Generalized Deterministic Languages and their Automata: A Characterization of Restricted Temporal Logic

Heinz Schmitz*
Theoretische Informatik
Universität Würzburg
email: *schmitz@informatik.uni-wuerzburg.de*

Abstract

Let $\mathrm{TL}[\mathbf{X},\mathbf{F}]$ ($\mathrm{TL}[\mathbf{F}]$) be the class of formulas of temporal logic, where as the only temporal operators $\mathbf{X}$ (*next*) and $\mathbf{F}$ (*eventually*) are allowed (only $\mathbf{F}$ is allowed, resp.). For a class $\Phi \subseteq \mathrm{TL}[\mathbf{X},\mathbf{F}]$ let $L(\Phi)$ be the class of Φ-definable languages. Among others, characterizations of $L(\mathrm{TL}[\mathbf{X},\mathbf{F}])$ and $L(\mathrm{TL}[\mathbf{F}])$ in terms of forbidden patterns in finite automata are known. Here we ask for every bound $k \geq 0$ on the number of nested uses of the next operator for the expressive power of the respective fragment of $\mathrm{TL}[\mathbf{X},\mathbf{F}]$. Denote by $\mathrm{TL}[\mathbf{X}(k),\mathbf{F}]$ the class of formulas in $\mathrm{TL}[\mathbf{X},\mathbf{F}]$ with nesting depth $\leq k$ in the next operator. Obviously, $L(\mathrm{TL}[\mathbf{X},\mathbf{F}]) = \bigcup_{k\geq 0} L(\mathrm{TL}[\mathbf{X}(k),\mathbf{F}])$ and $L(\mathrm{TL}[\mathbf{F}]) = L(\mathrm{TL}[\mathbf{X}(0),\mathbf{F}])$.

We prove a levelwise characterization of the above syntactical hierarchy (1) in terms of a certain pattern $\mathbf{S}_k^{rev}$ (cf. Figure 6) that *must not* appear in the transition graph of deterministic finite automata, and (2) in terms of a formal language representation involving k-leftdeterministic languages. With (1) we generalize the known forbidden pattern for $L(\mathrm{TL}[\mathbf{F}])$ which is $\mathbf{S}_0^{rev}$ and in (2) we carry over the notion of leftdeterministic languages to decompositions of words into consecutive blocks of length $k+1$. Our characterization enables us to give concise proofs of strictness and decidability results for the family of classes $\{\, L(\mathrm{TL}[\mathbf{X}(k),\mathbf{F}]) \mid k \geq 0 \,\}$. As a by-product, we reproof the known forbidden pattern characterization for $L(\mathrm{TL}[\mathbf{X},\mathbf{F}])$ in a novel way and we learn, how the pattern $\mathbf{S}_0^{rev}$ turns into the pattern for $L(\mathrm{TL}[\mathbf{X},\mathbf{F}])$ as level k increases.

All results hold dually for restricted past temporal logic, k-rightdeterministic languages and the pattern $\mathbf{S}_k$ (cf. Figure 5). We point out similarities and differences to languages having dot-depth one.

1 Introduction

Formulas of temporal logic (more precisely: propositional linear-time temporal logic) over an alphabet A are built up from the elements of A as atomic formulas, using the boolean connectives $\wedge, \vee, \neg$, the unary temporal operators $\mathbf{X}$ (*next*) and $\mathbf{F}$ (*eventually*) and the binary temporal operator $\mathbf{U}$ (*until*). The set of all such formulas is denoted as TL. It is known that the languages definable by formulas from TL are exactly the starfree regular languages, which play an important role in

*Supported by the Deutsche Forschungsgemeinschaft (DFG), grant Wa 847/4-1.

the theory of finite automata, the theory of finite semigroups, first-order theory and other fields of research, due to the various characterizations of starfree languages. For an overview on these rich theories and their surprising relations we refer to the articles [Brz76, Pin96a, Pin96b, Tho96] and to the textbooks [MP71, Pin86].

A fragment $\mathrm{TL}[.]$ of TL is a subset of TL where only the use of the temporal operators specified in brackets is allowed, e.g. $\mathrm{TL}[] \subseteq \mathrm{TL}[\mathbf{F}] \subseteq \mathrm{TL}[\mathbf{X},\mathbf{F}] \subseteq \mathrm{TL}[\mathbf{X},\mathbf{F},\mathbf{U}] = \mathrm{TL}$. For a class of formulas $\Phi \subseteq \mathrm{TL}$ we denote by $L(\Phi)$ the class of Φ-definable languages. An important question is the one asking for the expressive power of temporal operators, i.e. one studies the classes of definable languages when omitting one or the other such operator. For this end, characterizations in terms of *forbidden patterns* in transition graphs of deterministic finite automata are often aspired, since the absence of a subgraph in an automaton corresponds in a straightforward way to the fact that certain properties cannot be expressed with the restricted formalism in question. For more background on temporal logic and the study of strict fragments we refer to the recent survey paper [Wil99], where an up-to-date list of forbidden pattern results is given.

Here we are interested in the fragment $\mathrm{TL}[\mathbf{X},\mathbf{F}]$, when leaving out the until operator (also known as RTL, i.e. resticted temporal logic). The following characterization was given in [CPP93]: A language L belongs to $L(\mathrm{TL}[\mathbf{X},\mathbf{F}])$ if and only if the minimal deterministic finite automaton accepting L does not have pattern $\mathbf{S}^{rev}$ (cf. Figure 2). The proof rests on an algebraic characterization of the $\mathrm{TL}[\mathbf{X},\mathbf{F}]$-definable languages in terms of so-called locally $\mathcal{L}$-trivial monoids. It was reproved in [EW96] together with a result of the same type for $L(\mathrm{TL}[\mathbf{F}])$ involving pattern $\mathbf{S}_0^{rev}$ (cf. Figure 4). This time, completely different proofs were given, based on Ehrenfeucht-Fraissé games. So on one hand we are not allowed to specify the *next* event, on the other hand unrestricted use of the next operator is possible. We fill the room between these two positions and ask more detailed *how many* nested uses of the next operator $\mathbf{X}$ are needed to express a certain property with this formalism (the analog question concerning the nesting depth of the until operator was treated in [TW96]).

It turns out that we can give a comprehensive answer in two ways: (1) For each single bound $k \geq 0$ on the nesting depth of $\mathbf{X}$ we give a characterization in a forbidden pattern manner, and (2) we can give a concrete description in terms of formal languages. Let us first put a syntactical restriction on formulas and denote by $\mathrm{TL}[\mathbf{X}(k),\mathbf{F}]$ the class of formulas in $\mathrm{TL}[\mathbf{X},\mathbf{F}]$ with nesting depth $\leq k$ in the next operator. Then obviously we have $L(\mathrm{TL}[\mathbf{X},\mathbf{F}]) = \bigcup_{k\geq 0} L(\mathrm{TL}[\mathbf{X}(k),\mathbf{F}])$ and $L(\mathrm{TL}[\mathbf{F}]) = L(\mathrm{TL}[\mathbf{X}(0),\mathbf{F}])$. We will refer to the family of classes $\{\, L(\mathrm{TL}[\mathbf{X}(k),\mathbf{F}]) \mid k \geq 0 \,\}$ as the *next hierarchy* of restricted temporal logic.

Next we recall two notions from formal language theory, which – both combined – will be useful in our case. Let A be some finite alphabet. With a formula having next depth at most k we can make assertions about at most $k+1$ consecutive positions in a word. So we will look at a word w as a word over A^{k+1} by packing together each $k+1$ consecutive letters. Call this the k-decomposition of w. This approach was used before in several contexts, among them locally testable events (see [BS73, McN74]), dot-depth one languages (for which in [Ste85a] also a levelwise characterization was given), and it was mentioned recently in [TW98]. As a second concept we recall the notion of leftdeterministic languages. A language is called leftdeterministic if it is a finite union of languages $A_0^* a_1 A_1^* \cdots a_n A_n^*$ with $A_i \subseteq A$ and $a_i \notin A_{i-1}$. Such languages were used in [Eil76] (see also [Pin86]) to characterize the varieties of languages having $\mathcal{L}$-trivial syntactic monoids. A characterization in terms of a certain congruence was given in [BF80] and a forbidden pattern theorem was provided in [CPP93]. We adapt with little modifications the notion of leftdeterministic languages to k-decompositions of words and introduce k-leftdeterministic languages.

Let F be an arbitrary deterministic finite automaton accepting some language L. We prove for all $k \geq 0$ that L is a finite union of k-leftdeterministic languages if and only if F does not have pattern $\mathbf{S}_{\boldsymbol{k}}^{\boldsymbol{rev}}$ (cf. Figure 6) in its transition graph. Note that here the minimality condition is obsolet and that in case $k = 0$ we obtain pattern $\mathbf{S}_{\mathbf{0}}^{\boldsymbol{rev}}$ from above. Then we show that $L \in L(\mathrm{TL}[\mathbf{X}(k), \mathbf{F}])$ if and only if it is a finite union of k-leftdeterministic languages. Now the previous characterization in terms of automata enables us to give concise proofs of strictness and decidability results for the next hierarchy. As a by-product, we reproof in a novel way the known forbidden pattern $\mathbf{S}^{\boldsymbol{rev}}$ characterizing $L(\mathrm{TL}[\mathbf{X}, \mathbf{F}])$ and we learn, how the pattern $\mathbf{S}_{\mathbf{0}}^{\boldsymbol{rev}}$ turns into $\mathbf{S}^{\boldsymbol{rev}}$ as level k increases.

All results hold dually for resticted past temporal logic (RPTL), k-rightdeterministic languages and the pattern $\mathbf{S}_{\boldsymbol{k}}$ (cf. Figure 5). Patterns $\mathbf{S}_{\boldsymbol{k}}^{\boldsymbol{rev}}$ and $\mathbf{S}_{\boldsymbol{k}}$ appear to be very similar to two out of three patterns from [Ste85a] that levelwise characterize dot-depth one languages. We give a description of the classes defined by these patterns in terms of so-called *weak* k-leftdeterministic and *weak* k-rightdeterministic languages and show that even a slightly different pattern really makes a difference. Relations between $L(\mathrm{RTL})$ and $L(\mathrm{RPTL})$, and level one of the dot-depth hierarchy are pointed out in detail.

The paper is organized as follows. In Section 3 we review results on resticted temporal logic and recall notations. Moreover we make precise the notion of the next depth of a formula of resticted temporal logic and dually the notion of the previously depth of a formula of resticted past temporal logic. In Section 4 we turn to formal language theory and introduce generalized deterministic languages. For each $k \geq 0$ we define the classes of k-leftdetermistic languages and k-rightdeterministic languages. Section 5 is devoted prove for every k a forbidden pattern characterization of these classes. It follows that they are all different and that each of them is a Boolean algebra. In Section 6 we return to resticted temporal logic and show that k-rightdeterministic languages (k-leftdeterministic languages) are exactly the languages definable by formulas of next depth (previously depth, respectively) at most k. The decidability of these fragments can be easily followed. In Section 7 we compare our results with known characterizations of languages having dot-depth one.

2 Preliminaries

We consider languages over a finite alphabet A with $|A| \geq 2$. All definitions will be made for arbitrary but fixed alphabet A. For $k \geq 0$ we denote by A^k ($A^{\leq k}, A^{>k}, \ldots$) the set of words over A of length exactly k (less or equal to k, greater $k, \ldots$ respectively). We set $A^{<0} =_{\mathrm{def}} \emptyset$.

Let ϵ denote the empty word and let $|x|$ be the length of a word $x \in A^*$. Let $x, y \in A^*$. We call y a *prefix* (*suffix*) of x if there exists a word z such that $x = yz$ ($x = zy$, resp.). The *k-prefix* and the *k-suffix* of $x = a_1 \cdots a_n$ are given by

$$p_k(x) =_{\mathrm{def}} \begin{cases} a_1 a_2 \cdots a_k, & \text{if } n \geq k, \\ x, & \text{otherwise.} \end{cases}$$

and

$$s_k(x) =_{\text{def}} \begin{cases} a_{n-k+1}a_{n-k+2}\cdots a_n, & \text{if } n \geq k, \\ x, & \text{otherwise}. \end{cases}$$

Moreover y is a *factor* of x if there are word z_0, z_1 such that $x = z_0 y z_1$. We denote with x^R its *reverse*, i.e. $x^R =_{\text{def}} a_n a_{n-1} \cdots a_1$, and for a language L let $L^R =_{\text{def}} \{ x^R \mid x \in L \}$. Let $\alpha(x)$ be the set of all letters occuring in x.

As is standard, a deterministic finite automaton (dfa) F is given by $F = (A, S, \delta, s_0, S^{'})$, where A is its input alphabet, S is its set of states, $\delta : A \times S \to S$ is its transition function, $s_0 \in S$ is the starting state and $S^{'} \subseteq S$ is the set of accepting states. As usual we extend transition functions to input words, and we denote by $|F|$ the number of states of F.

3 Restricted Temporal Logic

We review results on restricted temporal logic and recall notations to get to the starting point of our investigations. With little modifications we follow [Wil99]. Moreover we introduce in the second part of this section the notion of the *next depth* of a formula.

Formulas of temporal logic (more precisely: propositional linear-time temporal logic) over alphabet A are built up from the elements of A as atomic formulas, using the boolean connectives $\wedge, \vee, \neg$, the unary temporal operators $\mathbf{X}$ (*next*) and $\mathbf{F}$ (*eventually*) and the binary temporal operator $\mathbf{U}$ (*until*). The set of all such formulas is denoted as TL. A fragment $\mathrm{TL}[.]$ of TL is a subset of TL where only the use of the temporal operators specified in brackets is allowed, e.g. $\mathrm{TL}[] \subseteq \mathrm{TL}[\mathbf{F}] \subseteq \mathrm{TL}[\mathbf{X}, \mathbf{F}, \mathbf{U}] = \mathrm{TL}$.

Here we are interested in the fragment $\mathrm{TL}[\mathbf{X}, \mathbf{F}]$ leaving out the *until* operator. This is known as resticed temporal logic (RTL) and we make all further definitions only for the restricted case. For a given word $w = a_1 a_2 \cdots a_n \in A^+$ and some $i \in \{1, \ldots, n\}$ we define what it means that an RTL formula φ is true in w at position i, in symbols $(w, i) \models \varphi$.

(1) $(w, i) \models a$ if and only if $a_i = a$, for all $a \in A$.

(2) $(w, i) \models \varphi \vee \psi$ if and only if $(w, i) \models \varphi$ or $(w, i) \models \psi$ (analogously for $\wedge, \neg$).

(3) $(w, i) \models \mathbf{X}\varphi$ if and only if $i < n$ and $(w, i + 1) \models \varphi$.

(4) $(w, i) \models \mathbf{F}\varphi$ if and only if there exists some j with $i < j \leq n$ and $(w, j) \models \varphi$.

Given an RTL formula φ and a word w we say that w is a model of φ, i.s. $w \models \varphi$, if and only if $(w, 1) \models \varphi$. We write $L(\varphi) =_{\text{def}} \{ w \in A^+ \mid w \models \varphi \}$ for the set of all models of φ and for a class of formulas $\Phi \subseteq \mathrm{RTL}$ we denote by $L(\Phi) =_{\text{def}} \{ L(\varphi) \mid \varphi \in \Phi \}$ the class of Φ-definable languages. Two formulas are called equivalent if they define the same language. Note that the *eventually* operator is defined here in such a way that the quantified position is strictly greater than the actual position.

So far we have defined the future version of restricted temporal logic. One can also think of a past version simply by reversing the ordering on the events. We will treat both cases in parallel and look also at formulas with a dual set of operators. If we substitute the temporal operators $\mathbf{X}$ by $\mathbf{Y}$ (*previously*) and $\mathbf{F}$ by $\mathbf{P}$ (*eventually in the past*) we obtain formulas of restricted past temporal logic (RPTL). Again, we define for a given word $w = a_1 a_2 \cdots a_n \in A^+$ and some $i \in \{1, \ldots, n\}$ what it means that such a formula φ is true in w at position i (with (1) and (2) as above).

(3') $(w, i) \models \mathbf{Y}\varphi$ if and only if $i > 1$ and $(w, i-1) \models \varphi$.

(4') $(w, i) \models \mathbf{P}\varphi$ if and only if there exists some j with $1 \leq j < i$ and $(w, j) \models \varphi$.

Given an RPTL formula φ and a word w we say that w is a model of φ, i.s. $w \models \varphi$, if and only if $(w, |w|) \models \varphi$. Furthermore we carry over all definitions introduced for RTL formulas in an obvious way to RPTL formulas, e.g. the definition of fragments and definable languages. We will ensure that it will be always clear from the context, e.g. indicated by the involved temporal operators, with what kind of formula we deal.

There is an easy way to turn an RTL formula into an RPTL formula and vice versa, such that the defined languages are just the reverse of one another and the syntactic structure of the formula is maintained. The *dual* of an RTL formula (RPTL formula) φ is the RPTL formula (RTL formula) $\overline{\varphi}$ where each $\mathbf{X}$ is substituted by $\mathbf{Y}$ ($\mathbf{Y}$ by $\mathbf{X}$) and each $\mathbf{F}$ by $\mathbf{P}$ ($\mathbf{P}$ by $\mathbf{F}$, respectively).

Proposition 3.1. *1. For any $w \in A^+$ and $\varphi \in \mathrm{RTL}$ it holds that $w \models \varphi \iff w^R \models \overline{\varphi}$.*

2. For any $w \in A^+$ and $\varphi \in \mathrm{RPTL}$ it holds that $w \models \varphi \iff w^R \models \overline{\varphi}$.

Proof. We only show 1 since 2 follows immediately from $\varphi = \overline{\overline{\varphi}}$. Let $\varphi \in \mathrm{RTL}$ and fix some $w = a_1 \cdots a_n \in A^+$. We show by induction on the structure of φ that for all $1 \leq i \leq n$ it holds that $(w, i) \models \varphi$ if and only if $(w^R, n-i+1) \models \overline{\varphi}$. Note that the i-th letter of w is the $(n-i+1)$-th letter of w^R, so if $\varphi = a$ for some $a \in A$ then $(w, i) \models a \iff a_i = a \iff (w^R, n-i+1) \models a$. If φ is a Boolean combination of formulas for which the induction hypothesis holds, there is nothing to prove. Now assume $\varphi = \mathbf{X}\psi$. Then

$$\begin{aligned} (w, i) \models \mathbf{X}\psi &\iff i < n \text{ and } (w, i+1) \models \psi \\ &\overset{\text{hyp}}{\iff} n - i > 0 \text{ and } (w^R, n-i) \models \overline{\psi} \\ &\iff (w^R, n-i+1) \models \mathbf{Y}\overline{\psi} = \overline{\varphi}. \end{aligned}$$

Finally, suppose $\varphi = \mathbf{F}\psi$. Then we can conclude

$$\begin{aligned} (w, i) \models \mathbf{F}\psi &\iff \text{there exists some } i < j \leq n \text{ with } (w, j) \models \psi \\ &\overset{\text{hyp}}{\iff} \text{there exists some } 1 \leq n-j+1 < n-i+1 \text{ with } (w^R, n-j+1) \models \overline{\psi} \\ &\iff (w^R, n-i+1) \models \mathbf{P}\overline{\psi} = \overline{\varphi}. \end{aligned}$$

In particular we have shown $(w, 1) \models \varphi$ if and only if $(w^R, n) \models \overline{\varphi}$. ❑

Let $\Phi \in \{\mathrm{TL}[\mathbf{F}], \mathrm{TL}[\mathbf{P}], \mathrm{TL}[\mathbf{X}, \mathbf{F}], \mathrm{TL}[\mathbf{Y}, \mathbf{P}]\}$. We restate known characterizations of the classes of Φ-definable languages in terms of forbidden patterns in the transition graphs of deterministic finite automata and define the involved patterns.

A dfa $F = (A, S, \delta, s_0, S')$ has pattern $\mathbf{S}$ if there are states $s_1, s_2, s, s' \in S$ and words $x, w, u, v, z \in A^*$ with $|w| \geq 1$, $s \in S' \Leftrightarrow s' \notin S'$, $\delta(s_0, x) = s_1 = \delta(s_1, w) = \delta(s_2, v)$, $\delta(s_1, u) = s_2 = \delta(s_2, w)$, $\delta(s_1, z) = s$ and $\delta(s_2, z) = s'$ (cf. Figure 1).

A dfa F has pattern $\mathbf{S}^{\boldsymbol{rev}}$ if there are states $s_1, s_2, s_3, s_4, s_5, s, s' \in S$ and words $x, w, u, v, z \in A^*$ with $|w| \geq 1$, $s \in S' \Leftrightarrow s' \notin S'$, $\delta(s_0, x) = s_1$, $\delta(s_i, w) = s_i$ for $1 \leq i \leq 5$, $\delta(s_1, u) = s_2 = \delta(s_4, u)$, $\delta(s_1, v) = s_3 = \delta(s_5, v)$, $\delta(s_2, v) = s_4$, $\delta(s_3, u) = s_5$, $\delta(s_2, z) = s$ and $\delta(s_5, z) = s'$ (cf. Figure 2).

A dfa F has pattern $\mathbf{S_0}$ if there are states $s_1, s_2, s_3, s_4, s, s' \in S$ and words $x, u, v, z \in A^*$ and $a \in A$ with $s \in S' \Leftrightarrow s' \notin S'$, $\delta(s_0, x) = s_1 = \delta(s_2, v)$, $\delta(s_1, u) = s_2$, $\delta(s_1, a) = s_3$, $\delta(s_2, a) = s_4$, $\delta(s_3, z) = s$ and $\delta(s_4, z) = s'$ (cf. Figure 3).

A dfa F has pattern $\mathbf{S_0^{rev}}$ if there are states $s_1, s_2, s_3, s_4, s_5, s_6, s, s' \in S$ and words $x, u, v, z \in A^*$ and $a \in A$ with $s \in S' \Leftrightarrow s' \notin S'$, $\delta(s_0, x) = s_1$, $\delta(s_1, a) = s_2$, $\delta(s_2, u) = s_3 = \delta(s_5, u)$, $\delta(s_2, v) = s_4 = \delta(s_6, v)$, $\delta(s_3, v) = s_5$, $\delta(s_4, u) = s_6$, $\delta(s_3, z) = s$ and $\delta(s_6, z) = s'$ (cf. Figure 4).

Statements 1 and 2 in the following theorem were first proved in [CPP93] and reproved in [EW96] together with a characterization of $L(\mathrm{TL}[\mathbf{F}])$.

Theorem 3.2 [CPP93, EW96]. *Let $L \subseteq A^+$ and F be the minimal dfa with $L(F) = L$. Then the following statements hold.*

1. *$L \in L(\mathrm{TL}[\mathbf{Y}, \mathbf{P}])$ if and only if F does not have pattern $\mathbf{S}$.*
2. *$L \in L(\mathrm{TL}[\mathbf{X}, \mathbf{F}])$ if and only if F does not have pattern $\mathbf{S^{rev}}$.*
3. *$L \in L(\mathrm{TL}[\mathbf{P}])$ if and only if F does not have pattern $\mathbf{S_0}$.*
4. *$L \in L(\mathrm{TL}[\mathbf{F}])$ if and only if F does not have pattern $\mathbf{S_0^{rev}}$.*

Originally, in [EW96] a combination of 3 and 4 was stated as $L \in L(\mathrm{TL}[\mathbf{F}])$ if and only if the minimal automaton accepting L^R does not have pattern $\mathbf{S_0}$. For systematic reasons we introduce here the pattern $\mathbf{S_0^{rev}}$ and give a proof later in connection with more general theorems (cf. Theorem 6.3 in combination with Theorem 5.6). Also, decision algorithms for the mentioned classes were pointed out by the authors.

We take a closer look at the fine structure of the class of $\mathrm{TL}[\mathbf{X}, \mathbf{F}]$-definable languages and introduce the notion of the *next depth* of an RTL formula, i.e. we count the number of nested uses of the $\mathbf{X}$ operator. If φ is an RTL formula then $nd(\varphi)$ denotes its next depth.

Definition 3.3. Let $\varphi \in \mathrm{RTL}$. We define inductively

- $nd(a) =_{\mathrm{def}} 0$ for all $a \in A$,
- $nd(\neg\varphi) =_{\mathrm{def}} nd(\varphi)$ and $nd(\mathbf{F}\varphi) =_{\mathrm{def}} nd(\varphi)$,
  $nd(\varphi \vee \psi) =_{\mathrm{def}} max\{nd(\varphi), nd(\psi)\}$ and $nd(\varphi \wedge \psi) =_{\mathrm{def}} max\{nd(\varphi), nd(\psi)\}$,
  $nd(\mathbf{X}\varphi) =_{\mathrm{def}} nd(\varphi) + 1$.

Let $k \geq 0$ and denote with $\mathrm{TL}[\mathbf{X}(k), \mathbf{F}] =_{\mathrm{def}} \{\varphi \in \mathrm{TL}[\mathbf{X}, \mathbf{F}] \mid nd(\varphi) \leq k\}$ the set of all RTL formulas φ having next depth at most k. Obviously $\mathrm{TL}[\mathbf{F}] = \mathrm{TL}[\mathbf{X}(0), \mathbf{F}]$ and $\mathrm{RTL} = \mathrm{TL}[\mathbf{X}, \mathbf{F}] = \bigcup_{k \geq 0} \mathrm{TL}[\mathbf{X}(k), \mathbf{F}]$.

Fortunately the next operator $\mathbf{X}$ shows a nice property: it commutes with the Boolean connectives and with $\mathbf{F}$ (even with $\mathbf{U}$). The following switching rules (cf.[Eme90]) will allow us to bring all $\mathbf{X}$ operators next to atomic formulas.

Proposition 3.4. *Let $\varphi, \psi \in \mathrm{RTL}$ and let $w = a_1 \cdots a_n \in A^+$. Then for all $1 \leq i \leq n$ the following holds.*

1. *$(w, i) \models \mathbf{XF}\varphi \Longleftrightarrow (w, i) \models \mathbf{FX}\varphi$.*
2. *$(w, i) \models \mathbf{X}(\varphi \vee \psi) \Longleftrightarrow (w, i) \models (\mathbf{X}\varphi \vee \mathbf{X}\psi)$.*
3. *$(w, i) \models \mathbf{X}(\varphi \wedge \psi) \Longleftrightarrow (w, i) \models (\mathbf{X}\varphi \wedge \mathbf{X}\psi)$.*

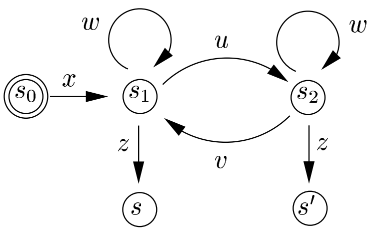

Figure 1: Pattern **S** with $|w| \geq 1$ and $s \in S^{'} \Leftrightarrow s' \notin S^{'}$.

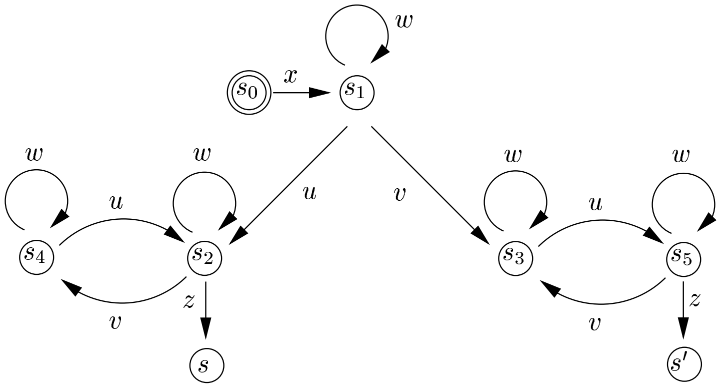

Figure 2: Pattern $\mathbf{S}^{\boldsymbol{rev}}$ with $|w| \geq 1$ and $s \in S^{'} \Leftrightarrow s' \notin S^{'}$.

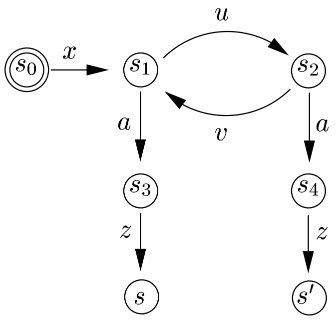

Figure 3: Pattern $\mathbf{S_0}$ with $a \in A$ and $s \in S^{'} \Leftrightarrow s' \notin S^{'}$.

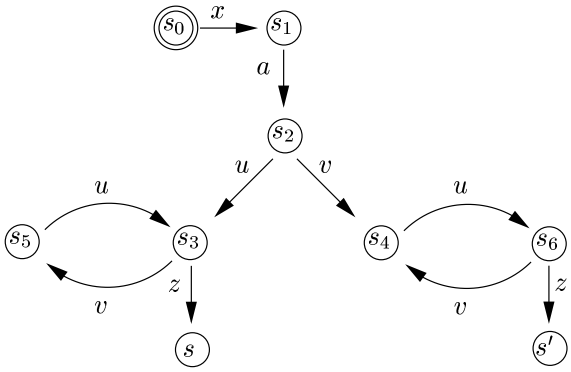

Figure 4: Pattern $\mathbf{S_0^{\boldsymbol{rev}}}$ with $a \in A$ and $s \in S^{'} \Leftrightarrow s' \notin S^{'}$.

4. $(w,i) \models \mathbf{X}\neg\varphi \iff (w,i) \models ((\neg\mathbf{X}\varphi) \wedge (\mathbf{X}\bigvee_{a\in A} a))$.

Proof. We show only 1 and 4, so let $\varphi \in \mathrm{RTL}$ and $w = a_1 \cdots a_n \in A^+$. Suppose $(w,i) \models \mathbf{XF}\varphi$ for some i with $1 \leq i \leq n$. So $i < n$ and there exists some j with $i+1 < j \leq n$ such that $(w,j) \models \varphi$. Then $j' =_{\text{def}} j-1$ witnesses that $(w,i) \models \mathbf{FX}\varphi$. On the other hand assume $(w,i) \models \mathbf{FX}\varphi$, then there is some j with $i < j \leq n$ such that $j < n$ and $(w, j+1) \models \varphi$. For $j' =_{\text{def}} j+1$ it holds that $i+1 < j' \leq n$ and $(w,j') \models \varphi$.

To see 4 suppose $(w,i) \models \mathbf{X}\neg\varphi$. Then $i < n$ and it is not the case that $(w,i+1) \models \varphi$. Additionally, since $i < n$ it holds that $(w,i) \models (\mathbf{X}\bigvee_{a\in A} a)$. On the other hand assume $(w,i) \models (\neg\mathbf{X}\varphi)$ and $(w,i) \models (\mathbf{X}\bigvee_{a\in A} a)$. From the first part we known that $i < n$ is not true or it is not the case that $(w,i+1) \models \varphi$. From the second part we conclude that $i < n$ so together we obtain $i < n$ and it is not the case that $(w,i+1) \models \varphi$. ❑

Note that in 4 we have to ensure that there is at least one more letter right to the actual position in order to establish the implication from right to left. It is important to see that these rules preserve the next depth of the formula in question. This gives rise to the definition of the set $X_k =_{\text{def}} \left\{ \underbrace{\mathbf{X}\cdots\mathbf{X}}_{l\text{-times}} a \;\middle|\; a \in A, 0 \leq l \leq k \right\}$ of all atomic formulas a with a prefix of at most k next operators. Denote by $\langle X_k \rangle$ the set of all RTL formulas built up from elements of X_k using $\vee, \wedge, \neg$ and $\mathbf{F}$.

Proposition 3.5. *For all $k \geq 0$ it holds that $L(\mathrm{TL}[\mathbf{X}(k), \mathbf{F}]) = L(\langle X_k \rangle)$.*

Proof. The inclusion from left to right holds since we can bring any RTL formula with next depth at most k with the help of Proposition 3.4 in the desired form and preserve the defined language. The inclusion from right to left is trivial. ❑

We will refer to $\left\{ L(\langle X_k \rangle) \;\middle|\; k \geq 0 \right\}$ as the *next hierarchy*. If we simply substitute in the definition of the next depth the operators $\mathbf{X}$ by $\mathbf{Y}$ and $\mathbf{F}$ by $\mathbf{P}$ we come to the notion of the *previously depth* $pd(\varphi)$ of an RPTL formula φ. In the same way one verifies that Propositions 3.4 and 3.5 hold as well with the appropriate definitions $\mathrm{TL}[\mathbf{Y}(k), \mathbf{P}]$, Y_k and also $\langle Y_k \rangle$ for RPTL formulas. With Proposition 3.1 we easily get a close connection between the classes of the next hierarchy and the classes of the previously hierarchy for each bound k on the next depth and the previously depth.

Proposition 3.6. *For all $k \geq 0$ and $L \subseteq A^+$ it holds that $L \in L(\langle X_k \rangle)$ if and only if $L^R \in L(\langle Y_k \rangle)$.*

4 k-Decomposition and Deterministic Languages

With a formula having next depth k we can make assertions about $k+1$ consecutive positions in a word. So it will be useful for us to look at a word $w \in A^+$ as a word over A^{k+1} by packing together each $k+1$ consecutive letters. A similar approach was used in [Ste85a] to obtain a levelwise characterization of languages of dot-depth one (see also Section 7) and was also mentioned in [TW98].

We start with defining what the k-decomposition of a word is. In order to avoid confusion we will denote elements from A^{k+1} as $\alpha, \beta, \gamma, \ldots$ and subsets of A^{k+1} as $\Sigma, \Gamma, \ldots$. Let $w = a_1 a_2 \cdots a_{k+l} \in A^+$ for some $l \geq 1$. We call

$$\widehat{w} =_{\text{def}} (\alpha_1, \alpha_2, \ldots, \alpha_l)$$

the k-*decomposition* of w if $\alpha_i = a_i \cdots a_{i+k}$ for $1 \leq i \leq l$. The number k indicates by how many letters from A consecutive α_i overlap. We set $\alpha(\widehat{w}) =_{\text{def}} \{\alpha_1, \alpha_2, \ldots, \alpha_l\}$ as the set of elements from A^{k+1} occuring in the k-decomposition of w.

Next we want to define languages of words of A^+ that admit the same k-decomposition with respect to given elements and subsets of A^{k+1}.

Definition 4.1. Let $k \geq 0$. Let $\alpha_1, \ldots, \alpha_{n+1} \in A^{k+1}$ and $\Sigma_1, \ldots, \Sigma_{n+1} \subseteq A^{k+1}$ for some $n \geq 0$. For every $w \in A^+$ we say $w \in (\alpha_1, \Sigma_1, \ldots, \alpha_n, \Sigma_n, \alpha_{n+1}, \Sigma_{n+1})_k$ if and only if $|w| \geq n + k + 1$, $\widehat{w} = (\beta_1, \ldots, \beta_l)$ and there exist $1 = j_1 < j_2 < \ldots < j_{n+1} \leq l$ such that for all $1 \leq i \leq n+1$ it holds that

(a) $\beta_{j_i} = \alpha_i$, and
(b) $\beta_j \in \Sigma_i$ for $j_i < j < j_{i+1}$(with $j_{n+2} =_{\text{def}} l + 1$).

So the language $(\alpha_1, \Sigma_1, \alpha_2, \Sigma_2, \ldots, \alpha_{n+1}, \Sigma_{n+1})_k$ consists of those words $w \in A^+$, whose k-decomposition starts with α_1, followed by a number (possibly zero) of elements from Σ_1, then α_2 and so on. Completely analogously one defines e.g. languages $(\Sigma_1, \alpha_1, \Sigma_2, \alpha_2, \ldots, \alpha_n, \Sigma_{n+1}, \alpha_{n+1})_k$ and we will even allow consecutive elements of A^{k+1} in a language description. So $(\alpha_1, \alpha_2, \Sigma_2, \alpha_3, \alpha_4, \alpha_5, \ldots, \Sigma_{n+1})_k$ is the set of words w, whose k-decomposition starts with α_1, immediately followed by α_2, then a number of elements from Σ_2, followed by α_3, α_4 and α_5 and so on. Note that in case $k = 0$ we deal with the usual concatenation, e.g. $(a_1, A_1, \ldots a_{n+1}, A_{n+1})_0 = a_1 A_1^* \cdots a_{n+1} A_{n+1}^*$.

Of special interest for us are languages that admit a unique decomposition of words with respect to their description over A^{k+1}. Since the above definition refers only to words of length at least $k+1$ we add in the following definition a finite set D of words of shorter length.

Definition 4.2. Let $k \geq 0$ and $L \subseteq A^+$.

1. We say that L is k-*leftdeterministic* if and only if there exist $\alpha_1, \ldots, \alpha_{n+1} \in A^{k+1}$ and $\Sigma_1, \ldots, \Sigma_{n+1} \subseteq A^{k+1}$ for some $n \geq 0$ and a set $D \subseteq A^{\leq k} \backslash \{\epsilon\}$ such that for $1 \leq i \leq n$ it holds that $\alpha_i \notin \Sigma_i$ and $L = (\Sigma_1, \alpha_1, \ldots, \Sigma_n, \alpha_n, \Sigma_{n+1}, \alpha_{n+1})_k \cup D$.

2. We say that L is k-*rightdeterministic* if and only if there exist $\alpha_1, \ldots, \alpha_{n+1} \in A^{k+1}$ and $\Sigma_1, \ldots, \Sigma_{n+1} \subseteq A^{k+1}$ for some $n \geq 0$ and a set $D \subseteq A^{\leq k} \backslash \{\epsilon\}$ such that for $2 \leq i \leq n+1$ it holds that $\alpha_i \notin \Sigma_i$ and $L = (\alpha_1, \Sigma_1, \alpha_2, \Sigma_2, \ldots, \alpha_{n+1}, \Sigma_{n+1})_k \cup D$

We will refer to these languages as k-deterministic languages. Note that the requirement $\alpha_i \notin \Sigma_i$ does not range over the last (first) index. Deterministic languages for the case $k = 0$ and sets of the form $A_0^* a_1 A_1^* \cdots a_n A_n^*$ were used in [Eil76] (see also [Pin86]) to characterize the varieties of languages having $\mathcal{R}$-trivial and $\mathcal{L}$-trivial syntactic monoids. As we will see, our generalization yields a levelwise characterization of RTL and RPTL-definable languages which are known to be exactly the languages having locally $\mathcal{R}$-trivial and locally $\mathcal{L}$-trivial syntactic semigroups (cf. [CPP93]). Moreover we can think of k-deterministic languages as a special case of unambiguous languages (cf. [PW97], see also [TW98]). We define the classes of k-leftdeterministic languages and k-rightdeterministic languages.

Definition 4.3. Let $k \geq 0$ and $L \subseteq A^+$.

1. $L \in \mathcal{D}_k^{\text{left}}$ if and only if L is a finite union of k-leftdeterministic languages.

2. $L \in \mathcal{D}_k^{\text{right}}$ if and only if L is a finite union of k-rightdeterministic languages.

A first easy to see property of k-deterministic languages is stated in the following proposition.

Proposition 4.4. *For every $k \geq 0$ and $L \subseteq A^+$ it holds that $L \in \mathcal{D}_k^{\text{left}}$ if and only if $L^R \in \mathcal{D}_k^{\text{right}}$.*

Proof. Let $L \in \mathcal{D}_k^{\text{left}}$. Then L is a finite union of k-leftdeterministic languages L'. Let $L' = (\Sigma_1, \alpha_1, \ldots, \Sigma_n, \alpha_n, \Sigma_{n+1}, \alpha_{n+1})_k \cup D$ with $n \geq 0$, $D \subseteq A^{\leq k} \backslash \{\epsilon\}$, $\alpha_1, \ldots, \alpha_{n+1} \in A^{k+1}$, $\Sigma_1, \ldots, \Sigma_{n+1} \subseteq A^{k+1}$ and $\alpha_i \notin \Sigma_i$ for $1 \leq i \leq n$. After a moment of reflection we see that

$$w \in L' \Longleftrightarrow w^R \in (\beta_1, \Gamma_1, \beta_2, \Gamma_2, \ldots, \beta_{n+1}, \Gamma_{n+1})_k \cup D^R$$

with $\beta_i =_{\text{def}} \alpha_{n-i+2}^R$ and $\Gamma_i =_{\text{def}} \Sigma_{n-i+2}^R$ for $1 \leq i \leq n+1$. Moreover it holds that $\beta_i \notin \Gamma_i$ for $2 \leq i \leq n+1$. The if-part follows analogously. ❑

In the sequel of this section we demonstrate the effect of the property $\alpha_i \notin \Sigma_i$ in k-deterministic languages. The following lemma basically sais that the partitioning of a word w witnessing $w \in L$, where L is such that the condition $\alpha_i \notin \Sigma_i$ ranges over all occuring α_i, must be the same for all $wv \in L$.

Lemma 4.5. *Let $k, s \geq 0$. Let $\alpha_1, \ldots, \alpha_s \in A^{k+1}$ and $\Sigma_1, \ldots, \Sigma_{s+1} \subseteq A^{k+1}$ with $\alpha_i \notin \Sigma_i$ for $1 \leq i \leq s$. Furthermore let $w, v \in A^+$ with $\widehat{w} = (\beta_1, \ldots, \beta_l)$ for some $l \geq 1$ and $\widehat{wv} = (\beta_1, \ldots, \beta_l, \beta_{l+1}, \ldots, \beta_{l+m})$ for some $m \geq 1$. If w and wv belong to $(\Sigma_1, \alpha_1, \ldots, \Sigma_s, \alpha_s, \Sigma_{s+1})_k$ then $\beta_{l+i} \in \Sigma_{s+1}$ for $1 \leq i \leq m$.*

Proof. Suppose all prerequisites of the lemma are given. Since $w \in L$ there are $1 \leq j_1 < j_2 < \ldots < j_s \leq l$ with $\beta_{j_i} = \alpha_i$ and $\beta_j \in \Sigma_i$ with $j_i < j < j_{i+1}$ for all $0 \leq i \leq s$, where we set $j_0 =_{\text{def}} 0$ and $j_{s+1} =_{\text{def}} l+1$. The same holds for $wv \in L$, so there are $1 \leq j'_1 < j'_2 < \ldots < j'_s \leq l+m$ with $\beta_{j'_i} = \alpha_i$ and $\beta_j \in \Sigma_i$ with $j'_i < j < j'_{i+1}$ for all $0 \leq i \leq s$. Again we set $j'_0 =_{\text{def}} 0$ and $j'_{s+1} =_{\text{def}} l+m+1$.

We show $j_i = j'_i$ for all $0 \leq i \leq s$ by induction on i. The induction base is clear by definition, so assume that $j_{i+1} \neq j'_{i+1}$ for some i with $0 \leq i < s$. W.l.o.g. we may suppose that $j'_{i+1} > j_{i+1}$. Moreover we know that $\beta_{j'_i+1}, \ldots, \beta_{j'_{i+1}-1} \in \Sigma_{i+1}$. Now since by hypothesis $j_i = j'_i$ and $j'_{i+1} > j_{i+1}$ this sequence in fact starts with elements $\beta_{j_i+1}, \ldots, \beta_{j_{i+1}-1}, \beta_{j_{i+1}}$ and we get $\beta_{j_{i+1}} = \alpha_{i+1} \in \Sigma_{i+1}$, a contradiction.

Especially we have $j_s = j'_s \leq l < l+i$ for $1 \leq i \leq m$. So from $\beta_j \in \Sigma_{s+1}$ for $j'_s < j \leq l+m$ we can conclude $\beta_{l+i} \in \Sigma_{s+1}$ for $1 \leq i \leq m$. ❑

As a special case we look at the situation when $\Sigma_{s+1} = \emptyset$, i.e. we require that the k-decomposition of w ends with α_s.

Lemma 4.6. *Let $k, s \geq 0$. Let $\alpha_1, \ldots, \alpha_s \in A^{k+1}$ and $\Sigma_1, \ldots, \Sigma_s \subseteq A^{k+1}$ with $\alpha_i \notin \Sigma_i$ for $1 \leq i \leq s$. Furthermore let $w \in A^+$. If $w \in (\Sigma_1, \alpha_1, \ldots, \Sigma_s, \alpha_s)_k$ then there exists no $v \in A^+$ with $wv \in (\Sigma_1, \alpha_1, \ldots, \Sigma_s, \alpha_s)_k$*

Proof. Again suppose all prerequisites of the lemma are given. Let $w \in (\Sigma_1, \alpha_1, \ldots, \Sigma_s, \alpha_s)_k$ and assume there exists some $v \in A^+$ with $wv \in (\Sigma_1, \alpha_1, \ldots, \Sigma_s, \alpha_s)_k$. Since $w, wv \in (\Sigma_1, \alpha_1, \ldots, \Sigma_s, \alpha_s, \emptyset)_k$ we can apply Lemma 4.5. Moreover $v \in A^+$ so the k-decomposition of wv is stricly longer than the k-decomposition of w, say by $\beta_{l+1}, \ldots, \beta_{l+m} \in A^{k+1}$ extra elements for some $m \geq 1$. Now it must be that $\beta_{l+1}, \ldots, \beta_{l+m} \in \Sigma_{s+1} = \emptyset$, which is a contradiction. ❑

When taking the reverse of the occuring elements and subsets of A^{k+1} as in the proof of Proposition 4.4 we can easily derive the dual versions of the previous two lemmata.

Lemma 4.7. *Let $k, s \geq 0$. Let $\alpha_1, \dots, \alpha_s \in A^{k+1}$ and $\Sigma_0, \Sigma_1, \dots, \Sigma_s \subseteq A^{k+1}$ with $\alpha_i \notin \Sigma_i$ for $1 \leq i \leq s$. Furthermore let $v, w \in A^+$ with $\widehat{vw} = (\beta_1, \dots, \beta_m, \beta_{m+1}, \dots, \beta_{m+l})$ for some $m, l \geq 1$ such that $\widehat{w} = (\beta_{m+1}, \dots, \beta_{m+l})$. If w and vw belong to $(\Sigma_0, \alpha_1, \Sigma_1, \dots, \alpha_s, \Sigma_s)_k$ then $\beta_i \in \Sigma_0$ for $1 \leq i \leq m$.*

Lemma 4.8. *Let $k, s \geq 0$. Let $\alpha_1, \dots, \alpha_s \in A^{k+1}$ and $\Sigma_1, \dots, \Sigma_s \subseteq A^{k+1}$ with $\alpha_i \notin \Sigma_i$ for $1 \leq i \leq s$. If $w \in (\alpha_1, \Sigma_1, \dots, \alpha_s, \Sigma_s)_k$ then there exists no $v \in A^+$ with $vw \in (\alpha_1, \Sigma_1, \dots, \alpha_s, \Sigma_s)_k$.*

5 Forbidden Pattern Characterizations of $\mathcal{D}_k^{\text{right}}$ and $\mathcal{D}_k^{\text{left}}$

We define more patterns. Let $F = (A, S, \delta, s_0, S')$ be a dfa and let $k \geq 0$. Then F has pattern $\mathbf{S}_k$ if there are states $s_1, \dots, s_6, s, s' \in S$ and words $x, u, v, z \in A^*$, $w \in A^k$ and $a \in A$ with $s \in S' \Leftrightarrow s' \notin S'$, $\delta(s_0, x) = s_1 = \delta(s_4, w)$, $\delta(s_1, u) = s_2$, $\delta(s_2, w) = s_3$, $\delta(s_3, v) = s_4$, $\delta(s_1, a) = s_5$, $\delta(s_3, a) = s_6$, $\delta(s_5, z) = s$ and $\delta(s_6, z) = s'$ (cf. Figure 5).

A dfa F has pattern $\mathbf{S}_k^{rev}$ if there are states $s_1, \dots, s_7, s, s' \in S$ and words $x, u, v, z \in A^*$, $w \in A^k$ and $a \in A$ with $s \in S' \Leftrightarrow s' \notin S'$, $\delta(s_0, x) = s_1$, $\delta(s_1, a) = s_2$, $\delta(s_2, w) = s_3$, $\delta(s_3, u) = s_4 = \delta(s_6, wu)$, $\delta(s_3, v) = s_5 = \delta(s_7, wv)$, $\delta(s_4, wv) = s_6$, $\delta(s_5, wu) = s_7$, $\delta(s_4, z) = s$ and $\delta(s_7, z) = s'$ (cf. Figure 6).

Note that for $k = 0$ these definitions are identical to the previous definitions of $\mathbf{S_0}$ and $\mathbf{S_0^{rev}}$.

Let $F = (A, S, \delta, s_0, S')$ be an arbitrary dfa and let $k \geq 0$. Our goal is to obtain a description of $L(F)$ as a finite union of k-leftdeterministic languages under the assumption that F does not have pattern $\mathbf{S}_k$. To do so we will proceed stepwise and start with a description of $L(F)$ as a finite union of languages $L = w_0 L_1 w_1 L_2 \cdots L_n w_n L_{n+1}$ where each L_i is a set of words that all remain in the same strongly connected component of the transition graph of F and that fulfill certain other requirements. We will fix two states p and q where each word enters and leaves the component and we will ensure that all words in L_i have at least length k, have the same k-prefix and have the same k-suffix.

A strongly connected component (SCC) of the transition graph of F is a maximal set of states $C \subseteq S$ such that there is a path between any two states in C. Let $p \in S$ and set $L(p) =_{\text{def}} \{ w \in A^+ \mid \delta(s_0, w) = p \}$. We immediately obtain

$$L(F) = \bigcup_{s \in S'} L(s). \tag{1}$$

If C is an SCC and $p, q \in C$ we define $L(p, C, q) =_{\text{def}} \{ w \in A^* \mid \delta(p, w) = q \}$ as the set of words that remain in C between two particular states. For any $p \in S$ it holds that

$$L(p) = \bigcup L(p_1, C_1, q_1) a_1 L(p_2, C_2, q_2) a_2 \cdots L(p_n, C_n, q_n) a_n L(p_{n+1}, C_{n+1}, q_{n+1}) \tag{2}$$

where the union goes over any sequence of distinct SCC's $C_1, \dots, C_{n+1}$ with $n \geq 0$ such that $p_1 = s_0$, $q_{n+1} = p$, $p_i, q_i \in C_i$ for $1 \leq i \leq n+1$ and $\delta(q_i, a_i) = p_{i+1}$ for $1 \leq i \leq n$. The inclusion from right to left holds by definition. For the other inclusion observe that for any $w \in L(p)$ we can find appropriate witnesses $C_1, \dots, C_{n+1}$ on the path from s_0 to p given by w. These SCC's are all distinct since the graph of SCC's is acyclic due to the maximality of an SCC.

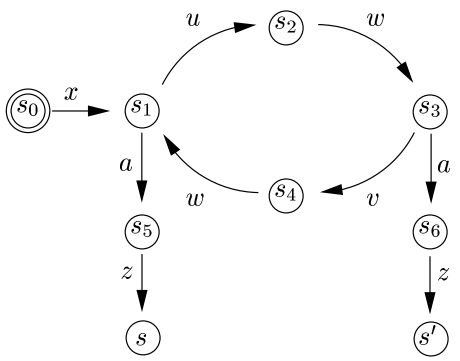

Figure 5: Pattern $\mathbf{S}_{\boldsymbol{k}}$ with $|w| = k$, $a \in A$ and $s \in S^{'} \Leftrightarrow s' \notin S^{'}$.

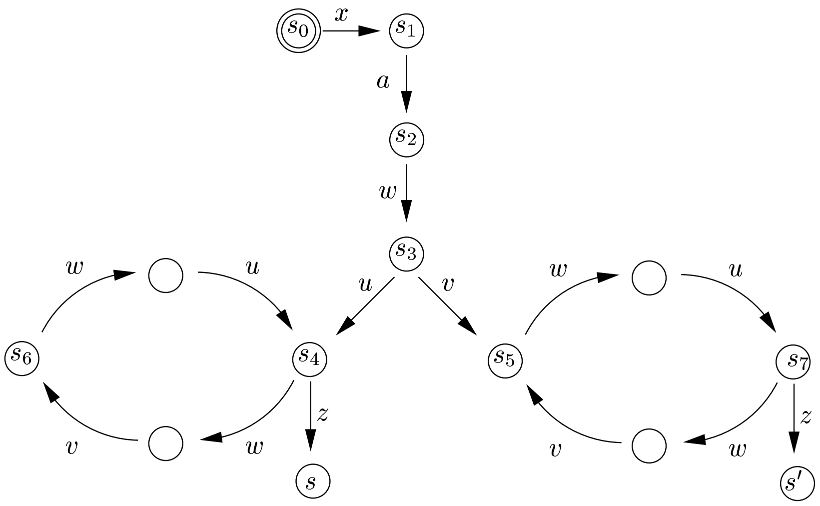

Figure 6: Pattern $\mathbf{S}_{\boldsymbol{k}}^{\boldsymbol{rev}}$ with $|w| = k$, $a \in A$ and $s \in S^{'} \Leftrightarrow s' \notin S^{'}$.

Moreover the union is finite since there is only a finite number of SCC's in the transition graph of F. Next we define $L^{<k}(p, C, q) =_{\text{def}} L(p, C, q) \cap A^{<k}$ and $L^{\geq k}(p, C, q) =_{\text{def}} L(p, C, q) \cap A^{\geq k}$. Obviously each set $L^{<k}(p, C, q)$ is finite and it holds that

$$L(p, C, q) = L^{<k}(p, C, q) \cup L^{\geq k}(p, C, q). \tag{3}$$

Now let $u \in A^*$ and define $L^{\geq k}(p, u, C, q) =_{\text{def}} L^{\geq k}(p, C, q) \cap uA^*$ as the set of all words in $L^{\geq k}(p, C, q)$ with prefix u. Then it holds that

$$L^{\geq k}(p, C, q) = \bigcup_{u \in A^k} L^{\geq k}(p, u, C, q). \tag{4}$$

Finally let $v \in A^*$ and define $L^{\geq k}(p, u, C, v, q) =_{\text{def}} L^{\geq k}(p, u, C, q) \cap A^*v$ as the set of all words in $L^{\geq k}(p, u, C, q)$ with suffix v. Again we have

$$L^{\geq k}(p, u, C, q) = \bigcup_{v \in A^k} L^{\geq k}(p, u, C, v, q). \tag{5}$$

Lemma 5.1. *Let $F = (A, S, \delta, s_0, S')$ be a dfa and let $k \geq 0$. Then $L(F)$ is a finite union of languages $w_0 L_1 a_1 w_1 L_2 a_2 w_2 \cdots L_n a_n w_n L_{n+1}$ with $n \geq 0$, $w_i \in A^*$ for $0 \leq i \leq n$ and there exist distinct SCC's $C_1, \ldots, C_{n+1}$ in the transition graph of F, states $p_i, q_i \in C_i$ for $1 \leq i \leq n+1$, words $u_i, v_i \in A^k$ for $1 \leq i \leq n$ and words $u_{n+1} \in A^k$ and $v_{n+1} \in A^{k+1}$ such that*

1. *$\delta(s_0, w_0) = p_1$ and $q_{n+1} \in S'$,*
2. *$\delta(q_i, a_i w_i) = p_{i+1}$ and $\delta(q_i, a_i) \notin C_i$ for $1 \leq i \leq n$,*
3. *$L_i = L^{\geq k}(p_i, u_i, C_i, v_i, q_i)$ for $1 \leq i \leq n$ and*
4. *$L_{n+1} = L^{<k+1}(p_{n+1}, C_{n+1}, q_{n+1}) \cup L^{\geq k+1}(p_{n+1}, u_{n+1}, C_{n+1}, v_{n+1}, q_{n+1})$.*

Proof. We start with equation (1), substitute each $L(s)$ using equation (2). For all $1 \leq i \leq n$ we rewrite each $L(p_i, C_i, q_i)$ using equation (3), then each $L^{\geq k}(p_i, C_i, q_i)$ using equation (4), each $L^{\geq k}(p_i, u, C_i, q_i)$ using equation (5) and finally each finite set $L^{<k}(p_i, C_i, q_i)$ as the finite union of its elements. Now we apply the identity $L(L' \cup L'') = LL' \cup LL''$ to obtain a union of languages of the desired form, while the set L_{n+1} in each language has still the form $L(p_{n+1}, C_{n+1}, q_{n+1})$. In this set we want to emphasize the prefix of length k and the suffix of length $k+1$. So again we start with equations (3) to (5) and end up with sets

$$L_{n+1} = L^{<k}(p_{n+1}, C_{n+1}, q_{n+1}) \cup L^{\geq k}(p_{n+1}, u_{n+1}, C_{n+1}, v_{n+1}, q_{n+1})$$

with $|u_{n+1}|, |v_{n+1}| = k$. Next we ensure that there are in $L^{\geq k}(p_{n+1}, u_{n+1}, C_{n+1}, v_{n+1}, q_{n+1})$ only words of length at least $k+1$ by taking u_{n+1} into the finite set if needed. We can then state

$$L_{n+1} = L^{<k+1}(p_{n+1}, C_{n+1}, q_{n+1}) \cup L^{\geq k+1}(p_{n+1}, u_{n+1}, C_{n+1}, v_{n+1}, q_{n+1}).$$

Finally we see that

$$L^{\geq k+1}(p_{n+1}, u_{n+1}, C_{n+1}, v_{n+1}, q_{n+1}) = \bigcup_{a \in A} L^{\geq k+1}(p_{n+1}, u_{n+1}, C_{n+1}, av_{n+1}, q_{n+1})$$

and obtain in a similar way as above unions of languages, where L_{n+1} has the form as stated in 4. Note that $|av_{n+1}| = k+1$ and that we maintain with all operations the finiteness of the union. ❑

We want to mention the strong existential character of the previous lemma. If a language L_i for some $1 \leq i \leq n+1$ is a set in a concatenation as stated in the lemma then in C_i there exists indeed a state p_i' such that $\delta(p_i, u_i) = p_i'$ and there is at least one state $q_i' \in C_i$ such that $\delta(q_i', v_i) = q_i$.

Now we proof for any $k \geq 0$ that a language $L \subseteq A^+$ belongs to $\mathcal{D}_k^{\mathrm{left}}$ if the minimal automaton accepting L does not have pattern $\mathbf{S}_{\boldsymbol{k}}$. We will drop the condition of the minimality of F later on.

Lemma 5.2. *Let $k \geq 0$, let F be a minimal dfa with $L(F) \subseteq A^+$. If F does not have pattern $\mathbf{S}_{\boldsymbol{k}}$ then $L(F) \in \mathcal{D}_k^{\mathrm{left}}$.*

Proof. Let $F = (A, S, \delta, s_0, S')$ be a minimal dfa accepting some language $L(F) \subseteq A^+$. We show that any language $L =_{\mathrm{def}} w_0 L_1 a_1 w_1 L_2 a_2 w_2 \cdots L_n a_n w_n L_{n+1}$ from Lemma 5.1 is a finite union of k-leftdeterministic languages if F does not have pattern $\mathbf{S}_{\boldsymbol{k}}$. We keep the notations of Lemma 5.1 and additionally fix for each $1 \leq i \leq n+1$ states $p_i', q_i' \in C_i$ such that $\delta(p_i, u_i) = p_i'$ and $\delta(q_i', v_i) = q_i$.

The absence of pattern $\mathbf{S}_{\boldsymbol{k}}$ in an automaton intuitively sais that within an SCC the automaton can not distinguish between different occurences of the same factor of length $k+1$ in the input. This enables us to make significant relaxations which will be our first step.

Step 1. In an SCC C_i with $1 \leq i \leq n$ there is no need to fix the state q_i since any word with a path in C_i labeled v_i and followed by the letter a_i must end in p_{i+1} or we find $\mathbf{S}_{\boldsymbol{k}}$. Therefore we define for $1 \leq i \leq n$

$$V_i =_{\mathrm{def}} \left\{ x \in A^{\geq k} \;\middle|\; \delta(p_i, x) \in C_i \right\} \cap (u_i A^* \cap A^* v_i).$$

In case of the SCC C_{n+1} we do the same thing just for the prefix of length k and the suffix of length $k+1$ and set

$$V_{n+1} =_{\mathrm{def}} \left\{ x \in A^{\geq k+1} \;\middle|\; \delta(p_{n+1}, x) \in C_{n+1} \right\} \cap (u_{n+1} A^* \cap A^* v_{n+1}).$$

We write $B =_{\mathrm{def}} L^{<k+1}(p_{n+1}, C_{n+1}, q_{n+1})$ as the set of words of short length in L_{n+1} and claim that

$$L = w_0 V_1 a_1 w_1 V_2 a_2 w_2 \cdots V_n a_n w_n (B \cup V_{n+1}).$$

Obviously we have $L^{\geq k}(p_i, u_i, C_i, v_i, q_i) \subseteq V_i$ for $1 \leq i \leq n$ and it holds that $L^{\geq k+1}(p_{n+1}, u_{n+1}, C_{n+1}, v_{n+1}, q_{n+1}) \subseteq V_{n+1}$.

We show the inclusion from right to left by contradiction. Assume that for some $1 \leq i \leq n$ there is some $wv_i \in V_i$ with $\delta(p_i, wv_i a_i) \neq \delta(q_i', v_i a_i)$. Then also $\delta(p_i, w) \neq q_i'$ and since $\delta(p_i, wv_i) \in C_i$ we found pattern $\mathbf{S}_{\boldsymbol{k}}$ witnessed by the states $\delta(p_i, w), \delta(p_i, wv_i), q_i', q_i \in C_i$ and the letter a_i. Note that there is some $z \in A^*$ such that $\delta(p_i, wv_i a_i z) \in S'$ and $\delta(q_i', v_i a_i z) \notin S'$ (or vice versa) because F is a minimal dfa. Now assume there is some $wv_{n+1} \in V_{n+1}$ with $\delta(p_{n+1}, wv_{n+1}) \neq q_{n+1}$) and let $v_{n+1} = va$ for some $a \in A$. Then $\delta(p_{n+1}, w) \neq q_{n+1}'$ and we found pattern $\mathbf{S}_{\boldsymbol{k}}$ witnessed by the states $\delta(p_{n+1}, w), \delta(p_{n+1}, wv), q_{n+1}', \delta(q_{n+1}', v) \in C_{n+1}$ and the letter a.

Step 2. We are interested in a description of the sets V_i over A^{k+1}. Let C be an SCC of the transition graph of F, then we set $S(C) =_{\mathrm{def}} \left\{ w \in A^{k+1} \;\middle|\; \text{there ex. some } q \in C \text{ such that } \delta(q, w) \in C \right\}$ as the set of words of length $k+1$ that occur as a label of a path within a particular SCC.

Moreover we introduce a shorthand notation to describe languages over A^{k+1}. Given some $w \in A^+$ with $|w| \geq k+1$ and k-decomposition $\widehat{w} =_{\mathrm{def}} (\alpha_1, \alpha_2, \ldots, \alpha_l)$ for suitable $l \geq 1$ and $\alpha_i \in A^{k+1}$ we simply write $(\ldots, \widehat{w}, \ldots)_k$ instead of consecutive $\alpha_1, \alpha_2, \ldots, \alpha_l$ in a language description. Furthermore we let for short

$$\widehat{y_i} =_{\mathrm{def}} \widehat{v_i a_i w_i u_{i+1}}$$

for $1 \leq i \leq n$ be the k-decomposition of $v_i a_i w_i u_{i+1}$. Assume for the moment that $w_0 \neq \epsilon$ and that $n > 0$. We claim that

$$L \;=\; \bigcup_{x \in B} \left(\widehat{w_0 u_1}, S(C_1), \widehat{y_1}, S(C_2), \widehat{y_2}, \ldots, S(C_n), v_n \widehat{a_n w_n} x\right)_k \;\cup \tag{1}$$
$$\left(\widehat{w_0 u_1}, S(C_1), \widehat{y_1}, S(C_2), \widehat{y_2}, \ldots, S(C_n), \widehat{y_n}, S(C_{n+1}), v_{n+1}\right)_k .$$

To see the inclusion from left to right let $1 \leq i \leq n$ and take some $w \in V_i$ with $|w| = k$. Then $w = u_i = v_i$ and we find w on the right hand side in $(\ldots \widehat{\ldots u_i}, \widehat{v_i \ldots} \ldots)_k$, i.e. no element from $S(C_i)$ occurs. Suppose $w \in V_i$ with $|w| \geq k+1$. Since by definition $\delta(p_i, w) \in C_i$ it holds that $\alpha(\widehat{w}) \subseteq S(C_i)$. Moreover w has k-prefix u_i and k-suffix v_i according to the definition of the words before and after each $S(C_i)$. Now take some $w \in V_{n+1}$. Clearly the k-decomposition of w contains only elements from $S(C_{n+1})$ and ends with v_{n+1}.

For the inclusion from right to left it suffices to show that for all $1 \leq i \leq n+1$ and for any w with prefix u_i, with suffix v_i and with $\alpha(\widehat{w}) \subseteq S(C_i)$ it holds that $\delta(p_i, w) \in C_i$. If w has length k then $\delta(p_i, w) = p'_i \in C_i$. Now assume that $|w| \geq k+1$ and $\delta(p_i, w) \notin C_i$. Then there is a state $p \in C_i$ and a word $va \in \alpha(\widehat{w})$ with $v \in A^k$ and $a \in A$ such that $\delta(p, v) \in C_i$ and $\delta(p, va) \notin C_i$. Moreover, since $va \in S(C_i)$ there is a state $p' \in C_i$ such that $\delta(p', v) \in C_i$ and also $\delta(p', va) \in C_i$. Hence $p \neq p'$ and we find pattern $\mathbf{S}_{\boldsymbol{k}}$ witnessed by the states $p, \delta(p, v), p', \delta(p', v)$ and the letter a.

It remains to show the cases when $w_0 = \epsilon$ or $n = 0$. With little modifications that deal with the length of words we can state equations similiar to equation (1). First we assume that $w_0 = \epsilon$ and $n > 0$. If $u_1 \neq v_1$ then V_1 contains only words of length $\geq k+1$ and we can state

$$L \;=\; \bigcup_{u_1 a \in S(C_1)} \Big[\bigcup_{x \in B} \left(u_1 a, S(C_1), \widehat{y_1}, S(C_2), \widehat{y_2}, \ldots, S(C_n), v_n \widehat{a_n w_n} x\right)_k \;\cup \tag{2}$$
$$\left(u_1 a, S(C_1), \widehat{y_1}, S(C_2), \widehat{y_2}, \ldots, S(C_n), \widehat{y_n}, S(C_{n+1}), v_{n+1}\right)_k \Big].$$

If $u_1 = v_1$ we add to the union above each set again, this time starting with $(\widehat{y_1}, S(C_2), \widehat{y_2}, \ldots)_k$. If $n = 0$ and $w_0 \neq \epsilon$ we have $L = w_0 B \cup w_0 V_1$ and see that

$$L \;=\; \left(w_0 B \cap A^{\leq k}\right) \;\cup \bigcup_{\substack{w \in w_0 B \\ |w| > k}} (\widehat{w})_k \;\cup \left(\widehat{w_0 u_1}, S(C_1), v_1\right)_k . \tag{3}$$

Finally suppose $n = 0$ and $w_0 = \epsilon$. Then $L = B \cup V_1$ and we have

$$L \;=\; B \;\cup \bigcup_{u_1 a \in S(C_1)} \left(u_1 a, S(C_1), v_1\right)_k . \tag{4}$$

Step 3. In any of the sets on the right hand side in (1) and (2) and for any $1 \leq i \leq n$ it holds that $v_i a_i \notin S(C_i)$. Suppose this is not the case, then we find immediately pattern $\mathbf{S}_{\boldsymbol{k}}$ in the component C_i with the same argument as in step 2 above. To see that L in all equations (1) to (4) is k-leftdeterministic observe that we can insert the empty set at any position in the description over A^{k+1}. So we can write e.g. $(\ldots, \alpha, \emptyset, \beta, \ldots)_k$ instead of $(\ldots, \alpha, \beta, \ldots)_k$ and have fulfilled $\beta \notin \emptyset$. Furthermore note that the last index is not subject to the requirement $\alpha_i \notin \Sigma_i$. In our construction we have indeed $v_{n+1} \in S(C_{n+1})$. ❑

Lemma 5.3. *Let $k \geq 0$ and F be dfa with $L(F) \subseteq A^+$. If $L(F) \in \mathcal{D}_k^{\mathrm{left}}$ then F does not have pattern $\mathbf{S}_{\boldsymbol{k}}$.*

Proof. Let $F = (A, S, \delta, s_0, S')$, suppose that $L(F)$ is a finite union of k-leftdeterministic languages and that F has pattern $\mathbf{S}_{\boldsymbol{k}}$. W.l.o.g. we may assume that $s \in S'$ and $s' \notin S'$. Then for all $l \geq 0$ it holds that $x(uwvw)^l az \in L(F)$ and $x(uwvw)^l uwaz \notin L(F)$. Since the union is finite there is a k-leftdeterministic language L such that $x(uwvw)^l az \in L$ for infinitely many l. By definition of k-leftdeterministic languages there exist $\alpha_1, \ldots, \alpha_{n+1} \in A^{k+1}$ and $\Sigma_1, \ldots, \Sigma_{n+1} \subseteq A^{k+1}$ for some $n \geq 0$ and a set $D \subseteq A^{\leq k} \backslash \{\epsilon\}$ such that for $1 \leq i \leq n$ it holds that $\alpha_i \notin \Sigma_i$ and $L = (\Sigma_1, \alpha_1, \ldots, \Sigma_n, \alpha_n, \Sigma_{n+1}, \alpha_{n+1})_k \cup D$. Then it must be that there is some i with $0 \leq i \leq n$ such that for infinitely many l it holds that

$$x(uwvw)^l \in (\Sigma_1, \alpha_1, \ldots, \Sigma_i, \alpha_i, \Sigma_{i+1})_k \quad \text{and} \quad waz \in (\Sigma_{i+1}, \alpha_{i+1}, \ldots, \Sigma_{n+1}, \alpha_{n+1})_k \tag{1}$$

or there is some i with $1 \leq i \leq n$ such that for infinitely many l it holds that

$$x(uwvw)^l \in (\Sigma_1, \alpha_1, \ldots, \Sigma_i, \alpha_i)_k \quad \text{and} \quad waz \in (\Sigma_{i+1}, \alpha_{i+1}, \ldots, \Sigma_{n+1}, \alpha_{n+1})_k. \tag{2}$$

Note that in (2) it can not be that $i = n + 1$ since $az \neq \epsilon$. Now Lemma 4.6 tells us that case (2) is in fact not possible. So let $1 < l_1 < l_2$ such that $x(uwvw)^{l_1} \in (\Sigma_1, \alpha_1, \ldots, \alpha_i, \Sigma_{i+1})_k$ and $x(uwvw)^{l_2} \in (\Sigma_1, \alpha_1, \ldots, \alpha_i, \Sigma_{i+1})_k$ as stated in (1). Since $x(uwvw)^{l_1}$ is a strict prefix of $x(uwvw)^{l_2}$ we have by Lemma 4.5 that $\alpha(\widehat{wuwvw}) \in \Sigma_{i+1}$. It follows that $x(uwvw)^{l_1} uwaz \in (\Sigma_1, \alpha_1, \ldots, \Sigma_{i+1}, \ldots, \Sigma_{n+1}, \alpha_{n+1})_k$ which is a contradiction. ❑

With the next two propositions we show that we can drop the condition of minimality of the dfa in question and see how patterns $\mathbf{S}_{\boldsymbol{k}}$ and $\mathbf{S}_{\boldsymbol{k}}^{\boldsymbol{rev}}$ are related.

Proposition 5.4. *Let $k \geq 0$ and $L \subseteq A^+$. Let F be some dfa with $L(F) = L$ and let $\hat{F}$ be the minimal dfa with $L(\hat{F}) = L$. Then F has pattern $\mathbf{S}_{\boldsymbol{k}}$ if and only if $\hat{F}$ has pattern $\mathbf{S}_{\boldsymbol{k}}$.*

Proof. Let $F = (A, S, \delta, s_0, S')$ be a dfa with $L(F) = L$ and let $\hat{F}$ be the minimal dfa for L. Assume $\hat{F}$ has pattern $\mathbf{S}_{\boldsymbol{k}}$ for some $k \geq 0$ witnessed by $x, u, v, z \in A^*$, $a \in A$ and $w \in A^k$. Now for all $i \geq 0$ it must that $\delta(s_0, x(uwvw)^i) \neq \delta(s_0, x(uwvw)^i uw)$ in F. Due to the finiteness of S we find a loop in $\delta(s_0, x), \delta(s_0, xuw), \delta(s_0, xuwvw), \delta(s_0, xuwvwuw), \ldots$ which shows that F also has $\mathbf{S}_{\boldsymbol{k}}$.

To see the only-if-part note that the witnessing states s_1 and s_3 for pattern $\mathbf{S}_{\boldsymbol{k}}$ in F are not equivalent and do not fuse in $\hat{F}$. ❑

Proposition 5.5. *Let $k \geq 0$ and $L \subseteq A^+$. Let F be a dfa with $L(F) = L$. Then for any dfa $\hat{F}$ with $L(\hat{F}) = L^R$ it holds that F has pattern $\mathbf{S}_{\boldsymbol{k}}$ if and only if $\hat{F}$ has pattern $\mathbf{S}_{\boldsymbol{k}}^{\boldsymbol{rev}}$.*

Proof. Let $F = (A, S, \delta, s_0, S')$ be a dfa with $L(F) = L$ and let $\hat{F} = (A, \hat{S}, \hat{\delta}, \hat{s}_0, \hat{S}')$ be some dfa accepting L^R.

Suppose F has pattern $\mathbf{S}_{\boldsymbol{k}}$ for some $k \geq 0$ witnessed by $x, u, v, z \in A^*$, $a \in A$ and $w \in A^k$. We assume w.l.o.g. that $s \in S'$ and $s' \notin S'$. Then for all $l \geq 0$ we have $x(uwvw)^l az \in L$ and $x(uwvw)^l uwaz \notin L$. So for all $l \geq 0$ we know that $z^R a(w^R v^R w^R u^R)^l x^R \in L^R$ and $z^R aw^R u^R (w^R v^R w^R u^R)^l x^R = z^R a(w^R u^R w^R v^R)^l w^R u^R x^R \notin L^R$.

Due to the finiteness of $\hat{S}$ there are $r_1, l_1, r_2, l_2 \geq 1$ such that $\hat{\delta}(\hat{s}_0, z^R a(w^R v^R w^R u^R)^{r_1}) = \hat{\delta}(\hat{s}_0, z^R a(w^R v^R w^R u^R)^{r_1+l_1})$ and $\hat{\delta}(\hat{s}_0, z^R a(w^R u^R w^R v^R)^{r_2}) = \hat{\delta}(\hat{s}_0, z^R a(w^R u^R w^R v^R)^{r_2+l_2})$. W.l.o.g. we may assume that $l_1 \geq r_1$ and $l_2 \geq r_2$. Let t be the smallest common multiple of l_1 and l_2. Then $\hat{\delta}(\hat{s}_0, z^R a(w^R v^R w^R u^R)^t) = \hat{\delta}(\hat{s}_0, z^R a(w^R v^R w^R u^R)^{rt})$ and $\hat{\delta}(\hat{s}_0, z^R a(w^R u^R w^R v^R)^t) = \hat{\delta}(\hat{s}_0, z^R a(w^R u^R w^R v^R)^{rt})$ for all $r \geq 1$. We define

$$\begin{aligned}
&\hat{x} =_{\text{def}} z^R,\\
&\hat{u} =_{\text{def}} (u^R w^R v^R w^R)^{2t-1} u^R,\\
&\hat{v} =_{\text{def}} (v^R w^R u^R w^R)^t v^R,\\
&\hat{w} =_{\text{def}} w^R \text{ and}\\
&\hat{z} =_{\text{def}} x^R.
\end{aligned}$$

Then $\hat{x}, \hat{u}, \hat{v}, \hat{z} \in A^*$, $a \in A$ and $\hat{w} \in A^k$ witness that $\hat{F}$ has pattern $\mathbf{S}_{\boldsymbol{k}}^{\boldsymbol{rev}}$.

On the other hand suppose that $\hat{F}$ has pattern $\mathbf{S}_{\boldsymbol{k}}^{\boldsymbol{rev}}$ witnessed by $x, u, v, z \in A^*$, $a \in A$ and $w \in A^k$. We assume w.l.o.g. that $s \in S'$ and $s' \notin S'$. Then for all $l \geq 0$ we have $xawu(wvwu)^l z \in L^R$ and $xa(wvwu)^{l+1} z \notin L^R$. So for all $l \geq 1$ we know that $z^R(u^R w^R v^R w^R)^l u^R w^R a x^R \in L$ and $z^R(u^R w^R v^R w^R)^l a x^R \notin L$.

Due to the finiteness of S there are $r, t \geq 1$ such that $\delta(s_0, z^R(u^R w^R v^R w^R)^r) = \delta(s_0, z^R(u^R w^R v^R w^R)^{r+t})$. We define

$$\begin{aligned}
&\hat{x} =_{\text{def}} z^R(u^R w^R v^R w^R)^r,\\
&\hat{u} =_{\text{def}} u^R,\\
&\hat{v} =_{\text{def}} (v^R w^R u^R w^R)^{t-1} v^R,\\
&\hat{w} =_{\text{def}} w^R \text{ and}\\
&\hat{z} =_{\text{def}} x^R.
\end{aligned}$$

Then $\hat{x}, \hat{u}, \hat{v}, \hat{z} \in A^*$, $a \in A$ and $\hat{w} \in A^k$ witness that F has pattern $\mathbf{S}_{\boldsymbol{k}}$. ❑

Theorem 5.6. *Let $k \geq 0$ and $L \subseteq A^+$. Let F be a dfa with $L(F) = L$.*

1. *$L \in \mathcal{D}_k^{\text{left}}$ if and only if F does not have pattern $\mathbf{S}_{\boldsymbol{k}}$.*
2. *$L \in \mathcal{D}_k^{\text{right}}$ if and only if F does not have pattern $\mathbf{S}_{\boldsymbol{k}}^{\boldsymbol{rev}}$.*

Proof. Statement 1 was proved in Lemma 5.2 in combination with Proposition 5.4 and Lemma 5.3. Let $L \in \mathcal{D}_k^{\text{right}}$, then by Proposition 4.4 this is equivalent to $L^R \in \mathcal{D}_k^{\text{left}}$ which in turn holds by statement 1 if and only if no dfa $\hat{F}$ accepting L^R has pattern $\mathbf{S}_{\boldsymbol{k}}$. Now by Proposition 5.5 the latter holds if and only if F does not have pattern $\mathbf{S}_{\boldsymbol{k}}^{\boldsymbol{rev}}$. ❑

Suppose a dfa F has pattern $\mathbf{S}_{\boldsymbol{k}}$ for some $k \geq 0$. Then it obviously has also the patterns $\mathbf{S}_{\boldsymbol{l}}$ for any $0 \leq l \leq k$. So for any $k \geq 0$ it holds that $\mathcal{D}_k^{\text{left}} \subseteq \mathcal{D}_{k+1}^{\text{left}}$ and also $\mathcal{D}_k^{\text{right}} \subseteq \mathcal{D}_{k+1}^{\text{right}}$. The forbidden pattern characterization from Theorem 5.6 enables us to give concise proofs of the facts that each class $\mathcal{D}_k^{\text{left}}$ and $\mathcal{D}_k^{\text{right}}$ is a Boolean algebra and that the hierarchies of classes $\left\{ \mathcal{D}_k^{\text{left}} \mid k \geq 0 \right\}$ and $\left\{ \mathcal{D}_k^{\text{right}} \mid k \geq 0 \right\}$ are indeed strict.

Proposition 5.7. *For any $k \geq 0$ it holds that $\mathcal{D}_k^{\text{left}}$ is closed under union, intersection and complementation. The same holds for $\mathcal{D}_k^{\text{right}}$.*

Proof. It suffices to show the closure under union and complementation. Each $\mathcal{D}_k^{\text{left}}$ and each $\mathcal{D}_k^{\text{right}}$ is closed under union by definition. Suppose $\overline{L} \notin \mathcal{D}_k^{\text{left}}$ for some $L \in \mathcal{D}_k^{\text{left}}$. In particular we may assume by Theorem 5.6 that the minimal dfa $F = (A, S, \delta, s_0, S')$ accepting $\overline{L}$ has pattern $\mathbf{S}_{\boldsymbol{k}}$ with $s \in S'$ and $s' \notin S'$. But then the minimal dfa accepting L has $\mathbf{S}_{\boldsymbol{k}}$ with $s \notin S'$ and $s' \in S'$, thus $L \notin \mathcal{D}_k^{\text{left}}$, a contradiction. The same argumentation is valid for $\mathcal{D}_k^{\text{right}}$ and $\mathbf{S}_{\boldsymbol{k}}^{\boldsymbol{rev}}$. ❑

Proposition 5.8. *For any* $k \geq 0$ *it holds that* $\mathcal{D}_k^{\text{left}}$ *is strictly included in* $\mathcal{D}_{k+1}^{\text{left}}$. *The same holds for* $\mathcal{D}_k^{\text{right}}$.

Proof. We define the witnessing languages $L_{k+1} =_{\text{def}} A^*b^{k+1}\backslash A^*b^{k+2}A^*$ for all $k \geq 0$. The minimal automaton accepting L_{k+1} has pattern $\mathbf{S}_{\boldsymbol{k}}$, but not pattern $\mathbf{S}_{\boldsymbol{k+1}}$. Thus $L_{k+1} \in \mathcal{D}_{k+1}^{\text{left}}\backslash\mathcal{D}_k^{\text{left}}$. Take L_{k+1}^R as a witness for $\mathcal{D}_k^{\text{right}} \subsetneq \mathcal{D}_{k+1}^{\text{right}}$. ❑

6 The Next Hierarchy

We come back to restricted temporal logic and show that k-deterministic languages are exactly the languages definable by formulas of restricted temporal logic having next depth (previously depth) at most k.

Lemma 6.1. *For all* $k \geq 0$ *it holds that* $L(\langle X_k \rangle) \subseteq \mathcal{D}_k^{\text{right}}$.

Proof. Let $\varphi \in \langle X_k \rangle$. We show $L(\varphi) \in \mathcal{D}_k^{\text{right}}$ by induction on the structure of φ. For the induction base let $\varphi \in X_k$, so $\varphi = \underbrace{\mathbf{X}\cdots\mathbf{X}}_{l\text{-times}}a$ for some $a \in A$ and $0 \leq l \leq k$. Then $L(\varphi) = A^l a A^*$. For $i, j \geq 0$ we define $B_{i,j} =_{\text{def}} A^i a A^j$. One verifies that

$$L(\varphi) \;=\; \bigcup_{0 \leq j < k-l} B_{l,j} \;\cup\; \bigcup_{w \in B_{l,k-l}} (w, A^{k+1})_k$$

where we have a finite union of k-rightdeterministic languages on the right hand side.

Since $\mathcal{D}_k^{\text{right}}$ is closed under Boolean operations (Proposition 5.7) it suffices to show the induction step for $\varphi = \mathbf{F}\psi$ with $L(\psi) \in \mathcal{D}_k^{\text{right}}$. It holds that $L(\mathbf{F}\psi) = AA^*L(\psi) = \bigcup_{a \in A} aA^*L(\psi)$. Suppose $L(\psi)$ is a finite union of k-rightdeterministic languages. We are done if we prove for one such language L that $aA^*L \in \mathcal{D}_k^{\text{right}}$. So let $L = (\alpha_1, \Sigma_1, \alpha_2, \Sigma_2, \ldots, \alpha_{n+1}, \Sigma_{n+1})_k \cup D$ with $D \subseteq A^{\leq k}\backslash\{\epsilon\}$ and with $\alpha_1, \ldots, \alpha_{n+1} \in A^{k+1}$ and $\Sigma_1, \ldots, \Sigma_{n+1} \subseteq A^{k+1}$ for some $n \geq 0$ such that for $2 \leq i \leq n+1$ it holds that $\alpha_i \notin \Sigma_i$.

We first show that $aA^*D \in \mathcal{D}_k^{\text{right}}$. Define $D' =_{\text{def}} \{ xz \in A^{k+1} \mid z \in D \}$ as the set of all words of length $k+1$ with a suffix from D. Then one verifies that

$$aA^*D \;=\; (aA^*D \cap A^{\leq k}) \;\cup\; \bigcup_{w \in aA^k} \; \bigcup_{v \in D'} (w, A^{k+1}, v)_k$$

where again we have a finite union of k-rightdeterministic languages on the right hand side.

Now we have to deal with the concatenation of aA^* and $(\alpha_1, \Sigma_1, \ldots, \alpha_{n+1}, \Sigma_{n+1})_k$ and obviously this is equal to

$$L' \;=_{\text{def}}\; \bigcup_{w \in aA^k} (w, A^{k+1}, \alpha_1, \Sigma_1, \ldots, \alpha_{n+1}, \Sigma_{n+1})_k.$$

Note that L' is not necessarily k-rightdeterministic since possibly $\alpha_1 \in \Sigma_1$. We define $\Sigma_1' =_{\text{def}} \Sigma_1\backslash\{\alpha_1\}$ and claim that

$$L' \;=\; \bigcup_{w \in aA^k} (w, A^{k+1}, \alpha_1, \Sigma_1', \ldots, \alpha_{n+1}, \Sigma_{n+1})_k.$$

To see the inclusion from left to right note that we can choose in the k-decomposition of a word from L' the rightmost occurence of α_1 in the sequence of elements from Σ_1. Since $\alpha_i \notin \Sigma_i$ for $2 \leq i \leq n+1$ and $\alpha_1 \notin \Sigma_1'$ we have that L' is a finite union of k-rightdeterministic languages. ❑

Lemma 6.2. *For all $k \geq 0$ it holds that $\mathcal{D}_k^{\text{right}} \subseteq L(\langle X_k \rangle)$.*

Proof. Let $k \geq 0$. It suffices to proof for a k-rightdeterministic language L that there is an RTL-formulas $\varphi \in \langle X_k \rangle$ with $L(\varphi) = L$. Obviously we can define a finite set D with words of length at most k by a formula of next depth $\leq k$. So let $L = (\alpha_1, \Sigma_1, \alpha_2, \Sigma_2, \ldots, \alpha_{n+1}, \Sigma_{n+1})_k$ with $\alpha_1, \ldots, \alpha_{n+1} \in A^{k+1}$ and $\Sigma_1, \ldots, \Sigma_{n+1} \subseteq A^{k+1}$ for some $n \geq 0$ such that for $2 \leq i \leq n+1$ it holds that $\alpha_i \notin \Sigma_i$.

The proof will be by induction on n. We will make available in each step a formula φ' defining the set $pre(L(\varphi))$ of all prefixes of words in $L(\varphi)$, i.e. $L(\varphi') = pre(L(\varphi)) =_{\text{def}} \{ w \in A^* \mid \text{there exists some } v \in A^* \text{ with } vw \in L(\varphi) \}$. For convenient notations we write $\mathbf{G}\phi$ instead of $\neg\mathbf{F}\neg\phi$ for all $\phi \in \text{RTL}$ and we define for all $\alpha = a_1 \cdots a_{k+1} \in A^{k+1}$ formulas

$$\psi(\alpha) =_{\text{def}} \bigwedge_{0 \leq l \leq k} \underbrace{\mathbf{X} \cdots \mathbf{X}}_{l\text{-times}} a_{l+1} \quad \text{with} \quad L(\psi(\alpha)) = \alpha A^*$$

and for each $\Sigma \subseteq A^{k+1}$ formulas

$$\psi(\Sigma) =_{\text{def}} \bigvee_{\alpha \in \Sigma} \psi(\alpha) \quad \text{with} \quad L(\psi(\Sigma)) = \Sigma A^*.$$

Let $n = 0$. Then $L = (\alpha_1, \Sigma_1)_k$ and we define

$$\varphi =_{\text{def}} \psi(\alpha_1) \wedge \mathbf{G}\psi(\Sigma_1) \quad \text{and} \quad \varphi' =_{\text{def}} (\psi(\alpha_1) \vee \psi(\Sigma_1)) \wedge \mathbf{G}\psi(\Sigma_1).$$

Then $\varphi, \varphi' \in \langle X_k \rangle$ and it holds that $L = L(\varphi)$ and $pre(L(\varphi)) = L(\varphi')$.

Now let $n \geq 1$ and $L = (\alpha_1, \Sigma_1, \alpha_2, \Sigma_2, \ldots, \alpha_{n+1}, \Sigma_{n+1})_k$. By hypothesis there exist $\varphi_1, \varphi_1' \in \langle X_k \rangle$ such that $L(\varphi_1) = (\alpha_2, \Sigma_2, \ldots, \alpha_{n+1}, \Sigma_{n+1})_k$ and $L(\varphi_1') = pre(L(\varphi_1))$. We define

$$\begin{aligned} \varphi &=_{\text{def}} \big(\psi(\alpha_1) \wedge \mathbf{F}\varphi_1\big) \wedge \mathbf{G}(\psi(\Sigma_1) \vee \varphi_1') \quad \text{and} \\ \varphi' &=_{\text{def}} \Big[\big((\psi(\alpha_1) \vee \psi(\Sigma_1)) \wedge \mathbf{F}\varphi_1\big) \wedge \mathbf{G}(\psi(\Sigma_1) \vee \varphi_1')\Big] \vee \varphi_1'. \end{aligned}$$

Obviously, $\varphi, \varphi' \in \langle X_k \rangle$ since $\varphi_1, \varphi_1' \in \langle X_k \rangle$. To show $L = L(\varphi)$ we have to argue for the inclusion from right to left. Let $w \in L(\varphi)$ with $\widehat{w} = (\beta_1, \ldots, \beta_l)$ for some $l \geq 1$. Then $\beta_1 = \alpha_1$ and there is some strict suffix $z \in L(\varphi_1)$ of w with $\widehat{z} = (\beta_i, \ldots, \beta_l)$ for some i with $2 \leq i \leq l$. We have to show that $\{\beta_2, \ldots, \beta_{i-1}\} \in \Sigma_1$. Suppose that this is not the case. Then there is some j with $2 \leq j \leq i-1$ and some $y \in A^+$ such that $\widehat{yz} = (\beta_j, \ldots, \beta_i, \ldots, \beta_l)$ and $yz \in L(\varphi_1')$. Since $L(\varphi_1') = pre(L(\varphi_1))$ there is some $x \in A^*$ with $xyz \in L(\varphi_1)$. So we obtain that z and xyz are in $L(\varphi_1) = (\alpha_2, \Sigma_2, \ldots, \alpha_{n+1}, \Sigma_{n+1})_k$. Note that for $2 \leq i \leq n+1$ we have $\alpha_i \notin \Sigma_i$ and since $|z| < |xyz|$ we get a contradiction to Lemma 4.8. The same argument proves that $pre(L(\varphi)) = L(\varphi')$. ❑

Theorem 6.3. *For all $k \geq 0$ it holds that*

1. $L(\text{TL}[\mathbf{X}(k), \mathbf{F}]) = \mathcal{D}_k^{\text{right}}$ *and*
2. $L(\text{TL}[\mathbf{Y}(k), \mathbf{P}]) = \mathcal{D}_k^{\text{left}}$.

Proof. Statement 1 can be followed from Lemmata 6.1 and 6.2 together with Proposition 3.5. Statement 2 follows from 1 with Propositions 3.6 and 4.4. ❑

We subsume our characterizations of the next hierarchy and the previously hierarchy.

Theorem 6.4. *Let $k \geq 0$ and $L \subseteq A^+$. Let F be a dfa with $L(F) = L$. Then the following statements are equivalent.*

(1) $L \in L(\mathrm{TL}[\mathbf{Y}(k), \mathbf{P}])$

(2) $L \in \mathcal{D}_k^{\mathrm{left}}$

(3) F does not have pattern $\mathbf{S_k}$.

Theorem 6.5. *Let $k \geq 0$ and $L \subseteq A^+$. Let F be a dfa with $L(F) = L$. Then the following statements are equivalent.*

(1) $L \in L(\mathrm{TL}[\mathbf{X}(k), \mathbf{F}])$

(2) $L \in \mathcal{D}_k^{\mathrm{right}}$

(3) F does not have pattern $\mathbf{S_k^{rev}}$.

In combination with Proposition 5.8 we see that the next hierarchy and the previously hierarchy are strict. Moreover we immediately obtain $L(\mathrm{RPTL}) = \bigcup_{k \geq 0} \mathcal{D}_k^{\mathrm{left}}$ and $L(\mathrm{RTL}) = \bigcup_{k \geq 0} \mathcal{D}_k^{\mathrm{right}}$. With the following proposition we give a complete reproof of Theorem 3.2 and even drop the minimality condition for the dfa.

Proposition 6.6. *Let F be a dfa.*

1. *F does not have pattern $\mathbf{S}$ if and only if there exists some $k \geq 0$ such that F does not have pattern $\mathbf{S_k}$.*
2. *F does not have pattern $\mathbf{S^{rev}}$ if and only if there exists some $k \geq 0$ such that F does not have pattern $\mathbf{S_k^{rev}}$.*

Proof. Let $F = (A, S, \delta, s_0, S')$. In light of Proposition 5.5 it suffices to prove statement 1. To see the if-part assume that F does not have pattern $\mathbf{S_k}$ for some $k \geq 0$, but there is pattern $\mathbf{S}$ in F witnessed by $x, u, v, z \in A^*$ and $w \in A^+$. Then x, u, v, wz and $w^k \in A^+$ show that F has pattern $\mathbf{S_l}$ for $l =_{\mathrm{def}} |w^k|$. Since $l \geq k$ there is also pattern $\mathbf{S_k}$ in F, a contradiction.

For the other implication suppose F does not have pattern $\mathbf{S}$, but for all $k \geq 0$ there is pattern $\mathbf{S_k}$ in F witnessed by $x_k, u_k, v_k, z_k \in A^*$, $a_k \in A$ and $w_k \in A^k$. Especially there are $s_1, \ldots, s_6, s, s' \in S$ that are involved in pattern $\mathbf{S_l}$ for $l =_{\mathrm{def}} |S|^2 + 1$. Look at the set of pairs $(p_i, q_i) \in S \times S$ for $0 \leq i \leq l$ with $p_i =_{\mathrm{def}} \delta(s_1, y_i)$ and $q_i =_{\mathrm{def}} \delta(s_3, y_i)$ if y_i is the i-prefix of w_l. Since there are only $|S|^2$ different such pairs there must be $i < j$ such that $(p_i, q_i) = (p_j, q_j)$. Let $w' \in A^+$ such that $y_j = y_i w'$ and let $w'' \in A^*$ such that $w = y_i w' w''$. Then $x_l u_l y_i, w'' v_l y_i, w'' u_l y_i, w'' a_l z_l \in A^*$ and $w' \in A^+$ show that there is pattern $\mathbf{S}$ in F, a contradiction. ❑

Notice that we can bound k by $|S|^2$. We turn to decidability questions. Using Theorems 6.4 and 6.5 we obtain a co-NL(=NL)-algorithm for the decision problem for $L(\mathrm{TL}[\mathbf{Y}(k), \mathbf{P}])$ and for $L(\mathrm{TL}[\mathbf{X}(k), \mathbf{F}])$ with fixed k simply by testing the occurence of the patterns $\mathbf{S_k}$ and $\mathbf{S_k^{rev}}$ in a given dfa. To see this note that we can guess the relevant states of a pattern and follow nondeterministically possible words between them. Computing strongly connected components can also be done in NL. Moreover there is a polynomial-time algorithm to determine for a given minimal dfa the level k in both hierarchies: We simply apply the previous algorithm for $k = 0, \dots, |S|^2$. This also decides $L(\mathrm{TL}[\mathbf{Y}, \mathbf{P}])$ and $L(\mathrm{TL}[\mathbf{X}, \mathbf{F}])$.

The mentioned algorithms are closely related to the ones known to decide level one of the dot-depth hierarchy (cf. [Ste85b, CH91]). We discuss this relation more detailed in the following section.

7 Relations to Dot-Depth One

We point out relations of the classes we studied to level one of the dot-depth hierarchy and give a definition of this hierarchy. For a class $\mathcal{C}$ of languages of A^+, let $\mathrm{BC}(\mathcal{C})$ be the boolean closure of $\mathcal{C}$, i.e. $\mathrm{BC}(\mathcal{C})$ is the smallest class containing $\mathcal{C}$ and being closed under union, intersection and complementation, where the latter is taken with respect to A^+. For a class $\mathcal{C}$ of languages of A^+, let $\mathrm{POL}(\mathcal{C})$ be its polynomial closure defined as follows: A language $L \subseteq A^+$ is in $\mathrm{POL}(\mathcal{C})$ if and only if L can be written as a finite union of languages $u_0 L_1 u_1 L_2 u_2 \cdots u_{n-1} L_n u_n$, where $u_i \in A^*$, $L_i \in \mathcal{C}$ and $n \geq 0$ (if $n = 0$ then $u_0 \neq \epsilon$). Then the dot-depth hierarchy can be defined as the following family of classes of languages of A^+:

1. $\mathcal{B}_0 =_{\mathrm{def}} \{\emptyset, A^+\}$,
2. $\mathcal{B}_{n+1/2} =_{\mathrm{def}} \mathrm{POL}(\mathcal{B}_n)$ for $n \geq 0$, and
3. $\mathcal{B}_{n+1} =_{\mathrm{def}} \mathrm{BC}(\mathcal{B}_{n+1/2})$ for $n \geq 0$.

For a language $L \subseteq A^+$ and a minimal n with $L \in \mathcal{B}_{n/2}$ we say that L *has dot-depth* $n/2$. To our knowledge, only levels 0, 1/2 and 1 are known to be decidable. For an overview see [Pin95, Pin96a], recent results can be found in [PW97]. In the first part of this section we recall some known facts about languages having dot-depth one and start with the characterization of $\mathcal{B}_1$ in terms of a certain congruence $\sim_{m,k}$ with $k, m \geq 0$ introduced by I. Simon in [Sim72]. Roughly speaking, two words x and y are equivalent under this congruence if they have the same first k letters, the same last k letters and if the same m-tuples of words of length $k+1$ appear in x and y. More precisely, let $m, k \geq 0$ and $W_{m,k} =_{\mathrm{def}} (\alpha_1, \alpha_2, \dots, \alpha_m)$ with $\alpha_i \in A^{k+1}$ be an m-tuple of words of length $k+1$. We say that $W_{m,k}$ *appears* in a word w, if $\alpha_1, \alpha_2, \dots, \alpha_m$ in this ordering are elements of the k-decomposition of w. Let $\mu_{m,k}(w)$ be the set of all $W_{m,k}$ that appear in a given word w. By convention, $\mu_{0,k}(w) = \emptyset$ for all k.

Definition 7.1 [Sim72]. Let $m, k \geq 0$ and $x, y \in A^*$. We say that x and y are (m,k)*-equivalent*, in symbols $x \sim_{m,k} y$, if and only if

(a) $x = y$, if $|x| < m + k$, or
(b) $p_k(x) = p_k(y), s_k(x) = s_k(y)$ and $\mu_{m,k}(x) = \mu_{m,k}(y)$, otherwise.

It is easy to see that $\sim_{m,k}$ is a congruence of finite index on A^*. With $[w]_{m,k}$ we denote the equivalence class of a word w modulo $\sim_{m,k}$. It is known that L is in $\mathcal{B}_1$ if and only if L can be written as a union of equivalence classes of $\sim_{m,k}$ for some integers m, k.

Theorem 7.2 [Sim72]. *Let $L \subseteq A^+$. Then $L \in \mathcal{B}_1$ if and only if L can be written as a union of equivalence classes modulo $\sim_{m,k}$ for some $m, k \geq 0$.*

We take a closer look at the structure of $\mathcal{B}_1$ and fix the "level" k of a language within this class.

Definition 7.3. Let $k \geq 0$. Then $\mathcal{B}_{1,k}$ is the class of languages $L \subseteq A^+$ that can be written as a union of equivalence classes modulo $\sim_{m,k}$ for some $m \geq 0$.

It holds that $\mathcal{B}_{1,k} \subseteq \mathcal{B}_{1,k+1}$ and $\mathcal{B}_1 = \bigcup_{k\geq 0} \mathcal{B}_{1,k}$. We restate the forbidden pattern characterization of the classes $\mathcal{B}_{1,k}$ obtained by J. Stern in [Ste85a]. For a decision algorithm for $\mathcal{B}_1$ derived from this characterization see [Ste85b, CH91].

Let $F = (A, S, \delta, s_0, S')$ be a dfa and let $k \geq 0$. Then F has pattern $\mathbf{T_k}$ if there are states $s_1, \ldots, s_4, s, s' \in S$ and words $x, u, v, z \in A^*$ and $w \in A^k$ with $s \in S' \Leftrightarrow s' \notin S'$, $\delta(s_0, x) = s_1 = \delta(s_4, w)$, $\delta(s_1, u) = s_2$, $\delta(s_2, w) = s_3$, $\delta(s_3, v) = s_4$, $\delta(s_1, z) = s$ and $\delta(s_3, z) = s'$ (cf. Figure 7).

A dfa F has pattern $\mathbf{T_k^{rev}}$ if there are states $s_1, \ldots, s_6, s, s' \in S$ and words $x, u, v, z \in A^*$ and $w \in A^k$ with $s \in S' \Leftrightarrow s' \notin S'$, $\delta(s_0, x) = s_1$, $\delta(s_1, w) = s_2$, $\delta(s_2, u) = s_3 = \delta(s_5, wu)$, $\delta(s_2, v) = s_4 = \delta(s_6, wv)$, $\delta(s_3, wv) = s_5$, $\delta(s_4, wu) = s_6$, $\delta(s_3, z) = s$ and $\delta(s_6, z) = s'$ (cf. Figure 8).

A dfa F has pattern $\mathbf{R_k}$ if there are states $s_1, \ldots, s_8, s, s' \in S$ and words $x, y, y', u, v, z \in A^*$ and $w, w' \in A^k$ with $s \in S' \Leftrightarrow s' \notin S'$, $\delta(s_0, x) = s_1 = \delta(s_4, w)$, $\delta(s_1, y) = s_2$, $\delta(s_2, w') = s_3$, $\delta(s_3, y') = s_4$, $\delta(s_1, u) = s_5 = \delta(s_7, wu)$, $\delta(s_3, v) = s_6 = \delta(s_8, w'v)$, $\delta(s_5, w'v) = s_7$, $\delta(s_6, wu) = s_8$, $\delta(s_5, z) = s$ and $\delta(s_8, z) = s'$ (cf. Figure 9).

Theorem 7.4 [Ste85a]. *Let $L \subseteq A^+$, $k \geq 0$ and let F be the minimal dfa accepting L. Then the following statements are equivalent:*

(1) $L \in \mathcal{B}_{1,k}$

(2) F does not have any of the patterns $\mathbf{T_k}$, $\mathbf{T_k^{rev}}$ and $\mathbf{R_k}$.

Originally these patterns were defined in [Ste85a] with an additional condition refering to the distinctness of strongly connected components. It is easy to show that we come to the same class with and without this condition if we do not allow all three patterns at the same time. Moreover we see that $\mathbf{T_k}$ and $\mathbf{T_k^{rev}}$ are special cases of $\mathbf{R_k}$ and we also obtain $\mathcal{B}_{1,k}$ if we only forbid $\mathbf{R_k}$. To get $\mathbf{T_k}$ set $u = v = \epsilon$ and $w = w'$ in $\mathbf{R_k}$ and to get $\mathbf{T_k^{rev}}$ let $s_1 = s_3$, $y = y' = \epsilon$ and $w = w'$. Additionally we may drop the minimality condition for the dfa as before.

The relation between pattern $\mathbf{T_k}$ and $\mathbf{T_k^{rev}}$ is the same as between $\mathbf{S_k}$ and $\mathbf{S_k^{rev}}$ (see Proposition 5.5). Let F be a dfa. Then for any dfa $\hat{F}$ with $L(\hat{F}) = L(F)^R$ it holds that F has pattern $\mathbf{T_k}$ if and only if $\hat{F}$ has pattern $\mathbf{T_k^{rev}}$. Patterns $\mathbf{S_k}$ and $\mathbf{T_k}$ on one hand, and $\mathbf{S_k^{rev}}$ and $\mathbf{T_k^{rev}}$ are very similar. Obviously for all $k \geq 0$ it holds that if a dfa F has $\mathbf{S_k}$ then F has $\mathbf{T_k}$ and if F has $\mathbf{T_{k+1}}$ then it has $\mathbf{S_k}$. Thus the patterns $\mathbf{T_k}$ and $\mathbf{T_k^{rev}}$ define possibly smaller classes than the ones we obtain with $\mathbf{S_k}$ and $\mathbf{S_k^{rev}}$.

Definition 7.5. Let $k \geq 0$, let $L \subseteq A^+$ and let F be a dfa acepting L.
We say that $L \in \hat{\mathcal{D}}_k^{\text{left}}$ if and only if F does not have $\mathbf{T_k}$.
We say that $L \in \hat{\mathcal{D}}_k^{\text{right}}$ if and only if F does not have $\mathbf{T_k^{rev}}$.

But is there really a difference between $\hat{\mathcal{D}}_k^{\text{left}}$ and $\mathcal{D}_k^{\text{left}}$ for each k? If we look again at the proof of Proposition 5.8 we see that the minimal automaton accepting $L_{k+1} = A^*b^{k+1}\backslash A^*b^{k+2}A^*$ has $\mathbf{S_k}$, but it not even has $\mathbf{T_{k+1}}$. On the other hand we observe that the minimal automaton accepting $L'_k = A^*b^{k+1}$ has $\mathbf{T_k}$, but it does not have $\mathbf{S_k}$. Analog statements hold for the reverse patterns and languages. This shows that $\mathbf{T_k}$ and $\mathbf{T_k^{rev}}$ define strict intermediate classes and refine the next and previously hierarchy. We obtain $\hat{\mathcal{D}}_k^{\text{left}} \subsetneq \mathcal{D}_k^{\text{left}} \subsetneq \hat{\mathcal{D}}_{k+1}^{\text{left}}$ and $\hat{\mathcal{D}}_k^{\text{right}} \subsetneq \mathcal{D}_k^{\text{right}} \subsetneq \hat{\mathcal{D}}_{k+1}^{\text{right}}$ for all $k \geq 0$ (cf. Figure 10).

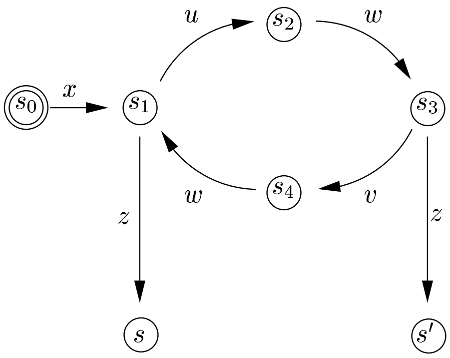

Figure 7: Pattern $\mathbf{T}_{\boldsymbol{k}}$ with $|w| = k$ and $s \in S^{'} \Leftrightarrow s' \notin S^{'}$.

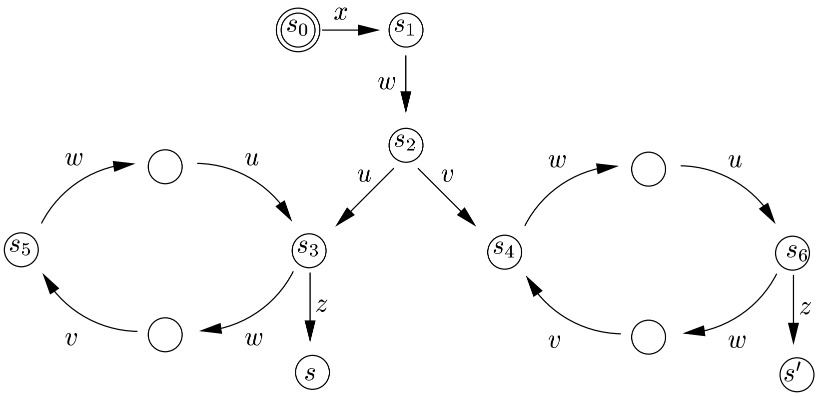

Figure 8: Pattern $\mathbf{T}_{\boldsymbol{k}}^{\boldsymbol{rev}}$ with $|w| = k$ and $s \in S^{'} \Leftrightarrow s' \notin S^{'}$.

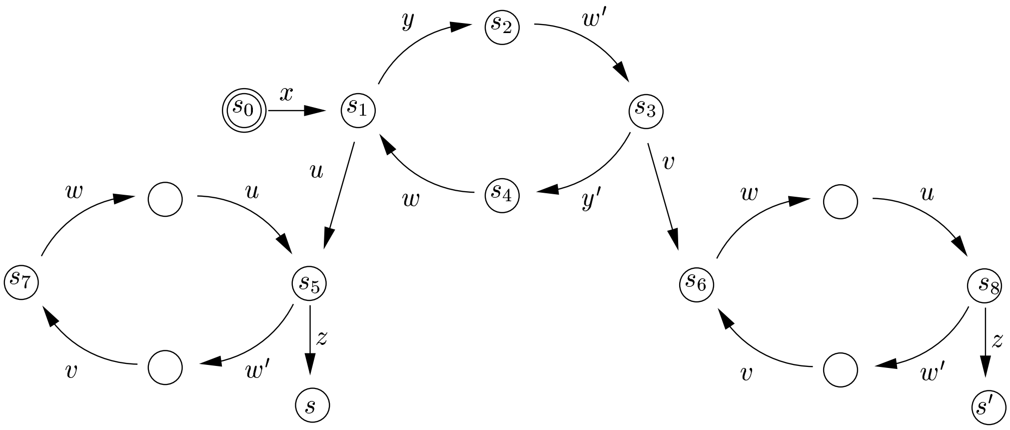

Figure 9: Pattern $\mathbf{R}_{\boldsymbol{k}}$ with $|w| = |w'| = k$ and $s \in S^{'} \Leftrightarrow s' \notin S^{'}$.

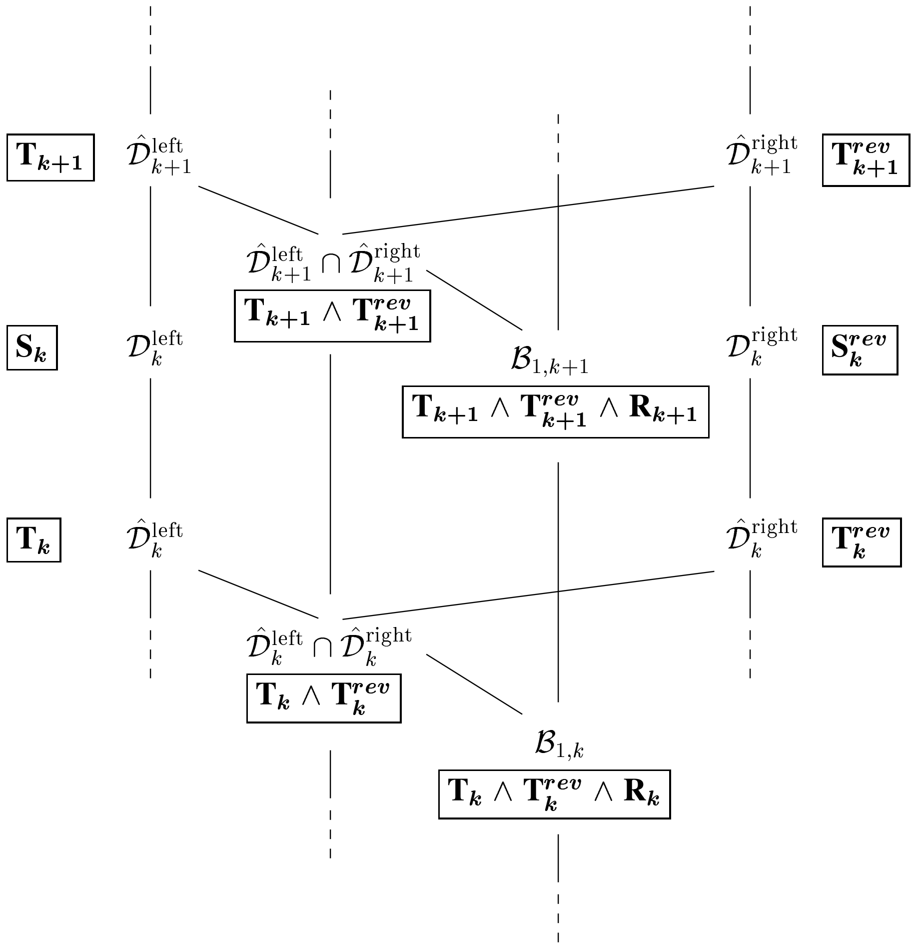

Figure 10: Levelwise inclusion structure.

Next we give characterizations of the classes $\hat{\mathcal{D}}^{\text{left}}_k$, $\hat{\mathcal{D}}^{\text{right}}_k$ and $\mathcal{B}_{1,k}$ over A^{k+1}. To capture $\hat{\mathcal{D}}^{\text{left}}_k$ and $\hat{\mathcal{D}}^{\text{right}}_k$ we only have to slightly modify the definitions of k-leftdeterministic and k-rightdeterministic languages. Instead of fixing the last (first) element in the k-decomposition we only need to fix the k-suffix (k-prefix) of some last (first) element. Let $w, v \in A^*$, $\alpha_1, \dots, \alpha_n \in A^{k+1}$ and $\Sigma_1, \dots, \Sigma_{n+1} \subseteq A^{k+1}$ for some $n \geq 0$. For a convenient notation we write $(w|\Sigma_1, \alpha_1, \Sigma_2, \dots, \alpha_n, \Sigma_{n+1}|v)_k$ instead of $\big((wA^* \cap A^*v) \cap (\Sigma_1, \alpha_1, \Sigma_2, \dots, \alpha_n, \Sigma_{n+1})_k\big)$.

Definition 7.6. Let $k \geq 0$ and $L \subseteq A^+$.

1. We say that L is *weak k-leftdeterministic* if and only if there exist $v \in A^k$, $\alpha_1, \dots, \alpha_{n+1} \in A^{k+1}$ and $\Sigma_1, \dots, \Sigma_{n+1} \subseteq A^{k+1}$ for some $n \geq 0$ and a set $D \subseteq A^{\leq k} \backslash \{\epsilon\}$ such that for $1 \leq i \leq n$ it holds that $\alpha_i \notin \Sigma_i$ and $L = (\Sigma_1, \alpha_1, \dots, \Sigma_n, \alpha_n, \Sigma_{n+1}|v)_k \cup D$.

2. We say that L is *weak k-rightdeterministic* if and only if there exist $w \in A^k$, $\alpha_2, \dots, \alpha_{n+1} \in A^{k+1}$ and $\Sigma_1, \dots, \Sigma_{n+1} \subseteq A^{k+1}$ for some $n \geq 0$ and a set $D \subseteq A^{\leq k} \backslash \{\epsilon\}$ such that for $2 \leq i \leq n+1$ it holds that $\alpha_i \notin \Sigma_i$ and $L = (w|\Sigma_1, \alpha_2, \Sigma_2, \dots, \alpha_{n+1}, \Sigma_{n+1})_k \cup D$

If we go carefully through the proof of Theorem 5.6 we can establish a similar theorem for $\hat{\mathcal{D}}^{\text{left}}_k$ and $\hat{\mathcal{D}}^{\text{right}}_k$.

Theorem 7.7. *Let $k \geq 0$ and $L \subseteq A^+$.*

1. *$L \in \hat{\mathcal{D}}_k^{\text{left}}$ if and only L is a finite union of weak k-leftdeterministic languages.*
2. *$L \in \hat{\mathcal{D}}_k^{\text{right}}$ if and only L is a finite union of weak k-rightdeterministic languages.*

Of course all these classes are Boolean algebras as well. What are their characterizations in terms of temporal logic? The subtle difference between deterministic languages and their weak versions is that in the first case we fix a block of $k+1$ letters at the beginning or end of each word, and in the latter case we only fix a block of k letters. In restricted temporal logic this translates to a next or previously depth of $k-1$ in the **F**-free or **P**-free part of the formula, and depth k otherwise. A depth of '-1' here means that there is no **F**-free or **P**-free part in the formula. We do not go into details here, but turn to the levels of dot-depth one.

Theorem 7.8. *Let $k \geq 0$ and $L \subseteq A^+$. Then $L \in \mathcal{B}_{1,k}$ if and only if L is in the Boolean algebra generated by languages of the form*

$$(w|A^{k+1}, \alpha_1, A^{k+1}, \alpha_2, A^{k+1}, \dots, A^{k+1}, \alpha_m, A^{k+1}|v)_k \cup D$$

where $m \geq 0, \alpha_1, \dots, \alpha_m \in A^{k+1}$, $w, v \in A^k$ and $D \subseteq A^{<k+m} \backslash \{\epsilon\}$.

Proof. Let $L \in \mathcal{B}_{1,k}$ and assume that L is a finite union of equivalence classes $[x]_{m,k}$ for some $m \geq 0$. If this class is a singleton then we can describe it with the finite set D. So assume $|x| \geq m+k$ and look at $[x]_{m,k}$ given by $p_k(x)$, $s_k(x)$ and $\mu_{m,k}(x)$. Then it holds that

$$\begin{aligned} [x]_{m,k} \quad = \quad & \bigcap \ (w|A^{k+1}, \alpha_1, A^{k+1}, \alpha_2, A^{k+1}, \dots, A^{k+1}, \alpha_m, A^{k+1}|v) \ \backslash \\ & \bigcup \ (w|A^{k+1}, \alpha_1, A^{k+1}, \alpha_2, A^{k+1}, \dots, A^{k+1}, \alpha_m, A^{k+1}|v) \end{aligned}$$

where $w =_{\text{def}} p_k(x)$, $v =_{\text{def}} s_k(x)$ and the intersection (union) goes over all tuples $(\alpha_1, \dots, \alpha_m)$ in $\mu_{m,k}(x)$ (not in $\mu_{m,k}(x)$, respectively).

Now let $L = (w|A^{k+1}, \alpha_1, A^{k+1}, \alpha_2, A^{k+1}, \dots, A^{k+1}, \alpha_m, A^{k+1}|v)_k \cup D$. We show that L is saturated by $\sim_{m,k}$. If $x \in D$ we have $\{x\} = [x]_{m,k}$. Let $x \in L \backslash D$ and let y such that $x \sim_{m,k} y$. Then $p_k(x) = p_k(y) = w$, $s_k(x) = s_k(y) = v$ and in particular $(\alpha_1, \dots, \alpha_m)$ appears in y. It follows that $y \in L$. ❑

The case $k = 0$ in the levelwise inclusion structure is somewhat special and was already studied in a different context. First note that if a minimal dfa F has pattern $\mathbf{R_0}$ then is also has pattern $\mathbf{T_0}$ or pattern $\mathbf{T_0^{rev}}$. It follows that $\mathcal{B}_{1,0} = \hat{\mathcal{D}}_0^{\text{right}} \cap \hat{\mathcal{D}}_0^{\text{left}}$. But this is a well known equation from the so-called Straubing-Thérien hierarchy. The class $\mathcal{B}_{1,0}$ is (up to the empty word) equal to level one of this hierarchy and the weak 0-rightdeterministic and weak 0-leftdeterministic languages are the langauges that form the varieties of languages having $\mathcal{R}$-trivial and $\mathcal{L}$-trivial syntactic monoids ([Eil76]).

As we have seen the 'level' k of a language can be bound by a constant which relies on the size of the given dfa. If we do not fix k we get in the limit of course the same patterns for $\mathbf{T_k}$ and $\mathbf{T_k^{rev}}$ as in the case of $\mathbf{S_k}$ and $\mathbf{S_k^{rev}}$, namely $\mathbf{S}$ and $\mathbf{S^{rev}}$. In case of pattern $\mathbf{R_k}$ we encounter a new pattern $\mathbf{R}$.

A dfa F has pattern $\mathbf{R}$ if there are states $s_1, \dots, s_6, s, s' \in S$ and words $x, y, y', w, w', u, v, z \in A^*$ with $|w|, |w'| \geq 1$, $s \in S' \Leftrightarrow s' \notin S'$, $\delta(s_0, x) = s_1$, $\delta(s_1, y) = s_2$, $\delta(s_2, y') = s_1$, $\delta(s_i, w) = s_i$ for $i \in \{1,4,5\}$, $\delta(s_i, w') = s_i$ for $i \in \{2,3,6\}$, $\delta(s_1, u) = s_3 = \delta(s_5, u)$, $\delta(s_2, v) = s_4 = \delta(s_6, v)$, $\delta(s_3, v) = s_5$, $\delta(s_4, u) = s_6$, $\delta(s_2, z) = s$ and $\delta(s_5, z) = s'$ (cf. Figure 11).

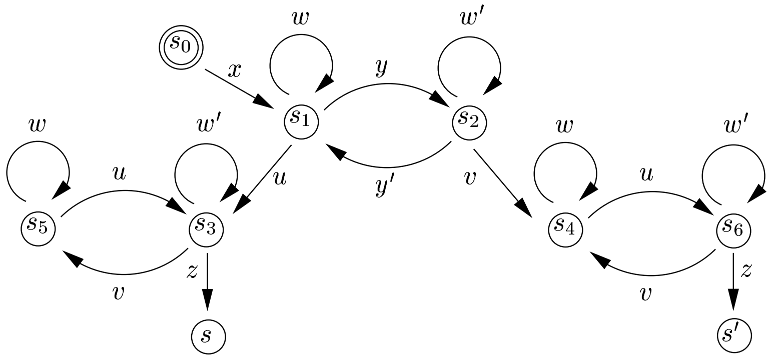

Figure 11: Pattern **R** with $|w|, |w'| \geq 1$ and $s \in S' \Leftrightarrow s' \notin S'$.

Proposition 7.9. *Let F be a dfa.*

1. *F does not have pattern* **S** *if and only if there exists some $k \geq 0$ such that F does not have pattern $\mathbf{T_k}$.*
2. *F does not have pattern* $\mathbf{S^{rev}}$ *if and only if there exists some $k \geq 0$ such that F does not have pattern $\mathbf{T_k^{rev}}$.*
3. *F does not have pattern* **R** *if and only if there exists some $k \geq 0$ such that F does not have pattern $\mathbf{R_k}$.*

The characterization of $\mathcal{B}_1$ in terms of these patterns is straightforward.

Theorem 7.10. *Let $L \subseteq A^+$, $k \geq 0$ and let F be the minimal dfa accepting L. Then the following statements are equivalent:*

(1) $L \in \mathcal{B}_1$

(2) F does not have any of the patterns **S**, $\mathbf{S^{rev}}$ *and* **R**.

(3) F does not have pattern **R**.

We notice that in a sense pattern **R** describes the difference in $\mathcal{B}_1 \subsetneq L(\mathrm{RTL}) \cap L(\mathrm{RPTL})$.

Conclusions

We have proved a detailed levelwise characterization of languages definable by formulas of restricted temporal logic in terms of the structure of their deterministic finite automata. We could give a language theoretical description of such languages by generalizing the notion of deterministic languages. Easy proofs of strictness and decidability results were given. We believe that this contributes to the better understanding of the classes $L(\mathrm{RTL})$ and $L(\mathrm{RPTL})$, which are located in the difficult and challenging area of the dot-depth hierarchy with its fundamental open questions.

In a following paper we will study these classes from a different point of view. It turns out that our characerizations in terms of deterministic languages has interesting consequences in the context of complexity theory. Along the lines of [Bor95, BKS98] we look at RTL-definable languages and their relation to Δ_2^{p}. Motivated from these investigations we would like to find a description of $L(\mathrm{RTL}) \cap L(\mathrm{RPTL})$ in terms of formal languages. Also a levelwise characterization would be helpful here.

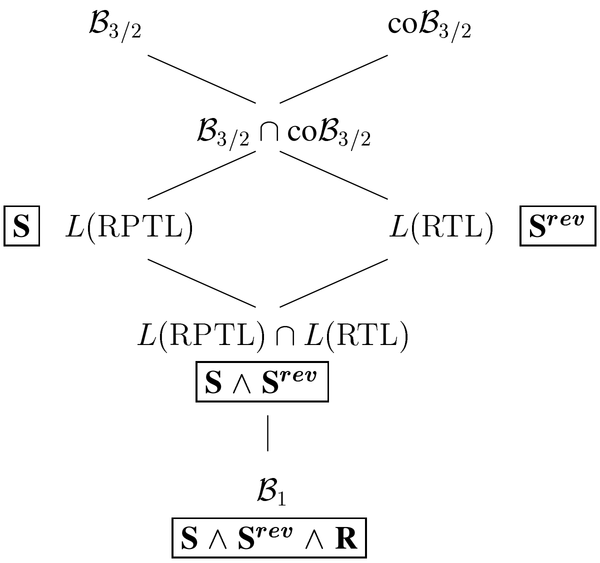

Figure 12: All inclusions are known to be strict.

Acknowledgements. Thanks are due to Klaus Wagner, Würzburg, who initiated this work by giving a proof of Theorem 6.3 for the case $k = 0$ and to Heribert Vollmer, Würzburg, for fruitful discussions on the topic.

References

[BF80] J. A. Brzozowski and F. E. Fich. Languages of r-trivial monoids. *Journal of Computer and System Sciences*, 20:32–49, 1980.

[BKS98] B. Borchert, D. Kuske, and F. Stephan. On existentially first-order definable languages and their relation to NP. In *Proceedings 25rd International Colloquium on Automata, Languages and Programming*, Lecture Notes in Computer Science. Springer Verlag, 1998. To appear.

[Bor95] B. Borchert. On the acceptance power of regular languages. *Theoretical Computer Science*, 148:207–225, 1995.

[Brz76] J. A. Brzozowski. Hierarchies of aperiodic languages. *RAIRO Inform. Theor.*, 10:33–49, 1976.

[BS73] J. A. Brzozowski and I. Simon. Characterizations of locally testable languages. *Discrete Math.*, 4:243–271, 1973.

[CH91] S. Cho and D. T. Huynh. Finite-automaton aperiodicity is PSPACE-complete. *Theoretical Computer Science*, 88:99–116, 1991.

[CPP93] R. S. Cohen, D. Perrin, and J. E. Pin. On the expressive power of temporal logic. *Journal of Computer and System Sciences*, 46:271–294, 1993.

[Eil76] S. Eilenberg. *Automata, languages and machines*, volume B. Academic Press, New York, 1976.

[Eme90] E. A. Emerson. Temporal and modal logic. In J. van Leeuwen, editor, *Handbook of theoretical computer science*, volume B, pages 995–1072. Springer, 1990.

[EW96] K. Etessami and T. Wilke. An until hierarchy for temporal logic. In *Proceedings 11th IEEE Symposium on Logic in Computer Science*, pages 108–117. IEEE Computer Society Press, 1996.

[McN74] R. McNaughton. Algebraic decision procedures for local testablility. *Math. Syst. Theor.*, 8:60–76, 1974.

[MP71] R. McNaughton and S. Papert. *Counterfree Automata*. MIT Press, Cambridge, 1971.

[Pin86] J. E. Pin. *Varieties of formal languages*. Plenum, New York, 1986.

[Pin95] J. E. Pin. Finite semigroups and recognizable languages: an introduction. In J. Fountain, editor, *NATO Advanced Study Institute: Semigroups, Formal Languages and Groups*, pages 1–32. Kluwer Academic Publishers, 1995.

[Pin96a] J. E. Pin. Logic, semigroups and automata on words. *Annals of Mathematics and Artificial Intelligence*, 16:343–384, 1996.

[Pin96b] J. E. Pin. Syntactic semigroups. In G.Rozenberg and A.Salomaa, editors, *Handbook of formal languages*, volume I, pages 679–746. Springer, 1996.

[PW97] J. E. Pin and P. Weil. Polynomial closure and unambiguous product. *Theory of computing systems*, 30:383–422, 1997.

[Sim72] I. Simon. *Hierarchies of events with dot-depth one*. PhD thesis, University of Waterloo, 1972.

[Ste85a] J. Stern. Characterizations of some classes of regular events. *Theoretical Computer Science*, 35:17–42, 1985.

[Ste85b] J. Stern. Complexity of some problems from the theory of automata. *Information and Control*, 66:163–176, 1985.

[Tho96] W. Thomas. Languages, automata, and logic. In G.Rozenberg and A.Salomaa, editors, *Handbook of formal languages*, volume III, pages 389–456. Springer, 1996.

[TW96] D. Thérien and Th. Wilke. Temporal logic and semidirect products: an effective characterization of the until hierarchy. In *Proceedings of the 37th Annual Symposium on Foundations of Computer Science*. IEEE Computer Society Press, 1996.

[TW98] D. Thérien and Th. Wilke. Over words, two variables are as powerful as one quantifier alternation: $\mathrm{FO}^{\mathbf{2}} = \Sigma_{\mathbf{2}} \cap \Pi_{\mathbf{2}}$. In *Proceedings of the 30th annual ACM symposium on Theory of computing*, pages 234–240, Dallas, Texas, 1998. ACM Press.

[Wil99] Th. Wilke. Classifying discrete temporal properties. In *Proceedings 16th Symposium on Theoretical Aspects of Computer Science*, volume 1563 of *Lecture Notes in Computer Science*, pages 32–46. Springer-Verlag, 1999.