Logical Methods in Computer Science
Volume 21, Issue 3, 2025, pp. 10:1–10:29
https://lmcs.episciences.org/

Submitted Aug. 22, 2024
Published Jul. 28, 2025

BOOLEAN BASIS, FORMULA SIZE, AND NUMBER OF MODAL OPERATORS

CHRISTOPH BERKHOLZ *, DIETRICH KUSKE, AND CHRISTIAN SCHWARZ

Technische Universität Ilmenau, Germany

Abstract. Is it possible to write significantly smaller formulae when using Boolean operators other than those of the De Morgan basis (and, or, not, and the constants)? For propositional logic, a negative answer was given by Pratt: formulae over one set of operators can always be translated into an equivalent formula over any other complete set of operators with only polynomial increase in size.

Surprisingly, for modal logic the picture is different: we show that elimination of bi-implication is only possible at the cost of an exponential number of occurrences of the modal operator $\Diamond$ and therefore of an exponential increase in formula size, i.e., the De Morgan basis and its extension with bi-implication differ in succinctness. Moreover, we prove that any complete set of Boolean operators agrees in succinctness with the De Morgan basis or with its extension with bi-implication. More precisely, these results are shown for the modal logic T (and therefore for K). We complement them by showing that the modal logic S5 behaves as propositional logic: the choice of Boolean operators has no significant impact on the size of formulae.

1. Introduction

Many classical logics such as propositional logic, first-order and second-order logic, temporal and modal logics incorporate a complete set G of Boolean operators in their definitions — often the De Morgan basis consisting of the set of operators $\mathrm{DMor} = \{\wedge, \vee, \neg, \top, \bot\}$. But there are certainly other options like $\{\rightarrow, \bot\}$ or $\{\mathrm{NAND}\}$. While for the expressivity it is clearly irrelevant which complete operator set is used, this choice may have an impact on how succinctly properties can be formulated. The main aim of this paper is to understand the influence of the set of operators on the succinctness of formulae.

Suppose we extend the De Morgan basis with an additional operator. If that additional operator defines a *read-once function*, i.e., it can be expressed in the De Morgan basis in such a way that every variable occurs at most once, then it can easily be eliminated without blowing-up the formula too much. Thus, read-once operators such as $x \rightarrow y \equiv \neg x \vee y$ are really just syntactic sugar. For operators that are not read-once, such as bi-implication $x \leftrightarrow y$ or the ternary majority operator $\mathrm{maj}(x, y, z)$, the situation is less clear, because mindlessly replacing them with any equivalent De Morgan formula may lead to an exponential explosion

Key words and phrases: succinctness, modal logic.

* Christoph Berkholz was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - project number 414325841.

DOI:10.46298/LMCS-21(3:10)2025

of the formula size. So can it be that such additional operators (or more generally, bases G other than the De Morgan basis) actually allow us to write exponentially more succinct formulae? For propositional logic, a negative answer was given by Pratt [Pra75]: for any two complete bases F and G, first balance the formula (that uses the basis F) so that it has logarithmic depth (this step may introduce operators from the De Morgan basis) and then replace all operators by any translation using the target basis G. This clearly leads to a linear increase in formula depth and therefore only to a polynomial increase in formula size.

Balancing a formula is, however, not possible for logics that contain quantifiers. For such logics it is still possible to efficiently replace certain operators that are not read-once by De Morgan formulae. We show that if an operator $\mathrm{op}(x_1, \ldots, x_k)$ is *locally read-once*, that is, has for every $i \in \{1, 2, \ldots, k\}$ an equivalent De Morgan formula in which x_i appears only once, then it can be efficiently eliminated. An example of an operator that is locally read-once, but not read-once, is $\mathrm{maj}(x, y, z)$ (see Example 3.5). The notion of a locally read-once operator and the algorithm of their replacement can be generalized to any complete basis in place of the De Morgan basis. Thus, our first result says that the succinctness does not differ much as long as operators of one basis are locally read-once in the other and vice versa.

We then give a decidable characterisation of the operators that are locally read-once in a given basis. This characterisation is based on the notion of locally monotone operators: an operator is *locally monotone* if fixing all but one argument defines a unary function that is increasing or decreasing *no matter how we fixed the remaining arguments* (e.g., bi-implication is not locally monotone, but ternary majority is). Then an operator is locally read-once in a given complete basis G if, and only if, it is locally monotone or some function from the basis G is not locally monotone.

As a result, for any complete basis G, there are just two possibilities: it allows to write formulae (up to a polynomial) as succinct as the De Morgan basis or as the De Morgan basis extended with the operator bi-implication $\leftrightarrow$.[1] In other words, there are at most two succinctness classes (recall that for propositional logic, there is just one such class).

So far, the techniques and results hold for many classical logics (but we spell them out in terms of modal logic). For modal logic, we proceed by showing that there are indeed two different succinctness classes. More precisely, we demonstrate that the use of operators that are not locally monotone can be avoided, but only at the cost of an exponential number of occurrences of the modal operator $\Diamond$ and therefore an exponential increase in formula size. Examples of such useful operators are bi-implication $x \leftrightarrow y$ and exclusive disjunction $x \,\mathrm{XOR}\, y$. In summary, there are exactly two succinctness classes that are exponentially separated: one containing standard modal logic and the other containing its extension with bi-implication.

Since this dichotomy is in contrast with propositional logic, where only one succinctness class exists, we also investigate what happens for fragments of modal logic defined by restrictions on the Kripke structures. Here we obtain the same dichotomy for structures with a reflexive accessibility relation. But upon considering equivalence relations only, we can show that the two succinctness classes collapse, as they do in propositional logic.

[1]In the conference version [BKS24] of this paper, this result was only shown for extensions G of the De Morgan basis.

Related work. It seems that this paper is the first to consider the influence of Boolean operators on the succinctness of modal logics. Other aspects have been studied in detail.

Pratt [Pra75] studied the effect of complete bases of binary operators on the size of propositional formulae and proved in particular that there are always polynomial translations. Wilke [Wil99] proved a succinctness gap between two branching time temporal logics, Adler and Immerman [AI03] developed a game-theoretic method and used it to improve Wilke's result and to show other succinctness gaps. The succinctness of further temporal logics was considered, e.g., in [EVW02, Mar03].

Lutz et al. [LSW01, Lut06] study the succinctness and complexity of several modal logics. French et al. [FvdHIK13] consider multi-modal logic with an abbreviation that allows to express "for all $i \in \Gamma$ and all i-successors, φ holds" where Γ is some set of modalities. Using Adler-Immerman-games, they prove (among other results in similar spirit) that this abbreviation allows exponentially more succinct formulae than plain multi-modal logic.

Grohe and Schweikardt [GS05] study the succinctness of first-order logic with a bounded number of variables and, for that purpose, develop extended syntax trees as an alternative view on Adler-Immerman-games. These extended syntax trees were used by van Ditmarsch et al. [DFH+14] to prove an exponential succinctness gap between a logic of contingency (public announcement logic, resp.) and modal logic.

Hella and Vilander [HV19] define a formula size game (modifying the Adler-Immerman-game) and use it to show that bisimulation invariant first-order logic is non-elementarily more succinct than modal logic.

Immerman [Imm81] defined separability games that characterise the number of quantifiers needed to express a certain property in first-order logic. These games were rediscovered and applied [FLRV21, FLVW22] and developed further (see [CFI$^+$24] where further applications can be found). Hella and Luosto [HL24] defined alternative and equivalent games. Vinall-Smeeth [Vin24] studied the interplay of the number of quantifiers and the number of variables needed to express certain properties.

Since the modal operator $\Diamond$ is a restricted form of quantification, our result on the number of occurrences needed to express certain properties is related to the works cited above.

2. Definitions

In this paper, $[n] = \{1, 2, \dots, n\}$ for all $n \geq 0$ and $0 \in \mathbb{N}$.

Boolean functions I. Let $\mathbb{B} = \{0, 1\}$ denote the Boolean domain. A *Boolean function* is a function $f \colon \mathbb{B}^n \to \mathbb{B}$ for some $n \geq 0$. We consider the following functions that are usually called disjunction, conjunction, implication, bi-implication, negation, falsum, verum, and majority:

$$\begin{aligned}
\vee \colon \mathbb{B}^2 \to \mathbb{B} \colon (a, b) &\mapsto \max(a, b)\\
\wedge \colon \mathbb{B}^2 \to \mathbb{B} \colon (a, b) &\mapsto \min(a, b)\\
\to \colon \mathbb{B}^2 \to \mathbb{B} \colon (a, b) &\mapsto \max(1 - a, b)\\
\leftrightarrow \colon \mathbb{B}^2 \to \mathbb{B} \colon (a, b) &\mapsto \max\bigl(\min(a, b), \min(1 - a, 1 - b)\bigr)\\
\neg \colon \mathbb{B}^1 \to \mathbb{B} \colon (a) &\mapsto 1 - a\\
\bot \colon \mathbb{B}^0 \to \mathbb{B} \colon () &\mapsto 0
\end{aligned}$$

$$\top\colon \mathbb{B}^0 \to \mathbb{B}\colon () \mapsto 1$$

$$\mathrm{maj}\colon \mathbb{B}^3 \to \mathbb{B}\colon (a,b,c) \mapsto \begin{cases} 1 & \text{if } a+b+c \geq 2 \\ 0 & \text{otherwise} \end{cases}$$

Propositional logic. Let $\mathcal{V} = \{p_i \mid i \geq 0\}$ be a countably infinite set of propositional variables. For a set F of Boolean functions, let the set of formulae of the propositional logic $\mathrm{PL}[F]$ be defined by

$$\varphi ::= \; p \mid f\langle\underbrace{\varphi,\ldots,\varphi}_{k \text{ times}}\rangle\,,$$

where p is some propositional variable and $f \in F$ is of arity k. Note that we do not allow $\bot$, $\top$, $\vee$, nor $\wedge$ in formulae unless they belong to F. An example of a $\mathrm{PL}[\{\neg,\vee\}]$-formula is $\neg\big\langle\vee\langle p_1, p_3\rangle\big\rangle$, which we usually write $\neg(p_1 \vee p_3)$. We also write $\bot$ for $\bot\langle\rangle$, but not in cases where we want to stress the distinction between the nullary function $\bot$ and the formula $\bot\langle\rangle$.

Let $\varphi \in \mathrm{PL}[F]$ and $y_1,\ldots,y_n$ be distinct propositional variables. We write $\varphi(y_1,\ldots,y_n)$ to emphasise that φ uses at most the variables $y_1,\ldots,y_n$. Let furthermore $\alpha_1,\ldots,\alpha_n \in \mathrm{PL}[F]$. Then $\varphi(\alpha_1,\ldots,\alpha_n)$ denotes the $\mathrm{PL}[F]$-formula obtained from φ by substituting the α_i for the y_i.

Let $\mathcal{I}$ be an *interpretation* of the variables, i.e., a map $\mathcal{I}\colon \mathcal{V} \to \mathbb{B}$. Inductively, we define the value $\mathcal{I}(\varphi) \in \mathbb{B}$ for formulae $\varphi \in \mathrm{PL}[F]$ via

$$\mathcal{I}\big(f\langle\varphi_1,\ldots,\varphi_k\rangle\big) = f\big(\mathcal{I}(\varphi_1),\ldots,\mathcal{I}(\varphi_k)\big)\,.$$

We call two formulae φ and ψ *equivalent* (denoted $\varphi \equiv \psi$) if, for all interpretations $\mathcal{I}$, we have $\mathcal{I}(\varphi) = \mathcal{I}(\psi)$. For instance, the formulae $x \vee y$ and $y \vee x$ are equivalent, but also the formulae $x \vee \neg x$ and $\top$ (for any propositional variables x and y).

For any sets of Boolean functions F and G, we have $\mathrm{PL}[F], \mathrm{PL}[G] \subseteq \mathrm{PL}[F \cup G]$, hence it makes sense to say that formulae from $\mathrm{PL}[F]$ are equivalent to formulae from $\mathrm{PL}[G]$.

Boolean functions II. Let $\varphi \in \mathrm{PL}[F]$ be a formula that uses, at most, the variables from $\{y_1,\ldots,y_n\}$. Then $\mathcal{I}(\varphi)$ depends on the values $\mathcal{I}(y_i)$ for $i \in [n]$, only (i.e., if $\mathcal{I}_1$ and $\mathcal{I}_2$ are interpretations with $\mathcal{I}_1(y_i) = \mathcal{I}_2(y_i)$ for all $i \in [n]$, then $\mathcal{I}_1(\varphi) = \mathcal{I}_2(\varphi)$). Hence the formula φ together with the sequence of variables $y_1,\ldots,y_n$ defines a function

$$[\![\varphi; y_1,\ldots,y_n]\!]\colon \mathbb{B}^n \to \mathbb{B}\colon (a_1,\ldots,a_n) \mapsto \mathcal{I}(\varphi)\,,$$

where $\mathcal{I}$ is any interpretation with $\mathcal{I}(y_i) = a_i$ for all $i \in [n]$.

Example 2.1. Suppose $\varphi, \psi \in \mathrm{PL}[F]$ are equivalent and use (at most) the variables $\{y_1,\ldots,y_n\}$. Then the functions $[\![\varphi; y_1,\ldots,y_n]\!]$ and $[\![\psi; y_1,\ldots,y_n]\!]$ are identical.

Now consider the equivalent formulae $\varphi = \top\langle\rangle$ and $\psi = (x \vee \neg x)$. By the above, $[\![\varphi; x]\!]$ and $[\![\psi; x]\!]$ are identical unary functions. But the nullary function $[\![\varphi;]\!]$ is defined while $[\![\psi;]\!]$ is undefined.

Let G be a set of Boolean functions. Then G is said to be *functionally complete* if for every Boolean function f of arity $k \geq 1$ there exists a $\mathrm{PL}[G]$-formula $\varphi(p_1,\ldots,p_k)$ such that $[\![\varphi; p_1,\ldots,p_k]\!] = f$. Standard examples of functionally complete sets of Boolean functions are the *De Morgan basis* $\mathrm{DMor} = \{\neg,\wedge,\vee,\top,\bot\}$ and the *extended De Morgan basis* $\mathrm{DMor}_\leftrightarrow = \mathrm{DMor} \cup \{\leftrightarrow\}$. In this paper, we use the following notion of completeness. A set G of Boolean functions is *complete* if for every $\mathrm{PL}[\mathrm{DMor}]$-formula there is an equivalent

PL[G]-formula (that may use more variables). Hence we consider DMor as the canonical complete set of Boolean functions and call G complete if PL[G] is equally expressive as PL[DMor]. It is not difficult to see that the two notions of completeness coincide:

- Assume that G is functionally complete and let $\psi(p_1,\dots,p_k) \in \mathrm{PL[DMor]}$. Then $f = [\![\psi; p_1,\dots,p_{k+1}]\!]$ is a Boolean function of arity $k+1 \geq 1$. Hence there is $\varphi(p_1,\dots,p_{k+1}) \in \mathrm{PL}[G]$ such that
$$\varphi(p_1,\dots,p_{k+1}) \equiv f\langle p_1,\dots,p_{k+1}\rangle \equiv \psi(p_1,\dots,p_k)\,,$$
i.e., G is complete.
- Conversely, assume that G is complete and let f be a Boolean function of arity $k \geq 1$. Since DMor is functionally complete and since G is complete, there are formulae $\psi(p_1,\dots,p_k) \in \mathrm{PL[DMor]}$ and $\varphi(p_1,\dots,p_\ell) \in \mathrm{PL}[G]$ with $\ell \geq k$ such that
$$f\langle p_1,\dots,p_k\rangle \equiv \psi(p_1,\dots,p_k) \equiv \varphi(p_1,\dots,p_\ell)\,.$$
Let $\varphi' = \varphi(p_1,\dots,p_k,p_1,\dots,p_1)$. Since $k \geq 1$, φ' uses at most the variables $p_1,\dots,p_k$. Furthermore, $\varphi' \equiv \varphi$ and therefore $[\![\varphi'; p_1,\dots,p_k]\!] = f$. This shows that G is functionally complete.

Modal logic. *Syntax.* For a set F of Boolean functions, let the set of formulae of the modal logic ML[F] be defined by
$$\varphi ::= \; p \;\big|\; f\langle\underbrace{\varphi,\dots,\varphi}_{k \text{ times}}\rangle \;\big|\; \Diamond\varphi\,,$$
where p is some propositional variable and $f \in F$ is of arity k. The *size* $|\varphi|$ of a formula φ is the number of nodes in its syntax tree.

Semantics. Formulae are interpreted over *pointed Kripke structures*, i.e., over tuples $S = (W, R, V, \iota)$, consisting of a set W of possible worlds, a binary accessibility relation $R \subseteq W \times W$, a valuation $V : \mathcal{V} \to \mathcal{P}(W)$, assigning to every propositional variable $p \in \mathcal{V}$ the set of worlds where p is declared to be true, and an initial world $\iota \in W$. The satisfaction relation $\models$ between a world w of S and an ML[F]-formula is defined inductively, where

- $S, w \models p$ if $w \in V(p)$,
- $S, w \models \Diamond\varphi$ if $S, w' \models \varphi$ for some $w' \in W$ with $(w, w') \in R$, and
- $S, w \models f\langle\alpha_1,\dots,\alpha_k\rangle$ if $f(b_1,\dots,b_k) = 1$ where, for all $i \in [k]$, $b_i = 1$ iff $S, w \models \alpha_i$.

A pointed Kripke structure S is a *model* of φ ($S \models \varphi$) if φ holds in its initial world, i.e., $S, \iota \models \varphi$.

Now let $\mathcal{S}$ be some class of pointed Kripke structures. A formula φ is *satisfiable in* $\mathcal{S}$ if it has a model in $\mathcal{S}$ and φ *holds in* $\mathcal{S}$ if every structure from $\mathcal{S}$ is a model of φ. The formula φ *entails* the formula ψ *in* $\mathcal{S}$ (written $\varphi \models_{\mathcal{S}} \psi$) if any model of φ from $\mathcal{S}$ is also a model of ψ; φ and ψ are *equivalent over* $\mathcal{S}$ (denoted $\varphi \equiv_{\mathcal{S}} \psi$) if $\varphi \models_{\mathcal{S}} \psi$ and $\psi \models_{\mathcal{S}} \varphi$.

Classes of Kripke structures. For different application areas (i.e., interpretations of the operator $\Diamond$), the following classes of Kripke structures have attracted particular interest. For convenience, we define them as classes of *pointed* Kripke structures.

- The class $\mathcal{S}_{\mathrm{K}}$ of all pointed Kripke structures.
- The class $\mathcal{S}_{\mathrm{T}}$ of all pointed Kripke structures with reflexive accessibility relation.
- The class $\mathcal{S}_{\mathrm{S5}}$ of all pointed Kripke structures where the accessibility relation is an equivalence relation.

Succinctness and translations. Suppose F and G are two sets of Boolean functions and G is complete. Then the logic ML[G] is at least as expressive as the logic ML[F], i.e., for any formula φ from ML[F], there exists an equivalent formula ψ from ML[G]. But what about the size of ψ? Intuitively, the logic ML[G] is at least as succinct as the logic ML[F] if the formula ψ is "not much larger" than the formula φ. This idea is formalized by the following definition.

Definition 2.2 (Translations)**.** Let F and G be sets of Boolean functions, $\mathcal{S}$ a class of pointed Kripke structures, and $\kappa\colon \mathbb{N} \to \mathbb{N}$ some function. Then ML[F] has *κ-translations wrt.* $\mathcal{S}$ in ML[G] if, for every formula $\varphi \in$ ML[F], there exists a formula $\psi \in$ ML[G] with $\varphi \equiv_{\mathcal{S}} \psi$ and $|\psi| \leq \kappa(|\varphi|)$.

The logic ML[F] has *polynomial translations wrt.* $\mathcal{S}$ in ML[G] if it has κ-translations wrt. $\mathcal{S}$ for some polynomial function κ. Finally, the logic ML[F] has sub-exponential translations wrt. $\mathcal{S}$ if it has κ-translations wrt. $\mathcal{S}$ for some function κ with $\lim_{n\to\infty} \frac{\log \kappa(n)}{n} = 0$.

We consider two logics ML[F] and ML[G] equally succinct if the former has polynomial translations in the latter and vice versa. It is easily seen that this notion "equally succinct" is an equivalence relation on the set of logics ML[F] for F a complete set of Boolean functions; we refer to the equivalence classes of this relation as *succinctness classes*. The following section will demonstrate that there are at most two such succinctness classes, namely those containing ML[DMor] and ML[$\text{DMor}_{\leftrightarrow}$], respectively.[2]

3. "All" logics have at most two succinctness classes

The aim of this section is to show that, for any finite and complete set of Boolean functions F, the logic ML[F] has the same succinctness as the logic ML[DMor] or as the logic ML[$\text{DMor}_{\leftrightarrow}$] (Corollary 3.3). Formally, one has to be more precise since the relation "equally succinct" depends on the class of pointed Kripke structures used to define the equivalence of formulae. In this section, we consider the largest such class, i.e., the class $\mathcal{S}_{\text{K}}$ of all pointed Kripke structures. For notational convenience, we will regularly omit the explicit reference to the class $\mathcal{S}_{\text{K}}$, e.g., "equivalent" means "equivalent over $\mathcal{S}_{\text{K}}$", $\varphi \models \psi$ means $\varphi \models_{\mathcal{S}_{\text{K}}} \psi$, and "$\kappa$-translations" means "$\kappa$-translations wrt. $\mathcal{S}_{\text{K}}$".

As to whether the logic ML[F] is in the succinctness class of ML[DMor] or of ML[$\text{DMor}_{\leftrightarrow}$] depends on whether all functions from F are locally monotone:

Definition 3.1 (Local monotonicity)**.** Let $f\colon \mathbb{B}^k \to \mathbb{B}$ be a Boolean function. We say that f is *monotone in the i-th argument* if

- for all $\overline{a} \in \mathbb{B}^{i-1}$ and $\overline{b} \in \mathbb{B}^{k-i}$, $f(\overline{a}, 0, \overline{b}) \leq f(\overline{a}, 1, \overline{b})$ or
- for all $\overline{a} \in \mathbb{B}^{i-1}$ and $\overline{b} \in \mathbb{B}^{k-i}$, $f(\overline{a}, 0, \overline{b}) \geq f(\overline{a}, 1, \overline{b})$.

The function f is *locally monotone* if it is monotone in every argument $i \in [k]$.

Hence f is monotone in the i-th argument, if, when changing the i-th argument from 0 to 1, while keeping the remaining ones fixed, the value of f uniformly increases or decreases (where, in both cases, the value may also remain unchanged). By this definition, conjunction, disjunction, negation, implication, as well as majority are locally monotone functions, while bi-implication is not.

The main result of this section is a consequence of the following theorem.

[2]Recall that $\text{DMor} = \{\neg, \vee, \wedge, \top, \bot\}$ and $\text{DMor}_{\leftrightarrow} = \{\neg, \vee, \wedge, \leftrightarrow, \top, \bot\}$ denote the (extended) De Morgan basis.

Theorem 3.2. *Let F and G be finite sets of Boolean functions such that G is complete. If all functions in F are locally monotone or some function in G is not locally monotone, then* $\mathrm{ML}[F]$ *has polynomial translations in* $\mathrm{ML}[G]$.

The proof of the theorem can be found at the end of this section, in the final Subsection 3.4. With Theorem 3.2, we can already establish that the class of logics $\mathrm{ML}[G]$ with G finite and complete has at most two succinctness classes. Recall that we consider two logics equally succinct if the former has polynomial translations in the later and vice versa.

Corollary 3.3. *Let G be some finite and complete set of Boolean functions. Then*

- $\mathrm{ML}[G]$ *and* $\mathrm{ML}[\mathrm{DMor}]$ *are equally succinct wrt.* $\mathcal{S}_{\mathrm{K}}$*, or*
- $\mathrm{ML}[G]$ *and* $\mathrm{ML}[\mathrm{DMor}_{\leftrightarrow}]$ *are equally succinct wrt.* $\mathcal{S}_{\mathrm{K}}$.

Proof. First recall that the De Morgan basis DMor and the extended De Morgan basis $\mathrm{DMor}_{\leftrightarrow} = \mathrm{DMor} \cup \{\leftrightarrow\}$ are complete. The above theorem implies the following:

- Suppose that all functions from G are locally monotone. Then $\mathrm{ML}[G]$ has polynomial translations in $\mathrm{ML}[\mathrm{DMor}]$. But also all functions from DMor are locally monotone; hence also $\mathrm{ML}[\mathrm{DMor}]$ has polynomial translations in $\mathrm{ML}[G]$. In other words, $\mathrm{ML}[G]$ and $\mathrm{ML}[\mathrm{DMor}]$ are equally succinct wrt. $\mathcal{S}_{\mathrm{K}}$ in this case.
- Suppose that some function from G is not locally monotone. Then the logic $\mathrm{ML}[\mathrm{DMor}_{\leftrightarrow}]$ has polynomial translations in $\mathrm{ML}[G]$. The extended De Morgan basis $\mathrm{DMor}_{\leftrightarrow}$ contains the non-locally monotone function $\leftrightarrow$; hence also $\mathrm{ML}[G]$ has polynomial translations in $\mathrm{ML}[\mathrm{DMor}_{\leftrightarrow}]$. In other words, $\mathrm{ML}[G]$ and $\mathrm{ML}[\mathrm{DMor}_{\leftrightarrow}]$ are equally succinct wrt. $\mathcal{S}_{\mathrm{K}}$ in this case. □

The remainder of this section is dedicated to the proof of Theorem 3.2, which can be divided into the following three steps.

Step 1 (cf. Section 3.1). We first show that $\mathrm{ML}[F]$ has polynomial translations in $\mathrm{ML}[G]$ if the set F of Boolean functions admits "$\mathrm{PL}[G]$-representations" (to be defined next). The idea is as follows:

Since G is complete and therefore also functionally complete, there is for every function $f \in F$ of arity $k \geq 1$ some formula $\omega(p_1, \ldots, p_k) \in \mathrm{PL}[G]$ such that $\omega(p_1, \ldots, p_k) \equiv f\langle p_1, \ldots, p_k\rangle$. Consequently, in order to translate a formula $\varphi \in \mathrm{ML}[F]$ into an equivalent formula $\psi \in \mathrm{ML}[G]$, we only need to replace every sub-formula $f\langle \alpha_1, \ldots, \alpha_k\rangle$ in φ by $\omega(\beta_1, \ldots, \beta_k)$ where β_i is the translation of α_i for $i \in [k]$.[3] In general, this translation leads to an exponential size increase. But if, in the formula ω, every variable p_i appears only once, we obtain a linear translation. The notion of representations is somewhat half-way between these two extremes.

Definition 3.4 (Representations)**.** Let G be a set of Boolean functions, f a Boolean function of arity $k > 0$, and $i \in [k]$.

A $\mathrm{PL}[G]$*-representation of* (f, i) is a formula $\omega_i(p_1, \ldots, p_k) \in \mathrm{PL}[G]$ that is equivalent to the formula $f\langle p_1, \ldots, p_k\rangle \in \mathrm{PL}[\{f\}]$ and uses the variable p_i at most once.

A set F of Boolean functions *has* $\mathrm{PL}[G]$*-representations* if there are $\mathrm{PL}[G]$-representations for all $f \in F$ of arity $k > 0$ and all $i \in [k]$.

[3] There is a small issue here if F contains a nullary function but G does not. However, we ignore this for now by assuming that G contains both $\top$ and $\bot$ and handle the problem after Step 3.3 in a natural manner.

Example 3.5. Consider the majority function $\text{maj}(p_1, p_2, p_3)$ that is true iff at least two arguments are true. Then $(\text{maj}, 1)$ has the $\text{PL}[\{\wedge, \vee\}]$-representation $\big(p_1 \wedge (p_2 \vee p_3)\big) \vee (p_2 \wedge p_3)$. Using the symmetry of maj, it follows that $\{\text{maj}\}$ has $\text{PL}[\{\wedge, \vee\}]$-representations.

Next, consider bi-implication $\leftrightarrow$. Recall that a $\text{PL}[\{\neg, \wedge, \vee\}]$-formula $\psi(p_1, p_2)$ that contains p_1 only under an even number of negations is monotonely increasing in p_1; analogously, if p_1 occurs only under an odd number of negations, then ψ is monotonely decreasing in p_1. Aiming at a contradiction, assume that $(\leftrightarrow, 1)$ had a $\text{PL}[\{\neg, \wedge, \vee\}]$-representation $\omega(p_1, p_2)$ that mentions p_1 only once (say, under an even number of negations). Then, by the previous observation, flipping p_1 from 0 to 1 does not decrease the truth value of the formula, hence

$$1 = (0 \leftrightarrow 0) = [\![\omega; p_1, p_2]\!](0, 0) \leq [\![\omega; p_1, p_2]\!](1, 0) = (1 \leftrightarrow 0) = 0\,,$$

a contradiction. Thus $\{\leftrightarrow\}$ does not have $\text{PL}[\{\neg, \wedge, \vee\}]$-representations.

Step 2 (cf. Section 3.2). Having established in step 1 that representations yield polynomial translations, we next show that any Boolean function has $\text{PL}[\text{DMor}_\leftrightarrow]$-representations and that any locally monotone function even has PL[DMor]-representations.

Step 3 (cf. Section 3.3). In this final step, we construct $\text{PL}[G]$-representations of all functions from DMor and $\text{DMor}_\leftrightarrow$ provided G is complete (and contains some non-locally monotone function for the case $\text{DMor}_\leftrightarrow$).

Summary (cf. Section 3.4). Now suppose that G is complete and that all functions from F are locally monotone or some function from G is not locally monotone. In the first case, F and DMor have PL[DMor]- and $\text{PL}[G]$-representations, respectively (step 2 and 3). Consequently, $\text{ML}[F]$ and ML[DMor] have polynomial translations in ML[DMor] and $\text{ML}[G]$, respectively (step 1). By transitivity, $\text{ML}[F]$ has polynomial translations in $\text{ML}[G]$.

In the second case (i.e., G contains some function that is not locally monotone), we can argue similarly but using $\text{DMor}_\leftrightarrow$ in place of DMor.

3.1. Representations yield polynomial translations.

Let F and G be two sets of Boolean functions such that F has $\text{PL}[G]$-representations. We will construct from a formula in $\text{ML}[F]$ an equivalent formula in $\text{ML}[G]$ of polynomial size. Since this will be done inductively, we will have to deal with formulae from $\text{ML}[F \cup G]$ and the task then is better described as elimination of functions $f \in F \setminus G$ from formulae in $\text{ML}[F \cup G]$.

Before we present the details of our construction, we briefly demonstrate the main idea behind the proof (for $F = \{\wedge, \vee, f\}$ and $G = \{\wedge, \vee, \neg\}$, i.e., we aim to eliminate some Boolean function f). The main results (Lemma 3.9 and Proposition 3.10) will appear at the end of the section.

Assume that f is of arity 2 and consider the two formulae

$$\varphi = \big(p_1 \vee f\big\langle\ f\langle p_2, p_3\rangle,\ p_4 \wedge p_4\ \big\rangle\big) \vee f\langle p_6, p_7 \wedge p_8\rangle \ \text{ and}$$
$$\psi = f\big\langle\ p_1 \vee f\big\langle\ f\langle p_2, p_3\rangle,\ p_4 \wedge p_5\ \big\rangle,\ f\langle p_6, p_7 \wedge p_8\rangle\ \big\rangle\,,$$

whose syntax trees are depicted in Fig. 1 (they only differ in the root node).

A distinguishing property of the left tree is that there is no f-node with left immediate successor ℓ and right immediate successor r such that the sub-trees with roots ℓ and r both contain an f-node. Assuming that f has $\text{PL}[G]$-representations, there exist Boolean

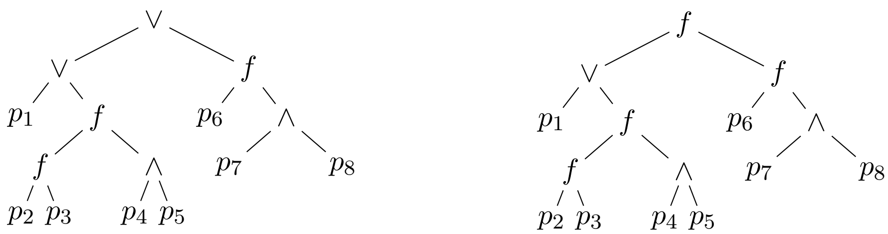

Figure 1: Syntax trees of the formulae $\varphi = \big(p_1 \vee f\langle f\langle p_2, p_3\rangle, p_4 \wedge p_5\rangle\big) \vee f\langle p_6, p_7 \wedge p_8\rangle$ and $\psi = f\langle p_1 \vee f\langle f\langle p_2, p_3\rangle, p_4 \wedge p_5\rangle, f\langle p_6, p_7 \wedge p_8\rangle\rangle$.

combinations $\omega_1(x,y) \equiv \omega_2(x,y) \equiv f\langle x,y\rangle$ of the variables x and y, such that x occurs only once in $\omega_1(x,y)$ and y only once in $\omega_2(x,y)$. Proceeding bottom-up, we now replace each f-node $f\langle\alpha,\beta\rangle$ in the syntax tree of φ by either $\omega_1(\alpha,\beta)$ or $\omega_2(\alpha,\beta)$, depending on whether we have previously modified the left or the right sub-tree (regarding, e.g., $f\langle p_2,p_3\rangle$, we are free to choose between ω_1 and ω_2). Note that, although ω_1 and ω_2 may duplicate some parts of φ, our choice ensures that we never duplicate such parts whose size has already changed. Consequently, this procedure results in a linear increase $\ell \cdot |\varphi|$ in the size of φ, where the coefficient ℓ essentially depends on how often y occurs in $\omega_1(x,y)$ and how often x occurs in $\omega_2(x,y)$. The resulting formula φ' belongs to $\mathrm{ML}[G]$ and is equivalent to φ. Hence $\mathrm{ML}[F \cup G]$-formulae where each f-node in the syntax tree has at most one immediate successor s such that the sub-tree with root s contains an f-node have $\mathrm{ML}[G]$-translations of linear size.

The formula $\psi = f\langle\alpha,\beta\rangle$ on the other hand does not have this property, but the two sub-formulae $\alpha = p_1 \vee f\langle f\langle p_2,p_3\rangle, p_4 \wedge p_5\rangle$ and $\beta = f\langle p_6, p_7 \wedge p_8\rangle$ do. Hence we can apply the above transformation to them separately, yielding equivalent $\mathrm{ML}[G]$-formulae α' and β' whose size increases at most by a factor of ℓ. Then $\psi' = f\langle\alpha',\beta'\rangle \equiv \psi$ is of size $|\psi'| \le \ell \cdot |\psi|$. Note that this step reduces the total number of f-vertices. In our example, $\alpha', \beta' \in \mathrm{ML}[G]$ and there remains only the single f-node which is the root of the syntax tree of $\psi' = f\langle\alpha',\beta'\rangle$. Applying the step again yields an equivalent $\mathrm{ML}[G]$-formula ψ'' of size $|\psi''| \le \ell \cdot |\psi'| \le \ell^2 \cdot |\psi|$.

Consider now an arbitrary formula $\chi \in \mathrm{PL}[G \cup F]$ and let D denote the number of steps that are required until all f-nodes have been removed from χ, i.e., until a $\mathrm{PL}[G]$-formula is obtained. Then the resulting formula has size at most $\ell^D \cdot |\chi|$. In this section, we will show that D is at most logarithmic in the size of χ. Both results together give a polynomial bound on the size of the equivalent formula from $\mathrm{ML}[G]$, which establishes the main part of the succinctness result.

We now formalize this idea and prove the result rigorously. Let F and G be disjoint sets of Boolean functions (we assume that F and G are disjoint for notational convenience). Let $N_{F,G} \subseteq \mathrm{ML}[F \cup G]$ be defined by

$$\varphi ::= \ \psi \mid f\langle\psi,\ldots,\psi,\varphi,\psi,\ldots,\psi\rangle \mid g\langle\varphi,\ldots\varphi\rangle \mid \Diamond\varphi$$

for $\psi \in \mathrm{ML}[G]$, $f \in F$, and $g \in G$. Note that a formula φ belongs to $N_{F,G}$ if, and only if, for any sub-formula $f\langle\alpha_1,\ldots,\alpha_k\rangle$ of φ with $f \in F$, at most one α_i contains some function from F. There is no such restriction on the sub-formulae $g\langle\alpha_1,\ldots,\alpha_k\rangle$ of φ where $g \in G$. Note that a formula in $N_{F,G} \setminus \mathrm{ML}[G]$ contains some sub-formula of the form $f\langle\alpha_1,\ldots,\alpha_k\rangle$ (and all but at most one α_i belong to $\mathrm{ML}[G]$).

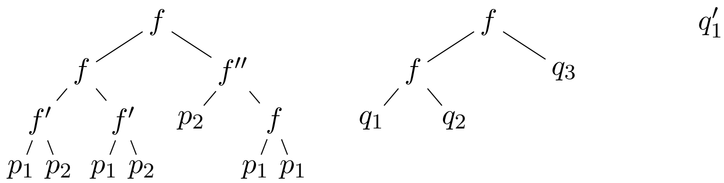

Figure 2: Syntax trees of $\varphi = f\big\langle f\big\langle f'\langle p_1, p_2\rangle, f'\langle p_1, p_2\rangle\big\rangle, f''\langle p_2, f\langle p_1, p_1\rangle\rangle\big\rangle$, $d_F(\varphi) = f\big\langle f\langle q_1, q_2\rangle, q_3\big\rangle$, and $d_F^2(\varphi) = q_1'$.

Definition 3.6 (Derivative)**.** Let F and G be disjoint sets of Boolean functions and let $\varphi = \varphi(p_1, \ldots, p_m) \in \mathrm{ML}[F \cup G]$. The *$F$-derivative $d_F(\varphi)$ of φ* is the smallest $\mathrm{ML}[F \cup G]$-formula $\gamma(p_1, \ldots, p_m, q_1, \ldots, q_n)$ (up to renaming of the variables $q_1, \ldots, q_n$), such that

- q_i occurs exactly once in $\gamma(p_1, \ldots, p_m, q_1, \ldots q_n)$ for all $i \in [n]$ and
- there exist $\alpha_1, \ldots, \alpha_n \in N_{F,G} \setminus \mathrm{ML}[G]$ such that φ and $\gamma(p_1, \ldots, p_m, \alpha_1, \ldots, \alpha_n)$ are identical.

Intuitively, $\gamma(p_1, \ldots, p_m, q_1, \ldots, q_n)$ is obtained from $\varphi(p_1, \ldots, p_m)$ by simultaneously replacing all "maximal $(N_{F,G} \setminus \mathrm{ML}[G])$-formulae" by distinct fresh variables $q_1, \ldots, q_n$ (where multiple occurrences of the same formula are replaced by different variables). An example is depicted in Fig. 2 (with $F = \{f, f', f''\}$ and $G = \{\neg, \wedge, \vee\}$).

Let $\varphi \in \mathrm{ML}[F \cup G]$ be a formula that contains some function from F. Then $d_F(\varphi)$ contains fewer occurrences of functions from F than φ. Hence there exists a smallest integer $r \geq 0$ for which the r-th derivative $d_F^r(\varphi) = d_F(d_F(\cdots d_F(\varphi)\cdots))$ is an $\mathrm{ML}[G]$-formula, where $d_F^0(\varphi) = \varphi$.

Definition 3.7 (Rank)**.** Let $\varphi \in \mathrm{ML}[F \cup G]$. The *$F$-rank $\mathrm{rk}_F(\varphi)$ of φ* is the smallest integer $r \geq 0$ for which $d_F^r(\varphi) \in \mathrm{ML}[G]$.

We first show that a formula with high F-rank must also be large.

Lemma 3.8. *Let F and G be disjoint sets of Boolean functions and let $\varphi \in \mathrm{ML}[F \cup G]$. Then $|\varphi| \geq 2^{\mathrm{rk}_F(\varphi)} - 1$.*

Proof. Throughout this proof, we will refer to the F-rank of a formula simply as rank.

We prove the stronger claim that a formula φ of positive rank has at least $2^{\mathrm{rk}_F(\varphi)-1}$ subformulae (counting multiplicity) of the form $f\langle \beta_1, \ldots, \beta_k\rangle$ with $f \in F$ and $\beta_1, \ldots, \beta_k \in \mathrm{ML}[G]$. In the following, we call such formulae *F-leaves*. Since counting the F-leaves in all derivatives of φ results in a lower bound on the total number of functions from F in φ, it follows that φ contains at least $2^0 + 2^1 + \ldots + 2^{\mathrm{rk}_F(\varphi)-1} = 2^{\mathrm{rk}_F(\varphi)} - 1$ functions from F.

It now remains to prove the bound on the number of F-leaves. Recall that we consider formulae of rank at least one. If $\mathrm{rk}_F(\varphi) = 1$, φ contains at least one function from F and hence has at least one F-leaf. Now, assume that $\varphi = \varphi(\overline{p})$ is of rank at least two, where $\overline{p} = (p_1, \ldots, p_m)$. Let $\gamma(\overline{p}, q_1, \ldots, q_n)$ be the derivative of φ and let $\alpha_1, \ldots, \alpha_n \in N_{F,G} \setminus \mathrm{ML}[G]$ with $\varphi = \gamma(\overline{p}, \alpha_1, \ldots, \alpha_n)$. Then, for every F-leaf $f\langle \beta_1, \ldots, \beta_k\rangle$ of $\gamma(\overline{p}, q_1, \ldots, q_n)$, there exist indices $1 \leq i < j \leq k$ such that $\beta_i, \beta_j \in \{q_1, \ldots, q_n\}$. By induction hypothesis, $\gamma(\overline{p}, q_1, \ldots, q_n)$ has at least $2^{\mathrm{rk}_F(\varphi)-2}$ F-leaves, each of which contains at least two of the fresh variables $\{q_1, \ldots, q_n\}$. Since each α_i contains at least one function from F, $\varphi = \gamma(\overline{p}, \alpha_1, \ldots, \alpha_n)$ has at least twice the number of F-leaves compared to $\gamma(\overline{p}, q_1, \dots, q_n)$, i.e., at least $2^{\mathrm{rk}_F(\varphi)-1}$ F-leaves. □

We can now turn to the main ingredient for our succinctness result.

Lemma 3.9. *Let F and G be disjoint sets of Boolean functions, $\top, \bot \in G$ (such that no function from F is nullary), and, for $f \in F$ and $i \in [\mathrm{ar}(f)]$, let $\omega_{f,i} \in \mathrm{PL}[G]$ be a $\mathrm{PL}[G]$-representation of (f, i). Let, furthermore, $\kappa\colon \mathbb{N} \to \mathbb{N}$ be a monotone function such that $1 \le \kappa(0)$ and $|\omega_{f,i}| \le \kappa(\mathrm{ar}(f))$ for any $f \in F$ and $i \in [\mathrm{ar}(f)]$. Finally, let $\kappa'\colon \mathbb{N} \to \mathbb{N}\colon n \mapsto \kappa(n)^{1+\log_2 n} \cdot n$.*

Then $\mathrm{ML}[F \cup G]$ has κ'-translations in $\mathrm{ML}[G]$.

Proof. For $\varphi \in \mathrm{ML}[F \cup G]$, let K_φ denote the maximal arity of any $f \in F$ occurring in φ (or 0 if $\varphi \in \mathrm{ML}[G]$).

We prove the following claim: every formula $\varphi(r_1, \dots, r_\ell) \in \mathrm{ML}[F \cup G]$ is equivalent to an $\mathrm{ML}[G]$-formula $\psi(r_1, \dots, r_\ell)$ of size at most $\kappa(K_\varphi)^{\mathrm{rk}_F(\varphi)} \cdot |\varphi|$. For notational simplicity, we write $\bar{r}$ for the tuple $(r_1, \dots, r_\ell)$. Since $K_\varphi \le |\varphi|$ and $\mathrm{rk}_F(\varphi) \le \log_2(|\varphi| + 1)$ by Lemma 3.8, this claim ensures that every $\mathrm{ML}[F \cup G]$-formula φ of size n has an equivalent $\mathrm{ML}[G]$-formula of size at most $\kappa(K_\varphi)^{\log_2(|\varphi|+1)} \cdot |\varphi| \le \kappa(n)^{1+\log_2 n} \cdot n = \kappa'(n)$.

The proof of the claim proceeds by induction on the F-rank of φ. Since F remains fixed, we will refer to the F-rank simply as rank.

Let $\varphi \in \mathrm{ML}[F \cup G]$ be of rank at most one. We show by induction on the structure of φ that there exists an equivalent $\mathrm{ML}[G]$-formula $\varphi'(\bar{r})$ of size $|\varphi'| \le \kappa(K_\varphi) \cdot |\varphi|$. If φ is a propositional variable, $|\varphi| = 1 \le \kappa(0) \le \kappa(1) \cdot |\varphi|$ by choice of the function κ.

Now, assume that $\varphi(\bar{r}) = g\langle \alpha_1, \dots, \alpha_k \rangle$ where $g \in G$ is of arity k. By induction hypothesis, there exists for each $j \in [k]$ an $\mathrm{ML}[G]$-formula $\beta_j(\bar{r})$ with $\beta_j \equiv \alpha_j$ and $|\beta_j| \le \kappa(K_{\alpha_j}) \cdot |\alpha_j|$. Since κ is monotone and $K_{\alpha_j} \le K_\varphi$, we obtain $|\beta_j| \le \kappa(K_\varphi) \cdot |\alpha_j|$ for $j \in [k]$. Set $\varphi'(\bar{r}) = g\langle \beta_1, \dots, \beta_k \rangle$. Then φ' is an $\mathrm{ML}[G]$-formula and equivalent to φ. Furthermore, the size of φ' satisfies

$$\begin{aligned} |\varphi'| &= 1 + |\beta_1| + \dots + |\beta_k| \\ &\le 1 + \kappa(K_\varphi) \cdot \big(|\alpha_1| + \dots + |\alpha_k|\big) \\ &\le \kappa(K_\varphi) \cdot \big(1 + |\alpha_1| + \dots + |\alpha_k|\big) && \text{since } \kappa(K_\varphi) \ge 1 \\ &= \kappa(K_\varphi) \cdot |\varphi|\,. \end{aligned}$$

A similar argument establishes the case $\varphi(\bar{r}) = \Diamond\alpha$.

Finally, assume that $\varphi(\bar{r}) = f\langle \alpha_1, \dots, \alpha_k \rangle$ where $f \in F$ is of arity k. Since F and G are disjoint and since all constant functions belong to G, we get $k \ge 1$. Recall that φ is of rank one and therefore a formula in $N_{F,G}$. Hence there is $i \in [k]$ such that, from among the arguments $\alpha_1, \dots, \alpha_k$, at most α_i contains a function from F, i.e., such that $\alpha_j \in \mathrm{ML}[G]$ for all $j \in [k] \setminus \{i\}$. By induction hypothesis, there is an $\mathrm{ML}[G]$-formula $\beta_i(\bar{r}) \equiv \alpha_i(\bar{r})$ of size $|\beta_i| \le \kappa(K_{\alpha_i}) \cdot |\alpha_i| \le \kappa(K_\varphi) \cdot |\alpha_i|$. Recall that $\omega_{f,i}(p_1, \dots, p_k)$ is a $\mathrm{PL}[G]$-representation of (f, i) that uses the variable p_i at most once and has size $|\omega_{f,i}| \le \kappa(k) \le \kappa(K_\varphi)$. Set $\varphi'(\bar{r}) = \omega_{f,i}(\beta_1, \dots, \beta_k)$ where $\beta_j(\bar{r}) = \alpha_j(\bar{r})$ for $j \in [k] \setminus \{i\}$. Then φ' is an $\mathrm{ML}[G]$-formula and equivalent to $f\langle \alpha_1, \dots, \alpha_k \rangle = \varphi$. Furthermore, φ' is obtained from $\omega_{f,i}(p_1, \dots, p_k)$ by replacing one variable with β_i and all others with formulae of size at most $\sum_{j \ne i} |\alpha_j|$. Hence

$$|\varphi'| = |\omega_{f,i}(\beta_1, \dots, \beta_k)|$$

$$\begin{aligned}
&\leq |\beta_i| + \kappa(k) \cdot \Big(1 + \sum_{j \neq i} |\alpha_j|\Big) \\
&\leq \kappa(K_\varphi) \cdot |\alpha_i| + \kappa(K_\varphi) \cdot \Big(1 + \sum_{j \neq i} |\alpha_j|\Big) \\
&= \kappa(K_\varphi) \cdot \Big(1 + \sum_{j} |\alpha_j|\Big) \\
&= \kappa(K_\varphi) \cdot |f\langle \alpha_1, \ldots, \alpha_k \rangle| = \kappa(K_\varphi) \cdot |\varphi|.
\end{aligned}$$

This shows the claim for formulae of rank at most one.

We proceed by induction on the rank of φ. Assume $\varphi(\bar{r}) \in \mathrm{ML}[F \cup G]$ is of rank $R \geq 2$. Let $\gamma(\bar{r}, q_1, \ldots, q_n)$ be the derivative of $\varphi(\bar{r})$ and let $\alpha_1, \ldots, \alpha_n \in N_{F,G}$ with $\varphi = \gamma(\bar{r}, \alpha_1, \ldots, \alpha_n)$. Since each of the formulae $\alpha_1(\bar{r}), \ldots, \alpha_k(\bar{r})$ is of rank one, there are $\beta_1(\bar{r}), \ldots, \beta_n(\bar{r}) \in \mathrm{ML}[G]$ with $\alpha_i \equiv \beta_i$ and $|\beta_i| \leq \kappa(K_{\alpha_i}) \cdot |\alpha_i| \leq \kappa(K_\varphi) \cdot |\alpha_i|$ for $i \in [k]$. Let $\psi(\bar{r}) = \gamma(\bar{r}, \beta_1, \ldots, \beta_k)$. Then ψ is equivalent to φ and of size $|\psi| \leq \kappa(K_\varphi) \cdot |\varphi|$. Intuitively, ψ is obtained from φ by replacing the "maximal $\mathrm{ML}[F \cup G]$ sub-formulae" of rank one by equivalent $\mathrm{ML}[G]$-formulae. Hence the rank of ψ is equal to the rank of γ, namely $R - 1$. It now follows by induction hypothesis that ψ is equivalent to a formula $\varphi' \in \mathrm{ML}[G]$ of size $|\varphi'| \leq \kappa(K_\psi)^{R-1} \cdot |\psi| \leq \kappa(K_\varphi)^R \cdot |\varphi|$. Since $\varphi \equiv \psi \equiv \varphi'$, this finishes the verification of the claim from the beginning of this proof. □

From Lemma 3.9, we can get the main result of this section, stating that $\mathrm{ML}[F \cup G]$ is not more succinct than $\mathrm{ML}[G]$, provided F is a finite set of Boolean functions with $\mathrm{PL}[G]$-representations.

Proposition 3.10. *Let F and G be sets of Boolean functions such that F is finite and has $\mathrm{PL}[G]$-representations and such that $\top, \bot \in G$. Then $\mathrm{ML}[F \cup G]$ has polynomial translations in $\mathrm{ML}[G]$.*

Noting that $\mathrm{ML}[F] \subseteq \mathrm{ML}[F \cup G]$, the above implies in particular that $\mathrm{ML}[F]$ has polynomial translations in $\mathrm{ML}[G]$. In view of Example 3.5, it ensures specifically that $\mathrm{ML}[\mathrm{DMor} \cup \{\mathrm{maj}\}]$ has polynomial translations in $\mathrm{ML}[\mathrm{DMor}]$.

Proof. The set $F' = F \setminus G$ is finite, has $\mathrm{PL}[G]$-representations, and, in addition, F' and G are disjoint with $F \cup G = F' \cup G$ and $\top, \bot \in G$.

Choose $\mathrm{PL}[G]$-representations $\omega_{f,i}$ for all $f \in F'$ and $i \in [k]$ (where $k > 0$ is the arity of f). Since F' is finite, there is some constant $c > 0$ such that all these formulae are of size at most c. Let $\kappa : \mathbb{N} \to \mathbb{N}$ be the constant function with $\kappa(n) = c$ for all $n \in \mathbb{N}$. By the previous lemma, any $\varphi \in \mathrm{ML}[F' \cup G]$ is thus equivalent to an $\mathrm{ML}[G]$-formula of size at most $\kappa(|\varphi|)^{1+\log_2 |\varphi|} \cdot |\varphi| = c^{1+\log_2 |\varphi|} \cdot |\varphi| = c \cdot |\varphi|^{1+\log_2 c} \leq c \cdot |\varphi|^d$ for some $d > 0$. Since $\mathrm{ML}[F' \cup G] = \mathrm{ML}[F \cup G]$, the claim follows. □

3.2. **Representations of F in $\mathrm{PL}[\mathrm{DMor}]$ and $\mathrm{PL}[\mathrm{DMor}_\leftrightarrow]$.** This section is devoted to the second step of our programme (see page 7), i.e., we will construct $\mathrm{PL}[\mathrm{DMor}_\leftrightarrow]$- and $\mathrm{PL}[\mathrm{DMor}]$-representations of arbitrary Boolean functions.

Proposition 3.11. *Let f be a Boolean function of arity $k > 0$. Then*

(1) *f has $\mathrm{PL}[\mathrm{DMor}_\leftrightarrow]$-representations and*
(2) *f has $\mathrm{PL}[\mathrm{DMor}]$-representations if, and only if, the function f is locally monotone.*

Proof. (1) For notational simplicity, we prove that there is some PL[DMor$_\leftrightarrow$]-representation of (f,k), i.e., that there is some PL[DMor$_\leftrightarrow$]-formula $\omega(p_1,\dots,p_k)$ that is equivalent to $f\langle p_1,\dots,p_k\rangle$ and that uses the variable p_k at most once. Over the De Morgan basis there is a formula in disjunctive normal form that is equivalent to $f\langle p_1,\dots,p_k\rangle$ (but that may use p_k more than once). Without loss of generality however, we can assume that every clause contains precisely one of p_k and $\neg p_k$. Thus, there are formulae $\alpha = \alpha(p_1,\dots,p_{k-1})$ and $\beta = \beta(p_1,\dots,p_{k-1})$ in disjunctive normal form such that

$$f\langle p_1,\dots,p_k\rangle \equiv (p_k \wedge \alpha) \vee (\neg p_k \wedge \beta)\,.$$

But this latter formula is equivalent to the formula

$$(\alpha \wedge \beta) \vee \big(\neg(\alpha \wedge \beta) \wedge (p_k \leftrightarrow (\alpha \wedge \neg\beta))\big)$$

that belongs to ML[DMor$_\leftrightarrow$] and uses p_k only once.

(2) First, let f be locally monotone. For notational simplicity, we will only construct a PL[DMor]-representation of (f,k). In addition, we assume that f is increasing in the k-th argument, i.e., $f(\overline{a},0) \le f(\overline{a},1)$ for all $\overline{a} \in \mathbb{B}^{k-1}$. Since DMor is complete, there is a PL[DMor]-formula $\omega(p_1,\dots,p_k)$ that is equivalent to $f\langle p_1,\dots,p_k\rangle$. Since $f(\overline{a},0) \le f(\overline{a},1)$ for any $\overline{a} \in \mathbb{B}^{k-1}$, it follows that

$$f\langle p_1,\dots,p_k\rangle \equiv \Big(\omega\big(p_1,\dots,p_{k-1},\top\big) \wedge p_k\Big) \vee \omega\big(p_1,\dots,p_{k-1},\bot\big)\,.$$

In particular, the formula on the right uses the variable p_k only once and therefore forms a PL[DMor]-representation of (f,k).

Now assume that f has PL[DMor]-representations. We prove by induction on the size of a formula $\omega(p_1,\dots,p_k) \in$ PL[DMor] the following: if the variable p_k appears at most once in ω, then the function $[\![\omega; p_1,\dots,p_k]\!]$ represented by ω is monotone in its k-th argument. The claim is trivial for formulae of the form $\top\langle\rangle$, $\bot\langle\rangle$, and p_i. In particular, we can assume $k > 0$.

For the induction step, let $\omega = \alpha_1 \vee \alpha_2$. Since the formula ω contains the variable p_k at most once, it appears in at most one of the arguments α_i; for notational simplicity, we assume it does not appear in α_1. By the induction hypothesis, we have one of the following:

- for all $\overline{a} \in \mathbb{B}^{k-1}$: $[\![\alpha_2; p_1,\dots,p_k]\!](\overline{a},0) \le [\![\alpha_2; p_1,\dots,p_k]\!](\overline{a},1)$ and therefore
  $[\![\alpha_1 \vee \alpha_2; p_1,\dots,p_k]\!](\overline{a},0) \le [\![\alpha_1 \vee \alpha_2; p_1,\dots,p_k]\!](\overline{a},1)$ or
- for all $\overline{a} \in \mathbb{B}^{k-1}$: $[\![\alpha_2; p_1,\dots,p_k]\!](\overline{a},0) \ge [\![\alpha_2; p_1,\dots,p_k]\!](\overline{a},1)$ and therefore
  $[\![\alpha_1 \vee \alpha_2; p_1,\dots,p_k]\!](\overline{a},0) \ge [\![\alpha_1 \vee \alpha_2; p_1,\dots,p_k]\!](\overline{a},1)$.

In either case, the formula ω represents a function that is locally monotone in its last argument (provided ω is of the form $\alpha_1 \vee \alpha_2$). The remaining cases $\omega = \alpha_1 \wedge \alpha_2$ or $\omega = \neg\alpha_1$ follow by a similar argument. This finishes the inductive proof.

Recall that f has PL[DMor]-representations. Since each PL[DMor]-representation of (f,i) describes a function (namely f) that is monotone in the i-th argument, we obtain that the function f is locally monotone, which completes the proof. □

3.3. **Representations of** DMor **and** DMor$_\leftrightarrow$ **in** PL[G]**.** This section is devoted to the third step of our programme (see page 7), i.e., given an arbitrary complete set G of Boolean functions, we will construct PL[G]-representations of the Boolean functions from DMor and DMor$_\leftrightarrow$, respectively. More precisely, we show the following.

Theorem 3.12. *Let G be a complete set of Boolean functions with $\top, \bot \in G$.*

- *Then* DMor *has* PL$[G]$*-representations.*
- *If* G *contains some non-locally monotone function, then* $\mathrm{DMor}_{\leftrightarrow}$ *has* PL$[G]$*-representations.*

Proof. The claims follow from Propositions 3.13, 3.14, and 3.15 below, as well as De Morgan's law $x \wedge y \equiv \neg(\neg x \vee \neg y)$. □

Let G be some complete set of Boolean functions with $\bot, \top \in G$. We now provide PL$[G]$-representations for each of the Boolean functions $\neg$, $\vee$, and $\leftrightarrow$. These constructions follow the proof from [PM90] (at least for disjunction, but also the handling of bi-implication is a variant of their handling of disjunction). Thus, in some sense, the following constructions can be understood as an analysis of the constructions by Pelletier and Martin with an additional emphasis on the condition that the variable p_1 is used only once.

Negation. We start with the construction of a PL$[G]$-representation of $(\neg, 1)$. Since G is complete, there is some PL$[G]$-formula $\alpha(p_1)$ that is equivalent to $\neg p_1$. However, α may use p_1 more than once, say, ℓ times. Let $q_1, \ldots, q_\ell$ be distinct variables and let $\alpha'(q_1, \ldots, q_\ell)$ be obtained from α by substituting each occurrence of p_1 by one of the variables $q_1, \ldots, q_\ell$ such that each q_i is used exactly once. Then α and α' may not be equivalent but

$$[\![\alpha'; q_1, \ldots, q_\ell]\!](0, \ldots, 0) = \neg(0) = 1 \quad \text{and} \quad [\![\alpha'; q_1, \ldots, q_\ell]\!](1, \ldots, 1) = \neg(1) = 0\,.$$

Now, consider the sequence

$$\begin{aligned} 1 = {} & [\![\alpha'; q_1, \ldots, q_\ell]\!](0, 0, \ldots, 0, 0), \\ & [\![\alpha'; q_1, \ldots, q_\ell]\!](1, 0, \ldots, 0, 0), \\ & \qquad \cdots \\ & [\![\alpha'; q_1, \ldots, q_\ell]\!](1, 1, \ldots, 1, 0), \\ & [\![\alpha'; q_1, \ldots, q_\ell]\!](1, 1, \ldots, 1, 1) = 0\,, \end{aligned}$$

starting with a value of 1 and ending with a value of 0. Hence there is $i \in [\ell]$ such that

$$[\![\alpha'; q_1, \ldots, q_\ell]\!](\underbrace{1, \ldots, 1}_{i-1 \text{ times}}, 0, \underbrace{0, \ldots, 0}_{\ell-i \text{ times}}) = 1 > 0 = [\![\alpha'; q_1, \ldots, q_\ell]\!](\underbrace{1, \ldots, 1}_{i-1 \text{ times}}, 1, \underbrace{0, \ldots, 0}_{\ell-i \text{ times}})\,.$$

In particular, $[\![\alpha'; q_1, \ldots, q_\ell]\!](1, \ldots, 1, a, 0, \ldots, 0) = 1 - a$, where the i-th argument is set to $a \in \mathbb{B}$. Set

$$\gamma_j() = \begin{cases} \top\langle\rangle & \text{if } j < i \\ \bot\langle\rangle & \text{if } j > i \end{cases}$$

for all $j \in [\ell] \setminus \{i\}$ and consider the PL$[G]$-formula

$$\mathtt{neg}(p_1) = \alpha'\big(\gamma_1, \ldots, \gamma_{i-1}, p_1, \gamma_{i+1}, \ldots, \gamma_\ell\big)$$

that uses p_1 only once since α' uses no variable twice. Then, for any interpretation $\mathcal{I}$, we get

$$\begin{aligned} \mathcal{I}(\mathtt{neg}) &= [\![\alpha'; q_1, \ldots, q_\ell]\!]\big(\mathcal{I}(\gamma_1), \ldots, \mathcal{I}(\gamma_{i-1}), \mathcal{I}(p_1), \mathcal{I}(\gamma_{i+1}), \ldots, \mathcal{I}(\gamma_k)\big) \\ &= [\![\alpha'; q_1, \ldots, q_\ell]\!]\big(1, \ldots, 1, \mathcal{I}(p_1), 0, \ldots, 0\big) \\ &= 1 - \mathcal{I}(p_1) \\ &= \mathcal{I}\big(\neg p_1\big)\,. \end{aligned}$$

Hence the formulae $\mathtt{neg}(p_1)$ and $\neg p_1$ are equivalent. This shows

Proposition 3.13. *Let G be a complete set of Boolean functions with $\bot, \top \in G$. Then $\mathtt{neg}(p_1)$ is a* PL$[G]$*-representation of* $(\neg, 1)$.

For the next step, we require the following concept. A Boolean function $f : \mathbb{B}^k \to \mathbb{B}$ is *affine* if there are $c_0, \ldots, c_k \in \mathbb{B}$ such that $f(a_1, \ldots, a_k) = \big(c_0 + \sum_{i\in[k]} c_i \cdot a_i\big) \bmod 2$ for all $a_1, \ldots, a_k \in \mathbb{B}$. Then f is affine if, and only if, for all $i \in [n]$, one of the following holds:

- $f(a_1, \ldots, a_{i-1}, 0, a_{i+1}, \ldots, a_k) = f(a_1, \ldots, a_{i-1}, 1, a_{i+1}, \ldots, a_k)$ for all $a_1, \ldots, a_n \in \mathbb{B}$ or
- $f(a_1, \ldots, a_{i-1}, 0, a_{i+1}, \ldots, a_k) \neq f(a_1, \ldots, a_{i-1}, 1, a_{i+1}, \ldots, a_k)$ for all $a_1, \ldots, a_n \in \mathbb{B}$.

Note that the functions $\neg$, $\bot$, $\top$, and $\leftrightarrow$ are affine while $\vee$ and $\wedge$ are not affine.

Disjunction. We now construct a PL$[G]$-representation of $(\vee, 1)$, i.e., a PL$[G]$-formula that is equivalent to $p_1 \vee p_2$ and uses p_1 at most once. In our construction, we will use the formula $\mathtt{neg}(p_1)$ from Proposition 3.13.

Since G is complete, there exists a PL$[G]$-formula $\alpha(p_1, p_2)$ that is equivalent to $p_1 \vee p_2$. Let ℓ be the number of occurrences of p_1 in α and let m be the number of occurrences of p_2 in α. Furthermore, let $q_1, \ldots, q_{\ell+m}$ be distinct fresh variables and let $\alpha'(q_1, \ldots, q_{\ell+m})$ be obtained from α by substituting each occurrence of p_1 by one of $q_1, \ldots, q_\ell$ and each occurrence of p_2 by one of $q_{\ell+1}, \ldots, q_{\ell+m}$ such that each new variable is used exactly once. For convenience, let $k = \ell + m$ and $f_{\alpha'}$ denote the k-ary Boolean function $[\![\alpha'; q_1, \ldots, q_k]\!]$.

For $i \in \mathbb{N}$, let $\bar{0}_i$ and $\bar{1}_i$ denote the i-tuples consisting solely of zeroes or solely of ones. Then

$$f_{\alpha'}(\bar{0}_\ell, \bar{0}_m) = \vee(0,0) = 0\,, \qquad f_{\alpha'}(\bar{1}_\ell, \bar{0}_m) = \vee(1,0) = 1\,,$$
$$f_{\alpha'}(\bar{0}_\ell, \bar{1}_m) = \vee(0,1) = 1\,, \text{ and} \qquad f_{\alpha'}(\bar{1}_\ell, \bar{1}_m) = \vee(1,1) = 1\,.$$

Aiming at a contradiction, assume that $f_{\alpha'}$ were affine, i.e., that there are $c_i \in \mathbb{B}$ for $i \in [k] \cup \{0\}$ such that for all $a_1, \ldots, a_k \in \mathbb{B}$

$$f_{\alpha'}(a_1, \ldots, a_k) = \Big(c_0 + \sum_{i\in[k]} c_i \cdot a_i\Big) \bmod 2\,.$$

Using the relationship between $f_{\alpha'}$ and $\vee$, we can thus readily observe the following:

- $c_0 = 0$ since $f_{\alpha'}(\bar{0}_\ell, \bar{0}_m) = 0$,
- the set $\{i \in [k] : i \le \ell \text{ and } c_i = 1\}$ is of odd size since $f_{\alpha'}(\bar{1}_\ell, \bar{0}_m) = 1$ and $c_0 = 0$, and
- the set $\{i \in [k] : i > \ell \text{ and } c_i = 1\}$ is also of odd size since $f_{\alpha'}(\bar{0}_\ell, \bar{1}_m) = 1$ and $c_0 = 0$.

But then the total number of non-zero coefficients is even, hence $\vee(1,1) = f_{\alpha'}(\bar{1}_\ell, \bar{1}_m) = 0$, a contradiction. Consequently, $f_{\alpha'}$ is not affine and there are $i \in [k]$ and $a_1, \ldots, a_k, b_1, \ldots, b_k \in \mathbb{B}$ such that

$$f_{\alpha'}(a_1, \ldots, a_{i-1}, 0, a_{i+1}, \ldots, a_k) = f_{\alpha'}(a_1, \ldots, a_{i-1}, 1, a_{i+1}, \ldots, a_k) \text{ and}$$
$$f_{\alpha'}(b_1, \ldots, b_{i-1}, 0, b_{i+1}, \ldots, b_k) \neq f_{\alpha'}(b_1, \ldots, b_{i-1}, 1, b_{i+1}, \ldots, b_k)\,.$$

Recall that we want to construct a representation of $\vee$ that uses p_1 only once. The general idea is as follows: the variable q_i will be substituted by p_1 (or $\mathtt{neg}(p_1)$, depending on the precise distribution of zeroes and ones in the above equalities). At this point it is important that α' uses q_i only once. The remaining variables q_j with $j \neq i$ will be substituted by formulae depending solely on q_2 such that, if q_2 is set to one, we arrive in the first line and if q_2 is set to zero, we arrive in the second line. More precisely, we define formulae γ_j for $j \in [k]$ as follows:

$$\gamma_j(p_1, p_2) = \begin{cases} \top\langle\rangle & \text{if } a_j = b_j = 1 \text{ and } j \neq i \\ \bot\langle\rangle & \text{if } a_j = b_j = 0 \text{ and } j \neq i \\ \texttt{neg}(p_2) & \text{if } a_j < b_j \text{ and } j \neq i \\ p_2 & \text{if } a_j > b_j \text{ and } j \neq i \\ p_1 & \text{if } j = i \end{cases}$$

Then each $\gamma_j(p_1, p_2) \in \mathrm{PL}[G]$ uses both p_1 and p_2 at most once. Furthermore, there is at most one $j \in [k]$ such that γ_j uses p_1 at all (p_2 can be used by more than one of these formulae since there may exist several $j \neq i$ with $a_j \neq b_j$). Then these formulae have indeed the properties described above, as for any interpretation $\mathcal{I}$ we have

$$\mathcal{I}(\gamma_j) = \begin{cases} a_j & \text{if } \mathcal{I}(p_2) = 1 \text{ and } j \neq i\,, \\ b_j & \text{if } \mathcal{I}(p_2) = 0 \text{ and } j \neq i\,, \\ \mathcal{I}(p_1) & \text{if } j = i\,. \end{cases}$$

Let $\texttt{predis}(p_1, p_2) = \alpha'(\gamma_1, \ldots, \gamma_k)$ and, for $a, b \in \mathbb{B}$, let $\mathcal{I}_{a,b}$ denote an interpretation with $\mathcal{I}_{a,b}(p_1) = a$ and $\mathcal{I}_{a,b}(p_2) = b$. Then

$$\begin{aligned} \mathcal{I}_{0,1}(\texttt{predis}) &= f_{\alpha'}(a_1, \ldots, a_{i-1}, 0, a_{i+1}, \ldots, a_k) \\ &= f_{\alpha'}(a_1, \ldots, a_{i-1}, 1, a_{i+1}, \ldots, a_k) = \mathcal{I}_{1,1}(\texttt{predis}) \end{aligned}$$

and

$$\begin{aligned} \mathcal{I}_{0,0}(\texttt{predis}) &= f_{\alpha'}(b_1, \ldots, b_{i-1}, 0, b_{i+1}, \ldots, b_k) \\ &\neq f_{\alpha'}(b_1, \ldots, b_{i-1}, 1, b_{i+1}, \ldots, b_k) = \mathcal{I}_{1,0}(\texttt{predis}) \end{aligned}$$

Furthermore, $\texttt{predis} = \alpha'(\gamma_1, \ldots, \gamma_k) \in \mathrm{PL}[G]$ uses p_1 only once since α' uses no variable twice and since at most one of the formulae γ_j uses p_1 at all.

Now, all that remains is to modify predis by negating, if necessary, p_1, the whole formula, or both, such that $\mathcal{I}_{1,1}(\texttt{predis}') = 1$ and $\mathcal{I}_{0,0}(\texttt{predis}') = 0$. The above equalities then imply $\mathcal{I}_{0,1}(\texttt{predis}') = 1$ and $\mathcal{I}_{1,0}(\texttt{predis}') = 1$. We accomplish this in two steps:

(1) If $\mathcal{I}_{1,1}(\texttt{predis}) = 1$, then set $\texttt{predis}_1 = \texttt{predis}$. Otherwise, set $\texttt{predis}_1 = \texttt{neg}(\texttt{predis})$, i.e., we obtain $\texttt{predis}_1$ from the formula $\texttt{neg}(p_1)$ by substituting p_1 by predis.[4] In this latter case, we obtain

$$\begin{aligned} &\mathcal{I}_{0,1}(\texttt{predis}_1) = 1 - \mathcal{I}_{0,1}(\texttt{predis}) = 1 - \mathcal{I}_{1,1}(\texttt{predis}) = \mathcal{I}_{1,1}(\texttt{predis}_1) \text{ and} \\ &\mathcal{I}_{0,0}(\texttt{predis}_1) = 1 - \mathcal{I}_{0,0}(\texttt{predis}) \neq 1 - \mathcal{I}_{1,0}(\texttt{predis}) = \mathcal{I}_{1,0}(\texttt{predis}_1)\,. \end{aligned}$$

Since $\mathcal{I}_{1,1}(\texttt{predis}) = 1$, this implies

$$\begin{aligned} &\mathcal{I}_{0,1}(\texttt{predis}_1) = \mathcal{I}_{1,1}(\texttt{predis}_1) = 1 \text{ and} \\ &\mathcal{I}_{0,0}(\texttt{predis}_1) \neq \mathcal{I}_{1,0}(\texttt{predis}_1)\,. \end{aligned}$$

Note that this also holds in case $\mathcal{I}_{1,1}(\texttt{predis}) = 1$.

[4]Since $\mathcal{I}(\texttt{predis})$ depends solely on $\mathcal{I}(p_1)$ and $\mathcal{I}(p_2)$, the formula dis is well-defined.

(2) If $\mathcal{I}_{0,0}(\texttt{predis}_1) = 0$, then set $\texttt{predis}_2 = \texttt{predis}_1$. Otherwise, set $\texttt{predis}_2(p_1, p_2) = \texttt{predis}_1(\texttt{neg}(p_1), p_2))$, i.e., we obtain $\texttt{predis}_2$ from the formula $\texttt{predis}_1$ by replacing all occurrences of p_1 by the formula $\texttt{neg}(p_1)$. Again, since $\mathcal{I}(\texttt{predis}_1)$ depends, at most, on $\mathcal{I}(p_1)$ and $\mathcal{I}(p_2)$, the formula $\texttt{predis}_2$ is well-defined. In the latter case, we obtain

$$\mathcal{I}_{1,1}(\texttt{predis}_2) = \mathcal{I}_{0,1}(\texttt{predis}_1) = \mathcal{I}_{1,1}(\texttt{predis}_1) = \mathcal{I}_{0,1}(\texttt{predis}_2) \text{ and}$$
$$\mathcal{I}_{0,0}(\texttt{predis}_2) = \mathcal{I}_{1,0}(\texttt{predis}_1) \neq \mathcal{I}_{0,0}(\texttt{predis}_1) = \mathcal{I}_{1,0}(\texttt{predis}_2)\,.$$

Since $\mathcal{I}_{0,0}(\texttt{predis}_1) = 1$, this implies

$$\mathcal{I}_{0,1}(\texttt{predis}_2) = \mathcal{I}_{1,1}(\texttt{predis}_2) = 1 \text{ and}$$
$$0 = \mathcal{I}_{0,0}(\texttt{predis}_2) \neq \mathcal{I}_{1,0}(\texttt{predis}_2) = 1\,.$$

Note that this also holds in case $\mathcal{I}_{0,0}(\texttt{predis}_1) = 0$.

In summary, we have $\mathcal{I}_{a,b}(\texttt{predis}_2) = 1$ if, and only if, $a = 1$ or $b = 1$. Hence, indeed, the formulae $p_1 \vee p_2$ and $\texttt{predis}_2(p_1, p_2)$ are equivalent. Since, in the above procedure, we did not duplicate any variables, the formula $\texttt{predis}_2$ uses the variable p_1 at most once. Hence choosing $\texttt{dis}(p_1, p_2) = \texttt{predis}_2(p_1, p_2)$ proves

Proposition 3.14. *Let $G \ni \top, \bot$ be a complete set of Boolean functions. Then* $\texttt{dis}$ *is a* PL$[G]$*-representation of* $(\vee, 1)$*; by symmetry, there is also a* PL$[G]$*-representation of* $(\vee, 2)$.

Bi-implication. We next construct a PL$[G]$-representation of $(\leftrightarrow, 1)$, i.e., a PL$[G]$-formula that is equivalent to $p_1 \leftrightarrow p_2$ and uses p_1 at most once. Here, we use the additional assumption that G contains some function g that is not locally monotone. In our construction, we will use the PL$[G]$-representations $\texttt{neg}(p_1)$ and $\texttt{dis}(p_1, p_2)$ of $(\neg, 1)$ and $(\vee, 1)$ from Propositions 3.13 and 3.14, respectively.

Since g is not locally monotone, its arity k is at least two and there are $i \in [k]$ and $a_1, \ldots, a_k, b_1, \ldots, b_k \in \mathbb{B}$ such that

$$g(a_1, \ldots a_{i-1}, 0, a_{i+1}, \ldots, a_k) < g(a_1, \ldots a_{i-1}, 1, a_{i+1}, \ldots, a_k) \text{ and}$$
$$g(b_1, \ldots b_{i-1}, 0, b_{i+1}, \ldots, b_k) > g(b_1, \ldots b_{i-1}, 1, b_{i+1}, \ldots, b_k)\,.$$

Let the formulae γ_j be defined as in the previous section. We consider the formula

$$\texttt{biimpl}(p_1, p_2) = g\langle \gamma_1, \gamma_2, \ldots, \gamma_k \rangle\,.$$

Since at most one of the formulae γ_j uses p_1 at all, and this formula uses p_1 at most once, the formula $\texttt{biimpl}$ uses p_1 at most once.

For $a, b \in \mathbb{B}$, let $\mathcal{I}_{a,b}$ denote an interpretation with $\mathcal{I}_{a,b}(p_1) = a$ and $\mathcal{I}_{a,b}(p_2) = b$. Then we have

$$\begin{aligned}
\mathcal{I}_{0,1}(\texttt{biimpl}) &= g\big(\mathcal{I}_{0,1}(\gamma_1), \ldots, \mathcal{I}_{0,1}(\gamma_k)\big) \\
&= g(a_1, \ldots, a_{i-1}, 0, a_{i+1}, \ldots, a_k) \\
&< g(a_1, \ldots, a_{i-1}, 1, a_{i+1}, \ldots, a_k) \\
&= g\big(\mathcal{I}_{1,1}(\gamma_1), \ldots, \mathcal{I}_{1,1}(\gamma_k)\big) \\
&= \mathcal{I}_{1,1}(\texttt{biimpl})
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{I}_{0,0}(\texttt{biimpl}) &= g\big(\mathcal{I}_{0,0}(\gamma_1), \ldots, \mathcal{I}_{0,0}(\gamma_k)\big) \\
&= g(b_1, \ldots, b_{i-1}, 0, b_{i+1}, \ldots, b_k)
\end{aligned}$$

$$\begin{aligned} &> g(b_1, \ldots, b_{i-1}, 1, b_{i+1}, \ldots, b_k) \\ &= g\big(\mathcal{I}_{1,0}(\gamma_1), \ldots, \mathcal{I}_{1,0}(\gamma_k)\big) \\ &= \mathcal{I}_{1,0}(\texttt{biimpl}) \end{aligned}$$

It follows that

$$\mathcal{I}_{a,b}(\texttt{biimpl}) = 1 \iff a = b\,,$$

which proves

Proposition 3.15. *Let $G \ni \top, \bot$ be a complete set of Boolean functions with a non-locally monotone function $g \in G$. Then $\texttt{biimpl}(p_1, p_2)$ is a $\mathrm{PL}[G]$-representation of $(\leftrightarrow, 1)$; by symmetry, there is also a $\mathrm{PL}[G]$-representation of $(\leftrightarrow, 2)$.*

3.4. **Proof of this section's main theorem.**

Proof of Theorem 3.2 (page 7). First, consider the case that all functions from F are locally monotone. By Proposition 3.11, all functions from F have PL[DMor]-representations. Hence, by Proposition 3.10, $\mathrm{ML}[F]$ has polynomial translations in ML[DMor]. In the previous section, we derived representations of $\neg$ and $\vee$ (for complete sets of Boolean functions containing $\top$ and $\bot$), but not explicitly of $\wedge$. However, replacing any sub-formula of the form $\alpha \wedge \beta$ by $\neg(\neg\alpha \vee \neg\beta)$ defines a polynomial translation of ML[DMor] in $\mathrm{ML}[\{\top, \bot, \neg, \vee\}]$. Since the concatenation of polynomial translations is again a polynomial translation, it follows that $\mathrm{ML}[F]$ has polynomial translations in $\mathrm{ML}[\{\top, \bot, \neg, \vee\}]$.

Now consider $G^+ = G \cup \{\top, \bot\}$. This set of Boolean functions is complete since G is complete. By Propositions 3.13 and 3.14, the functions $\neg$ and $\vee$ have $\mathrm{PL}[G^+]$-representations. Again by Proposition 3.10, $\mathrm{ML}[\{\top, \bot, \neg, \vee\}]$ has polynomial translations in $\mathrm{ML}[G^+]$.

Since G is complete, there are formulae $\texttt{tt}, \texttt{ff} \in \mathrm{PL}[G]$ that are equivalent to $\top$ and $\bot$, respectively (note that these formulae may contain some variables). Replacing any sub-formula $\top$ and $\bot$ by $\texttt{tt}$ and $\texttt{ff}$ defines a polynomial translation of $\mathrm{ML}[G^+]$ in $\mathrm{ML}[G]$. Using once more that the concatenation of polynomial translations is a polynomial translation, we find a polynomial translation of $\mathrm{ML}[F]$ in $\mathrm{ML}[G]$ provided all functions from F are locally monotone and G is complete.

Now suppose that G contains some function that is not locally monotone. Then we can argue as above, using

- the logic $\mathrm{ML}[\mathrm{DMor}_{\leftrightarrow}]$ instead of the logic ML[DMor],
- the logic $\mathrm{ML}[\{\top, \bot, \neg, \vee, \leftrightarrow\}]$ instead of the logic $\mathrm{ML}[\{\top, \bot, \neg, \vee\}]$, and
- Proposition 3.15 in addition to the Propositions 3.13 and 3.14. □

Recall that the main result of this section was a simple corollary of Theorem 3.2. Hence we have shown that the class of logics $\mathrm{ML}[G]$ with G finite and complete has at most two succinctness classes, namely that of DMor and that of $\mathrm{DMor}_{\leftrightarrow} \supseteq \mathrm{DMor}$.

Now let $\mathcal{S}$ be any class of pointed Kripke structures. Since formulae equivalent wrt. $\mathcal{S}_{\mathrm{K}}$ are trivially equivalent wrt. $\mathcal{S}$, Corollary 3.3 also holds for the succinctness wrt. $\mathcal{S}$ – which is the restricted meaning of "all" in the title of this section. But as the reader will have realized, the above proof technique carries over to many other logics like multi-modal logic, predicate calculus, temporal logics etc. – which gives a wider (and imprecise) meaning to the term "all".

4. Modal logics with one and with two succinctness classes

So far, we saw that for many logics $\mathcal{L}$, there are at most two succinctness classes (up to a polynomial), namely those of $\mathcal{L}[\text{DMor}]$ and $\mathcal{L}[\text{DMor}_{\leftrightarrow}]$, respectively. In this section, we will show that these classes may differ in some modal logics, but coincide in others. Recall the following classes of pointed Kripke structures with the corresponding properties of their accessibility relations

- $\mathcal{S}_{\text{K}}$, the class of all pointed Kripke structures,
- $\mathcal{S}_{\text{T}}$, the class of reflexive pointed Kripke structures, and
- $\mathcal{S}_{\text{S5}}$, the class of equivalence relations.

More precisely, we show that the succinctness classes differ for the modal logics T (and hence also for K) but coincide for the modal logic S5. We will therefore, from now on, be precise and return to the original notation, i.e., write $\equiv_{\mathcal{S}_{\text{K}}}$ instead of $\equiv$ etc.

4.1. Two distinct succinctness classes.

Throughout this section, we will consider the class $\mathcal{S}_{\text{T}}$ of pointed Kripke structures where the accessibility relation R is reflexive. We will demonstrate that $\text{ML}[\text{DMor}_{\leftrightarrow}]$ does not have polynomial translations in $\text{ML}[\text{DMor}]$ with respect to the class $\mathcal{S}_{\text{T}}$. To this aim, we consider the following sequence of formulae from $\text{ML}[\text{DMor}_{\leftrightarrow}]$:

$$\varphi_0 = p_0 \wedge \Diamond\neg p_0 \quad \text{and} \quad \varphi_{n+1} = p_{(n+1) \bmod 2} \wedge (p \leftrightarrow \Diamond\varphi_n)\,.$$

Note that there is $c > 0$ such that $|\varphi_n| \leq c \cdot n$ for all $n \geq 1$. For all $n \in \mathbb{N}$, we will prove that all formulae $\psi \in \text{ML}[\text{DMor}]$ with $\varphi_n \equiv_{\mathcal{S}_{\text{T}}} \psi$ satisfy $|\psi| \geq 2^n$. Consequently, $\text{ML}[\text{DMor}_{\leftrightarrow}]$ does not have sub-exponential translations in $\text{ML}[\text{DMor}]$ – the two succinctness classes of $\text{ML}[\text{DMor}]$ and $\text{ML}[\text{DMor}_{\leftrightarrow}]$ are distinct.

Actually, we show a bit more: The formula ψ is not only "large", but it even contains exponentially many occurrences of the modal operator $\Diamond$. We denote this number of occurrences of $\Diamond$ in the formula ψ by $|\psi|_{\Diamond}$.

In the course of this proof, we need the following concept. Let $\psi \in \text{ML}[\text{DMor}]$. Then ψ is a Boolean combination of atomic formulae p_i, $\top$, and $\bot$ and of formulae of the form $\Diamond\lambda$ with $\lambda \in \text{ML}[\text{DMor}]$. The set E_ψ consists of all formulae λ such that $\Diamond\lambda$ appears in the Boolean combination ψ with even negation depth; the set O_ψ consists of all $\lambda \in \text{ML}[\text{DMor}]$ such that $\Diamond\lambda$ appears with odd negation depth. Inductively, these two sets are defined as follows:

$$E_\psi = \begin{cases} \emptyset & \text{if } \psi \in \{\top,\bot\} \cup \mathcal{V} \\ \{\lambda\} & \text{if } \psi = \Diamond\lambda, \lambda \in \text{ML}[\text{DMor}] \\ O_\alpha & \text{if } \psi = \neg\alpha \\ E_\alpha \cup E_\beta & \text{if } \psi \in \{\alpha \wedge \beta, \alpha \vee \beta\} \end{cases}$$

$$O_\psi = \begin{cases} \emptyset & \text{if } \psi \in \{\top,\bot\} \cup \mathcal{V} \\ \emptyset & \text{if } \psi = \Diamond\lambda, \lambda \in \text{ML}[\text{DMor}] \\ E_\alpha & \text{if } \psi = \neg\alpha \\ O_\alpha \cup O_\beta & \text{if } \psi \in \{\alpha \wedge \beta, \alpha \vee \beta\} \end{cases}$$

The use of these sets is described by the following lemma that can be proved by structural induction.

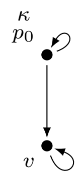

Figure 3: Schematic representation of the Kripke structure S_0. Vertices are labelled with the name of the world and the propositional variables holding there.

Lemma 4.1. *Let* $S, T \in \mathcal{S}_\mathrm{K}$ *and* $\psi \in \mathrm{ML}[\mathrm{DMor}]$ *such that the following properties hold:*

(i) $S \models p \iff T \models p$ *for all atomic formulae* p.
(ii) *If* $\lambda \in E_\psi$ *with* $S \models \Diamond\lambda$, *then* $T \models \Diamond\lambda$.
(iii) *If* $\lambda \in O_\psi$ *with* $T \models \Diamond\lambda$, *then* $S \models \Diamond\lambda$.

Then $S \models \psi$ *implies* $T \models \psi$.

Fix, for every formula $\alpha \in \mathrm{ML}[\mathrm{DMor}]$, a pointed Kripke structure $S_\alpha = (W_\alpha, R_\alpha, V_\alpha, \iota_\alpha)$ from $\mathcal{S}_\mathrm{T}$ such that $S_\alpha \models \alpha$ whenever α is satisfiable in $\mathcal{S}_\mathrm{T}$ (for unsatisfiable formulae α, fix an arbitrary structure S_α from $\mathcal{S}_\mathrm{T}$). We will always assume that these structures S_α are mutually disjoint.

We now come to the central result of this section.

Lemma 4.2. *For all* $n \in \mathbb{N}$ *and all formulae* $\psi \in \mathrm{ML}[\mathrm{DMor}]$ *with* $\psi \equiv_{\mathcal{S}_\mathrm{T}} \varphi_n$, *we have* $|\psi|_\Diamond \geq 2^n$.

Proof. Since we want to prove this lemma by induction on n, we first consider the case $n = 0$, i.e., the formula $\varphi_0 = p_0 \wedge \Diamond\neg p_0$. Let $\psi \in \mathrm{ML}[\mathrm{DMor}]$ with $\psi \equiv_{\mathcal{S}_\mathrm{T}} \varphi_0$. Towards a contradiction, assume $|\psi|_\Diamond < 2^n = 1$, i.e., assume ψ to be a Boolean combination of atomic formulae. Consider the Kripke structure $S_0 \in \mathcal{S}_\mathrm{T}$ from Fig. 3. Then $(S_0, \kappa) \models \varphi_0$ implies $(S_0, \kappa) \models \psi$ since φ_0 and ψ are assumed to be equivalent. Since ψ is a Boolean combination of atomic formulae, and since the worlds κ and ι satisfy the same atomic formulae, we obtain $(S_0, \iota) \models \psi$ and therefore $(S_0, \iota) \models \varphi_0$, which is not the case. Hence we showed $|\psi|_\Diamond \geq 1 = 2^0$, i.e., we verified the base case $n = 0$.

So let $n \geq 0$ and consider the formula $\varphi_{n+1} = p_{(n+1) \bmod 2} \wedge (p \leftrightarrow \Diamond\varphi_n)$. Let $\psi \in \mathrm{ML}[\mathrm{DMor}]$ be some formula such that $\psi \equiv_{\mathcal{S}_\mathrm{T}} \varphi_{n+1}$.

Aiming at a contradiction, assume $|\psi|_\Diamond < 2^{n+1}$. Although the sets E_ψ and O_ψ may have non-empty intersection, it follows that

$$\sum_{\lambda \in E_\psi} |\lambda|_\Diamond \; + \; \sum_{\lambda \in O_\psi} |\lambda|_\Diamond \leq |\psi|_\Diamond < 2^{n+1} \,.$$

Hence, for at least one of the sets E_ψ and O_ψ, the total number of modal operators must be less than 2^n; we consider these two cases separately.

Case 1, $\sum_{\lambda \in O_\psi} |\lambda|_\Diamond < 2^n$**.** Let O^*_ψ be the set of formulae λ from O_ψ with $\lambda \models_{\mathcal{S}_\mathrm{T}} \varphi_n$. Since $|\bigvee O^*_\psi|_\Diamond \leq \sum_{\lambda \in O_\psi} |\lambda|_\Diamond < 2^n$, the induction hypothesis ensures $\bigvee O^*_\psi \not\equiv_{\mathcal{S}_\mathrm{T}} \varphi_n$. On the other hand, by choice of O^*_ψ, we have $\bigvee O^*_\psi \models_{\mathcal{S}_\mathrm{T}} \varphi_n$ (this holds also in case $O^*_\psi = \emptyset$ since $\bigvee \emptyset \equiv_{\mathcal{S}_\mathrm{K}} \bot$). Hence $\varphi_n \not\models_{\mathcal{S}_\mathrm{T}} \bigvee O^*_\psi$ implying that the formula $\alpha = \varphi_n \wedge \neg \bigvee O^*_\psi$ is satisfiable

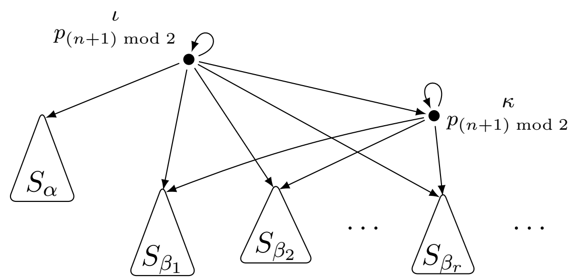

Figure 4: Schematic representation of the Kripke structure S.

in $\mathcal{S}_\mathrm{T}$. The choice of the structure S_α implies $S_\alpha \models \alpha$, i.e., $S_\alpha \models \varphi_n$ but $S_\alpha \not\models \bigvee O^*_\psi$. Finally, let B denote the set of formulae β with $S_\beta \models \beta \wedge \neg\varphi_n$.

We now define a Kripke structure $S = (W, R, V)$ as follows (cf. Fig. 4):

$$W = \{\iota, \kappa\} \uplus \bigcup_{\beta \in B} W_\beta \cup W_\alpha$$

$$R = \big\{(\iota,\iota), (\kappa,\kappa), (\iota,\kappa), (\iota,\iota_\alpha)\big\} \cup \Big(\{\iota,\kappa\} \times \{\iota_\beta \mid \beta \in B\}\Big) \cup \bigcup_{\beta \in B} R_\beta \cup R_\alpha$$

$$V(q) = \begin{cases} \bigcup_{\beta \in B} V_\beta(q) \cup V_\alpha(q) \cup \{\iota, \kappa\} & \text{if } q = p_{(n+1) \bmod 2} \\ \bigcup_{\beta \in B} V_\beta(q) \cup V_\alpha(q) & \text{otherwise} \end{cases}$$

Since the accessibility relations of the structures S_α and S_β for $\beta \in H$ are reflexive, the same applies to the accessibility relation of S, i.e., we obtain $S \in \mathcal{S}_\mathrm{T}$. Furthermore,

$$\text{for all } \lambda \in \mathrm{ML}[\mathrm{DMor}], \gamma \in B \cup \{\alpha\}, w \in W_\gamma \colon \Big((S_\gamma, w) \models \lambda \iff (S, w) \models \lambda\Big) \qquad (4.1)$$

since we only add edges originating in ι or κ.

We will proceed by verifying the following claims:

(A) $(S, \iota) \not\models \varphi_{n+1}$,
(B) $(S, \kappa) \models \varphi_{n+1}$,
(C) $(S, \kappa) \models \psi$, and
(D) $(S, \iota) \models \psi$,

thus showing $(S, \iota) \models \neg\varphi_{n+1} \wedge \psi$, which contradicts the equivalence of φ_{n+1} and ψ.

Proof of (A), $(S, \iota) \not\models \varphi_{n+1}$**.** Recall that $(S_\alpha, \iota_\alpha) \models \varphi_n$ implying $(S, \iota_\alpha) \models \varphi_n$ by (4.1). Hence we have $(S, \iota) \models \neg p \wedge \Diamond\varphi_n$ which ensures $(S, \iota) \not\models \varphi_{n+1}$.

Proof of (B), $(S, \kappa) \models \varphi_{n+1}$**.** Towards a contradiction, suppose $(S, \kappa) \models \Diamond\varphi_n$. Then there exists $w \in W$ with $(S, w) \models \varphi_n$ and $(\kappa, w) \in R$, i.e., $w \in \{\kappa\} \cup \{\iota_\beta \mid \beta \in B\}$. From $\varphi_n \models_{\mathcal{S}_\mathrm{K}} p_{n \bmod 2}$ and $(S, \kappa) \not\models p_{n \bmod 2}$, we get $(S, \kappa) \models \neg\varphi_n$ and therefore $w \neq \kappa$. Hence there is $\beta \in B$ with $w = \iota_\beta$. Now $(S, w) \models \varphi_n$ implies $(S_\beta, w) \models \varphi_n$ by (4.1). But this contradicts $\beta \in B$. Thus, indeed, $(S, \kappa) \not\models \Diamond\varphi_n$. Together with the observation $(S, \kappa) \models p_{(n+1) \bmod 2} \wedge \neg p$, we get $(S, \kappa) \models \varphi_{n+1}$.

Proof of (C), $(S, \kappa) \models \psi$**.** This is immediate by the above since $S \in \mathcal{S}_\mathrm{T}$ and $\psi \equiv_{\mathcal{S}_\mathrm{T}} \varphi_{n+1}$.

Proof of (D), $(S,\iota) \models \psi$. From $(S,\kappa) \models \psi$, we will now infer that ψ holds in the world ι, too. Recall that ψ is a Boolean combination of atomic formulae and of formulae $\Diamond\lambda$ with $\lambda \in O_\psi \cup E_\psi$. Note that (S,ι) and (S,κ) agree in the atomic formulae holding there. In view of Lemma 4.1, it therefore suffices to show the following:

(D1) If $\lambda \in E_\psi$ with $(S,\kappa) \models \Diamond\lambda$, then $(S,\iota) \models \Diamond\lambda$.
(D2) If $\lambda \in O_\psi$ with $(S,\iota) \models \Diamond\lambda$, then $(S,\kappa) \models \Diamond\lambda$.

Proof of (D1). Let $\lambda \in E_\psi$ with $(S,\kappa) \models \Diamond\lambda$. Then there is $w \in \{\kappa\} \cup \{\iota_\beta \mid \beta \in B\}$ such that $(S,w) \models \lambda$. In any case, $(\iota,w) \in R$. Hence we have $(S,\iota) \models \Diamond\lambda$.

Proof of (D2). Let $\lambda \in O_\psi$ with $(S,\iota) \models \Diamond\lambda$.

First, assume $\lambda \models_{\mathcal{S}_\mathrm{T}} \varphi_n$, i.e., $\lambda \in O^*_\psi$. From $(S,\iota) \models \Diamond\lambda$, we obtain that the formula λ holds in one of the worlds ι, κ, ι_α, or ι_β for some $\beta \in B$. But $(S,\iota_\alpha) \models \lambda$ implies $S_\alpha \models \lambda$ by (4.1), which is impossible since $S_\alpha \models \alpha$, $\alpha \models_{\mathcal{S}_\mathrm{K}} \neg \bigvee O^*_\psi$, and $\neg \bigvee O^*_\psi \models_{\mathcal{S}_\mathrm{K}} \neg\lambda$ since $\lambda \in O^*_\psi$. Next suppose $(S,\iota_\beta) \models \lambda$ for some $\beta \in B$. Then $S_\beta \models \lambda$ which, together with $\lambda \models_{\mathcal{S}_\mathrm{T}} \varphi_n$, implies $S_\beta \models \varphi_n$. Because of $\beta \in B$, we have $S_\beta \models \neg\varphi_n$, a contradiction.

Consequently, the formula λ holds in one of the worlds ι and κ. Again using $\lambda \models_{\mathcal{S}_\mathrm{T}} \varphi_n$, we obtain that also φ_n holds in ι or in κ. But this cannot be the case since $p_{n \bmod 2}$ does not hold in either of the two worlds – a contradiction.

Consequently, we have $\lambda \not\models_{\mathcal{S}_\mathrm{T}} \varphi_n$. But then the formula $\beta := (\lambda \wedge \neg\varphi_n)$ is satisfiable in $\mathcal{S}_\mathrm{T}$ implying $S_\beta \models \beta$. From $\beta \models_{\mathcal{S}_\mathrm{K}} \lambda$ and $\beta \models_{\mathcal{S}_\mathrm{K}} \neg\varphi_n$, we obtain $S_\beta \models \beta \wedge \neg\varphi_n \wedge \lambda$. Hence $\beta \in B$. Now $S_\beta \models \lambda$ implies $(S,\iota_\beta) \models \lambda$ by (4.1) and therefore $(S,\kappa) \models \Diamond\lambda$.

This finishes the proof of the claims (D1) and (D2). By Lemma 4.1 and $(S,\kappa) \models \psi$, we obtain $(S,\iota) \models \psi$.

Conclusion of Case 1. Through steps (A) to (D), we proved $(S,\iota) \models \neg\varphi_{n+1} \wedge \psi$ in case $\left|\bigvee O_\psi\right| < 2^n$, contradicting the equivalence of φ_{n+1} and ψ.

Case 2, $\sum_{\lambda \in E_\psi} |\lambda|_\Diamond < 2^n$. We consider the formula $\neg p_{(n+1) \bmod 2} \vee \big((\neg p) \leftrightarrow \Diamond\varphi_n\big) \equiv_{\mathcal{S}_\mathrm{K}} \neg\varphi_{n+1} \equiv_{\mathcal{S}_\mathrm{T}} \neg\psi$. Observe that $O_{\neg\psi} = E_\psi$, such that $|\bigvee O_{\neg\psi}| < 2^n$. Hence we can use the same argument as before for $\neg\psi$ to obtain a contradiction: simply force p to hold in the two worlds ι and κ. We provide the details in the appendix.

Having derived contradictions in both case 1 and case 2, it follows that $|\psi|_\Diamond < 2^{n+1}$ is not possible. But this finishes the inductive proof. □

From the above lemma, we immediately get that ML[DMor$_\leftrightarrow$] does not have sub-exponential translations wrt. $\mathcal{S}_\mathrm{T}$ in ML[DMor]. Now let G be any finite and complete set of Boolean functions such that at least one function from G is not locally monotone. Then, by Theorem 3.2, the logic ML[DMor$_\leftrightarrow$] has polynomial translations wrt. $\mathcal{S}_\mathrm{T}$ in ML[G]. Since the relation "has polynomial translations wrt. $\mathcal{S}_\mathrm{T}$" is transitive, the above lemma ensures that ML[G] does not have polynomial translations wrt. $\mathcal{S}_\mathrm{T}$ in ML[DMor] (for otherwise ML[DMor$_\leftrightarrow$] would have such translations).

The class of sub-exponential functions is not closed under composition (e.g., the functions $n \mapsto n^2$ and $n \mapsto 2^{\sqrt{n}}$ are sub-exponential, but their composition is $n \mapsto 2^n$). Hence we cannot use the above argument to show that ML[G] does not have sub-exponential translations wrt. $\mathcal{S}_\mathrm{T}$ in ML[DMor]. To show this stronger result, we adopt the above proof as follows.

Let $G^+ = G \cup \{\top, \bot\}$. From Propositions 3.13, 3.14, and 3.15, we obtain PL[G^+]-formulae $\texttt{iff}(x,y) \equiv_{\mathcal{S}_\mathrm{K}} x \leftrightarrow y$, $\texttt{and}(x,y) \equiv_{\mathcal{S}_\mathrm{K}} x \wedge y$, and $\texttt{neg}(y) \equiv_{\mathcal{S}_\mathrm{K}} \neg y$, each using the variable y only once. We can therefore express the ML[DMor$_\leftrightarrow$]-formulae φ_n in a natural way by equivalent ML[G^+]-formulae $\tilde{\varphi}_n$:

$$\tilde{\varphi}_0 = \texttt{and}\Big(\ p_0,\ \Diamond\texttt{neg}(p_0)\ \Big) \qquad \text{and}$$
$$\tilde{\varphi}_{n+1} = \texttt{and}\Big(\ p_{(n+1) \bmod 2},\ \texttt{iff}\big(p, \Diamond\varphi_n\big)\ \Big) \qquad \text{for } n \geq 0\,.$$

By choice of iff and and, the formula $\Diamond\varphi_n$ occurs only once in $\tilde{\varphi}_{n+1}$. Since the formulae iff, and, and neg remain fixed, the size of $\tilde{\varphi}_n$ is linear in n, as is the size of φ_n. Note that the formula $\tilde{\varphi}_n$ can contain the formulae $\top, \bot \in G^+$ which need not belong to G. Since G is complete however, we can express them by some fixed ML[G]-formulae possibly using a new propositional variable z. Let $\overline{\varphi_n}$ denote the result of these replacements. Then the size of $\overline{\varphi_n}$ is linear in that of $\tilde{\varphi}_n$ and therefore in n. Furthermore, $\overline{\varphi_n} \equiv_{\mathcal{S}_\mathrm{K}} \tilde{\varphi}_n \equiv_{\mathcal{S}_\mathrm{K}} \varphi_n$. Now the proof of Lemma 4.2 works for ML[G] and we obtain the main result of the section.

Theorem 4.3. *Let G be a finite and complete set of Boolean functions with some function that is not locally monotone. Then the logic* ML[G] *does not have sub-exponential translations wrt. $\mathcal{S}_\mathrm{T}$ in* ML[DMor].

4.2. **Just one succinctness class.** In this section we show that, wrt. the modal logic S5, there is only one succinctness class. Recall that $\mathcal{S}_\mathrm{S5}$ is the class of pointed Kripke structures whose accessibility is an equivalence relation.

Theorem 4.4. *Let F be a finite set of Boolean functions. Then* ML[F] *has polynomial translations wrt. $\mathcal{S}_\mathrm{S5}$ in* ML[DMor].

In view of Theorem 3.2, it suffices to show the claim for F = DMor$_\leftrightarrow$, i.e., that ML[DMor$_\leftrightarrow$] has polynomial translations wrt. $\mathcal{S}_\mathrm{S5}$ in ML[DMor] (for more details, we refer to the proof of Theorem 4.4 on page 26). Hence we will, for the major part of the section, consider the (extended) De Morgan basis.

Let $\varphi \in$ ML[DMor$_\leftrightarrow$]. Recall that $|\varphi|$ denotes the number of nodes in the syntax tree of φ. In this section, we use the following additional size measures of φ and its syntax tree:

- the *norm* $\|\varphi\|$ of φ, denoting the number of leaves in the syntax tree, i.e., the total number of occurrences of propositional variables and constants $\top$ and $\bot$, as well as
- the *depth* $d(\varphi)$ of φ, denoting the depth of the syntax tree, where $d(\varphi) = 0$ if, and only if, the tree consists of a single leaf, i.e., if $\varphi \in \mathcal{V} \cup \{\top, \bot\}$.

Since the arity of all functions that occur in φ is bounded by 2, the number of leaves, the size, and the depth of φ satisfy $1 \leq \|\varphi\| \leq |\varphi| \leq 2^{d(\varphi)+1} - 1$.

We first show that, wrt. $\mathcal{S}_\mathrm{S5}$, formulae can be balanced, i.e., that any formula in ML[DMor$_\leftrightarrow$] can be rewritten so as to have logarithmic depth in the norm of the initial formula. A similar result is already known for PL (see Remark 4.6).

Lemma 4.5. *For every $\varphi \in$ ML[DMor$_\leftrightarrow$] there exists a formula $\varphi' \in$ ML[DMor$_\leftrightarrow$] with $\varphi' \equiv_{\mathcal{S}_\mathrm{S5}} \varphi$ and $d(\varphi') \leq 8 \cdot (1 + \log_2 \|\varphi\|)$.*

Remark 4.6. Since PL[DMor$_\leftrightarrow$] $\subseteq$ ML[DMor$_\leftrightarrow$], the lemma implies that each PL[DMor$_\leftrightarrow$]-formula φ is equivalent to some PL[DMor$_\leftrightarrow$]-formula of depth logarithmic in $\|\varphi\|$. According to Gashkov and Sergeev [GS20], a more general form of this result for propositional logic was known to Khrapchenko in 1967, namely that it holds for any complete basis of Boolean functions (e.g., for $\{\wedge, \vee, \neg, \leftrightarrow\}$ as here). They also express their regret that the only source for this is a single paragraph in a survey article by Yablonskii and Kozyrev [YK68], see also [Juk12]. Often, it is referred to as Spira's theorem who published it in 1971 [Spi71], assuming that all at most binary Boolean functions are allowed in propositional formulae. Khrapchenko's general form was then published by Savage [Sav76].

Proof. Differently from other results and proofs in this paper, this lemma is only concerned with one logic, namely ML[$\mathrm{DMor}_{\leftrightarrow}$]. We will therefore simply speak of 'formulae' when actually referring to ML[$\mathrm{DMor}_{\leftrightarrow}$]-formulae.

The proof proceeds by induction on the number of leaves $\|\varphi\|$ in the syntax tree of the formula φ.

First, suppose that $\|\varphi\| = 1$. Then there is an integer $r \geq 0$, functions $f_1, \ldots, f_r \in \{\Diamond, \neg\}$, and an atomic formula $\lambda \in \mathcal{V} \cup \{\top, \bot\}$ such that $\varphi = f_1\langle f_2\langle \cdots f_r\langle\lambda\rangle \cdots\rangle\rangle$ – that is, φ is a sequence of negations and modal operators, terminating with the atomic formula λ. It is well known that, in $\mathcal{S}_{\mathrm{S5}}$, any such formula is equivalent to a sequence of at most three negations and/or modal operators that ends with λ, i.e., to a formula of depth at most $3 \leq 8 \cdot (1 + \log_2 \|\varphi\|)$ (see, e.g., Section 4.4 in [Che80]). This establishes the case $\|\varphi\| = 1$.

Otherwise, $\|\varphi\| \geq 2$ and φ contains at least one of the binary functions $\wedge$, $\vee$, and $\leftrightarrow$. Let $m = \|\varphi\|$. Intuitively, we split the formula φ into two parts, each containing about half the leaves from φ. Formally, there are formulae $\alpha = \alpha(x)$, with only one occurrence of x, and β such that

- $\varphi = \alpha(\beta)$,
- $\|\alpha\| \leq \frac{m}{2} < \|\beta\|$, and
- $\beta = f\langle\beta_1, \beta_2\rangle$ with $f \in \{\wedge, \vee, \leftrightarrow\}$ and $\|\beta_1\|, \|\beta_2\| \leq \frac{m}{2}$.

It is not difficult to find such formulae α and β: simply start at the root of the syntax tree of φ and, while possible, proceed towards the child that contains more than half the leaves of φ. The procedure stops in a node, corresponding to a sub-formula β of φ, which contains more than $\frac{\|\varphi\|}{2}$ leaves, but for which each child contains less than $\frac{\|\varphi\|}{2}$ leaves. In particular, there are two children, hence $\beta = f\langle\beta_1, \beta_2\rangle$ for some $f \in \{\wedge, \vee, \leftrightarrow\}$. Substituting this particular sub-formula β in φ by x results in the formula $\alpha(x)$.

Note that $\|\alpha\| = m - \|\beta\| + 1 < m - \frac{m}{2} + 1 = \frac{m}{2} + 1$, i.e., $\|\alpha\| \leq \frac{m}{2}$.

First assume that x does not occur within the scope of a $\Diamond$-operator in $\alpha(x)$. Then β is interpreted in the initial world implying

$$\varphi = \alpha(\beta) \equiv_{\mathcal{S}_{\mathrm{S5}}} \big\langle \alpha(\bot) \wedge \neg\beta \big\rangle \vee \big\langle \alpha(\top) \wedge \beta \big\rangle\,. \tag{4.2}$$

By induction hypothesis, there exist formulae $\alpha'(x) \equiv_{\mathcal{S}_{\mathrm{S5}}} \alpha(x)$, $\beta_1' \equiv_{\mathcal{S}_{\mathrm{S5}}} \beta_1$, and $\beta_2' \equiv_{\mathcal{S}_{\mathrm{S5}}} \beta_2$ of depth $\leq 8 \cdot (1 + \log_2 \frac{m}{2}) = 8 \cdot \log_2 m$. Set $\beta' = f\langle\beta_1', \beta_2'\rangle$. Then β' is equivalent to β over $\mathcal{S}_{\mathrm{S5}}$ and of depth $\leq 1 + 8 \cdot \log_2 m$. Together with (4.2), it follows that

$$\varphi \equiv_{\mathcal{S}_{\mathrm{S5}}} \big\langle \alpha'(\bot) \wedge \neg\beta' \big\rangle \vee \big\langle \alpha'(\top) \wedge \beta' \big\rangle\,.$$

Let φ' denote the formula on the right hand side of this equivalence. Then the depth of φ' satisfies

$$\begin{aligned} d(\varphi') &= \max\big\{2 + d\big(\alpha'\big),\ 3 + d\big(\beta'\big)\big\} \\ &\leq \max\{2 + 8 \cdot \log_2 m,\ 3 + 1 + 8 \cdot \log_2 m\} \leq 4 + 8 \cdot \log_2 m \\ &\leq 8 \cdot (1 + \log_2 m)\,. \end{aligned}$$

This completes the first case, where x does not occur within the scope of a $\Diamond$-operator.

Otherwise, we split $\alpha(x)$ at the last $\Diamond$-operator above x, i.e., there are formulae $\alpha_1(y)$ with only one occurrence of y and $\alpha_2(x)$ with only one occurrence of x that, furthermore, does not lie in the scope of a $\Diamond$-operator, such that $\alpha(x) = \alpha_1(\Diamond\alpha_2(x))$. In particular, $\|\alpha_1\|, \|\alpha_2\| \leq \|\alpha\| \leq \frac{m}{2}$. Hence, by induction hypothesis, there are formulae $\alpha'_1(y), \alpha'_2(x), \beta'_1$, and β'_2 with

- $\alpha'_i(z) \equiv_{\mathcal{S}_{\mathrm{S5}}} \alpha_i(z)$ and $d(\alpha'_i) \leq 8 \cdot \log_2 m$ for $i \in \{1, 2\}$, as well as
- $\beta'_i \equiv_{\mathcal{S}_{\mathrm{S5}}} \beta_i$ and $d(\beta'_i) \leq 8 \cdot \log_2 m$ for $i \in \{1, 2\}$.

As before, let $\beta' = f\langle\beta'_1, \beta'_2\rangle$ with $\beta' \equiv_{\mathcal{S}_{\mathrm{S5}}} \beta$ and $d(\beta') \leq 1 + 8 \cdot \log_2 m$.

Recall that, for a structure $K \in \mathcal{S}_{\mathrm{S5}}$, the accessibility relation R is an equivalence relation. Hence, for any two worlds v and w within the same equivalence class and any formula χ, we have

$$K, v \models \Diamond\chi \iff K, w \models \Diamond\chi\,, \tag{4.3}$$

i.e., whether or not $\Diamond\chi$ holds in K does not depend on the choice of the world (when considering only the relevant equivalence class containing the initial world).

In $\alpha_2(\beta)$, the sub-formula β is interpreted in the same world that the whole formula is interpreted in. With

$$\psi' = \big\langle \alpha'_2(\bot) \wedge \neg\beta' \big\rangle \vee \big\langle \alpha'_2(\top) \wedge \beta' \big\rangle \equiv_{\mathcal{S}_{\mathrm{S5}}} \big\langle \alpha_2(\bot) \wedge \neg\beta \big\rangle \vee \big\langle \alpha_2(\top) \wedge \beta \big\rangle$$

we therefore get $\alpha_2(\beta) \equiv_{\mathcal{S}_{\mathrm{S5}}} \psi'$. Setting

$$\varphi' = \big\langle \alpha'_1(\bot) \wedge \neg\Diamond\psi' \big\rangle \vee \big\langle \alpha'_1(\top) \wedge \Diamond\psi' \big\rangle\,,$$

we furthermore get

$$\begin{aligned} \varphi' &\equiv_{\mathcal{S}_{\mathrm{S5}}} \big\langle \alpha_1(\bot) \wedge \neg\Diamond\psi' \big\rangle \vee \big\langle \alpha_1(\top) \wedge \Diamond\psi' \big\rangle \\ &\equiv_{\mathcal{S}_{\mathrm{S5}}} \alpha_1\big(\Diamond\psi'\big) && \text{by (4.3)} \\ &\equiv_{\mathcal{S}_{\mathrm{S5}}} \alpha_1\big(\Diamond\alpha_2(\beta)\big) \\ &= \varphi\,. \end{aligned}$$

Then

$$\begin{aligned} d(\psi') &= \max\big\{2 + d\big(\alpha'_2(x)\big),\ 3 + d\big(\beta'\big)\big\} \leq 4 + 8 \cdot \log_2 m \quad \text{and} \\ d(\varphi') &= \max\big\{2 + d\big(\alpha'_1(y)\big),\ 4 + d\big(\psi'\big)\big\} \\ &\leq \max\{2 + 8 \cdot \log_2 m,\ 8 + 8 \cdot \log_2 m\} \leq 8 \cdot (1 + \log_2 m)\,, \end{aligned}$$

which completes the second case, where x occurs in the scope of a $\Diamond$-operator, and hence the inductive proof. □

Let $\varphi \in \mathrm{ML}[\mathrm{DMor}_{\leftrightarrow}] \setminus \mathrm{ML}[\mathrm{DMor}]$. Since φ contains at least one occurrence of bi-implication, we obtain $\|\varphi\| \geq 2$. By the previous lemma, there exists an $\mathrm{ML}[\mathrm{DMor}_{\leftrightarrow}]$-formula ψ that is equivalent to φ over $\mathcal{S}_{\mathrm{S5}}$ and of depth logarithmic in the norm of φ. Let φ' be obtained from ψ by replacing each sub-formula $\alpha \leftrightarrow \beta$ by $\langle\alpha \wedge \beta\rangle \vee \langle\neg\alpha \wedge \neg\beta\rangle$. Then $\varphi' \in \mathrm{ML}[\mathrm{DMor}]$ and $\varphi' \equiv_{\mathcal{S}_{\mathrm{S5}}} \varphi$. Furthermore, the depth of φ' increases at most by a factor of three and therefore remains logarithmic in $\|\varphi\|$; more precisely, $d(\varphi') \leq 3 \cdot d(\psi) \leq 3 \cdot 8 \cdot (1 + \log_2 \|\varphi\|) = 24 \cdot 2 \cdot \log_2 \|\varphi\|$ since $\|\varphi\| \geq 2$. Since all functions in $\mathrm{DMor}_{\leftrightarrow}$ are at most binary, it follows that

$$|\varphi'| \leq 2^{d(\varphi')+1} \leq 2^{48 \cdot \log_2 \|\varphi\| + 1} \leq c \cdot \|\varphi\|^d \leq c \cdot |\varphi|^d \quad \text{for some constants } c, d > 0.$$

Hence, we verified the following claim.

Lemma 4.7. ML[$\mathrm{DMor}_{\leftrightarrow}$] *has polynomial translations wrt.* $\mathcal{S}_{\mathrm{S5}}$ *in* ML[DMor].

Therefore, the modal logic S5 has only one "succinctness class", as stated in Theorem 4.4.

Proof of Theorem 4.4. Let F be a finite set of Boolean functions. Since bi-implication $\leftrightarrow$ is not locally monotone, it follows by Theorem 3.2 that ML[F] has polynomial translations (wrt. $\mathcal{S}_{\mathrm{K}}$) in ML[$\mathrm{DMor}_{\leftrightarrow}$] and hence, particularly, also wrt. $\mathcal{S}_{\mathrm{S5}}$. By Lemma 4.7 above, ML[$\mathrm{DMor}_{\leftrightarrow}$] again has polynomial translations wrt. $\mathcal{S}_{\mathrm{S5}}$ in ML[DMor]. Since the relation "has polynomial translations" is transitive, this establishes the claim. □

5. Conclusion

This paper considers the impact of the choice of a complete Boolean basis on the succinctness of logical formulae. It is show that the succinctness depends solely on the existence of some non-locally monotone operator in the basis: if there is no such function, then the logical formulae using this basis agree in succinctness (up to a polynomial) with those using the De Morgan basis, otherwise, they agree with the formulae using the extension of the De Morgan basis with bi-implication.

While this result is demonstrated for modal logic, the proof carries over to many other classical logics like first-order logic, temporal logic etc.

Regarding propositional logic, it was known before that all Boolean bases give rise to the same succinctness. We show the same for the modal logic S5, i.e., when we restrict to Kripke structures whose accessibility is an equivalence relation. When considering all reflexive Kripke structures (i.e., the modal logic T) however, this is no longer the case since then, the extended De Morgan basis allows formulae to be written exponentially more succinct than the plain De Morgan basis does.

It remains open, where exactly this dichotomy occurs, namely whether it holds when considering reflexive and transitive or reflexive and symmetric accessibility relations. Furthermore, it is not known whether the dichotomy holds for other logics like first-order or temporal logic.

References

[AI03] M. Adler and N. Immerman. An *n!* lower bound on formula size. *ACM Trans. Comput. Log.*, 4(3):296–314, 2003.

[BKS24] Ch. Berkholz, D. Kuske, and Ch. Schwarz. Modal logic is more succinct iff bi-implication is available in some form. In *STACS'24*, Leibniz International Proceedings in Informatics (LIPIcs) vol. 289, pages 12:1–12:17. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2024.

[CFI+24] M. Carmosino, R. Fagin, N. Immerman, P. Kolaitis, J. Lenchner, R. Sengupta, and R. Williams. Parallel play saves quantifiers. *CoRR*, abs/2402.10293, 2024.

[Che80] B.F. Chellas. *Modal Logic: An Introduction.* Cambridge University Press, 1980.

[EVW02] K. Etessami, M. Vardi, and T. Wilke. First-order logic with two variables and unary temporal logic. *Inf. Comput.*, 179(2):279–295, 2002.

[FLRV21] R. Fagin, J. Lenchner, K. W. Regan, and N. Vyas. Multi-structural games and number of quantifiers. In *LICS'21*, pages 1–13. IEEE, 2021.

[FLVW22] R. Fagin, J. Lenchner, N. Vyas, and R. Williams. On the number of quantifiers as a complexity measure. In *MFCS'22*, volume 241 of *LIPIcs*, pages 48:1–48:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.

[FvdHIK13] T. French, W. van der Hoek, P. Iliev, and B.P. Kooi. On the succinctness of some modal logics. *Artif. Intell.*, 197:56–85, 2013.

[GS05] M. Grohe and N. Schweikardt. The succinctness of first-order logic on linear orders. *Log. Methods Comput. Sci.*, 1(1), 2005.

[GS20] S.B. Gashkov and I.S. Sergeev. О значении работ В. М. Храпченко. *Прикладная дискретная математика*, 2020(48):109–124, 2020. doi:10.17223/20710410/48/10.

[HL24] L. Hella and K. Luosto. Game characterizations for the number of quantifiers. *Mathematical Structures in Computer Science*, page 1–20, 2024.

[HV19] L. Hella and M. Vilander. Formula size games for modal logic and μ-calculus. *J. Log. Comput.*, 29(8):1311–1344, 2019.

[Imm81] N. Immerman. Number of quantifiers is better than number of tape cells. *J. Comput. Syst. Sci.*, 22(3):384–406, 1981.

[Juk12] S. Jukna. *Boolean Function Complexity - Advances and Frontiers*, volume 27 of *Algorithms and combinatorics*. Springer, 2012.

[LSW01] C. Lutz, U. Sattler, and F. Wolter. Modal logic and the two-variable fragment. In *CSL'01*, Lecture Notes in Computer Science vol. 2142, pages 247–261. Springer, 2001.

[Lut06] C Lutz. Complexity and succinctness of public announcement logic. In *AAMAS'06*, pages 137–143. ACM, 2006.

[Mar03] N. Markey. Temporal logic with past is exponentially more succinct. *Bull. EATCS*, 79:122–128, 2003.

[PM90] P.J. Pelletier and N.M. Martin. Post's functional completeness theorem. *Notre Dame Journal of Formal Logic*, 31(2):462–475, 1990.

[Pra75] V.R. Pratt. The effect of basis on size of Boolean expressions. In *16th Annual Symposium on Foundations of Computer Science*, pages 119–121, 1975. doi:10.1109/SFCS.1975.29.

[Sav76] J.E. Savage. *The Complexity of Computing*. Wiley, 1976.

[Spi71] P.M. Spira. On time-hardware tradeoffs for Boolean functions. In *Proc. of 4th Hawaii Intern. Symp. on System Sciences*, pages 525–527, 1971.

[Vin24] H. Vinall-Smeeth. From quantifier depth to quantifier number: Separating structures with k variables. In *LICS'24*, pages 71:1–71:14. ACM, 2024.

[DFH+14] H. van Ditmarsch, Jie Fan, W. van der Hoek, and P. Iliev. Some exponential lower bounds on formula-size in modal logic. In Rajeev Goré, Barteld P. Kooi, and Agi Kurucz, editors, *Advances in Modal Logic 2014*, pages 139–157. College Publications, 2014.

[Wil99] T. Wilke. CTL$^+$ is exponentially more succinct than CTL. In *FSTTCS'99*, Lecture Notes in Computer Science vol. 1738, pages 110–121. Springer, 1999.

[YK68] S.V. Yablonskii and V.P. Kozyrev. Математические бопросы кибернетики. *Информационные материалы, Академия наук СССР научный совет по комплексной проблеме "кибернетика"*, 19a:3–15, 1968.

Appendix A. Appendix

In the induction step of the proof of Lemma 4.2, we only spelled out one case and said that the second one is handled similarly. Here, we provide the details of this missing case.

Case 2, $\sum_{\lambda\in E_\psi} |\lambda|_\Diamond < 2^n$. Let E^*_ψ be the set of formulae λ from E_ψ with $\lambda \models_{\mathcal{S}_\mathrm{T}} \varphi_n$. Since $|\bigvee E^*_\psi|_\Diamond \leq \sum_{\lambda\in E_\psi} |\lambda|_\Diamond < 2^n$, the induction hypothesis ensures $\bigvee E^*_\psi \not\equiv_{\mathcal{S}_\mathrm{T}} \varphi_n$. On the other hand, $\bigvee E^*_\psi \models_{\mathcal{S}_\mathrm{T}} \varphi_n$ by choice of E^*_ψ. Hence the formula $\alpha = \varphi_n \wedge \neg \bigvee E^*_\psi$ is satisfiable in $\mathcal{S}_\mathrm{T}$ implying $S_\alpha \models \alpha$, i.e., $S_\alpha \models \varphi_n$ but $S_\alpha \not\models \bigvee E^*_\psi$.

Let B denote the set of formulae β with $S_\beta \models \beta \wedge \neg\varphi_n$.

We now define a Kripke structure $S = (W, R, V)$ as follows (cf. Fig. 5):

$$W = \{\iota, \kappa\} \uplus \bigcup_{\beta\in B} W_\beta \cup W_\alpha$$

$$R = \big\{(\iota,\iota), (\kappa,\kappa), (\iota,\kappa), (\iota,\iota_\alpha)\big\} \cup \Big(\{\iota,\kappa\} \times \{\iota_\beta \mid \beta \in B\}\Big) \cup \bigcup_{\beta\in B} R_\beta \cup R_\alpha$$

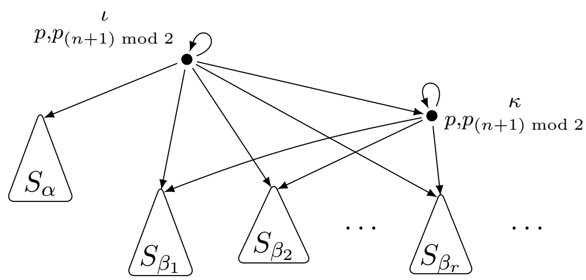

Figure 5: Schematic representation of the Kripke structure S.

$$V(q) = \begin{cases} \bigcup_{\beta\in B} V_\beta(q) \cup V_\alpha(q) \cup \{\iota,\kappa\} & \text{if } q \in \{p, p_{(n+1) \bmod 2}\} \\ \bigcup_{\beta\in B} V_\beta(q) \cup V_\alpha(q) & \text{otherwise} \end{cases}$$

Since the accessibility relations of the structures S_α and S_β for $\beta \in H$ are reflexive, the same applies to the accessibility relation of S, i.e., we obtain $S \in \mathcal{S}_{\mathrm{T}}$. Furthermore,

$$\text{for all } \lambda \in \mathrm{ML[DMor]}, \gamma \in B \cup \{\alpha\}, w \in W_\gamma\colon \Big((S_\gamma, w) \models \lambda \iff (S, w) \models \lambda\Big) \tag{A.1}$$

since we only add edges originating in ι or κ.

We will proceed by verifying the following claims:

(A) $(S,\iota) \models \varphi_{n+1}$,
(B) $(S,\kappa) \not\models \varphi_{n+1}$,
(C) $(S,\kappa) \not\models \psi$, and
(D) $(S,\iota) \not\models \psi$,

thus showing $(S,\iota) \models \varphi_{n+1} \wedge \neg\psi$, which contradicts the equivalence of φ_{n+1} and ψ.

Proof of (A), $(S,\iota) \models \varphi_{n+1}$**.** Recall that $(S_\alpha, \iota_\alpha) \models \varphi_n$ implying $(S, \iota_\alpha) \models \varphi_n$ by (A.1). Hence we have $(S,\iota) \models p \wedge \Diamond\varphi_n$ which ensures $(S,\iota) \models \varphi_{n+1}$.

Proof of (B), $(S,\kappa) \not\models \varphi_{n+1}$**.** Towards a contradiction, suppose $(S,\kappa) \models \Diamond\varphi_n$. Then there exists $w \in W$ with $(S,w) \models \varphi_n$ and $(\kappa, w) \in R$, i.e., $w \in \{\kappa\} \cup \{\iota_\beta \mid \beta \in B\}$. From $\varphi_n \models_{\mathcal{S}_{\mathrm{K}}} p_{n \bmod 2}$ and $(S,\kappa) \not\models p_{n \bmod 2}$, we get $(S,\kappa) \models \neg\varphi_n$ and therefore $w \neq \kappa$. Hence there is $\beta \in B$ with $w = \iota_\beta$. Now $(S,w) \models \varphi_n$ implies $(S_\beta, w) \models \varphi_n$. But this contradicts $\beta \in B$. Thus, indeed, $(S,\kappa) \not\models \Diamond\varphi_n$. Together with the observation $(S,\kappa) \models p_{(n+1) \bmod 2} \wedge p$, we get $(S,\kappa) \not\models \varphi_{n+1}$.

Claim (C), $(S,\kappa) \not\models \psi$**.** This is immediate by the above since $S \in \mathcal{S}_{\mathrm{T}}$ and $\psi \equiv_{\mathcal{S}_{\mathrm{T}}} \varphi_{n+1}$.

Claim (D), $(S,\iota) \not\models \psi$**.** Towards a contradiction, assume $(S,\iota) \models \psi$.

Recall that ψ is a Boolean combination of atomic formulae and of formulae $\Diamond\lambda$ with $\lambda \in O_\psi \cup E_\psi$. We now prove the following two claims:

(D1) If $\lambda \in E_\psi$ with $(S,\iota) \models \Diamond\lambda$, then $(S,\kappa) \models \Diamond\lambda$.
(D2) If $\lambda \in O_\psi$ with $(S,\kappa) \models \Diamond\lambda$, then $(S,\iota) \models \Diamond\lambda$.

Proof of (D1). Let $\lambda \in E_\psi$ with $(S, \iota) \models \Diamond\lambda$.

First, assume $\lambda \models_{\mathcal{S}_\mathrm{T}} \varphi_n$, i.e., $\lambda \in E^*_\psi$. From $(S, \iota) \models \Diamond\lambda$, we obtain that the formula λ holds in one of the worlds ι, κ, ι_α, or ι_β for some $\beta \in B$. But $(S, \iota_\alpha) \models \lambda$ implies $S_\alpha \models \lambda$ by (A.1), which is impossible since $S_\alpha \models \alpha$, $\alpha \models_{\mathcal{S}_\mathrm{K}} \neg \bigvee E^*_\psi$, and $\neg \bigvee E^*_\psi \models_{\mathcal{S}_\mathrm{K}} \neg\lambda$ since $\lambda \in E^*_\psi$. Next suppose $(S, \iota_\beta) \models \lambda$ for some $\beta \in B$. Then $S_\beta \models \lambda$ which, together with $\lambda \models_{\mathcal{S}_\mathrm{T}} \varphi_n$, implies $S_\beta \models \varphi_n$. Because of $\beta \in B$, we have $S_\beta \models \neg\varphi_n$, a contradiction.

Consequently, the formula λ holds in one of the worlds ι and κ. Again using $\lambda \models_{\mathcal{S}_\mathrm{T}} \varphi_n$, we obtain that also φ_n holds in ι or in κ. But this cannot be the case since $p_{n \bmod 2}$ does not hold in either of the two worlds – a contradiction.

Hence, we have $\lambda \not\models_{\mathcal{S}_\mathrm{T}} \varphi_n$. But then the formula $\beta := (\lambda \wedge \neg\varphi_n)$ is satisfiable in $\mathcal{S}_\mathrm{T}$ implying $S_\beta \models \beta$. From $\beta \models_{\mathcal{S}_\mathrm{K}} \lambda$ and $\beta \models_{\mathcal{S}_\mathrm{K}} \neg\varphi_n$, we obtain $S_\beta \models \beta \wedge \neg\varphi_n \wedge \lambda$. Hence $\beta \in B$. Now $S_\beta \models \lambda$ implies $(S, \iota_\beta) \models \lambda$ and therefore $(S, \kappa) \models \Diamond\lambda$.

Proof of (D2). Let $\lambda \in O_\psi$ with $(S, \kappa) \models \Diamond\lambda$. Then there is $w \in \{\kappa\} \cup \{\iota_\beta \mid \beta \in B\}$ such that $(S, w) \models \lambda$. In any case, $(\iota, w) \in R$. Hence we have $(S, \iota) \models \Diamond\lambda$.

This finishes the proof of the claims (D1) and (D2). Note that (S, ι) and (S, κ) agree in the atomic formulae holding there. From Lemma 4.1 and our assumption $(S, \iota) \models \psi$, we therefore get $(S, \kappa) \models \psi$, contrary to what we saw before. Hence, indeed, $(S, \kappa) \not\models \psi$.

Conclusion of Case 2. Through steps (A) to (D), we proved $(S, \iota) \models \varphi_{n+1} \wedge \neg\psi$ also in case $\left|\bigvee E_\psi\right| < 2^n$, contradicting the equivalence of φ_{n+1} and ψ.