Uniformly scattered factors

Lucian Ilie

Turku Centre for Computer Science TUCS
FIN-20520 Turku, Finland

Ion Petre

Turku Centre for Computer Science TUCS
FIN-20520, Turku, Finland
and
Department of Mathematics, University of Turku
FIN-20014, Turku, Finland

Grzegorz Rozenberg

Department of Computer Science, Leiden University
P.O.Box 9512, 2300 RA Leiden, The Netherlands
and
Department of Computer Science
University of Colorado at Boulder
Boulder, CO 80309, U.S.A.

Turku Centre for Computer Science
TUCS Technical Report No 243
February 1999

ISBN 952-12-0389-7
ISSN 1239-1891

Abstract

A word u appears as a factor of another word v as it is; in one piece. When u is a subword of v, u may be scattered as several factors. We consider the in-between case and put some restrictions on the number of factors as which u is allowed to be scattered. A large class of partial orders which are generalizations of factors and subwords is obtained. For all of them, the property of being well partial order is fully characterized. The result generalizes Higman's theorem.

TUCS Research Group
Theory Group: Mathematical Structures in Computer Science

1 Introduction

Two of the most important partial orders on words are the factor and the subword partial orders. If a word u is a factor of a word v, then u appears inside v exactly as it is. If u is a subword of v, then u may appear inside v scattered as several factors. It naturally arises the idea of considering restrictions on the number of factors as which u is allowed to be scattered. This idea is not new. It appeared first in a paper of Haines, cf. [Ha], who just mentioned the possibility of putting a fixed upper bound on the numer of factors and notices that, in this way, one obtains a relation admitting infinite antichains (sets containing pairwise incomparable elements only). A deeper investigation of the problem is performed for the first time in [Il], where some general restrictions are imposed and some necessary and sufficient conditions for the existence of the infinite antichains. In this paper, we consider a different approach of the problem. Our restrictions on the number of factors are similar with the ones of [Il] but the way we consider the factors is different: all factors but the last one have the same length. In this way, all relations obtained are partial orders and, moreover, we can iterate the generalization process, obtaining a very large class of partial orders which are generalizations of factors and subwords. We then fully chracterize those which are well partial orders. The result generalizes Higman's theorem.

2 Basic partial orders

In this section, we perform the first step of our generalization. We fix first some notations and concepts.

For an alphabet Σ, the set of all finite words over Σ is denoted Σ^*; the empty word is ε. For a word $w \in \Sigma^*$, $|w|$ denotes the length of w. For two words $u, v \in \Sigma^*$, we say that u is a *factor* of v, denoted $u \leqslant_f v$, if there are $x, y \in \Sigma^*$ such that $v = xuy$; u is a subword of v, denoted $u \leqslant_s v$, if $u = a_1 a_2 \cdots a_n$ for some $a_i \in \Sigma, 1 \leq i \leq n$, and $v = v_1 a_1 v_2 a_2 \cdots v_n a_n v_{n+1}$, for some $v_i \in \Sigma^*, 1 \leq i \leq n+1$.

Consider a set A and a partial order $\leqslant$ on A. A *chain* of $\leqslant$ is a set $B \subseteq A$ such that, for any $a, b \in B$, either $a \leqslant b$ or $b \leqslant a$. An *antichain* of $\leqslant$ is a set $B \subseteq A$ such that, for any $a, b \in B$, $a \not\leqslant b$. The partial order $\leqslant$ is called *well founded* if it has no infinite descending sequence $a_1 \leqslant^{-1} a_2 \leqslant^{-1} a_3 \leqslant^{-1} \cdots$ such that, for no $i \geq 1$, $a_i \leqslant a_{i+1}$, where $\leqslant^{-1}$ is the inverse of $\leqslant$; $\leqslant$ is called *well partial order* if it is well founded and has no infinite antichain.

It is easy to see that both the factor partial order $\leqslant_f$ and the subword partial order $\leqslant_s$ are well founded. Also, the factor partial order is not well partial order as soon as $\mathrm{card}(\Sigma) \geq 2$, since it has infinite antichains, e.g., $\{ab^n a \mid n \geq 1\}$. The famous theorem of Higman, cf. [Hi], states that this is not the case for $\leqslant_s$. (In fact Higman [Hi] proved a much more general result; Theorem 1 is a consequence of Higman's result in the case of words.)

Theorem 1 (Higman [Hi]). *The subword partial order $\leqslant_s$ is well partial order.*

All the bounds we are to put on the number of factors as which a certain word is allowed to be scattered are given by functions having as argument the length of the word. We now construct the family $\mathcal{G}$ of all these functions. It contains all $g : \mathbb{N} \longrightarrow \mathbb{R}$ with the following properties ($[r]$ denotes the integer part or r):

(i) $g(0) = 0, 1 \leq g(n) \leq n$, for any $n \geq 1$;

(ii) the function $[g(n)]$ is increasing;

(iii) the function $[\overline{g}(n)]$ is increasing, where $\overline{g}(0) = 0$ and $\overline{g}(n) = \frac{n}{g(n)}$, for any $n \geq 1$.

For a non-empty word $w \in \Sigma^+$ and a function $g \in \mathcal{G}$, we define the *g-factorization* of w as the factorization

$$w = w_1.w_2.\cdots.w_{[\overline{g}(|w|)]}.w_{[\overline{g}(|w|)]+1}$$

where

$$|w_1| = |w_2| = \cdots = |w_{[\overline{g}(|w|)]}| = [g(|w|)].$$

From the definition of the family $\mathcal{G}$, the g-factorization of a non-empty word is well defined. We may assume, for the sake of completeness, that the g-factorization of the empty word has just one factor.

Next, for a function $g \in \mathcal{G}$, we define the binary relation $\leqslant_{f,g}$ on Σ^* as follows.

(a) $\varepsilon \leqslant_{f,g} \varepsilon$; for any $w \in \Sigma^*$, $\varepsilon \leqslant_{f,g} w$; $w \leqslant_{f,g} \varepsilon$ implies $w = \varepsilon$;

(b) for two non-empty words $u, v \in \Sigma^+$, if their g-factorizations are, respectively,

$$u = u_1.u_2.\cdots.u_{[\overline{g}(|u|)]}.u_{[\overline{g}(|u|)]+1},$$
$$v = v_1.v_2.\cdots.v_{[\overline{g}(|v|)]}.v_{[\overline{g}(|v|)]+1},$$

then $u \leqslant_{f,g} v$ if and only if (i)-(iii) below are fulfilled:

(i) $|u| \leq |v|$;

(ii) there is a subsequence $1 \leq k_1 < k_2 < \cdots < k_{[\overline{g}(|u|)]} \leq [\overline{g}(|v|)]$ such that $u_i \leqslant_f v_{k_i}$, for any $1 \leq i \leq [\overline{g}(|u|)]$;

(iii) if $|u| = |v|$, then $u_{[\overline{g}(|u|)]+1} = v_{[\overline{g}(|v|)]+1}$.

Remark 2. For any $u, v \in \Sigma^*$, if $u \leqslant_{f,g} v$ and $|u| = |v|$, then $u = v$. Also, for any $g \in \mathcal{G}$, $\leqslant_{f,g}$ is well founded.

Consider next a few examples.

Example 3. Our first example shows that indeed our relations introduced above (and which are partial orders as proved below) are generalizations of the factor and subword partial orders. Consider the two functions

$$\mathbf{1}, \mathbf{1}_{\mathbb{N}} : \mathbb{N} \longrightarrow \mathbb{N},$$

given by

$$\mathbf{1}(0) = 0, \mathbf{1}(n) = 1, \quad \mathbf{1}_{\mathbb{N}}(n) = n,$$

for any $n \geq 0$. It is clear that we have

$$\leqslant_{f,\mathbf{1}} = \leqslant_s, \quad \leqslant_{f,\mathbf{1}_{\mathbb{N}}} = \leqslant_f .$$

Example 4. Consider the function $g_1(n) = \min(n, 2)$, for any $n \geq 0$. Then, for the words $u = aabab$ and $v = bbaaabba$ we have the g_1-factorizations respectively $aa.ba.b$ and $bb.aa.ab.ba$ and hence $u \leqslant_{f,g_1} v$. Also $aa.ba \not\leqslant_{f,g_1} ba.ab.ab$.

Example 5. For the function

$$g_2 = \begin{cases} 1, & \text{if } 1 \leq n \leq 3, \\ \frac{n}{3}, & \text{if } n \geq 4, \end{cases}$$

we have that $aab.abb.aba.ab \leqslant_{f,g_2} baab.abbb.abaa$ but $ab.aa.bb.a \not\leqslant_{f,g_2} ab.aa.bb.b$.

Example 6. Consider next a little bit more strange function, $g_3(n) = \sqrt{n}$. Clearly, $g_3 \in \mathcal{G}$. We have then $aba.aba.bba.bbbaba \leqslant_{f,g_3} abba.abaa.baba.abba$.

Lemma 7. *For any $g \in \mathcal{G}$, the relation $\leqslant_{f,g}$ is a partial order.*

Proof. The reflexivity of the relation is obvious and the antisymmetry follows from Remark 2. For the transitivity, we just have to use the transitivity of the factor partial order $\leqslant_f$. ■

3 Characterization of well partial orders

We give in this section a characterization of those partial orders introduced above which are well partial orders. As they are all well founded, we shall be concerned only with the finiteness of antichains.

We first notice that, for any $g \in \mathcal{G}$, since g is increasing, the limit $\lim_{n\to\infty} g(n)$ always exists in $\mathbb{R} \cup \{\infty\}$.

Theorem 8. *If $g \in \mathcal{G}$ such that $\lim_{n\to\infty} g(n) = \infty$, then $\leqslant_{f,g}$ is not well partial order as soon as* $\mathrm{card}(\Sigma) \geq 2$.

Proof. All we have to do is to find an infinite antichain for $\leqslant_{f,g}$. As $g(n)$ goes to infinity with n, it follows that there is a subsequence $(k_n)_{n\geq 1}$ of $(n)_{n\geq 1}$ such that

$$2 \leq g(k_1) < g(k_2) - 1 < g(k_3) - 2 < g(k_4) - 3 < \cdots .$$

Consider then the following set of words over the two-letter alphabet $\{a, b\}$:

$$A = \{w_n = (ab^{[g(k_n)]-2}a)^{[\overline{g}(k_n)]}a^{k_n - [g(k_n)][\overline{g}(k_n)]} \mid n \geq 1\}.$$

We claim that A is an antichain of $\leqslant_{f,g}$. Assume that there $1 \leq n < m$ such that $w_n \leqslant_{f,g} w_m$. For any $n \geq 1$, the g-factorizations of w_n is

$$w_n = ab^{[g(k_n)]-2}a.ab^{[g(k_n)]-2}a.\cdots.ab^{[g(k_n)]-2}a.a^{k_n-[g(k_n)][\overline{g}(k_n)]},$$

and so, from the fact that $w_n \leqslant_{f,g} w_m$, we get that $ab^{[g(k_n)]-2}a \leqslant_f ab^{[g(k_m)]-2}a$, which is imposible since $g(k_n) < g(k_m) - 1$ implies that $[g(k_n)] < [g(k_m)]$. The result is proved. ■

We now prove that the converse of Theorem 8 hold true as well, thus obtaining a characterization of the relations in the family $(\leqslant_{f,g})_{g\in\mathcal{G}}$ which are well partial orders.

Theorem 9. *If $g \in \mathcal{G}$ such that $\lim_{n\to\infty} g(n) < \infty$, then $\leqslant_{f,g}$ is well partial order.*

Proof. Assume that $\lim_{n\to\infty} g(n) = d$, for some $d \in \mathbb{R}$. Therefore, as g is increasing, there are $c, n_0 \in \mathbb{N}, c \geq 1$, such that, for any $n \geq n_0$, $[g(n)] = c$. (In fact, $c \in \{[d], [d] - 1\}$.)

We argue by contradiction. Assume that $\leqslant_{f,g}$ is not a well quasi order and consider an antichain of it, say $\{w_1, w_2, w_3, \ldots\}$. We may assume that $c \leq |w_1| < |w_2| < |w_3| < \cdots$.

Consider a new alphabet Δ which has as letters the words of length c over Σ, that is, $\Delta = \{w \in \Sigma^* \mid |w| = c\}$. For any $i \geq 1$, consider the g-factorizations of w_i $w_i = w_{i,1}.w_{i,2}.\cdots.w_{i,[\overline{g}(|w_i|)]}.w_{i,[\overline{g}(|w_i|)]+1}$ and denote $v_i = w_{i,1}w_{i,2}\cdots w_{i,[\overline{g}(|w_i|)]}$. Then the set $\{v_1, v_2, v_3, \dots\} \subseteq \Delta^*$, where each word v_i is viewed as a word over Δ (its letters being the factors in the corresponding g-factorization) is an infinite antichain of the subword partial order $\leqslant_s$ on Δ^*. But this contradicts Higman's theorem. ■

From Theorems 8 and 9 we obtain the characterization of the relations in the family $(\leqslant_{f,g})_{g\in\mathcal{G}}$ which are well partial orders.

Theorem 10. *For any $g \in \mathcal{G}$, $\leqslant_{f,g}$ is well partial order if and only if $\lim_{n\to\infty} g(n) < \infty$.*

4 Further generalization

It is not difficult to see that in defining our general relations we may start with any partial order instead of the factor partial order $\leqslant_f$ and still the obtained relations are partial orders. In this way, a natural question arises: what are the well partial orders if we start with the subword partial order $\leqslant_s$ instead of the factor one $\leqslant_f$? More general, we can iterate the process of defining our partial orders above. So, as $\leqslant_{f,g_1}$ is a partial order, for any $g_1 \in \mathcal{G}$, we can use it as a start instead of $\leqslant_f$. If the function of $\mathcal{G}$ we use is g_2, then denote the obtained partial order by $\leqslant_{f,g_1,g_2}$. We then iterate this process and get the most general partial orders which are of the form

$$\leqslant_{f,g_1,g_2,\dots,g_n}, \tag{1}$$

where $n \geq 1, g_i \in \mathcal{G}$, for any $1 \leq i \leq n$. We notice that starting with $\leqslant_s$ is just a particular case here, namely, when $g_1(n) = \mathbf{1}_{\mathbb{N}}$.

In this section we make all preliminaries for the main result of the paper, in the next section, which characterizes those partial orders in (1) which are well partial orders. As all partial orders in (1) are well founded, we shall investigate only the existence of infinite antichains.

We prove first a result on arbitrary partial orders.

Lemma 11. *Let A be an infinite set and $\leqslant$ a partial order on A. Then A contains either an infinite antichain or an infinite chain of $\leqslant$.*

Proof. Consider the set M of all maximal elements of A w.r.t. $\leqslant$. M is antichain of $\leqslant$. It it is infinite, then we are done. It it is empty, then, clearly, A contains an infinite chain of $\leqslant$. Assume M is finite non-empty. Then there is $x \in M$ such that the set $N = \{y \in A \mid y \leqslant x\}$ is infinite. We now repeat the above reasoning with $N - \{x\}$ instead of A, namely, we consider the cardinality of the set of maximal elements of $N - \{x\}$. If the above reasoning can be repeated indefinitely, that is, at each step, the respective set of maximal elements is finite non-empty, then, again, we can construct an infinite chain of $\leqslant$ in A. ■

The next corollary follows immediately from Lemma 11.

Corollary 12. *Let A be an infinite set and $\leqslant$ a partial order on A. If $\leqslant$ is well partial order, then A contains an infinite chain of $\leqslant$.*

For any partial order $\leqslant$ on Σ^* and any $g \in \mathcal{G}$, we denote by $\leqslant_g$ the relation obtained as before, using $\leqslant$ instead of $\leqslant_f$.

For any positive integer $c \geq 1$, consider the function $\mathbf{c} : \mathbb{N} \longrightarrow \mathbb{N}$, defined by $\mathbf{c}(n) = \min(n, c)$, for any $n \geq 0$. Obviously, $\mathbf{c} \in \mathcal{G}$.

Lemma 13. *For any $c \geq 1$ and any partial order $\leqslant$ on Σ^*, the relation $\leqslant_{\mathbf{c}}$ is well partial order.*

Proof. Our idea is similar with the one used by Conway [Co] for proving Higman's theorem and which is originally due to Nash-Williams, cf. [NW]. Clearly, it is enough to prove that the relation

$$\leqslant'_{\mathbf{c}} = \leqslant_{\mathbf{c}} - \{(u, v) \mid u \leqslant_{\mathbf{c}} v, |u| \leq c\}$$

has no infinite antichain. We argue by contradiction. Assume that there is an infinite antichain of $\leqslant'_{\mathbf{c}}$, say

$$w_1, w_2, w_3, \ldots . \tag{2}$$

It follows that

$$\text{for any } i, j \geq 1, \text{if } i < j, \text{then } w_i \not\leqslant'_{\mathbf{c}} w_j. \tag{3}$$

We may assume that (2) is an "earliest" sequence of words such that (3) is fulfilled, i.e., w_1 is a shortest word beginning a sequence satisfying (3), w_2 is a shortest word such that w_1, w_2 begin a sequence satisfying (3), and so on and so forth.

Now, there exists a word $w \in \Sigma^*, |w| = c$, such that infinitely many of the words of (2) start with w as a prefix. Assume that the indices of all such words are $(i_j)_{j\geq 1}$ such that $1 \leq i_1 < i_2 < i_3 < \cdots$ and put, for any $j \geq 1$, $w_{i_j} = wu_j$. Construct then the sequence

$$w_1, w_2, \ldots , w_{i_1-1}, u_1, u_2, \ldots . \tag{4}$$

It is not difficult to check that the sequence (4) satisfies (3). Also, as $c \geq 1$, it is "earlier" than (2), a contradiction. The lemma is proved. ■

The next result generalizes Theorem 9.

Lemma 14. *For any partial order* $\leqslant$ *and any* $g \in \mathcal{G}$ *such that* $\lim_{n\to\infty} g(n) < \infty$, *the relation* $\leqslant_g$ *is well partial order.*

Proof. Assume $\lim_{n\to\infty} g(n) = d$, for some $d \in \mathbb{R}$. Since g is increasing, there are $c, n_0 \in \mathbb{N}, c \geq 1$, such that, for any $n \geq n_0$, $[g(n)] = c$. (In fact, $c \in \{[d], [d]-1\}$.)

Assume that $\leqslant_g$ is not a well partial order. Then, there exists an infinite antichain for it, say $A = \{w_1, w_2, w_3, \dots\}$. By possibly eliminating some of the words of A, we may assume that $n_0 \leq |w_1| < |w_2| < |w_3| < \cdots$.

Consider, for each $w_i, i \geq 1$, its g-factorization, say

$$w_i = w_{i,1}.w_{i,2}.\cdots.w_{i,[\overline{g}(|w_i|)]}.w_{i,[\overline{g}(|w_i|)]+1}$$

and denote, for any $i \geq 1$,

$$v_i = w_{i,1}w_{i,2}\cdots w_{i,[\overline{g}(|w_i|)]}.$$

Notice that $|w_{i,j}| = c$, for any $1 \leq j \leq [\overline{g}(|w_i|)]$.

Fix a leter $a \in \Sigma$. We then claim that

$$B = \{v_1 a, v_2 a, v_3 a, \dots\}$$

is an antichain of $\leqslant_{\mathbf{c}}$. To prove this, consider $v_i a, v_j a \in B$ and assume that $i < j$. Then, the $\mathbf{c}$-factorizations of these two words are clearly

$$\begin{aligned} v_i a &= w_{i,1}.w_{i,2}.\cdots.w_{i,[\overline{g}(|w_i|)]}.a, \\ v_j a &= w_{j,1}.w_{j,2}.\cdots.w_{j,[\overline{g}(|w_j|)]}.a \end{aligned}$$

and hence, using the fact that $|w_i| < |w_j|$, $v_i a \leqslant_{\mathbf{c}} v_j a$ would imply $w_i \leqslant_g w_j$, a contradiction. Therefore, B is an infinite antichain of $\leqslant_{\mathbf{c}}$, which contradicts Lemma 13. ■

For the rest of the paper, we shall make the following *assumption*: any partial order $\leqslant$ on words we shall consider has the property that $u \leqslant v$ implies $|u| \leq |v|$. Remark that all partial orders $\leqslant_{f,g_1,\dots,g_n}$ do have this property.

Consider an alphabet Σ and a partial order $\leqslant$ on Σ^*. Denote by $\overline{\Sigma^*}$ the set of all tuples of words of equal length over Σ, that is,

$$\overline{\Sigma^*} = \{(w_1, w_2, \dots, w_n) \mid n \geq 1, w_i \in \Sigma^*, |w_1| = |w_2| = \cdots = |w_n|\}.$$

We define the partial order $\overline{\leqslant}$ on $\overline{\Sigma^*}$ as follows: for any $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_m) \in \overline{\Sigma^*}$, we have $x \overline{\leqslant} y$ if and only if (i)-(iii) below are fulfilled:

(i) $n \leq m$;

(ii) $|x_1| \leq |y_1|$;

(iii) there is a subsequence $1 \leq k_1 < k_2 < \cdots < k_n \leq m$ such that $x_i \leqslant y_{k_i}$, for any $1 \leq i \leq n$.

It is easy to see that $\overline{\leqslant}$ is a partial order on $\overline{\Sigma^*}$. The next lemma shows that the property of being well is inherited from $\leqslant$ to $\overline{\leqslant}$.

Lemma 15. *If $\leqslant$ is a well partial order on Σ^*, then $\overline{\leqslant}$ is a well partial order on $\overline{\Sigma^*}$.*

Proof. By contradiction. Assume $\overline{\leqslant}$ is not well partial order. We shall construct an infinite antichain of $\leqslant$, thus contradicting the hypothesis. This infinite antichain is constructed inductively, the main tool for the construction being the next claim.

Claim. Assume we have $B \subseteq \Sigma^*$ and $C \subseteq \overline{\Sigma^*}$ which verify $(*)$-$(***)$ below:

$(*)$ B is an antichain of $\leqslant$;

$(**)$ C is an infinite antichain of $\overline{\leqslant}$;

$(***)$ for any $u \in B$ and $v = (v_1, v_2, \ldots, v_n) \in C$, we have $u \not\leqslant v_i$, for any $1 \leq i \leq n$.

Then, there is $u \in \Sigma^*$ and $C' \subseteq \overline{\Sigma^*}$ such that $(*)$-$(***)$ above are fulfilled with $B' = B \cup \{u\}$ and C' instead of B and C, respectively.

Before proving the claim, we notice that an infinite antichain B of $\leqslant$ is constructed starting with $B = \emptyset$ and C an infinite antichain of $\overline{\leqslant}$ (there exists such a C by our assumption) and applying indefinitely the Claim.

Proof of Claim. Assume $C \subseteq \overline{\Sigma^*}$ is an infinite antichain of $\overline{\leqslant}$ and denote

$$C = \{w_n = (w_{n,1}, w_{n,2}, \ldots, w_{n,l(n)}) \mid n \geq 1\}.$$

In virtue of Lemma 13, we may assume that, for any $n < m$, $|w_{n,i}| < |w_{m,j}|$.

For any $n \geq 1$, denote by $i(n)$ the smallest $i, 1 \leq i \leq l(n)$, such that $w_{1,1} \leqslant w_{n,l(n)}$; if there is no such i, then set $i(n) = 0$. (Of course, $i(1) = 1$.)

If there are infinitely many n such that $i(n) = 0$, then $B' = B \cup \{w_{1,1}\}$ and $C' = \{w_n \in C \mid i(n) = 0\}$ are good choices. It is important to notice that when $l(1) = 1$, i.e., $w_1 = w_{1,1}$, then $i(n) = 0$, for all $n \geq 2$, and we can construct B' and C' as required.

If there are finitely many n with $i(n) = 0$, then consider the set

$$C_1 = \{(w_{n,1}, w_{n,2}, \ldots, w_{n,i(n)-1}) \mid i(n) \neq 0\}.$$

By Lemma 11, C_1 contains either an infinite antichain or an infinite chain of $\overline{\leqslant}$; denote it C_2 in either case. In the former case, take $B' = B \cup \{w_{1,1}\}$ and $C' = C_2$; clearly, B' and C' satisfy $(*)$-$(***)$.

Consider the latter case and put

$$C_3 = \{w_{n,i(n)} \mid (w_{n,1}, w_{n,2}, \ldots, w_{n,i(n)-1}) \in C_2\}.$$

Applying Corollary 12 to C_3 and $\leqslant$ we get an infinite chain of $\leqslant$ in C_3, say C_4. Therefore, the set

$$C_5 = \{(w_{1,2}, w_{1,3}, \ldots, w_{1,l(1)})\} \cup \{(w_{n,i(n)+1}, w_{n,i(n)+2}, \ldots, w_{n,l(n)}) \mid w_{n,i(n)} \in C_4\}$$

is an infinite antichain of $\overline{\leqslant}$ and also $(*)$-$(***)$ are fulfilled with C_5 instead of C. Notice the essential difference between C and C_5: $w_{1,1}$ has been eliminated. We next repeat the same reasoning with C_5 instead of C and continue in this way until either some good B' and C' are found or $w_{1,2}, w_{1,3}, \ldots, w_{1,l(1)-1}$ are all eliminated, one by one; in the latter case, B' and C' as required can be again constructed, as we already noticed. The proof is concluded. ■

Corollary 16. *For any well partial order $\leqslant$ on Σ^* and any $g \in \mathcal{G}$, the relation $\leqslant_g$ is well partial order.*

Proof. Assume that $\leqslant_g$ is not well partial order and consider an infinite antichain $\{w_n \mid n \geq 1\}$ of it. Without loss of generality we can assume that $|w_n| < |w_{n+1}|$ for any $n \geq 1$. If the g-factorization of w_n is $w_n = w_{n,1}.w_{n,2}.\cdots.w_{n,[\overline{g}(|w_n|)]}.w_{n,[\overline{g}(|w_n|)]+1}$, then the set $\{(w_{n,1}, w_{n,2}, \ldots, w_{n,[\overline{g}(|w_n|)]}) \mid n \geq 1\}$ is an infinite antichain of $\overline{\leqslant}$. This contradicts Lemma 15. ■

5 The general characterization

We are now ready to prove the main result of the paper: the characterization of the partial orders in (1) which are well partial orders. The part (i) in Theorem 17 is a generalization of Higman's theorem.

Theorem 17. *Let $m \geq 1$ and $g_1, \ldots, g_m \in \mathcal{G}$.*

(i) *If $\lim\limits_{n\to\infty} g_k(n) < \infty$ for some $1 \leq k \leq m$, then $\leqslant_{f,g_1,\ldots,g_m}$ is well partial order.*

(ii) *If* $\lim\limits_{n\to\infty} g_k(n) = \infty$ *for all* $1 \leq k \leq m$*, then* $\leqslant_{f,g_1,\ldots,g_m}$ *is not well partial order as soon as* $\mathrm{card}(\Sigma) \geq 2$*.*

Proof. (i) By Lemma 14 we obtain that $\leqslant_{f,g_1,\ldots,g_k}$ is well partial order. Then, by Corrolary 16, all relations $\leqslant_{f,g_1,\ldots,g_l}$ with $k \leq l \leq m$, are well partial orders.

(ii) For any $g \in \mathcal{G}$, let $[g]$ denote the function given, for any $n \geq 0$, by $[g](n) = [g(n)]$. It is clear from the hypothesis that $\lim\limits_{n\to\infty}([g_1] \circ \cdots \circ [g_m])(n) = \infty$. Then, there is a subsequence $(k_n)_{n\geq 1}$ of $(n)_{n\geq 1}$ such that $2 \leq ([g_1] \circ \cdots \circ [g_m])(k_1)$ and $([g_1] \circ \ldots \circ [g_m])(k_n) < ([g_1] \circ \ldots \circ [g_m])(k_{n+1}) - 1$ for all $n \geq 1$.

Denote $k_n^{(i)} = ([g_i] \circ \ldots \circ [g_m])(k_n)$ for any $n \geq 1$ and i, $1 \leq i \leq m$. We have then $2 \leq [g_1](k_1^{(2)}) < [g_1](k_2^{(2)}) - 1 < [g_1](k_3^{(2)}) - 2 < \cdots$. This, by the proof of Theorem 8, implies that the set

$$A_1 = \{(ab^{[g_1(k_n^{(2)})]-2}a)^{[\overline{g_1}(k_n^{(2)})]}a^{k_n^{(2)}-[g_1(k_n^{(2)})][\overline{g_1}(k_n^{(2)})]} \mid n \geq 1\}$$

is an antichain of $\leqslant_{f,g_1}$.

Assume that we have constructed an infinite antichain of $\leqslant_{f,g_1,\ldots,g_{i-1}}$, say $A_{i-1} = \{u_n \mid n \geq 1\}$, such that $|u_n| = k_n^{(i)}$, for all $n \geq 1$. We claim that we can construct an infinite antichain $A_i = \{v_n \mid n \geq 1\}$ of $\leqslant_{f,g_1,\ldots,g_i}$ such that $|v_n| = k_n^{(i+1)}$, for all $n \geq 1$. Notice that $1 \leq k_n^{(i)} = ([g_i] \circ [g_{i+1}] \circ \ldots \circ [g_m])(k_n) = [g_i](k_n^{(i+1)}) \leq k_n^{(i+1)}$.

We choose, for every $n \geq 1$, the word $v_n = (u_n)^{[k_n^{(i+1)}/k_n^{(i)}]}w_n$, where w_n is a prefix of u_n such that $|v_n| = k_n^{(i+1)}$. (Actually, w_n can be any other word of the same length.)

Suppose that A_i is not an antichain of $\leqslant_{f,g_1,\ldots,g_i}$, i.e. for some $r < s$ we have $v_r \leqslant_{f,g_1,\ldots,g_i} v_s$. Since $[g_i](|v_n|) = [g_i](k_n^{i+1}) = k_n^{(i)} = |u_n|$, for all $n \geq 1$, the g_i-factorizations of v_r and v_s are

$$v_r = u_r. \cdots .u_r.w_r, \qquad v_s = u_s. \cdots .u_s.w_s.$$

Thus, $v_r \leqslant_{f,g_1,\ldots,g_i} v_s$ implies that $u_r \leqslant_{f,g_1,\ldots,g_{i-1}} u_s$, which contradicts the fact that A_{i-1} is an antichain of $\leqslant_{f,g_1,\ldots,g_{i-1}}$.

For $i = m$ we obtain an infinite antichain A_m of $\leqslant_{f,g_1,\ldots,g_m}$, which concludes the proof of the theorem. ∎

References

[Co] Conway, J.H., *Regular Algebra and Finite Machines* (Chapman and Hall, 1971).

[Ha] Haines, L., On free monoids partially ordered by embedding, *J. Combin. Theory* **6** (1969) 94 – 98.

[Hi] Higman, G., Ordering by divisibility in abstract algebras, *Proc. London Math. Soc.* **2**(3) (1952) 326 – 336.

[Il] Ilie, L., Generalized factors of words, *Fund. Inform.* **33** (1998) 239 – 247.

[NW] Nash-Williams, C., On well quasi-ordering finite trees, *Proc. Cambridge Philos. Soc.* **59** (1963) 833 – 835.

Turku Centre for Computer Science
Lemminkäisenkatu 14
FIN-20520 Turku
Finland

http://www.tucs.abo.fi

University of Turku
- **Department of Mathematical Sciences**

Åbo Akademi University
- **Department of Computer Science**
- **Institute for Advanced Management Systems Research**

Turku School of Economics and Business Administration
- **Institute of Information Systems Science**