Модел. и анализ информ. систем. Т.**14**, №**2** (2007) 12–16

УДК 512.552.4+519.115.1

О числе запретов, задающих периодическую последовательность

Челноков Г.Р. [1]
Ярославский государственный университет,
150 000, Ярославль, Советская, 14,

получена 29 апреля 2007

Аннотация

Рассматриваются последовательности W периода u над алфавитом из l букв. Требуется однозначно определить последовательность W, указывая слова, не являющиеся ее подсловами. Для $n \in \mathbb{N}$ обозначим за U_n множество слов u длины n, не являющихся степенями (т.е. не представимых в виде $u = v^k$, $k > 1$). Пусть $T(u^\infty)$ — минимальное число запретов, задающих последовательность u^∞. Обозначим

$$m_n = \max_{u \in U_n} T(u^\infty), \quad r_n = \min_{u \in U_n} T(u^\infty).$$

Доказаны следующие теоремы:

Теорема 1. $m_n \leq n(l-1)$.

Отметим, что оценка точна при бесконечно многих n и реализуется, например, для периода, содержащего все слова некоторой фиксированной длины t (т.е. $n = l^t$).

Теорема 2. $r_n \geq \log_2 n + 1$.

Теорема 3. *Существует возрастающая последовательность n_i, такая, что*

$$r_{n_i} \leq \log_\phi n_i, \quad \text{где} \quad \phi = \frac{1+\sqrt{5}}{2}.$$

1. Введение

Исследование комбинаторных свойств периодических последовательностей играет важную роль в проблемах бернсайдовского типа. Соответствующие вопросы исследовались рядом авторов, см. [1-3].

При изучении мономиальных алгебр (алгебр, заданных соотношениями вида: моном равен 0) важную роль играют алгебры A_u, заданные соотношениями $v = 0$, где v — не подслово u^∞, см. [3]. Это — первичная конечно определенная PI-алгебра, и все первичные конечно определенные мономиальные алгебры имеют такой вид. Можно показать, что все определяющие соотношения алгебры A_u имеют вид $v = 0$, где $|v| \leq |u| - 1$.

Представляет интерес более точное исследование структуры соотношений, задающих A_u. Этому и посвящена настоящая работа.

2. Основные результаты

Далее все рассматриваемые слова есть над фиксированным алфавитом A из l букв.

Ниже некоторые множества слов будут называться *системами запретов*. Определим понятия, связанные с системой запретов.

Пусть дана некоторая система запретов $V = \{v_i\}$. Будем говорить, что *w удовлетворяет системе запретов V*, если слово w не содержит ни одно из v_i в качестве подслова; *система запретов определяет бесконечное в обе стороны слово w*, если w есть единственное бесконечное слово, удовлетворяющее системе запретов v_i.

Легко показать, что только периодические слова могут быть определены конечной системой запретов, см. [1].

Под *минимальной* системой запретов, определяющей слово w, ниже будет пониматься минимальная по количеству слов система.

Рассмотрим слово u, не содержащее ни один из запретов в качестве подслова. Слово uX, где $X \in A$, называется *продолжением вправо* слова u, если uX также не содержит ни один из запретов в качестве подслова. Слово u называется *однозначно продолжаемым вправо*, если u имеет ровно одно продолжение

[1]Исследование выполнено при поддержке РФФИ, грант №06-01-00648.

вправо, *неоднозначно продолжаемым*, если имеет больше одного продолжения, и *непродолжаемым*, если не имеет ни одного продолжения. Аналогичным образом определяются продолжения слова на любое количество букв, а также бесконечные продолжения. Продолжения влево определяются тоже аналогично.

Слово u называется *началом* слова w, если слово w представимо в виде us, *собственным началом*, если слово s непусто. Аналогично определяется *конец* и *собственный конец*.

Для $n \in \mathbb{N}$ обозначим за U_n множество слов u длины n, не являющихся степенями (т.е. не представимых в виде $u = v^k$, $k > 1$). Очевидно, что эти и только эти слова являются периодами последовательностей, наименьший период которых есть n. Обозначим через $T(u^\infty)$ минимальное число запретов, задающих последовательность u^∞. Будем обозначать

$$m_n = \max_{u \in U_n} T(u^\infty), \quad r_n = \min_{u \in U_n} T(u^\infty).$$

Рассмотрим бесконечное в обе стороны слово w. Множество всех слов v_i, таких, что v_i не есть подслово w, но любое собственное подслово v_i является подсловом w, будем называть *канонической системой запретов* и обозначать $C(w)$. Легко показать, что $C(u^\infty)$ конечна, см. [1].

Лемма 1. *Одна из минимальных систем запретов, задающих слово u^∞, есть $C(u^\infty)$.*

Доказательство. Пусть дана система запретов V, задающая слово u^∞. Каждый запрет является невстречающимся подсловом. Если есть запреты, содержащие невстречающиеся подслова в качестве собственных подслов, то заменим каждый из таких запретов на его минимальное по включению невстречающееся подслово, и выкинем из системы повторяющиеся запреты, если они появились. Полученная система V_1 не слабее исходной. Теперь докажем, что каждое минимальное невстречающееся подслово входит в систему V_1 в качестве запрета. Пусть слово v является минимальным невстречающимся в u^∞, но не входит в V_1. Возможны 2 случая.

1. Пусть существует бесконечное влево слово s и бесконечное вправо слово t, такие, что слова sv и vt не содержат запретов из V_1. Тогда рассмотрим бесконечное в обе стороны слово svt. Любое его подслово есть или подслово одного из слов sv и vt, и тогда не может быть запретом V_1 по предположению, или содержит подсловом слово v, и тогда не может быть запретом из V_1, ибо тогда этот запрет или не минимален, или совпадает с v. Итак, в этом случае существует удовлетворяющее системе запретов бесконечное в обе стороны слово svt, следовательно, система запретов V_1 не задает слово u^∞.

2. Пусть, без ограничения общности, не существует бесконечного вправо слова t, такого, что vt не содержит запретов. Пусть первая буква слова v есть X, обозначим $v = Xs$. Заметим, что s есть подслово в u^∞ из минимальности v, следовательно, s допускает бесконечное продолжение вправо st (такое, как в u^∞). Слово vt содержит запрет $v_i \in V_1$, следовательно, этот запрет содержит первую букву слова v, тогда v_i или содержит слово v, что противоречит минимальности v_i, или содержится в слове v, что противоречит минимальности слова v. $\square$

Теорема 1. $m_n \le n(l-1)$.

Доказательство. Рассмотрим слово u^∞, минимальный период которого n. Если у произвольного запрета $v_i \in C(u^\infty)$ отрезать последнюю букву, он станет подсловом u^∞. Таким образом, каждому запрету соответствует пара из буквы алфавита (являющейся последней буквой этого запрета) и места в периоде, на которое попадает правый конец этого запрета без последней буквы. Заметим, что если бы двум запретам соответствовала одна пара, то один из них был бы подсловом другого, значит, запретов не больше, чем пар. Всего позиций в периоде n, для каждой запрещено может быть не больше $l-1$ продолжений, значит, всего запретов не больше $n(l-1)$.

$\square$

В качестве примера u, для которого оценка становится точной, построим такое u, что при некотором фиксированном натуральном k каждое слово длины k встречается ровно 1 раз на периоде u^∞. Тогда период этого слова имеет длину l^k. Минимальных невстречающихся слов есть ровно $l^k(l-1)$, так как для каждого слова длины k есть его единственное продолжение, а все остальные продолжения не встречаются. Но все эти слова длины $k+1$ есть минимальные невстречающиеся, ибо каждое слово длины k встречается.

Построим такое u. Рассмотрим ориентированный граф, вершины которого суть слова длины $k-1$ над A, то есть количество вершин l^{k-1}. Из вершины X в вершину Y ведет стрелка, если последние $k-2$ буквы слова X есть первые $k-2$ буквы слова Y. Таким образом, стрелки этого графа биективно соответствуют словам длины k. Будем говорить, что на каждой стрелке написана буква — последняя буква слова, соответствующего вершине, в которую входит эта стрелка. Входящая степень каждой вершины в этом графе равна выходящей, ибо обе равны l, кроме того, граф, очевидно, связен. Тогда в нем есть эйлеров цикл (см., например, [4]). Последовательность букв, соответствующих стрелкам цикла, и есть искомое слово u.

Теорема 2. $r_n \geq \log_2 n + 1$.

Доказательство.
Докажем сначала несколько лемм.

Лемма 2. *Для натуральных чисел $k_1, \ldots, k_i$ из $k_1 k_2 \cdots k_i \geq n$ следует*

$$(k_1 - 1) + \cdots + (k_i - 1) \geq \log_2 n.$$

Доказательство. Предположим противное и рассмотрим контрпример с минимальным значением суммы $\max(0; k_1 - 2) + \max(0; k_2 - 2) + \cdots + \max(0; k_i - 2)$. Если все k_j для $j \in [1; i]$ равны 1 или 2 — утверждение очевидно. Если некоторое $k_j \geq 3$, то заменим его на $k_j - 1$ и добавим $k_{i+1} = 2$. Произведение $k_1, \ldots, k_{i+1}$ увеличилось, сумма $(k_1 - 1) + \cdots + (k_{i+1} - 1)$ не изменилась, значение суммы $\max(0; k_1 - 2) + \max(0; k_2 - 2) + \cdots + \max(0; k_{i+1} - 2)$ уменьшилось, что противоречит минимальности контрпримера. □

Рассмотрим бесконечное в обе стороны слово w. Слово v будем называть *развилкой*, если оно относительно системы запретов $C(w)$ неоднозначно продолжается как влево, так и вправо. Количество продолжений слова v вправо назовем *правой кратностью* или просто *кратностью* развилки v.

Два слова, ни одно из которых не является подсловом другого, назовем *несравнимыми.*

Лемма 3. *Для каждого слова v, являющегося подсловом u^∞, существует наименьшая развилка w, содержащая v в качестве подслова, причем единственного, если w не есть u^∞.*

Доказательство. Пусть есть две несравнимые развилки w_1 и w_2, содержащие v, обозначим $w_1 = s_1 v t_1$ и $w_2 = s_2 v t_2$. Пусть s — наибольший общий конец слов s_1 и s_2, а t — наибольшее общее начало слов t_1 и t_2. Тогда $w = svt$ — развилка, меньшая w_1 и w_2. В самом деле, t или есть собственное начало t_1 и t_2, и тогда svt продолжается вправо минимум двумя способами: так, как оно продолжается в w_1, и так, как оно продолжается в w_2; или t совпадает, не ограничивая общности, с t_1, и тогда svt продолжается вправо неоднозначно, потому что w_1 неоднозначно продолжается вправо, а svt есть конец w_1. Аналогично svt неоднозначно продолжается влево. □

Лемма 4. $k_1 k_2 \cdots k_i \geq n$.

Пусть все развилки в u^∞ занумерованы $v_1, \ldots, v_i$, их кратности $k_1, \ldots, k_i$, кроме того, для каждой развилки v_j пронумерованы все ее продолжения вправо x_{jm} $m \in [1; k_j]$ (так, к примеру, нумерация продолжений пустой развилки есть просто нумерация всех букв, встречающихся в u). Рассмотрим некоторый циклический сдвиг слова u, то есть подслово w слова u^∞, такое, что $|w| = |u|$. Будем считать пустую развилку v_1 началом слова w. Посмотрим, продолжением с каким номером для развилки v_1 является первая буква слова w, обозначим номер за x_1. По лемме 3 для первой буквы однозначно определена минимальная развилка v_j, содержащая ее, то есть имеем $v_j = s_j t_j$, $w = t_j r_j$. Пусть первая буква слова r_j в списке продолжений развилки v_j имеет номер x_j, тогда по лемме 3 однозначно определена минимальная развилка, содержащая слово $v_j x_j$. Будем повторять этот процесс, пока минимальной развилкой не станет u^∞. Таким образом, для некоторых номеров $j \in J \subset [1; i]$ мы определили соответствующие им продолжения x_j, доопределим произвольным образом x_j, $j \notin J$. Из алгоритма построения очевидно, что выбором x_j однозначно определен сдвиг w слова u, а поскольку различных сдвигов n, лемма доказана. □

Рассмотрим дерево (связный граф без циклов), некоторую его вершину назовем корнем, ориентируем все ребра в направлении от корня (поскольку граф без циклов, направление определено корректно). Такой граф будем называть *ориентированным деревом.*

Лемма 5. *Количество запретов не меньше* $(k_1 - 1) + \cdots + (k_i - 1) + 1$.

Доказательство. Построим ориентированное дерево. Множество его вершин есть объединение множеств развилок и некоторых (возможно, не всех) запретов канонической системы, а исходящая степень каждой развилки в этом дереве есть в точности кратность этой развилки; вершины, соответствующие запретам — тупиковые.

Для построения дерева рассмотрим развилку v_j, $j \in [1; i]$ и ее продолжение x_m, $m \in [1; k_j]$, а также два различных продолжения влево y_1 и y_2 развилки v_j. Если слово $v_j x_m$ однозначно продолжается влево, то одно из слов $y_1 v_j x_m$ и $y_2 v_j x_m$ не встречается. Пусть без ограничения общности это $y_1 v_j x_m$. Тогда, поскольку и $y_1 v_j$, и $v_j x_m$ встречаются, то $y_1 v_j x_m$ — запрет. Каждый построенный таким образом запрет соответствует только одной паре (v_j, x_m). Если слово $v_j x_m$ неоднозначно продолжается влево, то минимальная развилка v_h, содержащая слово $v_j x_m$, имеет его своим началом.

Следовательно, v_h однозначно определяется парой (v_j, x_m). В самом деле, пусть при $v_{j_1} \neq v_j$ мы построили ту же развилку v_h, тогда из двух слов v_j и v_{j_1} одно является началом другого. Пусть v_j, тогда и $v_j x_m$ является началом v_{j_1}, что противоречит минимальности v_h.

Проведем из каждой развилки стрелки, соответствующие ее продолжениям вправо, в построенные ранее запреты или развилки.

В полученном дереве сумма исходящих степеней есть $k_1 + k_2 + \cdots + k_i$, v_1 имеет входящую степень 0, остальные $i-1$ развилка имеют входящую степень 1, значит, есть

$$k_1 + k_2 + \cdots + k_i - (i-1) = (k_1 - 1) + \cdots + (k_i - 1) + 1$$

висячих вершин, соответствующих запретам. $\square$

Приступим к доказательству теоремы 2. По лемме 4 кратности развилок $k_1, k_2, \ldots, k_i$ слова u^∞ удовлетворяют неравенству $k_1 k_2 \cdots k_i \geq n$. Откуда по лемме 2 имеем

$$(k_1 - 1) + \cdots + (k_i - 1) \geq \log_2 n.$$

А по лемме 5 число запретов не меньше

$$(k_1 - 1) + \cdots + (k_i - 1) + 1 \geq \log_2 n + 1.$$

$\square$

Теорема 3. *Существует возрастающая последовательность n_i такая, что*

$$r_n \leq \log_\phi n_i, \quad \text{где} \quad \phi = \frac{1+\sqrt{5}}{2}.$$

Доказательство. Построим бесконечную серию слов u, для которых $T(u^\infty) \leq \log_\phi n_i$, где $\phi = \frac{1+\sqrt{5}}{2}$. Рассмотрим алфавит из двух букв 0 и 1. Будем строить две вспомогательные последовательности слов s_i и t_i. Обозначим $s_1 = 0$, $t_1 = 1$; $t_{i+1} = s_i t_i$, $s_{i+1} = s_i s_i t_i$ $\forall i \in \mathbb{N}$. Рассмотрим систему запретов $v_{2i-1} = t_1 t_2 \cdots t_{i-1} t_i t_i$ и $v_{2i-1} = t_1 t_2 \cdots t_{i-1} s_i s_i s_i$ при $i \in [1; n-1]$ для некоторого $n \in \mathbb{N}$, и $v_{2n-1} = t_1 t_2 \cdots t_{n-1} t_n t_n$, $v_{2n} = t_1 t_2 \cdots t_{n-1} s_n s_n$.

Докажем, что система запретов $v_1, \ldots, v_{2n}$ задает слово $(t_{n+1})^\infty$.

Лемма 6. *Слово $t_1 t_2 \cdots t_{i-1}$ является концом любого слова, представимого в виде произведения слов s_i и t_i.*

Доказательство. Обоснование леммы проведем индукцией по i. База индукции при $i = 2$ очевидна. Выполним шаг индукции. Пусть слово w представимо в виде произведения слов s_i и t_i. Тогда оно также представимо в виде произведения слов s_{i-1} и t_{i-1}, причем последним словом в произведении будет t_{i-1}. Обозначим $w = w_1 t_{i-1}$, тогда то, что слово $t_1 t_2 \cdots t_{i-1}$ является концом слова w, равносильно тому, что слово $t_1 t_2 \cdots t_{i-2}$ является концом слова w_1 (представимого в виде произведения слов s_{i-1} и t_{i-1}, как отмечалось выше), но это в точности индукционное предположение. $\square$

Лемма 7. *Всякое бесконечное слово, не содержащее запретов $v_1, \ldots, v_n$, есть $(t_{n+1})^\infty$.*

Доказательство. Индукцией по i докажем, что любое бесконечное слово, удовлетворяющее запретам $v_1, \ldots, v_{2i}$, $i < n$, разбивается на слова s_{i+1} и t_{i+1}. База очевидна. Пусть утверждение доказано для $i-1$, докажем его для i. Бесконечное слово w разбивается на слова s_i и t_i по предположению индукции. Докажем, что разбиение не может содержать два слова t_i подряд. Предположив противное, обозначим $w = w_1 t_i t_i w_2$, где w_1 само некоторым образом разбито на слова s_i и t_i, тогда по лемме 6 слово $t_1 t_2 \cdots t_{i-1}$ есть конец w_1, а значит, слово $v_{2i-1} = t_1 t_2 \cdots t_{i-1} t_i t_i$ есть подслово w, что противоречит условию.

Аналогично доказывается, что разбиение слова w на слова s_i и t_i не может содержать три слова s_i подряд. Следовательно, в разбиении слова w слова s_i и t_i можно объединить в слова s_{i+1} и t_{i+1}.

Аналогично доказывается, что бесконечное слово, не содержащее запретов $v_1, \ldots, v_n$, разбивается на слова (t_{n+1}), то есть что оно есть $(t_{n+1})^\infty$. $\square$

Очевидно, что слово $(t_{n+1})^\infty$ удовлетворяет системе запретов $v_1, \ldots, v_{2n}$. Заметим, что слова t_i и s_i являются конечными шагами в итерационном построении равномернорекуррентного слова Штурма (Sturmian prime word, см. [5],[6]), связанного с Bergman's gap, см. [3],[6].

Рассмотрим последовательность f_i, заданную условиями $f_1 = f_2 = 1$, $f_i = f_{i-1} + f_{i-2}$ при $i > 2$. Легко видеть, что $\forall i \quad f_i$ и f_{i+1} взаимнопросты. Поскольку слово t_{n+1} содержит f_{2n-1} букв 1 и f_{2n} букв 0 (это утверждение легко доказывается индукцией по n вместе с утверждением, что s_{n+1} содержит f_{2n} букв 1 и f_{2n+1} букв 0), t_{n+1} не является степенью большей 1 никакого слова, а значит, t_{n+1} есть минимальный период слова $(t_{n+1})^\infty$. Длина периода, таким образом, равна $f_{2n+1} = \frac{1}{\sqrt{5}}(\phi^{2n+1} + \psi^{2n+1})$, где $\phi = \frac{1+\sqrt{5}}{2}$, $\psi = \frac{1-\sqrt{5}}{2}$.

При этом, поскольку последовательность $(t_{n+1})^\infty$ задана $2n$ запретами, $T((t_{n+1})^\infty) \leq 2n$.

□

Автор выражает благодарность А. Я. Белову за постановку задачи, И. И. Богданову, за ценные замечания, приведшие к существенному упрощению многих доказательств, а также В. Л. Дольникову за помощь в работе над статьей.

Список литературы

1. *Уфнаровский, В.А.* Комбинаторные и асимптотические методы в алгебре / *В.А. Уфнаровский.* // Итоги науки и техники. Сер. Совр. пробл. математики. Фундаментальные направления. Т.57. — М.: ВИНИТИ, 1990. — С. 5 – 177.

2. *Курош, А.Г.* Проблемы теории колец, связанные с проблеммой Бернсайда о периодических группах / *А.Г. Курош* // Изв. АН СССР сер. мат. — 1941. — Т. 5. — С. 233 – 240.

3. *Белов, А.Я.* Мономиальные алгебры / *А.Я. Белов, В.В. Борисенко, В.Н. Латышев* // Итоги науки и техники. Сер. Современная математика и ее приложения. Тематические обзоры. Т.26. М.: ВИНИТИ, 2002. — С.35 – 214.

4. *Уилсон, Р.* Введение в теорию графов / *Р. Уилсон.* — М.:Мир, 1977. — 208 с.

5. *Allouche, J.-P.* Automatic sequences. Theory, applications, generalizations / *J.-P. Allouche, J. Shallit.* — Cambridge: Cambridge Univercity Press, 2003. — 571 p.

6. *Bell, J.P.* Examples in finite Gel'fand-Kirilov dimension / *J.P. Bell* // J. Algebra. — 2003. — 263, no. 1. — P. 159 – 175.

On the Number of Restrictions Determining a Periodical Sequence

Chelnokov G.R.

We consider sequences W of the period u over an alphabet consisting of l letters. It is required to determine unambiguously the sequence W picking out words which are not subwords of the sequence. For $n \in \mathbb{N}$ we denote by U_n the set of words u of length n, which are not powers (i.e. are not represented in form $u = v^k$ $k > 1$).

Let $T(u^\infty)$ be the minimal number of restrictions determining the sequence u^∞.

Denote

$$m_n = \max_{u \in U_n} T(u^\infty), \quad r_n = \min_{u \in U_n} T(u^\infty).$$

We prove that

1. $m_n \leq n(l-1)$.

The estimate is precise for infinite values of n. For instance, it takes place for a period which contains all the words of some given length t (i.e. $n = l^t$).

2. $r_n \geq \log_2 n + 1$.

3. There exists an increasing sequence n_i so that

$$r_{n_i} \leq \log_\phi n_i, \quad \text{where} \quad \phi = \frac{1+\sqrt{5}}{2}.$$