{i+2} & \dots \\ \shortparallel & \shortparallel & \shortparallel & \dots & \shortparallel & \shortparallel & \shortparallel & \shortparallel & & \\ \underline{u_1} & \underline{\pi_1} & \underline{u_2} & \dots & \underline{u_i} & \underline{\pi_i} & \underline{u_{i+1}} & \underline{\pi_{i+1}} & \underline{u_{i+2}} & \dots \end{array}$$

Let us consider $\underline{u_{i+2}}$. Again, we have the following possibilities.

- If $\underline{u_{i+1}} = u_{i+1} = \grave{x}_j$ then $\pi_{i+1} = \underline{\pi_{i+1}} \in \{a, b\}$ and $\underline{u_{i+1}}$ is $\underline{\pi_{i+1}}$-connected to $\grave{y}_j$. Hence, $\underline{u_{i+2}} = u_{i+1} = \grave{x}_j$ or $\underline{u_{i+2}} = u_{i+2} = \grave{y}_j$;
- If $\underline{u_{i+1}} = u_{i+1} = \grave{y}_j$, where $1 < j$, then $\underline{\pi_{i+1}} = \pi_{i+1} = c$ and $\underline{u_{i+1}}$ is $\underline{\pi_{i+1}}$-connected to $\grave{z}_j$. Therefore, $\underline{u_{i+2}} = u_{i+1} = \grave{y}_j$ or $\underline{u_{i+2}} = u_{i+2} = \grave{z}_j$
- If $\underline{u_{i+1}} = u_{i+1} = \grave{z}_j$, then $\underline{\pi_{i+1}} = \pi_{i+1} = d$ and $\underline{u_{i+1}}$ is d-connected to $\grave{x}_{j-1}$. Therefore, $\underline{u_{i+2}} = u_{i+1} = \grave{z}_j$ or $\underline{u_{i+2}} = u_{i+2} = \grave{x}_{j-1}$.

This completes the proof.

□

Proposition 5.18. *For any model $\mathscr{N}_w^n$, the canonical path π_w is the only path over $\mathscr{N}_w^n$ that does not contain a sub-string of the form xx or xyx.*

Proof. Suppose that there is another path $\underline{\pi} = \underline{\pi_1} \dots \underline{\pi_p}$ over $\mathscr{N}_w^n$, such that $\pi_w \neq \underline{\pi}$ and $\underline{\pi}$ does not contain a sub-string of the form xx or xyx. Since π_w is the shortest path, if $\pi_w \neq \underline{\pi}$, we must consider two cases.

1. π_w is a proper prefix of $\underline{\pi}$;
2. π_w is not a prefix of $\underline{\pi}$, i.e., there is an i such that $\pi_i \neq \underline{\pi_i}$.

Again, we assume that $(\mathscr{N}_w^n, \grave{x}_n)$ is one of $(\mathscr{A}_w^n, \grave{x}_n)$, $(\mathscr{C}_w^n, \grave{x}_n)$ or $(\mathscr{G}_w^n, \grave{x}_n)$. The case for $(\mathscr{E}_w^n, x_n)$ is similar.

Suppose that $\pi_w = \pi_1\pi_2 \dots \pi_{n+2(n-1)}$. Hence, the extended canonical path $ext(\pi_w)$ is

$$\begin{array}{rcccccccc} ext(\pi) = & u_1 & \pi_1 & u_2 & \dots & d & u_{n+2(n-1)} & \pi_{n+2(n-1)} & u_{n+2(n-1)+1}. \\ & \shortparallel & & \shortparallel & & & \shortparallel & & \shortparallel \\ & \grave{x}_n & & \grave{y}_n & & & \grave{x}_1 & & \grave{y}_1 \end{array}$$

If π_w is a proper prefix of $\underline{\pi}$, Proposition 5.17 implies that any extended path $ext(\underline{\pi})$ over $\underline{\pi}$ has one of the following forms

$$\begin{array}{rccccccccc} ext(\pi) = & u_1 & \pi_1 & u_2 & \dots & d & u_{n+2(n-1)} & \pi_{n+2(n-1)} & u_{n+2(n-1)+1} & \\ & \shortparallel & & \shortparallel & & & \shortparallel & & \shortparallel & \\ & \grave{x}_n & & \grave{y}_n & & & \grave{x}_1 & & \grave{y}_1 & \\ & \shortparallel & & \shortparallel & & & \shortparallel & & \shortparallel & \\ ext(\underline{\pi}) = & \underline{u_1} & \pi_1 & \underline{u_2} & \dots & d & \underline{u_{n+2(n-1)}} & \pi_{n+2(n-1)} & \underline{u_{n+2(n-1)+1}} & \dots \end{array}$$

or

$$\begin{array}{rccccccccc} ext(\pi) = & u_1 & \pi_1 & u_2 & \dots & d & u_{n+2(n-1)} & \pi_{n+2(n-1)} & u_{n+2(n-1)+1} \\ & \shortparallel & & \shortparallel & & & \shortparallel & & \shortparallel \\ & \grave{x}_n & & \grave{y}_n & & & \grave{x}_1 & & \grave{y}_1 \\ & \shortparallel & & \shortparallel & & & \shortparallel & & \\ ext(\underline{\pi}) = & \underline{u_1} & \pi_1 & \underline{u_2} & \dots & d & \underline{u_{n+2(n-1)}} & \pi_{n+2(n-1)} & \grave{x}_1 \dots \end{array}$$

In the first case, Definition 5.11 implies that $\underline{\pi_{n+2(n-1)+1}} = \pi_{n+2(n-1)}$ and, therefore, $\underline{\pi}$ contains a sub-string of the form xx. If the second case is true, then using again Definition 5.11, we see that either $\underline{\pi_{n+2(n-1)+1}} = \pi_{n+2(n-1)}$ or $\underline{\pi_{n+2(n-1)+1}} = d$. Hence, $\underline{\pi}$ contains either a sub-string of the form xx or xyx.

Let us suppose that π_w is not a proper prefix of $\underline{\pi}$. Therefore, there is an index j such that $\pi_j \neq \underline{\pi_j}$. Let i be the smallest such index. Note that it follows from Definition 5.11 that for any two paths π and $\underline{\pi}$ over $\mathscr{N}^n_w$, it is true that $\pi_1 = \underline{\pi_1}$, hence $1 < i$. As before, Proposition 5.17 implies that we have the following two possibilities.

$$\begin{array}{rcccccccc} ext(\pi) = & u_1 & \pi_1 & u_2 & \dots & u_{i-1} & \pi_{i-1} & u_i & \pi_i & \dots \\ & \shortparallel & \shortparallel & \shortparallel & & \shortparallel & \shortparallel & \shortparallel & \nparallel & \\ ext(\underline{\pi}) = & \underline{u_1} & \pi_1 & \underline{u_2} & \dots & \underline{u_{i-1}} & \underline{\pi_{i-1}} & \underline{u_i} & \underline{\pi_i} & \dots \end{array}$$

or

$$\begin{array}{rcccccccc} ext(\pi) = & u_1 & \pi_1 & u_2 & \dots & u_{i-1} & \pi_{i-1} & u_i & \pi_i & \dots \\ & \shortparallel & \shortparallel & \shortparallel & & \shortparallel & \shortparallel & & \nparallel & \\ ext(\underline{\pi}) = & \underline{u_1} & \pi_1 & \underline{u_2} & \dots & \underline{u_{i-1}} & \underline{\pi_{i-1}} & \underline{u_{i-1}} & \underline{\pi_i} & \dots \end{array}$$

If the first case is true then we have three possibilities for u_i.

($u_i = \grave{y}_j$), where $1 < j$: Definition 5.11 implies that $\pi_i = c$ and $\pi_{i-1}, \underline{\pi_{i-1}} \in \{a, b\}$ is such that $u_{i-1} = \grave{x}_j$ is an π_{i-1}-successor of $\grave{y}_j$. Since $\pi_i = c \neq \underline{\pi_i}$ and $\underline{\pi_{i-1}} = \pi_{i-1}$, again using Definition 5.11, we see that $\underline{\pi_i} = \underline{\pi_{i-1}}$ and, therefore, we have a sub-string of the form xx in $\underline{\pi}$.

($u_i = \grave{x}_j$), where $n > j$: Using reasoning similar to case above, we see that $\pi_i \in \{a, b\}$ and $u_{i+1} = \grave{y}_j$ is a π_i-successor of $u_i = \grave{x}_j$, whereas $\pi_{i-1} = \underline{\pi_{i-1}} = d$. Since $\pi_i \neq \underline{\pi_i}$, Definition 5.11, implies that $\underline{\pi_i} = \underline{\pi_{i-1}}$.

($u_i = \grave{z}_j$), where $1 < j < n$: According to Definition 5.11, $\pi_{i-1} = c$ and $\pi_i = d$. Since $\underline{\pi_i} \neq \pi_i = d$ and $\underline{\pi_{i-1}} = \pi_{i-1} = c$, again using Definition 5.11, we obtain that $\underline{\pi_i} = \underline{\pi_{i-1}} = c$.

If the second case is true, then, again we have three possibilities for u_i. We consider only the case $u_i = \grave{x}_j$, where $1 \leq j < n$. The other two cases are completely analogous.

Let us suppose that $u_i = \grave{x}_j$. Then $\underline{u_{i-1}} = \grave{z}_{j+1}$ and $\underline{\pi_{i-1}} = \pi_{i-1} = d$. According to Definition 5.11, we have two possibilities for $\underline{\pi_i}$, namely, $\underline{\pi_i} = d$ and in this case

$\underline{\pi_{i-1}} = \underline{\pi_i} = d$, hence, we have a sub-string of the form xx in $\underline{\pi}$ or $\pi_i = d = \pi_{i-2}$ and in this case, we have a sub-string of the form xyx in $\underline{\pi}$.

□

We proceed to verifying the property of the rewriting procedure described in the next step.

Step 2.

Proposition 5.19. *For any path π over $\mathscr{N}_w^n$, if π contains a sub-string of the form xx or xyx, then replacing this sub-string with x or xx respectively results in a new path π_1 over $\mathscr{N}_w^n$.*

Proof. Let us assume that π contains a sub-string xx. Then any extended path $ext(\pi)$ over π contains a sub-string $s = u_{i-1}\ x\ u_i\ x\ u_{i+1}$. Using Definition 5.11, we see that we have one of the 4 possibilities below.

$(u_{i-1} = u_i = u_{i+1})$: In this case s has the following shape, $s = u\ x\ u\ x\ u$. Therefore, we can replace s in $ext(\pi)$ with $s_1 = u\ x\ u$ and, thus, obtain a new path π_1 over $\mathscr{N}_w^n$ by replacing the string xx in π with x;

$(u_{i-1} = u_i)$: If this is true, then $s = u\ x\ u\ x\ u_{i+1}$. Hence s can be replaced with $s_1 = u\ x\ u_{i+1}$ in $ext(\pi)$ and, therefore, we obtain a new path π_1 over $\mathscr{N}_w^n$ by replacing the string xx in π with x;

$(u_{i-1} = u_{i+1})$: In this case s looks as follows $s = u\ x\ u_i\ x\ u$. This sub-string can be replaced in $ext(\pi)$ with $s_1 = u\ x\ u$, and we obtain again a new path over $\mathscr{N}_w^n$ by replacing the string xx in π with x;

$u_i = u_{i+1}$: In this situation we have that $s = u_{i-1}\ x\ u\ x\ u$ and, yet again, we can replace s in $ext(\pi)$ with $s_1 = u_{i-1}\ x\ u$ and obtain a new path over $\mathscr{N}_w^n$ by replacing the string xx in π with x.

If π contains a sub-string xyx, then, as in the previous case above, any extended path $ext(\pi)$ over π contains a sub-string of the form $s = u_i\ x\ u_{i+1}\ y\ u_{i+2}\ x\ u_{i+3}$. Again using Definition 5.11, we see that we have one of the 4 possibilities

$(u_i = u_{i+1} = u_{i+2} = u_{i+3})$: In this case s has the following shape, $s = u\ x\ u\ y\ u\ x\ u$ and, therefore, can be replaced in $ext(\pi)$ with $s_1 = u\ x\ u\ x\ u$ and, thus, we can obtain a new path π_1 over $\mathscr{N}_w^n$ by replacing the string xyx in π with xx;

$(u_i = u_{i+1} = u_{i+2})$: If this is true, then $s = u\ x\ u\ y\ u\ x\ u_{i+3}$. Hence s can be replaced in $ext(\pi)$ with $s_1 = u\ x\ u\ x\ u_{i+3}$ and we obtain a new path π_1 over $\mathscr{N}_w^n$ by replacing the string xyx in π with xx ;

$(u_{i+1} = u_{i+2} = u_{i+3})$: In this case s looks as follows $s = u_i\ x\ u\ y\ u\ x\ u$. This sub-string can be replaced in π with $s_1 = p\ x\ p\ x\ p$, and, therefore, we obtain again a new path π_1 over $\mathscr{H}_w^n$ by replacing the string xyx in π with xx;

$u_i = u_{i+3},\ u_{i+1} = u_{i+2}$**:** In this situation we have that $s = u_i\ x\ u_{i+1}\ y\ u_{i+1}\ x\ u_i$ and we can replace s with $s_1 = u_i\ x\ u_{i+1}\ x\ u_i$ in $ext(\pi)$ and obtain a new path π_1 over $\mathscr{N}_w^n$ by replacing the string xyx in π with xx.

This completes the proof of Proposition 5.19 and, therefore, Lemma 5.16 and the main theorem of this chapter, namely Theorem 5.1 □

Chapter 6

Formulae as Directed Acyclic Graphs

The study of lower bounds on formula-size is concerned with, among other things, making explicit (by providing succinctness results) the advantages offered by the different ways of representing formulae of given logics. In the previous chapters, we represented formulae as trees. Frank Wolter suggested (personal communication) that it would be interesting to study lower bounds on formula-size in the more general framework of representing formulae as directed acyclic graphs (*DAGs*). This makes the problem considerably more difficult and that is why we were able to obtain some results on the relative succinctness of two logics only when the formulae of one of them are represented as *DAGs* while the formulae of the other are still represented as syntax trees. We will be particularly interested in comparing the relative succinctness of *DAGs* for the formulae in Φ_{ML} and syntax trees for the logics $[\cup]$ML, $[\exists]$ML, $[\cap]$ML, and $[\varphi]$ML. In the present chapter, we will show that none of the sets of formulae we used to prove our succinctness results in Chapters 4 and 5 can be used to prove exponential lower bounds in this new setting; what is more, we will prove that there are sets of *DAGs* for Φ_{ML} that are exponentially more succinct than syntax trees for the formulae of the logics $[\cup]$ML and $[\exists]$ML.

We are not going to give a precise definition of a directed acyclic graph. The only fact about such graphs (see for example [47] p. 21) that we are going to use is that

> in every finite *DAG*, there are nodes with no predecessors, called roots, and nodes with no successors called leaves.

All the *DAGs* we are going to consider will have just one leaf.

As is well-known, with every modal formula $\varphi \in \Phi_{\mathrm{ML}}$, we can associate a *DAG* in which every vertex corresponds to one of the operators $\neg, \vee, [i]$ or to a propositional symbol. A *DAG* corresponding to a formula φ will be denoted $DAG(\varphi)$. Let $DAG(\Phi_{\mathrm{ML}})$ denote the set of ML-formulae represented as DAGs. As usual (see for example [30] p.5), the size of a directed acyclic graph G, denoted $||G||$, is defined as the number of edges in G.

Proposition 6.1. *For all formulae defined in Tables 4.1, 4.2, 5.1, and 5.2, there are equivalent on the relevant classes of models* ML*-formulae that can be represented as DAGs of linear size.*

Proof.
(I) We begin with the formulae from Tables 4.1 and 4.2 and their ML-equivalents which we give again for convenience.

- The formulae in the set $\Xi_{[\cup]}$ and their ML-equivalents on the class $\mathbf{K}$ are:

$$\xi_1 \stackrel{\text{def}}{=} \neg[\cup_{\{a,d\}}]\neg\mathbf{b} \text{ equivalent to } \xi'_1 \stackrel{\text{def}}{=} \langle a\rangle\mathbf{b} \vee \langle d\rangle\mathbf{b};$$
$$\xi_n \stackrel{\text{def}}{=} \neg \underbrace{[\cup_{\{a,d\}}]\dots[\cup_{\{a,d\}}]}_{n \text{ times}}\neg\mathbf{b} \text{ equivalent to } \xi'_n \stackrel{\text{def}}{=} \langle a\rangle\xi'_{n-1} \vee \langle d\rangle\xi'_{n-1} \text{ for } n > 1.$$

The linearly growing *DAGs* corresponding to the formulae ξ'_n are shown in the figures below.

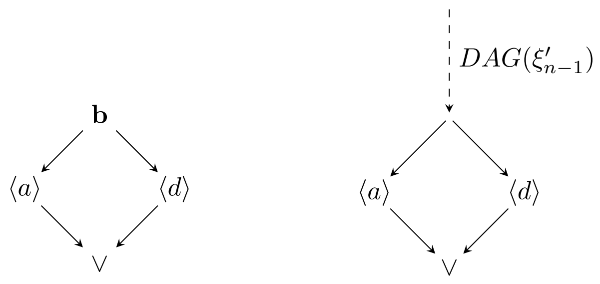


FIGURE 6.1: The $DAG(\xi'_1)$ (left) and the $DAG(\xi'_n)$ (right).

The $DAG(\xi'_n)$ is formed by taking the $DAG(\xi'_{n-1})$ (symbolised by the dashed arrow with label $DAG(\xi'_{n-1})$) and connecting its leaf to the nodes $\langle a\rangle$, $\langle d\rangle$, and $\vee$ as shown. For example, the $DAG(\xi'_2)$ looks as follows.

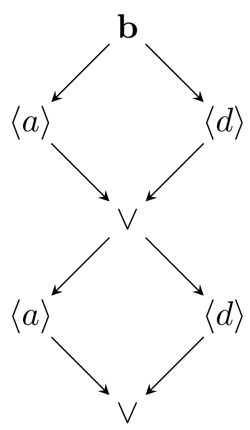


FIGURE 6.2: The $DAG(\xi'_2)$.

- The formulae in the set $\Delta_{[\varphi]}$ and their ML-equivalents on $\mathbf{K}$ are:

$\delta_1 \stackrel{\text{def}}{=} \langle a\rangle\mathbf{b} \vee \langle d\rangle\mathbf{b}$ equivalent to $\delta'_1 \stackrel{\text{def}}{=} \langle a\rangle\mathbf{b} \vee \langle d\rangle\mathbf{b}$;

$\delta_n \stackrel{\text{def}}{=} \langle\delta_{n-1}\rangle\delta_1$ equivalent to $\delta'_n \stackrel{\text{def}}{=} \delta'_{n-1} \wedge (\langle a\rangle(\mathbf{b} \wedge \delta'_{n-1}) \vee \langle d\rangle(\mathbf{b} \wedge \delta'_{n-1}))$ where $n > 1$.

The *DAGs* for the ML-formulae δ'_n are shown below.

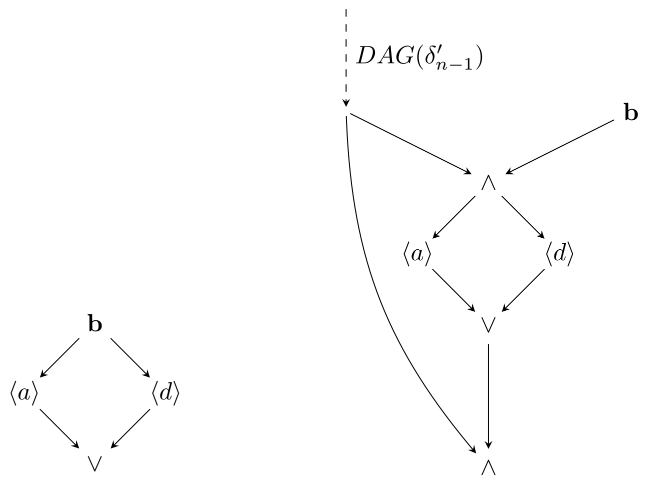


FIGURE 6.3: The $DAG(\delta'_1)$ (left) and the $DAG(\delta'_n)$ (right).

We continue with the formulae from Table 4.2.

- The formulae in the set $\Theta_{[\exists]}$ and their ML-equivalents on the class $\mathbf{K}$ are:

$\theta_1 \stackrel{\text{def}}{=} [\exists_{\{a,d\}}]\mathbf{b}$ equivalent to $\theta'_1 \stackrel{\text{def}}{=} [a]\mathbf{b} \vee [d]\mathbf{b}$;

$\theta_n \stackrel{\text{def}}{=} \underbrace{[\exists_{\{a,d\}}]\ldots[\exists_{\{a,d\}}]}_{n \text{ times}}\mathbf{b}$ equivalent to $\theta'_n \stackrel{\text{def}}{=} [a]\theta_{n-1} \vee [d]\theta_{n-1}$ where $n > 1$.

The linearly growing *DAGs* for these formulae are shown in Figure 6.4.

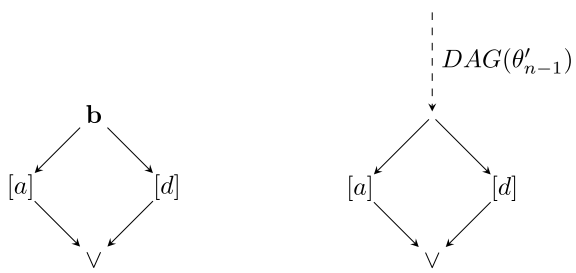


FIGURE 6.4: The $DAG(\theta'_1)$ (left) and the $DAG(\theta'_n)$ (right).

We would like to remind the reader that the formulae θ_n, and therefore the formulae θ'_n, are equivalent to the formulae ψ_n from the set of formulae $\Psi_{[\cap]}$ (see Table 4.2) on the set of models $\mathbb{O} \cup \mathbb{P}$ defined in Chapter 4. Therefore, Figures 6.4 and provide us with linearly growing ML-*DAGs* for the formulae ψ_n, too.

- The formulae in the set $\Omega_{[\varphi]}$ and their ML-equivalents on the class $\mathbf{K}$ are:

$\omega_1 \stackrel{\text{def}}{=} [a]\mathbf{b} \vee [d]\mathbf{b}$ equivalent to $\omega'_1 \stackrel{\text{def}}{=} [a]\mathbf{b} \vee [d]\mathbf{b}$;

$\omega_n \stackrel{\text{def}}{=} \langle \omega_{n-1} \rangle (\langle a \rangle \mathbf{b} \wedge \langle d \rangle \mathbf{b})$ equivalent to $\omega'_n \stackrel{\text{def}}{=} \omega'_{n-1} \wedge (\langle a \rangle (\mathbf{b} \wedge \omega'_{n-1}) \wedge \langle d \rangle (\mathbf{b} \wedge \omega'_{n-1}))$.

The *DAGs* for the ML-formulae ω'_n are shown below in Figure 6.5.

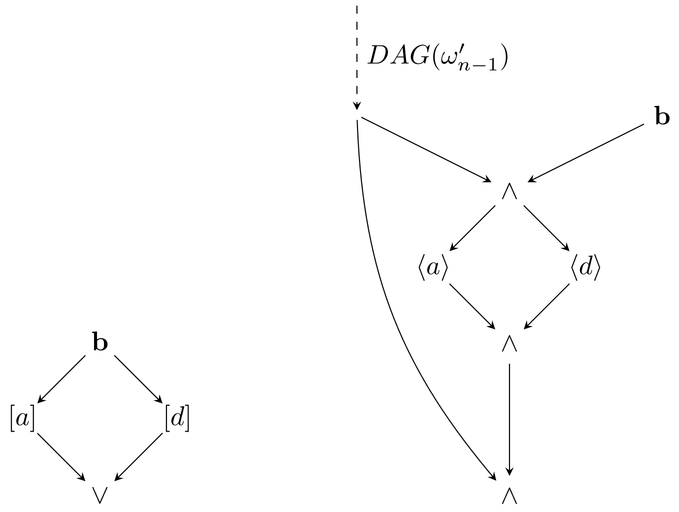


FIGURE 6.5: The $DAG(\omega'_1)$ (left) and the $DAG(\omega'_n)$ (right).

(II) We proceed with the formulae from Tables 5.1 and 5.2.

- The formulae in the set $\Delta_{[\cup]}$ and their ML-equivalents are:

$\delta_1 \stackrel{\text{def}}{=} \neg[\cup_{\{a,b\}}]\neg\mathbf{c}$ equivalent to $\delta'_1 \stackrel{\text{def}}{=} \langle a \rangle \mathbf{c} \vee \langle d \rangle \mathbf{c}$;

$\delta_n \stackrel{\text{def}}{=} \neg[\cup_{\{a,b\}}][c][d]\neg\delta_{n-1}$ equivalent to $\delta'_n \stackrel{\text{def}}{=} \langle a \rangle \langle c \rangle \langle d \rangle \delta_{n-1} \vee \langle b \rangle \langle c \rangle \langle d \rangle \delta_{n-1}$.

The *DAGs* corresponding to the formulae δ'_n are shown in Figure 6.9 below.

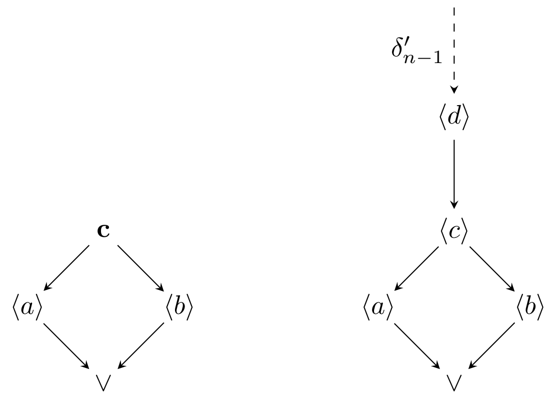


FIGURE 6.6: The $DAG(\delta'_1)$ (left) and the $DAG(\delta'_n)$ (right).

- The formulae in the set $\Sigma_{[\exists]}$ and their respective ML-equivalents are:

$\sigma_1 \stackrel{\text{def}}{=} [\exists_{\{a,b\}}]\mathbf{c}$ equivalent to $\sigma'_1 \stackrel{\text{def}}{=} [a]\mathbf{c} \vee [b]\mathbf{c}$;

$\sigma_n \stackrel{\text{def}}{=} [\exists_{\{a,b\}}](\mathbf{c} \wedge [c][d]\sigma'_{n-1})$ equivalent to $\sigma'_n \stackrel{\text{def}}{=} [a](\mathbf{c} \wedge [c][d]\sigma_j) \vee [b](\mathbf{c} \wedge [c][d]\sigma_j)$.

$DAG(\sigma_1)$ and DAG(σ_n) look as follows.

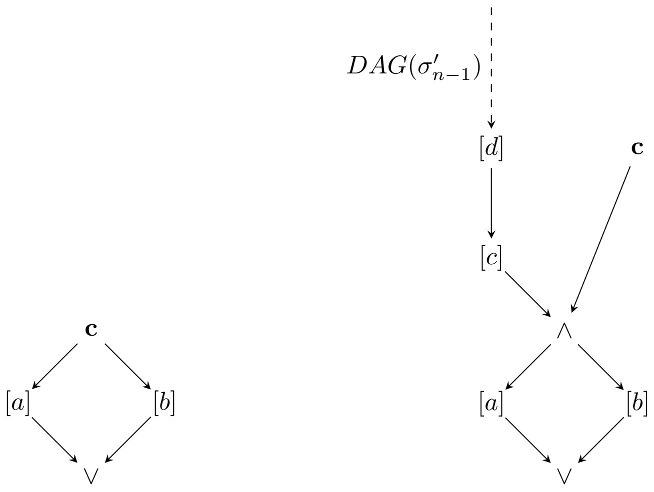


FIGURE 6.7: $DAG(\sigma_1)$ (left) and $DAG(\sigma_n)$ (right).

- The formulae in set $\Theta_{[\cap]}$ and the respective equivalent on the set of models $\mathbb{G}$ (see Definition 5.9) ML-formulae are:

$\theta_1 \stackrel{\text{def}}{=} \langle\cap_{\{a,b\}}\rangle\mathbf{c}$ equivalent to $\theta'_1 \stackrel{\text{def}}{=} \langle a\rangle\mathbf{c} \wedge \langle b\rangle\mathbf{c}$;

$\theta_n \stackrel{\text{def}}{=} \langle\cap_{\{a,b\}}\rangle\langle c\rangle\langle d\rangle\theta_{n-1}$ equivalent to $\theta'_n \stackrel{\text{def}}{=} \langle a\rangle\langle c\rangle\langle d\rangle\theta'_{n-1} \wedge \langle b\rangle\langle c\rangle\langle d\rangle\theta'_{n-1}$.

$DAG(\theta'_1)$ and DAG(θ'_n) are shown below.

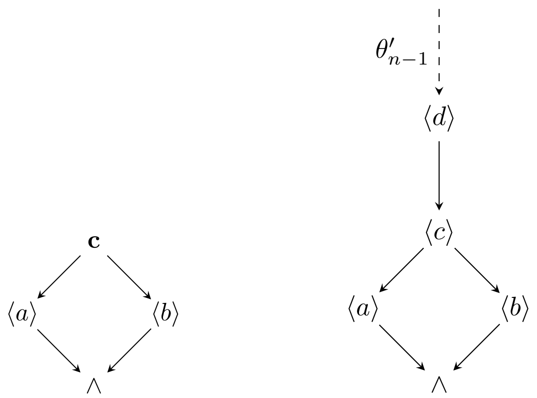


FIGURE 6.8: $DAG(\theta'_1)$ (left) and DAG(θ'_n) (right).

- The formulae in set $\Theta_{[\cap]}$ and the respective equivalent ML-formulae are:

$\omega_1 \stackrel{\text{def}}{=} \langle c\rangle(\mathbf{c} \wedge \langle d\rangle(\mathbf{d} \wedge (\langle a\rangle\mathbf{a} \vee \langle b\rangle\mathbf{b})))$ equivalent to $\omega'_1 \stackrel{\text{def}}{=} \langle c\rangle(\mathbf{c} \wedge \langle d\rangle(\mathbf{d} \wedge (\langle a\rangle\mathbf{a} \vee \langle b\rangle\mathbf{b})))$;

$\omega_n \stackrel{\text{def}}{=} \langle\omega_{n-1}\rangle\omega_1$ equivalent to the formula

$$\omega'_n \stackrel{\text{def}}{=} \omega_{n-1} \wedge \langle c\rangle(\omega_{n-1} \wedge \mathbf{c} \wedge \langle d\rangle(\omega_{n-1} \wedge \mathbf{d} \wedge (\langle a\rangle(\omega_{n-1} \wedge \mathbf{a}) \vee \langle b\rangle(\omega_{n-1} \wedge \mathbf{b})))).$$

The respective linearly growing *DAGs* are shown below.

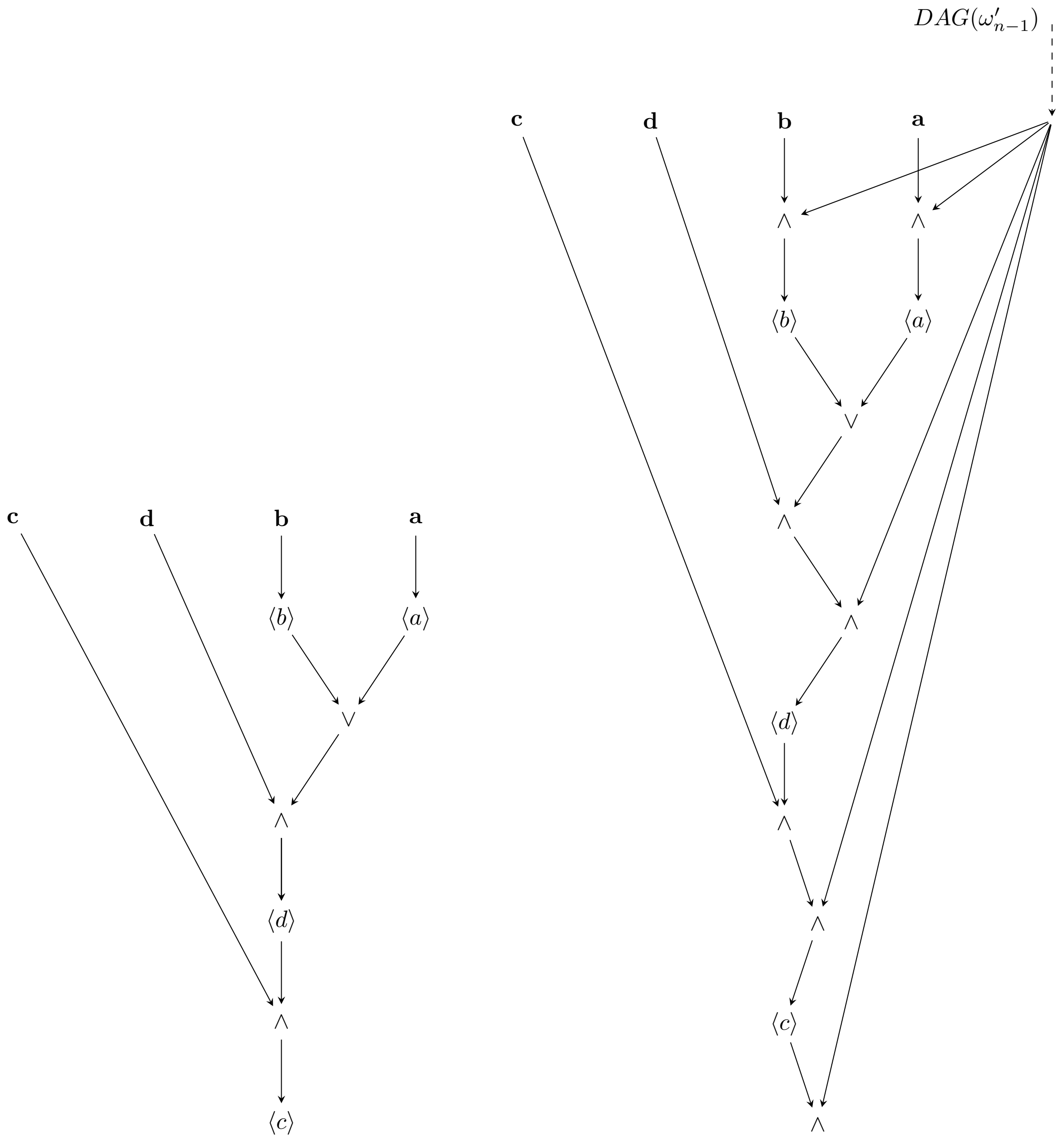


FIGURE 6.9: $DAG(\omega_1)$ (left) and $DAG(\omega_n)$ (right).

□

We can use Proposition 6.1 and our previous results from Chapters 4 and 5 to obtain the next theorem.

Theorem 6.2. *$DAG(\Phi_{\mathrm{ML}})$ is at least exponentially more succinct than Φ_{ML} on the class of models $\mathbf{S}_5$.*

Proof. On way of proving this statement is to consider the set of formulae $\Delta_{[\cup]}$ defined in Table 5.1 and the set of models $\mathbb{A}$ from Definition 5.3. According to Theorem 5.1, the tree-size of every $\psi_n \in \Phi_{\mathrm{ML}}$ that is equivalent to $\delta_n \in \Delta_{[\cup]}$ is at least 2^n. However, for every n, the size of the $DAG(\delta'_n) \in DAG(\Phi_{\mathrm{ML}})$ from Figure 6.6 is linear in n. □

Of course, using Proposition 3.8 and the above Theorem 6.2, we see that $DAG(\Phi_{\mathrm{ML}})$ is at least exponentially more succinct than Φ_{ML} on the class $\mathbf{K}$, too[1].

Intuitively, Proposition 6.1 says that the exponential shrinkage of space with respect to Φ_{ML} offered by the formulae from Tables 4.1, 4.2, 5.1, and 5.2 is lost if Φ_{ML} is replaced with $DAG(\Phi_{\mathrm{ML}})$ and we face the natural question whether the operators $[\cup_\Gamma]$, $[\exists_\Gamma]$, $[\cap_\Gamma]$, and $[\varphi]$ could still cause any space compression on some class of models but this time with respect to $DAG(\Phi_{\mathrm{ML}})$. As far as the first two operators are concerned, we show in Theorem 6.4 below that, in fact, $DAG(\Phi_{\mathrm{ML}})$ is at least exponentially more succinct on $\mathbf{K}$ than both $\Phi_{[\cup]\mathrm{ML}}$ and $\Phi_{[\exists]\mathrm{ML}}$. For this we need some preliminary work

When we defined the length $|\varphi|$ of a formula φ in Chapter 3 (see Definition 3.12), we did not take into account the size of the set Γ in formulae like $[\cup_\Gamma]\varphi$, $[\exists_\Gamma]\varphi$, and $[\cap_\Gamma]\varphi$. This was done mainly in order to connect smoothly the number of nodes in the syntax tree of φ and the size $|\varphi|$ (as defined in Definition 3.12) of φ. Of course, defining $|[\cup_{\{i_1,\ldots,i_n\}}]\varphi|$ as $n + |\varphi|$ instead of $1 + \varphi$ can increase the length of the formula considerably if n is a large number but note that, even if we had defined the length of the formulae in Tables 4.1, 4.2, 5.1, and 5.2 in this new way, it would still have been linear in their indices because, in all modal operators of the type $[\cup_\Gamma], [\exists_\Gamma]$, and $[\cap_\Gamma]$, the set Γ contains just two elements.

For the purposes of the next proposition however we are going to define the length of formulae of the form $[\cup_{\{i_1,\ldots,i_n\}}]\varphi$ and $[\exists_{\{i_1,\ldots,i_n\}}]\varphi$ as $n + |\varphi|$.

Proposition 6.3. *For every class of models $\mathbf{M} \subseteq \mathbf{K}$ and every formula $\varphi \in \Phi_{[\cup]\mathrm{ML}}$ (or $\varphi \in \Phi_{[\exists]\mathrm{ML}}$), there is an equivalent on $\mathbf{M}$ formula $\psi \in \Phi_{\mathrm{ML}}$ such that $||DAG(\psi)|| \leq 2 \times |\varphi|$.*

Proof. The (almost trivial) proof is by induction of the structure of φ. Let $\mathbf{M}$ be given.

Base case. If φ is a propositional symbol p, then $\psi \stackrel{\text{def}}{=} p$ and the statement is obvious since $DAG(p)$ consists of a single vertex (corresponding to p) and no edges.

Induction hypothesis. Let us assume that for $i \in \{1, 2\}$ and $\varphi_i \in \Phi_{[\cup]\mathrm{ML}}$ (or $\varphi_i \in \Phi_{[\exists]\mathrm{ML}}$), there are equivalent on $\mathbf{M}$ formulae $\psi_i \in \Phi_{\mathrm{ML}}$ such that $||DAG(\psi_i)|| \leq 2 \times |\varphi_i|$.

Induction step. We have to consider the following cases.

[1]This result can be proven indipendantly with the help of, e.g, the set of formulae $\Xi_{[\cup]}$ from Table 4.1.

($\varphi \stackrel{\text{def}}{=} \varphi_1 \vee \varphi_2$) Let $\psi \stackrel{\text{def}}{=} \psi_1 \vee \psi_2$, where, ψ_1 and ψ_2 are as described in the Induction hypothesis. $DAG(\psi)$ is shown in Figure 6.10. As in Proposition 6.1, $DAG(\psi_1 \vee \psi_2)$ is constructed by taking the directed acyclic graphs for ψ_1 and ψ_2 (represented by the dashed arrows with labels $DAG(\psi_i)$) and connecting the leaf of each $DAG(\psi_i)$ to the vertex $\vee$ as shown.

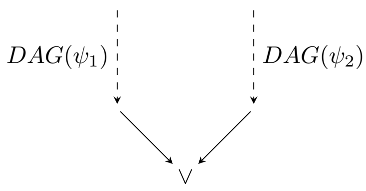


FIGURE 6.10: $DAG(\psi_1 \vee \psi_2)$.

We have

$$||DAG(\psi)|| = (||DAG(\psi_1)|| + ||DAG(\psi_2)|| + 2) \leq 2 \times (|\varphi_1| + |\varphi_2| + 1).$$

($\varphi \stackrel{\text{def}}{=} \neg\varphi_1$) Let $\psi \stackrel{\text{def}}{=} \neg\psi_1$, where, ψ_1 is as in the Induction hypothesis. $DAG(\psi)$ is shown in Figure 6.11.

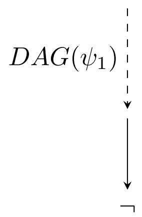


FIGURE 6.11: $DAG(\neg\psi_1)$.

Again, we have

$$||DAG(\psi)|| = (||DAG(\psi_1)|| + 1) < 2 \times (|\varphi_1| + 1).$$

($\varphi \stackrel{\text{def}}{=} [i]\varphi_1$) Let $\psi \stackrel{\text{def}}{=} [i]\psi_1$, where, ψ_1 is as described in the Induction hypothesis. The reasoning in this case is identical to the one above but with the use of $DAG(\psi)$ shown in Figure 6.12.

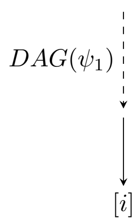


FIGURE 6.12: $DAG([i]\psi_1)$.

($\varphi \stackrel{\textbf{def}}{=} [\cup_{i_1,\ldots,i_n}]\varphi_1$) Let us suppose that $\psi_1 \in \Phi_{\mathrm{ML}}$ is as described in the Induction hypothesis and let $\psi = [i_1]\psi_1 \wedge [i_2]\psi_1 \wedge \ldots \wedge [i_n]\psi_1$. The directed acyclic graph $DAG(\psi)$ associated with ψ is shown in Figure 6.13 below.

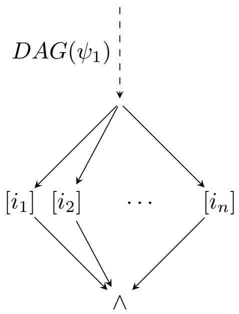


FIGURE 6.13: $DAG(\psi)$.

It is obvious that

$$||DAG(\psi)|| = (||DAG(\psi_1)|| + 2 \times n) \leq 2 \times (|\varphi_1| + n).$$

($\varphi \stackrel{\textbf{def}}{=} [\exists_{i_1,\ldots,i_n}]\varphi_1$) The reasoning in this case is absolutely identical to the one above modulo the fact that we set $\psi \stackrel{\text{def}}{=} [i_1]\psi_1 \vee [i_2]\psi_1 \vee \ldots \vee [i_n]\psi_1$ for some ψ_1 as in the Induction hypothesis and we use Figure 6.14 below.

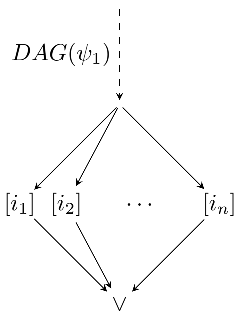


FIGURE 6.14: $DAG(\psi)$.

□

Theorem 6.4.

- $DAG(\Phi_{\mathrm{ML}})$ *is at least exponentially more succinct on* $\mathbf{K}$ *than* $\Phi_{[\cup]\mathrm{ML}}$*;*
- $DAG(\Phi_{\mathrm{ML}})$ *is at least exponentially more succinct than* $\Phi_{[\exists]\mathrm{ML}}$ *on* $\mathbf{K}$.

Proof. The first item is proven as follows. Let us consider the set of formulae $\Theta_{[\exists]}$ from Table 4.2. According to Theorem 4.6, we have that $\Theta_{[\exists]}$ is exponentially more succinct than $\Phi_{[\cup]\mathrm{ML}}$ on the set of models $\mathbb{O} \cup \mathbb{P} \subset \mathbf{K}$ from Definition 4.3. Since there are ML-formulae such that their DAG representation is linear in the length of the formulae in $\Theta_{[\exists]}$ (this follows from both Proposition 6.1 or the more general Proposition 6.3) the statement is immediate from Proposition 3.8.

The reasoning in the proof of the second item is identical. Indeed, let us consider the set of formulae $\Xi_{[\cup]}$ from Table 4.1. We know from Theorem 4.5 that $\Xi_{[\cup]}$ is exponentially more succinct than $\Phi_{[\exists]\mathrm{ML}}$ on the models $\mathbb{A} \cup \mathbb{B} \subset \mathbf{K}$ from Definition 4.1. Again, Proposition 6.1 or Proposition 6.3) imply that there are ML-formulae such that their *DAG* representation is linear in the length of the formulae in $\Xi_{[\cup]}$ and the statement follows from Proposition 3.8. □

Chapter 7

Conclusion and Open Problems

In this chapter we are going to summarise our results, list some open problems, and suggest some possible ways of generalising our work. We will be less formal than in the previous chapters and rely on intuitive explanations rather than on precise formulations.

7.1 Summary of the Results from Chapters 4 and 5 and Related Open Problem

Provided that the length of formulae of the logics we considered is identified with the number of nodes in their syntax trees as defined in Definition 3.19, our results can be summarised as shown in Figures 7.1 and 7.2. Both figures depict graphs in which the nodes correspond to the logics we studied while an arrow starting at one logic L_1 and pointing to another L_2 means that L_1 is exponentially more succinct than L_2 on the relevant class of models defined in Chapter 4 and Chapter 5.

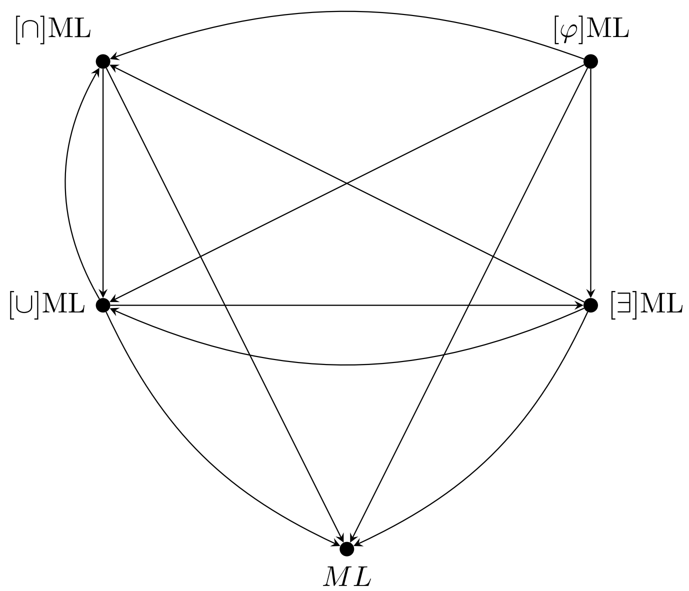


FIGURE 7.1: Summary of the succinctness results from Chapter 4.

It is obvious that the succinctness results in the case of $\mathbf{S}_5$-models are much weaker than the ones proven in Chapter 4. For example, we have that [∪]ML is exponentially more succinct than [∃]ML and vice versa on **K** but whether this is true on the class $\mathbf{S}_5$ is unknown.

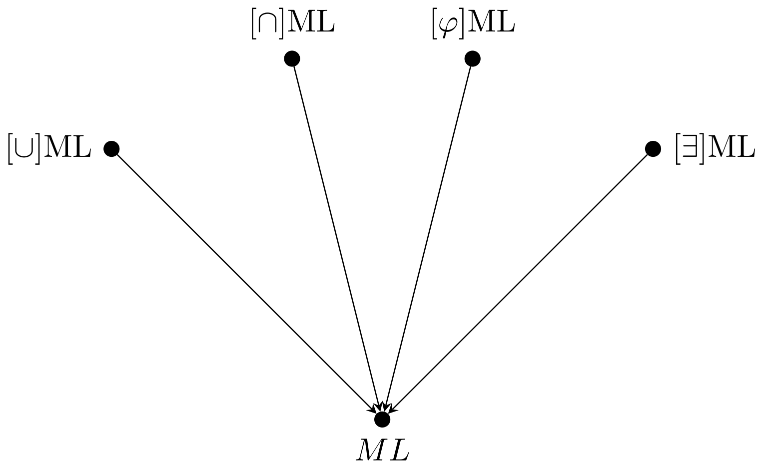


FIGURE 7.2: Summary of succinctness results from Chapter 5.

7.1.1 Open Problems

The open problems directly related to the work presented in Chapters 4 and 5 are summarised in Figures 7.3 and 7.4 respectively. Again, both figures, depict graphs where the nodes are the logics we studied but this time a dashed arrow starting from L_1 and pointing to L_2 means that it is an open problem whether L_1 is exponentially more succinct than L_2 on the relevant class of models.

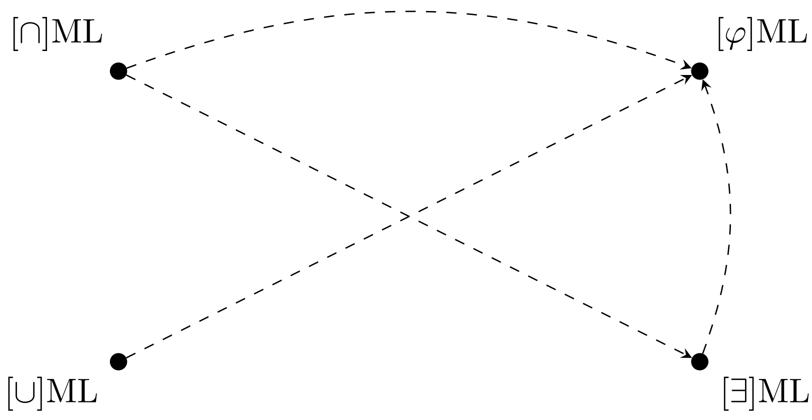


FIGURE 7.3: Summary of the open problems on the class of models **K**.

Of course, when the relative succinctness of the logic [∩]ML and another logic L is concerned, we are actually asking whether there is a class of models $\mathbb{M}$ that is a subset of **K** or $\mathbf{S}_5$ such that some semantic properties of the models in $\mathbb{M}$ are expressed

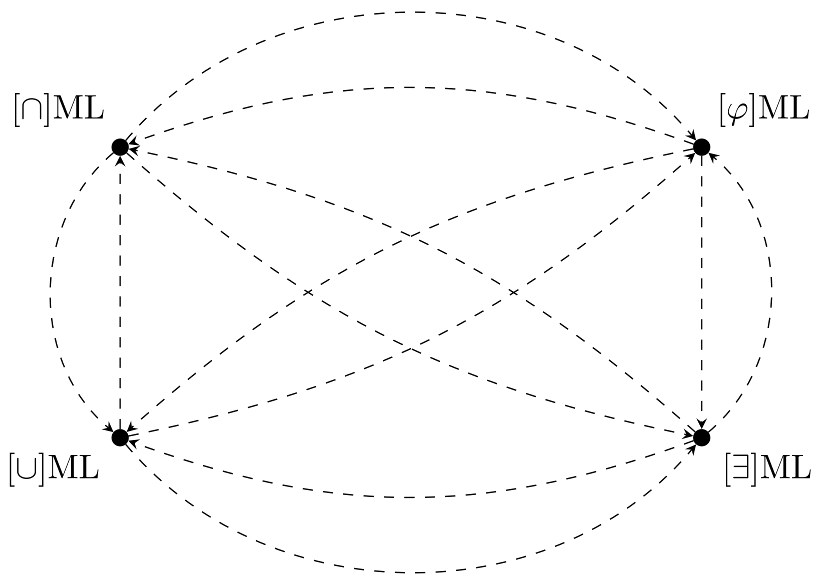


FIGURE 7.4: Summary of the open problems on the class of models $\mathbf{S}_5$.

exponentially more efficiently with formulae from [∩]ML (L) than with formulae from L ([∩]ML) while at the same time we assume that the properties in question are expressible in both L and [∩]ML.

We conjecture that [∪]ML is exponentially more succinct than [∃]ML and vice versa on $\mathbf{S}_5$-models. What is more, we believe that this result will probably require just a slight modification of the methods and the models used in Chapter 5. However, the question whether there is a class of models $\mathbb{M}$ on which at least one of the logics [∪]ML, [∃]ML or [∩]ML is exponentially more succinct than [φ]ML is considerably more difficult and we cannot offer any conjectures in this case. Intuitively, it is unclear whether there is some semantic property that can be more efficiently expressed in some of the former logics than in [φ]ML. If there is such a property, an argument based on extended syntax trees for [φ]ML seems to involve very difficult combinatorial considerations due to the semantics of the [φ] modality. Of course, it is perfectly conceivable that none of the logics [∪]ML, [∃]ML or [∩]ML is exponentially more succinct than [φ]ML on any class of Kripke models. In this case however, it is not very clear how an equivalence-preserving polynomial translation from the logics [∪]ML, [∃]ML, [∩]ML to [φ]ML can be obtained.

7.2 Summary of the Results from Chapter 7 and Some Open Problem

To the best of our knowledge, the fact that modal formulae represented as *DAGs* are at least exponentially more succinct than modal formulae in tree-representation seems to be taken for granted by many researchers (see, for example, [28] and the discussion there on p. 10) but we were unable to find a published proof despite considerable efforts. Nevertheless, we consider our Theorem 6.2 to be yet another confirmation of a widely-known folklore fact rather than an original new result. It is perhaps worth pointing out that the corresponding problem in the Boolean setting, namely, whether Boolean circuits are more succinct (and by how much) than Boolean formulae is still open and believed to require significant new mathematical ideas for its solution (for the connection between this question and a fundamental open problem in computational complexity see [49] p. 386). The results from Theorem 6.4, namely, that $DAG(\Phi_{\mathrm{ML}})$ is at least exponentially more succinct than both $\Phi_{[\cup]\mathrm{ML}}$ and $\Phi_{[\exists]\mathrm{ML}}$, seem to be new.

7.2.1 Open Problems

The material found in Chapter 7 is best viewed as the initial stage of a work in progress rather than a polished exposition of a solid body of results. There are many open questions here. One such family of problems is whether, if we consider the formulae of the logics $[\cup]$ML, $[\exists]$ML, $[\cap]$ML, $[\varphi]$ML as *DAGs*, our exponential succinctness results from Chapters 4 and 5 are still valid in this new setting. These questions seem difficult at present and we cannot offer any meaningful conjectures or possible ways of attacking them. Another question which might be easier to solve is whether $DAG(\Phi_{\mathrm{ML}})$ is at least exponentially more succinct than the tree representation of $[\varphi]$ML-formulae which we think is possible.

Bibliography

[1] M. Adler and N. Immerman, *An n! lower bound on formula size*, ACM Trans. Comput. Logic **4** (2003), no. 3, 296–314.

[2] A. Andreev, *On a method for obtaining more than quadratic effective lower bounds for the complexity of π-schemes*, Vestnik Moskovskogo Universiteta, Seriya 1 Matematika **1** (1987), 70–73, In Russian.

[3] S. Arora and B. Barak, *Computational complexity: A modern approach*, Cambridge University Press, 2009.

[4] F. Baader, D. Calvanese, D. McGuinness, D. Nardi, and P. Patel-Schneider (eds), *The description logic handbook: Theory, implementation, and applications*, Cambridge University Press, 2003.

[5] P. Blackburn, M. de Rijke, and Y. Venema, *Modal logic*, Cambridge University Press, 2001.

[6] M. Cadoli and F. Donini, *A survey on knowledge compilation*, AI Communications **10** (1998), 137–150.

[7] M. Cadoli, F. Donini, P. Liberatore, and M. Schaerf, *The size of a revised knowledge base*, Artificial Intelligence (1995), 151–162.

[8] ———, *The size of a revised knowledge base*, Artificial Intelligence **115** (1999), no. 1, 25–64.

[9] M. Cadoli, F. Donini, and M. Schaerf, *Is intractability of non-monotonic reasoning a real drawback*, Proceedings of the Twelfth National Conference on Artificial Intelligence (AAAI) (B. Hayes-Roth and R. Korf, eds.), vol. 2, 1994, pp. 946–951.

[10] ———, *On compact representation of propositional circumscription*, Proceedings of the Twelfth Symposium on Theoretical Aspects of Computer Science (STACS) (E. Mayr and C. Puech, eds.), 1995, pp. 205–216.

[11] J. Cai, M. Fürer, and N. Immerman, *An optimal lower bound on the number of variables for graph identification*, Combinatorica **12:4** (1992), 389–410.

[12] T. Cormen, C. Leiserson, R. Rivest, and C. Stein, *Introduction to algorithms 3 ed*, The MIT Press, 2009.

[13] A. Darwiche and P. Marquis, *A knowledge compilation map*, Journal of Artificial Intelligence Research **17** (2002), 229–264.

[14] A. Dawar, M. Grohe, S. Kreutzer, and N. Schweikardt, *Model theory makes formulas large*, Proceedings of the Thirty Fourth International Colloquium on Automata, Languages and Programming (ICALP 2007) (L. Arge, C. Cachin, and T. Jurdzinski, eds.), Springer, 2007, pp. 913–924.

[15] P. Dunne, *The complexity of Boolean networks*, Academic Press, 1988.

[16] A. Ehrenfeucht, *An application of games to the completeness problem for formalized theories*, Fundamenta Mathematicae **49** (1961), 129–141.

[17] K. Etessami, M. Vardi, and T. Wilke, *First-order logic with two variables and unary temporal logic*, Proceedings of the Twelfth Annual IEEE Symposium on Logic in Computer Science (LICS), 1997, pp. 228–236.

[18] R. Fagin, J. Halpern, Y. Moses, and M. Vardi, *Reasoning about knowledge*, The MIT Press, 1995.

[19] S. Figueira and D. Gorín, *On the size of shortest modal descriptions*, Advances in Modal Logic **8** (2010), 114–132.

[20] R. Fraïssé, *Sur quelques classifications des systémes de relations*, Publications Scientifiques de l'Université d'Alger, series A1, 1954, pp. 35–182.

[21] T. French, W. van der Hoek, P. Iliev, and B. Kooi, *Succinctness of epistemic languages*, Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI 2011) (T. Walsh, ed.), 2011, pp. 881–886.

[22] ______, *On the succinctness of some modal logics*, Artificial Intelligence **197** (2013), 56–85.

[23] M. Frick and M. Grohe, *The complexity of first-order and monadic second-order logic revisited*, Annals of Pure and Applied Logic **130** (2004), 3–31.

[24] G. Gargov, S. Passy, and T. Tinchev, *Modal environment for boolean speculations*, Proceedings of The Advanced International Summer School and Conference on Mathematical Logic and Its Applications (D. Skordev, ed.), 1987, pp. 480–490.

[25] G. Gogic, C. Papadimitriou, B. Selman, and H. Kautz, *The comparative linguistics of knowledge representation*, Proceedings of the Ninth International Joint Conference on Artificial Intelligence (IJCAI), 1995, pp. 862–869.

[26] M. Grohe and N. Schweikardt, *Comparing the succinctness of monadic query languages over finite trees*, Theoretical Informatics and Applications **38** (2004), no. 4, 343–373.

[27] ———, *The succinctness of first-order logic on linear orders*, Logical Methods in Computer Science **1** (2005), 1–25.

[28] J. Halpern and B. Kapron, *Zero-one laws for modal logic*, Annals of Pure and Applied Logic **69** (1994), no. 2, 157–193.

[29] I. Harel, D. Kozen, and J. Tiuryn, *Dynamic logic*, MIT Press, 200.

[30] J. Harris, J. Hirst, and M. Mossinghoff, *Combinatorics and graph theory*, Springer-Verlag, 2000.

[31] P. Hinman, *Fundamentals of mathematical logic*, A K Peters Ltd, 2005.

[32] J. Håstad, *The shrinkage exponent is 2*, SIAM Journal of Computing **27** (1998), 48–64.

[33] N. Immerman, *Length of predicate calculus formulas as a new complexity measure*, Proceedings of Twentieth Annual Symposium on Foundations of Computer Science (FOCS), 1979, pp. 337–347.

[34] ———, *Upper and lower bounds for first order expressibility*, Proceedings of Twenty First Annual Symposium on Foundations of Computer Science (FOCS), 1980, pp. 74–82.

[35] ———, *Descriptive complexity*, Springer, 1999.

[36] N. Immerman, J. Buss, and D. Barrington, *Number of variables is equivalent to space*, Journal of Symbolic Logic **66(3)** (2001), 1217–1230.

[37] N. Immerman and D. Kozen, *Definability with bounded number of bound variables*, Information and Computation **83** (1989), 121–139.

[38] S. Jukna, *Boolean function complexity*, Springer-Verlag Berlin Heidelberg, 2012.

[39] J. Lang and B. Zanuttini, *Knowledge-based programs as plans: Succinctness and the complexity of plan existence (extended abstract)*, Proceedings of Theoretical Aspects of Rationality and Knowledge (?, ed.), 2013, p. ?

[40] F. Laroussinie, N. Markey, and P. Schnoebelen, *Temporal logic with forgettable past*, Proceedings of Seventeenth Annual Symposium on Logic in Computer Science (LICS), 2002, pp. 383–392.

[41] H. Liu, C. Lutz, M. Milicic, and F. Wolter, *Foundations of instance level updates in expressive description logics*, Artificial Intelligence **175** (2011), no. 18, 2170–2197.

[42] O. Lupanov, *On the realization of functions of logical algebra by formulae of finite classes (formulae of limited depth) in the basis* $\bullet, +, -$, Problemi Kibernetiki **6** (1965), 1–14, In Russian.

[43] C. Lutz, *Complexity and succinctness of public announcement logic*, Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2006) (P. Stone and G. Weiss, eds.), 2006, Extended version at `http://lat.inf.tu-dresden.de/research/ reports.html`, pp. 137–144.

[44] C. Lutz, U. Sattler, and F. Wolter, *Modal logic and the two-variable fragment*, Proceedings of the Fifteenth Annual Conference of the European Association for Computer Science Logic (CSL) (L. Fribourg, ed.), 2001, pp. 247–261.

[45] N. Markey, *Temporal logic with past is exponentially more succinct*, Bulletin of the EATCS **79** (2003), 122–128.

[46] B. Nebel, *On the compilability and expressive power of propositional planning formalisms*, Journal of Artificial Intelligence Research **12** (2000), 271–315.

[47] R. Nigmatullin, *The complexity of Boolean functions*, Moskva Nauka, 1991, In Russian.

[48] M. Otto, *Bisimulation invariance and finite models*, Lecture Notes in Logic, Logic Colloquium 2002 (W. Pohlers Z. Chatzidakis, P. Koepke, ed.), 2006, pp. 276–298.

[49] C. Papadimitriou, *Computational complexity*, Addison-Wesley, 1993.

[50] J. Plaza, *Logics of public communications*, Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems: Poster Session Program (M.L. Emrich, M.S. Pfeifer, M. Hadzikadic, and Z.W. Ras, eds.), Oak Ridge National Laboratory, 1989, pp. 201–216.

[51] A. Rabinovich, *Succinctness gap between monadic logic and duration calculus*, Fundamenta Informaticae **42** (2000), no. 2, 201–210.

[52] A. Razborov, *On submodular complexity measures*, Boolean Function Complexity. London Math. Soc., Lecture Note Series 169 (M. Paterson, ed.), Cambridge University Press, 1992, pp. 76–83.

[53] J. Riordan and C. Shannon, *The number of two terminal series-parallel networks*, Journal of Mathematics and Physics **21** (1942), 83–93.

[54] B. Rossman, *Average-case complexity of detecting cliques*, Ph.D. thesis, Massachusetts Institute of Technology, 2010, Available at `http://www.mit.edu/ brossman/mit-thesis.pdf`.

[55] K. Segerberg, *A model existence theorem in infinitary propositional modal logic*, Journal of Philosophical Logic **23** (1994), 337–367.

[56] C. Shannon, *The synthesis of two-terminal switching circuits*, Bell System Technical Journal **28** (1948), 59–98.

[57] B. Subbotovskaya, *Realizations of linear functions by formulas using* $+$*,* $\times$*,* $-$, Dokladi Akademii Nauk SSSR **136** (1961), 535–555, In Russian; English translation in Soviet Mathematics Doklady 2 (1961) 110–112.

[58] B. ten Cate and M. Marx, *Navigational xpath: calculus and algebra*, SIGMOD Record, Database Principles Column **36** (2007), no. 2, 19–26.

[59] D. Vakarelov, *Modal definability in languages with a finite number of propositional variables and a new extension of the Sahlqvist's class*, Advances in Modal Logic **4** (2003), 499–518.

[60] J. van Benthem, *Modal logic and classical logic*, Bibliopolis, 1983.

[61] W. van der Hoek, P. Iliev, and B. Kooi, *On the relative succinctness of two extensions by definitions of multimodal logic.*, Proceedings of the Turing Centenary Conference and Eighth Conference on Computability in Europe (CiE 2012) (S. Barry Cooper, A. Dawar, and B. Löwe, eds.), 2012, pp. 323–333.

[62] H. van Ditmarsch, W. van der Hoek, and B. Kooi, *Dynamic epistemic logic*, Springer, 2007.

[63] ———, *Dynamic epistemic logic*, Springer, 2008.

[64] F. van Harmelen, V. Lifschitz, and B. Porter (eds.), *Handbook of knowledge representation*, Elsevier, 2008.

[65] I. Wegener, *The complexity of Boolean functions*, John Wiley and Sons Ltd and B. G. Teubner, Stuttgart, 1987, Online version available at `http://eccc.hpi-web.de/resources/pdf/cobf.pdf`.

[66] T. Wilke, *CTL^+ is exponentially more succinct than CTL*, Proceedings of Nineteenth Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS), 1999, pp. 110–121.

[67] S. Yablonskii, *Realisation of the linear function in the class of* π*-schemes*, Dokladi Akademii Nauk SSSR **94(5)** (1954), 805–806, In Russian.

[68] ———, *On algorithmic difficulties of synthesis of minimal contact schemes*, Problemy Kibernetiki **2** (1959), 75–121, In Russian.