COMBINATORICA
Bolyai Society – Springer-Verlag

COMBINATORICA **23** (4) (2003) 571–584

THE FIXING BLOCK METHOD IN COMBINATORICS ON WORDS

JAMES D. CURRIE, CAMERON W. PIERCE

Received February 1, 2000

We give an overview of the method of fixing blocks introduced by Shelton. We apply the method to words which are nonrepetitive up to mod k.

1. Introduction

A word is **repetitive** if it contains two consecutive identical blocks. For example, $barbarian = bar\,bar\,ian$ is repetitive, while *civilized* is non-repetitive. A word containing k consecutive identical blocks is said to contain a k **power**. Thus a repetitive word contains a 2 power. Dejean [4] introduced the study of words containing **fractional** k **powers**. For example, the word *civilized* contains $ivi=(iv)^{3/2}$, a 3/2 power.

Thue showed [9] that there are infinite words over the three letter alphabet $\{a,b,c\}$ which are non-repetitive. Infinite non-repetitive words (sequences) have been used to build counter-examples in algebra [7], ordered sets [10], symbolic dynamics [6] and other areas. In constructing generalizations of the 1-dimensional non-repetitive sequences to higher dimensions, Currie and Simpson [5] introduced sequences which are non-repetitive up to mod r. A sequence $\{s_n\}_{n=1}^{\infty}$ is **non-repetitive up to mod** r if each of its mod k subsequences $\{s_{nk+j}\}_{n=1}^{\infty}$ is non-repetitive, $1\le k\le r$, $0\le j\le k-1$.

A non-empty set L of infinite words is **perfect** if for any $u\in L$ and any n there is a word $v\in L$, $v\neq u$ such that u and v have a common prefix of

Mathematics Subject Classification (2000): 05-XX, 68Q45

0209–9683/103/$6.00

length at least n.[1] On infinite words u and v, for any alphabet Σ, and any positive integer k the set of infinite k power free words over Σ is perfect whenever it is nonempty [3]. Related results have been proved for fractional k [2].

Problem 1.1. Let L be the set of infinite words over Σ which are non-repetitive up to mod r. Is L perfect?

For $r=1,2,3,5$ one can find a sequence over an $r+2$ letter alphabet which is non-repetitive up to mod r.

Problem 1.2. Is there a sequence over a 6 letter alphabet which is non-repetitive up to mod 4?

Using the method of fixing block inequalities introduced by Shelton, we show that for a given r Problem 1.1 has a positive answer if Σ is large enough. We outline a method of attack for Problem 1.2 above that would use fixing blocks. In the process we seek to give a gentle explanation of the fixing block method, which deserves to be more widely known.

2. Bottlenecks

Let L be a set of words over a finite alphabet Σ. Suppose that L is closed under taking prefixes. We consider the partial order $\leq$ on L where $u \leq v$ if and only if u is a prefix of v. The diagram of L under this order is a tree with root ϵ, the empty word. We will identify L with this tree. (See Figure 1 for an example.) The **meet** of words u, v is their longest common prefix, denoted by $u \wedge v$. Thus $2122 \wedge 211 = 21$, for example. As usual, the **length** of a word u is denoted by $|u|$, and is the number of letters in it, eg. $|12321| = 5$.

We use the notion of **upper cover** from ordered sets: v is an upper cover of u if $u < v$, but there is no $z \in L$ such that $u < z < v$. Given words $u \leq v$ in L, the closed interval $[u, v]$ is the set $\{w \in L : u \leq w \leq v\}$. For notational convenience, we also define $[u, \infty] = \{w \in L : u \leq w\}$. Suppose that $u \leq v$ and

1. $[\hat{v}, \infty]$ is infinite for at most one upper cover $\hat{v}$ of v.
2. $[u, \infty] \setminus [v, \infty]$ is finite.

In this case any path in L from u to ∞ must traverse the vertices of $[u, \hat{v}]$. The situation is analogous to the bottlenecks arising in rush hour traffic

[1] This corresponds to the usual topological notion when we give an appropriate metric. Infinite words u and v are at distance $1/(d+1)$ if there maximal common prefix has length d.

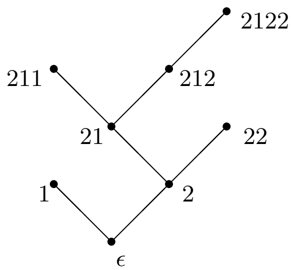

Fig. 1. The tree associated with the language $L = \{\epsilon, 1, 2, 21, 22, 211, 212, 2122\}$

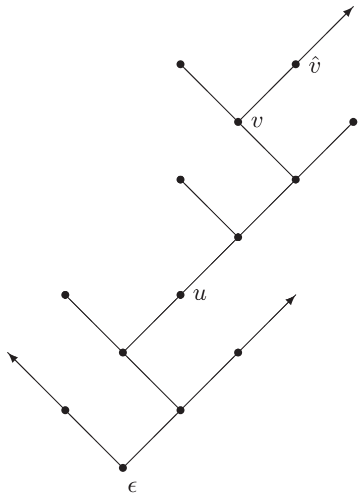

Fig. 2. The bottleneck $B(u,v)$ has length 4, index 3.

when one must leave the downtown via a certain bridge. We refer to the set $B_{\hat{v}}(u,v) = B(u,v) = [u,\infty] \setminus [\hat{v},\infty]$ as a **bottleneck** with **core** $[u,v]$.[2] The **length** of $B = B(u,v)$ is $|B| = |v| - |u| + 1$. The **index** of B is $\iota(B) = |u|$. In Figure 2 $|B| = 4$ and $\iota(B) = 3$. Suppose that $B(u,v)$ is a bottleneck and that $\hat{u}$ and $\hat{v}$ are elements of $[u,v]$ with $\hat{u} \leq \hat{v}$. It follows that at most one cover of $\hat{v}$ has an infinite extension and we can form a bottleneck $B(\hat{u},\hat{v})$.

3. Inequalities involving bottlenecks

Our purpose is to understand the structure of L in the case where L is the set of words over Σ avoiding certain patterns. In particular, we ask whether

[2] If $[v,\infty]$ is finite, let $B(u,v) = [u,\infty]$.

L is perfect, or even infinite. Our approach in particular cases is to show that long bottlenecks in L must occur far out, i.e., that for a bottleneck B of L

$$\iota(B) \geq f(|B|), \text{ whenever } |B| > N_0. \tag{1}$$

Here N_0 is some constant, while f is eventually increasing and unbounded. Suppose now that (1) holds. The existence of such an inequality gives information about the structure of L. Let $g > N_0$ be a function such that for any non-negative integer M we have $f(x) > M$ whenever $x > g(M)$.

Theorem 3.1. *Language L is infinite if and only if it contains a word of length $g(0)$.*

This is a special case ($u = \epsilon$) of the following:

Theorem 3.2. *Word $u \in L$ is a prefix of infinitely many words in L if and only if L contains a word $v > u$, with $|v| - |u| = g(|u|)$.*

Proof. It suffices to prove the if direction. Suppose that L contains a word $v > u$ with $|v| - |u| = g(|u|)$, but that $[u, \infty]$ is finite. Since $[u, \infty]$ is finite, $B = B(u, v) = [u, \infty]$ is a bottleneck. We have $\iota(B) = |u|$, and $|B| = |v| - |u| + 1 > g(|u|)$. By inequality (1),

$$\begin{aligned} |u| &= \iota(B) \\ &\geq f(|B|) \\ &= f(|v| - |u| + 1) \\ &> |u|, \text{ since } |v| - |u| + 1 > g(|u|). \end{aligned}$$

This is a contradiction. ∎

Remark 3.3. We see that if g is computable and L is recursive, we can decide which words of L have infinite extensions in L.

Inequality (1) can also be used to show that L is perfect.

Theorem 3.4. *If L is infinite, then L is perfect.*

Proof. Suppose that L is infinite but not perfect. There is some word $u \in L$ with exactly one infinite extension. Let $v > u$ be any finite prefix of the unique infinite extension of u. There will correspond a bottleneck $B = B(u, v)$ of index $|u|$, length $|v| - |u| + 1$. Because $|u|$ is fixed, but $|v|$ can be made arbitrarily large, this violates Inequality (1). ∎

Thus the infinite tree L is constantly branching. The last proof is readily sharpened to put a bound on how far L can go without branching. We state the result without proof:

Theorem 3.5. *Suppose that u has an infinite extension in L. Then there is a word $v > u$, $|v| \leq |u| + g(|u|)$ such that v is the meet of two infinite extensions of u in L.*

In the case that L is the set of non-repetitive words over $\{1,2,3\}$, f can be taken to have the form $f(x) = ax^{3/2}$, some positive constant a. (See [1, 3].) Thus g can be taken to be $g(x) = \max(N_0+1, (x/a)^{2/3})$. This implies the following:

Corollary 3.6. *The set of non-repetitive words over $\{1,2,3\}$ of length n grows exponentially.*

A striking aspect is that all of this structural information about L is demonstrated non-constructively!

4. Fixing Block Inequalities

We see the usefulness of inequalities of the form of (1). How does one get such an inequality? Our approach is to show that at least one word w in a long bottleneck B ends in a long **fixing block** - a special sort of suffix. Say $w = uv$, v the fixing block. Suffix v will be so long that prefix u is not inside the bottleneck B. This forces the bottleneck out to a position of high index.

If we consider the bottleneck schematically illustrated in Figure 2, we see that it contains several leaves or **maximal words**, i.e., words of L with no extension in L. For concreteness, let us consider the case when L is the set of non-repetitive words over $\Sigma = \{1,2,3\}$. An example of a maximal word in this case is $w = 23121323121$. Working over $\{1,2,3\}$, any word has 3 extensions. However, the extension $w1$ of w is not in L because the suffix $w_1 = 1$ of w is nearly the square 1 1. Similarly, $w2$ is not in L because the suffix $w_2 = 12$ 1 of w is nearly the square 12 12. Finally, $w_3 = w = 231213$ 23121 is nearly the square 231213 231213. Although w_3 is not a square, it is periodic; a word $u = a_1 a_2 \cdots a_n$ has a **period** t if $a_i = a_{i+t}$ for $1 \leq i < i+t \leq n$. A word with a period is **periodic**. Thus w_3 is periodic with period 6. Again, w_2 is periodic, with period 2, and w_1 has period 1. A maximal non-repetitive word has several periodic suffixes which are 'near squares'. The same situation can also occur with a word w which is not maximal. For each letter $s \in \Sigma$ such that $ws \notin L$, w must have a periodic suffix. We call such periodic suffixes **fixing blocks**; they restrict the possible extensions of w in L. In the present setting fixing blocks are exactly suffixes of non-repetitive words of the form ycy where y is a word of L, c a letter of Σ. The **period** of such a fixing block is $I = |yc|$.

In a word with several periods, interference patterns can arise. For example, if an infinite periodic word has periods α and β, it also has period $\gcd(\alpha,\beta)$. Of course if α is a multiple of β, this will not be a new period, but simply β. Suppose that a non-repetitive word w has suffixes with periods α and β. If the two suffixes 'interfere', w will have some suffix $\hat{w}$ with period $\gamma=\gcd(\alpha,\beta)$. Since w is non-repetitive, we must have $|\hat{w}|<2\gamma$. A more complicated version of this interaction happens when two distinct non-repetitive words v_1 and v_2 end in fixing blocks, and we consider the possible interference of these blocks in the common prefix $v_1\wedge v_2$. Our discussion will have motivated and made plausible the following lemma from [1]:

Lemma 4.1. *For $i=1,2$, let v_i be a word of L with a fixing block suffix of period I_i. Suppose that*

(2) $$I_2 \geq I_1$$

(3) $$\text{Not both } I_1 = I_2 \text{ and } |v_1| = |v_2|$$

(4) $$I_2 > |v_2| - |v_1 \wedge v_2|$$

Then

$$I_2 \geq 2I_1 - (|v_1| - |v_1 \wedge v_2|)$$

Condition (3) makes v_1 and v_2 distinct, while Condition (4) forces interference to occur. Note that this lemma holds when L is the set of non-repetitive words over any fixed alphabet Σ.

A variation of this lemma [3] holds in the case of words avoiding k powers, k a positive integer. A second variation [2] holds if k is some fraction, $1<k<2$. Using induction on Lemma 4.1 we can already show that reasonably long fixing blocks occur:

Corollary 4.2. *For $i=1,\dots,r$, let v_i be a word in L with a fixing block of period I_i. Suppose that for $i=1,\dots,r-1$ we have*

(5) $$I_{i+1} \geq I_i$$

(6) $$\text{Not both } I_i = I_{i+1} \text{ and } |v_i| = |v_{i+1}|$$

(7) $$I_{i+1} > |v_{i+1}| - |v_i \wedge v_{i+1}|$$

Then

$$I_r \geq 2^{r-1}I_1 - \sum_{j=1}^{r-1} 2^{r-1-j}(|v_j| - |v_j \wedge v_{j+1}|).$$

Corollary 4.3. *If v is a word of L with exactly d upper covers, $d < |\Sigma|$, then v ends in a fixing block of period $\geq 2^{|\Sigma|-d-1}$.*

We begin to get a glimmer of how we arrive at the very long fixing blocks in bottlenecks promised at the start of this section. A bottleneck, since it offers only one path to infinity, must feature many dead ends, i.e., maximal words. Such words offer sources of many fixing blocks.

Before moving on to look at how we inductively get long fixing blocks in long bottlenecks, we give the fixing block inequalities for words which are non-repetitive up to mod r. In this context, a fixing block of a word w will again be a minimal suffix v of w such that for some $s \in \Sigma$, word vs is repetitive mod k for some $k \leq r$. The period of v will be $(|v| + k)/2$. To give a concrete example, let $r = 3$. The word $w = 7123456123$ is non-repetitive up to mod 3. A fixing block of w is $w_4 = 123456123$, since its subsequence 141 means that $w4$ is repetitive mod 3. The period of w_4 is 6. One notices that w contains several 'near squares' mod 3: 141, 252, 363. However, these subsequences do not interfere with one another since their indices are in different congruence classes mod 3. We will require more than simply 2 fixing blocks in close proximity before interference starts to have an effect. However, if we have $r+1$ fixing blocks, then two must lie in the same congruence class mod r, and Lemma 4.1 may be applied to that pair:

Lemma 4.4. *Let L be the set of words over a fixed alphabet Σ which are non-repetitive up to mod r, some fixed r. For $i = 1, 2, \ldots r+1$, let v_i be a word of L with a fixing block suffix of period I_i. Suppose that for $i = 1, \ldots, r$ we have*

$$I_{i+1} \geq I_i \tag{8}$$

$$I_i \neq I_j \text{ or } |v_i| \neq |v_j| \text{ when } i \neq j \tag{9}$$

$$I_i > |v_i| - |v_i \wedge v_j| \text{when } i > j \tag{10}$$

Then

$$I_r \geq 2I_1 - (\max_{i \leq r+1} |v_i| - |\wedge_{i \leq r+1} v_i|) \tag{11}$$

This too has an extended version:

Corollary 4.5. *For $i = 1, \ldots, (s-1)r+1$, let v_i be a word in L with a fixing block of period I_i. Suppose that for $i = 1, \ldots, (s-1)r$*

$$I_{i+1} \geq I_i \tag{12}$$

$$I_i \neq I_j \text{ or } |v_i| \neq |v_j| \text{ when } i \neq j \tag{13}$$

$$I_i > |v_i| - |v_i \wedge v_j| \text{ when } i > j \tag{14}$$

Then

$$I_{(s-1)r+1} \geq 2^{s-1} I_1 - \left(\max_{i \leq (s-1)r+1} |v_i| - |\wedge_{i \leq jr} v_i| \right) \left(2^{s-2} - 1\right).$$

Corollary 4.6. *If v is a word of L with exactly d upper covers, $d < |\Sigma|$, then v ends in a fixing block of period $\geq 2^{(|\Sigma|-d-1)/r}$.*

Analogous lemmas can be written down for words avoiding k powers mod r.

5. Inductive Lemmas and Brush Lemmas

For our strategy to work we will need to restrict bottlenecks somewhat. So far there is no reason a bottleneck of fixed length and index cannot contain arbitrarily long words; we do not directly get a contradiction by having long fixing blocks in a given bottleneck. We call bottleneck $B = B(u,v)$ **regular** when

$$|B| \geq |w| - |w \wedge v| + 1 \text{ for any } w \in B.$$

A regular bottleneck of fixed index cannot contain words of arbitrary fixing block periods. Any bottleneck B contains a regular bottleneck of length $|B|$. (Just take the length B suffix of the longest maximal word in B.)

The following lemma is proved in [1], where L is the language of non-repetitive words over a fixed alphabet Σ.

Lemma 5.1. *Suppose there exist numbers m and α, $\alpha > 5$, such that that every regular bottleneck of length at least m contains a word with fixing block period at least αm. Then each bottleneck of length at least $4m$ contains a word with a fixing block period of at least $4m(2\alpha - 5)$.*

Using the fact that $2\alpha - 5 > 5$, induction gives the following:

Lemma 5.2. *If every regular bottleneck of length at least m contains a word with a fixing block period at least αm, for some $\alpha > 5$, then each bottleneck of length $\geq 4^n m$ must contain a word with a fixing block period at least $4^n m(2^n(\alpha - 5))$.*

Suppose that we can exhibit α and m such that every regular bottleneck of length at least m contains a word with a fixing block period of at least αm. Then Lemma 5.2 gives an inequality of the form of Inequality 1. When $|\Sigma|>4$, the hypothesis of Lemma 5.2 can be shown to hold with $m=1$, $\alpha=2^{|\Sigma|-2}$, using Corollary 4.3. As noted in Section 3, this gives the interesting result that for such Σ, the set of infinite non-repetitive words over Σ is perfect. Also, by Theorem 3.2, we can decide which words extend to infinite non-repetitive words over Σ.

When $|\Sigma| = 3$ a brute force search verifies that hypothesis holds for $m=13, \alpha=66/13$. The search used for [3] was difficult, using over 14 days of CPU time. This gave an alternative, non-constructive, proof that there are infinite non-repetitive words over 3 letters.

How is Lemma 5.1 proved? We can apply the fixing block inequality of Corollary 4.2 of the previous section. Some setup is necessary, after which the proof becomes a finite problem in combinatorial optimization.

Call a word with a fixing block period of at least αm a **critical word**. The hypothesis of the lemma is that every regular bottleneck of length at least m contains a critical word. Now consider a bottleneck of length at least $4m$. This will contain a regular bottleneck $B(u,v)$ of length exactly $4m$. Let $u_1=u$, and let w_1 be the least word of $[u,v]$ such that $B_1=B[u_1,w_1]$ contains a critical word w. We will have $w_1 \leq w$, and by the hypothesis of the lemma we will have $|w_1|-|u_1|<m$. Let u_2 be the upper cover of w_1 in $[u,v]$. Choose w_2 be the least word of $[u,v]$ such that $B[u_2,w_2]$ contains a critical word. Similarly choose u_3,w_3,u_4,w_4 and bottlenecks $B_i=B(u_i,w_i)$. Each B_i will contain at least one critical word, and each critical word in B_i will be an extension of w_i. For each i we will have $w_i - w_{i-1} \leq m$.

Consider the bottleneck B_1. This may not be regular, but the irregularity cannot be too bad; if $w \in B_1$, then $w \in B$, so that $|w| - |w \wedge w_1| + 1 = |w|-|w \wedge v|+1 \leq |B|=4m$, because B is regular. Pick the critical word $w^* \in B_1$ which gives the worst irregularity, viz, so that $|w^*|-|w_1|+1$ is maximal. Let $t_1=\max(1, \lfloor(|w^*|-|w_1|+1)/m\rfloor)$. We have $1 \leq t_1 \leq 4$. If $t_1>1$, partition an initial segment of $[w_1,w^*]$ into t_1 pieces of length m. Each of these segments will give a bottleneck of length m, and hence a regular bottleneck of length m. In each regular bottleneck, we can find a critical word. In this way B_1 gives us t_1 critical words $v_1,v_2,\dots,v_{t_1}$ each with $|v_j|-|w_1| \leq t_1 m-1$. We will use these critical words in the inequality of Corollary 4.2.

Proceeding similarly, we get values t_1,t_2,t_3,t_4, $1 \leq t_i \leq 4$, such that B_i contains t_i critical words v_j, each $v_j > w_i$, and with $|v_j|-|w_i| \leq t_i m-1$. In applying Corollary 4.2 we will have occasion to consider values $|v_j|-|v_j \wedge v_k|$.

These values can be bounded easily: suppose that $v_j \in B_i$, $v_k \in B_h$. Then

$$v_j \wedge v_k \;=\; \begin{cases} w_i & \text{if } i \leq h \\ w_h & \text{otherwise.} \end{cases}$$

Observe that $w_h - w_i \leq (h-i)m$. Thus

$$|v_j| - |v_j \wedge v_k| \;=\; \begin{cases} |v_j| - |w_i| & \text{for } i < h \\ |v_j| - |w_h| & \text{otherwise} \end{cases} \tag{15}$$

$$\;=\; \begin{cases} |v_j| - |w_i| & \text{for } i < h \\ |v_j| - |w_i| + |w_i| - |w_h| & \text{otherwise.} \end{cases} \tag{16}$$

$$\;\leq\; \begin{cases} t_j m & \text{for } i < h \\ t_j m + (i-h)m & \text{otherwise.} \end{cases} \tag{17}$$

So far we have found a set of $t_1+t_2+t_3+t_4$ critical words in B. We will find a subset V of these critical words such that if $v_1, v_2, \ldots, v_r$ is any ordering of the words of V, with I_j the fixing block period of v_j, then condition (7) of Corollary 4.2 will hold; in fact, we will have $\alpha m > |v_{i+1}| - |v_i \wedge v_{i+1}|$. The elements of V were chosen to be distinct, and thus the v_j will satisfy condition (6) of Corollary 4.2. Rearranging if necessary, we can ensure that condition (5) holds. Applying Corollary 4.2 to $v_1, v_2, \ldots, v_r$ then gives a fixing block period of

$$I_r \;\geq\; 2^{r-1} I_1 - \sum_{j=1}^{r-1} 2^{r-1-j}(|v_j| - |v_j \wedge v_{j+1}|).$$

To organize our thinking, let us isolate some key properties of our set of critical words V. We use the abbreviation $d_i = \lceil (|v_i| - |v_i \wedge v_{i+1}|)/m \rceil$.

- There are 4 B_i.
- Each B_i contributes t_i critical vertices v_j, $1 \leq t_i \leq 4$.
- If $v_i \in B_j$ and $v_{i+1} \in B_k$ then

$$d_i \;\leq\; \begin{cases} t_j & \text{for } j < k \\ t_j + (j-k)) & \text{otherwise.} \end{cases}$$

- We require $\alpha > d_i$ for each $i = 1, 2, 3, 4$.
- We wish to show that $2^{r-1}\alpha - \sum_{j=1}^{r-1} 2^{r-1-j} d_j \geq 4\alpha$.

We have now reached the finite problem of combinatorial optimization spoken of earlier:

Problem 5.3. Let sets V_i be given, $|V_i|\le 4$, $1\le i\le 4$. Let $V=\cup_{i=1}^{4}V_i$. Define a function $d:V\times V\to\mathbb{Z}$ as follows: If $a\in V_j$, $b\in V_k$ then

$$d(a,b) \;=\; \begin{cases} |V_j| & \text{for } j<k \\ |V_j|+(j-k)) & \text{otherwise.} \end{cases}$$

Let $u=(u_1,u_2,\dots,u_r)$ be an r-tuple of distinct elements of V. For $i=1,2,\dots,r-1$, define $d_i=d(u_{j+1},u_j)$. Define

$$\text{bound}(u) \;=\; \begin{cases} 2^{r-1}\alpha-\sum_{j=1}^{r-1}2^{r-1-j}d_j & \text{if } d_1,\dots,d_{r-1}\le\alpha \\ 0 & \text{otherwise.} \end{cases}$$

Given $U\subseteq V$ with $|U|=r$, denote by $\mathfrak{S}(U)$ the set of all ordered r-tuples of U. Define

$$\text{bound}(U)=\min_{u\in\mathfrak{S}(U)}\text{bound}(u).$$

Define

$$\text{bound}(V)=\max_{U\subseteq V}\text{bound}(U).$$

Prove that $\text{bound}(V)\ge 4\alpha$.

This problem (although involved) is finite; for the V_i, only their sizes are relevant. Once we have specified one of 4^4 possible shapes for the V_i, we look at all possible orderings of all possible subsets of V. This is again a finite problem. For each ordering u we can compute bound(u). Notice that the proof of Lemma 5.1 is a special case of this problem where each V_i is the set of critical words of B_i.

In [1,3] V above is presented in an alternate form called a **brush**, and solving Problem 5.3 becomes a 'brush lemma'. In [1] the brush lemma was proved analytically, and such a lemma was implicitly proved in [8], again analytically. However for [3] the necessary brush lemmas were simply verified mechanically by computer search.

The number 4 of the V_i in Problem 5.3, and in Lemma 5.1 can be varied at will, as long as the corresponding brush lemma can be proven. In fact in [8] the author used 5 instead of 4. This complicates the proof of the Brush Lemma in [8], since the search space for V is that much larger.[3]

Corollary 4.5 is noticeably weaker than Corollary 4.2. Preliminary analysis seems to indicate that a corresponding brush lemma will involve α of an unwieldy size, and at require at least 7 sets V_i. This would be the direction

[3] In fact an error creeps into that paper's 'brush lemma'.

to pursue to crack Problem 1.2 by our method. We can use a weaker method to say something about Problem 1.1:

For the rest of this section, let L be the language of words over an alphabet Σ which are non-repetitive up to mod r. Let $s \in \mathbb{N}$ be chosen so that $(2^{s-1} - (r(s-1)+1)) > 0$. Let α be chosen so that

$$\alpha > 2(r(s-1)+1)$$

and

$$\alpha > \frac{2(r(s-1)+1)(2^{s-2}-1)}{2^{s-1} - (r(s-1)+1)}$$

Lemma 5.4. *Suppose that every regular bottleneck of length m_0 has a vertex with a fixing block of period at least αm_0. Then every regular bottleneck of length at least $(r(s-1)+1)m_0$ has a vertex with a fixing block of period at least $\alpha(r(s-1)+1)m_0$.*

Proof. Let B be a regular bottleneck of length $(r(s-1)+1)m_0$. Divide the core of B into $r(s-1)+1$ disjoint paths, each of length m_0. Each of these paths is the core of a bottleneck of length m_0. Since every bottleneck of length m_0 contains a regular bottleneck of length m_0, we can find $r(s-1)+1$ distinct vertices in B, each having a fixing block of period at least αm_0.

Let these vertices in order of non-decreasing fixing block periods be $v_1, v_2, \dots, v_r$. Clearly conditions (12) and (13) of Corollary 4.5 hold. By the choice of α condition (14) holds also; no two words in B can have lengths differing by more than $2|B|-1$, since B is regular. Thus for any v_i, v_j we have

$$\begin{aligned} |v_i| - |v_i \wedge v_j| &\leq 2|B| - 1 \\ &= 2(r(s-1)+1)m_0 - 1 \\ &< \alpha m_0 \\ &\leq I_i \end{aligned}$$

and (14) of Corollary 4.5 holds.

Thus Corollary 4.4 applies and

$$\begin{aligned} I_{(s-1)r+1} &\geq 2^{s-1} I_1 - \left(\max_{i \leq (s-1)r+1} |v_i| - |\wedge_{i \leq jr} v_i| \right) \left(2^{s-2} - 1\right) \\ &\geq 2^{s-1}\alpha m_0 - (2(r(s-1)+1)m_0 - 1)\left(2^{s-2}-1\right) \\ &\geq \alpha m_0 (r(s-1)+1) \end{aligned}$$

by the choice of α. ∎

If we can find α and m_0 as in Lemma 5.4 we can get an inequality of form (1). This will enable us to show that L is perfect, and to decide which words extend infinitely in L. However, Lemma 5.4 requires s and α to be rather large. For example, if $r=4$, then s is at least 6. However $r=4$, $s=6$ would require $\alpha \geq 630/13$. Nevertheless, an induction can be started with $r=3$, $s=5$, $\alpha=630/13$, $m=1$, if $|\Sigma| \geq 35$; this follows from Corollary 4.6. We get the following results:

Theorem 5.5. *The set of infinite words over an alphabet Σ which are non-repetitive up to mod 4 is perfect if $|\Sigma| \geq 35$. For such a Σ we can also decide which words over Σ extend to infinite words over Σ which are non-repetitive up to mod 4.*

Theorem 5.6. *The set of infinite words over alphabet Σ which are non-repetitive up to mod r is perfect if $|\Sigma|$ is sufficiently large. We can decide which words over Σ extend to infinite words over Σ which are non-repetitive up to mod r.*

'Sufficiently large' can be replaced by a constructive condition based on first choosing s, then α, then Σ. It would be good to sharpen the bound 35 on $|\Sigma|$ mentioned above, and to give a sharpconstructive version of Theorem 5.6.

Acknowledgements

ΜΟΝΩ ΣΟΦΩ ΘΕΩ, ΔΙΑ ΙΗΣΟΥ ΧΡΙΣΤΟΥ Ω Η ΔΟΞΑ ΕΙΣ ΤΟΥΣ ΑΙΩΝΑΣ ΤΩΝ ΑΙΩΝΩΝ, ΑΜΗΝ.[4]

The first author was supported by an NSERC operating grant, and the second author was supported by an NSERC Undergraduate Summer Research Award.

References

[1] J. D. Currie: On the structure and extendibility of k-powerfree words, *Eur. J. Comb.* **16** (1995), 111–124.

[2] J. D. Currie and R. O. Shelton: Cantor sets and Dejean's conjecture, *J. Automata, Languages and Combin.* **1** (1996), 113–128.

[3] J. D. Currie and R. O. Shelton: The set of k-powerfree words over Σ is empty or perfect, *Eur. J. Comb.* **24** (2003), 573–580.

[4] Romans 16:27

[4] F. Dejean: Sur un théorème de Thue, *J. Combin. Theory Ser. A* **13** (1972), 90–99.
[5] J. D. Currie and R. J. Simpson: Nonrepetitive tilings, *Electron J. Combin.* **9** (2002), no.I, R28.
[6] M. Morse and G. A. Hedlund: Symbolic dynamics I, II, *Amer. J. Math.* **60** (1938), 815–866; **62** (1940) 142; *MR* **1**, 123d.
[7] P. S. Novikov and S. I. Adjan: Infinite periodic groups I, II, III, *Izv. Akad. Nauk. SSSR Ser. Mat.* **32** (1968), 212–244; 251–524; 709–731; *MR* **39** #1532a–c.
[8] R. O. Shelton: On the structure and extendibility of square-free words, *Combinatorics on Words: Progress and Perspectives* (1983), Academic Press, 101–188.
[9] A. Thue: Über unendliche Zeichenreihen, *Norske Vid. Selsk. Skr. I. Mat. Nat. Kl.* Christiana (1906), 1–22.
[10] W. T. Trotter and P. Winkler: Arithmetic progressions in partially ordered sets, *Order* **4** (1987), 37–42.

James D. Currie

Department of Mathematics and Statistics
University of Winnipeg
Winnipeg, Manitoba R3B 2E9
Canada
currie@io.uwinnipeg.ca

Cameron W. Pierce

Department of Mathematics and Statistics
University of Winnipeg
Winnipeg, Manitoba R3B 2E9
Canada
cpierce@io.uwinnipeg.ca