ISSN 0037–4466, Siberian Mathematical Journal, 2026, Vol. 67, No. 5, pp. 1106–1115. © Pleiades Publishing, Ltd., 2026.
Russian Text © The Author(s), 2026, published in Sibirskii Matematicheskii Zhurnal, 2026, Vol. 67, No. 5, pp. 891–902.

OBJECT-ORIENTED DATA VIA PREFIX REWRITING. PART III: RECURRENCE

A. E. Gutman

UDC 519.682.1+519.683+519.7+519.1

Abstract—For an arbitrary prefix rewriting system, algorithms are described for recognizing recurrent words and verifying their existence, a new finite recurrence basis is proposed, and conditions for the existence of weakly recurrent words are obtained, together with a relation between such words and infinite chains of weak rewriting.

DOI: 10.1134/S0037446626050101

Keywords: prefix rewriting, term rewriting, information system, object-oriented data

This paper continues the planned series of publications devoted to the systematic study, announced in [1] and initiated in [2], of prefix rewriting as a tool for representing and analyzing object-oriented data systems.

The central object of study is the mechanism of rewriting prefixes of words over a finite alphabet $\mathbb{A}$ by means of an arbitrary function $\to$ defined on some finite subset $\mathbb{E} \subseteq \mathbb{A}^+$ and acting into $\mathbb{A}^+$.

A rewriting function can be regarded as a system of explicit definitions by interpreting the relation $X \to Y$ as the explicit definition "X is Y" and the rewriting $X \to Y$ of the prefix X in a word XS as the definition "the detail XS of X is YS." A system of definitions gives rise to the corresponding relation of conceptual dependence between the concepts being defined: if the definition of X explicitly or implicitly uses Y, then the concept X conceptually depends on Y. The correctness of a system of definitions presupposes the absence of cycles in the conceptual scheme. If successive application of definitions to a concept X leads to a "vicious circle," so that at some step X itself or a detail XS of X reappears, then X cannot be regarded as well-defined. In the case of a deterministic rewriting system, where only the longest prefixes are subject to rewriting, such a word X is said to be *recurrent*, whereas in the general case of prefix rewriting it is *weakly recurrent*.

A similar application of prefix rewriting to the formalization of program modules occurs in studies on representing semiautomata by canonical words. In particular, [3] studies conditions under which the corresponding prefix-rewriting systems have no infinite chains.

The material of the paper is grouped into four sections. Section 1 contains the basic definitions and notation concerning prefix rewriting. Section 2 states some results of previous papers used in the sequel. Section 3 is devoted to studying properties of recurrent words and, in particular, describes algorithms for verifying whether a given word is recurrent and whether recurrent words exist for a given rewriting function. The same section describes a recurrence basis having considerably fewer elements than the only finite basis known so far. Section 4 introduces the notions of weak rewriting and weak recurrence, obtains a criterion for the absence of weakly recurrent words, and proves that the existence of an infinite chain of weak rewriting implies the existence of a weakly recurrent word.

1. Definitions and Notation

1.1. Throughout the paper, $\mathbb{A}$ is a nonempty finite alphabet, and $\mathbb{A}^*$ (respectively, $\mathbb{A}^+$) is the set of all words (respectively, all nonempty words) over $\mathbb{A}$. The elements of $\mathbb{A}$ are called *letters* and are conventionally identified with the corresponding single-letter words. In the examples below, the symbols ω, $\mathtt{X}$, $\mathtt{Y}$, $\mathtt{A}$, $\mathtt{B}$ are used as elements of $\mathbb{A}$.

The length of a word X is denoted by $|X|$. For $X \in \mathbb{A}^+$ and $n \geqslant 0$, the symbol X^n denotes the concatenation $XX \cdots X$ consisting of n occurrences of the word X; moreover, X^0 is assumed to be the empty word $\varnothing$.

Original article submitted July 15, 2026; accepted August 3, 2026.

For brevity, in the sequel we use "*word*" as shorthand for "nonempty word over $\mathbb{A}$" and call the elements of $\mathbb{A}^*$ *strings*.

The symbol $\mathbb{N}$ denotes the set of naturals $\{1, 2, \dots\}$. We write ordered tuples and sequences using angle brackets: $\langle x_1, \dots, x_n \rangle$, $\langle x_n \rangle_{n \in \mathbb{N}}$.

1.2. Say that X is a *prefix* of a word $Y \in \mathbb{A}^+$ and write $X \sqsubseteq Y$ or $Y \sqsupseteq X$ if $X \in \mathbb{A}^+$ and $Y = XS$ for some string $S \in \mathbb{A}^*$. The suffix S of the word $Y = XS$ determined by the prefix $X \sqsubseteq Y$ will be denoted by $\frac{Y}{X}$. Thus, $\frac{XS}{X} = S$ and $X\frac{Y}{X} = Y$ whenever $X \sqsubseteq Y$. If $n \in \mathbb{N}$, $X \in \mathbb{A}^+$, and $|X| \geqslant n$, then $X{\upharpoonright}_n$ denotes the prefix of X having length n.

1.3. If a string S ends with T, then the initial fragment of S obtained by deleting the suffix T from S will be denoted by S/T. Thus, $(RT)/T = R$ for all $R, T \in \mathbb{A}^*$.

1.4. If $\rightsquigarrow$ is a symbol denoting a binary relation, then $\overset{+}{\rightsquigarrow}$ denotes the transitive closure of $\rightsquigarrow$. Thus, $x \overset{+}{\rightsquigarrow} y$ means that $x = z_0 \rightsquigarrow z_1 \rightsquigarrow \cdots \rightsquigarrow z_n = y$ for some $n \in \mathbb{N}$ and $z_0, z_1, \dots, z_n$.

1.5. By a *rewriting function* we mean an arbitrary function $\to$ defined on a finite nonempty subset of $\mathbb{A}^+$ and acting into $\mathbb{A}^+$. We treat such a function $\to$ as a binary relation on $\mathbb{A}^+$ (i.e., as a subset of $\mathbb{A}^+ \times \mathbb{A}^+$) and use the infix notation $X \to Y$ instead of the membership $\langle X, Y \rangle \in \to$ or the equality $\to(X) = Y$.

In what follows, when giving definitions and formulating statements, we assume by default that an arbitrary rewriting function $\to$ is fixed, and all introduced notions, as well as the terms and notation used, are interpreted with respect to this function.

1.6. Denote by $\mathbb{E}$ the nonempty finite subset of $\mathbb{A}^+$ which is the domain of the rewriting function $\to$ under consideration. The elements of $\mathbb{E}$ will be called *explicit words*.

Define μ as the length of the longest explicit word: $\mu := \max\{|E| : E \in \mathbb{E}\}$.

Denote by $\mathbb{A}_{\mathbb{E}}$ the *explicit alphabet*, i.e., the set of all letters occurring in explicit words.

A word $S \in \mathbb{A}^+$ is an *explicit complement* of a word X if $XS \in \mathbb{E}$.

Say that a word E is an *explicit prefix* of a word X if $E \sqsubseteq X$ and $E \in \mathbb{E}$. If X has an explicit prefix, denote by $\mathbb{E}(X)$ the longest explicit prefix of X. (In [2], the notation $\mathrm{lep}(X)$ is used instead of $\mathbb{E}(X)$.)

1.7. The value of the function $\to$ on an explicit word $E \in \mathbb{E}$ will be denoted by E' and called the *explicit rewrite* of E. Thus, $E \to E'$ for all $E \in \mathbb{E}$. A word X is *rewritable* if X has at least one explicit prefix. Given a rewritable word X, consider its longest explicit prefix $\mathbb{E}(X)$, determine the corresponding suffix $S \in \mathbb{A}^*$ such that $X = \mathbb{E}(X)S$, and introduce the notation

$$X' := \mathbb{E}(X)'S.$$

We call X' the *rewrite* of X.

1.8. Introduce the binary relation $\Rightarrow$ on $\mathbb{A}^+$ by setting $X \Rightarrow Y$ if and only if the word X is rewritable and $X' = Y$. Clearly, $X \to Y$ implies $X \Rightarrow Y$. We may read the formula $X \overset{+}{\Rightarrow} Y$ as "X *rewrites to* Y," and the formula $X \to Y$ as "X *explicitly rewrites to* Y."

1.9. For all integers $n \geqslant 0$, define the sets $\mathbb{W}_n$ of words rewritable n times and the nth rewrites $X^{(n)}$ of words $X \in \mathbb{W}_n$ by the following joint recursion:

$$\begin{aligned} &\mathbb{W}_0 := \mathbb{A}^+, && X^{(0)} := X \text{ for } X \in \mathbb{W}_0; \\ &\mathbb{W}_n := \{X \in \mathbb{W}_{n-1} : X^{(n-1)} \text{ is rewritable}\}, && X^{(n)} := (X^{(n-1)})' \text{ for } X \in \mathbb{W}_n, \quad n \geqslant 1. \end{aligned}$$

Put

$$\mathbb{W}_\infty := \bigcap_{n \in \mathbb{N}} \mathbb{W}_n, \quad \mathbb{W}_{\text{fin}} := \mathbb{A}^+ \setminus \mathbb{W}_\infty.$$

The elements of $\mathbb{W}_\infty$ and $\mathbb{W}_{\text{fin}}$ will be called *infinitely rewritable* and *finitely rewritable* words, respectively. Given a word X, call the maximal (finite or infinite) sequence of the form $\langle X^{(0)}, X^{(1)}, X^{(2)}, \dots \rangle$ the *rewriting sequence* of X. Thus, a word is finitely rewritable if and only if it has a finite rewriting sequence.

2. Preliminary Results

For subsequent use, we state without proof some results established in [2] and partially announced earlier in [1].

2.1. Proposition [2, 2.1, 2.2]. *Let $X \in \mathbb{A}^+$, let $\alpha \in \mathbb{A}$, and let $S \in \mathbb{A}^*$.*

(a) *If $X\alpha \notin \mathbb{E}$, then the conditions $X \in \mathbb{W}_1$ and $X\alpha \in \mathbb{W}_1$ are equivalent, and if either of them holds, then $(X\alpha)' = X'\alpha$.*

(b) *If X has no explicit complements, then the conditions $X \in \mathbb{W}_1$ and $XS \in \mathbb{W}_1$ are equivalent, and if either of them holds, then $(XS)' = X'S$, i.e., $XS \Rightarrow X'S$.*

2.2. Lemma [2, 2.3]. *Let $n \in \mathbb{N}$, let $X \in \mathbb{W}_n$, let $\left|X^{(1)}\right|, \ldots, \left|X^{(n)}\right| \geqslant |X|$, and let $Y \in \mathbb{A}^+$ and $S \in \mathbb{A}^*$ be such that $X = YS$ and $|Y| \geqslant \mu$. Then $Y \in \mathbb{W}_i$, $X^{(i)} = Y^{(i)}S$, and $\left|Y^{(i)}\right| \geqslant \mu$ for all $i = 0, \ldots, n$.*

2.3. Lemma [2, 2.4]. *For every sequence of naturals $\langle x_n \rangle_{n\in\mathbb{N}}$, there exists a strictly increasing sequence of naturals $\langle n_i \rangle_{i\in\mathbb{N}}$ such that, for all $n, i \in \mathbb{N}$, the inequality $n \geqslant n_i$ implies $x_n \geqslant x_{n_i}$.*

2.4. Theorem [1, 5.1; 2, 3.1]. *A word X is infinitely rewritable if and only if one of the following two mutually exclusive conditions holds:*

(a) *there are integers $n \geqslant 0$ and $r > 0$ such that $X^{(n)} = X^{(n+r)}$;*

(b) *there are integers $n \geqslant 0$ and $r > 0$ such that*

$$\mu \leqslant \left|X^{(n)}\right| \leqslant \left|X^{(n+1)}\right|, \ldots, \left|X^{(n+r)}\right|,$$
$$X^{(n)} \neq X^{(n+r)}, \quad X^{(n)}\!\upharpoonright_{\mu} = X^{(n+r)}\!\upharpoonright_{\mu}.$$

2.5. Let $\mathcal{P}$ be some set of "parameters" whose elements are admissible as inputs for algorithms. Given a subset $C \subseteq \mathbb{A}^+ \times \mathcal{P}$, we write the inclusion $\langle Y, p\rangle \in C$ as $C(Y, p)$ and call it a "condition." A condition $C(Y, p)$ is *cyclically decidable* (see [2, 3.3]) if there is an algorithm verifying the relation $(\exists\, m \in \mathbb{N})\ C(YR^mS, p)$ for arbitrary inputs $Y \in \mathbb{A}^+$, $R, S \in \mathbb{A}^*$, and $p \in \mathcal{P}$. Every cyclically decidable condition is decidable, whereas the converse implication does not hold in general (see [2, 3.4]).

Theorem [1, 5.3; 2, 3.5, 3.6]. *If $C(Y, p)$ is a cyclically decidable condition, then there exists an algorithm that, given a rewriting function, a word X, and a parameter $p \in \mathcal{P}$, verifies the existence of a word Y satisfying the conditions $X \overset{+}{\Rightarrow} Y$ and $C(Y, p)$.*

2.6. Corollary [1, 5.4; 2, 3.7]. *There exists an algorithm that, given a rewriting function and a word X, verifies whether X is finitely rewritable.*

2.7. Theorem [1, 2.12; 2, 3.8]. *If $\mathbb{E} \subseteq \mathbb{W}_{\mathrm{fin}}$, then $\mathbb{W}_{\mathrm{fin}} = \mathbb{A}^+$. In other words, if all explicit words are finitely rewritable, then all words are finitely rewritable.*

2.8. Corollary [1, 5.4; 2, 3.9]. *There exists an algorithm that, given a rewriting function, verifies whether all words are finitely rewritable.*

3. Recurrence

A word X is *recurrent* if $X \overset{+}{\Rightarrow} XS$ for some string S, i.e., if there exist $n \in \mathbb{N}$ and $S \in \mathbb{A}^*$ such that $X \in \mathbb{W}_n$ and $X^{(n)} = XS$.

Similar cyclic behavior is traditionally considered in the general theory of rewriting systems, where reductions of the form $X \overset{+}{\rightsquigarrow} RXS$ are called loops. The relation between loops and the existence of infinite rewriting sequences is studied, for instance, in [4].

The main purpose of this section is to describe algorithms for recognizing recurrent words and verifying their existence for a given rewriting function. Algorithmic problems for other variants of prefix rewriting are considered, for instance, in [5].

3.1. Theorem. *There exists an algorithm that, given a rewriting function and a word X, verifies whether X is recurrent.*

Proof. As is easily seen, the condition $C(Y,X) := (Y \sqsupseteq X)$ is cyclically decidable (see 2.5). Indeed, if $R = \varnothing$, then the relation $(\exists\, m \in \mathbb{N})\ C(YR^mS, X)$ is equivalent to $YS \sqsupseteq X$ and can be verified directly, whereas if $R \neq \varnothing$, it can be verified by a simple exhaustive search over the values of m from 1 to $\lfloor |X|/|R| \rfloor + 1$, not to mention the possibility of a more efficient verification. Therefore, the corresponding specialization of the algorithm mentioned in Theorem 2.5 verifies whether a given word X is recurrent. $\square$

3.2. Theorem. *If all prefixes of explicit words are nonrecurrent, then all words are nonrecurrent.*

Proof.[1] For brevity, denote by $\mathcal{R}$ the set of all recurrent words and by $\mathcal{P}$ the set of all prefixes of explicit words. Suppose that $\mathcal{R} \cap \mathcal{P} = \varnothing$ and prove by induction that, for all $m \in \mathbb{N}$, the set $\mathcal{R}$ contains no words of length m.

A single-letter word cannot belong to $\mathcal{R}$, since otherwise, being rewritable, it would coincide with its unique explicit prefix and hence would belong to $\mathcal{R} \cap \mathcal{P}$.

Suppose now that $\mathcal{R}$ contains no words of length m, consider a word $X \in \mathcal{R}$ of length $m+1$, and derive a contradiction.

The recurrence of X means that there exist $n \in \mathbb{N}$ and $S \in \mathbb{A}^*$ such that $X \in \mathbb{W}_n$ and

$$X^{(0)} \Rightarrow X^{(1)} \Rightarrow \cdots \Rightarrow X^{(n)} = X^{(0)}S. \tag{1}$$

Note at once that $X^{(0)} \notin \mathcal{P}$ and $X^{(n)} \notin \mathcal{P}$. (The first relation follows from the assumption $\mathcal{R} \cap \mathcal{P} = \varnothing$, while the second follows from the first by virtue of the equality $X^{(n)} = X^{(0)}S$.)

We demonstrate that $X^{(1)}, \dots, X^{(n-1)} \notin \mathcal{P}$. Indeed, suppose otherwise and let j be the least index $i \in \{1, \dots, n-1\}$ for which $X^{(i)} \in \mathcal{P}$. Since $X^{(0)}, X^{(1)}, \dots, X^{(j-1)} \notin \mathcal{P}$, Proposition 2.1(b) gives

$$X^{(0)}S \Rightarrow X^{(1)}S \Rightarrow \cdots \Rightarrow X^{(j)}S$$

and, in particular, $(X^{(0)}S)^{(j)} = X^{(j)}S$. On the other hand,

$$(X^{(j)})^{(n)} = X^{(j+n)} = (X^{(n)})^{(j)} = (X^{(0)}S)^{(j)} = X^{(j)}S,$$

i.e., $X^{(j)} \in \mathcal{R}$, contrary to the assumption $\mathcal{R} \cap \mathcal{P} = \varnothing$.

Thus, $X^{(0)}, X^{(1)}, \dots, X^{(n-1)}$ do not belong to $\mathcal{P}$. In particular, these words are not explicit and therefore differ from their longest explicit prefixes. Consequently, every rewrite in the chain (1) leaves at least the last letter of the word unchanged, and hence all the words $X^{(0)}, X^{(1)}, \dots, X^{(n)}$ end with the same letter. Thus,

$$X^{(0)} = Y_0\, \alpha, \quad X^{(1)} = Y_1\, \alpha, \quad \dots, \quad X^{(n)} = Y_n\, \alpha$$

for some $Y_0, \dots, Y_n \in \mathbb{A}^+$ and $\alpha \in \mathbb{A}$. Moreover, $|Y_0| = |X^{(0)}| - 1 = m$. We show that $Y_0 \in \mathcal{R}$, contrary to the induction hypothesis.

For each $i = 0, \dots, n-1$, the rewritable word $Y_i\, \alpha = X^{(i)}$ is not explicit, and hence Proposition 2.1(a) implies that $Y_i \in \mathbb{W}_1$ and

$$Y_i'\alpha = (Y_i\, \alpha)' = (X^{(i)})' = X^{(i+1)} = Y_{i+1}\, \alpha,$$

so that $Y_i' = Y_{i+1}$. Thus, $Y_0 \Rightarrow Y_1 \Rightarrow \cdots \Rightarrow Y_n$ and, in particular, $Y_0 \overset{+}{\Rightarrow} Y_n$.

If $S = \varnothing$, then $Y_n\, \alpha = X^{(n)} = X^{(0)}S = X^{(0)} = Y_0\, \alpha$, whence $Y_n = Y_0$ and therefore $Y_0 \in \mathcal{R}$.

Suppose now that $S \neq \varnothing$. Since the word $X^{(0)}S = X^{(n)} = Y_n\, \alpha$ ends with the letter α, the string S also ends with α, and hence $S = T\alpha$ for some string $T \in \mathbb{A}^*$. Then

$$Y_n\, \alpha = X^{(n)} = X^{(0)}S = X^{(0)}T\alpha = Y_0\, \alpha\, T\alpha,$$

whence $Y_n = Y_0\, \alpha\, T$ and therefore again $Y_0 \in \mathcal{R}$. $\square$

[1] The outline of this proof was suggested by `GPT-5.6 Sol`.

3.3. The following assertion was announced in [1, Theorem 2.9].

Corollary. *If all rewritable words $X \in \mathbb{A}^+_{\mathbb{E}}$ of length $|X| \leqslant \mu$ are nonrecurrent, then all words are nonrecurrent.*

Proof. The assertion of the corollary follows from Theorem 3.2, since every prefix of an explicit word belongs to $\mathbb{A}^+_{\mathbb{E}}$ and has length at most μ, while nonrewritable prefixes are necessarily nonrecurrent. $\square$

3.4. The following assertion was announced in [1, Theorem 2.13].

Theorem. *If there are no recurrent words, then all words are finitely rewritable.*

Proof. Suppose that all words are nonrecurrent. Arguing by contradiction, assume that there exists an infinitely rewritable word X. By Theorem 2.4, the word X satisfies one of the conditions 2.4(a), (b).

If condition 2.4(a) holds, i.e., $X^{(n)} = X^{(n+r)}$ for some $n \geqslant 0$ and $r > 0$, then $X^{(n)}$ is recurrent.

In case 2.4(b), there exist $n \geqslant 0$ and $r > 0$ such that

$$|X^{(n+1)}|, \ldots, |X^{(n+r)}| \geqslant |X^{(n)}| \geqslant \mu, \quad X^{(n)}\restriction_\mu = X^{(n+r)}\restriction_\mu.$$

Consider $Y \in \mathbb{A}^+$ and $S, T \in \mathbb{A}^*$ such that $|Y| = \mu$, $X^{(n)} = YS$, and $X^{(n+r)} = YT$. Applying Lemma 2.2 to $X^{(n)}$ and r (in the roles of X and n), we conclude that $X^{(n+i)} = Y^{(i)}S$ and $|Y^{(i)}| \geqslant \mu$ for all $i = 0, \ldots, r$. In particular, $YT = X^{(n+r)} = Y^{(r)}S$ and $|Y^{(r)}| \geqslant \mu$. Since $|Y| = \mu \leqslant |Y^{(r)}|$, the equality $YT = Y^{(r)}S$ implies that $Y \sqsubseteq Y^{(r)}$, and hence the word Y is recurrent. $\square$

3.5. Remark. The converse of Theorem 3.4 does not hold in general. For instance, each of the rewriting functions (a)–(c) described below makes all words finitely rewritable (see Theorem 2.7) but admits recurrent words:

$$\text{(a)}\ \begin{aligned}&\mathtt{X} \to \mathtt{XA},\\ &\mathtt{XA} \to \omega;\end{aligned} \qquad \text{(b)}\ \begin{aligned}&\mathtt{X} \to \mathtt{Y},\\ &\mathtt{XA} \to \omega,\\ &\mathtt{Y} \to \mathtt{XA};\end{aligned} \qquad \text{(c)}\ \begin{aligned}&\mathtt{X} \to \omega,\\ &\mathtt{XA} \to \mathtt{YAB},\\ &\mathtt{Y} \to \mathtt{X},\\ &\mathtt{YAB} \to \omega.\end{aligned}$$

Indeed, in the first two cases $\mathtt{X} \overset{+}{\Rightarrow} \mathtt{XA}$, whereas in the third case $\mathtt{YA} \overset{+}{\Rightarrow} \mathtt{YAB}$.

3.6. Remark. Let $B(\to)$ be a set of words defined by some condition depending on a rewriting function $\to$. We call $B(\to)$ a *recurrence basis* if, for an arbitrary rewriting function $\to$, nonrecurrence of all words belonging to $B(\to)$ implies the complete absence of recurrent words. According to Theorem 3.2, the set

$$\mathcal{B}(\to) := \{X \in \mathbb{A}^+ : X \sqsubseteq E \text{ for some } E \in \mathbb{E}\} \tag{2}$$

of all prefixes of explicit words is a recurrence basis.[2]

Until now, the only known example of a finite recurrence basis has been the set $\{X \in \mathbb{A}^+_{\mathbb{E}} : |X| \leqslant \mu\}$ given in [1, 2.10], which contains the basis $\mathcal{B}(\to)$ and, in general, has considerably more elements.

As Example 3.5(c) shows, the set of explicit words is not a recurrence basis even in the class of functions for which all words are finitely rewritable. Nor is the set of proper prefixes of explicit words a recurrence basis. Indeed, for the rewriting function $\{\mathtt{X} \to \mathtt{Y},\ \mathtt{YA} \to \mathtt{XA}\}$, recurrent words exist (for instance, $\mathtt{XA} \Rightarrow \mathtt{YA} \to \mathtt{XA}$ and $\mathtt{YA} \to \mathtt{XA} \Rightarrow \mathtt{YA}$), although the word $\mathtt{Y}$ is nonrecurrent.

3.7. Theorem. *There exists an algorithm that, given a rewriting function, verifies whether recurrent words exist.*

Proof. By Theorem 3.2, given a rewriting function $\to$, one can algorithmically determine the finite recurrence basis $\mathcal{B}(\to)$ (see (2)). To verify whether recurrent words exist, it remains to apply Algorithm 3.1 to all elements of this basis.

[2] In [1, 2.10], it is stated that the set of all prefixes of explicit words cannot serve as a recurrence basis. This is an error: the word "proper" is missing from that statement.

Another possible approach does not use Algorithm 3.1 and is based on Theorem 3.4, according to which, in the absence of recurrent words, all words are finitely rewritable. First, Algorithm 2.8 can be used to determine whether infinitely rewritable words exist. If such words exist, then, by Theorem 3.4, a recurrent word exists as well. (Moreover, the proof of Theorem 3.4 shows how, in this case, an example of a recurrent word can be found algorithmically.) Otherwise, it suffices, for each word $X \in \mathcal{B}(\to)$, to construct its necessarily finite rewriting sequence, checking the condition $X \sqsubseteq X^{(n)}$ at each step $n \geqslant 1$. This procedure may be more efficient than applying Algorithm 3.1 to X. $\square$

4. Weak Recurrence

We now consider a weakened version of the rewriting relation that allows arbitrary prefixes to be replaced and study the corresponding notion of recurrence. As will be shown in subsequent studies, weak rewriting and weak recurrence are closely related to the notions of conceptual dependence and conceptual consistency. The main purpose of this section is to determine conditions for the existence of weakly recurrent words and to establish the relation between such words and infinite chains of weak rewriting.

4.1. For $X, Y \in \mathbb{A}^+$ and $E \in \mathbb{E}$, write $X \xRightarrow{E}_{\mathrm{w}} Y$ if $E \sqsubseteq X$ and $Y = E' \frac{X}{E}$, or, equivalently, if $X = ES$ and $Y = E'S$ for some $S \in \mathbb{A}^*$.

Introduce the binary relation $\Rightarrow_{\mathrm{w}}$ on $\mathbb{A}^+$ by setting $X \Rightarrow_{\mathrm{w}} Y$ if and only if $X \xRightarrow{E}_{\mathrm{w}} Y$ for some $E \in \mathbb{E}$. Thus, the relation $\Rightarrow_{\mathrm{w}}$ is the rewriting corresponding to the function $\to$ regarded as an ordinary prefix-rewriting system rather than a longest-prefix rewriting system. Clearly, $X \Rightarrow Y$ implies $X \Rightarrow_{\mathrm{w}} Y$, since $X \Rightarrow Y$ if and only if X is rewritable and $X \xRightarrow{\mathbb{E}(X)}_{\mathrm{w}} Y$. We call the relation $\Rightarrow_{\mathrm{w}}$ *weak rewriting* and read the formula $X \overset{+}{\Rightarrow}_{\mathrm{w}} Y$ as "X *weakly rewrites to* Y."

4.2. A word X is *weakly recurrent* if $X \overset{+}{\Rightarrow}_{\mathrm{w}} XS$ for some string S. Clearly, every recurrent word is weakly recurrent.

4.3. The following assertion was announced in [1, Theorem 2.6(3),(4)].

Theorem. *If there are no weakly recurrent explicit words, then no word is weakly recurrent.*

PROOF. Suppose that no explicit word is weakly recurrent but, contrary to the assertion to be proved, there exists a weakly recurrent word X_0. Then

$$X_0 \xRightarrow{E_1}_{\mathrm{w}} X_1 \xRightarrow{E_2}_{\mathrm{w}} \cdots \xRightarrow{E_n}_{\mathrm{w}} X_n = X_0 S$$

for some $n \in \mathbb{N}$, $X_0, X_1, \ldots, X_n \in \mathbb{A}^+$, $E_1, \ldots, E_n \in \mathbb{E}$, and $S \in \mathbb{A}^*$. By definition 4.1, each link $X_{i-1} \xRightarrow{E_i}_{\mathrm{w}} X_i$ has the form $E_i S_i \Rightarrow_{\mathrm{w}} E_i' S_i$, where $S_i \in \mathbb{A}^*$. Thus, there are strings $S_1, \ldots, S_n$ having the following properties:

$$E_i' S_i = E_{i+1} S_{i+1} \text{ for all } i = 1, \ldots, n-1, \quad E_n' S_n = (E_1 S_1) S.$$

Let $m \in \{1, \ldots, n\}$ be an index such that

$$|S_m| = \min\{|S_1|, \ldots, |S_n|\}.$$

We derive a contradiction by proving that the string $S_m S$ ends with S_m and

$$E_m \overset{+}{\Rightarrow}_{\mathrm{w}} E_m\big((S_m S)/S_m\big),$$

so that the explicit word E_m is weakly recurrent.

The calculations below use the relations $E_i T \Rightarrow_{\mathrm{w}} E_i' T$, valid for all strings T, as well as the identities

$$(E_i S_i)/S_m = E_i(S_i/S_m), \quad \big((E_i S_i)S\big)/S_m = E_i\big((S_i S)/S_m\big),$$

which hold owing to the inequalities $|S_m| \leqslant |S_i|$ whenever at least one side of the corresponding identity is defined.

If $1 = m = n$, then

$$E_1 \Rightarrow_{\mathrm{w}} E_1' = (E_1'S_1)/S_1 = \big((E_1S_1)S\big)/S_1 = E_1\big((S_1S)/S_1\big).$$

If $1 = m < n$, then

$$\begin{aligned}
E_1 \Rightarrow_{\mathrm{w}} E_1' &= (E_1'S_1)/S_1 \\
&= (E_2S_2)/S_1 \qquad = E_2(S_2/S_1) \\
\Rightarrow_{\mathrm{w}} E_2'(S_2/S_1) &= (E_2'S_2)/S_1 \\
&= (E_3S_3)/S_1 \qquad = E_3(S_3/S_1) \\
\Rightarrow_{\mathrm{w}} E_3'(S_3/S_1) &= (E_3'S_3)/S_1 \\
&= (E_4S_4)/S_1 \qquad = E_4(S_4/S_1) \\
\dots \\
\Rightarrow_{\mathrm{w}} E_n'(S_n/S_1) &= (E_n'S_n)/S_1 \\
&= \big((E_1S_1)S\big)/S_1 = E_1\big((S_1S)/S_1\big).
\end{aligned}$$

If $1 < m = n$, then

$$\begin{aligned}
E_n \Rightarrow_{\mathrm{w}} E_n' &= (E_n'S_n)/S_n \\
&= \big((E_1S_1)S\big)/S_n = E_1\big((S_1S)/S_n\big) \\
\Rightarrow_{\mathrm{w}} E_1'\big((S_1S)/S_n\big) &= \big((E_1'S_1)S\big)/S_n \\
&= \big((E_2S_2)S\big)/S_n = E_2\big((S_2S)/S_n\big) \\
\Rightarrow_{\mathrm{w}} E_2'\big((S_2S)/S_n\big) &= \big((E_2'S_2)S\big)/S_n \\
&= \big((E_3S_3)S\big)/S_n = E_3\big((S_3S)/S_n\big) \\
\dots \\
\Rightarrow_{\mathrm{w}} E_{n-1}'\big((S_{n-1}S)/S_n\big) &= \big((E_{n-1}'S_{n-1})S\big)/S_n \\
&= \big((E_nS_n)S\big)/S_n = E_n\big((S_nS)/S_n\big).
\end{aligned}$$

If $1 < m < n$, then

$$\begin{aligned}
E_m \Rightarrow_{\mathrm{w}} E_m' &= (E_m'S_m)/S_m \\
&= (E_{m+1}S_{m+1})/S_m = E_{m+1}(S_{m+1}/S_m) \\
\Rightarrow_{\mathrm{w}} E_{m+1}'(S_{m+1}/S_m) &= (E_{m+1}'S_{m+1})/S_m \\
&= (E_{m+2}S_{m+2})/S_m = E_{m+2}(S_{m+2}/S_m) \\
\Rightarrow_{\mathrm{w}} E_{m+2}'(S_{m+2}/S_m) &= (E_{m+2}'S_{m+2})/S_m \\
&= (E_{m+3}S_{m+3})/S_m = E_{m+3}(S_{m+3}/S_m) \\
\dots \\
\Rightarrow_{\mathrm{w}} E_n'(S_n/S_m) &= (E_n'S_n)/S_m \\
&= \big((E_1S_1)S\big)/S_m \quad = E_1\big((S_1S)/S_m\big) \\
\Rightarrow_{\mathrm{w}} E_1'\big((S_1S)/S_m\big) &= \big((E_1'S_1)S\big)/S_m \\
&= \big((E_2S_2)S\big)/S_m \quad = E_2\big((S_2S)/S_m\big) \\
\Rightarrow_{\mathrm{w}} E_2'\big((S_2S)/S_m\big) &= \big((E_2'S_2)S\big)/S_m \\
&= \big((E_3S_3)S\big)/S_m \quad = E_3\big((S_3S)/S_m\big) \\
\dots \\
\Rightarrow_{\mathrm{w}} E_{m-1}'\big((S_{m-1}S)/S_m\big) &= \big((E_{m-1}'S_{m-1})S\big)/S_m \\
&= \big((E_mS_m)S\big)/S_m \quad = E_m\big((S_mS)/S_m\big).
\end{aligned}$$

Thus, in each case the explicit word E_m is weakly recurrent. The resulting contradiction completes the proof. $\square$

4.4. REMARK. Since recurrence implies weak recurrence, Theorem 4.3 shows that the absence of weakly recurrent explicit words implies the absence of recurrent words. The converse implication does not hold in general, even under the additional assumption that all words are finitely rewritable. For instance, for each of the rewriting functions (a) and (b) described below, all words are finitely rewritable and nonrecurrent (see Theorems 3.2 and 3.4), but weakly recurrent words exist:

$$\text{(a)}\quad \begin{aligned} &\mathtt{X} \to \omega, \\ &\mathtt{XA} \to \mathtt{YA}, \\ &\mathtt{Y} \to \mathtt{X}, \\ &\mathtt{YA} \to \omega; \end{aligned} \qquad \text{(b)}\quad \begin{aligned} &\mathtt{X} \to \omega, \\ &\mathtt{XA} \to \mathtt{YAB}, \\ &\mathtt{Y} \to \mathtt{X}, \\ &\mathtt{YA} \to \omega, \\ &\mathtt{YAB} \to \omega. \end{aligned}$$

Indeed, in the first case $\mathtt{XA} \Rightarrow_{\mathrm{w}} \mathtt{YA} \Rightarrow_{\mathrm{w}} \mathtt{XA}$, whereas in the second case $\mathtt{YA} \Rightarrow_{\mathrm{w}} \mathtt{XA} \Rightarrow_{\mathrm{w}} \mathtt{YAB}$.

4.5. Lemma. *Let $n \in \mathbb{N}$, let $X_0, Y_0, X_1, \dots, X_n \in \mathbb{A}^+$, let $E_1, \dots, E_n \in \mathbb{E}$, let $S \in \mathbb{A}^*$, and suppose that*

$$\begin{aligned} &X_0 = Y_0 S, \quad |Y_0| \geqslant \mu, \\ &X_0 \xRightarrow{E_1}_{\mathrm{w}} X_1 \xRightarrow{E_2}_{\mathrm{w}} \cdots \xRightarrow{E_n}_{\mathrm{w}} X_n, \\ &|X_0| \leqslant |X_1|, \dots, |X_n|. \end{aligned} \tag{3}$$

(a) *There exist words $Y_1, \dots, Y_n \in \mathbb{A}^+$ such that*

$$Y_0 \xRightarrow{E_1}_{\mathrm{w}} Y_1 \xRightarrow{E_2}_{\mathrm{w}} \cdots \xRightarrow{E_n}_{\mathrm{w}} Y_n. \tag{4}$$

(b) *The words $Y_1, \dots, Y_n$ are uniquely determined by the relation (4) and, for all $i = 1, \dots, n$, have the properties*

$$X_i = Y_i S, \quad |Y_i| \geqslant \mu. \tag{5}$$

PROOF. (b): The uniqueness of the sequence $Y_1, \dots, Y_n$ satisfying condition (4) is clear, since this condition means that

$$Y_1 = E_1' \tfrac{Y_0}{E_1}, \quad Y_2 = E_2' \tfrac{Y_1}{E_2}, \quad \dots, \quad Y_n = E_n' \tfrac{Y_{n-1}}{E_n}.$$

Relations (5) for all $i = 0, 1, \dots, n$ are easily established by induction: the induction basis for $i = 0$ is contained in the hypotheses of the lemma, while at the induction step for $i > 0$ we have

$$X_i = E_i' \tfrac{X_{i-1}}{E_i} = E_i' \tfrac{Y_{i-1} S}{E_i} = E_i' \tfrac{Y_{i-1}}{E_i} S = Y_i S$$

and, moreover,

$$|Y_i| + |S| = |Y_i S| = |X_i| \geqslant |X_0| = |Y_0 S| = |Y_0| + |S| \geqslant \mu + |S|,$$

whence $|Y_i| \geqslant \mu$.

(a): We prove the assertion by induction on n. First note that it holds for $n = 1$. Indeed, if $n = 1$, the hypotheses of the lemma imply that $E_1 \sqsubseteq X_0 = Y_0 S$ and $|E_1| \leqslant \mu \leqslant |Y_0|$, whence $E_1 \sqsubseteq Y_0$, and therefore $Y_0 \xRightarrow{E_1}_{\mathrm{w}} Y_1$ for $Y_1 := E_1' \tfrac{Y_0}{E_1}$.

Suppose that, for some $m \in \mathbb{N}$, assertion (a) holds under the hypotheses of the lemma for $n = m$, and prove that it remains valid for $n = m + 1$. Let X_0, Y_0, $X_1, \dots, X_{m+1}$, $E_1, \dots, E_{m+1}$, and S satisfy (3) for $n = m + 1$. Then X_0, Y_0, $X_1, \dots, X_m$, $E_1, \dots, E_m$, and S satisfy (3) for $n = m$, and hence, by the induction hypothesis, there exist words $Y_1, \dots, Y_m$ satisfying (4). Moreover, $X_m = Y_m S$ and $|Y_m| \geqslant \mu$ by assertion (b) proved above. Since $E_{m+1} \sqsubseteq X_m = Y_m S$ and $|E_{m+1}| \leqslant \mu \leqslant |Y_m|$, we have $E_{m+1} \sqsubseteq Y_m$, and therefore $Y_m \xRightarrow{E_{m+1}}_{\mathrm{w}} Y_{m+1}$ for $Y_{m+1} := E_{m+1}' \tfrac{Y_m}{E_{m+1}}$. $\square$

4.6. Lemma. *Consider a word X_0, a sequence of words $\langle X_i\rangle_{i\in\mathbb{N}}$, and a sequence of explicit words $\langle E_i\rangle_{i\in\mathbb{N}}$ such that*

$$X_0 \xRightarrow{E_1}_{\mathrm{w}} X_1 \xRightarrow{E_2}_{\mathrm{w}} \cdots \xRightarrow{E_i}_{\mathrm{w}} X_i \xRightarrow{E_{i+1}}_{\mathrm{w}} X_{i+1} \xRightarrow{E_{i+2}}_{\mathrm{w}} \cdots,$$

and suppose that the words $X_0, X_1, X_2, \ldots$ are pairwise distinct.

(a) *There exist $n, m \in \mathbb{N}$ such that*

$$\mu \leqslant |X_n| \leqslant |X_{n+1}|, \ldots, |X_{n+m}|, \quad X_n\restriction_\mu = X_{n+m}\restriction_\mu. \tag{6}$$

(b) *If $n, m \in \mathbb{N}$ satisfy* (6) *and $Y_0 := X_n\restriction_\mu$, then there exist words $Y_1, \ldots, Y_m$ and S such that*

$$Y_0 \xRightarrow{E_{n+1}}_{\mathrm{w}} Y_1 \xRightarrow{E_{n+2}}_{\mathrm{w}} \cdots \xRightarrow{E_{n+m}}_{\mathrm{w}} Y_m = Y_0 S.$$

PROOF. Let X_0, $\langle X_i\rangle_{i\in\mathbb{N}}$, and $\langle E_i\rangle_{i\in\mathbb{N}}$ satisfy the hypotheses of the lemma.

(a): First note that there exists $n_0 \in \mathbb{N}$ such that

$$|X_n| \geqslant \mu \quad \text{for all } n \geqslant n_0. \tag{7}$$

Indeed, if no such n_0 exists, then there is a strictly increasing sequence of naturals $\langle n_i\rangle_{i\in\mathbb{N}}$ such that $|X_{n_i}| < \mu$ for all $i \in \mathbb{N}$, and then the words X_{n_i} cannot be pairwise distinct, since there are finitely many words X with $|X| < \mu$.

By Lemma 2.3, there exists a strictly increasing sequence of naturals $\langle n_i\rangle_{i\in\mathbb{N}}$ such that, for all $n, i \in \mathbb{N}$,

$$\text{if } n \geqslant n_i, \text{ then } |X_n| \geqslant |X_{n_i}|. \tag{8}$$

Without loss of generality, we may assume that $n_1 \geqslant n_0$. Then, by (7), we have

$$|X_{n_i}| \geqslant \mu \quad \text{for all } i \in \mathbb{N}. \tag{9}$$

Since the set of words of length μ is finite, the words $X_{n_i}\restriction_\mu$ cannot be pairwise distinct. Hence there exist $i < j$ such that $X_{n_i}\restriction_\mu = X_{n_j}\restriction_\mu$. Put $n := n_i$ and $m := n_j - n_i$. Then $X_n\restriction_\mu = X_{n+m}\restriction_\mu$, and relations (8) and (9) yield the inequalities

$$\mu \leqslant |X_n| \leqslant |X_{n+1}|, \ldots, |X_{n+m}|.$$

(b): Consider arbitrary $n, m \in \mathbb{N}$ satisfying conditions (6) and put $Y_0 := X_n\restriction_\mu = X_{n+m}\restriction_\mu$, $T := \frac{X_n}{Y_0}$, and $R := \frac{X_{n+m}}{Y_0}$. Thus, $|Y_0| = \mu$, $X_n = Y_0 T$, and $X_{n+m} = Y_0 R$.

Applying Lemma 4.5 to the words X_n and Y_0 (in the roles of X_0 and Y_0), to the string T (in the role of S), to the words $X_{n+1}, \ldots, X_{n+m}$ (in the roles of $X_1, \ldots, X_n$), and to the explicit words $E_{n+1}, \ldots, E_{n+m}$ (in the roles of $E_1, \ldots, E_n$), we conclude that there exist words $Y_1, \ldots, Y_m$ such that

$$Y_0 \xRightarrow{E_{n+1}}_{\mathrm{w}} Y_1 \xRightarrow{E_{n+2}}_{\mathrm{w}} \cdots \xRightarrow{E_{n+m}}_{\mathrm{w}} Y_m,$$

while $X_{n+i} = Y_i T$ and $|Y_i| \geqslant \mu$ for all $i = 1, \ldots, m$. In particular, $Y_m T = X_{n+m} = Y_0 R$ and $|Y_m| \geqslant \mu$. On the other hand, $|Y_0| = \mu$, and hence $Y_0 \sqsubseteq Y_m$, i.e.,

$$Y_m = Y_0 S$$

for some string S. It remains to note that $S \neq \varnothing$, since otherwise $X_n = Y_0 T = Y_m T = X_{n+m}$. $\square$

4.7. Theorem. *If there exist words $X_1, X_2, X_3, \ldots$ forming an infinite chain $X_1 \Rightarrow_{\mathrm{w}} X_2 \Rightarrow_{\mathrm{w}} X_3 \Rightarrow_{\mathrm{w}} \cdots$, then there exists a weakly recurrent word.*

PROOF. If $X_n = X_m$ for some $n < m$, then the word X_n is weakly recurrent by definition. If the words $X_1, X_2, X_3, \ldots$ are pairwise distinct, then the existence of a weakly recurrent word follows from Lemma 4.6. $\square$

4.8. Remark. Note that, for general string-rewriting systems, an infinite reduction need not exhibit looping behavior (see [4]). Theorem 4.7 shows that the situation is substantially more rigid for the weak prefix rewriting considered here: the existence of an infinite chain always implies the existence of a loop.

Theorem 4.7 will be useful, in particular, in justifying the algorithm for verifying the existence of weakly recurrent words (see [1, Algorithm 5.10]). The absence of weakly recurrent words is closely related to conceptual consistency of the prefix-rewriting system under consideration (see [1, Theorem 2.6]). A subsequent paper in the planned series of publications will be devoted to a detailed study of this relation.

FUNDING

The work was carried out in the framework of the State Task to the Sobolev Institute of Mathematics (project FWNF–2026–0022).

CONFLICT OF INTEREST

The author of this work declares that he has no conflicts of interest.

References

1. Gutman A.E., “Object-oriented data via prefix rewriting. Part I: Overview and main results,” Sib. Math. J., vol. 67, no. 3, 531–547 (2026).
2. Gutman A.E., “Object-oriented data via prefix rewriting. Part II: Effective analysis of rewriting,” Sib. Math. J., vol. 67, no. 4, 785–792 (2026).
3. Brzozowski J.A. and Jürgensen H., “Representation of semiautomata by canonical words and equivalences,” Internat. J. Found. Comput. Sci., vol. 16, no. 5, 831–850 (2005).
4. Geser A. and Zantema H., “Non-looping string rewriting,” RAIRO Theor. Inform. Appl., vol. 33, no. 3, 279–301 (1999).
5. Lohrey M. and Petersen H., “Complexity results for prefix grammars,” RAIRO Theor. Inform. Appl., vol. 39, no. 2, 391–401 (2005).

Publisher’s Note

Pleiades Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
AI tools may have been used in the translation or editing of this article.

A. E. Gutman
Sobolev Institute of Mathematics, Novosibirsk, Russia
https://orcid.org/0000-0003-2030-7459
gutman@math.nsc.ru