Available online at www.sciencedirect.com
SCIENCE DIRECT®
Theoretical Computer Science 339 (2005) 49–60
Theoretical Computer Science
www.elsevier.com/locate/tcs

Generalised fine and Wilf's theorem for arbitrary number of periods

Sorin Constantinescu, Lucian Ilie*,[1]
Department of Computer Science, University of Western Ontario, London, Ont., Canada N6A 5B7

Abstract

The well known Fine and Wilf's theorem for words states that if a word has two periods and its length is at least as long as the sum of the two periods minus their greatest common divisor, then the word also has as period the greatest common divisor. We generalise this result for an arbitrary number of periods. Our bound is strictly better in some cases than previous generalisations. Moreover, we prove it optimal. We show also that any extremal word is unique up to letter renaming and give an algorithm to compute both the bound and a corresponding extremal word.

Keywords: Combinatorics on words; Periods; Fine and Wilf's theorem

1. Introduction

Fine and Wilf's theorem for words is one of the most widely used and known results on words. It was initially proved by Fine and Wilf [7] in connection with real functions but then adopted as a natural result for words, see [5,9,10]. We say that a word (string) has a certain integer as period if the word repeats itself after that period; e.g., the word *abaabaaba* has periods 3, 6, and 8. It is not difficult to see that, given a set of periods, any long enough word which has those periods will have also their greatest common divisor as period. The essential question is how long the word should be. Fine and Wilf's theorem states that length for two periods: it is the sum of the two periods minus their greatest common divisor. We

* Corresponding author.
E-mail addresses: sorinco@csd.uwo.ca (S. Constantinescu), ilie@csd.uwo.ca (L. Ilie).
[1] Research supported in part by NSERC.

0304-3975/$ - see front matter
doi:10.1016/j.tcs.2005.01.007

are interested in both upper and lower bounds for this length. While the bound stated by this theorem is a lower bound, it has been also proved to be an upper bound, that is, it has been proved to be optimal. The optimality has been rigorously proved by Choffrut and Karhumäki [5].

The problem of finding the (optimal) bound for arbitrary number of period has not been settled yet. The first generalisation was given by Castelli et al. [3] for three periods. Then, following the same ideas, a generalisation for an arbitrary number of periods was given by Justin [8]. Their bounds are proved to be lower bounds and were claimed to be optimal in some loose sense, see below. Further extensions and generalisations of Fine and Wilf's theorem are given in [1,4,2,11].

First of all, we need to make it clear what we are looking for. Given a set of periods, we want the optimal bound (i.e., shortest length) which imposes the greatest common divisor as period. The above mentioned generalisations gave a lower bound which can be strictly improved in some cases. The loose optimality given by them essentially shows that for some, but not all, sets of periods, if 1 is subtracted from their bound, then it will no longer impose the greatest common divisor as period.

We shall give a new bound, in the general case of any arbitrarily fixed number of periods, and prove it optimal in the natural (strong) sense mentioned above. Our construction closely follows the previous generalisations and then modifies those by considering a case when their bound can be strictly improved. The modification proves to be essential as it brings the optimality. While the proof that the new bound is a good lower bound is not difficult, the optimality is a bit more involved. We give also an algorithm which computes simultaneously the bound and a word realising it.

The paper is organised as follows. Section 2 gives the basic definitions and the formal statement of Fine and Wilf's theorem together with its counterpart for the optimality. Section 3 introduces the new bound which is proved to be good in Section 4. Section 5 introduces graphs associated with bounds and sets of periods and gives several results about those which are used in the optimality proof of Section 6. The results for the associated graphs are interesting by themselves and, in particular, after the optimality of the bound is proved the uniqueness of the extremal words follows immediately. The last section contains the algorithm which is a straightforward application of the results on graphs.

2. Fine and Wilf's theorem

An *alphabet* is a finite non-empty set. For an alphabet A, the set of all finite words over A is denoted by A^*. For a word $w \in A^*$, the *length* of w, that is, the number of letters in w, is denoted by $|w|$. If $w = a_1a_2 \ldots a_n$, where $a_i \in A$, for all $1 \leqslant i \leqslant n$, we say that $p \geqslant 1$ is a *period* of w iff $a_i = a_{i+p}$, for all $1 \leqslant i \leqslant n - p$. Notice that any $p \geqslant |w|$ is a period of w.

Given an n-tuple of positive integers $\mathbf{p} = (p_1, \ldots, p_n)$ and a positive integer k, we say that k is a *good bound* for $\mathbf{p}$ if any word of length k which has periods $p_1, \ldots, p_n$ has also period $d = \gcd(p_1, \ldots, p_n)$; k is the *optimal bound* for $\mathbf{p}$ if it is a good bound whereas $k - 1$ is not, that is, there exists a word w of length $k - 1$ which has the periods in $\mathbf{p}$ but not d. Notice that the notion of optimal bound makes sense only if d is not among the elements of $\mathbf{p}$.