Theoretical Computer Science 8 (1979) 161–175.
© North-Holland Publishing Company

LANGUAGE OPERATORS RELATED TO INIT

Helmut PRODINGER and Friedrich J. URBANEK

Institut für Mathematische Logik und Formale Sprachen, Technische Universität Wien, Austria

Communicated by Maurice Nivat
Received November 1977
Revised March 1978

Abstract. The language operator $\text{Init}(L) = \{x : \text{there is a } y, \text{ such that } xy \in L\}$ is well-known. Replacing "there is a $y, \ldots$" by "there are infinitely many $y, \ldots$" a new operator $\text{Anf}(L) = \{x : \text{there are infinitely many } y, \text{ such that } xy \in L\}$ is obtained.

Furthermore, there are language operators, for which definitions structurally similar to those of Init and Anf can be given.

0. Introduction

The language operator $\text{Init}(L) = \{x\colon \text{there is a } y, \text{ such that } xy \in L\}$ is well-known. Replacing "there is a $y, \ldots$" by "there are infinitely many $y, \ldots$" a new operator $\text{Anf}(L) = \{x\colon \text{there are infinitely many } y, \text{ such that } xy \in L\}$ is obtained.

There are common properties of Init and Anf, caused by the fact, that Init as well as Anf pick out initial subwords of words in L, discussed in Chapter 2.

On the other hand the behaviour of Anf is quite different from that of Init, e.g. the recursive (recursively enumerable) languages are not closed under Anf (Chapter 3).

In Chapter 4, the definition of Init and Anf is generalized, and the class of the resulting operators is examined.

1. Basic Definitions

An *alphabet* Σ is a finite nonempty set of *symbols*. Σ^* denotes the set of all words over Σ, $\Sigma^+ = \Sigma^* - \{\varepsilon\}$, where ε denotes the empty word. For $x \in \Sigma^*$ $|x|$ is the *length* of x and x^R is the mirror image of x ($\varepsilon^R = \varepsilon$, $a_1 \cdots a_n^R = a_n \cdots a_1$). If $a \in \Sigma$, then $n_a(x)$ is the number of symbols a in x.

A *grammar* is a quadruple $G = (\Phi, \Sigma, P, S)$, where Φ is the alphabet of the *nonterminals*, Σ is the alphabet of the terminals, P is a finite set of *rules* $\alpha \to \beta$,

$\alpha \in (\Phi \cup \Sigma)^+$, $\beta \in (\Phi \cup \Sigma)^*$ and $S \in \Phi$ is the *start symbol*. Derivations in G are defined in the usual way and $L(G)$ denotes the language generated by G.

If for all $\alpha \to \beta \in P$ $|\alpha| \leq |\beta|$, then G and $L(G)$ are called *context-sensitive*. If for all $\alpha \to \beta \in P$ $\alpha \in \Phi$ and $\beta \in (\Phi \cup \Sigma)^*$, then G and $L(G)$ are called *contextfree*. If furthermore $\beta \in \Sigma\Phi^*$, then G is called a *Greibach-normalform-grammar*.

A *finite automaton* is a quintuple $M = (K, \Sigma, \delta, q_0, F)$, where K is a finite nonempty set of *states*, Σ is the *input alphabet*, δ maps $K \times \Sigma$ into K, $q_0 \in K$ is the *initial state* and $F \subseteq K$ is the set of *final states*. The language accepted by M is denoted by $T(M)$.

In this paper the following model of a Turing-machine is used: A *Turing-machine* is a sextuple $T = (K, \Sigma, \Gamma, \delta, q_0, F)$, where K is the finite set of *states*, Σ is the *input alphabet*, Γ is the alphabet of *tape symbols*, containing b (the *blank symbol*). It is assumed that $\Sigma \subseteq \Gamma - \{b\}$. δ is a partial mapping from $K \times \Gamma$ to $K \times \Gamma \times \{L, R\}$, $q_0 \in K$ is the *initial state*, $F \subseteq K$ is the set of *final states*.

A *configuration of* T is word in $\Gamma^* K \Gamma^*$. The relation $\vdash$ on the set of all configurations is defined as follows: $\alpha \vdash \beta$ iff

(i) $\alpha = uqCv$, $\beta = uDpv$ and $\delta(q, C) = (p, D, R)$ or
(ii) $\alpha = uEqCv$, $\beta = upEDv$ and $\delta(q, C) = (p, D, L)$ or
(iii) $\alpha = uq$, $\beta = uDp$ and $\delta(q, b) = (p, D, R)$ or
(iv) $\alpha = uEq$, $\beta = upED$ and $\delta(q, b) = (p, D, L)$ or
(v) $\alpha = qCv$, $\beta = pbDv$ and $\delta(q, C) = (p, D, L)$ or
(vi) $\alpha = q$, $\beta = pbD$ and $\delta(q, b) = (p, D, L)$, where $u, v \in \Gamma^*$; $q, p \in K$; $C, D, E \in \Gamma$.

$\overset{*}{\vdash}$ is the reflexive and transitive closure of $\vdash$. The language accepted by T is $L(T) = \{w \in \Sigma^*$: there are $u, v \in \Gamma^*$, $q \in F$, such that $q_0 w \overset{*}{\vdash} uqv\}$.

If $L \subseteq \Sigma^*$, then $\mathrm{Init}(L) = \{x$: there is a y, such that $xy \in L\}$.

If $L_1, L_2 \subseteq \Sigma^*$, then $L_1 \backslash L_2 = \{x$: there is a $y \in L_1$ such that $yx \in L_2\}$.

For each $L \subseteq \Sigma^*$ the binary relation $\underset{L}{\sim}$ on Σ^* is defined by $x \underset{L}{\sim} y$, iff for all $z \in \Sigma^*$ $xz \in L$ iff $yz \in L$.

If there is no danger of confusion a language $\{w\}$ is denoted by w.

All results of the theory of formal languages, which are used in this paper, can be found in [1].

2. The operator Anf and some of its Properties

Definition 2.1. For each language L let $\mathrm{Anf}(L) = \{x$: there are infinitely many y, such that $xy \in L\}$.

(Anf is an abbreviation of the German word Anfang (i.e. begin).)

As $\mathrm{Init}(L) = \{x$: there exists a y such that $xy \in L\}$, $\mathrm{Anf}(L) \subseteq \mathrm{Init}(L)$ always holds.

First a property of Anf is established, which is later used:

Theorem 2.2. *L is infinite if and only if* Anf(L) *is infinite. L is finite if and only if* Anf(L) $= \emptyset$.

Proof. Let $L \subseteq \Sigma^*$ be infinite and $n \geq 0$ be fixed. Let $L' = L \cap \Sigma^n \Sigma^*$. Then L' is infinite. Let $\psi : L' \to \Sigma^n$ be defined by $\psi(xy) = x$ if $|x| = n$. Then ψ maps an infinite set into a finite set. By the pigeon hole principle there is a $x \in \Sigma^n$, such that $\psi^{-1}(x)$ is infinite, that means, that there are infinitely many y, such that $xy \in L'$ $(\subseteq L)$ and therefore $x \in$ Anf(L).

Therefore, for each $n \geq 0$ Anf(L)$\cap \Sigma^n \neq \emptyset$ and hence Anf(L) is infinite.

Let conversely Anf(L) be infinite and $x \in$ Anf(L). Then for infinitely many y is $xy \in L$ and therefore L is infinite.

As for a finite L Anf(L) $= \emptyset$, the second statement of the theorem is obvious.

Many language operators T are defined in the following manner: First $T(w)$ is defined for all $w \in \Sigma^*$, and then T is extended to $\mathfrak{P}(\Sigma^*)$ by $T(L) = \bigcup_{w \in L} T(w)$. The following corollary shows that Anf does not belong to this sort of operators.

Corollary 2.3. *There is no word operator T, such that for all L* Anf(L) $= \bigcup_{w \in L} T(w)$.

Proof. If there would be such an operator T, then for all w Anf($\{w\}$) $= T(w) = \emptyset$ would hold. If L is infinite, then Anf(L) is infinite but $\bigcup_{w \in L} T(w) = \emptyset$, which is a contradiction.

Now some examples are given.

Example 2.4. Let $L \subseteq a^*$ be infinite. Then Anf(L) $= a^*$. (By Theorem 2.2 for each $n \geq 0$ Anf(L)$\cap a^n \neq \emptyset$.)

Example 2.5. Let $L = \{a^n b^n : n \geq 0\}$. Then Anf($L$) $= a^*$. (The word a^i can be completed to a word in L by concatenation with each of the infinitely many words $a^n b^{n+i}$, $n \geq 0$. Conversely, a word $a^i b^j$ where $0 < j \leq i$ can be completed to a word in L only by concatenation with b^{i-j}.)

Example 2.6. Let $L = \{w \in \{a, b\}^* : n_a(w) = n_b(w)\}$. Then for each $w \in \{a, b\}^*$ $wa^{n_b(w)} b^{n_a(w)} (ab)^* \subseteq L$ and therefore Anf(L) $= \{a, b\}^*$.

Example 2.7. Let $L \subseteq \Sigma^*$ an arbitrary language and $c \notin \Sigma$. Then Anf(Lc^*) $=$ Init(Lc^*). (If $x \in$ Init(Lc^*), then there is a y, such that $xy \in Lc^*$. Then $xyc^* \subseteq Lc^*$ and therefore $x \in$ Anf(Lc^*).)

Now some properties of Anf, concerning $\cup, \cap, \cdot, {}^*$, Init are given:

Theorem 2.8. Anf($L_1 \cup L_2$) $=$ Anf(L_1)$\cup$ Anf(L_2).

Proof. Let $w \in \text{Anf}(L_1 \cup L_2)$. Then there are infinitely many y, such that $wy \in L_1 \cup L_2$. Then there are infinitely many z, such that (without loss of generality) $wz \in L_1$, hence $w \in \text{Anf}(L_1) \subseteq \text{Anf}(L_1) \cup \text{Anf}(L_2)$.

Let conversely (without loss of generality) be $w \in \text{Anf}(L_1)$. Then there are infinitely many y, such that $wy \in L_1 \subseteq L_1 \cup L_2$, hence $w \in \text{Anf}(L_1 \cup L_2)$.

Theorem 2.9. $\text{Anf}(L_1 \cap L_2) \subseteq \text{Anf}(L_1) \cap \text{Anf}(L_2)$.

Proof. Let $w \in \text{Anf}(L_1 \cap L_2)$. Then there are infinitely many y, such that $wy \in L_1 \cap L_2$. As $L_1 \cap L_2 \subseteq L_i$, $w \in \text{Anf}(L_i)$ holds for $i = 1, 2$.

The following example shows, that equality as well as inequality can hold in Theorem 2.9:

Example 2.10. $\text{Anf}((a^2)^* \cap (a^2)^*a) = \emptyset$, but $\text{Anf}((a^2)^*) \cap \text{Anf}((a^2)^*a) = a^*$. If $L_1 \Delta L_2$ $(= (L_1 - L_2) \cup (L_2 - L_1))$ is finite, then $\text{Anf}(L_1 \cap L_2) = \text{Anf}(L_1) \cap \text{Anf}(L_2)$.

A relation analogous to $\text{Init}(L_1L_2) = \text{Init}(L_1) \cup L_1\text{Init}(L_2)$ $(L_2 \neq \emptyset)$ shows the following theorem:

Theorem 2.11. *Let $L_2 \neq \emptyset$. Then*

$$\text{Anf}(L_1L_2) = \begin{cases} \text{Init}(L_1) \cup L_1\text{Anf}(L_2) & \text{if } L_2 \text{ is infinite} \\ \text{Anf}(L_1) & \text{if } L_2 \text{ is finite.} \end{cases}$$

Proof. First let L_2 be infinite. Let $w \in \text{Anf}(L_1L_2)$. If $w \in \text{Init}(L_1)$, then $w \in \text{Init}(L_1) \cup L_1\text{Anf}(L_2)$. If $w \notin \text{Init}(L_1)$, then there are $y_1, y_2, \ldots$, such that $wy_i \in L_1L_2$ $(i = 1, 2, \ldots)$. Hence $wy_i = u_iv_i$, $u_i \in L_1$, $v_i \in L_2$. As $w \notin \text{Init}(L_1)$, $w = u_iu_i'$, $u_i' \neq \varepsilon$, $i = 1, 2, \ldots$. By the pigeon hole principle there are a u_{i_0} $(w = u_{i_0}u_{i_0}'$, $u_{i_0} \in L_1)$ and infinitely many indices $i_1, i_2, \ldots$, such that $u_{i_0}v_{i_k} = wy_{i_k} = u_{i_0}u_{i_0}'y_{i_k}$. As $u_{i_0}'y_{i_k} \in L_2$, $u_{i_0}' \in \text{Anf}(L_2)$ and therefore $w = u_{i_0}u_{i_0}' \in L_1\text{Anf}(L_2)$.

Let conversely be $w \in \text{Init}(L_1) \cup L_1\,\text{Anf}(L_2)$. If $w \in \text{Init}(L_1)$, then there is a z, such that $wz \in L_1$ and as zL_2 is infinite and $wzL_2 \subseteq L_1L_2$, $w \in \text{Anf}(L_1L_2)$ holds.

If $w \in L_1\text{Anf}(L_2)$, then $w = xy$, $x \in L_1$, $y \in \text{Anf}(L_2)$. There are infinitely many z, such that $yz \in L_2$ and therefore $wz = xyz \in L_1L_2$. Hence $w \in \text{Anf}(L_1L_2)$.

Now let $L_2 = \{v_1, \ldots, v_n\}$ $(n \geq 1)$ be finite and $K = \max_{1 \leq i \leq n} |v_i|$. Let $w \in \text{Anf}(L_1L_2)$. Then there are infinitely many y (without loss of generality $|y| \geq k$), such that $wy \in L_1L_2$. As L_2 is finite, there is a $v \in L_2$ (pigeon hole principle), such that infinitely many y can be written as $y = zv$, where $wz \in L_1$. Hence, there are infinitely many z, such that $wz \in L_1$, therefore $w \in \text{Anf}(L_1)$.

If conversely $w \in \text{Anf}(L_1)$, then there are infinitely many y, such that $wy \in L_1$. Then $wyv \in L_1L_2$ always holds, and therefore $w \in \text{Anf}(L_1L_2)$.

Theorem 2.11 does not hold for $L_2=\emptyset$ ($\text{Anf}(\Sigma^*\emptyset)=\text{Anf}(\emptyset)=\emptyset\neq\Sigma^*=\text{Anf}(\Sigma^*)$).

Theorem 2.12. *Let* $L\neq\emptyset$ *and* $L\neq\{\varepsilon\}$. *Then* $\text{Anf}(L^*)=\text{Init}(L^*)$.

Proof. It is only to show that $\text{Init}(L^*)\subseteq\text{Anf}(L^*)$. Let $w\in\text{Init}(L^*)$, then there is a x, such that $wx\in L^*$. Let $z\in L$ and $z\neq\varepsilon$. Then for all i $wxz^i\in L^*$ and therefore $w\in\text{Anf}(L^*)$.

Corollary 2.13. *Let L be infinite. Then* $Anf(L^*)\supseteq L^*\text{Anf}(L)$.

Proof. By Theorem 2.11 and Theorem 2.12 $\text{Anf}(L^*)=\text{Anf}(\varepsilon\cup L^*L)=\text{Anf}(\varepsilon)\cup\text{Anf}(L^*L)=\text{Init}(L^*)\cup L^*\text{Anf}(L)=\text{Anf}(L^*)\cup L^*\text{Anf}(L)$.

A further relation between Anf and Init, as well as the idempotence law for Anf shows the following theorem:

Theorem 2.14. $\text{Anf}(L)=\text{Init}(\text{Anf}(L))=\text{Anf}(\text{Init}(L))=\text{Anf}(\text{Anf}(L))$.

Proof. (i) As $L'\subseteq\text{Init}(L')$ for arbitrary L', $\text{Anf}(L)\subseteq\text{Init}(\text{Anf}(L))$ holds.

(ii) Let $w\in\text{Init}(\text{Anf}(L))$, then there is a x, such that $wx\in\text{Anf}(L)$. Then there are infinitely many y, such that $wxy\in L(\subseteq\text{Init}(L))$, therefore $w\in\text{Anf}(\text{Init}(L))$. Hence $\text{Init}(\text{Anf}(L))\subseteq\text{Anf}(\text{Init}(L))$ holds.

(iii) Let $w\in\text{Anf}(\text{Init}(L))$ Then there are infinitely many y, such that $wy\in\text{Init}(L)$. Therefore there are infinitely many pairs y, z, such that $wyz\in L$. Hence $w\in\text{Anf}(L)$. Let $M=\{y: wy\in L\}$. Then M and also $\text{Anf}(M)$ are infinite (Theorem 2.2). For each $y\in\text{Anf}(M)$ there are infinitely many z, such that $yz\in M$ and therefore $wyz\in L$. Hence $wy\in\text{Anf}(L)$ for infinitely many y and so $w\in\text{Anf}(\text{Anf}(L))$. Hence $\text{Anf}(\text{Init}(L))\subseteq\text{Anf}(\text{Anf}(L))$ holds.

(iv) Let $w\in\text{Anf}(\text{Anf}(L))$. Then there are infinitely many pairs y, z, such that $wyz\in L$ and therefore is $w\in\text{Anf}(L)$. Hence $\text{Anf}(\text{Anf}(L))\subseteq\text{Anf}(L)$, which completes the proof.

Finally those languages are characterized, which can occur as images under Anf.

Theorem 2.15. *For given L there exists a L' such that* $L=\text{Anf}(L')$ *if and only if* (i) *and* (ii) *hold*:

(i) $\text{Init}(L)\subseteq L$,

(ii) *For each* $w\in L$ *there is a* $x\neq\varepsilon$, *such that* $wx\in L$.

Proof. First L is assumed to have properties (i) and (ii), and $L=\text{Anf}(L)$ is shown:

Let $w\in L$, then, by (ii), there is a $x_1\neq\varepsilon$, such that $wx_1\in L$, then again by (ii), there is a $x_2\neq\varepsilon$, such that $wx_1x_2\in L$ and so on. Therefore $\{wx_1, wx_1x_2, wx_1x_2x_3, \ldots\}\subseteq L$ and so $w\in\text{Anf}(L)$.

Let conversely be $w \in \text{Anf}(L)$. Then there are infinitely many y, such that $wy \in L$. Then $w \in \text{Init}(L)$ and by (i), $w \in L$.

Now, $L = \text{Anf}(L')$ is assumed. As $\text{Init}(L) = \text{Init}(\text{Anf}(L')) = \text{Anf}(L') = L$, (i) holds for L.

Let $w \in L$. As $L = \text{Anf}(L') = \text{Anf}(\text{Anf}(L'))$, there are infinitely many y, such that $wy \in \text{Anf}(L')$. Therefore there is at least one $y \neq \varepsilon$, such that $wy \in \text{Anf}(L') = L$ and so (ii) holds.

3. Families of languages and Anf

First, it will be shown, that the family of the regular languages and the family of the contextfree languages are closed under Anf; next, that the family of the contextsensitive languages and the family of the recursively enumerable languages are not closed under Anf.

Theorem 3.1. *If L is regular, then* Anf(L) *is regular.*

Proof. Let $M = (K, \Sigma, \delta, q_0, F)$ be a finite automaton, such that $T(M) = L$. Let $M_1 = (K, \Sigma, \delta, q_0, F_1)$ where $F_1 = \{q \in K$: there are infinitely many $x \in \Sigma^*$, such that $\delta(q, x) \in F\}$.

Then $w \in T(M_1)$ holds if and only if $\delta(q_0, w) \in F_1$. This is equivalent with the fact, that there are infinitely many x, such that $\delta(\delta(q_0, w), x) \in F$, i.e. $wx \in T(M) = L$, and that means $w \in \text{Anf}(L)$.

From this it follows, that $T(M_1) = \text{Anf}(L)$, and therefore $\text{Anf}(L)$ is regular.

As the question, whether a given finite automaton accepts infinitely many words, is decidable, for each M a finite automaton M_1, such that $T(M_1) = \text{Anf}(T(M))$, can be effectively constructed.

Corollary 3.2. *Let $L \subseteq \Sigma^*$ be a regular language and M a finite automaton, which accepts L.*

Then $\text{Anf}(L) = \Sigma^*$ *holds if and only if M contains no trap. (A trap is a state q, such that there is no $x \in \Sigma^*$, for which $\delta(q, x) \in F$.)*

Proof. For M let be M_1 constructed as in the proof of Theorem 3.1.

If there would be a trap q in M, then for all x $\delta(q, x) \notin F$ would hold, and therefore $q \notin F_1$. Let w be a word such that $\delta(q_0, w) = q$. Then $w \notin T(M_1) = \text{Anf}(L)$ and hence $\text{Anf}(L) \neq \Sigma^*$ holds.

Conversely it is assumed, that $\text{Anf}(L) \neq \Sigma^*$. Then there exists a word w, such that $w \notin \text{Anf}(L)$, i.e. $\delta(q_0, w) \notin F_1$. Therefore there are only finitely many x, such that

$\delta(\delta(q_0, w), x) \in F$. Let y be a word which is longer than the longest of these words x. Then $\delta(\delta(q_0, wy), z) \notin F$ for all z and therefore $\delta(q_0, wy)$ is a trap.

Theorem 3.3. *If L is contextfree, then* Anf(L) is contextfree.

Proof. Let be $G = (\Phi, \Sigma, P, S)$ a contextfree grammar in Greibach-normalform, such that $L(G) = L$.

Let $L' = \{\alpha \in (\Phi \cup \Sigma)^*: S \overset{*}{\underset{L}{\Rightarrow}} \alpha\}$. Then L' is contextfree. Furthermore let be $\Phi' = \{A \in \Phi$: there exist infinitely many words x, such that $A \overset{*}{\underset{L}{\Rightarrow}} x\}$.

Then $w \in \text{Anf}(L)$ holds if and only if there is a word $w\beta$ in L', such that there is a nonterminal in β, which is also in Φ'.

Therefore the set $L' \cap \Sigma^*\Phi^*\Phi'\Phi^*$ contains a word of the form $w\beta$ if and only if $w \in \text{Anf}(L)$ holds.

Now the homomorphism $h: (\Phi \cup \Sigma)^* \to \Sigma^*$ is defined by $h(a) = a$, $a \in \Sigma$ *and* $h(A) = \varepsilon$, $A \in \Phi$, then $h(L' \cap \Sigma^*\Phi^*\Phi'\Phi^*) = \text{Anf}(L)$ and thus Anf(L) is contextfree.

As the question, whether a given contextfree grammar generates an infinite language is decidable, for each contextfree grammar G a contextfree grammar G_1, for which $L(G_1) = \text{Anf}(L(G))$ can be effectively constructed.

If $L \subseteq \Sigma^*$ is contextfree but not regular and $c \notin \Sigma$, then Lc^* is also contextfree, and from Example 2.7 it follows that $\text{Anf}(Lc^*) \cap \Sigma^*c = Lc$. This language is contextfree but not regular. Therefore Anf(L) is for given contextfree L not necessarily regular.

Theorem 3.4. *The family of contextsensitive languages is not closed under* Anf.

Proof. Let $T = (K, \Sigma, \Gamma, \delta, q_0, F)$ be a Turing-machine, which accepts a language L, such that $\Sigma^* - L$ is not contextsensitive (such a language does exist). Without loss of generality it can be assumed that $\delta(q, a)$ is not defined iff $q \in F$ (T halts on a word w, iff $w \in L$).

Let $L_1 = \{\$\alpha_1\$ \cdots \$\alpha_n\$\$: n \geq 1$ and there is a $w \in \Sigma^*$, such that $\alpha_1 = q_0 w$, $\alpha_i \vdash \alpha_{i+1} (1 \leq i \leq n-1)\}$ $(\$ \notin \Gamma)$.

Then $\$q_0w\$ \in \text{Anf}(L_1)$ iff $w \notin L$.

Therefore $\text{Anf}(L_1) \cap \$q_0\Sigma^*\$ = \$q_0(\Sigma^* - L)\$$ and $\text{Anf}(L_1)$ is not contextsensitive.

It remains to show that L_1 is contextsensitive:

Let $G = (\{S, S_1, T, T', U\} \cup \{T_X: X \in \Gamma \cup K\} \cup \{X': X \in \Gamma \cup K\}, \Gamma \cup K \cup \{\$\}, P, S)$, where P contains the following rules (for all $a \in \Sigma$; $X, Y \in \Gamma \cup K$; $p, q \in K$; $C, D, E \in \Gamma$):

(1) $S \to \$q_0S_1$, $S_1 \to aS_1$, $S_1 \to U\$$;

(2) $U\$ \to \$\$$, $U\$ \to T\$\$$;

(3) $X'T \to TX'$, $XT \to X'T_X$, $\$T \to \$T'$, $T_XY \to YT_X$, $T_X\$ \to T\$X$, $T'X' \to XT'$, $T'\$ \to \$U$;

(4) $UC \to CU$;

(5) $UqC \to DpU$ if $\delta(q, C) = (p, D, R)$, $Uq\$ \to DpU\$$ *if* $\delta(q, b) = (p, D, R)$, $EUqC \to pEDU$ if $\delta(q, C) = (p, D, L)$, $EUq\$ \to pEDU\$$ *if* $\delta(q, b) = (p, D, L)$, $\$UqC \to \$pbDU$ if $\delta(q, C) = (p, D, L)$, $\$Uq\$ \to \$pbDU\$$ *if* $\delta(q, b) = (p, D, L)$.

Then G is contextsensitive and generates L_1 as follows:

(i) By rules (1) all derivations $S \overset{*}{\Rightarrow} \$q_0wU\$$, $w \in \Sigma^*$ are possible. Applying $U\$ \to \$\$$ to $\$q_0wU\$$, $\$q_0w\$\$ = \$\alpha_1\$\$ \in L_1$ can be derived.

(ii) Suppose there is a derivation $S \overset{*}{\Rightarrow} \$\alpha_1\$ \cdots \$\alpha_nU\$$. Then there are two possibilities:

(a) $U\$ \to \$\$$ leads to $\$\alpha_1\$ \cdots \$\alpha_n\$\$ \in L_1$.

(b) $U\$ \to T\$\$$ leads to $\$\alpha_1\$ \cdots \$\alpha_nT\$\$$ from which, using rules (3) $\$\alpha_1\$ \cdots \$\alpha_n\$U\alpha_n\$$ is obtained.

If $\alpha_n = uqv$ and $q \notin F$, then there is a unique configuration α_{n+1} such that $\alpha_n \vdash \alpha_{n+1}$ and rules (4), (5) guarantee that $U\alpha_n \overset{*}{\Rightarrow} \alpha_{n+1}U$. Hence $\$\alpha_1\$ \cdots \$\alpha_n\$\alpha_{n+1}U\$$ is obtained.

If $\alpha_n = uqv$ and $q \in F$ then $\delta(q, C)$ is not defined and no word consisting only of terminals can be obtained.

Corollary 3.5. *The family of recursively enumerable languages is not closed under* Anf.

Proof. Choosing T as in Theorem 3.4, but so that $\Sigma^* - L$ is not recursively enumerable (such a language exists), the statement is obtained.

There are also contextsensitive (recursively enumerable) languages, such that Anf(L) is contextsensitive but not contextfree (recursively enumerable but not contextsensitive). (For this see Example 2.7.)

In Theorem 2.15 a criterion, when a language can be an image under Anf, is given. Yet this condition need not hold for arbitrary languages. Therefore the inclusion $\mathfrak{L} \subseteq \text{Anf}(\mathfrak{L})$ cannot be expected for an arbitrary family of languages. But a result in this direction is:

Theorem 3.6. *For all families of languages* $\mathfrak{L}$ $\mathfrak{L} \subseteq \mathfrak{F}(\text{Anf}(\mathfrak{F}(\mathfrak{L})))$ *holds, where for given* $\mathfrak{L}'$ $\mathfrak{F}(\mathfrak{L}')$ *is the smallest family of languages containing* $\mathfrak{L}'$ *and being closed under homomorphism, concatenation with regular sets and intersection with regular sets.*

Proof. Let $L \in \mathfrak{L}$, $L \subseteq \Sigma^*$, $c \notin \Sigma$. Then $Lc^* \in \mathfrak{F}(\mathfrak{L})$, $\text{Anf}(Lc^*) \cap \Sigma^*c = Lc$ (see for this Example 2.7), and therefore $L \in \mathfrak{F}(\text{Anf}(\mathfrak{F}(\mathfrak{L})))$. ($L = h(Lc)$, where the homomorphism h is defined by $h(a) = a$, $a \in \Sigma$, $h(c) = \varepsilon$).

4. A generalization of Anf

Throughout the rest of this paper Σ denotes a fixed alphabet.

If $\mathfrak{U}$ denotes the set of all infinite languages, Anf(L) can be redefined as follows: $\mathrm{Anf}(L)=\{x: x\backslash L \in \mathfrak{U}\}$.

This representation causes the following generalization of Anf:

Definition 4.1. For each set $\mathfrak{L}$ of languages over Σ the operator $\mathrm{Anf}_{\mathfrak{L}} : \mathfrak{P}(\Sigma^*) \to \mathfrak{P}(\Sigma^*)$ is defined as follows:

$$\mathrm{Anf}_{\mathfrak{L}}(L)=\{x \in \Sigma^*: x\backslash L \in \mathfrak{L}\}.$$

Remark. In language theory a family $\mathfrak{L}$ of languages is never empty. In this paper it is convenient to allow $\mathfrak{L}$ to be empty. Therefore in Definition 4.1 "set" is used instead of "family".

As mentioned above, $\mathrm{Anf}_{\mathfrak{U}} = \mathrm{Anf}$ holds.

It is possible to represent some other language operators as $\mathrm{Anf}_{\mathfrak{L}}$; this is shown in the following theorems.

Lemma 4.2. $\mathrm{Anf}_{\mathfrak{L}}(L)=[\mathrm{Anf}_{\mathfrak{L}^c}(L)]^c$. *($L^c$ ($\mathfrak{L}^c$) means complementation with respect to* Σ^* ($\mathfrak{P}(\Sigma^*)$).)

Proof. $w \in \mathrm{Anf}_{\mathfrak{L}}(L)$ holds iff $w\backslash L \in \mathfrak{L}$, i.e. $w\backslash L \notin \mathfrak{L}^c$. This is equivalent to $w \notin \mathrm{Anf}_{\mathfrak{L}^c}(L)$.

Theorem 4.3. *There are* $\mathfrak{L}_1$, $\mathfrak{L}_2$, *such that for all* L $\mathrm{Anf}_{\mathfrak{L}_1}(L)=\emptyset$ *and* $\mathrm{Anf}_{\mathfrak{L}_2}(L)=L$.

Proof. Let $\mathfrak{L}_1=\emptyset$. Then there is no w such that $w\backslash L \in \emptyset$. Let $\mathfrak{L}_2=\{L \subseteq \Sigma^*: \varepsilon \in L\}$. Then $w\backslash L \in \mathfrak{L}_2$ is equivalent to $w \in L$, and thus $\mathrm{Anf}_{\mathfrak{L}_2}(L)=\{w: w\backslash L \in \mathfrak{L}_2\}=\{w: w \in L\}=L$.

Corollary 4.4. *There are* $\mathfrak{L}_3$, $\mathfrak{L}_4$, *such that for all* L $\mathrm{Anf}_{\mathfrak{L}_3}(L)=\Sigma^*$ *and* $\mathit{Anf}_{\mathfrak{L}_4}(L)=L^c$.

Proof. Let $\mathfrak{L}_3=\mathfrak{L}_1^c$ and $\mathfrak{L}_4=\mathfrak{L}_2^c$ ($\mathfrak{L}_1$, $\mathfrak{L}_2$ from Theorem 4.3). Then the statement follows from Lemma 4.2.

The following theorem shows that $L \mapsto \emptyset$ and $L \mapsto \Sigma^*$ are the only constant mappings that can be represented as $\mathrm{Anf}_{\mathfrak{L}}$.

Theorem 4.5. *Let* A *be a language, such that* $0 \subsetneq A \subsetneq \Sigma^*$. *Then there is no* $\mathfrak{L}$ *such that* $\mathrm{Anf}_{\mathfrak{L}}(L)=A$ *for all* L.

Proof. It is assumed that the contrary holds. As $\emptyset \subsetneq A \subsetneq \Sigma^*$, there are x, y, such that $x \in A$ and $y \notin A$. Let L be an arbitrary language. As $x \in A = \mathrm{Anf}_{\mathfrak{L}}(L)$, it follows that $x\backslash L \in \mathfrak{L}$. Then $y\backslash(y(x\backslash L)) = (x\backslash L) \in \mathfrak{L}$ and therefore $y \in \mathrm{Anf}_{\mathfrak{L}}(y(x\backslash L)) = A$, which is a contradiction.

It is also possible to represent Init as $\mathrm{Anf}_{\mathfrak{L}}$.

Theorem 4.6. $\mathrm{Anf}_{\mathfrak{P}(\Sigma^*)-\{\emptyset\}}(L) = \mathrm{Init}(L)$.

Proof. $x \in \mathrm{Init}(L)$ holds iff $x\backslash L \neq \emptyset$. This is equivalent to $x\backslash L \in \mathfrak{P}(\Sigma^*) - \{\emptyset\}$ and this to $x \in \mathrm{Anf}_{\mathfrak{P}(\Sigma^*)-\{\emptyset\}}(L)$.

The following theorem shows that beside of the the language operators above, there is a non countable set of operators of the form $\mathrm{Anf}_{\mathfrak{L}}$.

Theorem 4.7. *If* $\mathfrak{L}_1 \neq \mathfrak{L}_2$, *then there is a L, such that* $\mathrm{Anf}_{\mathfrak{L}_1}(L) \neq \mathrm{Anf}_{\mathfrak{L}_2}(L)$.

Proof. Let $\mathfrak{L}_1 \neq \mathfrak{L}_2$ and without loss of generality it is assumed that there is a L such that $L \in \mathfrak{L}_1$ and $L \notin \mathfrak{L}_2$. Then $\varepsilon \in \mathrm{Anf}_{\mathfrak{L}_1}(L)$ but $\varepsilon \notin \mathrm{Anf}_{\mathfrak{L}_2}(L)$. (This holds as $\varepsilon\backslash L = L \in \mathfrak{L}_1$ and $\varepsilon\backslash L = L \notin \mathfrak{L}_2$.)

The relations $\mathrm{Anf}(L_1 \cup L_2) = \mathrm{Anf}(L_1) \cup \mathrm{Anf}(L_2)$ and $\mathrm{Anf}(L_1 \cap L_2) \subseteq \mathrm{Anf}(L_1) \cap \mathrm{Anf}(L_2)$ which are valid for Anf cannot be generalized to $\mathrm{Anf}_{\mathfrak{L}}$ for arbitrary sets $\mathfrak{L}$. Yet it is possible to characterize those sets $\mathfrak{L}$, for which these relations are valid.

Lemma 4.8. *For all* L_1, L_2 $\mathrm{Anf}_{\mathfrak{L}}(L_1 \cup L_2) \supseteq \mathrm{Anf}_{\mathfrak{L}}(L_1) \cup \mathrm{Anf}_{\mathfrak{L}}(L_2)$ *holds iff* $\mathfrak{L}$ *has the following property*:

If $A \in \mathfrak{L}$ *and* $A \subseteq B$, *then* $B \in \mathfrak{L}$.

Proof. It is assumed that $\mathfrak{L}$ has the given property. Let $x \in \mathrm{Anf}_{\mathfrak{L}}(L_1)$ $(\mathrm{Anf}_{\mathfrak{L}}(L_2))$, i.e. $x\backslash L_1 \in \mathfrak{L}$ $(x\backslash L_2 \in \mathfrak{L})$. Then also $x\backslash L_1 \cup x\backslash L_2 = x\backslash(L_1 \cup L_2) \in \mathfrak{L}$ and therefore $x \in \mathrm{Anf}_{\mathfrak{L}}(L_1 \cup L_2)$. Hence $\mathrm{Anf}_{\mathfrak{L}}(L_1) \cup \mathrm{Anf}_{\mathfrak{L}}(L_2) \subseteq \mathrm{Anf}_{\mathfrak{L}}(L_1 \cup L_2)$.

To show the converse let A $(=\varepsilon\backslash A) \in \mathfrak{L}$, $A \subseteq B$. Then $\varepsilon \in \mathrm{Anf}_{\mathfrak{L}}(A) \subseteq \mathrm{Anf}_{\mathfrak{L}}(A) \cup \mathrm{Anf}_{\mathfrak{L}}(B) \subseteq \mathrm{Anf}_{\mathfrak{L}}(A \cup B) = \mathrm{Anf}_{\mathfrak{L}}(B)$ and thus $\varepsilon\backslash B = B \in \mathfrak{L}$.

Lemma 4.9. *For all* L_1, L_2 $\mathrm{Anf}_{\mathfrak{L}}(L_1 \cap L_2) \subseteq \mathrm{Anf}_{\mathfrak{L}}(L_1) \cap \mathrm{Anf}_{\mathfrak{L}}(L_2)$ *holds iff* $\mathfrak{L}$ *has the following property*:

If $A \in \mathfrak{L}$ *and* $A \subseteq B$, *then* $B \in \mathfrak{L}$.

Proof. It is assumed that $\mathfrak{L}$ has the given property: If $x \in \mathrm{Anf}_{\mathfrak{L}}(L_1 \cap L_2)$, then $x\backslash(L_1 \cap L_2) = (x\backslash L_1) \cap (x\backslash L_2) \in \mathfrak{L}$. Therefore all sets containing $(x\backslash L_1) \cap (x\backslash L_2)$ as a subset are in $\mathfrak{L}$. Especially $x\backslash L_1 \in \mathfrak{L}$ and $x\backslash L_2 \in \mathfrak{L}$ holds. and thus $x \in \mathrm{Anf}_{\mathfrak{L}}(L_1) \cap \mathrm{Anf}_{\mathfrak{L}}(L_2)$.

Conversely it is assumed that $\mathrm{Anf}_{\mathfrak{L}}(L_1 \cap L_2) \subseteq \mathrm{Anf}_{\mathfrak{L}}(L_1) \cap \mathrm{Anf}_{\mathfrak{L}}(L_2)$ is always valid. Let $A \in \mathfrak{L}$ and $A \subseteq B$. Then $\varepsilon \in \mathrm{Anf}_{\mathfrak{L}}(A) = \mathrm{Anf}_{\mathfrak{L}}(A \cap B) \subseteq \mathrm{Anf}_{\mathfrak{L}}(A) \cap \mathrm{Anf}_{\mathfrak{L}}(B) \subseteq \mathrm{Anf}_{\mathfrak{L}}(B)$ and thus $B \in \mathfrak{L}$.

Corollary 4.10. *The following two properties are equivalent:*

(i) *for all* L_1, L_2 $\mathrm{Anf}_{\mathfrak{L}}(L_1 \cup L_2) \supseteq \mathrm{Anf}_{\mathfrak{L}}(L_1) \cup \mathrm{Anf}_{\mathfrak{L}}(L_2)$,

(ii) *for all* L_1, L_2 $\mathrm{Anf}_{\mathfrak{L}}(L_1 \cap L_2) \subseteq \mathrm{Anf}_{\mathfrak{L}}(L_1) \cap \mathrm{Anf}_{\mathfrak{L}}(L_2)$.

Proof. The condition for $\mathfrak{L}$ in Lemma 4.8 is the same as the second condition in Lemma 4.9.

Lemma 4.11. *For all* L_1, L_2 $\mathrm{Anf}_{\mathfrak{L}}(L_1 \cup L_2) \subseteq \mathrm{Anf}_{\mathfrak{L}}(L_1) \cup \mathrm{Anf}_{\mathfrak{L}}(L_2)$ *holds iff* $\mathfrak{L}$ *has the following property:*

If $A \cup B \in \mathfrak{L}$ *then* $A \in \mathfrak{L}$ *or* $B \in \mathfrak{L}$.

Proof. It is assumed that $\mathfrak{L}$ has the given property. $x \in \mathrm{Anf}_{\mathfrak{L}}(L_1 \cup L_2)$ iff $x\backslash L_1 \cup x\backslash L_2 \in \mathfrak{L}$. Then $x\backslash L_1 \in \mathfrak{L}$ or $x\backslash L_2 \in \mathfrak{L}$ and therefore $x \in \mathrm{Anf}_{\mathfrak{L}}(L_1) \cup \mathrm{Anf}_{\mathfrak{L}}(L_2)$.

To show the converse let $A \cup B \in \mathfrak{L}$. Then $\varepsilon \in \mathrm{Anf}_{\mathfrak{L}}(\mathrm{A} \cup \mathrm{B}) \subseteq \mathrm{Anf}_{\mathfrak{L}}(A) \cup \mathrm{Anf}_{\mathfrak{L}}(B)$. Therefore $A \in \mathfrak{L}$ or $B \in \mathfrak{L}$ must hold.

Lemma 4.12. *For all* L_1, L_2 $\mathrm{Anf}_{\mathfrak{L}}(L_1 \cap L_2) \supseteq \mathrm{Anf}_{\mathfrak{L}}(L_1) \cap \mathrm{Anf}_{\mathfrak{L}}(L_2)$ *holds iff* $\mathfrak{L}$ *has the following property:*

If $A \in \mathfrak{L}$ *and* $B \in \mathfrak{L}$, *then* $A \cap B \in \mathfrak{L}$.

Proof. First, it is assumed, that $\mathfrak{L}$ has the given property. If $x \in \mathrm{Anf}_{\mathfrak{L}}(L_1) \cap \mathrm{Anf}_{\mathfrak{L}}(L_2)$, then $x\backslash L_1 \in \mathfrak{L}$ and $x\backslash L_2 \in \mathfrak{L}$, therefore $(x\backslash L_1) \cap (x\backslash L_2) = x\backslash(L_1 \cap L_2) \in \mathfrak{L}$ and thus $x \in \mathrm{Anf}_{\mathfrak{L}}(L_1 \cap L_2)$.

Conversely, $\mathrm{Anf}_{\mathfrak{L}}(L_1 \cap L_2) \supseteq \mathrm{Anf}_{\mathfrak{L}}(L_1) \cap \mathrm{Anf}_{\mathfrak{L}}(L_2)$ is assumed to hold. Let $A \in \mathfrak{L}$ and $B \in \mathfrak{L}$. Then $\varepsilon \in \mathrm{Anf}_{\mathfrak{L}}(A) \cap \mathrm{Anf}_{\mathfrak{L}}(B) \subseteq \mathrm{Anf}_{\mathfrak{L}}(A \cap B)$, and thus $A \cap B \in \mathfrak{L}$.

Theorem 4.13. *For all* L_1, L_2 $\mathrm{Anf}_{\mathfrak{L}}(L_1 \cup L_2) = \mathrm{Anf}_{\mathfrak{L}}(L_1) \cup \mathrm{Anf}_{\mathfrak{L}}(L_2)$ *holds iff* $\mathfrak{L}$ *has the following two properties:*

(i) *If* $A \cup B \in \mathfrak{L}$, *then* $A \in \mathfrak{L}$ *or* $B \in \mathfrak{L}$

(ii) *If* $A \in \mathfrak{L}$ *and* $A \subseteq B$, *then* $B \in \mathfrak{L}$.

Proof. Lemmata 4.8 and 4.11.

Theorem 4.14. *For all* L_1, L_2 $\mathrm{Anf}_{\mathfrak{L}}(L_1 \cap L_2) = \mathrm{Anf}_{\mathfrak{L}}(L_1) \cap \mathrm{Anf}_{\mathfrak{L}}(L_2)$ *holds iff* $\mathfrak{L}$ *has the following two properties:*

(i) *If* $A \in \mathfrak{L}$ *and* $B \in \mathfrak{L}$, *then* $A \cap B \in \mathfrak{L}$

(ii) *If* $A \in \mathfrak{L}$ *and* $A \subseteq B$, *then* $B \in \mathfrak{L}$.

Proof. Lemmata 4.9 and 4.12.

For arbitrary $\mathfrak{L}$ $\text{Anf}_{\mathfrak{L}}$ need not be idempotent. However, the following theorem holds:

Theorem 4.15. $\text{Anf}_{\mathfrak{L}}(\text{Anf}_{\mathfrak{L}}(L)) = \text{Anf}_{\mathfrak{L}}(L)$ *holds for all* L *iff* $\mathfrak{L}$ *has the following property*:

$$A \in \mathfrak{L} \textit{ iff } \text{Anf}_{\mathfrak{L}}(A) \in \mathfrak{L}.$$

Proof. First it will be shown, that for arbitrary $\mathfrak{L}$, x, L $\text{Anf}_{\mathfrak{L}}(x\backslash L) = x\backslash\text{Anf}_{\mathfrak{L}}(L)$ holds.

Let $w \in \text{Anf}_{\mathfrak{L}}(x\backslash L)$. This is equivalent to $w\backslash(x\backslash L) = (xw)\backslash L \in \mathfrak{L}$ and this yields $xw \in \text{Anf}_{\mathfrak{L}}(L)$. But this is equivalent to $w \in x\backslash\text{Anf}_{\mathfrak{L}}(L)$.

$\mathfrak{L}$ is assumed to have the given property. $x \in \text{Anf}_{\mathfrak{L}}(\text{Anf}_{\mathfrak{L}}(L))$ means $x\backslash\text{Anf}_{\mathfrak{L}}(L) \in \mathfrak{L}$, which is equivalent to $\text{Anf}_{\mathfrak{L}}(x\backslash L) \in \mathfrak{L}$ as shown above. Therefore $x\backslash L \in \mathfrak{L}$, i.e. $x \in \text{Anf}_{\mathfrak{L}}(L)$.

Conversely it is assumed that $\text{Anf}_{\mathfrak{L}}(\text{Anf}_{\mathfrak{L}}(L)) = \text{Anf}_{\mathfrak{L}}(L)$ is always true. $A \in \mathfrak{L}$ holds iff $\varepsilon \in \text{Anf}_{\mathfrak{L}}(A) = \text{Anf}_{\mathfrak{L}}(\text{Anf}_{\mathfrak{L}}(A))$. This is equivalent to $\text{Anf}_{\mathfrak{L}}(A) \in \mathfrak{L}$.

The idempotency law for Anf (of Chapter 2) therefore follows directly from Theorem 2.2.

The following theorem can be regarded as a generalization of Theorem 2.11.

Theorem 4.16. $\mathfrak{L}$ *is assumed to have the following properties*:

(i) $\emptyset \notin \mathfrak{L}$,

(ii) *If* $A \in \mathfrak{L}$ *and* $B \neq \emptyset$, *then* $AB \in \mathfrak{L}$ *and* $BA \in \mathfrak{L}$,

(iii) *If* $AB \in \mathfrak{L}$ *and* $B \notin \mathfrak{L}$, *then* $A \in \mathfrak{L}$,

(iv) *If* $A \cup B \in \mathfrak{L}$, *then* $A \in \mathfrak{L}$ *or* $B \in \mathfrak{L}$,

(v) *If* $A \in \mathfrak{L}$ *and* $A \subseteq B$, *then* $B \in \mathfrak{L}$.

Then for all L_1, L_2, $L_2 \neq \emptyset$

$$\text{Anf}_{\mathfrak{L}}(L_1L_2) = \begin{cases} \text{Init}(L_1) \cup L_1\,\text{Anf}_{\mathfrak{L}}(L_2) & \textit{if } L_2 \in \mathfrak{L}, \\ \text{Anf}_{\mathfrak{L}}(L_1) & \textit{if } L_2 \notin \mathfrak{L}. \end{cases}$$

Proof. The cases $L_2 \in \mathfrak{L}$ and $L_2 \notin \mathfrak{L}$ are distinguished.

(a) $L_2 \in \mathfrak{L}$. Let $w \in \text{Anf}_{\mathfrak{L}}(L_1L_2)$. Then

$$w\backslash(L_1L_2) = (w\backslash L_1)L_2 \cup \bigcup_{\substack{w=uv \\ u \in L_1}} (v\backslash L_2) \in \mathfrak{L}.$$

If $w \in \text{Init}(L_1)$, then $w \in \text{Init}(L_1) \cup L_1\,\text{Anf}_{\mathfrak{L}}(L_2)$. If $w \notin \text{Init}(L_1)$, then $w\backslash L_1 = \emptyset$. By (i) $\bigcup_{w=uv;u\in L_1}(v\backslash L_2) \neq \emptyset$ and, by (iv), there is a v, such that $w = uv$, $u \in L_1$, $v\backslash L_2 \in \mathfrak{L}$. Therefore $v \in \text{Anf}_{\mathfrak{L}}(L_2)$ and $w = uv \in L_1\,\text{Anf}_{\mathfrak{L}}(L_2)$.

Hence $\mathrm{Anf}_{\mathfrak{L}}(L_1L_2) \subseteq \mathrm{Init}(L_1) \cup L_1 \,\mathrm{Anf}_{\mathfrak{L}}(L_2)$.

Conversely, let $w \in \mathrm{Init}(L_1) \cup L_1 \,\mathrm{Anf}_{\mathfrak{L}}(L_2)$.

If $w \in \mathrm{Init}(L_1)$, then $w \backslash L_1 \neq \emptyset$ and, as $L_2 \in \mathfrak{L}$ and (ii) holds, $(w\backslash L_1)L_2 \in \mathfrak{L}$. As $(w\backslash L_1)L_2 \subseteq w\backslash(L_1L_2)$, by (v), $w\backslash(L_1L_2) \in \mathfrak{L}$ and therefore $w \in \mathrm{Anf}_{\mathfrak{L}}(L_1L_2)$.

If $w \in L_1 \,\mathrm{Anf}_{\mathfrak{L}}(L_2)$, then $w = uv$, $u \in L_1$, $v \in \mathrm{Anf}_{\mathfrak{L}}(L_2)$, therefore $v\backslash L_2 \in \mathfrak{L}$. As $v\backslash L_2 \subseteq (uv)\backslash(L_1L_2)$, by (v), $(uv)\backslash(L_1L_2) \in \mathfrak{L}$ and so $uv = w \in \mathrm{Anf}_{\mathfrak{L}}(L_1L_2)$.

Hence $\mathrm{Init}(L_1) \cup L_1 \,\mathrm{Anf}_{\mathfrak{L}}(L_2) \subseteq \mathrm{Anf}_{\mathfrak{L}}(L_1L_2)$.

(b) $L_2 \notin \mathfrak{L}$. Let $w \in \mathrm{Anf}_{\mathfrak{L}}(L_1L_2)$. Then

$$w\backslash(L_1L_2) = (w\backslash L_1)L_2 \cup \bigcup_{\substack{w=uv \\ u \in L_1}} (v\backslash L_2) \in \mathfrak{L}.$$

If $v\backslash L_2 \in \mathfrak{L}$ would hold, then, by (ii) $v(v\backslash L_2) \in \mathfrak{L}$ and, as $L_2 \supseteq v(v\backslash L_2)$, by (v) $L_2 \in \mathfrak{L}$, which is excluded. Therefore, by (iv), $(w\backslash L_1)L_2 \in \mathfrak{L}$. As $L_2 \notin \mathfrak{L}$ and (iii) holds, $w\backslash L_1 \in \mathfrak{L}$, and therefore $w \in \mathrm{Anf}_{\mathfrak{L}}(L_1)$.

Hence $\mathrm{Anf}_{\mathfrak{L}}(L_1L_2) \subseteq \mathrm{Anf}_{\mathfrak{L}}(L_1)$.

Conversely, let $w \in \mathrm{Anf}_{\mathfrak{L}}(L_1)$. Then $w\backslash L_1 \in \mathfrak{L}$ and, by (ii), $(w\backslash L_1)L_2 \in \mathfrak{L}$. As $(w\backslash L_1)L_2 \subseteq w\backslash(L_1L_2)$ and (v) holds, $w\backslash(L_1L_2) \in \mathfrak{L}$, therefore $w \in \mathrm{Anf}_{\mathfrak{L}}(L_1L_2)$.

Hence $\mathrm{Anf}_{\mathfrak{L}}(L_1) \subseteq \mathrm{Anf}_{\mathfrak{L}}(L_1L_2)$.

It should be noted, that the family $\mathfrak{U}$ of all infinite sets has properties (i)–(v). Therefore the property of Anf, stated in Theorem 2.11, is an immediate consequence of Theorem 4.16.

The necessity of properties (i), (ii) and (iii) of $\mathfrak{L}$ in Theorem 4.16 is shown in the following theorem:

Theorem 4.17. *For all* $L_1, L_2, L_2 \neq \emptyset$ *let*

$$\mathrm{Anf}_{\mathfrak{L}}(L_1L_2) = \begin{cases} \mathrm{Init}(L_1) \cup L_1 \,\mathrm{Anf}_{\mathfrak{L}}(L_2) & \text{if } L_2 \in \mathfrak{L}, \\ \mathrm{Anf}_{\mathfrak{L}}(L_1) & \text{if } L_2 \notin \mathfrak{L}. \end{cases}$$

Then $\mathfrak{L}$ *has the following properties*:

(i) $\emptyset \notin \mathfrak{L}$,

(ii) *If* $A \in \mathfrak{L}$ *and* $B \neq \emptyset$, *then* $BA \in \mathfrak{L}$ *and* $AB \in \mathfrak{L}$,

(iii) *If* $AB \in \mathfrak{L}$ *and* $B \notin \mathfrak{L}$, *then* $A \in \mathfrak{L}$.

Proof. First $\mathfrak{L} \neq \{\emptyset\}$ is shown. Assuming the contrary, because of $\{a\} \notin \mathfrak{L}$, $\mathrm{Anf}_{\{\emptyset\}}(\{a\}) = \mathrm{Anf}_{\{\emptyset\}}(\{\varepsilon\}\{a\}) = \mathrm{Anf}_{\{\emptyset\}}(\{\varepsilon\})$ would hold. As $a \in \mathrm{Anf}_{\{\emptyset\}}(\{\varepsilon\})$, but $a \notin \mathrm{Anf}_{\{\emptyset\}}(\{a\})$, $\mathfrak{L} = \{\emptyset\}$ is not possible.

Would $\mathfrak{L}$ contain $\emptyset$ and a $L \neq \emptyset$, then $\mathrm{Anf}_{\mathfrak{L}}(\emptyset L) = \mathrm{Init}(\emptyset) \cup \emptyset \mathrm{Anf}_{\mathfrak{L}}(L)$ would hold. As $\mathrm{Anf}_{\mathfrak{L}}(\emptyset L) = \mathrm{Anf}_{\mathfrak{L}}(\emptyset) = \Sigma^*$, but $\mathrm{Init}(\emptyset) \cup \emptyset \,\mathrm{Anf}_{\mathfrak{L}}(L) = \emptyset$, this is impossible.

Hence $\emptyset \notin \mathfrak{L}$ and so (i) holds.

Let $A \in \mathfrak{L}$ and $B \neq \emptyset$. Then $\mathrm{Anf}_{\mathfrak{L}}(BA) = \mathrm{Init}(B) \cup B \,\mathrm{Anf}_{\mathfrak{L}}(A)$. As $B \neq \emptyset$, $\varepsilon \in \mathrm{Init}(B) \subseteq \mathrm{Anf}_{\mathfrak{L}}(BA)$ and therefore $\varepsilon\backslash(BA) = BA \in \mathfrak{L}$.

If $B \in \mathfrak{L}$, then $AB \in \mathfrak{L}$ by a similar argument.

If $B \notin \mathfrak{L}$, then $\mathrm{Anf}_{\mathfrak{L}}(AB) = \mathrm{Anf}_{\mathfrak{L}}(A)$ and therefore $\varepsilon \in \mathrm{Anf}_{\mathfrak{L}}(A) \subseteq \mathrm{Anf}_{\mathfrak{L}}(AB)$ and so $AB \in \mathfrak{L}$.

Hence (ii) holds.

Let $AB \in \mathfrak{L}$ and $B \notin \mathfrak{L}$. Then $\mathrm{Anf}_{\mathfrak{L}}(AB) = \mathrm{Anf}_{\mathfrak{L}}(A)$. As $\varepsilon \in \mathrm{Anf}_{\mathfrak{L}}(AB) \subseteq \mathrm{Anf}_{\mathfrak{L}}(A)$, $A \in \mathfrak{L}$. Hence (iii) holds.

For certain sets $\mathfrak{L}$, the languages L, for which $\mathrm{Anf}_{\mathfrak{L}}(L) \neq \emptyset$ holds, can be characterized:

Theorem 4.18. *The following two conditions are equivalent*:
(1) *If* $w \backslash A \in \mathfrak{L}$, *then* $A \in \mathfrak{L}$,
(2) $\mathrm{Anf}_{\mathfrak{L}}(L) \neq \emptyset$ *iff* $L \in \mathfrak{L}$.

Proof. Let (1) be valid. If $\mathrm{Anf}_{\mathfrak{L}}(A) \neq \emptyset$, then there is a w, such that $w \backslash A \in \mathfrak{L}$. By (1), $A \in \mathfrak{L}$. Let conversely by $A \in \mathfrak{L}$. Then $\varepsilon \in \mathrm{Anf}_{\mathfrak{L}}(A)$ and therefore $\mathrm{Anf}_{\mathfrak{L}}(A) \neq \emptyset$. Hence (2) holds.

Let (2) be valid and $w \backslash A \in \mathfrak{L}$. Then $w \in \mathrm{Anf}_{\mathfrak{L}}(A)$ and therefore $\mathrm{Anf}_{\mathfrak{L}}(A) \neq \emptyset$. By (2) is $A \in \mathfrak{L}$ and so (1) holds.

Remark. (1) holds especially if (ii) and (v) of Theorem 4.16 hold.

The following theorem shows, that the family of all regular languages is closed under $\mathrm{Anf}_{\mathfrak{L}}$ for arbitrary $\mathfrak{L}$.

Theorem 4.19. *For all* $\mathfrak{L}$ *the following holds*:
If L *is regular, then* $\mathrm{Anf}_{\mathfrak{L}}(L)$ *is regular.*

Proof. Let $M = (K, \Sigma, \delta, q_0, F)$ be a finite automaton accepting L and $M_1 = (K, \Sigma, \delta, q_0, F)$, where $F_1 = \{q \in K : \{x \in \Sigma^* : \delta(q, x) \in F\} \in \mathfrak{L}\}$.

Then $w \in T(M_1)$ iff $\delta(q_0, w) \in F_1$, that means $\{x : \delta(\delta(q_0, w), x) \in F\} = \{x : wx \in L\} = w \backslash L \in \mathfrak{L}$. Hence $w \in \mathrm{Anf}_{\mathfrak{L}}(L)$.

If L is non regular, then, for suitable $\mathfrak{L}$, $\mathrm{Anf}_{\mathfrak{L}}(L)$ can be any arbitrary language. This is shown in the following theorems.

Theorem 4.20. *Let* $L \subseteq \Sigma^*$. *Then for all* $L_1 \subseteq \Sigma^*$ *there is a set* $\mathfrak{L} \subseteq \mathfrak{P}(\Sigma^*)$ *such that* $\mathrm{Anf}_{\mathfrak{L}}(L) = L_1$ *if and only if each class in* $\Sigma^* / \underset{L}{\sim}$ *consists of a single word.*

Proof. Let each class of $\Sigma^* / \underset{L}{\sim}$ consist of a single word. Then for $x \neq y$ there is always a z, such that (without loss of generality) $xz \in L$ and $yz \notin L$. Therefore $x \backslash L \neq y \backslash L$ iff $x \neq y$. For given L_1, let $\mathfrak{L} = \{x \backslash L : x \in L_1\}$. Then $x \in \mathrm{Anf}_{\mathfrak{L}}(L)$ iff $x \backslash L \in \mathfrak{L}$, that means $x \in L_1$. Hence $\mathrm{Anf}_{\mathfrak{L}}(L) = L_1$.

Let conversely be assumed that a class of $\Sigma^*/\underset{L}{\sim}$ contains more than one word, that means that there are x, y, $x \neq y$, but $x \underset{L}{\sim} y$. Then $x\backslash L = y\backslash L$ and for all $\mathfrak{L}$, for which $x \in \mathrm{Anf}_{\mathfrak{L}}(L)$ holds also $y \in \mathrm{Anf}_{\mathfrak{L}}(L)$. Hence there is no $\mathfrak{L}$, such that $\mathrm{Anf}_{\mathfrak{L}}(L) = \{x\}$.

Corollary 4.21. *Let $\Sigma = \{a\}$. Then there is a contextsensitive language $L \subseteq \Sigma^*$, such that for all $L_1 \subseteq \Sigma^*$ there is a $\mathfrak{L}$, such that* $\mathrm{Anf}_{\mathfrak{L}}(L) = L_1$.

Proof. Let $L = \{a^{2^n} : n \geqslant 0\}$. Then L is contextsensitive and for $i \neq j$ is a $a^i \underset{L}{\not\sim} a^j$. Therefore each class of $\Sigma^*/\underset{L}{\sim}$ consists of a single word and Theorem 4.17 can be applied.

Corollary 4.22. *Let $|\Sigma| \geqslant 2$. Then there is a contextfree language $L \subseteq \Sigma^*$, such that for all $L_1 \subseteq \Sigma^*$ there is a $\mathfrak{L}$, such that* $\mathrm{Anf}_{\mathfrak{L}}(L) = L_1$.

Proof. Let $L = \{ww^R : w \in \Sigma^*\}$. Then L is contextfree and the classes of $\Sigma^*/\underset{L}{\sim}$ consist of single words.

References

[1] J.E. Hopcroft and J.D. Ullmann, *Formal Languages and their Relation to Automata* (Addison-Wesley, Reading, MA., 1969).