

Separation of Test-Free Propositional Dynamic Logics over Context-Free Languages

Markus Latte*

Department of Computer Science
University of Munich, Germany

For a class \mathcal{L} of languages let $\text{PDL}[\mathcal{L}]$ be an extension of Propositional Dynamic Logic which allows programs to be in a language of \mathcal{L} rather than just to be regular. If \mathcal{L} contains a non-regular language, $\text{PDL}[\mathcal{L}]$ can express non-regular properties, in contrast to pure PDL.

For regular, visibly pushdown and deterministic context-free languages, the separation of the respective PDLs can be proven by automata-theoretic techniques. However, these techniques introduce non-determinism on the automata side. As non-determinism is also the difference between DCFL and CFL, these techniques seem to be inappropriate to separate $\text{PDL}[\text{DCFL}]$ from $\text{PDL}[\text{CFL}]$. Nevertheless, this separation is shown but for programs without test operators.

1 Introduction

Propositional Dynamic Logic (PDL) [9] is a logical formalism to specify and verify programs [12, 16, 11]. These tasks rely on the satisfiability and model-checking problems. Applications in the field are supported by their relatively low complexities: EXPTIME- and PTIME-complete, respectively [9].

Formulas in PDL are interpreted over labeled transition systems. For instance, the formula $\langle p \rangle \varphi$ means that after executing the program p the formula φ shall hold. In this context, programs and formulas are defined mutually inductively. This mixture allows programs to test whether or not a formula holds at the current state. Additionally, programs are required to be regular over the set of atomic programs and test operations. For instance, the program `while (b) do p;` can be rendered as $\langle (b?; p)^* ; \neg b \rangle \varphi$ to ensure that the loop is finite and that φ holds when the loop terminates [9].

The small model property of PDL [9] cuts both ways. First, it admits a decision procedure for satisfiability, but secondly it restricts the expressivity to regular properties. As a consequence counting properties and, in particular, the nature of execution stacks cannot be expressed. The last consequence runs contrary to the verification of recursive programs.

A natural way to enhance the expressivity is to relax the regularity requirement. For a class \mathcal{L} of languages let $\text{PDL}[\mathcal{L}]$ denote the variation which requires that any program belongs to \mathcal{L}^1 . For instance, we write a diamond as $\langle L \rangle \varphi$ for $L \in \mathcal{L}$. This leads to a hierarchy of logics. Obviously, $\text{PDL}[\mathcal{L}] \leq \text{PDL}[\mathcal{M}]$ holds for $\mathcal{L} \subseteq \mathcal{M}$. Besides regular languages, we consider the variations for the class of visibly pushdown

*Supported by the DFG Graduiertenkolleg 1480 (PUMA).

¹ If test operations and deterministic languages are involved, the test operations also must behave deterministically. In the case of DCFLs the additional restriction reads as follows (using the notation in [15]).

- For any state q , at most one of $\delta(q, a, X)$ (for $a \in \Sigma$), $\delta(q, \varepsilon, X)$ and $\delta(q, \varphi?, X)$ (for some $\text{PDL}[\mathcal{L}]$ -formula φ) is not empty.
- For any state q and two distinct $\text{PDL}[\mathcal{L}]$ -formulas φ_1 and φ_2 , we have that if $\delta(q, \varphi_1?, X) \neq \emptyset$ and $\delta(q, \varphi_2?, X) \neq \emptyset$ then φ_1 and φ_2 are semantically disjoint, that is $\models \neg(\varphi_1 \wedge \varphi_2)$.

Otherwise, it would be possible to simulate a non-deterministic choice by inserting a test for “true” for every possible choice and vary each test syntactically in a different way. Note that “true” has infinitely many synonyms. A non-example is $\Delta\text{PDL}^2[\text{CFL}] = \Delta\text{PDL}^2[\text{DCFL}]$ in [2].

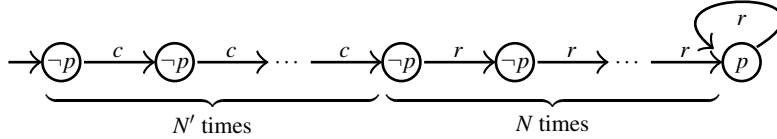
languages [1], VPL, the class of deterministic context-free languages, DCFL, and context-free languages, CFL. The inclusion order continues on the logics' side.

$$\text{PDL} = \text{PDL}[\text{REG}] \leq \text{PDL}[\text{VPL}] \leq \text{PDL}[\text{DCFL}] \leq \text{PDL}[\text{CFL}]. \quad (1)$$

Harel et al. discussed the effect of adding *single* (deterministic) context-free programs to PDL [13, 14, 12]. The logic PDL[VPL] were introduced by Löding et al. [17].

To handle the respective decision problems, the languages are represented by a machine model for the respective class. For each of these logics, any of its formula φ can be translated into an ω -tree-automaton which recognizes exactly all tree-like models of φ where the out-degree of any node is globally bounded. Such a model exists iff φ is satisfiable. For PDL and PDL[REG] these tree-automata are finite-state [23], for PDL[VPL] they are visibly pushdown tree-automata [14, 17] and for PDL[DCFL] and PDL[CFL] they are tree-automata with unbounded number of stacks. The last notion is rather artificial. However, the stacks are used, first, to accumulate unfulfilled eventualities and to simulate the complementation of programs given as pushdown automata. Note that in the setting of visibly pushdown automata, only one stack suffices as ω -VPLs are closed under complementation [1] and under determinisation (for stair-parity conditions) [18].

The first two inequalities in (1) are strict. In this paragraph we sketch the proofs for the first two inequalities. Consider the language $L := \{c^n r^n \mid n \in \mathbb{N}\}$ over an alphabet $\Sigma \supseteq \{c, r\}$. Hence, we have $L \in \text{VPL}$ if we take c for a call and r for a return in a visibly pushdown alphabet for Σ . Now, we claim that $\vartheta := \langle L \rangle p$ is not expressible in PDL[REG] where p is a proposition. For the sake of contradiction, assume that there were such a formula. Restricted to linear models, the previous translation leads to a finite-state Büchi-automaton \mathcal{A} which recognizes those models. Let N be sufficiently large—which depends on the pumping length and the, here omitted, encoding. Consider the following model of ϑ for $N' = N$.



As \mathcal{A} accepts this model, it also accepts this transition system for $N' < N$ due to the pumping lemma. However, this structure is not a model of ϑ . The separation of PDL[VPL] and PDL[DCFL] can be achieved in similar fashion. Take as program $L := \{w \# w^R \mid w \in (\Sigma \setminus \{\#\})^*\} \in \text{DCFL}$ over an alphabet $\Sigma \ni \#$. For any visibly pushdown alphabet for Σ its return-part is not empty in general. Using such a letter for the w -part in L , an assumed visibly pushdown automaton for $\langle L \rangle p$ operates on that part like a finite-state automaton. The same argumentation applies as for the first separation.

The separation for the last inequality in (1) is more cumbersome and intrinsic: For the satisfiability problem, the emptiness problem for finite-state and for visibly pushdown tree-automata is decidable [19, 23, 22][18]. The emptiness problem for the tree-automata with an unbounded number of stacks can be considered as the halting problem for Büchi-Turing machines [20]. Indeed, the satisfiability problems for PDL[DCFL] and PDL[CFL] are Σ_1^1 -complete [13]. Hence, both logics are not distinct by a “trivial” reason.

The standard translation [23, 14, 17] from formulas to tree-automata bases on Hintikka-sets. For a fixed formula ϑ and for every node of the given transition system the automaton for ϑ guesses—among other things—the set of those subformulas of ϑ which hold at that node. Informally speaking, the non-determinism is required to handle disjunctions in the given formula and to recognize the termination of a program in an expression such as $\langle L \rangle \varphi$. Note that a language in DCFL might be not prefix-free. However, non-determinism is also the difference between DCFL and CFL. Hence, the translation seems not to suffice to separate PDL[DCFL] from PDL[CFL].

In this paper we make a step towards the separation of PDL[CFL] from PDL[DCFL]. For technical reasons we consider PDL[\mathcal{L}] without the test operations like $\varphi?$ —call the logic PDL₀[\mathcal{L}]—and prove the separation of the corresponding logics. This restriction is proper as PDL₀ is weaker than PDL [5]. Note that PDL₀[\mathcal{L}] is exactly the EF/AG-fragment of CTL[\mathcal{L}] [2, 3]. This logic is obtained from CTL by restricting the moments of until- and release-operations by languages in \mathcal{L} . The separation of CTL[DCFL] and CTL[CFL] is unknown as well.

2 Preliminaries

Let Σ be an alphabet. For a finite word $w \in \Sigma^*$ we write $|w|$ for its length and $w[i..j]$ for its subword starting at index i and ending at index j where $0 \leq i \leq j < |w|$. Both indices are zero-based. For words $u, v \in \Sigma^* \cup \Sigma^\omega$ their concatenation is written as uv and the reversal of u as u^R . Concatenation is extended to sets in the usual way. The empty word is denoted by ε . A word $u \in \Sigma^* \cup \Sigma^\omega$ is a (proper) prefix of $w \in \Sigma^* \cup \Sigma^\omega$ iff there is $v \in \Sigma^* \cup \Sigma^\omega$ such that $uv = w$ (and $v \neq \varepsilon$). The notation of a suffix is defined similarly. For two languages L_1 and L_2 their left quotient $L_1 \setminus L_2$ is $\{v \mid \exists u \in L_1. uv \in L_2\}$. If one of both languages is a singleton we may replace the language by its single word. Standard notations are used [15] for (deterministic) pushdown automata on finite words, DPDA and PDA, and (deterministic) context-free languages. Deterministic pushdown automata on ω -words, ω DPDA, are equipped with Büchi-acceptance conditions [20].

Let $\text{Prop} = \{p, q, \dots\}$ be a set of *propositions*. A *labeled transition system*, LTS, is a triple $\mathcal{T} = (\mathcal{S}, \longrightarrow, \ell)$ consisting of a set of states \mathcal{S} , of a labeled edge relation $\longrightarrow \subseteq \mathcal{S} \times \Sigma \times \mathcal{S}$ and of an evaluation function $\ell : \mathcal{S} \rightarrow 2^{\text{Prop}}$. We write $s \xrightarrow{a} t$ instead of $(s, a, t) \in \longrightarrow$. A *path* is a sequence $s_0, a_1, s_1, a_1, \dots, a_{n-1}, s_n$ for some $n \in \mathbb{N}$ such that $s_i \xrightarrow{a_{i+1}} s_{i+1}$ for all $i \in \{0, \dots, n-1\}$. For such a path we may write $s_0 \xrightarrow{a_0} s_1 \xrightarrow{a_1} \dots \xrightarrow{a_{n-1}} s_n$. A *structure* is a pair $\mathcal{M} = (\mathcal{T}, s)$ of an LTS and a state in it, called *root*. Previous notations for LTSS are also used for structures. A structure $\mathcal{M} = ((\mathcal{S}', \longrightarrow', \ell'), s')$ is an *extension* of \mathcal{M} , written as $\mathcal{M} \leq \mathcal{M}'$, iff $\mathcal{S} \subseteq \mathcal{S}'$, $\longrightarrow \subseteq \longrightarrow'$, ℓ is the restriction of ℓ' to \mathcal{S} , and $s = s'$.

Let \mathcal{L} be a class of languages. We define the logic PDL₀[\mathcal{L}] in negation normal form using a CTL-like syntax [2]—that is, $\text{EF}^L \varphi$ stands for the PDL-expression $\langle L \rangle \varphi$ for instance. The formulas are given by the grammar

$$\varphi ::= \text{ff} \mid \text{tt} \mid p \mid \neg p \mid \varphi \vee \varphi \mid \varphi \wedge \varphi \mid \text{EF}^L \varphi \mid \text{AG}^L \varphi$$

where $p \in \text{Prop}$ and $L \in \mathcal{L}$. Such formulas are denoted by φ, ψ, ϑ , and δ . The atoms ff and tt are called *constants*, and p and \neg are called *literals*. Implication and equivalence are definable. A formula $\text{EF}^L \varphi$ is called *EF-formula*. An *AG-formula* is meant analogously. A formula is interpreted over a structure as follows.

$$\begin{aligned} \mathcal{T}, s \not\models \text{ff} & \quad \mathcal{T}, s \models \text{tt} & \quad \mathcal{T}, s \models p \text{ iff } p \in \ell(s) & \quad \mathcal{T}, s \models \neg p \text{ iff } p \notin \ell(s) \\ \mathcal{T}, s \models \varphi_1 \vee \varphi_2 & \text{ iff } \mathcal{T}, s \models \varphi_1 \text{ or } \mathcal{T}, s \models \varphi_2 & \quad \mathcal{T}, s \models \varphi_1 \wedge \varphi_2 & \text{ iff } \mathcal{T}, s \models \varphi_1 \text{ and } \mathcal{T}, s \models \varphi_2 \\ \mathcal{T}, s \models \text{EF}^L \varphi & \text{ iff there is path } s_0 \xrightarrow{a_0} s_1 \xrightarrow{a_1} \dots \xrightarrow{a_{n-1}} s_n & \text{ with } s = s_0, a_0 \dots a_{n-1} \in L & \text{ and } \mathcal{T}, s_n \models \varphi \\ \mathcal{T}, s \models \text{AG}^L \varphi & \text{ iff for all paths } s_0 \xrightarrow{a_0} s_1 \xrightarrow{a_1} \dots \xrightarrow{a_{n-1}} s_n & \text{ with } s = s_0 \text{ and } a_0 \dots a_{n-1} \in L: & \mathcal{T}, s_n \models \varphi \end{aligned}$$

If $\mathcal{T}, s \models \varphi$ then the structure (\mathcal{T}, s) is a *model* of φ . A structure (\mathcal{T}, s) is *tree-like* iff \mathcal{T} forms a tree with root s . Since PDL₀[\mathcal{L}] is closed under bisimulation, every satisfiable formula has a tree-like structure as a model. A formula φ is a *tautology*, written as $\models \varphi$, iff every structure is a model of φ .

3 Outline of the Proof

For the following parts, fix an alphabet Σ which at least contains 0 and 1 but not \$, and set $\Sigma_{\$} := \Sigma \cup \{\$\}$. The language of palindromes is denoted by $\text{Palindromes} := \{w \in \Sigma^* \mid w = w^R\}$. We will show that there is no $\text{PDL}_0[\text{DCFL}]$ -formula which is equivalent to the *reference* $\text{PDL}_0[\text{CFL}]$ -formula $\text{EF}^{\text{Palindromes}} \tt . As the reference formula does not contain propositions we may assume that neither does any equivalent formula. Equivalently, we may assume that $\text{Prop} = \emptyset$.

For the sake of contradiction, let $\vartheta \in \text{PDL}_0[\text{DCFL}]$ be a *candidate* formula which is assumed to be equivalent to $\text{EF}^{\text{Palindromes}} \tt . To illustrate the main problem about provoking a contradiction, we begin with a simpler setting in which ϑ does not contain any conjunctions or AG-formulas. As we have the equivalences

$$\text{EF}^L \text{ff} \leftrightarrow \text{ff}, \quad \text{EF}^{L_1} \text{EF}^{L_2} \psi \leftrightarrow \text{EF}^{L_1 L_2} \psi, \quad \text{and} \quad \bigvee_i \text{EF}^{L_i} \psi \leftrightarrow \text{EF}^{\bigcup_i L_i} \psi$$

the formula ϑ can be rewritten as

$$\text{EF}^{\bigcup_i L_{i,1} \dots L_{i,n_i}} \text{tt}$$

where $L_{i,j}$ are DCFLs over $\Sigma_{\$}$. In general, an equivalence $\text{EF}^{\text{Palindromes}} \$\text{tt} \leftrightarrow \text{EF}^L \text{tt}$ implies

$$\text{Palindromes} = \{w \in \Sigma^* \mid w\$ \text{ is a prefix of a word in } L\}$$

for $L \subseteq \Sigma_{\* . Therefore, we have that Palindromes would be expressible as a finite union over a finite concatenation over DCFLs over Σ . Some combinatorial argument shows that this is impossible.

Back to the real world, we are also faced with conjunctions and AG-formulas in ϑ . A natural attempt is to eliminate these subformulas. Indeed, a conjunction seems not to support a statement which speaks about a single path only. Instead, it speaks about a bunch of paths. Similarly, an AG-formula is not monotone with respect to models but the reference formula is monotone. To turn off such formulas, one could saturate the considered structures with substructures which falsify AG-formulas and which do not affect the desired property $\text{EF}^{\text{Palindromes}} \tt . However on such a new structure, the attached substructures could be recognized by other EF-subformulas. But these subformulas need not to be concerned with palindromes in any reasonable way. Moreover, Bojańczyk proved [6]—for the dual setting—that such an elimination procedure is only possible if—in our setting—palindromes were expressible as a finite union of languages of the form $A_0^* a_1 A_1^* a_2 \dots A_{n-1}^* a_n A_n^*$ for $a_1, \dots, a_n \in \Sigma$ and $A_0, \dots, A_n \subseteq \Sigma$. Obviously, this is not the case.

Therefore, our strategy is different. First, we show that topmost AG-formulas and topmost conjunctions can be eliminated (§ 6 and 7). This renders the candidate formula ϑ as $\bigvee_i \text{EF}^{L_i} \psi_i$ for some $L_i \subseteq \Sigma_{\$}^+$ and some formulas ψ_i with unknown structure. Secondly, if L_i is not a singleton language then the formula $\text{EF}^{L_i} \psi_i$ per se provides all the information required for a contradiction. Either it under- or over-approximates palindromes. And if L_i is a singleton we proceed in a similar way with the left-quotient of ϑ with the only word in L_i . The whole procedure (§ 8) terminates through a sophisticated measure (§ 5). The case that L_i is not a singleton give rise to a characterization of languages which will bridge between the formula and the language part of the separation proof.

Definition 1. A language $L \subseteq \Sigma^*$ is good iff $L = \bigcup_{i \in I} L_i R_i$ such that I is finite, and for each $i \in I$, the language L_i is a DCFL, $|L_i| \geq 2$ and $R_i \subseteq \Sigma^*$.

In the view of Bojańczyk's result, our iterated elimination is non-uniform compared to the preferable approach in the previous paragraph. Finally, we show on the language-theoretical level that palindromes are not good (§ 4).

4 On Palindromes and DCFLs

In this section it is proven that the language of palindromes is not good. For this purpose we first show that this language is not expressible as a union of DCFLs (Theorem 3). Although it is known that the set of palindromes is not deterministic context-free, the standard proof [10, Cor. 1] does not seem to be adaptable because the applied min-operator does not commute with the union. As a second step, it is shown that if palindromes are underapproximated by a concatenation then the components of the concatenation follow a very simple pattern (Lemma 5).

Lemma 2 (Pumping lemma). *Let $u \in \Sigma^\omega$ be accepted by an ω DPDA \mathcal{A} . There are words $u_0 \in \Sigma^*$, $u_1 \in \Sigma^+$ and $u_2 \in \Sigma^\omega$ such that $u_0u_1u_2 = u$, and u_0u_2 is accepted by \mathcal{A} .*

Proof. Firstly, we may assume that \mathcal{A} only erases or pushes symbols from or on the stack and never changes the topmost symbol. Indeed, an ω DPDA can keep the topmost element of the stack in its control state [15, Sect. 10.1]. By this restriction, in any run the stacks of two consecutive configurations are comparable with respect to the prefix-order. Secondly, consider the infinitely many stair positions in the accepting run of \mathcal{A} on u . By a stair position [18] we understand a position such that the current stack content is a prefix of all further stack contents in this run. As the set of states is finite, there are two different stair positions which name the same state. We may assume that a non-empty part of u , say u_1 with $u = u_0u_1u_2$, fits into their gap. Hence, this part can be removed. By the definition of stairs, the obtained sequence of configurations is a run of \mathcal{A} on u_0u_2 . As the modification affects a prefix of u only, \mathcal{A} also accepts u_0u_2 . \square

Theorem 3. *Let $v \in \Sigma^*$, $n \in \mathbb{N}$, and L_1, \dots, L_n be DCFLs over Σ . Then $\bigcup_{i=1}^n L_i \neq v \setminus \text{Palindromes}$.*

Proof. Define the sequence $(w_i)_{i \in \mathbb{N}}$ of strictly prefix-ordered words as follows.

$$\begin{aligned} w_0 &:= v^R \\ w_{i+1} &:= w_i 10^i 1 w_i^R v^R \end{aligned} \quad (i \in \mathbb{N})$$

For all $i \in \mathbb{N}$ we have $w_i \in v \setminus \text{Palindromes}$. For the sake of contradiction, assume that

$$\bigcup_{i=1}^n L_i = v \setminus \text{Palindromes}. \quad (2)$$

We sample the candidate on the left of Eq. 2 with the words $\{w_i\}_{i \in \mathbb{N}}$. Since the union is finite, there is an infinite $I \subseteq \mathbb{N}$ and an $i \in \{1, \dots, n\}$ such that the words $\{w_i\}_{i \in I}$ belong to L_i . Let \mathcal{A} be a DPDA for L_i . Additionally, we consider \mathcal{A} as an ω DPDA where the final states are the Büchi-states. Hence, as \mathcal{A} is a deterministic device it accepts

$$w := \lim_{i \in \omega} w_i = \lim_{i \in I} w_i \in \Sigma^\omega. \quad (3)$$

Apply Lemma 2 to \mathcal{A} and w . Let u_0, u_1, u_2 be the obtained factors. We run \mathcal{A} on w for at least $|u_0u_1|$ steps until it processes some subword $10^k 1$ for the first time. Note that the function which maps $i \in \mathbb{N}$ to the first occurrence of $10^i 1$ in w is unbounded. Let ℓ be the first index in w after that subword. So far, \mathcal{A} has seen the first ℓ letters in w . We keep \mathcal{A} running for at least another $\ell + |v|$ steps until it reaches a final state. Such a run is always possible as \mathcal{A} accepts infinitely many prefixes of w . Let u' be the word constructed in this way. Hence, $u' \in v \setminus \text{Palindromes}$ as \mathcal{A} accepts u' .

Let u'' be the word u' where the u_1 -block is removed. That is $u'' := u'[0 .. |u_0| - 1] u' [|u_0 u_1| .. |u'| - 1]$. Again by construction and Lemma 2, \mathcal{A} accepts u'' . Thus, $u'' \in v \setminus \text{Palindromes}$. Let \hat{u} be the word between u_1 and the block $10^\kappa 1$, that is $\hat{u} = w [|u_0 u_1| .. \ell - 3 - \kappa]$. As vu' is a palindrome, it ends in the word $(vu_0 u_1 \hat{u} 10^\kappa 1)^R$ of length $\ell + |v|$. The modification leading to u'' affects at most the first ℓ positions only. Hence, as $|u'| \geq 2\ell + |v|$, u'' also ends in $(vu_0 u_1 \hat{u} 10^\kappa 1)^R$. As vu'' is also a palindrome, $u_0 \hat{u} 10^\kappa 1$ is a prefix of $u_0 u_1 \hat{u} 10^\kappa 1$. Since u_1 is not the empty word, this is a contradiction to the choice of $10^\kappa 1$. \square

Lemma 4. *If $LR \subseteq \text{Palindromes}$ and R is infinite then L is prefix-ordered.*

Proof. Let $\ell_0, \ell_1 \in L$ with $|\ell_0| \leq |\ell_1|$. Take $r \in R$ such that $|r| \geq |\ell_1|$. This is possible as R is infinite. Since $\ell_0 r$ and $\ell_1 r$ are palindromes, ℓ_0^R and ℓ_1^R are suffixes of r . Therefore, ℓ_0 is a prefix of ℓ_1 . \square

Lemma 5. *Suppose $LR \subseteq \text{Palindromes}$, $|L| \geq 2$ and R is infinite. Then*

$$R \subseteq \hat{u}^* \hat{U}$$

for some word $\hat{u} \in \Sigma^*$ and a finite language $\hat{U} \subset \Sigma^*$.

Proof. Let u_0, u_1 be two distinct words in L . By the Lemma 4 we may assume that u_0 is a proper prefix of u_1 . Define $\hat{u} := u_0 \setminus u_1$. Note that $u_0 \hat{u} = u_1$.

Claim 5-1. *For $w \in R$ and $n \in \mathbb{N}$ we have*

- (i) \hat{u}^n is a prefix of w , and
- (ii) $(\hat{u}^R)^n u_0^R$ is a suffix of w

if $n|\hat{u}| + |u_0| \leq |w|$.

Proof of claim. By induction on n for a fixed $w \in R$. If $n = 0$, u_0^R is a suffix of w as $u_0 w$ is a palindrome. For the step case from n to $n + 1$ assume that

$$(n + 1)|\hat{u}| + |u_0| \leq |w|. \quad (4)$$

The word $v := u_0 \hat{u}^{n+1} = u_1 \hat{u}^n$ is prefix of $u_1 w$ by IH(i). As $u_1 w$ is palindrome, v^R is a suffix of w because of (4). This proves the second item. Since $u_0 w$ is also a palindrome, it has v is a prefix. Hence \hat{u}^{n+1} is a prefix of w —this is the first item. \square

Let $w \in R$. For $N_w := \lfloor (|w| - |u_0|) / |\hat{u}| \rfloor$, the claim yields $w = \hat{u}^{N_w} \hat{w}$ where \hat{w} are the $r_w := |w| - N_w |\hat{u}|$ last letters of w . Since $r_w \leq |u_0| + |\hat{u}|$ is bounded independently of w , there is a finite set \hat{U} such that $R \subseteq \hat{u}^* \hat{U}$. \square

Lemma 6. *Let*

$$L := \bigcup_{i \in I} u_{i,0} u_{i,1}^* u_{i,2}^* u_{i,3} \quad (5)$$

for I finite and $u_{i,j} \in \Sigma^*$ for all suitable indices. Then there is a word $w \in \Sigma^*$ which is not a prefix of any word in L .

Proof. Consider the tree Σ^ω . For each $i \in I$, the word $u_{i,0} u_{i,1}^\omega$ defines a (finite or infinite) path in the tree. As I is finite, there is a $w_0 \in \Sigma^*$ which is not on these paths. There are at most $|w_0| \cdot |I|$ paths of the form $u_{i,0} u_{i,1}^j u_{i,2}^\omega$ for $i \in I$ and $j \in \mathbb{N}$ which pass w_0 . By the same argument, we get a word w_1 which extends w_0 and cannot be reached by these paths. A final application to w_1 and $u_{i,0} u_{i,1}^j u_{i,2}^k u_{i,3}$ for $i \in I$ and $j, k \in \mathbb{N}$ yields the claimed word w . \square

Corollary 7. *The set Palindromes is not good.*

Proof. For the sake of contradiction, assume the contrary, that is

$$\text{Palindromes} = \bigcup_{i \in I} L_i R_i \tag{6}$$

where I is finite, and for any $i \in I$ the language L_i is a DCFL, and $|L_i| \geq 2$. Set $I^+ := \{i \in I \mid R_i \text{ is finite}\}$, and $I^- := I \setminus I^+$. For any $i \in I^+$, we have that $L_i R_i$ is a DCFL [10, Thm. 3.3] as R_i is finite in particular.

Let $i \in I^-$. Since $|L_i| \geq 2$ and R_i is infinite, Lemma 5 shows that $R_i \subseteq r_i^* \widehat{R}_i$ for some $r_i \in \Sigma^*$ and for a finite language $\widehat{R}_i \subset \Sigma^*$. Depending on the size of L_i we can bound $L_i R_i$. If L_i is infinite then the very same lemma shows by reversal that $L_i \subseteq \widehat{L}_i \ell_i^*$ for some $\ell_i \in \Sigma^*$ and a finite $\widehat{L}_i \subset \Sigma^*$. Hence,

$$L_i R_i \subseteq \widehat{L}_i \ell_i^* r_i^* \widehat{R}_i = \bigcup_{x \in \widehat{L}_i, y \in \widehat{R}_i} x \ell_i^* r_i^* y.$$

In the other case— $|L_i|$ is finite—one obtains

$$L_i R_i \subseteq \bigcup_{x \in L_i, y \in \widehat{R}_i} x r_i^* y.$$

In both cases, the unions are finite. All in all, we have

$$\bigcup_{i \in I} L_i R_i = \bigcup_{i \in I^+} \underbrace{L_i R_i}_{\text{DCFL}} \cup Q'$$

where $Q' \subseteq Q := \bigcup_{i \in J} u_{i,0} u_{i,1}^* u_{i,2}^* u_{i,3}$ for some finite set J , and some words $u_{i,0}, \dots, u_{i,3} \in \Sigma^*$. By Lemma 6, there is a finite word w which is not a prefix of any word in Q . Using (6), we get

$$w \setminus \text{Palindromes} = \bigcup_{i \in I^+} w \setminus (L_i R_i).$$

The left quotient with a single word w is the inverse of the gsm mapping which sends a word u to wu . As DCFLs are closed under the inverse of gsm mappings [10, Thm. 3.2], the language $w \setminus (L_i R_i)$ is a DCFL for $i \in I^+$. But this a contradiction to Theorem 3. \square

5 A Measure for the Extraction

Informally, the measure of a formula is a set of vectors. Each vector measures the languages annotated to EF-subformulas along a path from the root of the formula to its atoms. For the measure of a language, the size of its only word is considered if the language is a singleton.

Definition 8. Let \mathbb{M} be the set of all finite subsets of $(\omega + 1)^*$ where $\omega + 1 = \{0, 1, 2, \dots, \omega\}$. The second argument of the cons-operator $\cdot :: \cdot$ on $(\omega + 1)^*$ is extended to sets. The empty list is written as *nil*. The measure of a formula is defined by

$$\begin{aligned} \mu(\ell) &:= \{\text{nil}\} && \text{for } \ell \text{ a literal or a constant} \\ \mu(\varphi_0 \circ \varphi_1) &:= \mu(\varphi_0) \cup \mu(\varphi_1) && \text{for } \circ \in \{\wedge, \vee\} \\ \mu(Q^L \varphi) &:= ||L|| :: \mu(\varphi) && \text{for } Q \in \{\text{EF}, \text{AG}\} \end{aligned}$$

where

$$||L|| = \begin{cases} |w| & \text{if } L = \{w\} \text{ for some } w \in \Sigma^*, \\ \omega & \text{otherwise.} \end{cases}$$

Lemma 9. The lexicographic order [4, Sect. 2.4], $>_{\text{lex}}$, on $(\omega + 1)^*$ is defined by

$$(\omega + 1)^n \ni (u_1, \dots, u_n) >_{\text{lex}} (v_1, \dots, v_m) \in (\omega + 1)^m$$

iff $n > m \vee (n = m \wedge \exists k < n. u_k = v_k \wedge \forall i < k. u_i > v_i)$, where $>$ is the natural order on $\omega + 1$, that is $\omega > \dots > 1 > 0$.

Definition 10. The binary relation $>_{\mathbb{M}}$ on \mathbb{M} is defined as follows.

$$\begin{aligned} M >_{\mathbb{M}} N \text{ iff } & \text{there are } X, Y \in \mathbb{M} \text{ such that } \emptyset \neq X \subseteq M, \\ & N = (M \setminus X) \cup Y, \text{ and } \forall y \in Y \exists x \in X. x >_{\text{lex}} y. \end{aligned}$$

Lemma 11. The relation $>_{\mathbb{M}}$ is a strict and terminating order.

Proof. We follow Baader and Nipkow [4]. The natural order on $\omega + 1$ is strict and terminating. Hence, so is $>_{\text{lex}}$ [4, Lemma 2.4.3]. Therefore, the multiset order on $(\omega + 1)^*$ is also strict and terminating [4, Lemma 2.5.4 and Theorem 2.5.5]. Due to the natural embedding of \mathbb{M} into the set of finite multisets on $(\omega + 1)^*$, the relation $>_{\mathbb{M}}$ is dominated by the multiset order. Hence $>_{\mathbb{M}}$ is terminating. Thanks to the same embedding, $>_{\mathbb{M}}$ is a strict order \square

We write $\geq_{\mathbb{M}}$ for the reflexive closure of $>_{\mathbb{M}}$. Similarly, $\leq_{\mathbb{M}}$ and $<_{\mathbb{M}}$ are meant.

6 ε -Free Formulas

Formulas like $\text{EF}^L \psi$ and $\text{AG}^L \psi$ can speak about the current state if $\varepsilon \in L$. We intend to combine structures at their roots—in the proof to Thm. 17 and 18—, such that formulas should not realize this modification. Nonetheless, formulas can be transformed accordingly.

Definition 12. The property being ε -free is inductively defined on $\text{PDL}_0[\cdot]$ -formulas.

- (i) Any literal is ε -free.
- (ii) A conjunction and a disjunction is ε -free if both conjuncts or both disjuncts, respectively, are ε -free.
- (iii) $\text{EF}^L \varphi$ and $\text{AG}^L \varphi$ are ε -free iff $\varepsilon \notin L$ and φ is ε -free.

Definition 13. The function \cdot^{ε} is defined on $\text{PDL}_0[\cdot]$ -formulas

$$\begin{aligned} \ell^{\varepsilon} &:= \ell && \text{where } \ell \text{ literal or a constant} \\ \varphi_0 \circ \varphi_1^{\varepsilon} &:= \varphi_0^{\varepsilon} \circ \varphi_1^{\varepsilon} && \text{for } \circ \in \{\wedge, \vee\} \\ Q^L \varphi^{\varepsilon} &:= \begin{cases} Q^L \varphi^{\varepsilon} & \text{if } \varepsilon \notin L \\ \varphi^{\varepsilon} \vee Q^{L \setminus \{\varepsilon\}} \varphi^{\varepsilon} & \text{otherwise if } Q = \text{EF} \\ \varphi^{\varepsilon} \wedge Q^{L \setminus \{\varepsilon\}} \varphi^{\varepsilon} & \text{otherwise if } Q = \text{AG} \end{cases} && \text{for } Q \in \{\text{EF}, \text{AG}\} \end{aligned}$$

Lemma 14. *For every PDL₀[·]-formula φ we have,*

- (i) φ and φ^ε are equivalent,
- (ii) φ^ε is ε -free, and
- (iii) $\mu(\varphi^\varepsilon) \leq_{\mathbb{M}} \mu(\varphi)$.

Proof. Each item can be proven by induction on φ . We detail on the last item for the second case of $Q^L\varphi^\varepsilon$. As the IH yields $\mu(\varphi^\varepsilon) \leq_{\mathbb{M}} \mu(\varphi)$, we have $\mu(\varphi^\varepsilon) <_{\mathbb{M}} \mu(Q^{L \setminus \{\varepsilon\}}\varphi^\varepsilon) \leq_{\mathbb{M}} \mu(Q^{L \setminus \{\varepsilon\}}\varphi) \leq_{\mathbb{M}} \mu(Q^L\varphi)$. Hence, the claim follows by $\mu(\varphi^\varepsilon \vee Q^{L \setminus \{\varepsilon\}}\varphi^\varepsilon) \leq_{\mathbb{M}} \mu(Q^{L \setminus \{\varepsilon\}}\varphi^\varepsilon)$. \square

7 Elimination of Outermost AG-Formulas and Conjunctions

Although it is impossible to eliminate conjuncts and AG-formulas in general, the topmost ones can be removed (Thm. 17 and 18). Hence, if ϑ is equivalent to $EF^L\text{tt}$ for some language L then ϑ can be rearranged to a disjunction of EF-formulas only. However, these EF-formulas might contain conjunctions and AG-formulas in turn.

Definition 15. *A formula ϑ is in disjunctive normal form (DNF for short) iff it has the shape*

$$\bigvee_{i \in I} \left(\bigwedge_{j \in J_i^A} \alpha_{i,j} \wedge \bigwedge_{j \in J_i^E} \varepsilon_{i,j} \right)$$

where I , J_i^A and J_i^E are finite sets, $\alpha_{i,j}$ is an ε -free AG-formula, and $\varepsilon_{i,j}$ is an ε -free EF-formula (for all suitable indices). The completion of ϑ is

$$\vartheta^\bullet := \vartheta \quad \vee \quad \bigvee_{\substack{\Psi' \subseteq \Psi \\ \models \bigwedge \Psi' \rightarrow \vartheta}} \bigwedge \Psi'$$

where $\Psi := \{\varepsilon_{i,j} \mid i \in I, j \in J_i^E\}$. A formula ϑ' is complete iff it is ϑ^\bullet for some ϑ . The term “DNF” and “complete” shall be applied up to associativity and commutativity of the Boolean connectives².

Lemma 16. *For any ε -free formula ϑ we have*

- (i) *an equivalent formula ϑ' in DNF such that $\mu(\vartheta') \leq_{\mathbb{M}} \mu(\vartheta)$, and*
- (ii) *that ϑ^\bullet is a DNF, ϑ^\bullet and ϑ are equivalent, and $\mu(\vartheta^\bullet) \leq_{\mathbb{M}} \mu(\vartheta)$.*

Proof. To get a DNF, the distributive law is applied where AG- and EF-formulas are taken as atoms. This application might rearranging (positive) Boolean connectives and might duplicate atoms. However, the measure is defined in terms of unions for these cases.

For the second item, the implication to ϑ follows from the definition of the additional disjuncts. The other direction is weakening. As the additional terms are build only of top-level EF-formulas in ϑ , their measure is already subsumed in $\mu(\vartheta)$. Note that μ is just the union in the case of the (positive) Boolean connectives. \square

²Note that this is well-defined when the measure μ is taken.

For two structures \mathcal{M}_1 and \mathcal{M}_2 we define $\mathcal{M}_1 \oplus \mathcal{M}_2$ as the *disjoint sum* of both structures but with the root shared. The evaluation of the root is fixed as $\text{Prop} = \emptyset$ for our purposes. The notation is extended to sequences of structures, say $(\mathcal{M}_i)_{i \in I}$, in the usual way, written as $\oplus_{i \in I} \mathcal{M}_i$.

A formula ψ is *structurally monotone* iff for any model of ψ any of its extension is also a model of ψ . An example is $\text{EF}^L \text{tt}$ for any language L .

Theorem 17 (Elimination of AG-formulas). *Let*

$$\psi := \bigvee_{i \in I} \underbrace{\left(\alpha_i \wedge \bigwedge_{j \in J_i} \varepsilon_{i,j} \right)}_{=: \tau_i} \quad (7)$$

be complete where I, J_i are finite, each α_i is a (possibly empty) conjunction of ε -free AG-formulas, and each $\varepsilon_{i,j}$ is a ε -free EF-formula. If ψ is structurally monotone, then ψ is equivalent to

$$\psi' := \bigvee_{i \in I} \underbrace{\left(\bigwedge_{j \in J_i} \varepsilon_{i,j} \right)}_{\substack{=: \alpha_i \\ =: \tau'_i}}. \quad (8)$$

Note that $\mu(\psi') \leq_{\mathbb{M}} \mu(\psi)$.

Proof. $\models \psi' \rightarrow \psi$ is obvious. As the considered logic is closed under bisimulation, we consider tree-like structures in the following only. For the other direction, let \mathcal{M} be a model of ψ . We have to show that \mathcal{M} is also a model of ψ' . If there is an $i \in I$ such that $\models \alpha_i$ and $\mathcal{M} \models \tau_i$ then $\mathcal{M} \models \tau'_i$ and we are done. Otherwise, there is an $i_0 \in I$ such that $\not\models \alpha_{i_0}$ and $\mathcal{M} \models \bigwedge_{j \in J_{i_0}} \varepsilon_{i_0,j}$, as $\mathcal{M} \models \psi$. For $i \in I$ define

$$J_i^+ := \{j \in J_i \mid \mathcal{M} \models \varepsilon_{i,j}\}, \text{ and} \quad (9)$$

$$J_i^- := J_i \setminus J_i^+. \quad (10)$$

There are two cases. Either

$$\bigvee_{i \in I} \alpha_i \wedge \bigwedge_{j \in J_i^-} \varepsilon_{i,j} \quad (11)$$

is a tautology or not. If (11) is a tautology then so is

$$\bigwedge_{i \in I} \bigwedge_{j \in J_i^+} \varepsilon_{i,j} \rightarrow \psi \quad (12)$$

as a simple case distinction on (11) shows. Indeed, let $\widetilde{\mathcal{M}}$ be a model of the left side of (12). Then there is an $i \in I$ such that $\widetilde{\mathcal{M}} \models \alpha_i \wedge \bigwedge_{j \in J_i^-} \varepsilon_{i,j}$. Both together lead to $\widetilde{\mathcal{M}} \models \alpha_i \wedge \bigwedge_{j \in J_i} \varepsilon_{i,j}$ and finally to $\widetilde{\mathcal{M}} \models \psi$. Hence, the left hand side of (12) is a term in ψ as the latter is complete. But, by definition, this term is modeled by \mathcal{M} .

Otherwise (11) is not a tautology. So there is a structure \mathcal{M}' with

$$\mathcal{M}' \not\models \alpha_i \wedge \bigwedge_{j \in J_i^-} \varepsilon_{i,j} \quad (13)$$

for all $i \in I$. We will exclude this situation. As $J_{i_0}^- = \emptyset$ we have $\mathcal{M}' \not\models \alpha_{i_0}$ in particular. Now let $\mathcal{M}'' := \mathcal{M} \oplus \mathcal{M}'$. We claim that $\mathcal{M}'' \not\models \psi$ which is a contradiction to the assumption that ψ is structurally monotone. For the sake of contradiction, suppose that there is an $i \in I$ such that $\mathcal{M}'' \models \tau_i$. Among the EF-formulas only those indexed by J_i^+ are already fulfilled in \mathcal{M} . Hence, $\alpha_i \wedge \bigwedge_{j \in J_i^-} \varepsilon_{i,j}$ must be satisfied by \mathcal{M}' . This is a contradiction to the choice of \mathcal{M}' . Note that we used implicitly that EF-formulas are ε -free. \square

The theorem requires a syntactical presence of formulas called α_i . Note that minor changes make the proof also working if not all such parts are present. On the other hand, inserting such an empty conjunction does not increase the measure as atoms—such as tt —have the lowest measure anyway.

Theorem 18 (Elimination of \bigwedge EF-formulas). *Suppose*

$$\text{EF}^L \text{tt} = \delta \vee \bigwedge_{i \in I} \text{EF}^{L_i} \psi_i \quad (14)$$

where $\varepsilon \notin L_i$ for all $i \in I$. If $I \neq \emptyset$ then there is an $i \in I$ such that

$$\text{EF}^L \text{tt} = \delta \vee \text{EF}^{L_i} \psi. \quad (15)$$

Note that the measure of (15.r) is bounded by that of (14.r), trivially.

Proof. For any $i \in I$, (14.r) implies (15.r). If there is an $i \in I$ with $\models \text{EF}^{L_i} \psi_i \rightarrow \text{EF}^L \text{tt}$, this i suffices for the other direction. To exclude the other case, assume that we have tree-like structures \mathcal{M}_i for all $i \in I$ such that

- (i) $\mathcal{M}_i \models \text{EF}^{L_i} \psi_i$ but
- (ii) $\mathcal{M}_i \not\models \text{EF}^L \text{tt}$.

Let $w_i \in L_i$ be the witness for the first item. Set $\mathcal{M} := \bigoplus_{i \in I} \mathcal{M}_i$. The root of \mathcal{M} might satisfy different formulas than the root of \mathcal{M}_i , but this change is invisible to $\text{EF}^{L_i} \psi_i$ since $|w_i| > 0$. Hence, $\mathcal{M} \models \text{EF}^{L_i} \psi_i$. For the sake of a contradiction, assume that $\mathcal{M} \models \text{EF}^L \text{tt}$. This property depends only on a path in \mathcal{M} . The path is inherited from some \mathcal{M}_i for $i \in I$. Since $\text{EF}^L \text{tt}$ does not depend on the evaluation of the root, $\mathcal{M}_i \models \text{EF}^L \text{tt}$ which is a contradiction to the second property of \mathcal{M}_i . Therefore, $\mathcal{M} \not\models \text{EF}^L \text{tt}$. By construction we have $\mathcal{M} \models \bigwedge_{i \in I} \text{EF}^{L_i} \psi_i$ but $\mathcal{M} \not\models \text{EF}^L \text{tt}$. This property contradicts (14). \square

8 Extraction

In the proof of Theorem 22 we apply previous elimination techniques to show that the candidate formula is equivalent to $\bigvee_i \text{EF}^{L_i} \psi_i$. In the case that L_i is not a singleton set, we cannot decompose ψ_i any further. Indeed, the proof relies on the property $w(w \setminus L) \subseteq L$ for any language L and word w . However, this inequality is false when w is replaced by a non-singleton language. Nevertheless if the term $\text{EF}^{L_i} \psi_i$ accepts a linear structure then the term factorises the word on the structure. The left factor is L_i , surely, and the right one can be read off as follows.

Definition 19. *Let φ be a $\text{PDL}_0[\cdot]$ -formula. Its language is $\mathcal{L}(\varphi) := \{w \in \Sigma^* \mid \pi_w \$_ \models \varphi\}$ where $\pi_{a_1 \dots a_n}$ is a path labeled with a_1 to a_n for $a_1, \dots, a_n \in \Sigma$. The node reached after n steps has no successor. In each node, no proposition holds.*

Lemma 20. $\mathcal{L}(\text{EF}^{L^S} \text{tt}) = L$ for any $L \subseteq \Sigma^*$.

Proof. Let $w \in \mathcal{L}(\text{EF}^{L\$}\text{tt})$. By definition, we have $\pi_{w,\$} \models \text{EF}^{L\$}\text{tt}$. Hence there is a word $w' \in L$ such that $\pi_{w,\$} \models \text{EF}^{w'\$}\text{tt}$. As $w, w' \in \Sigma^*$ and $\$ \notin \Sigma$, we have $w' = w$. For the converse, let $w \in L$, then $\pi_{w,\$} \models \text{EF}^{L\$}\text{tt}$. Hence $w \in \mathcal{L}(\text{EF}^{L\$}\text{tt})$. \square

Lemma 21. *Let $L_0 \subseteq \Sigma_\* , $L \subseteq \Sigma^*$, let δ be a formula, and let ψ be a satisfiable formula. Suppose*

$$\text{EF}^{L\$}\text{tt} = \delta \vee \text{EF}^{L_0}\psi. \quad (16)$$

Define $L_1 := L_0 \cap \Sigma^$ and $L_2 := \{w \in \Sigma^*\$ \mid w \text{ is a prefix of a word in } L_0\}$. Then*

$$\delta \vee \text{EF}^{L_0}\psi = \delta \vee \text{EF}^{L_1}\psi \vee \text{EF}^{L_2}\text{tt}. \quad (17)$$

Additionally, the measure of (17.r) is weakly bounded by that of (17.l).

Proof. Case \rightarrow : Note that $\models \text{EF}^{L_0}\psi \rightarrow \text{EF}^{L_1}\psi \vee \text{EF}^{L_2}\text{tt}$ since for every word $w \in L_0 \setminus L_1$ there is a prefix of w in L_2 . **Case \leftarrow :** Since $L_1 \subseteq L_0$, $\models \text{EF}^{L_1}\psi \rightarrow \text{EF}^{L_0}\psi$ holds. So, we assume a model \mathcal{M} of EF^{L_2}tt . Hence there is a path π in \mathcal{M} labeled with a word $u\$ \in L_2$. Let $v \in \Sigma_\* such that $u\$v \in L_0$. At the end of the path, we attach a path labeled with v and on that one a model of ψ —note that ψ is assumed to be satisfiable. The new structure, say \mathcal{M}' , is a model of $\text{EF}^{L_0}\psi$. and also, by (16), of $\text{EF}^{L\$}\text{tt}$.

All (rooted) finite paths in \mathcal{M}' which not yet occur in \mathcal{M} passes the labels $u\$$. For the sake of contradiction, assume that \mathcal{M} is not a model of $\text{EF}^{L\$}\text{tt}$. Hence $u\$$ is a prefix of a word in $L\$$. So, $u \in L$ because $L \subseteq \Sigma^*$. Contradiction.

And as for the measure,

$$\begin{aligned} \mu(\text{EF}^{L_1}\psi) &= \|L_1\| \cdot \mu(\psi) \leq_{\mathbb{M}} \|L_0\| \cdot \mu(\psi) = \mu(\text{EF}^{L_0}\psi) \text{ and} \\ \mu(\text{EF}^{L_2}\text{tt}) &= \|L_2\| \leq_{\mathbb{M}} \|L_0\| \leq_{\mathbb{M}} \|L_0\| \cdot \mu(\psi) \end{aligned}$$

hold, and imply $\mu(\text{EF}^{L_1}\psi \vee \text{EF}^{L_2}\text{tt}) \leq_{\mathbb{M}} \mu(\text{EF}^{L_0}\psi)$. \square

Two remarks, to previous lemma: (1) If $\varepsilon \notin L_0$ then it is neither in L_1 nor in L_2 —but $\$ \in L_2$ might be. (2) If L_0 is a DCFL then so are L_1 and L_2 .

Theorem 22. *Let $P \subseteq \Sigma^*$, and let φ a $\text{PDL}_0[\text{DCFL}]$ -formula over $\Sigma_\$$. If $\varphi = \text{EF}^{P\$}\text{tt}$ then $\mathcal{L}(\varphi)$ is good.*

Proof. We apply to φ several transformations in sequence. Each transformation leads to a formula which is equivalent to φ and whose measure is weakly bounded by $\mu(\varphi)$ from above. The transformations are the following ones.

- Make the formula ε -free by Definition 13 and Lemma 14.
- Transform it into a DNF and complete the formula: Definition 15 and Lemma 16.
- Eliminate the outermost AG-quantifiers using Theorem 17.
- Apply Theorem 18 to each term of the DNF gotten from the previous transformation. Note that the applied formula is still ε -free.
- Apply Lemma 21 and its remarks to the outermost EF-formulas

Finally, we obtain a formula

$$\varphi' := \bigvee_{i \in I} \text{EF}^{L_i}\psi_i = \varphi \quad (18)$$

such that $\mu(\varphi') \leq_{\mathbb{M}} \mu(\varphi)$ holds. In addition, I is finite and for each $i \in I$ we have that

- $L_i \subseteq \Sigma^+$, or $L_i \subseteq \Sigma^*\$$ and $\psi_i = \text{tt}$,
- $L_i \neq \emptyset$, and
- L_i is a DCFL.

If $L_i \subseteq \Sigma^+$ then set $L_i^\natural := L_i$ and $R_i := \mathcal{L}(\psi_i)$, and else, $L_i^\natural := L_i/\$$ and $R_i := \{\varepsilon\}$. Note that L_i^\natural is DCFL in any case.

Claim 22-2. $\mathcal{L}(\varphi') = \bigcup_{i \in I} (L_i^\natural R_i)$.

Proof of claim. \subseteq : Let $w \in \mathcal{L}(\varphi')$. By (18), $\pi_w\$ \models \bigvee_{i \in I} \text{EF}^{L_i} \psi_i$. By case distinction and by using that $\$$ is not part of w , we have $w \in \bigcup_{i \in I} (L_i^\natural R_i)$. \supseteq : Let $w \in \bigcup_{i \in I} (L_i^\natural R_i)$. We have to show that $\pi_w\$ \models \bigvee_{i \in I} \text{EF}^{L_i} \psi_i$. There are two cases. First, if $w \in L_i^\natural R_i$ for some $i \in I$ with $L_i \subseteq \Sigma^*\$,$ then $\psi_i = \text{tt}$. Hence $\pi_w\$ \models \text{EF}^{L_i} \psi_i$. Second, if $w \in L_i^\natural R_i$ for some $i \in I$ with $L_i \subseteq \Sigma^+$, then $w = uv$ with $u \in L_i$ and $v \in R_i$. So, $\pi_v\$ \models \psi_i$ and thus $\pi_{uv\$} \models \text{EF}^{L_i} \psi_i$. \square

Thus, $\mathcal{L}(\varphi')$ is almost good. We have to exclude that there is an $i \in I$ with $|L_i^\natural| = 1$. Let $I^+ := \{i \in I \mid |L_i| > 1\}$, $I^- := \{i \in I \mid |L_i| = 1\}$, and $I_a^- := \{i \in I^- \mid a \text{ is a prefix of the sole word in } L_i\}$ for $a \in \Sigma$. Let $\Sigma^- := \{a \in \Sigma \mid I_a^- \neq \emptyset\}$. Note that $I = I^+ \cup I^-$ and that $\{I_a^-\}_{a \in \Sigma^-}$ forms a partitioning of I^- . For $a \in \Sigma^-$, set

$$\varphi_a := \bigvee_{i \in I^+} \text{EF}^{a \setminus L_i} \psi_i \vee \bigvee_{i \in I_a^-} \text{EF}^{a \setminus L_i} \psi_i.$$

As $a \setminus L_i = \emptyset$ for all $i \in I_b^-$ for $b \neq a$, the formula φ_a is equivalent to $\text{EF}^{a \setminus P\$} \text{tt}$. To apply the IH for $a \in \Sigma^-$, we have to ensure that $\mu(\varphi_a) <_{\mathbb{M}} \mu(\varphi')$. Indeed, $\mu(\text{EF}^{a \setminus L_i} \psi_i) \leq_{\mathbb{M}} \mu(\text{EF}^{L_i} \psi_i)$ for $i \in I^+$, and $\mu(\text{EF}^{a \setminus L_i} \psi_i) <_{\mathbb{M}} \mu(\text{EF}^{L_i} \psi_i)$ for $i \in I_a^- \neq \emptyset$. All in all, $\mu(\varphi_a) <_{\mathbb{M}} \mu(\varphi') \leq_{\mathbb{M}} \mu(\varphi)$ holds. We use the outcome of the IHs to replace the contributions of I^- to $\mathcal{L}(\varphi')$ by good languages.

$$\begin{aligned} P &= \mathcal{L}(\varphi') && \text{(by Lemma 20)} \\ &= \bigcup_{i \in I} (L_i^\natural R_i) && \text{(by Claim 22-2)} \\ &= \bigcup_{i \in I^+} (L_i^\natural R_i) \cup \bigcup_{a \in \Sigma^-} a \left[\bigcup_{i \in I^+} a \setminus L_i^\natural R_i \cup \bigcup_{i \in I_a^-} a \setminus L_i^\natural R_i \right] \\ &= \bigcup_{i \in I^+} (L_i^\natural R_i) \cup \bigcup_{a \in \Sigma^-} a [a \setminus P] \\ &= \bigcup_{i \in I^+} (L_i^\natural R_i) \cup \bigcup_{a \in \Sigma^-} a \mathcal{L}(\varphi_a) && \text{(by IH)} \end{aligned}$$

Also by IH, $\mathcal{L}(\varphi_a)$ is good. So, $\mathcal{L}(\varphi')$ is also good using the definition of I^+ . \square

Corollary 23. *Let $\varphi \in \text{PDL}_0[\text{DCFL}]$. If $\varphi = \text{EF}^{\text{Palindromes}\$} \text{tt}$ then the language Palindromes is good.*

Proof. By Theorem 22 and Lemma 20. \square

Corollary 24. $\text{PDL}_0[\text{DCFL}] \not\leq \text{PDL}_0[\text{CFL}]$.

Proof. By Corollaries 23 and 7. \square

9 Conclusion and Further Work

We proved that $\text{PDL}_0[\text{DCFL}]$ is distinct from $\text{PDL}_0[\text{CFL}]$ by means of model and language theory. Similar results—such as CTL vs. Fairness [8], PDL_0 vs. PDL [5], and unary CTL vs. unary CTL^+ [7]—uses two sequences of transition systems which are indistinguishable for the smaller logic. Their proofs are pretty compact. So, is it possible to reformulate our proof in a similar way? The main difficulty should be the incorporation of Theorem 3 into transition systems.

The considered logic is exactly the EF-/AG-fragment of the Extended Computation Tree Logic [2], say $\text{CTL}[\mathcal{L}]$. This observation poses at least two questions. First, is it possible to extend the separation from the unary fragment to the binary EU-/AR-fragment? Here, the main challenge is the interpretation of $\text{E}(\psi_1 \text{U}^L \psi_2)$ in the sense of Definition 19 as ψ_1 could prohibit linear models: take $\text{E}((p \wedge (\text{EF}^\Sigma \neg p)) \text{U}^L \psi_2)$ for instance. Secondly, one could go from one of these fragments to the whole logic to obtain a separation of $\text{CTL}[\text{DCFL}]$ and $\text{CTL}[\text{CFL}]$. In addition to the mentioned difficulties, one is faced with the alternating quantifiers EG and AF. To achieve such a goal, note that the Theorems 17 and 18 also hold for arbitrary path quantifications as long as ε -freedom is guaranteed. An iteration of these tools along a given ω -word could unravel an ω -sequence of disjunctions of E-formulas. Such a sequence could be a subject for a pumping lemma similar to Lemma 2. The ω -word could follow the lines of Theorem 3.

Finally, a separation of the full PDL (i.e. with tests) and of the Δ -variants of PDL [21, 17] could provide more insight into the difference between the non-determinism in CFLs and the non-determinism used in the translation of formulas into automata.

References

- [1] Rajeev Alur & Parthasarathy Madhusudan (2004): *Visibly pushdown languages*. In: *Proc. 36th Ann. ACM Symp. on Theory of Computing, STOC'04*, pp. 202–211, doi:10.1145/1007352.1007390.
- [2] Roland Axelsson, Matthew Hague, Stephan Kreutzer, Martin Lange & Markus Latte (2010): *Extended Computation Tree Logic*. In Christian Fermüller & Andrei Voronkov, editors: *Proceedings of the 17th International Conference on Logic for Programming, Artificial Intelligence and Reasoning (LPAR 2010, Yogyakarta, Indonesia)*, *Lecture Notes in Computer Science* 6397, Springer, pp. 67–81, doi:10.1007/978-3-642-16242-8_6.
- [3] Roland Axelsson, Matthew Hague, Stephan Kreutzer, Martin Lange & Markus Latte (2010): *Extended Computation Tree Logic*. *arXiv.org* Available at <http://arxiv.org/abs/1006.3709>.
- [4] Franz Baader & Tobias Nipkow (1998): *Term Rewriting and All That*. Cambridge University Press.
- [5] Francine Berman & Michael Paterson (1981): *Propositional dynamic logic is weaker without tests*. *Theoretical Computer Science* 16(3), pp. 321–328, doi:10.1016/0304-3975(81)90102-X.
- [6] Mikołaj Bojańczyk (2008): *The common fragment of ACTL and LTL*. In: *Proceedings of the Theory and practice of software, 11th international conference on Foundations of software science and computational structures, FOSSACS'08/ETAPS'08*, Springer-Verlag, Berlin, Heidelberg, pp. 172–185, doi:10.1007/978-3-540-78499-9_13.
- [7] E. Allen Emerson & Joseph Y. Halpern (1985): *Decision Procedures and Expressiveness in the Temporal Logic of Branching Time*. *Journal of Computer and System Sciences* 30, pp. 1–24, doi:10.1016/0022-0000(85)90001-7.
- [8] E. Allen Emerson & Joseph Y. Halpern (1986): “Sometimes” and “Not Never” Revisited: *On Branching versus Linear Time Temporal Logic*. *Journal of the ACM* 33, pp. 151–178, doi:10.1145/4904.4999.
- [9] Michael J. Fischer & Richard E. Ladner (1979): *Propositional Dynamic Logic of Regular Programs*. *Journal of Computer and System Sciences* 18(2), pp. 194–211.

- [10] Seymour Ginsburg & Sheila Greibach (1966): *Deterministic Context Free Languages*. *Information and Control* 9(6), pp. 620–648, doi:10.1016/S0019-9958(66)80019-0.
- [11] David Harel (1984): *Dynamic Logic*. In Dov M. Gabbay & F. Guentner, editors: *Handbook of Philosophical Logic, Vol. II: Extensions of Classical Logic*, chapter 10, D. Reidel Publishing Company, Dordrecht, Holland, pp. 497–604.
- [12] David Harel, Dexter Kozen & Jerzy Tiuryn (2000): *Dynamic Logic*. Foundations of Computing, MIT Press, Cambridge, Massachusetts.
- [13] David Harel, Amir Pnueli & Jonathan Stavi (1983): *Propositional Dynamic Logic of Nonregular Programs*. *Journal of Computer and System Sciences* 26(2), pp. 222–243.
- [14] David Harel & Danny Raz (1990): *Deciding properties of nonregular programs*. *Foundations of Computer Science, Annual IEEE Symposium on* , pp. 652–661.
- [15] John E. Hopcroft & Jeffrey D. Ullman (1979): *Introduction to Automata Theory, Languages and Computation*. Addison Wesley.
- [16] Dexter Kozen & Jerzy Tiuryn (1990): *Logics of Programs*. In: *Handbook of Theoretical Computer Science, Volume B: Formal Models and Semantics (B)*, Elsevier and MIT Press, pp. 789–840.
- [17] Christof Löding, Carsten Lutz & Olivier Serre (2007): *Propositional dynamic logic with recursive programs*. *J. Log. Algebr. Program.* 73(1-2), pp. 51–69, doi:10.1016/j.jlap.2006.11.003.
- [18] Christof Löding, Parthasarathy Madhusudan & Olivier Serre (2004): *Visibly Pushdown Games*. In: *Proc. 24th Int. Conf. on Foundations of Software Technology and Theoretical Computer Science, FSTTCS'04, LNCS 3328*, Springer, pp. 408–420, doi:10.1007/978-3-540-30538-5_34.
- [19] Michael O. Rabin (1970): *Weakly definable relations and special automata*. In Yehoshua Bar-Hillel, editor: *Mathematical Logic and Foundations of Set Theory*, North-Holland, Amsterdam, pp. 1–23.
- [20] Ludwig Staiger (1997): *ω -languages*. In Grzegorz Rozenberg & Arto Salomaa, editors: *Handbook of formal languages*, chapter 6, 3: beyond words, Springer, New York, NY, USA, pp. 339–387.
- [21] Robert S. Streett (1982): *Propositional Dynamic Logic of Looping and Converse is Elementarily Decidable*. *Information and Control* 54(1-2), pp. 121–141, doi:10.1016/S0019-9958(82)91258-X.
- [22] Wolfgang Thomas (1990): *Automata On Infinite Objects*. In: *Handbook of Theoretical Computer Science, Volume B: Formal Models and Semantics (B)*, Elsevier and MIT Press, pp. 133–191.
- [23] Moshe Y. Vardi & Pierre Wolper (1986): *Automata-Theoretic Techniques for Modal Logics of Programs*. *Journal of Computer and System Sciences* 32(2), pp. 183–221, doi:10.1016/0022-0000(86)90026-7.