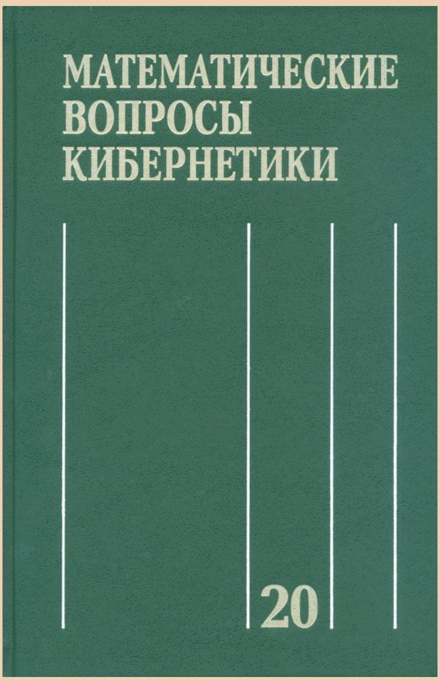

Р. М. Колпаков

Некоторые результаты о возможном числе периодических факторов в словах

Рекомендуемая форма библиографической ссылки: Колпаков Р. М. Некоторые результаты о возможном числе периодических факторов в словах // Математические вопросы кибернетики. Вып. 20. — М.: ФИЗМАТЛИТ, 2022. — С. 25–40. URL: http://library.keldysh.ru/mvk.asp?id=2022-25
DOI: 10.20948/mvk-2022-25

НЕКОТОРЫЕ РЕЗУЛЬТАТЫ О ВОЗМОЖНОМ ЧИСЛЕ ПЕРИОДИЧЕСКИХ ФАКТОРОВ В СЛОВАХ

Р. М. КОЛПАКОВ

(МОСКВА)

Пусть $w = w[1]w[2]\ldots w[n]$ — произвольное формальное слово длины $|w|=n$. Фрагмент $w[i]w[i+1]\ldots w[j]$ слова w, где $1\leqslant i\leqslant j\leqslant n$, называется *фактором* слова w и обозначается через $w[i..j]$. Для любого $i=1,\ldots,n$ фактор $w[1..i]$ ($w[i..n]$) называется *префиксом* (*суффиксом*) слова w. Если два фактора v', v'' слова w являются фактически одним и тем же фрагментом слова w, мы обозначаем это через $v'\equiv v''$. Отметим, что любые два фактора $w[i'..j']$, $w[i''..j'']$ такие, что $i'\neq i''$ или $j'\neq j''$, рассматриваются как различные факторы слова w даже если эти факторы являются одинаковыми словами. Таким образом, через $w[i'..j']\equiv w[i''..j'']$ мы обозначаем то, что $w[i'..j']$ и $w[i''..j'']$ являются одним и тем же фактором, т. е. $i'=i''$ и $j'=j''$, в то время как через $w[i'..j']=w[i''..j'']$ мы обозначаем то, что $w[i'..j']$ и $w[i''..j'']$ являются одинаковыми словами. Если некоторое слово u совпадает с некоторым фактором v слова w, то v называется *вхождением* слова u в w. Натуральное число $p\leqslant|w|$ называется *периодом* слова w, если $w[i]=w[i+p]$ для каждого $i=1,\ldots,n-p$. Мы обозначаем через $p(w)$ минимальный период слова w и через $e(w)$ отношение $|w|/p(w)$, которое называется *порядком* слова w. Значение $re(w)=e(w)-1$ называется *сокращенным порядком* слова w. Например, для слова $w=aabaa$ мы имеем $p(w)=3$, $e(w)=5/3$ и $re(w)=2/3$. Слово называется *периодическим*, если его порядок не меньше, чем 2. Слово называется *примитивным*, если это слово не является периодическим словом целого порядка. Вхождения периодических слов в некотором слове называются *периодичностями* в этом слове.

Пусть r — некоторая периодичность. Любой фактор длины $p(r)$ в r называется *циклическим корнем* периодичности r. Например, периодичность $abaabaab$ имеет различные циклические корни aba, baa и aab. Две периодичности называются *периодичностями с одинаковыми циклическими корнями*, если они имеют одинаковые множества различных циклических корней. Например, периодичности $abaabaab$ и $aabaabaaba$ являются периодичностями с одинаковыми циклическими корнями.

Периодичность в некотором слове называется *максимальной*, если эта периодичность не может быть расширена в этом слове ни на один символ ни вправо, ни влево с сохранением ее минимального периода. Более строго,

DOI: 10.20948/mvk-2022-25

периодичность $r \equiv w[i..j]$ в w называется *максимальной*, если она удовлетворяет следующим условиям:

1) если $i > 1$, то $w[i-1] \neq w[i-1+p(r)]$;

2) если $j < n$, то $w[j+1-p(r)] \neq w[j+1]$.

Например, слово, показанное на рис. 1, имеет 6 различных максимальных периодичностей, которые подчеркнуты на этом рисунке линиями. Посколь-

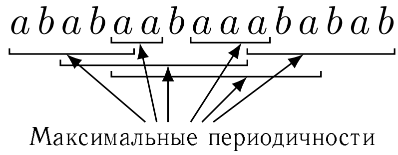

Рис. 1. Максимальные периодичности в слове

ку максимальные периодичности в слове содержат в себе любые другие периодичности произвольного типа в этом слове, множество всех максимальных периодичностей в слове является компактным представлением всех периодичностей в слове, имеющем много полезных приложений (см., например, [6]). Для любого слова w обозначим через $R(w)$ ($E(w)$) число (сумму порядков) всех максимальных периодичностей в слове w, и для любого n положим $R(n) = \max\limits_{|w|=n} R(w)$ ($E(n) = \max\limits_{|w|=n} E(w)$). Линейная верхняя оценка для $E(n)$ была получена в [16].

Теорема 1 [16]. $E(n) = O(n)$.

Поскольку, очевидно, $2R(n) \leqslant E(n)$, из данной теоремы вытекает следующий факт.

Следствие 1. $R(n) = O(n)$.

В результате многочисленных исследований [2–4, 7, 14, 21, 23, 24] в данной области был получен ряд более точных верхних оценок для $E(n)$ и $R(n)$. Революционный результат в данном направлении был получен в [1], где были доказаны оценки $R(n) < n$ и $E(n) < 3n$. Насколько нам известно, наилучшая к настоящему времени верхняя оценка $R(w) \leqslant \frac{22}{23}n$ для бинарных слов w длины n установлена в [10]. Схожая проблема оценки возможного числа *кубических* максимальных периодичностей, т. е. максимальных периодичностей порядка не меньшего, чем 3, в словах фиксированной длины рассматривалась в [5]. Для этого числа получена верхняя оценка $n/2$. В [22] была вычислена предельная асимптотика ожидаемого числа максимальных периодичностей в слове длины n при $n \to \infty$. По нашим сведениям, наилучшие к настоящему времени нижние оценки $2{,}035n$ для $E(n)$ и $0{,}9445757n$ для $R(n)$ были получены соответственно в [7] и [25].

Значения $E(n)$ и $R(n)$ могут быть обобщены на случай максимальных периодичностей с достаточно большими минимальными периодами следующим образом. Пусть $\lambda = 1, 2, \ldots$. Обозначим через $R_{\geqslant\lambda}(w)$ число максимальных периодичностей с минимальными периодами, не меньшими, чем λ, в слове w и положим $R_{\geqslant\lambda}(n) = \max\limits_{|w|=n} R_{\geqslant\lambda}(w)$, т. е. $R_{\geqslant\lambda}(n)$ является максимальным возможным числом максимальных периодичностей, минимальные периоды которых не меньше, чем λ, в словах длины n. Очевидно, что $R(n) = R_{\geqslant 1}(n) \geqslant R_{\geqslant 2}(n) \geqslant \ldots$. Для максимальных периодичностей могут быть естественным образом предположены следующие гипотезы:

1. $R_{\geqslant\lambda}(n) \leqslant cn$, где $c \to 0$, при $\lambda \to \infty$.

2. Число максимальных периодичностей, содержащих одну и ту же букву*), в слове w равно $o(|w|)$ при $|w| \to \infty$.

К сожалению, обе гипотезы являются неверными. В качестве контрпримера можно рассмотреть слово $w_k = (01)^k0(01)^k0 = (01)^k00(10)^k$ длины $4k+2$. Нетрудно увидеть, что w_k содержит $k+3$ максимальных периодичностей, указанных на рис. 2. Легко заметить, что из этих максимальных периодично-

$$w_k = 010101\ldots0101001010\ldots101010$$

$$\cdots\cdots\cdots\cdots\cdots$$

Рис. 2. Максимальные периодичности в слове w_k

стей $k+1-\lfloor\lambda/2\rfloor$ периодичностей имеют минимальные периоды не меньше, чем λ, поэтому**) $R_{\geqslant\lambda}(w_k) = k+1-\lfloor\lambda/2\rfloor \gtrsim k \gtrsim |w_k|/4$ при $k \to \infty$. Таким образом,

$$R_{\geqslant 1}(n) \geqslant R_{\geqslant 2}(n) \geqslant \ldots \gtrsim n/4.$$

Кроме того, средние буквы слова w_k содержатся в $k+2 > |w_k|/4$ максимальных периодичностях. С другой стороны, легко увидеть, что в действительности w_k имеет две смежные максимальные периодичности $(01)^k0$, в то время как другие максимальные периодичности в w_k образуют серию вложенных друг в друга периодичностей, начинающихся в левой периодичности $(01)^k0$ и заканчивающихся в правой периодичности $(01)^k0$. Это может быть представлено в терминах порождения периодичностей, введенного в [13], в частности порождения серии максимальных периодичностей максимальными периодичностями $(01)^k0$. Строго говоря, мы говорим, что в слове w две смежные и, возможно, пересекающиеся максимальные периодичности $r' \equiv w[i'..j']$, $r'' \equiv w[i''..j'']$ с одинаковыми циклическими корнями *порождают* максимальную периодичность $r \equiv w[i..j]$, если $p(r) \geqslant 3p$, $i' < i \leqslant j'$ и $i'' \leqslant j < j''$, где p — минимальный период периодичностей r' и r'' (см. рис. 3). Максимальная периодичность называется *вторичной*, если она порождается другими максимальными периодичностями в слове. Максимальная периодичность называется *первичной*, если она не является вторичной. Например, на рис. 4 указаны вторичные периодичности в слове w_k. Можно показать, что любая вторичная периодичность порождается единственной парой первичных периодичностей. Заметим, что первичные периодичности в слове однозначно определяют все вторичные периодичности и тем самым однозначно определяют все остальные максимальные периодичности в слове. Для

*) Здесь буквы, находящиеся в различных позициях слова, считаются различными, даже если эти буквы являются одинаковыми.

**) Напомним, что $f(n) \gtrsim g(n)$ означает, что $f(n) \geqslant g(n) - o(g(n))$ при $n \to \infty$.

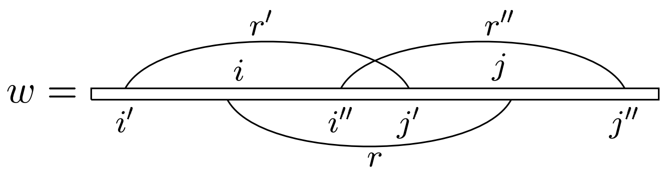

Рис. 3. Порождение максимальных периодичностей

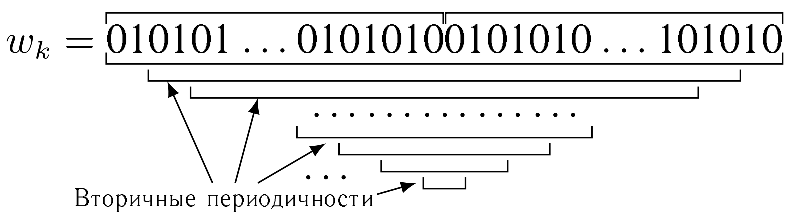

Рис. 4. Вторичные периодичности в w_k

любого натурального λ обозначим через $Rp_{\geqslant\lambda}(n)$ максимальное возможное число первичных периодичностей с минимальными периодами не меньшими, чем λ, в словах длины n, через $Ep_{\geqslant\lambda}(n)$ — максимальную возможную сумму порядков первичных периодичностей с минимальными периодами, не меньшими, чем λ, в словах длины n, и через $Eps_{\geqslant\lambda}(n)$ — максимально возможную сумму порядков первичных периодичностей с минимальными периодами, не меньшими, чем λ, и вторичных периодичностей, порождаемых этими первичными периодичностями, в словах длины n. Очевидно, что $Eps_{\geqslant\lambda}(n)\geqslant Ep_{\geqslant\lambda}(n)\geqslant 2Rp_{\geqslant\lambda}(n)$. В [18] доказано, что $Eps_{\geqslant\lambda}(n)=O(n/\lambda)$, так $Ep_{\geqslant\lambda}(n)=O(n/\lambda)$ и $Rp_{\geqslant\lambda}(n)=O(n/\lambda)$. Таким образом, число первичных периодичностей в слове удовлетворяет гипотезе 1. Более того, в [18] также доказано, что в слове длины n одна и та же буква содержится в не более чем $O\left(\log\frac{n}{\lambda}\right)$ первичных периодичностей с минимальными периодами, не меньшими, чем λ, тем самым первичные периодичности в слове удовлетворяют также гипотезе 2. Таким образом, первичные периодичности в слове являются более удобными для рассмотрения и анализа, чем все максимальные периодичности в слове. Также было показано, что порядок любой вторичной периодичности меньше, чем $7/3$, т. е. любая максимальная периодичность, порядок которой не меньше, чем $7/3$, является первичной. Таким образом, мы получаем, что в слове длины n содержится $O(n/\lambda)$ максимальных периодичностей, минимальные периоды которых не меньше, чем λ, и порядки которых не меньше, чем $7/3$, при этом не более чем $O\left(\log\frac{n}{\lambda}\right)$ таких периодичностей могут содержать один и тот же символ в данном слове.

Естественным обобщением периодичностей являются факторы слова, порядки которых строго меньше, чем 2, но не меньше некоторого порогового значения, большего единицы. Мы называем такие факторы *субпериодичностями*. Более строго, фактор r называется *субпериодичностью* (*δ-subrepetition*), если $1 < e(r) < 2$ $(1+\delta\leqslant e(r) < 2)$. Заметим, что аналогично понятию максимальной периодичности мы также можем ввести понятие максимальной субпериодичности: субпериодичность в слове называется *максимальной*, если она не может быть расширена в этом слове ни на один символ ни вправо, ни влево с сохранением ее минимального периода (см. рис. 5).

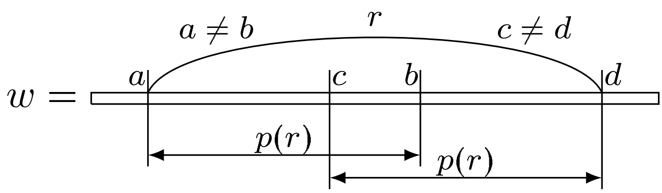

Рис. 5. Максимальная субпериодичность

Другим обобщением периодичностей являются *разрывные повторы*, являющиеся тесно связанными с субпериодичностями. Разрывными повторами являются факторы вида uvu, где u и v — непустые слова. Первый (второй) фактор u в разрывном повторе uvu называется *левой (правой) копией*, а фактор v называется *разрывом*. Под *периодом* разрывного повтора uvu понимается величина $|u|+|v|$ (заметим, что, вообще говоря, период разрывного повтора uvu может быть больше минимального периода слова uvu). Для разрывного повтора σ мы обозначаем длину копий повтора σ через $c(\sigma)$ и период повтора σ через $p(\sigma)$ (см. рис. 6). Под *порядком* $\widehat{e}(\sigma)$ разрывного

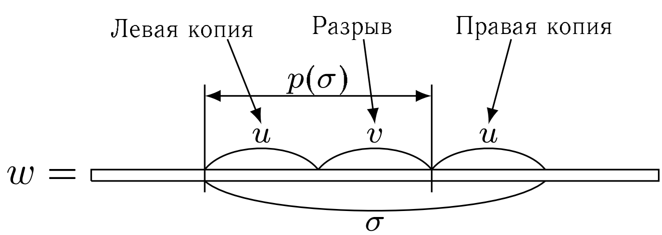

Рис. 6. Разрывный повтор σ в слове w

повтора $\sigma = uvu$ понимается величина $\frac{|\sigma|}{p(\sigma)} = 1 + \frac{c(\sigma)}{p(\sigma)}$, а под *сокращенным порядком* $r\widehat{e}(\sigma)$ повтора σ понимается величина $\widehat{e}(\sigma) - 1 = \frac{c(\sigma)}{p(\sigma)}$ (поскольку $p(\sigma)$ может быть больше минимального периода слова uvu, величина $\widehat{e}(\sigma)$ может быть меньше порядка слова uvu). Отметим, что различные разрывные повторы, имеющие разные периоды, могут образовывать один и тот же фактор в слове. Такие повторы рассматриваются как различные. Для любого вещественного $\alpha > 1$ разрывный повтор σ называется *α-разрывным*, если $p(\sigma) \leqslant \alpha c(\sigma)$.

Аналогично периодичностям разрывный повтор в слове называется *максимальным*, если его копии не могут быть расширены в этом слове ни на один символ ни вправо, ни влево с сохранением периода повтора (см. рис. 7). Например, в слове, изображенном на рис. 8, разрывные повторы $u'v'u'$ и $u''v''u''$ являются максимальными.

Заметим, что любой α-разрывный повтор $\sigma = uvu$ однозначным образом расширяется либо до максимального α-разрывного повтора $\sigma' = u'v'u'$ с тем

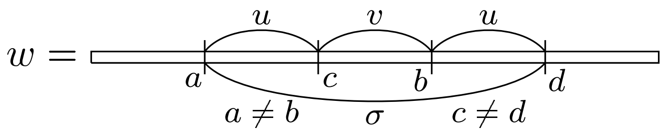

Рис. 7. Максимальный разрывный повтор σ в слове w

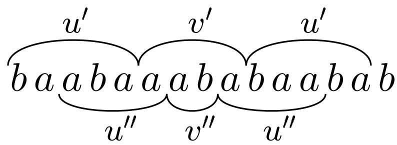

Рис. 8. Максимальные повторы в слове

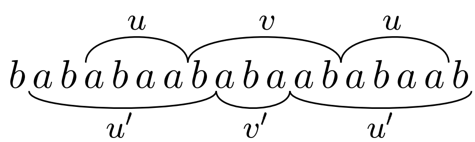

Рис. 9. Повтор, расширяемый до максимального разрывного повтора

же самым периодом (см. рис. 9), либо до максимальной периодичности r такой, что $p(r)$ является делителем периода $p(\sigma)$ (см. рис. 10).

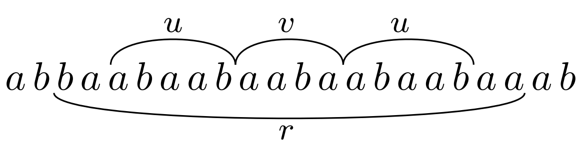

Рис. 10. Повтор, расширяемый до максимальной периодичности

Пусть r — некоторая максимальная δ-субпериодичность в слове w. Заметим, что мы можем также рассматривать r как разрывный повтор σ с периодом $p(r)$, левая (правая) копия которого является префиксом (суффиксом) длины $|r| - p(r)$ в слове w (см. рис. 11). Легко увидеть, что σ

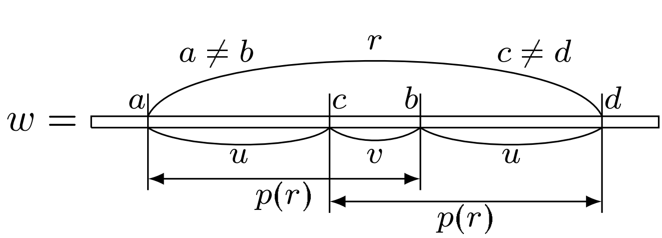

Рис. 11. Повтор, сопоставленный максимальной субпериодичности

является максимальным $\frac{1}{\delta}$-разрывным повтором. Мы называем субпериодичность r и повтор σ *сопоставленными* друг другу. Заметим, что сопоставленные друг другу субпериодичность r и повтор σ однозначно определяются друг из друга и $p(\sigma)=p(r)$, поэтому $\widehat{e}(\sigma)=e(r)$ $(r\widehat{e}(\sigma)=re(r))$.Таким образом, любая максимальная δ-субпериодичность соответствует сопоставленному ей максимальному $\frac{1}{\delta}$-разрывному повтору. Вместе с тем для максимального разрывного повтора может не быть сопоставленной ему максимальной субпериодичности. Например, в слове, изображенном на рис. 12, разрывный повтор uvu сопоставляется максимальной субпериодичности r, в то время как разрывный повтор $u'v'u'$ не сопоставлен r, поэтому повтор $u'v'u'$ не имеет сопоставленной ему максимальной субпериодичности. Мак-

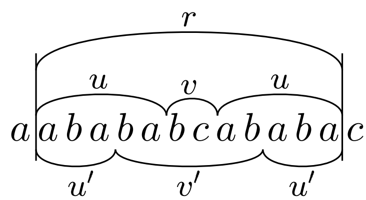

Рис. 12. Принципиальные и непринципиальные повторы

симальные разрывные повторы сопоставленные максимальным субпериодичностям называются *принципиальными*, в противном случае они называются *непринципиальными*. Таким образом, на рис. 12 повтор uvu является принципиальным, в то время как повтор $u'v'u'$ является непринципиальным. Поскольку имеется взаимно-однозначное соответствие между максимальными δ-субпериодичностями и принципиальными $\frac{1}{\delta}$-разрывными повторами, мы получаем следующий факт.

Утверждение 1. *В любом слове число максимальных субпериодичностей (δ-субпериодичностей) не превосходит числа максимальных разрывных повторов ($\frac{1}{\delta}$-разрывных повторов).*

Разрывный повтор uvu называется *примитивным*, если слово uv является примитивным. Например, на рис. 13 повтор $\sigma' \equiv u'v'u'$ является максимальным непримитивным разрывным повтором, а повтор $\sigma'' \equiv u''v''u''$ — максимальным примитивным разрывным повтором. Легко заметить, что для

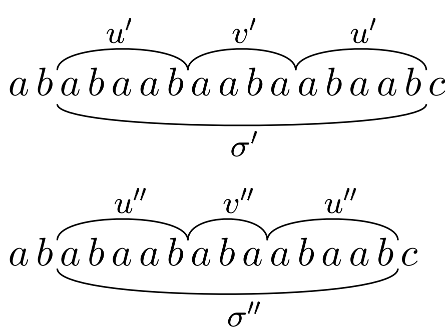

Рис. 13. Примитивные и непримитивные повторы в двух различных словах

любого непримитивного разрывного повтора $\sigma' \equiv u'v'u'$ фактор $u'v'u'$ является максимальной периодичностью с минимальным периодом, равным $p(u'v')$.

Мы говорим, что повтор σ' *индуцируется* данной максимальной периодичностью. Например, на рис. 14 повтор $\sigma' \equiv u'v'u'$ индуцируется периодичностью r. Таким образом, любой непримитивный разрывный повтор индуцируется некоторой максимальной периодичностью. С другой стороны, очевидно,

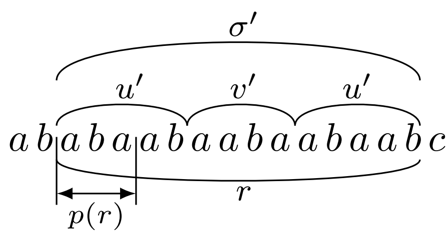

Рис. 14. Непримитивный повтор и индуцирующая его периодичность

что любая максимальная периодичность r индуцирует не более чем $\lceil e(r)/2\rceil$ максимальных непримитивных разрывных повторов. Таким образом, из теоремы 1 вытекает, что слово длины n содержит не более чем $O(n)$ максимальных непримитивных разрывных повторов. Отметим, что любой принципиальный максимальный разрывный повтор является примитивным, однако примитивный максимальный разрывный повтор также может быть непринципиальным (см., например, повтор σ' на рис. 12).

Для любого слова w обозначим через $K^m_\delta(w)$ $(K^m(w))$ множество всех максимальных $\frac{1}{\delta}$-разрывных повторов (разрывных повторов) в w, через $K^p_\delta(w)$ $(K^p(w))$ множество всех максимальных примитивных $\frac{1}{\delta}$-разрывных повторов (разрывных повторов) в w и через $K^s_\delta(w)$ $(K^s(w))$ множество всех принципиальных максимальных $\frac{1}{\delta}$-разрывных повторов (разрывных повторов) в w. Как было отмечено выше, имеются взаимно-однозначные соответствия между множествами $K^s_\delta(w)$ и множествами всех максимальных δ-субпериодичностей и между множеством $K^s(w)$ и множеством всех максимальных субпериодичностей в w. Из сделанных нами выше наблюдений мы заключаем, что

$$K^s_\delta(w) \subseteq K^p_\delta(w) \subseteq K^m_\delta(w) \quad (K^s(w) \subseteq K^p(w) \subseteq K^m(w)). \tag{1}$$

Для любого слова w обозначим через $RE^m(w)$ сумму сокращенных порядков всех максимальных разрывных повторов в w, через $RE^p(w)$ сумму сокращенных порядков всех максимальных примитивных разрывных повторов в w, и через $RE^s(w)$ сумму сокращенных порядков всех максимальных субпериодичностей w. Напомним, что для любых сопоставленных друг другу разрывного повтора σ и субпериодичности r мы имеем $\widehat{re}(\sigma) = re(r)$, поэтому $RE^s(w)$ фактически является суммой сокращенных порядков всех принципиальных максимальных разрывных повторов в w. Таким образом, из соотношений (1) вытекает, что

$$RE^s(w) \leqslant RE^p(w) \leqslant RE^m(w).$$

Согласно теореме 1 из [17], $RE^p(w) \leqslant n\ln n$ для любого слова w длины n. Аналогичным образом можно доказать такую же верхнюю оценку для $RE^m(w)$.

Т е о р е м а 2. *$RE^m(w) < n\ln n$ для любого слова w длины n.*

Д о к а з а т е л ь с т в о. Для каждого максимального разрывного повтора σ из $RE^m(w)$ мы разлагаем величину $\widehat{re}(\sigma) = \frac{c(\sigma)}{p(\sigma)}$ на $c(\sigma)$ слагаемых, соответствующих парам $(w[i], w[i+p(\sigma)])$ одинаковых символов в повторе σ. Каждое такое слагаемое дает в сумму $RE^m(w) = \sum\limits_{\sigma\in K^m(w)} \widehat{re}(\sigma)$ вклад, равный $\frac{1}{p(\sigma)}$. Рассмотрим в w две позиции i и j, где $1 \leqslant i < j \leqslant n$. Если $w[i] = w[j]$, то эта пара одинаковых символов соответствует слагаемому, включенному в величину $\widehat{re}(\sigma)$ для некоторого однозначным образом определенного максимального разрывного повтора σ такого, что $w[i]$ и $w[j]$ являются соответствующими символами левой и правой копий повтора σ соответственно. Заметим, что в этом случае $p(\sigma) = j - i$, поэтому пара $(w[i], w[j])$ дает вклад в $RE^m(w)$, равный величине $\frac{1}{p(\sigma)} = \frac{1}{j-i}$. Учитывая всевозможные пары различных символов в w, мы получаем верхнюю оценку

$$RE^m(w) \leqslant \sum_{1\leqslant i<j\leqslant n} \frac{1}{j-i} = \sum_{p=1}^{n-1} \frac{n-p}{p} = n\sum_{p=1}^{n-1}\frac{1}{p} - (n-1).$$

Поскольку функция $f(x) = \frac{1}{x}$ является выпуклой вниз для $x > 0$, мы имеем

$$\sum_{p=2}^{n-1}\frac{1}{p} \leqslant \int\limits_{3/2}^{n-1/2} \frac{dx}{x}.$$

Кроме того, поскольку $f'(1) = -1$, мы также имеем $\int\limits_{1}^{3/2} \frac{dx}{x} \geqslant 3/8$. Таким образом,

$$\sum_{p=1}^{n-1}\frac{1}{p} \leqslant \frac{5}{8} + \int\limits_{1}^{n-1/2} \frac{dx}{x} = \frac{5}{8} + \ln(n-1/2) < \frac{5}{8} + \ln n.$$

Следовательно,

$$n\sum_{p=1}^{n-1}\frac{1}{p} - (n-1) < n\ln n + (\frac{5}{8}n - (n-1)) = n\ln n + (1 - \frac{3}{8}n) < n\ln n$$

для $n > 2$. Таким образом, $RE^m(w) < n\ln n$.

Отметим, что разрывный повтор σ является α-разрывным повтором, если $\frac{p(\sigma)}{c(\sigma)} \leqslant \alpha$, т. е. $\widehat{re}(\sigma) = \frac{c(\sigma)}{p(\sigma)} \geqslant 1/\alpha$. Поэтому из теоремы 2 вытекает следующий факт.

С л е д с т в и е 2. *Число всех максимальных α-разрывных повторов в слове длины n меньше, чем $\alpha n\ln n$.*

Из утверждения 1 и следствия 2 мы получаем следующую верхнюю оценку для числа максимальных δ-субпериодичностей в слове.

С л е д с т в и е 3. *Число всех максимальных δ-субпериодичностей в слове длины n меньше, чем $(n\ln n)/\delta$.*

Теорема 2 может быть обобщена на случай максимальных разрывных повторов с ограниченными периодами. Пусть $\lambda = 2, 3, \ldots$. Обозначим через $RE^m_{\geqslant\lambda}(w)$ $(RE^m_{\leqslant\lambda}(w))$ сумму сокращенных порядков всех максимальных разрывных повторов с периодами, не меньшими, чем λ (не большими, чем λ) в слове w. Обозначим также через $RE^p_{\geqslant\lambda}(w)$ ($RE^p_{\leqslant\lambda}(w)$) сумму сокращенных порядков всех максимальных примитивных разрывных повторов с периодами, не меньшими, чем λ (не большими, чем λ) в слове w. В ([17], следствия 2 и 3) были получены верхние оценки для $RE^p_{\geqslant\lambda}(w)$ и $RE^p_{\leqslant\lambda}(w)$ для любого слова w. Используя аналогичное доказательство, мы можем получить те же самые верхние оценки для $RE^m_{\geqslant\lambda}(w)$ и $RE^m_{\leqslant\lambda}(w)$.

Т е о р е м а 3. $RE^m_{\geqslant\lambda}(w) < n\ln(n/\lambda)$ *для любого слова* w *длины* $n > \lambda$.

Д о к а з а т е л ь с т в о. Заметим, что для любого максимального разрывного повтора σ с периодом, не меньшим, чем λ, каждая пара одинаковых символов $(w[i], w[j])$, соответствующая слагаемому, включенному в величину $\widehat{re}(\sigma)$, удовлетворяет ограничению $j - i = p(\sigma) \geqslant \lambda$. Таким образом, аналогично доказательству теоремы 2, верхняя оценка для $RE^m_{\geqslant\lambda}(w)$ может быть модифицирована следующим образом:

$$RE^m_{\geqslant\lambda}(w) \leqslant \sum_{1\leqslant i<i+\lambda\leqslant j\leqslant n} \frac{1}{j-i} = \sum_{p=\lambda}^{n-1} \frac{n-p}{p} = n\sum_{p=\lambda}^{n-1} \frac{1}{p} - (n-\lambda).$$

Поскольку функция $f(x) = \frac{1}{x}$ является выпуклой вниз для $x > 0$, мы имеем

$$\sum_{p=\lambda+1}^{n-1} \frac{1}{p} \leqslant \int\limits_{\lambda+1/2}^{n-1/2} \frac{dx}{x}.$$

Кроме того, поскольку $f'(\lambda) = -\frac{1}{\lambda^2}$, мы также имеем

$$\int\limits_{\lambda}^{\lambda+1/2} \frac{dx}{x} \geqslant \frac{1}{2}\cdot\frac{1}{2}\left(\frac{1}{\lambda} + \left(\frac{1}{\lambda} - \frac{1}{2\lambda^2}\right)\right) = \frac{4\lambda-1}{8\lambda^2}.$$

Таким образом,

$$\sum_{p=\lambda}^{n-1} \frac{1}{p} \leqslant \frac{4\lambda+1}{8\lambda^2} + \int\limits_{\lambda}^{n-1/2} \frac{dx}{x} = \frac{4\lambda+1}{8\lambda^2} + \ln\left(\frac{n-1/2}{\lambda}\right) < \frac{4\lambda+1}{8\lambda^2} + \ln\left(\frac{n}{\lambda}\right).$$

Следовательно,

$$n\sum_{p=\lambda}^{n-1} \frac{1}{p} - (n-\lambda) < n\ln\left(\frac{n}{\lambda}\right) + n\frac{4\lambda+1}{8\lambda^2} - (n-\lambda) = n\ln\left(\frac{n}{\lambda}\right) + \left(\lambda - n\frac{8\lambda^2-4\lambda-1}{8\lambda^2}\right).$$

Легко проверить, что $\lambda - n\frac{8\lambda^2-4\lambda-1}{8\lambda^2} \leqslant 0$ для $n > \lambda > 1$. Таким образом, $RE^m_{\geqslant\lambda}(w) < n\ln\left(\frac{n}{\lambda}\right)$.

Т е о р е м а 4. $RE^m_{\leqslant\lambda}(w) < n(1 + \ln\lambda)$ *для любого слова* w *длины* n.

Д о к а з а т е л ь с т в о. Заметим, что в этом случае каждая пара одинаковых символов $(w[i], w[j])$, соответствующая слагаемому, которое вносит вклад в $RE^m_{\leqslant\lambda}(w)$, удовлетворяет ограничению $j - i \leqslant \lambda$. Таким образом, аналогично доказательству теоремы 2, верхняя оценка для $RE^m_{\leqslant\lambda}(w)$ может быть модифицирована следующим образом:

$$RE^m_{\leqslant\lambda}(w) \leqslant \sum_{1\leqslant i<j\leqslant \min(i+\lambda,n)} \frac{1}{j-i} = \sum_{p=1}^{\lambda} \frac{n-p}{p} = n\sum_{p=1}^{\lambda}\frac{1}{p} - \lambda < n\sum_{p=1}^{\lambda}\frac{1}{p}.$$

Поскольку функция $f(x) = \frac{1}{x}$ является убывающей для $x > 0$, мы имеем

$$\sum_{p=2}^{\lambda}\frac{1}{p} \leqslant \int_1^{\lambda}\frac{dx}{x} = \ln\lambda.$$

Таким образом,

$$n\sum_{p=1}^{\lambda}\frac{1}{p} \leqslant n(1+\ln\lambda),$$

поэтому $RE^m_{\leqslant\lambda}(w) < n(1+\ln\lambda)$.

Аналогично следствиям 2 и 3 из теоремы 2 мы имеем следующие следствия из теорем 3 и 4.

Следствие 4. *Число всех максимальных α-разрывных повторов с периодами, не меньшими, чем λ, в слове длины n меньше, чем $\alpha n\ln(n/\lambda)$.*

Следствие 5. *Число всех максимальных δ-субпериодичностей с минимальными периодами, не меньшими, чем λ, в слове длины n меньше, чем $n\ln(n/\lambda)/\delta$.*

Следствие 6. *Число всех максимальных α-разрывных повторов с периодами, не большими, чем λ, в слове длины n меньше, чем $\alpha n(1+\ln\lambda)$.*

Следствие 7. *Число всех максимальных δ-субпериодичностей с минимальными периодами, не большими, чем λ, в слове длины n меньше, чем $n(1+\ln\lambda)/\delta$.*

Чтобы получить нижние оценки для сумм сокращенных порядков всех максимальных разрывных порядков для общего случая неограниченного алфавита, положим $RE^m(n) = \max_{|w|=n} RE^m(w)$, $RE^p(n) = \max_{|w|=n} RE^p(w)$ и $RE^s(n) = \max_{|w|=n} RE^s(w)$. Тогда мы очевидным образом получаем $RE^s(n) \leqslant RE^p(n) \leqslant RE^m(n)$. Рассмотрим слово $w'_k = ab_1ab_2ab_3\ldots ab_ka$ такое, что $|w'_k| = 2k+1$. Легко проверить, что

$$RE^m(w'_k) = RE^p(w'_k) = RE^s(w'_k) > \frac{1}{2}[(k+1)\ln(k+1) - k] \gtrsim \frac{1}{4}|w'_k|\ln|w'_k|,$$

поэтому

$$\frac{1}{4}n\ln n \lesssim RE^s(n) \leqslant RE^p(n) \leqslant RE^m(n) \leqslant n\ln n.$$

Таким образом, $RE^s(n) = \Theta(n\log n)$, $RE^p(n) = \Theta(n\log n)$, $RE^m(n) = \Theta(n\log n)$.

Заметим, что w'_k содержит не меньше $\lfloor\alpha/2\rfloor((k+1) - \lceil\alpha/2\rceil)$ максимальных примитивных α-разрывных повторов и не меньше $\lfloor\frac{1}{2\delta}\rfloor[(k+1) - \lceil\frac{1}{2\delta}\rceil]$

максимальных δ-субпериодичностей. Таким образом, для достаточно большого α и достаточно малого δ слово w'_k содержит $\Omega(\alpha|w'_k|)$ максимальных примитивных α-разрывных повторов и $\Omega(|w'_k|/\delta)$ максимальных δ-субпериодичностей. Заметим также, что w'_k содержит в общей сложности $\Theta(|w'_k|^2)$ максимальных субпериодичностей.

Случай ограниченного алфавита является более сложным для исследования. Для простоты мы рассматриваем случай бинарного алфавита $\{0,1\}$. Положим $RE^m_{bin}(n) = \max\limits_{w\in\{0,1\}^n} RE^m(w)$, $RE^p_{bin}(n) = \max\limits_{w\in\{0,1\}^n} RE^p(w)$ и $RE^s_{bin}(n) = \max\limits_{w\in\{0,1\}^n} RE^s(w)$. Возьмем слово $w''_k = (0011)^k = \underbrace{00110011\ldots 0011}_{k}$ длины $4k$. Легко заметить, что все максимальные примитивные разрывные повторы в w''_k являются повторами с однобуквенными копиями 0 или 1. Разрывные повторы в w''_k с однобуквенными копиями 0 разбиваются на две подгруппы. Первая подгруппа состоит из всех разрывных повторов, позиции левых однобуквенных копий которых тождественны 1 по модулю 4, а позиции правых однобуквенных копий которых тождественны 2 по модулю 4. Общая сумма сокращенных порядков этих повторов равна

$$\sum_{i=1}^{k-1}\sum_{j=1}^{i}\frac{1}{4j+1}.$$

Вторая подгруппа состоит из всех разрывных повторов, позиции левых однобуквенных копий которых тождественны 2 по модулю 4, а позиции правых однобуквенных копий которых тождественны 1 по модулю 4. Общая сумма сокращенных порядков этих повторов равна

$$\sum_{i=1}^{k-1}\sum_{j=1}^{i}\frac{1}{4j-1}.$$

Поэтому общая сумма сокращенных порядков всех разрывных повторов с однобуквенными копиями 0 в w''_k равна

$$\sum_{i=1}^{k-1}\sum_{j=1}^{i}\frac{1}{4j+1}+\sum_{i=1}^{k-1}\sum_{j=1}^{i}\frac{1}{4j-1}=\sum_{i=1}^{k-1}\sum_{j=1}^{2i}\frac{1}{2j+1}.$$

Симметричным образом мы имеем ту же самую сумму сокращенных порядков всех разрывных повторов с однобуквенными копиями 1 в w''_k. Таким образом, общая сумма сокращенных порядков всех максимальных примитивных разрывных повторов в w''_k равна

$$2\sum_{i=1}^{k-1}\sum_{j=1}^{2i}\frac{1}{2j+1}.$$

Сумма $\sum\limits_{j=1}^{2i}\dfrac{1}{2j+1}$ может быть ограничена снизу следующим образом:

$$\sum_{j=1}^{2i}\frac{1}{2j+1}\geqslant\int\limits_1^{2i+1}\frac{1}{2x+1}\,dx=\frac{1}{2}\ln(2x+1)\Big|_1^{2i+1}=\frac{1}{2}(\ln(4i+3)-\ln 3).$$

Следовательно,

$$2\sum_{i=1}^{k-1}\sum_{j=1}^{2i}\frac{1}{2j+1} \geqslant \sum_{i=1}^{k-1}(\ln(4i+3)-\ln 3) = \sum_{i=1}^{k-1}\ln(4i+3)-(k-1)\ln 3.$$

Более того,

$$\sum_{i=1}^{k-1}\ln(4i+3) = \sum_{i=2}^{k}\ln(4i-1) > \int\limits_{1}^{k}\ln(4x-1)\,dx =$$
$$= [(x-\tfrac{1}{4})\ln(4x-1)-x]|_1^k = (k-\tfrac{1}{4})\ln(4k-1)-k-\tfrac{3}{4}\ln 3+1.$$

Таким образом,

$$RE^p(w_k'') > (k-\tfrac{1}{4})[\ln(4k-1)-\ln 3]-(k-1) \gtrsim \tfrac{1}{4}|w_k''|\ln|w_k''|,$$

поэтому, учитывая теорему 2, мы имеем

$$\tfrac{1}{4}n\ln n \lesssim RE_{bin}^p(n) \leqslant RE_{bin}^m(n) \leqslant n\ln n.$$

Таким образом, $RE_{bin}^p(n)=\Theta(n\log n)$ и $RE_{bin}^m(n)=\Theta(n\log n)$. Кроме того, нетрудно показать, что w_k'' содержит не меньше, чем $\lfloor\frac{\alpha-1}{4}\rfloor(4k-\lceil\alpha\rceil)$ максимальных примитивных α-разрывных повторов, поэтому для достаточно большого α число максимальных примитивных α-разрывных повторов в w_k'' равняется $\Omega(\alpha|w_k''|)$. Также легко показать, что w_k'' содержит в общей сложности $\Theta(|w_k''|^2)$ максимальных примитивных разрывных повторов. С другой стороны, отметим, что w_k'' содержит $2k+1$ максимальных периодичностей и только $2k-2$ максимальных субпериодичностей. Все эти субпериодичности имеют вид $abba$, где $a,b\in\{0,1\}$, $a\neq b$), т. е. все эти субпериодичности имеют сокращенный порядок $1/3$, поэтому $RE^s(w_k'')=(2k-2)/3\sim|w_k''|/6$.

Пусть w — слово длины n. Заметим, что $RE^m(w)-RE^p(w)$ является суммой сокращенных порядков всех максимальных непримитивных разрывных повторов в слове w, которая, очевидно, меньше числа всех максимальных непримитивных разрывных повторов в слове w. Следовательно, поскольку w содержит не более $O(n)$ максимальных непримитивных разрывных повторов, мы получаем, что $RE^m(w)-RE^p(w)=O(n)$, поэтому $RE^m(n)-RE^p(n)=O(n)$. Таким образом, из $RE^p(n)=\Theta(n\log n)$, $RE^m(n)=\Theta(n\log n)$ мы заключаем, что $RE^m(n)\sim RE^p(n)$. Таким же образом мы имеем $RE_{bin}^m(n)\sim RE_{bin}^p(n)$.

Задача оценки максимального возможного числа максимальных α-разрывных повторов в словах была сформулирована в [19] и активно изучалась в последние годы. В [9, 11] было независимо установлено, что число максимальных α-разрывных повторов в слове длины n не превосходит $O(\alpha n)$ (в [11, 12] доказано, что на самом деле это число не превосходит $18\alpha n$). Используя в качестве примера слово w_k'', можно видеть, что полученная верхняя оценка для числа максимальных α-разрывных повторов является точной по порядку для достаточно больших α в случае слов над произвольным алфавитом. Позднее в [15] полученная верхняя оценка для числа максимальных α-разрывных повторов была усилена до $3(\pi^2/6+5/2)\alpha n$.

Используя утверждение 1 и исходя из данной оценки, можно получить, что число максимальных δ-субпериодичностей в слове длины n не превосходит $3(\pi^2/6+5/2)n/\delta$. Используя пример слова w'_k, можно увидеть, что полученная оценка для числа максимальных δ-субпериодичностей является точной по порядку для достаточно малых δ в случае слов над произвольным алфавитом.

В [20] полученная верхняя оценка для числа максимальных разрывных повторов была обобщена на случай максимальных разрывных повторов с разрывами произвольного размера. Чтобы сформулировать данную оценку, введем следующие понятия и обозначения.

Для любого вещественного x положим

$$|x|^+ = \begin{cases} x, \text{ если } x>0; \\ 0 \text{ в противном случае;} \end{cases} \qquad |x|^- = \begin{cases} -x, \text{ если } x<0; \\ 0 \text{ в противном случае.} \end{cases}$$

Пусть $f(x)$ — некоторая функция из $\mathbb{N}$ в $\mathbb{R}$. Для каждого $x\in\mathbb{N}$ положим $\partial_f^+(x)=|f(x+1)-f(x)|^+$ и $\partial_f^-(x)=|f(x+1)-f(x)|^-$. Обозначим также $\sup_x\{\partial_f^+(x)\}$ $(\sup_x\{\partial_f^-(x)\})$ через ∂_f^+ (∂_f^-), если этот супремум существует. Пусть $f(x)$, $g(x)$ — две функции из $\mathbb{N}$ в $\mathbb{R}$ такие, что $f(x)\geqslant g(x)>0$ для любого $x\in\mathbb{N}$. Если обе величины ∂_f^+ и ∂_g^- существуют, обозначим $\max\{\partial_f^+,\partial_g^-\}$ через $\partial_{f,g}^a$. Если обе величины ∂_f^- и ∂_g^+ существуют, обозначим $\max\{\partial_f^-,\partial_g^+\}$ через $\partial_{f,g}^b$. Если хотя бы одна из величин $\partial_{f,g}^a$ или $\partial_{f,g}^b$ существует, то мы определяем $\partial_{f,g}$ как $\min\{\partial_{f,g}^a,\partial_{f,g}^b\}$, если обе величины $\partial_{f,g}^a$, $\partial_{f,g}^b$ существуют; в противном случае мы определяем значение $\partial_{f,g}$ равным существующей величине из величин $\partial_{f,g}^a$, $\partial_{f,g}^b$. Положим также $\Delta_{f,g}(x)=\frac{1}{x}(f(x)-g(x))\geqslant 0$ для каждого $x\in\mathbb{N}$, и $\Delta_{f,g}=\sup_x\{\Delta_{f,g}(x)\}$, если этот супремум существует.

Пусть f, g — некоторые функции из $\mathbb{N}$ в $\mathbb{R}$ такие, что $f(x)\geqslant g(x)>0$ для всех $x\in\mathbb{N}$. Мы называем разрывный повтор uvu *f,g-разрывным повтором*, если $g(|u|)\leqslant|v|\leqslant f(|u|)$. Например, α-разрывные повторы можно рассматривать как частный случай f,g-разрывных повторов для $f(x)=(\alpha-1)x$ и $g(x)=\min\{1,\alpha-1\}$. В [20] получена следующая верхняя оценка для числа максимальных f,g-разрывных повторов.

Теорема 5 [20]. *Пусть $f(x)$, $g(x)$ — некоторые функции такие, что обе величины $\partial_{f,g}$, $\Delta_{f,g}$ существуют. Тогда слово длины n содержит не более $O(n(1+\max\{\partial_{f,g},\Delta_{f,g}\}))$ максимальных f,g-разрывных повторов.*

Например, пусть $f(x)=\alpha_f+\beta_f x$ и $g(x)=\alpha_g+\beta_g x$, где $0<\alpha_g\leqslant\alpha_f$ и $0\leqslant\beta_g\leqslant\beta_f$. Тогда для каждого $x\in\mathbb{N}$ мы имеем $\partial_f^+(x)=\beta_f$, $\partial_f^-(x)=0$, $\partial_g^+(x)=\beta_g$, $\partial_g^-(x)=0$, поэтому $\partial_f^+=\beta_f$, $\partial_f^-=0$, $\partial_g^+=\beta_g$, $\partial_g^-=0$. Таким образом, $\partial_{f,g}^a=\beta_f$, $\partial_{f,g}^b=\beta_g$, и $\partial_{f,g}=\min\{\beta_f,\beta_g\}=\beta_g$. Мы также имеем $\Delta_{f,g}(x)=\frac{1}{x}(\alpha_f-\alpha_g)+(\beta_f-\beta_g)$, поэтому $\Delta_{f,g}=\Delta_{f,g}(1)=(\alpha_f-\alpha_g)+(\beta_f-\beta_g)$. Следовательно, из теоремы 5 в этом случае мы имеем, что число максимальных f,g-разрывных повторов в слове длины n ограничено сверху значением

$$O(n(1+\max\{\beta_g,(\alpha_f-\alpha_g)+(\beta_f-\beta_g)\}))= \\ =O(n(1+\max\{\beta_g,(\alpha_f-\alpha_g),(\beta_f-\beta_g)\})).$$

В заключение мы формулируем некоторые открытые проблемы в данной области.

1. Мы предполагаем, что $R_{\geqslant\lambda}(n)-R_{\geqslant\lambda+1}(n)=\Omega(n)$ для любого λ. Можно ли доказать эту гипотезу?

2. Выше на примере слов w_k показано, что $R_{\geqslant\lambda}(n)\gtrsim n/4$ для любого λ. Вместо слов w_k мы также можем рассмотреть слова $\widehat{w}_k=((01)^k0)^k$ такие, что $R_{\geqslant\lambda}(\widehat{w}_k)>(k-\lfloor\lambda/2\rfloor)(k-1)$ для $k>\lambda$, поэтому $R_{\geqslant\lambda}(\widehat{w}_k)\gtrsim k^2\gtrsim|\widehat{w}_k|/2$ при $k\to\infty$. Таким образом, мы на самом деле имеем

$$R_{\geqslant 1}(n)\geqslant R_{\geqslant 2}(n)\geqslant R_{\geqslant 3}(n)\geqslant\ldots\gtrsim n/2.$$

Является ли нижняя оценка $n/2$ строгой?

3. Обозначим через $Rp(n)$ максимальное возможное число первичных периодичностей и через $Ep(n)$ максимальную возможную сумму порядков первичных периодичностей в словах длины n. Верно ли, что $Rp(n)=R(n)$? Если $Rp(n)=R(n)$, то верно ли, что $Ep(n)=E(n)$?

4. Верно ли, что $RE^s_{bin}(n)=\Theta(n\log n)$?

5. Является ли верхняя оценка $O(n/\delta)$ для числа максимальных δ-субпериодичностей точной по порядку в случае слов над бинарным алфавитом?

6. Верно ли, что слово длины n содержит не более αn максимальных α-разрывных повторов?

7. Пусть $f(x)=\alpha x$, $g(x)=\beta x$, где $\beta<\alpha$. Верно ли, что в этом случае слово длины n содержит $O(n(1+\alpha-\beta))$ максимальных f,g-разрывных повторов?

8. Пусть на символах словарного алфавита определен некоторый линейный порядок. Этот порядок определяет лексикографический порядок на словах над рассматриваемым алфавитом. Тогда под *корнем Линдона* максимальной периодичности r в заданном слове понимается циклический корень периодичности r, который является словом Линдона для рассматриваемого лексикографического порядка, т. е. циклический корень, который является наименьшим в рассматриваемом лексикографическом порядке циклическим корнем среди всех циклических корней периодичности r. Самый левый корень Линдона периодичности r называется *левым корнем Линдона* периодичности r. В [8] доказано, что для любого фактора длины l в слове число содержащихся в этом факторе левых корней Линдона максимальных периодичностей в данном слове не превосходит $3l/2$. Верно ли, что данное число не превосходит l?

СПИСОК ЛИТЕРАТУРЫ

1. Bannai H., I T., Inenaga S., Nakashima Y., Takeda M., Tsuruta K. The “Runs” Theorem // SIAM J. Comput. — 2017. — V. 46, No 5. — P. 1501–1514.
2. Crochemore M., Ilie L. Analysis of maximal repetitions in strings // Lecture Notes in Comput. Sci. — 2007. — V. 4708. — P. 465–476.
3. Crochemore M., Ilie L. Maximal repetitions in strings // J. Comput. Syst. Sci. — 2008. — V. 74, No 5. — P. 796–807.
4. Crochemore M., Ilie L., Tinta L. Towards a solution to the “Runs” conjecture // Lecture Notes in Comput. Sci. — 2008. — V. 5029. — P. 290–302.
5. Crochemore M., Iliopoulos C., Kubica M., Radoszewski J., Rytter W., Walen T. The maximal number of cubic runs in a word // J. Comput. Syst. Sci. — 2012. — V. 78, No 6. — P. 1828–1836.

6. Crochemore M., Iliopoulos C., Kubica M., Radoszewski J., Rytter W., Walen T. Extracting powers and periods in a word from its runs structure // Theoret. Comput. Sci. — 2014. — V. 521. — P. 29–41.
7. Crochemore M., Kubica M., Radoszewski J., Rytter W., Walen T. On the maximal sum of exponents of runs in a string // J. Discrete Algorithms — 2012. — V. 14. — P. 29–36.
8. Crochemore M., Mercas R. On the density of Lyndon roots in factors // Theoret. Comput. Sci. — 2016. — V. 656. — P. 234–240.
9. Crochemore M., Kolpakov R., Kucherov G. Optimal Bounds for Computing α-gapped Repeats // Lecture Notes in Comput. Sci. — 2016. — V. 9618. — P. 245–255.
10. Fischer J., Holub S., I T., Lewenstein M. Beyond the Runs Theorem // Lecture Notes in Comput. Sci. — 2015. — V. 9309. — P. 277–286.
11. Gawrychowski P., TI T., Inenaga S., Köppl D., Manea F. Efficiently Finding All Maximal α-gapped Repeats // Proceedings of 33rd Symposium on Theoretical Aspects of Computer Science (STACS 2016). — 2016. — 39:1–39:14.
12. Gawrychowski P., I T., Inenaga S., Köppl D., Manea F. Tighter Bounds and Optimal Algorithms for All Maximal α-gapped Repeats and Palindromes // Theory Comput. Syst. — 2018. — V. 62, No 1. — P. 162–191.
13. Gasieniec L., Kolpakov R., Potapov I. Space efficient search for maximal repetitions // Theoret. Comput. Sci. — 2005. — V. 339. — P. 35–48.
14. Giraud M. Not so many runs in strings // Lecture Notes in Comput. Sci. — 2008. — V. 5196. — P. 232–239.
15. I T., Köppl D. Improved upper bounds on all maximal α-gapped repeats and palindromes // Theoret. Comput. Sci. — 2019. — V. 753. — P. 1–15.
16. Kolpakov R., Kucherov G. On Maximal Repetitions in Words // J. Discrete Algorithms — 2000. — V. 1, No 1. — P. 159–186.
17. Kolpakov R., Kucherov G. Ochem P. On maximal repetitions of arbitrary exponent // Inform. Process. Lett. — 2010. — V. 110, No 7. — P. 252–256.
18. Kolpakov R. On primary and secondary repetitions in words // Theoret. Comput. Sci. — 2012. — V. 418. — P. 71–81.
19. Kolpakov R., Podolskiy M., Posypkin M., Khrapov N. Searching of gapped repeats and subrepetitions in a word // J. Discrete Algorithms — 2017. — V. 46–47. — P. 1–15.
20. Kolpakov R. On the number of gapped repeats with arbitrary gap // Theoret. Comput. Sci. — 2018. — V. 723. — P. 11–22.
21. Puglisi S., Simpson J., Smyth W. How many runs can a string contain? // Theoret. Comput. Sci. — 2008. — V. 401. — P. 165–171.
22. Puglisi S., Simpson J. The expected number of runs in a word // Australasian J. Combinatorics — 2008. — V. 42. — P. 45–54.
23. Rytter W. The number of runs in a string: improved analysis of the linear upper bound // Lecture Notes in Comput. Sci. — 2006. — V. 3884. — P. 184–195.
24. Rytter W. The Number of Runs in a String // Inform. and Comput. — 2007. — V. 205, No 9. — P. 1459–1469.
25. Simpson J. Modified Padovan words and the maximum number of runs in a word // Australasian J. of Combinatorics — 2010. — V. 46. — P. 129–145.

Поступило в редакцию 6 II 2022