

ОБЪЕКТНО–ОРИЕНТИРОВАННЫЕ ДАННЫЕ И ПРЕФИКСНАЯ ПЕРЕЗАПИСЬ. ЧАСТЬ III: РЕКУРРЕНТНОСТЬ

А. Е. Гутман

Аннотация. Для произвольной префиксной перезаписывающей системы описаны алгоритмы распознавания рекуррентных слов и проверки их существования, предложен новый конечный базис рекуррентности, а также выяснены условия существования слабо рекуррентных слов и установлена связь таких слов с бесконечными цепями слабой перезаписи.

DOI 10.33048/smzh.2026.67.510

Ключевые слова: префиксная перезапись, полутуэвская система, информационная система, объектно-ориентированная система данных.

Эта статья продолжает запланированную серию публикаций, посвященную анонсированному в [1] и начатому в [2] систематическому исследованию префиксной перезаписи как инструмента для представления и анализа объектно-ориентированных систем данных.

Центральным предметом исследования является механизм перезаписи префиксов слов конечного алфавита A посредством произвольной функции $\rightarrow$, определенной на некотором конечном подмножестве $E \subseteq A^+$ и действующей в A^+ .

Функцию перезаписи можно считать системой явных определений, трактуя соотношение $X \rightarrow Y$ как явное определение « X — это Y », а перезапись $X \rightarrow Y$ префикса X в слове XS — как определение «деталь XS объекта X — это YS ». Система определений порождает соответствующее отношение концептуальной зависимости между определяемыми понятиями: если определение слова X явно или неявно задействует Y , то понятие X концептуально зависит от Y . Корректность системы определений предполагает отсутствие циклов в концептуальной схеме. Если при последовательном применении определений к понятию X возникает «порочный круг» и на каком-то шаге вновь появляется сам объект X или его деталь XS , то X нельзя считать корректно определенным. В случае детерминированной перезаписывающей системы, когда перезаписи подлежат лишь самые длинные префиксы, такое слово X называется *рекуррентным*, а в общем случае префиксной перезаписи — *слабо рекуррентным*.

Сходное применение префиксной перезаписи к формализации программных модулей встречается в работах о представлении полуавтоматов каноническими словами. В частности, в [3] исследуются условия, при которых в соответствующих системах префиксной перезаписи отсутствуют бесконечные цепи.

Работа выполнена в рамках государственного задания ИМ СО РАН (проект №FWNF–2026–0022).

Материал статьи распределен по четырем параграфам. Параграф 1 содержит основные определения и обозначения, связанные с префиксной перезаписью. В параграфе 2 сформулированы некоторые результаты предшествующих работ, используемые в дальнейшем изложении. Параграф 3 посвящен исследованию свойств рекуррентных слов и, в частности, содержит описание алгоритмов, проверяющих рекуррентность данного слова и наличие рекуррентных слов для данной функции перезаписи. В том же параграфе описан базис рекуррентности, существенно меньший по числу элементов, чем единственный известный до сих пор конечный базис. В параграфе 4 введены понятия слабой перезаписи и слабой рекуррентности, получен критерий отсутствия слабо рекуррентных слов и доказано, что наличие бесконечной цепи слабой перезаписи влечет существование слабо рекуррентного слова.

§ 1. Определения и обозначения

1.1. На протяжении всей статьи $\mathbb{A}$ — непустой конечный алфавит, $\mathbb{A}^*$ (соответственно, $\mathbb{A}^+$) — множество всех (соответственно, всех непустых) слов над $\mathbb{A}$. Элементы $\mathbb{A}$ называются *буквами* и традиционно отождествляются с соответствующими однобуквенными словами. В приводимых ниже примерах в роли элементов $\mathbb{A}$ используются символы ω, X, Y, A, B .

Длина слова X обозначается через $|X|$. Для $X \in \mathbb{A}^+$ и $n \geq 0$ символом X^n обозначается соединение $XX \cdots X$, составленное из n вхождений слова X ; при этом X^0 полагается равным пустому слову $\emptyset$.

Для краткости будем в дальнейшем говорить «слово» вместо «непустое слово над $\mathbb{A}$ », а элементы $\mathbb{A}^*$ будем называть *строками*.

Символ $\mathbb{N}$ обозначает множество натуральных чисел $\{1, 2, \dots\}$. Упорядоченные кортежи и последовательности записываются с использованием угловых скобок: $\langle x_1, \dots, x_n \rangle$, $\langle x_n \rangle_{n \in \mathbb{N}}$.

1.2. Слово X будем называть *префиксом* слова $Y \in \mathbb{A}^+$ и писать $X \sqsubseteq Y$ или $Y \supseteq X$, если $X \in \mathbb{A}^+$ и $Y = XS$ для некоторой строки $S \in \mathbb{A}^*$. Суффикс S слова $Y = XS$, определяемый префиксом $X \sqsubseteq Y$, условимся записывать в виде $\frac{Y}{X}$. Таким образом, $\frac{XS}{X} = S$ и $X \frac{Y}{X} = Y$ при $X \sqsubseteq Y$. Если $n \in \mathbb{N}$, $X \in \mathbb{A}^+$ и $|X| \geq n$, то символом $X|_n$ обозначается префикс слова X , имеющий длину n .

1.3. Если строка S оканчивается на T , то начальный фрагмент строки S , полученный из S удалением суффикса T , будет обозначаться через S/T . Таким образом, $(RT)/T = R$ для любых $R, T \in \mathbb{A}^*$.

1.4. Если символ $\rightsquigarrow$ обозначает какое-либо бинарное отношение, то символом $\rightsquigarrow^+$ обозначается транзитивное замыкание отношения $\rightsquigarrow$. Таким образом, запись $x \rightsquigarrow^+ y$ означает, что $x = z_0 \rightsquigarrow z_1 \rightsquigarrow \cdots \rightsquigarrow z_n = y$ для некоторых $n \in \mathbb{N}$ и $z_0, z_1, \dots, z_n$.

1.5. *Функцией перезаписи* назовем произвольную функцию $\rightarrow$, определенную на конечном непустом подмножестве $\mathbb{A}^+$ и действующую в $\mathbb{A}^+$. Такую функцию $\rightarrow$ условимся считать бинарным отношением на $\mathbb{A}^+$ (т. е. подмножеством $\mathbb{A}^+ \times \mathbb{A}^+$) и использовать инфиксную запись $X \rightarrow Y$ вместо принадлежности $\langle X, Y \rangle \in \rightarrow$ или равенства $\rightarrow(X) = Y$.

В дальнейшем, давая определения и формулируя утверждения, по умолчанию считаем фиксированной произвольную функцию перезаписи $\rightarrow$, относительно которой интерпретируются вводимые понятия, а также используемые термины и обозначения.

1.6. Непустое конечное подмножество $\mathbb{A}^+$, являющееся областью определения рассматриваемой функции перезаписи $\rightarrow$, обозначим символом $\mathbb{E}$. Элементы множества $\mathbb{E}$ будем называть *явными словами*. Определим число μ как длину самого длинного явного слова:

$$\mu := \max\{|E| : E \in \mathbb{E}\}.$$

Множество всех букв, встречающихся в явных словах, назовем *явным алфавитом* и обозначим символом $\mathbb{A}_{\mathbb{E}}$:

$$\mathbb{A}_{\mathbb{E}} := \min\{A \subseteq \mathbb{A} : \mathbb{E} \subseteq A^+\}.$$

Слово $S \in \mathbb{A}^+$ назовем *явным дополнением* слова X , если $XS \in \mathbb{E}$.

Будем говорить, что слово E является *явным префиксом* слова X , если $E \sqsubseteq X$ и $E \in \mathbb{E}$. В случае, когда X имеет явный префикс, условимся обозначать символом $\mathbb{E}(X)$ самый длинный явный префикс слова X . (В [2] вместо $\mathbb{E}(X)$ используется обозначение $\text{lep}(X)$.)

1.7. Значение функции $\rightarrow$ на явном слове $E \in \mathbb{E}$ будем обозначать символом E' и называть *явной перезаписью* слова E . Таким образом, $E \rightarrow E'$ для всех $E \in \mathbb{E}$. Слово X назовем *перезаписываемым*, если X имеет хотя бы один явный префикс. Для перезаписываемого слова X рассмотрим его самый длинный явный префикс $\mathbb{E}(X)$, определим соответствующий суффикс $S \in \mathbb{A}^*$, для которого $X = \mathbb{E}(X)S$, и введем обозначение

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

Слово X' назовем *перезаписью* слова X .

1.8. Определим бинарное отношение $\Rightarrow$ на $\mathbb{A}^+$, полагая $X \Rightarrow Y$ тогда и только тогда, когда слово X перезаписываемо и $X' = Y$. Ясно, что из $X \rightarrow Y$ следует $X \Rightarrow Y$. Формулу $X \stackrel{\pm}{\Rightarrow} Y$ можно прочитывать как « X перезаписывается в Y », а формулу $X \rightarrow Y$ можно снабдить прочтением « X явно перезаписывается в Y ».

1.9. Для всех целых $n \geq 0$ определим множества $\mathbb{W}_n$ n -кратно перезаписываемых слов и n -кратные перезаписи $X^{(n)}$ слов $X \in \mathbb{W}_n$ посредством следующей совместной рекурсии:

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

Положим

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

Элементы множеств $\mathbb{W}_{\infty}$ и $\mathbb{W}_{\text{fin}}$ будем называть соответственно *бесконечно перезаписываемыми* и *конечно перезаписываемыми* словами. Для данного слова X максимальную (конечную или бесконечную) последовательность вида

$$\langle X^{(0)}, X^{(1)}, X^{(2)}, \dots \rangle$$

назовем *последовательностью перезаписи* слова X . Таким образом, слово конечно перезаписываемо тогда и только тогда, когда оно имеет конечную последовательность перезаписи.

§ 2. Предварительные результаты

Для последующего использования приведем без доказательств некоторые результаты, установленные в [2] и частично анонсированные ранее в [1].

2.1. Предложение [2, 2.1, 2.2]. Пусть $X \in \mathbb{A}^+$, $\alpha \in \mathbb{A}$ и $S \in \mathbb{A}^*$.

- (а) Если $X\alpha \notin \mathbb{E}$, то условия $X \in \mathbb{W}_1$ и $X\alpha \in \mathbb{W}_1$ равносильны, и при выполнении любого из них имеет место равенство $(X\alpha)' = X'\alpha$.
- (б) Если X не имеет явных дополнений, то условия $X \in \mathbb{W}_1$ и $XS \in \mathbb{W}_1$ равносильны, и при выполнении любого из них имеет место равенство $(XS)' = X'S$, т. е. $XS \Rightarrow X'S$.

2.2. Лемма [2, 2.3]. Пусть $n \in \mathbb{N}$, $X \in \mathbb{W}_n$, $|X^{(1)}|, \dots, |X^{(n)}| \geq |X|$, и пусть $Y \in \mathbb{A}^+$ и $S \in \mathbb{A}^*$ таковы, что $X = YS$ и $|Y| \geq \mu$. Тогда $Y \in \mathbb{W}_i$, $X^{(i)} = Y^{(i)}S$ и $|Y^{(i)}| \geq \mu$ для всех $i = 0, \dots, n$.

2.3. Лемма [2, 2.4]. Для всякой последовательности натуральных чисел $\langle x_n \rangle_{n \in \mathbb{N}}$ существует строго возрастающая последовательность натуральных чисел $\langle n_i \rangle_{i \in \mathbb{N}}$ такая, что для любых $n, i \in \mathbb{N}$ из $n \geq n_i$ следует $x_n \geq x_{n_i}$.

2.4. Теорема [1, 5.1; 2, 3.1]. Слово X является бесконечно перезаписываемым тогда и только тогда, когда выполняется одно из следующих двух взаимоисключающих условий:

- (а) существуют целые числа $n \geq 0$ и $r > 0$ такие, что $X^{(n)} = X^{(n+r)}$;
- (б) существуют целые числа $n \geq 0$ и $r > 0$ такие, что

$$\mu \leq |X^{(n)}| \leq |X^{(n+1)}|, \dots, |X^{(n+r)}|, \\ X^{(n)} \neq X^{(n+r)}, \quad X^{(n)} \downarrow_\mu = X^{(n+r)} \downarrow_\mu.$$

2.5. Пусть $\mathcal{P}$ — некоторое множество «параметров», элементы которого допустимы в качестве входных данных алгоритмов. Рассматривая какое-либо подмножество $C \subseteq \mathbb{A}^+ \times \mathcal{P}$, будем записывать включение $\langle Y, p \rangle \in C$ в виде $C(Y, p)$ и называть его «условием». Условие $C(Y, p)$ называется *циклически разрешимым* (см. [2, 3.3]), если существует алгоритм, который для любых входных данных $Y \in \mathbb{A}^+$, $R, S \in \mathbb{A}^*$ и $p \in \mathcal{P}$ проверяет соотношение $(\exists m \in \mathbb{N}) C(YR^mS, p)$. Всякое циклически разрешимое условие является разрешимым, а обратная импликация в общем случае не имеет места (см. [2, 3.4]).

Теорема [1, 5.3; 2, 3.5, 3.6]. Если $C(Y, p)$ — циклически разрешимое условие, то существует алгоритм, который для данной функции перезаписи, данного слова X и данного параметра $p \in \mathcal{P}$ проверяет наличие слова Y , удовлетворяющего условиям $X \stackrel{\pm}{\Rightarrow} Y$ и $C(Y, p)$.

2.6. Следствие [1, 5.4; 2, 3.7]. Существует алгоритм, который для данной функции перезаписи и данного слова X проверяет, является ли слово X конечно перезаписываемым.

2.7. Теорема [1, 2.12; 2, 3.8]. Если $\mathbb{E} \subseteq \mathbb{W}_{\text{fin}}$, то $\mathbb{W}_{\text{fin}} = \mathbb{A}^+$. Иначе говоря, если все явные слова конечно перезаписываемы, то конечно перезаписываемыми являются все слова.

2.8. Следствие [1, 5.4; 2, 3.9]. Существует алгоритм, который для данной функции перезаписи проверяет, являются ли все слова конечно перезаписываемыми.

§ 3. Рекуррентность

Слово X называется *рекуррентным*, если $X \stackrel{\pm}{\Rightarrow} XS$ для некоторой строки S , т. е. существуют такие $n \in \mathbb{N}$ и $S \in \mathbb{A}^*$, что $X \in \mathbb{W}_n$ и $X^{(n)} = XS$.

Подобная цикличность традиционно рассматривается и в общей теории перезаписывающих систем, где редукции вида $X \stackrel{\pm}{\Rightarrow} RXS$ называются петлями. Связь петель с существованием бесконечных последовательностей перезаписи исследуется, например, в [4].

Основная цель этого параграфа — описание алгоритмов распознавания рекуррентных слов и проверки их существования для данной функции перезаписи. Алгоритмические вопросы для других вариантов префиксной перезаписи рассматриваются, например, в [5].

3.1. Теорема. *Существует алгоритм, который для данной функции перезаписи и данного слова X проверяет, является ли слово X рекуррентным.*

ДОКАЗАТЕЛЬСТВО. Как легко видеть, условие

$$C(Y, X) := Y \supseteq X$$

циклически разрешимо (см. п. 2.5). Действительно, при $R = \emptyset$ соотношение $(\exists m \in \mathbb{N}) C(YR^mS, X)$ равносильно $YS \supseteq X$ и проверяется непосредственно, а в случае $R \neq \emptyset$ его можно проверить простым перебором значений m от 1 до целой части $\frac{|X|}{|R|} + 1$, не говоря уже о возможности более эффективной проверки. Поэтому соответствующая специализация алгоритма, упомянутого в теореме 2.5, проверяет, является ли данное слово X рекуррентным. $\square$

3.2. Теорема. *Если все префиксы явных слов нерекуррентны, то все слова нерекуррентны.*

ДОКАЗАТЕЛЬСТВО.¹ Для краткости обозначим символом $\mathcal{R}$ множество всех рекуррентных слов, а символом $\mathcal{P}$ — множество всех префиксов явных слов. Предположим, что $\mathcal{R} \cap \mathcal{P} = \emptyset$, и, используя индукцию, докажем, что для всех $m \in \mathbb{N}$ множество $\mathcal{R}$ не содержит слов длины m .

Однобуквенное слово не может принадлежать $\mathcal{R}$, поскольку в противном случае оно, будучи перезаписываемым, совпало бы со своим единственным явным префиксом и тем самым оказалось бы элементом $\mathcal{R} \cap \mathcal{P}$.

Предположим теперь, что $\mathcal{R}$ не содержит слов длины m , рассмотрим слово $X \in \mathcal{R}$ длины $m + 1$ и получим противоречие.

Рекуррентность X означает существование $n \in \mathbb{N}$ и $S \in \mathbb{A}^*$ таких, что $X \in \mathbb{W}_n$ и

$$X^{(0)} \Rightarrow X^{(1)} \Rightarrow \dots \Rightarrow X^{(n)} = X^{(0)}S. \quad (1)$$

Сразу заметим, что $X^{(0)} \notin \mathcal{P}$ и $X^{(n)} \notin \mathcal{P}$. (Первое соотношение следует из условия $\mathcal{R} \cap \mathcal{P} = \emptyset$, а второе вытекает из первого в силу равенства $X^{(n)} = X^{(0)}S$.)

Покажем, что $X^{(1)}, \dots, X^{(n-1)} \notin \mathcal{P}$. Действительно, допустив противное, положим j равным наименьшему из номеров $i \in \{1, \dots, n-1\}$, для которых $X^{(i)} \in \mathcal{P}$. Поскольку $X^{(0)}, X^{(1)}, \dots, X^{(j-1)} \notin \mathcal{P}$, по предложению 2.1(b) имеем

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

и, в частности, $(X^{(0)}S)^{(j)} = X^{(j)}S$. С другой стороны,

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

т. е. $X^{(j)} \in \mathcal{R}$ вопреки условию $\mathcal{R} \cap \mathcal{P} = \emptyset$.

¹Схема этого доказательства была предложена моделью GPT-5.6 Sol.

Итак, $X^{(0)}, X^{(1)}, \dots, X^{(n-1)}$ не принадлежат $\mathcal{P}$. В частности, эти слова не являются явными и поэтому отличны от своих самых длинных явных префиксов. Следовательно, каждая перезапись в цепи (1) оставляет неизменной по крайней мере последнюю букву слова, а значит, все слова $X^{(0)}, X^{(1)}, \dots, X^{(n)}$ оканчиваются на одну и ту же букву. Таким образом,

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

для некоторых $Y_0, \dots, Y_n \in \mathbb{A}^+$ и $\alpha \in \mathbb{A}$. При этом $|Y_0| = |X^{(0)}| - 1 = m$. Покажем, что $Y_0 \in \mathcal{R}$ вопреки индукционному предположению.

Для каждого $i = 0, \dots, n-1$ перезаписываемое слово $Y_i \alpha = X^{(i)}$ не является явным, откуда согласно предложению 2.1(a) следует, что $Y_i \in \mathbb{W}_1$ и

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

и поэтому $Y'_i = Y_{i+1}$. Таким образом, $Y_0 \Rightarrow Y_1 \Rightarrow \dots \Rightarrow Y_n$ и, в частности,

$$Y_0 \stackrel{+}{\Rightarrow} Y_n.$$

Если $S = \emptyset$, то $Y_n \alpha = X^{(n)} = X^{(0)} S = X^{(0)} = Y_0 \alpha$, откуда $Y_n = Y_0$ и тем самым $Y_0 \in \mathcal{R}$.

Пусть теперь $S \neq \emptyset$. Поскольку слово $X^{(0)} S = X^{(n)} = Y_n \alpha$ оканчивается буквой α , строка S также оканчивается буквой α , а значит,

$$S = T \alpha$$

для некоторой строки $T \in \mathbb{A}^*$. Тогда

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

откуда $Y_n = Y_0 \alpha T$ и тем самым вновь $Y_0 \in \mathcal{R}$. $\square$

3.3. Следующее утверждение анонсировано в [1, теорема 2.9].

Следствие. Если все перезаписываемые слова $X \in \mathbb{A}_{\mathbb{E}}^+$ длины $|X| \leq \mu$ нерекуррентны, то все слова нерекуррентны.

ДОКАЗАТЕЛЬСТВО. Утверждение следствия вытекает из теоремы 3.2, поскольку всякий префикс явного слова принадлежит $\mathbb{A}_{\mathbb{E}}^+$ и имеет длину не более μ , причем непerezаписываемые префиксы заведомо нерекуррентны. $\square$

3.4. Следующее утверждение анонсировано в [1, теорема 2.13].

Теорема. Если нет рекуррентных слов, то все слова конечно перезаписываемы.

ДОКАЗАТЕЛЬСТВО. Пусть все слова нерекуррентны. Рассуждая от противного, допустим, что существует бесконечно перезаписываемое слово X . По теореме 2.4 слово X удовлетворяет одному из условий 2.4(a), (b).

Если выполнено условие 2.4(a), т.е. $X^{(n)} = X^{(n+r)}$ для некоторых $n \geq 0$ и $r > 0$, то слово $X^{(n)}$ рекуррентно.

В случае 2.4(b) имеются такие $n \geq 0$ и $r > 0$, что $|X^{(n+1)}|, \dots, |X^{(n+r)}| \geq |X^{(n)}| \geq \mu$ и $X^{(n)} \upharpoonright_{\mu} = X^{(n+r)} \upharpoonright_{\mu}$. Рассмотрим слово Y и строки S и T , для которых $|Y| = \mu$, $X^{(n)} = YS$ и $X^{(n+r)} = YT$. Применяя лемму 2.2 к $X^{(n)}$ и r (в роли X и n), заключаем, что $X^{(n+i)} = Y^{(i)}S$ и $|Y^{(i)}| \geq \mu$ для всех $i = 0, \dots, r$. В частности, $YT = X^{(n+r)} = Y^{(r)}S$ и $|Y^{(r)}| \geq \mu$. Поскольку $|Y| = \mu \leq |Y^{(r)}|$, из равенства $YT = Y^{(r)}S$ следует, что $Y \subseteq Y^{(r)}$, а значит, слово Y рекуррентно. $\square$

3.5. ЗАМЕЧАНИЕ. Импликация, обратная утверждению теоремы 3.4, в общем случае не имеет места. Например, каждая из описанных ниже функций перезаписи (a)–(c) обеспечивает конечную перезаписываемость любого слова (см. теорему 2.7), но допускает наличие рекуррентных слов:

$$\begin{array}{lll}
 \text{(a)} & X \rightarrow XA, & \text{(b)} & X \rightarrow Y, \\
 & XA \rightarrow \omega; & & XA \rightarrow \omega, \\
 & & & Y \rightarrow XA; \\
 \text{(c)} & X \rightarrow \omega, & & X \rightarrow YAB, \\
 & XA \rightarrow YAB, & & Y \rightarrow X, \\
 & YAB \rightarrow \omega. & &
 \end{array}$$

Действительно, в первых двух случаях $X \stackrel{+}{\Rightarrow} XA$, а в третьем случае $YA \stackrel{+}{\Rightarrow} YAB$.

3.6. ЗАМЕЧАНИЕ. Пусть $B(\rightarrow)$ — множество слов, определяемое некоторым условием, зависящим от функции перезаписи $\rightarrow$. Назовем $B(\rightarrow)$ *базисом рекуррентности*, если для произвольной функции перезаписи $\rightarrow$ из нерекуррентности слов, принадлежащих $B(\rightarrow)$, следует полное отсутствие рекуррентных слов. Согласно теореме 3.2 множество

$$\mathcal{B}(\rightarrow) := \{X \in \mathbb{A}^+ : X \sqsubseteq E \text{ для некоторого } E \in \mathbb{E}\} \quad (2)$$

всех префиксов явных слов служит базисом рекуррентности.²

До сих пор единственным известным примером конечного базиса рекуррентности было приведенное в [1, 2.10] множество $\{X \in \mathbb{A}_E^+ : |X| \leq \mu\}$, которое содержит базис $\mathcal{B}(\rightarrow)$ и в общем случае значительно превосходит его по количеству элементов.

Как показывает пример 3.5(c), множество явных слов не является базисом рекуррентности даже в классе функций, обеспечивающих конечную перезаписываемость всех слов. Базисом рекуррентности не является и множество собственных префиксов явных слов. Действительно, для функции перезаписи $\{X \rightarrow Y, YA \rightarrow XA\}$ имеются рекуррентные слова (например, $XA \Rightarrow YA \rightarrow XA$ и $YA \rightarrow XA \Rightarrow YA$), хотя слово Y нерекуррентно.

3.7. Теорема. Существует алгоритм, который для данной функции перезаписи проверяет наличие рекуррентных слов.

ДОКАЗАТЕЛЬСТВО. Благодаря теореме 3.2 по данной функции перезаписи $\rightarrow$ можно алгоритмически определить конечный базис рекуррентности $\mathcal{B}(\rightarrow)$ (см. (2)). Для проверки наличия рекуррентных слов остается применить алгоритм 3.1 ко всем элементам такого базиса.

Другой возможный подход не использует алгоритм 3.1 и основывается на теореме 3.4, согласно которой при отсутствии рекуррентных слов все слова являются конечно перезаписываемыми. Сначала с помощью алгоритма 2.8 можно выяснить, существуют ли бесконечно перезаписываемые слова. Если такие слова существуют, то по теореме 3.4 существует и рекуррентное слово. (Более того, из доказательства теоремы 3.4 видно, как в этом случае можно алгоритмически найти пример рекуррентного слова.) В противном случае достаточно построить для всякого слова $X \in \mathcal{B}(\rightarrow)$ его заведомо конечную последовательность перезаписи, проверяя условие $X \sqsubseteq X^{(n)}$ на каждом шаге $n \geq 1$. Такая процедура может оказаться эффективнее применения к X алгоритма 3.1. $\square$

²В [1, 2.10] сказано, что множество всех префиксов явных слов не может служить базисом рекуррентности. Это ошибка: в том утверждении пропущено слово «собственных».

§ 4. Слабая рекуррентность

Далее рассматривается ослабленный вариант отношения перезаписи, допускающий замену произвольных префиксов, и изучается соответствующее ему понятие рекуррентности. Как покажут дальнейшие исследования, слабая перезапись и слабая рекуррентность тесно связаны с понятиями концептуальной зависимости и непротиворечивости. Основная задача этого параграфа состоит в выяснении условий существования слабо рекуррентных слов и связи таких слов с бесконечными цепями слабой перезаписи.

4.1. Для $X, Y \in \mathbb{A}^+$ и $E \in \mathbb{E}$ будем писать $X \xrightarrow{E}_w Y$, если $E \subseteq X$ и $Y = E' \frac{X}{E}$ или, что то же самое, $X = ES$ и $Y = E'S$ для некоторого $S \in \mathbb{A}^*$.

Введем в рассмотрение бинарное отношение $\Rightarrow_w$ на $\mathbb{A}^+$, полагая $X \Rightarrow_w Y$ тогда и только тогда, когда $X \xrightarrow{E}_w Y$ для некоторого $E \in \mathbb{E}$. Таким образом, отношение $\Rightarrow_w$ представляет собой перезапись, которая соответствует функции $\rightarrow$, рассматриваемой как обычная префиксная перезаписывающая система, а не как система перезаписи самых длинных префиксов. Ясно, что из $X \Rightarrow Y$ следует $X \Rightarrow_w Y$, поскольку $X \Rightarrow Y$ тогда и только тогда, когда X перезаписываемо и $X \xrightarrow{\mathbb{E}(X)}_w Y$. Отношение $\Rightarrow_w$ назовем *слабой перезаписью* и условимся прочитывать формулу $X \xrightarrow{\pm}_w Y$ как « X слабо перезаписывается в Y ».

4.2. Слово X назовем *слабо рекуррентным*, если $X \xrightarrow{\pm}_w XS$ для некоторой строки S . Очевидно, всякое рекуррентное слово является слабо рекуррентным.

4.3. Следующее утверждение анонсировано в [1, теорема 2.6(3),(4)].

Теорема. Если среди явных слов нет слабо рекуррентных, то ни одно слово не является слабо рекуррентным.

ДОКАЗАТЕЛЬСТВО. Предположим, что явные слова не являются слабо рекуррентными, но вопреки доказываемому утверждению существует слабо рекуррентное слово X_0 . Тогда

$$X_0 \xrightarrow{E_1}_w X_1 \xrightarrow{E_2}_w \dots \xrightarrow{E_n}_w X_n = X_0 S$$

для некоторых $n \in \mathbb{N}$, $X_0, X_1, \dots, X_n \in \mathbb{A}^+$, $E_1, \dots, E_n \in \mathbb{E}$ и $S \in \mathbb{A}^*$. По определению 4.1 каждое звено $X_{i-1} \xrightarrow{E_i}_w X_i$ имеет вид $E_i S_i \Rightarrow_w E'_i S_i$, где $S_i \in \mathbb{A}^*$. Таким образом, имеются строки $S_1, \dots, S_n$, удовлетворяющие следующим условиям:

$$\begin{aligned} E'_i S_i &= E_{i+1} S_{i+1} \text{ для всех } i = 1, \dots, n-1, \\ E'_n S_n &= (E_1 S_1) S. \end{aligned}$$

Пусть $m \in \{1, \dots, n\}$ — такой номер, что

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

Получим противоречие, доказав, что строка $S_m S$ оканчивается на S_m и

$$E_m \xrightarrow{\pm}_w E_m((S_m S)/S_m),$$

а значит, явное слово E_m оказывается слабо рекуррентным.

В приводимых ниже выкладках используются соотношения $E_i T \Rightarrow_w E'_i T$, имеющие место для любых строк T , а также тождества

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

справедливые благодаря неравенствам $|S_m| \leq |S_i|$ в тех случаях, когда имеет смысл хотя бы одна из частей соответствующего тождества.

В случае $1 = m = n$ имеем

$$E_1 \Rightarrow_w E'_1 = (E'_1 S_1)/S_1 = ((E_1 S_1)S)/S_1 = E_1((S_1 S)/S_1).$$

В случае $1 = m < n$ имеем

$$\begin{aligned} E_1 \Rightarrow_w E'_1 &= (E'_1 S_1)/S_1 \\ &= (E_2 S_2)/S_1 = E_2(S_2/S_1) \\ \Rightarrow_w E'_2(S_2/S_1) &= (E'_2 S_2)/S_1 \\ &= (E_3 S_3)/S_1 = E_3(S_3/S_1) \\ \Rightarrow_w E'_3(S_3/S_1) &= (E'_3 S_3)/S_1 \\ &= (E_4 S_4)/S_1 = E_4(S_4/S_1) \\ &\dots \\ \Rightarrow_w E'_n(S_n/S_1) &= (E'_n S_n)/S_1 \\ &= ((E_1 S_1)S)/S_1 = E_1((S_1 S)/S_1). \end{aligned}$$

В случае $1 < m = n$ имеем

$$\begin{aligned} E_n \Rightarrow_w E'_n &= (E'_n S_n)/S_n \\ &= ((E_1 S_1)S)/S_n = E_1((S_1 S)/S_n) \\ \Rightarrow_w E'_1((S_1 S)/S_n) &= ((E'_1 S_1)S)/S_n \\ &= ((E_2 S_2)S)/S_n = E_2((S_2 S)/S_n) \\ \Rightarrow_w E'_2((S_2 S)/S_n) &= ((E'_2 S_2)S)/S_n \\ &= ((E_3 S_3)S)/S_n = E_3((S_3 S)/S_n) \\ &\dots \\ \Rightarrow_w E'_{n-1}((S_{n-1} S)/S_n) &= ((E'_{n-1} S_{n-1})S)/S_n \\ &= ((E_n S_n)S)/S_n = E_n((S_n S)/S_n). \end{aligned}$$

В случае $1 < m < n$ имеем

$$\begin{aligned} E_m \Rightarrow_w E'_m &= (E'_m S_m)/S_m \\ &= (E_{m+1} S_{m+1})/S_m = E_{m+1}(S_{m+1}/S_m) \\ \Rightarrow_w E'_{m+1}(S_{m+1}/S_m) &= (E'_{m+1} S_{m+1})/S_m \\ &= (E_{m+2} S_{m+2})/S_m = E_{m+2}(S_{m+2}/S_m) \\ \Rightarrow_w E'_{m+2}(S_{m+2}/S_m) &= (E'_{m+2} S_{m+2})/S_m \\ &= (E_{m+3} S_{m+3})/S_m = E_{m+3}(S_{m+3}/S_m) \\ &\dots \\ \Rightarrow_w E'_n(S_n/S_m) &= (E'_n S_n)/S_m \\ &= ((E_1 S_1)S)/S_m = E_1((S_1 S)/S_m) \\ \Rightarrow_w E'_1((S_1 S)/S_m) &= ((E'_1 S_1)S)/S_m \\ &= ((E_2 S_2)S)/S_m = E_2((S_2 S)/S_m) \\ \Rightarrow_w E'_2((S_2 S)/S_m) &= ((E'_2 S_2)S)/S_m \\ &= ((E_3 S_3)S)/S_m = E_3((S_3 S)/S_m) \\ &\dots \\ \Rightarrow_w E'_{m-1}((S_{m-1} S)/S_m) &= ((E'_{m-1} S_{m-1})S)/S_m \\ &= ((E_m S_m)S)/S_m = E_m((S_m S)/S_m). \end{aligned}$$

Таким образом, в каждом случае явное слово E_m оказывается слабо рекуррентным. Полученное противоречие завершает доказательство. $\square$

4.4. ЗАМЕЧАНИЕ. Поскольку рекуррентность влечет слабую рекуррентность, из теоремы 4.3 следует, что отсутствие слабо рекуррентных явных слов влечет отсутствие рекуррентных слов. Обратная импликация в общем случае не имеет места даже при дополнительном предположении конечной перезаписываемости всех слов. Например, для каждой из описанных ниже функций перезаписи (а) и (b) все слова являются конечно перезаписываемыми и нерекуррентными (см. теоремы 3.2 и 3.4), но имеются слабо рекуррентные слова:

$$\begin{array}{ll} \text{(a)} & \begin{array}{l} X \rightarrow \omega, \\ XA \rightarrow YA, \\ Y \rightarrow X, \\ YA \rightarrow \omega; \end{array} & \text{(b)} & \begin{array}{l} X \rightarrow \omega, \\ XA \rightarrow YAB, \\ Y \rightarrow X, \\ YA \rightarrow \omega, \\ YAB \rightarrow \omega. \end{array} \end{array}$$

Действительно, в первом случае $XA \Rightarrow_w YA \Rightarrow_w XA$, а во втором — $YA \Rightarrow_w XA \Rightarrow_w YAB$.

4.5. Лемма. Пусть $n \in \mathbb{N}$, $X_0, Y_0, X_1, \dots, X_n \in \mathbb{A}^+$, $E_1, \dots, E_n \in \mathbb{E}$, $S \in \mathbb{A}^*$, и пусть

$$\begin{aligned} X_0 &= Y_0 S, \quad |Y_0| \geq \mu, \\ X_0 &\xrightarrow{E_1}_w X_1 \xrightarrow{E_2}_w \dots \xrightarrow{E_n}_w X_n, \\ |X_0| &\leq |X_1|, \dots, |X_n|. \end{aligned} \quad (3)$$

(а) Существуют такие слова $Y_1, \dots, Y_n \in \mathbb{A}^+$, что

$$Y_0 \xrightarrow{E_1}_w Y_1 \xrightarrow{E_2}_w \dots \xrightarrow{E_n}_w Y_n. \quad (4)$$

(б) Слова $Y_1, \dots, Y_n$ однозначно определяются соотношением (4) и для всех $i = 1, \dots, n$ удовлетворяют условиям

$$X_i = Y_i S, \quad |Y_i| \geq \mu. \quad (5)$$

ДОКАЗАТЕЛЬСТВО. (б) Единственность последовательности $Y_1, \dots, Y_n$, удовлетворяющей условию (4), очевидна, поскольку это условие означает, что

$$Y_1 = E'_1 \frac{Y_0}{E_1}, \quad Y_2 = E'_2 \frac{Y_1}{E_2}, \quad \dots, \quad Y_n = E'_n \frac{Y_{n-1}}{E_n}.$$

Соотношения (5) для всех $i = 0, 1, \dots, n$ легко установить индукцией: база индукции при $i = 0$ содержится в условиях леммы, а на индукционном шаге для $i > 0$ имеем

$$X_i = E'_i \frac{X_{i-1}}{E_i} = E'_i \frac{Y_{i-1} S}{E_i} = E'_i \frac{Y_{i-1}}{E_i} S = Y_i S$$

и, кроме того,

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

откуда $|Y_i| \geq \mu$.

(а) Докажем утверждение индукцией по n . Сначала заметим, что оно справедливо в случае $n = 1$. Действительно, при $n = 1$ из условий леммы следует, что $E_1 \subseteq X_0 = Y_0 S$ и $|E_1| \leq \mu \leq |Y_0|$, откуда $E_1 \subseteq Y_0$, а значит, $Y_0 \xrightarrow{E_1}_w Y_1$ для $Y_1 := E'_1 \frac{Y_0}{E_1}$.

Допустим, что для некоторого $m \in \mathbb{N}$ утверждение (а) справедливо в условиях леммы при $n = m$, и докажем, что оно сохраняет силу при $n = m + 1$. Пусть $X_0, Y_0, X_1, \dots, X_{m+1}, E_1, \dots, E_{m+1}$ и S удовлетворяют условиям (3) при $n = m + 1$. Тогда $X_0, Y_0, X_1, \dots, X_m, E_1, \dots, E_m$ и S удовлетворяют условиям (3) при $n = m$, а значит, по предположению индукции имеются слова $Y_1, \dots, Y_m$, удовлетворяющие (4). При этом $X_m = Y_m S$ и $|Y_m| \geq \mu$ согласно доказанному выше утверждению (б). Поскольку $E_{m+1} \subseteq X_m = Y_m S$ и $|E_{m+1}| \leq \mu \leq |Y_m|$, имеем $E_{m+1} \subseteq Y_m$, а значит, $Y_m \xrightarrow{E_{m+1}}_w Y_{m+1}$ для $Y_{m+1} := E'_{m+1} \frac{Y_m}{E_{m+1}}$. $\square$

4.6. Лемма. Пусть слово X_0 , последовательность слов $\langle X_i \rangle_{i \in \mathbb{N}}$ и последовательность явных слов $\langle E_i \rangle_{i \in \mathbb{N}}$ таковы, что

$$X_0 \xRightarrow{E_1}_w X_1 \xRightarrow{E_2}_w \dots \xRightarrow{E_i}_w X_i \xRightarrow{E_{i+1}}_w X_{i+1} \xRightarrow{E_{i+2}}_w \dots,$$

причем слова $X_0, X_1, X_2, \dots$ попарно различны.

(а) Существуют $n, m \in \mathbb{N}$ такие, что

$$\mu \leq |X_n| \leq |X_{n+1}|, \dots, |X_{n+m}|, \quad X_n \upharpoonright_\mu = X_{n+m} \upharpoonright_\mu. \quad (6)$$

(б) Если $n, m \in \mathbb{N}$ удовлетворяют условиям (6) и $Y_0 := X_n \upharpoonright_\mu$, то найдутся такие слова $Y_1, \dots, Y_m$ и S , что

$$Y_0 \xRightarrow{E_{n+1}}_w Y_1 \xRightarrow{E_{n+2}}_w \dots \xRightarrow{E_{n+m}}_w Y_m = Y_0 S.$$

Доказательство. Пусть $X_0, \langle X_i \rangle_{i \in \mathbb{N}}$ и $\langle E_i \rangle_{i \in \mathbb{N}}$ удовлетворяют условиям леммы.

(а) Прежде всего заметим, что существует число $n_0 \in \mathbb{N}$, для которого

$$|X_n| \geq \mu \quad \text{при всех } n \geq n_0. \quad (7)$$

Действительно, если такого числа n_0 нет, то имеется строго возрастающая последовательность натуральных чисел $\langle n_i \rangle_{i \in \mathbb{N}}$ такая, что $|X_{n_i}| < \mu$ для всех $i \in \mathbb{N}$, а тогда слова X_{n_i} не могут быть попарно различными, так как множество $\{X \in \mathbb{A}^+ : |X| < \mu\}$ конечно.

Согласно лемме 2.3 существует строго возрастающая последовательность натуральных чисел $\langle n_i \rangle_{i \in \mathbb{N}}$ такая, что для любых $n, i \in \mathbb{N}$

$$\text{из } n \geq n_i \text{ следует } |X_n| \geq |X_{n_i}|. \quad (8)$$

Без ограничения общности можно считать, что $n_1 \geq n_0$. Тогда в силу (7) имеем

$$|X_{n_i}| \geq \mu \quad \text{для всех } i \in \mathbb{N}. \quad (9)$$

Поскольку множество слов длины μ конечно, слова $X_{n_i} \upharpoonright_\mu$ не могут быть попарно различными. Следовательно, найдутся такие числа $i < j$, что $X_{n_i} \upharpoonright_\mu = X_{n_j} \upharpoonright_\mu$. Положим $n := n_i$ и $m := n_j - n_i$. Тогда $X_n \upharpoonright_\mu = X_{n+m} \upharpoonright_\mu$, причем благодаря соотношениям (8) и (9) справедливы неравенства

$$\mu \leq |X_n| \leq |X_{n+1}|, \dots, |X_{n+m}|.$$

(б) Рассмотрим произвольные $n, m \in \mathbb{N}$, удовлетворяющие условиям (6), и положим $Y_0 := X_n \upharpoonright_\mu = X_{n+m} \upharpoonright_\mu$, $T := \frac{X_n}{Y_0}$ и $R := \frac{X_{n+m}}{Y_0}$. Таким образом, $|Y_0| = \mu$, $X_n = Y_0 T$ и $X_{n+m} = Y_0 R$.

Применяя лемму 4.5 к словам X_n и Y_0 (в роли X_0 и Y_0), строке T (в роли S), словам $X_{n+1}, \dots, X_{n+m}$ (в роли $X_1, \dots, X_n$) и явным словам $E_{n+1}, \dots, E_{n+m}$ (в роли $E_1, \dots, E_n$), заключаем, что имеются слова $Y_1, \dots, Y_m$, для которых

$$Y_0 \xRightarrow{E_{n+1}}_w Y_1 \xRightarrow{E_{n+2}}_w \dots \xRightarrow{E_{n+m}}_w Y_m,$$

причем $X_{n+i} = Y_i T$ и $|Y_i| \geq \mu$ при всех $i = 1, \dots, m$. В частности, $Y_m T = X_{n+m} = Y_0 R$ и $|Y_m| \geq \mu$. С другой стороны, $|Y_0| = \mu$, а значит, $Y_0 \subseteq Y_m$, т. е.

$$Y_m = Y_0 S$$

для некоторой строки S . Остается заметить, что $S \neq \emptyset$, так как в противном случае $X_n = Y_0 T = Y_m T = X_{n+m}$. $\square$

4.7. Теорема. Если существуют слова $X_1, X_2, X_3, \dots$, образующие бесконечную цепь $X_1 \Rightarrow_w X_2 \Rightarrow_w X_3 \Rightarrow_w \dots$, то существует слабо рекуррентное слово.

ДОКАЗАТЕЛЬСТВО. Если $X_n = X_m$ для некоторых $n < m$, то слово X_n слабо рекуррентно по определению. Если же слова $X_1, X_2, X_3, \dots$ попарно различны, то существование слабо рекуррентного слова вытекает из леммы 4.6. $\square$

4.8. ЗАМЕЧАНИЕ. Отметим, что для общих систем перезаписи строк бесконечная редукция не обязана сопровождаться петлевым поведением (см. [4]). Теорема 4.7 показывает, что для рассматриваемой слабой префиксной перезаписи ситуация существенно жестче: наличие бесконечной цепи всегда влечет существование петли.

Теорема 4.7 окажется полезной, в частности, при обосновании алгоритма проверки существования слабо рекуррентных слов (см. [1, алгоритм 5.10]). Отсутствие слабо рекуррентных слов тесно связано с концептуальной непротиворечивостью рассматриваемой системы префиксной перезаписи (см. [1, теорема 2.6]). Соответствующему детальному исследованию будет посвящена очередная работа в рамках запланированной серии публикаций.

ЛИТЕРАТУРА

1. Гутман А. Е. Объектно-ориентированные данные и префиксная перезапись. Часть I: Обзор и основные результаты // Сиб. мат. журн. 2026. Т. 67, № 3. С. 392–413.
2. Гутман А. Е. Объектно-ориентированные данные и префиксная перезапись. Часть II: Эффективный анализ перезаписи // Сиб. мат. журн. 2026. Т. 67, № 4. С. 616–625.
3. Brzozowski J. A., Jürgensen H. Representation of semiautomata by canonical words and equivalences // Internat. J. Found. Comput. Sci. 2005. V. 16, N 5. P. 831–850.
4. Geser A., Zantema H. Non-looping string rewriting // RAIRO Theor. Inform. Appl. 1999. V. 33, N 3. P. 279–301.
5. Lohrey M., Petersen H. Complexity results for prefix grammars // RAIRO Theor. Inform. Appl. 2005. V. 39, N 2. P. 391–401.

Поступила в редакцию 15 июля 2026 г.

После доработки 15 июля 2026 г.

Принята к публикации 3 августа 2026 г.

Гутман Александр Ефимович (ORCID 0000-0003-2030-7459)
Институт математики им. С. Л. Соболева СО РАН,
пр. Академика Коптюга, 4, Новосибирск 630090
gutman@math.nsc.ru