The Boolean Hierarchy over Dot-Depth 1/2

Christian Glaßer

Theoretische Informatik
Universität Würzburg
Am Exerzierplatz 3
97072 Würzburg, Germany
glasser@informatik.uni-wuerzburg.de

6th June 1999

Abstract

For some fixed alphabet A with $|\mathrm{A}| \geq 2$, a language $L \subseteq \mathrm{A}^+$ is in the class $\mathcal{B}_{1/2}$ of the dot-depth hierachy if and only if it can be written as a finite union of languages $u_0\mathrm{A}^+u_1\mathrm{A}^+u_2\cdots\mathrm{A}^+u_n$, where $u_i \in \mathrm{A}^*$ and $n \geq 0$.

Using an automata-theoretic approach, we show a membership criterion for the single classes of the Boolean hierachy over $\mathcal{B}_{1/2}$. From this criterion it follows that the Boolean hierachy over $\mathcal{B}_{1/2}$ is proper and decidable. In finite model theory this implies the decidability of the classes of the Boolean hierachy over the class Σ_1 of the FOX[<]-logic. Furthermore, as a corollary we obtain a reproof of the forbidden pattern for $\mathcal{B}_{1/2}$.

Classification: finite automata, concatenation hierarchies, Boolean hierarchy, decidability

1 Introduction

Let A be a finite alphabet with $|\mathrm{A}| \geq 2$, by A^+ we denote the set of all non-empty words. Regular languages can be defined from alphabet letters by using Boolean operations, concatenation and iteration. Ignoring iterations, the class of *starfree regular languages* of A^+ can be defined as the smallest class containing the atomic languages and being closed under finite Boolean operations and concatenation. A systematic way to examine this class of languages is to count the number of alternating uses of Boolean operations on the one hand and concatenations on the other hand. Grouping together languages of similar concatenation complexity leads in a natural way to the definition of concatenation hierarchies that exhaust the class of starfree regular languages.

A prominent example of such an hierarchy is the dot-depth hierarchy, first studied in [CB71] by Brzozowski and Cohen. In [BK78] Brzozowski and Knast showed that the dot-depth hierarchy is strict. Several characterisations for this hierarchy are known, e.g. in the theory of formal languages, in the theory of finite automata, in algebra and in logic. Furthermore, this hierarchy has a close connection to complexity theory, via the leaf languages approach. Unfortunately, the decidability of its levels is still an open problem, although a lot of effort via different approaches has been invested. It is only known that the levels 1/2 and 1 are decidable (cf. [Kna83, Ste85b, PW97]). For an overview of concatenation hierarchies see for example the articles [Brz76, Pin96a, Pin96b, Tho96, Sch98].

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 w.r.t. A^+. For a class $\mathcal{C}$ of languages of A^+, let $\mathrm{Pol}(\mathcal{C})$ be its *polynomial closure*, i.e. the class of languages $L \subseteq \mathrm{A}^+$ that can be written as a finite union of languages of the form $u_0L_1u_1\cdots L_nu_n$, where $u_i \in \mathrm{A}^*$, $L_i \in \mathcal{C}$ and $n \geq 0$. Now the dot-depth hierarchy can be defined in terms of the theory of formal languages as the following family of classes of languages of A^+.

$$\begin{aligned}
\boldsymbol{\mathcal{B}_0} &\stackrel{df}{=} \{\emptyset, \mathrm{A}^+\} \\
\boldsymbol{\mathcal{B}_{n+\frac{1}{2}}} &\stackrel{df}{=} \mathrm{Pol}\,(\mathcal{B}_n) \quad \text{for } n \geq 0 \\
\boldsymbol{\mathcal{B}_{n+1}} &\stackrel{df}{=} \mathrm{BC}\left(\mathcal{B}_{n+\frac{1}{2}}\right) \quad \text{for } n \geq 0
\end{aligned}$$

For $n \geq 0$ we will also use the classes $\mathrm{co}\mathcal{B}_{n+1/2}$, where $\mathrm{co}\mathcal{C} =_{df} \{\mathrm{A}^+ \setminus L : L \in \mathcal{C}\}$ for a class $\mathcal{C}$.

Another way to define the dot-depth hierarchy is to use logical formulas. To define formulas of the class FOX[$<$] we use the binary relation symbols $<$, S, P and the unary relation symbols min, max and π_a for each letter $a \in \mathrm{A}$. Then atomic formulas are of the type $x = y$, $x < y$, $x\mathrm{S}y$, $x\mathrm{P}y$, $\min x$, $\max x$ and $\pi_a x$, with variables x, y. Further formulas are constructed from atomic formulas by using the connectives $\neg, \vee, \wedge$ and quantifiers $\exists, \forall$ bounding variables. The formulas 0 (false) and 1 (true) are considered as well. We say a language $L \subseteq \mathrm{A}^+$ is FOX[$<$]-definable if there exists a sentence ϕ (i.e. a formula of the above type without free variables) such that $w \in L$ if and only if w satisfies ϕ when variables are interpreted as positions in w, where the following holds:

- $\pi_a x$ means that the letter at position x is a
- $\min x$ ($\max x$) means that x is the first (last, resp.) position in w
- $x\mathrm{S}y$ ($x\mathrm{P}y$) means that position y is the successor (predecessor, resp.) of x
- $<$ and $=$ are the usual relations on $\{1, \ldots, |w|\}$

Thomas proved in [Tho82] that there is a natural hierarchy within FOX[$<$] in one-one correspondence with the classes of the dot-depth hierarchy. Let Σ_k (Π_k) be the subclass of FOX[$<$]-formulas which have at most $k-1$ quantifier alternations, starting with an existential (universal, resp.) quantifier.

Theorem 1.1 [Tho82] *Let $k \geq 1$ and $L \subseteq \mathrm{A}^+$.*

1. *$L \in \mathcal{B}_{k-1/2}$ if and only if L is* FOX[$<$]*-definable by a Σ_k formula.*
2. *$L \in \mathrm{co}\mathcal{B}_{k-1/2}$ if and only if L is* FOX[$<$]*-definable by a Π_k formula.*
3. *$L \in \mathcal{B}_k$ if and only if L is* FOX[$<$]*-definable by a Boolean combination of Σ_k formulas.*

Later we will prove that the Boolean hierarchy over dot-depth 1/2 is decidable. From Theorem 1.1 it follows that the Boolean hierachy over the class Σ_1 of the FOX[$<$]-logic is also decidable.

The paper is organized as follows. In Section 2 we give some basic definitions, and in Section 3 we generalize the notion of alternating word extension chains used in [SW98]. In Sections 4 and 5 we define and investigate structured words, which is a central notion of this paper. Sections 6 and 7 establish the correspondence between a certain measure depending on a given automaton F and the level of $\mathrm{L}\,(F)$ in the Boolean hierarchy over $\mathcal{B}_{1/2}$. Finally, in Section 8 we conclude the main results of this paper.

2 Preliminaries

First of all we summarize some notations. We denote the set of natural numbers, positive natural numbers by $\mathbb{N}$, $\mathbb{N}^+$, respectively. We fix an alphabet A with $|\mathrm{A}| \geq 2$. The empty word is denoted by ε, the set of all words over A is denoted by A^* and the set of all non-empty words over A is denoted by A^+. Let $k \in \mathbb{N}$, $w \in \mathrm{A}^*$ with $w = a_1 a_2 \cdots a_k$ and $a_1, \ldots, a_k \in \mathrm{A}$. We define the set of letters of w as $\alpha(w) =_{df} \{a_1, a_2, \ldots, a_k\}$.

We consider only deterministic, finite automata. For a given finite automaton $F = (\mathrm{A}, S, \delta, s_0, S')$ we extend the transition function δ to input words. The size of F is defined as $|F| =_{df} |S| > 0$, and the language accepted by F is denoted by $\mathrm{L}(F)$. We say that a state $s \in S$ has a loop $l \in \mathrm{A}^+$ (has an l-loop) if and only if $\delta(s, l) = s$. A total mapping $\delta' : S \to S$ leads to an l-loop if and only if $\delta'(s)$ has an l-loop for all $s \in S$. Every $w \in \mathrm{A}^*$ induces a total mapping $\delta^w : S \to S$ with $\delta^w(s) =_{df} \delta(s, w)$. Let $v, w \in \mathrm{A}^*$ then we define $v \sim_F w$ if and only if $\delta^v = \delta^w$.

Throughout the paper (if no other definition is made) we consider languages as subsets of A^+. For such a language $L \subseteq \mathrm{A}^+$, we take the complement w.r.t. A^+. For a class of languages $\mathcal{C}$ which is closed under union and intersection, the Boolean hierarchy over $\mathcal{C}$ is the family of classes $\mathcal{C}(k)$ and $\mathrm{co}\mathcal{C}(k)$ with $k \geq 1$,

$$\begin{aligned} \mathcal{C}(k) &\overset{df}{=} \overbrace{\mathcal{C} \oplus \mathcal{C} \oplus \cdots \oplus \mathcal{C}}^{k \text{ times}} \quad \text{and} \\ \mathrm{co}\mathcal{C}(k) &\overset{df}{=} \left\{ \mathrm{A}^+ \setminus L : L \in \mathcal{C}(k) \right\}, \end{aligned}$$

where $\mathcal{C} \oplus \mathcal{C} =_{df} \{A \triangle B : A, B \in \mathcal{C}\}$ and $A \triangle B =_{df} (A \setminus B) \cup (B \setminus A)$. The properties of Boolean hierarchies are well studied in [KSW87, CGH$^+$88], the following lemma states some known results.

Lemma 2.1 *Let $k \geq 1$ and $\mathcal{C}$ be a class of languages being closed under union and intersection.*

1. $\mathrm{BC}(\mathcal{C}) = \bigcup_{k \geq 1} \mathcal{C}(k)$
2. $\mathcal{C}(k) \cup \mathrm{co}\mathcal{C}(k) \subseteq \mathcal{C}(k+1) \cap \mathrm{co}\mathcal{C}(k+1)$
3. $L \in \mathcal{C}(2k) \iff L = \bigcup_{i=0}^{k-1} (L_{2i} \setminus L_{2i+1})$ *for suitable* $L_i \in \mathcal{C}$
4. $L \in \mathcal{C}(2k-1) \iff L = \bigcup_{i=0}^{k-1} (L_{2i} \setminus L_{2i+1})$ *for suitable* $L_i \in \mathcal{C}$ *with* $L_{2k-1} = \emptyset$
5. $L \in \mathcal{C}(2k) \iff L = \bigcup_{i=0}^{k-1} (L_{2i} \setminus L_{2i+1})$ *for suitable* $L_i \in \mathcal{C}$ *with* $L_0 \supseteq L_1 \supseteq \cdots \supseteq L_{2k-1}$
6. $L \in \mathcal{C}(2k-1) \iff L = \bigcup_{i=0}^{k-1} (L_{2i} \setminus L_{2i+1})$ *for suitable* $L_i \in \mathcal{C}$ *with* $L_0 \supseteq L_1 \supseteq \cdots \supseteq L_{2k-1} = \emptyset$

Later we will show, that words being sufficiently long w.r.t. to the size of a given automata, can be divided into basic components. For this we will use the following functions:

Definition 2.2 *For $n, k \in \mathbb{N}$ and $a \in \mathbb{N}^+$ let*

$$\begin{aligned} \boldsymbol{\mathcal{K}(n)} &\overset{df}{=} (n+1)^{(n+1)^{(n+1)}}, \\ \boldsymbol{\mathcal{M}(a,n)} &\overset{df}{=} \mathcal{K}(n)^3 \cdot a^{\mathcal{K}(n)+1} \quad \textit{and} \\ \boldsymbol{\mathcal{N}(a,n,k)} &\overset{df}{=} (\mathcal{M}(a,n) + \mathcal{K}(n) + 1)^{k+1}. \end{aligned}$$

Lemma 2.3 *Let $n \in \mathbb{N}^+$, $m_1 =_{df} \lfloor \mathcal{K}(n)/2 \rfloor$ and $m_{i+1} =_{df} \lfloor m_i / n^n \rfloor - 1$ for $i \geq 1$. Then the following holds for $1 \leq i \leq n^n + 1$:*

$$m_i \geq (2n^n)^{(n^n+3-i)}$$

Proof We will prove the lemma by induction on i with $1 \leq i \leq n^n + 1$. First of all we want to prove the induction base, i.e. the case $i = 1$. By definition of $\mathcal{K}(n)$, for $n = 1$ we have $m_1 = 8 = (2n^n)^{(n^n+3-1)}$. For $n > 1$ we obtain the following estimation

$$\begin{aligned}
(2n^n)^{(n^n+3-1)} = \left(n^n + n \cdot n^{n-1}\right)^{(n^n+3-1)} &\leq (n+1)^{n(n^n+2)} && (1)\\
&\leq (n+1)^{\left((n+1)^{(n+1)}-2\right)} && (2)\\
&\leq (n+1)^{\left((n+1)^{(n+1)}-1\right)} - 1 && (3)\\
&\leq \frac{(n+1)^{\left((n+1)^{(n+1)}\right)}}{2} - 1 && (4)\\
&\leq \lfloor \mathcal{K}(n)/2 \rfloor = m_1 && (5)
\end{aligned}$$

where Equation (2) is obtained from Equation (1) by using the binomial theorem, i.e. $(n+1)^{n+1} \geq n^{n+1} + (n+1)\, n^n + (n+1)\, n + 1 \geq n^{n+1} + 2n + 2$ (here we need $n > 1$). This shows the induction base.

Suppose now that we have already shown $m_k \geq (2n^n)^{(n^n+3-k)}$ with $1 \leq k < n^n + 1$. By definition, $m_{k+1} = \lfloor m_k / n^n \rfloor - 1$. By induction hypothesis we obtain

$$m_{k+1} \geq \frac{(2n^n)^{(n^n+3-k)}}{n^n} - 2.$$

Since $n > 0$ and $k < n^n + 1$ we have $(2n^n)^{(n^n+3-k)}/n^n \geq 8$. It follows that

$$\frac{(2n^n)^{(n^n+3-k)}}{n^n} - 2 = 2 \cdot (2n^n)^{(n^n+3-(k+1))} - 2 \geq (2n^n)^{(n^n+3-(k+1))}.$$

This completes the induction. $\square$

3 General Alternating Chains

In this section we will generalize the notion of alternating word extension chains used in [SW98]. We will see that the essential point of a word extension is the notion of a quasiordering. From this algebraic point of view we can define general alternating chains without using any facts from the theory of finite automata. Furthermore, we generalize basic facts of alternating word extension chains.

Definition 3.1 *Let M be a set and let $\leq$ be a binary, reflexive, transitive relation on M, then we call $\mathcal{Q} = (M, \leq)$ a* **quasiordering**.

Definition 3.2 *Let $\mathcal{Q} = (M, \leq)$ be a quasiordering, $n \in \mathbb{N}$ and $a_0, a_1, \ldots, a_n \in M$ such that $a_0 \leq a_1 \leq \cdots \leq a_n$, then we call $a_0 \leq a_1 \leq \cdots \leq a_n$ a* **chain of length n**.

Definition 3.3 *Let $\mathcal{Q} = (M, \leq)$ be a quasiordering, $N \subseteq M$ and let $a_0 \leq \cdots \leq a_n$ be a chain. We call this chain* **N-alternating** *if and only if $a_i \in N \iff i \equiv 0(2)$ for all $0 \leq i \leq n$.*

Analogously to [SW98], for a given $k \geq 0$, we are interested in all words which can be obtained by an alternating chain of length k. The following definition generalizes the measure which was used in [SW98].

Definition 3.4 *For a quasiordering $\mathcal{Q} = (M, \leq)$, a set $N \subseteq M$, $b \in M$ and $k \geq 0$ we define:*

$$\begin{aligned}
\boldsymbol{L_{\mathcal{Q}}(N, k)} &\overset{df}{=} \{a \in M : \text{there exists an } N\text{-alternating chain } a_0 \leq \cdots \leq a_k \text{ with } a_k \leq a\}\\
\boldsymbol{m_{\mathcal{Q}}(N, b)} &\overset{df}{=} \sup\{k : b \in L_{\mathcal{Q}}(N, k)\}\\
\boldsymbol{m_{\mathcal{Q}}(N)} &\overset{df}{=} \sup\{k : L_{\mathcal{Q}}(N, k) \neq \emptyset\}
\end{aligned}$$

Definition 3.5 *Let $\mathcal{Q} = (M, \leq)$ be a quasiordering and $N_1, N_2 \subseteq M$, then* **$\mathbf{N_1}$ is (finitely) generated by $\mathbf{N_2}$** *if and only if (N_2 is finite and)*

$$N_1 \stackrel{df}{=} \left\{ m \in M : \bigvee_{n \in N_2} n \leq m \right\}.$$

Lemma 3.6 *For every quasiordering $\mathcal{Q} = (M, \leq)$ and every $N \subseteq M$ we have:*

1. $L_{\mathcal{Q}}(N, 0) \supseteq L_{\mathcal{Q}}(N, 1) \supseteq L_{\mathcal{Q}}(N, 2) \supseteq \cdots$
2. *For $k \geq 0$, $L_{\mathcal{Q}}(N, k)$ is closed under $\leq$, i.e. if $a \in L_{\mathcal{Q}}(N, k)$ and $a \leq b$ then $b \in L_{\mathcal{Q}}(N, k)$.*
3. $N \subseteq L_{\mathcal{Q}}(N, 0)$

Proof The lemma is an easy consequence from Definition 3.4. □

Lemma 3.7 *Let $\mathcal{Q} = (M, \leq)$ be a quasiordering and $N \subseteq M$ with $m_{\mathcal{Q}}(N, a) < \infty$ for all $a \in M$, then the following holds:*

$$N = \bigcup_{k \in \mathbb{N}} (L_{\mathcal{Q}}(N, 2k) \setminus L_{\mathcal{Q}}(N, 2k+1))$$

Proof "$\subseteq$" Let $a \in N$, by Lemma 3.6 we have $a \in L_{\mathcal{Q}}(N, 0)$. Furthermore, there exists a minimal $j \in \mathbb{N}$ with $a \in L_{\mathcal{Q}}(N, j) \setminus L_{\mathcal{Q}}(N, j+1)$. Otherwise, we would have $a \in L_{\mathcal{Q}}(N, l)$ for all $l \in \mathbb{N}$. Particularly, $m_{\mathcal{Q}}(N, a) = \infty$, a contradiction to our assumption. Now we want to see that $j \equiv 0(2)$. Assume that $j = 2k + 1$, thus there exists an N-alternating chain $a_0 \leq \cdots \leq a_{2k+1}$ with $a_{2k+1} \leq a$. It follows that $a_{2k+1} \notin N$. Therefore, $a_0 \leq \cdots \leq a_{2k+1} \leq a$ is an N-alternating chain with $a \leq a$. Hence $a \in L_{\mathcal{Q}}(N, j+1)$, a contradiction to the definition of j. We conclude $j \equiv 0(2)$, thus there exists a $k \in \mathbb{N}$ with $a \in L_{\mathcal{Q}}(N, 2k) \setminus L_{\mathcal{Q}}(N, 2k+1)$.

"$\supseteq$" Now let $k \in \mathbb{N}$, $a \in L_{\mathcal{Q}}(N, 2k) \setminus L_{\mathcal{Q}}(N, 2k+1)$ and suppose that $a \notin N$. There exists an N-alternating chain $a_0 \leq \cdots \leq a_{2k}$ with $a_{2k} \leq a$. Since $a_{2k} \in N$, it follows that $a_0 \leq \cdots \leq a_{2k} \leq a$ is an N-alternating chain and $a \in L_{\mathcal{Q}}(N, 2k+1)$. This is a contradiction to our assumption, thus we have $a \in N$. □

4 Alternating Chains of Structured Words

In this section we will define a central notion of this paper, i.e. the notion of structured words. Informally, given a finite automaton we can divide every sufficiently long word into small blocks such that every second block u has the following property: If we start in an arbitrary state of the automaton and read the word u, then we reach a state having an u-loop. Furthermore, we will define extensions of structured words and we will see that those extensions induces a quasiordering on structured words.

Definition 4.1 *For a finite automaton $F = (\mathrm{A}, S, \delta, s_0, S')$ we define* **the set of structured words over F***:*

$$\mathcal{A}_{\mathcal{F}} \stackrel{df}{=} \left\{ w_0 \boxed{u_1} w_1 \boxed{u_2} w_2 \cdots \boxed{u_k} w_k : k \in \mathbb{N},\ w_i, u_i \in \mathrm{A}^*,\ |w_i|, |u_i| \leq \mathcal{K}(|S|),\ |u_i| \geq 1,\ \delta^{u_i} = \delta^{u_i u_i} \right\}$$

The following definition allows us to transform structured words into "normal" words. With this definition we are able to adapt all notions of A^* to $\mathcal{A}_{\mathcal{F}}$. For instance, this defines the length of structured words and the binary relation $\sim_F$ for them. The next section deals with the reverse direction, i.e. the transformation of "normal" words into structured words.

Definition 4.2 *Let* $w \in \mathcal{A}_{\mathcal{F}}$ *with* $w = w_0 \boxed{u_1} w_1 \boxed{u_2} w_2 \cdots \boxed{u_k} w_k$*, then* $\overline{\boldsymbol{w}} =_{df} w_0 u_1 w_1 \cdots u_k w_k$*. Furthermore,* $\overline{\boldsymbol{L}} =_{df} \{\overline{w} : w \in L\}$ *for every* $L \subseteq \mathcal{A}_{\mathcal{F}}$*.*

To avoid confusion with the notation for complementation of sets, troughout the paper we will never use bars for complementation (i.e. we will write $\mathrm{A}^+ \setminus L$ instead of the conventional notation).

Definition 4.3 *Let* $v, w \in \mathcal{A}_{\mathcal{F}}$ *then we define the following* **extensions of structured words**

- $v \leq_{\mathcal{F}} w \overset{df}{\Longleftrightarrow}$ *there exist* $a, b, z \in \mathcal{A}_{\mathcal{F}}$*,* $u \in \mathrm{A}^+$ *such that* $v = a\boxed{u}b$ *and* $w = a\boxed{u}z\boxed{u}b$*.*
- $v \preceq_{\mathcal{F}} w \overset{df}{\Longleftrightarrow}$ *there exist* $m \in \mathbb{N}$*,* $v_0, \dots, v_m \in \mathcal{A}_{\mathcal{F}}$ *such that* $v = v_0 \leq_{\mathcal{F}} \cdots \leq_{\mathcal{F}} v_m = w$*.*

The first extension is also called **elementary extension of structured words***.*

Lemma 4.4 *Let* F *be a finite automaton and* $v, w \in \mathcal{A}_{\mathcal{F}}$*, then the following holds:*

$v \preceq_{\mathcal{F}} w \Longleftrightarrow$ *there exist words* $v_0, \dots, v_m$*,* $z_1, \dots, z_m \in \mathcal{A}_{\mathcal{F}}$*,* $u_1, \dots, u_m \in \mathrm{A}^+$ *such that*

$$\begin{array}{rcl} v & = & v_0 \boxed{u_1} \qquad v_1 \boxed{u_2} \qquad v_2 \cdots \boxed{u_m} \qquad v_m \\ w & = & v_0 \boxed{u_1} z_1 \boxed{u_1} v_1 \boxed{u_2} z_2 \boxed{u_2} v_2 \cdots \boxed{u_m} z_m \boxed{u_m} v_m \end{array}$$

Proof The lemma can be easily proved by induction on the number of elementary extensions which are needed to transform v into w. □

Lemma 4.5 *Let* F *be a finite automaton, then* $\mathcal{F} =_{df} (\mathcal{A}_{\mathcal{F}}, \preceq_{\mathcal{F}})$ *is a quasiordering with* $m_{\mathcal{F}}(L, w) \leq |w| < \infty$ *for all* $L \subseteq \mathcal{A}_{\mathcal{F}}, w \in \mathcal{A}_{\mathcal{F}}$*.*

Proof It is easy to see that $\preceq_{\mathcal{F}}$ is reflexive and transitive. Let $L \subseteq \mathcal{A}_{\mathcal{F}}$, $w \in \mathcal{A}_{\mathcal{F}}$ and let $w_0 \preceq_{\mathcal{F}} \cdots \preceq_{\mathcal{F}} w_m$ be an L-alternating chain with $w_m \preceq_{\mathcal{F}} w$. Since the chain is alternating, we have $w_i \neq w_{i+1}$ for $0 \leq i < m$. Therefore, $|w_i| < |w_{i+1}|$ for $0 \leq i < m$. It follows that $m \leq |w|$, hence $m_{\mathcal{F}}(L, w) \leq |w|$. □

Definition 4.6 *Let* F *be a finite automaton,* $m \in \mathbb{N}$ *and* $\boldsymbol{\mathcal{F}} =_{df} (\mathcal{A}_{\mathcal{F}}, \preceq_{\mathcal{F}})$*.*

$$\begin{array}{rcl} \boldsymbol{\mathcal{L}_{\mathcal{F}}} & \overset{df}{=} & \{w \in \mathcal{A}_{\mathcal{F}} : \overline{w} \in \mathrm{L}\,(F)\} \\ \boldsymbol{L_{\mathcal{F}}(m)} & \overset{df}{=} & L_{\mathcal{F}}(\mathcal{L}_{\mathcal{F}}, m) \\ \boldsymbol{m_{\mathcal{F}}} & \overset{df}{=} & m_{\mathcal{F}}(\mathcal{L}_{\mathcal{F}}) \end{array}$$

Furthermore, an $\mathcal{L}_{\mathcal{F}}$*-alternating chain is simply called an* **alternating chain***.*

5 How to transform Words into Structured Words

Definition 4.1 shows the canonical way to transform structured words into "normal" words. In this section we will learn more about the reverse direction. More precisely, we will show that every "normal" word can be structured.

Lemma 5.1 *Let* $F = (\mathrm{A}, S, \delta, s_0, S')$ *be a finite automaton,* $m \in \mathbb{N}$*,* $u_1, u_2, \dots, u_m \in \mathrm{A}^+$*,* $v, w \in \mathrm{A}^*$ *and* $n =_{df} \lfloor m/(|S|^{|S|}) \rfloor$*. Furthermore, let* M *be a set of total mappings* $\delta' : S \to S$ *such that* δ' *leads to an* u_i*-loop for all* $\delta' \in M$*,* $1 \leq i \leq m$*. Then there exist words* $v', w' \in \mathrm{A}^*$*,* $u_1', \dots, u_n' \in \mathrm{A}^+$ *such that:*

1. $v'u'_1u'_2\cdots u'_nw' = vu_1u_2\cdots u_mw$

2. δ' *leads to an* u'_1*-loop for all* $\delta' \in M$, *(if* $n > 0$*)*

3. δ' *leads to an* u'_i*-loop for all* $\delta' \in M \cup \{\delta^{u'_1}\}$, $1 < i \le n$

Proof If $n = 0$ we set $v' = vu_1u_2\cdots u_mw$ and $w' = \varepsilon$. If $n = 1$ we set $v' = v$, $u'_1 = u_1$, $w' = u_2\cdots u_mw$ and we are done. Suppose now $n \ge 2$ and define $\delta_i =_{df} \delta^{u_1u_2\cdots u_i}$ for $1 \le i \le m$. Since there are at most $|S|^{|S|}$ total mappings $S \to S$, there exist mappings appearing several times on the list $\delta_1, \ldots, \delta_m$. Among these mappings we choose one which appears most frequently, say mapping $\overline{\delta}$ which appears n' times. Hence $n' \ge \lfloor m/(|S|^{|S|})\rfloor = n$ (by pigeon hole principle) and there are indices $i_1, \ldots, i_{n'}$ such that $\overline{\delta} = \delta_{i_1} = \delta_{i_2} = \cdots = \delta_{i_{n'}}$. Now define the following:

$$\begin{aligned}
v' &\stackrel{df}{=} v\\
u'_1 &\stackrel{df}{=} u_1u_2\cdots u_{i_1}\\
u'_2 &\stackrel{df}{=} u_{i_1+1}u_{i_1+2}\cdots u_{i_2}\\
u'_3 &\stackrel{df}{=} u_{i_2+1}u_{i_2+2}\cdots u_{i_3}\\
&\vdots\\
u'_n &\stackrel{df}{=} u_{(i_{n-1}+1)}u_{(i_{n-1}+2)}\cdots u_{i_n}\\
w' &\stackrel{df}{=} u_{i_n+1}u_{i_n+2}\cdots u_mw
\end{aligned}$$

Clearly, with this definition we have $v'u'_1u'_2\cdots u'_nw' = vu_1u_2\cdots u_mw$. By assumption, δ' leads to an u_i-loop for all $\delta' \in M$, $1 \le i \le m$. Thus for all $\delta' \in M$ and $1 \le i \le n$, δ' leads to an u'_i-loop, because u'_i is a concatenation of words u_j.

It remains to show that $\delta^{u'_1}$ leads to an u'_i-loop for all $1 < i \le n$. Since $\delta^{u'_1\cdots u'_{j-1}} = \overline{\delta} = \delta^{u'_1\cdots u'_j}$ for all $1 < j \le n$, we obtain that $\overline{\delta}$ leads to an u'_j-loop for all $1 < j \le n$. Since $\delta^{u'_1} = \overline{\delta}$, the lemma is proved. □

Lemma 5.2 *For every finite automaton* $F = (\mathrm{A}, S, \delta, s_0, S')$ *with* $n =_{df} |S|$ *and every* $w \in \mathrm{A}^*$ *there exist words* $w_0, \ldots, w_m \in \mathrm{A}^*$, $u_1, \ldots, u_m \in \mathrm{A}^+$ *with lengths* $\le \mathcal{K}(n)$ *such that* $w = w_0u_1w_1\cdots u_mw_m$ *and* $\delta^{u_i} = \delta^{u_iu_i}$ *for all* $1 \le i \le m$.

Proof If $|w| \le \mathcal{K}(n)$ then we are done. Suppose now that $|w| > \mathcal{K}(n)$. We divide w into parts of length $\lfloor\mathcal{K}(n)/2\rfloor$, i.e. $w = r_1\cdots r_k$ with $r_1, \ldots, r_k \in \mathrm{A}^*$, $|r_1| = \cdots = |r_{k-1}| = \lfloor\mathcal{K}(n)/2\rfloor$ and $|r_k| < \lfloor\mathcal{K}(n)/2\rfloor$. It remains to show that each r_i with $1 \le i < k$ can be represented as $r_i = v'_iu'_iw'_i$ with $v'_i, w'_i \in \mathrm{A}^*$, $u'_i \in \mathrm{A}^+$ and $\delta^{u'_i} = \delta^{u'_iu'_i}$.

Now we will repeatedly make use of Lemma 5.1. Let $r \in \mathrm{A}^*$ with $m_1 =_{df} \lfloor\mathcal{K}(n)/2\rfloor$, $r = u_{1,1}u_{1,2}\cdots u_{1,m_1}$ and $u_{1,1}, \ldots, u_{1,m_1} \in \mathrm{A}$. With $v_1 =_{df} \varepsilon$, $w_1 =_{df} \varepsilon$ and $M_1 =_{df} \emptyset$ we can apply Lemma 5.1 for a first time. For $n_1 =_{df} \lfloor m_1/n^n\rfloor$ we obtain words $v'_1, w'_1 \in \mathrm{A}^*$, $u'_{1,1}, \ldots, u'_{1,n_1} \in \mathrm{A}^+$ such that

1. $r = v'_1u'_{1,1}u'_{1,2}\cdots u'_{1,n_1}w'_1$

2. δ' leads to an $u'_{1,1}$-loop for all $\delta' \in M_1$, (if $n_1 > 0$)

3. δ' leads to an $u'_{1,j}$-loop for all $\delta' \in M_1 \cup \{\delta^{u'_{1,1}}\}$, $1 < j \le n_1$

With $m_2 =_{df} n_1 - 1$, $v_2 =_{df} v'_1u'_{1,1}$, $u_{2,1} =_{df} u'_{1,2}$, $u_{2,2} =_{df} u'_{1,3}$, $u_{2,3} =_{df} u'_{1,4}$, $\ldots$, $u_{2,m_2} =_{df} u'_{1,m_2+1}$, $w_2 =_{df} w'_1$, $M_2 =_{df} M_1 \cup \{\delta^{u'_{1,1}}\}$ we can apply Lemma 5.1 again. In general after the i-th application of Lemma 5.1 we obtain words $v'_i, w'_i \in \mathrm{A}^*$, $u'_{i,1}, \ldots, u'_{i,n_i} \in \mathrm{A}^+$ with $n_i = \lfloor m_i/n^n\rfloor$ such that

1. $r = v'_i u'_{i,1} u'_{i,2} \cdots u'_{i,n_i} w'_i$

2. δ' leads to an $u'_{i,1}$-loop for all $\delta' \in M_i$, (if $n_i > 0$)

3. δ' leads to an $u'_{i,j}$-loop for all $\delta' \in M_i \cup \{\delta^{u'_{i,1}}\}$, $1 < j \leq m_i$

With $m_{i+1} =_{df} n_i - 1$, $v_{i+1} =_{df} v'_i u'_{i,1}$, $u_{i+1,1} =_{df} u'_{i,2}$, $u_{i+1,2} =_{df} u'_{i,3}$, $\ldots$, $u_{i+1,m_{i+1}} =_{df} u'_{i,m_{i+1}+1}$, $w_{i+1} =_{df} w'_i$, $M_{i+1} =_{df} M_i \cup \{\delta^{u'_{i,1}}\}$ we can go to the $(i+1)$-st application of Lemma 5.1.

By Lemma 2.3, $m_i \geq (2n^n)^{(n^n+3-i)}$ for $1 \leq i \leq n^n + 1$. It follows that $n_i = \lfloor m_i / n^n \rfloor \geq (2n^n)^{(n^n+2-i)} - 1 \geq 1$ for $1 \leq i \leq n^n + 1$, particularly $u'_{i,1}$ exists for $1 \leq i \leq n^n + 1$. Remember that $M_{i+1} = M_i \cup \{\delta^{u'_{i,1}}\}$ and that δ' leads to an $u'_{i,1}$-loop for all $\delta' \in M_i$, $1 \leq i \leq n^n + 1$. Since there are at most n^n total mappings $S \to S$, there exists an h with $1 \leq h \leq n^n + 1$ such that $\delta^{u'_{h,1}}$ is already an element of M_h. It follows that $\delta^{u'_{h,1}}$ leads to an $u'_{h,1}$-loop, i.e. $\delta^{u'_{h,1} u'_{h,1}} = \delta^{u'_{h,1}}$. With $v' =_{df} v'_h$, $u' =_{df} u'_{h,1}$ and $w' =_{df} u'_{h,2} \cdots u'_{h,n_h} w'_h$ we have $r = v'u'w'$ with $v', w' \in \mathrm{A}^*$, $u' \in \mathrm{A}^+$ and $\delta^{u'u'} = \delta^{u'}$. We have shown that each r_i with $1 \leq i < k$ can be represented as $r_i = v'_i u'_i w'_i$ with $v'_i, w'_i \in \mathrm{A}^*$, $u'_i \in \mathrm{A}^+$ and $\delta^{u'_i} = \delta^{u'_i u'_i}$.

Finally, with $w_0 =_{df} v'_1$, $u_1 =_{df} u'_1$, $w_1 =_{df} w'_1 v'_2$, $u_2 =_{df} u'_2$, $w_2 =_{df} w'_2 v'_3$, $\ldots$, $u_{k-1} =_{df} u'_{k-1}$, $w_{k-1} =_{df} w'_{k-1} r_k$ we obtain the promised representation of w. □

Lemma 5.3 *For every finite automaton $F = (\mathrm{A}, S, \delta, s_0, S')$ and every $w \in \mathrm{A}^*$ there exists an $w' \in \mathcal{A}_{\mathcal{F}}$ with $w = \overline{w'}$.*

Proof Let $n =_{df} |S|$. By Lemma 5.2, w can be represented as $w = w_0 u_1 w_1 \cdots u_m w_m$ with $w_0, \ldots, w_m \in \mathrm{A}^*$, $u_1, \ldots, u_m \in \mathrm{A}^+$, $|w_0|, \ldots, |w_m|, |u_1|, \ldots, |u_m| \leq \mathcal{K}(n)$ and $\delta^{u_i} = \delta^{u_i u_i}$ for all $1 \leq i \leq m$. Therefore, with $w' =_{df} w_0 \boxed{u_1} w_1 \cdots \boxed{u_m} w_m$ we have $w' \in \mathcal{A}_{\mathcal{F}}$ and $w = \overline{w'}$. □

6 Small Measures induce Low Levels

We will prove that the language of an automaton with measure $< k$ is in Level k of the Boolean hierarchy over $\mathcal{B}_{1/2}$. For this we will show that for every structured word w which can be generated by an alternating chain of length n there exists a short structured word w' which can also be generated by an alternating chain of length n such that $w' \preceq_{\mathcal{F}} w$. It follows that $L_{\mathcal{F}}(m)$ is finitely generated for every $m \in \mathbb{N}$ and every finite automaton F. This fact can be understood as a kind of variation of a theorem by G. Higman [Hig52, SS83] which shows that $(\mathrm{A}^*, \preceq)$ is a well partial ordering (i.e. every ideal is finitely generated), where $\preceq$ stands for the "normal" word extension.

Lemma 6.1 *For every finite automaton $F = (\mathrm{A}, S, \delta, s_0, S')$ with $n =_{df} |S| > 0$ and every $w \in \mathcal{A}_{\mathcal{F}}$ there exists a $w' \in \mathcal{A}_{\mathcal{F}}$ such that $|w'| < \mathcal{M}(|\mathrm{A}|, n)$, $w' \sim_F w$ and $w' \preceq_{\mathcal{F}} w$.*

Proof Let $w \in \mathcal{A}_{\mathcal{F}}$ with $w = w_0 \boxed{u_1} w_1 \boxed{u_2} w_2 \cdots \boxed{u_k} w_k$ and $w_0, \ldots, w_k, u_1, \ldots, u_k \in \mathrm{A}^*$. Suppose $|w| \geq \mathcal{M}(|\mathrm{A}|, n) = \mathcal{K}(n)^3 \cdot |\mathrm{A}|^{\mathcal{K}(n)+1}$. It follows from an easy estimation that

$$|w| \geq 2 \cdot \mathcal{K}(n) \cdot \left(|\mathrm{A}|^{\mathcal{K}(n)+1} \cdot (n^n + 2) + 1 \right).$$

We obtain $k \geq |\mathrm{A}|^{\mathcal{K}(n)+1} \cdot (n^n + 2)$, because $|w_i|, |u_i| \leq \mathcal{K}(n)$. Let $m =_{df} n^n + 2$. Since $|u_i| \leq \mathcal{K}(n)$ for $1 \leq i \leq k$, there exist at most $|\mathrm{A}|^{\mathcal{K}(n)+1}$ different words u_i. Thus there exist indices $1 \leq i_1 < i_2 < \cdots < i_m \leq k$ with $u_{i_1} = u_{i_2} = \cdots = u_{i_m}$. For $2 \leq j \leq m$ we define $v_j =_{df} w_{i_1} \boxed{u_{i_1+1}} w_{i_1+1} \cdots \boxed{u_{i_j-1}} w_{i_j-1} \boxed{u_{i_j}}$. On the one hand there are at most n^n equivalence classes w.r.t. $\sim_F$, on the other hand we have defined $m - 1 = n^n + 1$ words v_j. Therefore, there exist indices $2 \leq a < b \leq m$ with $v_a \sim_F v_b$. Now we cut out the range $w_{i_a} \cdots \boxed{u_{i_b}}$ from w. We define $w'_1 =_{df} w_0 \boxed{u_1} w_1 \cdots \boxed{w_{i_a-1}}$, $w'_2 =_{df} w_{i_b} \cdots \boxed{u_k} w_k$ and $w' =_{df} w'_1 \boxed{u_{i_a}} w'_2$. Since we obtained w' from w by replacing the word v_b with v_a, and $v_a \sim_F v_b$, we have $w \sim_F w'$. Furthermore $|w'| < |w|$ and $w' \preceq_{\mathcal{F}} w$, because $w = w'_1 \boxed{u_{i_a}} w_{i_a} \cdots \boxed{u_{i_b}} w'_2$ and $u_{i_a} = u_{i_b}$.

We have shown that for every word $w \in \mathcal{A}_{\mathcal{F}}$ with $|w| \geq \mathcal{M}(|\mathrm{A}|, n)$ there exists a shorter word $w' \in \mathcal{A}_{\mathcal{F}}$ such that $w \sim_F w'$ and $w' \preceq_{\mathcal{F}} w$. Since $\preceq_{\mathcal{F}}$ and $\sim_F$ are transitive, we can repeat this procedure until we obtain a word w' with $|w'| < \mathcal{M}(|\mathrm{A}|, n)$. This proves the lemma. $\square$

Lemma 6.2 *Let $F = (\mathrm{A}, S, \delta, s_0, S')$ be a finite automaton, $n =_{df} |S| > 0$ and $k \in \mathbb{N}$. For every alternating chain $v_0 \preceq_{\mathcal{F}} \cdots \preceq_{\mathcal{F}} v_k$, there exists an alternating chain $w_0 \preceq_{\mathcal{F}} \cdots \preceq_{\mathcal{F}} w_k$ such that $|w_k| \leq \mathcal{N}(|\mathrm{A}|, n, k)$, $w_k \sim_F v_k$ and $w_k \preceq_{\mathcal{F}} v_k$.*

Proof We will prove the lemma by induction on k. The induction base, i.e. the case $k = 0$, follows from Lemma 6.1, because $\mathcal{M}(|\mathrm{A}|, n) \leq \mathcal{M}(|\mathrm{A}|, n) + \mathcal{K}(n) + 1 = \mathcal{N}(|\mathrm{A}|, n, 0)$. Suppose now that we have already proved the lemma for the case $k = m$. Let $v_0 \preceq_{\mathcal{F}} \cdots \preceq_{\mathcal{F}} v_{m+1}$ be an alternating chain. By induction hypothesis there exists an alternating chain $w_0 \preceq_{\mathcal{F}} \cdots \preceq_{\mathcal{F}} w_m$ such that $|w_m| \leq \mathcal{N}(|\mathrm{A}|, n, m)$, $w_m \sim_F v_m$ and $w_m \preceq_{\mathcal{F}} v_m$. Since $v_m \preceq_{\mathcal{F}} v_{m+1}$, we have $w_m \preceq_{\mathcal{F}} v_{m+1}$. By Lemma 4.4, there exist words $w'_0, \ldots, w'_l, z_1, \ldots, z_l \in \mathcal{A}_{\mathcal{F}}$, $u_1, \ldots, u_l \in \mathrm{A}^+$ such that $w_m = w'_0 \boxed{u_1} w'_1 \cdots \boxed{u_l} w'_l$ and $v_{m+1} = w'_0 \boxed{u_1} z_1 \boxed{u_1} w'_1 \cdots \boxed{u_l} z_l \boxed{u_l} w'_l$. Now we apply Lemma 6.1 to the words z_i. For every z_i with $1 \leq i \leq l$ we obtain a word z'_i such that $|z'_i| < \mathcal{M}(|\mathrm{A}|, n)$, $z'_i \sim_F z_i$ and $z'_i \preceq_{\mathcal{F}} z_i$. Therefore, with $w_{m+1} =_{df} w'_0 \boxed{u_1} z'_1 \boxed{u_1} w'_1 \cdots \boxed{u_l} z'_l \boxed{u_l} w'_l$ the following holds:

- $w_m \preceq_{\mathcal{F}} w_{m+1}$, by Lemma 4.4.
- $w_{m+1} \preceq_{\mathcal{F}} v_{m+1}$, because $z'_i \preceq_{\mathcal{F}} z_i$ for all $1 \leq i \leq l$.
- $w_{m+1} \sim_F v_{m+1}$, because we obtain v_{m+1} from w_{m+1} by replacing the words z'_i with z_i, and $z'_i \sim_F z_i$ for all $1 \leq i \leq l$.
- $|w_{m+1}| \leq \mathcal{N}(|\mathrm{A}|, n, m+1)$, because

$$\begin{aligned} |w_{m+1}| &\leq |w_m| + \overbrace{l\mathcal{K}(n)}^{\text{duplication of the } u_i} + \overbrace{l\mathcal{M}(|\mathrm{A}|, n)}^{\text{length of the } z'_i} \\ &\leq |w_m| \cdot (\mathcal{M}(|\mathrm{A}|, n) + \mathcal{K}(n) + 1) \\ &\leq (\mathcal{M}(|\mathrm{A}|, n) + \mathcal{K}(n) + 1)^{m+2} = \mathcal{N}(|\mathrm{A}|, n, m+1). \end{aligned}$$

Hence $w_0 \preceq_{\mathcal{F}} \cdots \preceq_{\mathcal{F}} w_{m+1}$ is an alternating chain with $|w_{m+1}| \leq \mathcal{N}(|\mathrm{A}|, n, m+1)$, $w_{m+1} \sim_F v_{m+1}$ and $w_{m+1} \preceq_{\mathcal{F}} v_{m+1}$. This completes the induction. $\square$

Lemma 6.3 *$L_{\mathcal{F}}(m)$ is finitely generated for every $m \in \mathbb{N}$ and every finite automaton F. More precisely, it is generated by elements of $\mathcal{A}_{\mathcal{F}}$ of lengths $\leq \mathcal{N}(|\mathrm{A}|, |F|, m)$.*

Proof This Lemma is an easy consequence from Lemma 6.2. $\square$

Lemma 6.4 $\mathrm{L}(F) = \bigcup_{k \in \mathbb{N}} \left(\overline{L_{\mathcal{F}}(2k)} \setminus \overline{L_{\mathcal{F}}(2k+1)} \right)$ *for every finite automaton F.*

Proof "$\subseteq$" Let $w' \in \mathrm{L}(F)$, from Lemma 5.3 we obtain $w' \in \overline{\mathcal{L}_{\mathcal{F}}} \subseteq \overline{L_{\mathcal{F}}(0)}$. Furthermore, let $w \in \mathcal{L}_{\mathcal{F}}$ with $w' = \overline{w}$ and let $w_0 \preceq_{\mathcal{F}} \cdots \preceq_{\mathcal{F}} w_m$ be an alternating chain with $w_m \preceq_{\mathcal{F}} w$. By Lemma 4.5, $m \leq |w| = |w'|$, hence $w' \notin \overline{L_{\mathcal{F}}(|w'|+1)}$. Thus there exists an index n with $0 \leq n \leq |w'|$ such that $w' \in \overline{L_{\mathcal{F}}(n)} \setminus \overline{L_{\mathcal{F}}(n+1)}$. Suppose n is odd, then there exist an alternating chain $v_0 \preceq_{\mathcal{F}} \cdots \preceq_{\mathcal{F}} v_n$ and a $v \in \mathcal{A}_{\mathcal{F}}$ such that $v_n \preceq_{\mathcal{F}} v$ and $w' = \overline{v}$. Since n is odd and $w' \in \mathrm{L}(F)$, we obtain $v_n \notin \mathcal{L}_{\mathcal{F}}$ and $v \in \mathcal{L}_{\mathcal{F}}$. Therefore, $v_0 \preceq_{\mathcal{F}} \cdots \preceq_{\mathcal{F}} v_n \preceq_{\mathcal{F}} v$ is an alternating chain of length $n+1$. It follows that $w' \in \overline{L_{\mathcal{F}}(n+1)}$, a contradiction to our assumption. Thus $w' \in \overline{L_{\mathcal{F}}(2k)} \setminus \overline{L_{\mathcal{F}}(2k+1)}$ for a suitable $k \in \mathbb{N}$ with $1 \leq 2k \leq |w'|$.

"$\supseteq$" Let $k \in \mathbb{N}$ and $w' \in \overline{L_{\mathcal{F}}(2k)} \setminus \overline{L_{\mathcal{F}}(2k+1)}$. Thus there exist a $v \in \mathcal{A}_{\mathcal{F}}$ and an alternating chain $v_0 \preceq_{\mathcal{F}} \cdots \preceq_{\mathcal{F}} v_{2k}$ such that $v_{2k} \preceq_{\mathcal{F}} v$ and $w' = \overline{v}$. Suppose $w' \notin \mathrm{L}(F)$, i.e. $v \notin \mathcal{L}_{\mathcal{F}}$. Hence $v_0 \preceq_{\mathcal{F}} \cdots \preceq_{\mathcal{F}} v_{2k} \preceq_{\mathcal{F}} v$ is an alternating chain of length $2k+1$. It follows that $w' \in \overline{L_{\mathcal{F}}(2k+1)}$, a contradiction. We conclude $w' \in \mathrm{L}(F)$. $\square$

Lemma 6.5 $\overline{L_{\mathcal{F}}(m)} \in \mathcal{B}_{1/2}$ *for every* $m \in \mathbb{N}$ *and every finite automaton* F *with* $\varepsilon \notin \mathrm{L}(F)$.

Proof By Lemma 6.3, $L_{\mathcal{F}}(m)$ is finitely generated, i.e. there exist $n \in \mathbb{N}$ and non-empty words $w_1, \ldots, w_n \in \mathcal{A}_{\mathcal{F}}$ such that

$$L_{\mathcal{F}}(m) = \bigcup_{1 \leq i \leq n} \{w : w_i \preceq_{\mathcal{F}} w\}.$$

It follows that

$$\overline{L_{\mathcal{F}}(m)} = \bigcup_{1 \leq i \leq n} \{\overline{w} : w_i \preceq_{\mathcal{F}} w\}.$$

Since $\mathcal{B}_{1/2}$ is closed under finite unions, it suffices to show $\{\overline{w} : v \preceq_{\mathcal{F}} w\} \in \mathcal{B}_{1/2}$ for every non-empty word $v \in \mathcal{A}_{\mathcal{F}}$. By Lemma 4.4, $v \preceq_{\mathcal{F}} w$ if and only if there exist words $v_0, \ldots v_k, z_1, \ldots, z_k \in \mathcal{A}_{\mathcal{F}}, u_1, \ldots, u_k \in \mathrm{A}^+$ such that $v = v_0 \boxed{u_1} v_1 \cdots \boxed{u_k} v_k$ and $w = v_0 \boxed{u_1} z_1 \boxed{u_1} v_1 \cdots \boxed{u_k} z_k \boxed{u_k} v_k$. Thus we have

$$\{w : v \preceq_{\mathcal{F}} w\} = \bigcup_{\substack{k \in \mathbb{N},\ v_0, \ldots v_k \in \mathcal{A}_{\mathcal{F}} \\ u_1, \ldots, u_k \in \mathrm{A}^+ \text{ with} \\ v = v_0 \boxed{u_1} v_1 \cdots \boxed{u_k} v_k}} v_0 \boxed{u_1} \mathcal{A}_{\mathcal{F}} \boxed{u_1} v_1 \boxed{u_2} \mathcal{A}_{\mathcal{F}} \boxed{u_2} v_2 \cdots \boxed{u_k} \mathcal{A}_{\mathcal{F}} \boxed{u_k} v_k.$$

From Lemma 5.3 it follows that $\overline{\mathcal{A}_{\mathcal{F}}} = \mathrm{A}^*$ and we obtain

$$\{\overline{w} : v \preceq_{\mathcal{F}} w\} = \bigcup_{\substack{k \in \mathbb{N},\ v_0, \ldots v_k \in \mathcal{A}_{\mathcal{F}} \\ u_1, \ldots, u_k \in \mathrm{A}^+ \text{ with} \\ v = v_0 \boxed{u_1} v_1 \cdots \boxed{u_k} v_k}} \overline{v_0} u_1 \mathrm{A}^* u_1 \overline{v_1} u_2 \mathrm{A}^* u_2 \overline{v_2} \cdots u_k \mathrm{A}^* u_k \overline{v_k}.$$

Notice that the number of different representations of v is finite, i.e. the union above is finite. Since v is non-empty, we obtain $\varepsilon \notin \{\overline{w} : v \preceq_{\mathcal{F}} w\}$. Now it is easy to see that $\{\overline{w} : v \preceq_{\mathcal{F}} w\} \in \mathcal{B}_{1/2}$ (replace each A^* with $(\mathrm{A}^+ \cup \{\varepsilon\})$ and apply DeMorgan's laws). $\square$

Lemma 6.6 *For every* $k \in \mathbb{N}^+$ *and every finite automaton* F *with* $\varepsilon \notin \mathrm{L}(F)$ *we have*

$$m_{\mathcal{F}} < k \Longrightarrow \mathrm{L}(F) \in \mathcal{B}_{1/2}(k).$$

Proof This lemma follows from the Lemmas 2.1, 6.4 and 6.5. $\square$

7 Low Levels induce Small Measures

We will prove that an automaton whose language is in Level k of the Boolean hierarchy over $\mathcal{B}_{1/2}$ has measure $< k$. For this we need the following lemma which reproofs one direction of the forbidden pattern for $\mathcal{B}_{1/2}$. This result was originally obtained by J. E. Pin and P. Weil in [PW97].

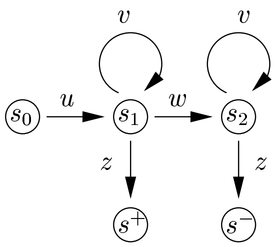

Figure 1: Pattern P with $v \in \mathrm{A}^+$, $s^+ \in F$ and $s^- \notin F$.

Lemma 7.1 *If $L \in \mathcal{B}_{1/2}$ and F is the minimal automaton with $\mathrm{L}(F) = L$, then F does not have Pattern P (cf. Figure 1). Particularly for every $l \in \mathrm{A}^+$, if a state of F has an ll-loop then it even has an l-loop.*

Proof Let $L \in \mathcal{B}_{1/2}$ and let $F = (\mathrm{A}, S, \delta, s_0, S')$ be the minimal finite automaton with $\mathrm{L}(F) = L$. By definition of $\mathcal{B}_{1/2}$, L is of the form

$$L = \bigcup_{i=1}^{k} w_{i,0}\mathrm{A}^+ w_{i,1} \cdots \mathrm{A}^+ w_{i,n_i}$$

with $k, n_i \in \mathbb{N}$, $w_{i,j} \in \mathrm{A}^*$ and $w_{i,0} \neq \varepsilon$ if $n_i = 0$. Suppose that automaton F has Pattern P, then $uv^i z \in L$ and $uv^i wv^j z \notin L$ for all $i, j \in \mathbb{N}$. Particularly $uv^n z \in L$ with $n =_{df} \max\{|w_{i,0}w_{i,1} \cdots w_{i,n_i}| + 4n_i : 1 \leq i \leq k\}$. Hence there exists a j with $uv^n z \in w_{j,0}\mathrm{A}^+ w_{j,1} \cdots \mathrm{A}^+ w_{j,n_j}$, i.e. there exist $r_1, \ldots, r_{n_j} \in \mathrm{A}^+$ such that $uv^n z = w_{j,0} r_1 w_{j,1} \cdots r_{n_j} w_{j,n_j}$. For $1 \leq i \leq n$ we mark the positions in $w_{j,0} r_1 w_{j,1} \cdots r_{n_j} w_{j,n_j}$ where the prefix uv^i ends. Since $|v| \geq 1$ we obtain n different marked positions. In $w_{j,0} r_1 w_{j,1} \cdots r_{n_j} w_{j,n_j}$ there exist at most $|w_{j,0} w_{j,1} \cdots w_{j,n_j}| + n_j$ positions being not located within an r_i. Since $n \geq |w_{j,0} w_{j,1} \cdots w_{j,n_j}| + 4n_j$, we find an l such that there exist 3 markers located in r_l. For suitable $r', r'' \in \mathrm{A}^*$ we have $r_l = r'vvr''$ and we have markers after r', $r'v$ and $r'vv$. Therefore, $w_{j,0} r_1 w_{j,1} \cdots r_{l-1} w_{j,l-1} r'v = uv^h$ for a suitable $h \in \mathbb{N}$. On the one hand, with $r'_l =_{df} r'vwvr''$ we obtain

$$x =_{df} w_{j,0} r_1 w_{j,1} \cdots r_{l-1} w_{j,l-1} r'_l w_{j,l} r_{l+1} w_{j,l+1} \cdots r_{n_j} w_{j,n_j} \in w_{j,0}\mathrm{A}^+ w_{j,1} \cdots \mathrm{A}^+ w_{j,n_j} \subseteq L,$$

because $r'_l \in \mathrm{A}^+$. On the other hand x is of the form $x = uv^h wv^{n-h} z$, thus $x \notin L$, since F has Pattern P. This is a contradiction, it follows that F does not have Pattern P. This shows the first assertion of our lemma.

Now let $s_1 \in S$ and $l \in \mathrm{A}^+$ such that s_1 has an ll-loop. Suppose $s_1 \neq s_2$ with $s_2 =_{df} \delta(s_1, l)$. Notice that both s_1 and s_2 have an ll-loop. Since F is a minimal automaton there exists a $z \in \mathrm{A}^*$ such that $\delta(s_1, z) \in S' \iff \delta(s_2, z) \notin S'$, w.l.o.g. we assume $\delta(s_1, z) \in S'$ and $\delta(s_2, z) \notin S'$. There exists an $u \in \mathrm{A}^*$ with $\delta(s_0, u) = s_1$, because F is a minimal automaton. Thus we have found Pattern P with $w =_{df} l$ and $v =_{df} ll$. By the first assertion of the lemma, this is a contradiction to $L \in \mathcal{B}_{1/2}$. Therefore, $s_1 = s_2$ and s_1 even has an l-loop. □

Lemma 7.2 *For every minimal automaton F and every $k \in \mathbb{N}^+$ we have*

$$\mathrm{L}(F) \in \mathcal{B}_{1/2}(k) \Longrightarrow m_{\mathcal{F}} < k.$$

Proof Let F be a minimal automaton with $\mathrm{L}(F) \in \mathcal{B}_{1/2}(k)$ and $k \in \mathbb{N}^+$. By Lemma 2.1, there exist minimal automata $F_0, \ldots, F_{k-1}$ such that $\mathrm{L}(F_0), \ldots, \mathrm{L}(F_{k-1}) \in \mathcal{B}_{1/2}$, $\mathrm{L}(F_0) \supseteq \cdots \supseteq \mathrm{L}(F_{k-1})$ and

$$\mathrm{L}(F) = \bigcup_{i=0}^{\lceil \frac{k}{2} \rceil - 1} \left(\mathrm{L}(F_{2i}) \setminus \mathrm{L}(F_{2i+1})\right), \tag{6}$$

where F_k is the minimal automaton rejecting all words. Let $t =_{df} \max\{|F_i| : 0 \le i < k\}$. Suppose $m_{\mathcal{F}} \ge k$, i.e. there exists an alternating chain $w_0 \preceq_{\mathcal{F}} \cdots \preceq_{\mathcal{F}} w_k$. Hence there exist a chain of elementary extensions $v_0 \le_{\mathcal{F}} \cdots \le_{\mathcal{F}} v_m$ and indices $0 \le i_0 < \cdots < i_{k-1} < m$ such that $v_0 \in \mathcal{L}_{\mathcal{F}}$ and $v_{i_j} \in \mathcal{L}_{\mathcal{F}} \iff v_{i_j+1} \notin \mathcal{L}_{\mathcal{F}}$ for $0 \le j < k$. Since $v_i \le_{\mathcal{F}} v_{i+1}$ for $0 \le i < m$, each v_i can be represented as $v_i = v_i' \boxed{u_i} v_i''$ such that $v_{i+1} = v_i' \boxed{u_i} z_i \boxed{u_i} v_i''$ with $v_i', v_i'', z_i \in \mathcal{A}_{\mathcal{F}}$, $u_i \in \mathrm{A}^+$ and $0 \le i < m$. We obtain:

$$
\begin{aligned}
v_0 &= v_0' \boxed{u_0} v_0'' \\
v_1 &= v_0' \boxed{u_0} z_0 \boxed{u_0} v_0'' = v_1' \boxed{u_1} v_1'' \\
v_2 &= v_1' \boxed{u_1} z_1 \boxed{u_1} v_1'' = v_2' \boxed{u_2} v_2'' \\
&\vdots \\
v_{m-1} &= v_{m-2}' \boxed{u_{m-2}} z_{m-2} \boxed{u_{m-2}} v_{m-2}'' = v_{m-1}' \boxed{u_{m-1}} v_{m-1}'' \\
v_m &= v_{m-1}' \boxed{u_{m-1}} z_{m-1} \boxed{u_{m-1}} v_{m-1}''
\end{aligned}
$$

In the following we will replace every $\boxed{u}$ with u^t in the chain $v_0 \le_{\mathcal{F}} \cdots \le_{\mathcal{F}} v_m$. More precisely, for every $v \in \mathcal{A}_{\mathcal{F}}$ with $v = w_0' \boxed{u_1'} w_1' \cdots \boxed{u_{k'}'} w_{k'}'$, $w_j' \in \mathrm{A}^*$ and $u_j' \in \mathrm{A}^+$, we define $\widehat{v} =_{df} w_0' u_1'^{\,t} w_1' \cdots u_{k'}'^{\,t} w_{k'}'$. Notice that $v \in \mathcal{L}_{\mathcal{F}} \iff \widehat{v} \in \mathrm{L}(F)$, because $u_j' \sim_F u_j' u_j'$ for $1 \le j \le k'$ by the definition of $\mathcal{A}_{\mathcal{F}}$. Let $a_i =_{df} \widehat{v_i}$, $b_j' =_{df} \widehat{v_j'}$, $b_j'' =_{df} \widehat{v_j''}$ and $c_j =_{df} \widehat{z_j}$ for $0 \le i \le m$ and $0 \le j < m$. With $a_0, \ldots, a_m$ we have obtained a sequence of elements of A^* with $a_i = b_i' u_i^t b_i''$ and $a_{i+1} = b_i' u_i^t c_i u_i^t b_i''$ for $0 \le i < m$. Furthermore, $a_i \in \mathrm{L}(F) \iff v_i \in \mathcal{L}_{\mathcal{F}}$ for $0 \le i \le m$. It follows that

$$
a_0 \in \mathrm{L}(F) \qquad \text{and} \qquad a_{i_j} \in \mathrm{L}(F) \iff a_{i_j+1} \notin \mathrm{L}(F) \;\text{ for } 0 \le j < k \tag{7}
$$

Now we want to show $a_i \in \mathrm{L}(F_j) \Longrightarrow a_{i+1} \in \mathrm{L}(F_j)$ for every $0 \le j < k$ and $0 \le i < m$. Let $a_i \in \mathrm{L}(F_j)$ and $a_{i+1} \notin \mathrm{L}(F_j)$ with $0 \le j < k$ and $0 \le i < m$. Since $t \ge |F_i|$, the word $b_i' u_i^t$ leads to an $u_i^{t'}$-loop in F_i for a suitable $t' \le t$. By Lemma 7.1, this is even an u_i-loop. The same argument shows that $b_i' u_i^t c_i u_i^t$ leads to an u_i-loop in F_i. With $u =_{df} b_i' u_i^t$, $v =_{df} u_i$, $z =_{df} b_i''$ and $w =_{df} c_i u_i^t$ we have found Pattern P. By Lemma 7.1 this is a contradiction to $\mathrm{L}(F_i) \in \mathcal{B}_{1/2}$. It follows that:

$$
a_i \in \mathrm{L}(F_j) \Longrightarrow a_{i+1} \in \mathrm{L}(F_j) \;\text{ for every } 0 \le j < k \text{ and } 0 \le i < m \tag{8}
$$

From (6), (7) and (8) it follows that $a_0 \in \mathrm{L}(F_0)$ and $a_{i_j} \in \mathrm{L}(F_j)$ for $0 \le j < k$. Thus we have $a_{i_{k-1}} \in \mathrm{L}(F_{k-1})$. By Equations (7) and (8), $a_{i_{k-1}+1}$ ought to be an element of $\mathrm{L}(F_k)$. This is a contradiction, because $\mathrm{L}(F_k) = \emptyset$. Hence our assumption was false and we conclude $m_{\mathcal{F}} < k$. □

8 Decidability of the Boolean Hierarchy over Dot-Depth 1/2

Here we combine the results obtained in the previous sections. The main result is the levelwise characterization of the Boolean hierarchy over dot-depth 1/2 by the measure $m_{\mathcal{F}}$. From the decidability of $m_{\mathcal{F}}$ it follows that the Boolean hierarchie over $\mathcal{B}_{1/2}$ is proper and decidable. Furthermore, we obtain a reproof for the forbidden pattern for $\mathcal{B}_{1/2}$ which was originally proved in [PW97].

Theorem 8.1 *For every $k \in \mathbb{N}^+$ and every minimal automaton F with $\varepsilon \notin \mathrm{L}(F)$ we have*

$$
\mathrm{L}(F) \in \mathcal{B}_{1/2}(k) \iff m_{\mathcal{F}} < k.
$$

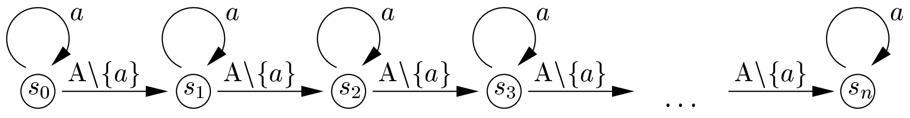

Figure 2: Automaton F_n with $n \geq 0$, $a \in \mathrm{A}$ and s_i is accepting if and only if i is even.

Proof This follows from the Lemmas 6.6 and 7.2. $\square$

Corollary 8.2 *On input of an alphabet* A*, a finite automaton* F' *and a* $k \in \mathbb{N}^+$*, it is decidable whether* $\mathrm{L}(F') \in \mathcal{B}_{1/2}(k)$*.*

Proof First of all we minimize automaton F', we obtain the minimal automaton F. If $\varepsilon \in \mathrm{L}(F)$ then $\mathrm{L}(F) \notin \mathcal{B}_{1/2}(k)$. Otherwise, by Theorem 8.1 it suffices to decide whether $m_{\mathcal{F}} < k$. By Lemma 6.2, $m_{\mathcal{F}} \geq k$ implies the existence of an alternating chain $w_0 \preceq_{\mathcal{F}} \cdots \preceq_{\mathcal{F}} w_k$ such that $|w_k| \leq \mathcal{N}(|\mathrm{A}|, |F|, k)$. Therefore, by testing (the finite number of) all possible chains $v_0 \preceq_{\mathcal{F}} \cdots \preceq_{\mathcal{F}} v_k$ with $|v_k| \leq \mathcal{N}(|\mathrm{A}|, |F|, k)$, we can decide whether $m_{\mathcal{F}} \geq k$. $\square$

Corollary 8.3 *The Boolean hierarchy over* $\mathcal{B}_{1/2}$ *is proper.*

Proof Let $a, b \in \mathrm{A}$ with $a \neq b$, and let $F = F_n$ be the automaton defined in Figure 2 for a fixed $n \geq 0$. We want to show $m_{\mathcal{F}} = n$. For $0 \leq i \leq n$ we define

$$w_i =_{df} \boxed{a}\ \overbrace{b\boxed{a} \cdots b\boxed{a}}^{i \text{ times } b\boxed{a}}.$$

Since every state has an a-loop, we have $w_i \in \mathcal{A}_{\mathcal{F}}$ for $0 \leq i \leq n$. Furthermore, it is easy to see that $w_i \leq_{\mathcal{F}} w_{i+1}$ and $w_j \in \mathcal{L}_{\mathcal{F}} \iff j \equiv 0(2)$ for all $0 \leq i < n$ and $0 \leq j \leq n$. Therefore, $m_{\mathcal{F}} \geq n$.

Suppose $m_{\mathcal{F}} > n$, i.e. there exists an alternating chain $v_0 \preceq_{\mathcal{F}} v_1 \preceq_{\mathcal{F}} \cdots \preceq_{\mathcal{F}} v_m$ with $m > n$. For $0 \leq i \leq m$ we define the number of letters in v_i which are different from a:

$$\|v_i\| =_{df} |v_i| - \text{number of letters } a \text{ in } v_i$$

Since every state has an a-loop, the insertion of letters a does not lead to an alternation. Thus every alternation is induced by the insertion of at least one letter being different from a. It follows that $\|v_i\| < \|v_{i+1}\|$ and $j \leq \|v_j\|$ for $0 \leq i < m$ and $0 \leq j \leq m$. Observe that the automaton stops in state s_n after reading an arbitrary word $v \in \mathcal{A}_{\mathcal{F}}$ with $\|v\| \geq n$. Since $n \leq m-1 \leq \|v_{m-1}\| < \|v_m\|$, we have $v_{m-1} \in \mathcal{L}_{\mathcal{F}} \iff v_m \in \mathcal{L}_{\mathcal{F}}$. This is a contradiction, and we obtain $m_{\mathcal{F}} = n$. From Theorem 8.1 we conclude that the Boolean hierarchy over $\mathcal{B}_{1/2}$ is proper. $\square$

The following corollary is due to J. E. Pin and P. Weil [PW97]. Here we give an automata theoretic reproof of this result.

Corollary 8.4 *A language of* A^+ *is in* $\mathcal{B}_{1/2}$ *if and only if its minimal automaton does not have Pattern* P *(cf. Figure 1).*

Proof By Lemma 7.1, it suffices to show that automaton $F = (\mathrm{A}, S, \delta, s_0, S')$ has Pattern P, if F is a minimal automaton with $\mathrm{L}(F) \subseteq \mathrm{A}^+$ and $\mathrm{L}(F) \notin \mathcal{B}_{1/2}$. In this case, from Theorem 8.1 we obtain $m_{\mathcal{F}} \geq 1$. Hence there exists an alternating chain of length ≥ 1. Therefore, there exist $x, y \in \mathcal{A}_{\mathcal{F}}$ with $x \in \mathcal{L}_{\mathcal{F}}$, $y \notin \mathcal{L}_{\mathcal{F}}$ and $x \leq_{\mathcal{F}} y$. Thus we can choose suitable $u', w', z \in \mathcal{A}_{\mathcal{F}}$, $v \in \mathrm{A}^+$ such that $x = u'\boxed{v}z$ and $y = u'\boxed{v}w'\boxed{v}z$. By the definition of $\mathcal{A}_{\mathcal{F}}$, we have $\delta(s, v) = \delta(s, vv)$

for every $s \in S$. It follows that the automaton reaches a v-loop after reading the words $u'v$ and $u'vw'v$. With $u =_{df} u'v$ and $w =_{df} w'v$ we have found Pattern P. $\square$

Acknowledgements. I am grateful to Heinz Schmitz, University of Würzburg, for many enjoyable discussions and helpful comments.

References

[Arf87] M. Arfi. Polynomial operations on rational languages. In *Proceedings 4th Symposium on Theoretical Aspects of Computer Science*, volume 247 of *Lecture Notes in Computer Science*, pages 198-206, Springer-Verlag, 1987.

[Arf91] M. Arfi. Opérations polynomials et hiérarchies de concaténation. *Theoretical Computer Science*, 91:71-84, 1991.

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

[BK78] J. A. Brzozowski and R. Knast. The dot-depth hierarchy of star-free languages is infinite. *Journal of Computer and System Science*, 16:37-55, 1978.

[CB71] R. S. Cohen and J. A. Brzozowski. Dot-depth of star-free events. *Journal of Computer and System Science*, 5:1-16, 1971.

[CGH+88] J.-Y. Cai, T. Gundermann, J. Hartmanis, L. A. Hemachandra, V. Sewelson, K. W. Wagner, and G. Wechsung. The Boolean hierarchy I: Structural properties. *SIAM Journal on Computing*, 17:1232-1252, 1988.

[Gla98] C. Glaßer. A Normalformen for Classes of Concatenation Hierarchies. Technical Report 216, Institut für Informatik, Universität Würzburg, 1998.

[Hig52] G. Higman. Ordering by divisibility in abstract algebras. *Proc. London Math. Soc.*, 3:326-336, 1952.

[Kna83] R. Knast. A semigroup characterization of dot-depth one languages. *RAIRO Inform. Theor.*, 17:321-330, 1983.

[KSW87] J. Köbler, U. Schöning, and K. W. Wagner. The difference and the truth-table hierarchies for NP. *RAIRO Inform. Theor.*, 21:419-435, 1987.

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

[Ner58] A. Nerode. Linear automata transformation. In *Proceedings of AMS*, volume 9, pages 541-544, 1958.

[Pin95] J. E. Pin. A variety theorem without complementation. *Russian Math.*, 39:74-83, 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.

[PP86] D. Perrin, J. E. Pin. First-order logic and star-free sets. *Journal of Computer and System Science*, 32:393-406, 1986.

[PW96] J. E. Pin and P. Weil. A reiterman theorem for pseudovarieties of finite first-order structures. *Algebra Universalis*, 35:577-595, 1996.

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

[Sch98] H. Schmitz. Relationships between some Hierarchies - An Introduction to the Dot-Depth-Problem. unpublished script, Institut für Informatik, Universität Würzburg, 1998.

[Sch65] M. P. Schützenberger. On finite monoids having only trivial subgroups. *Information and Control*, 8:190-194, 1965.

[SS83] J. Sakarovitch and I. Simon. Subwords. In M. Lothaire, editor, *Combinatorics on Words*, pages 105-142, Addison-Wesley, 1983.

[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.

[Str81] H. Straubing. A generalization of the Schützenberger product of finite monoids. *Theoretical Computer Science*, 13:137-150, 1981.

[Str85] H. Straubing. Finite semigroups varieties of the form V * D. *J.Pure Appl.Algebra*, 36:53-94, 1985.

[SW98] H. Schmitz and K. W. Wagner. The Boolean Hierarchy over Level 1/2 of the Straubing-Thérien-Hierarchy. Technical Report 201, Institut für Informatik, Universität Würzburg, 1998.

[Thé81] D. Thérien. Classification of finite monoids: the language approach. *Theoretical Computer Science*, 14:195-208, 1981.

[Tho82] W. Thomas. Classifying regular events in symbolic logic. *Journal of Computer and System Science*, 25:360-376, 1982.

[Tho96] W. Thomas. Languages, automata, and logic. Technical Report 9607, Institut für Informatik und praktische Mathematik, Universität Kiel, 1996.