Journal of Automata, Languages and Combinatorics **15** (2010) 3/4, 269–283

f-WORDS AND BINARY SOLID CODES [1]

STAVROS KONSTANTINIDIS and JOSHUA YOUNG
Department of Mathematics and Computing Science, Saint Mary's University
Halifax, Nova Scotia, B3H 3C3 Canada
e-mail: s.konstantinidis@smu.ca, jyo04@hotmail.com

ABSTRACT

Given any unbounded and non-decreasing sequence f of positive integers, we define an infinite set of binary words, called f-words, which constitute an overlap-free language. We investigate some properties of this language and then use these properties to define new classes of finite and infinite binary solid codes – solid codes have the strongest synchronization and error-delimiting capabilities in the hierarchies of codes. The finite class improves on an earlier construction of solid codes in terms of average word length (or information rate), without sacrificing their encoding complexity. The infinite class concerns maximal solid codes and builds on an earlier work on maximal binary solid codes. This work constitutes another step towards a systematic structural characterization of binary maximal solid codes.

Keywords: Construction, encoding complexity, maximal, overlap-free language, solid code, word

1. Introduction

From a language theoretic point of view, a solid code is an infix code (no codeword is contained in another codeword) and an overlap-free language (no proper prefix of a codeword is also a proper suffix of some codeword, unless it is empty).

Solid codes constitute a proper subclass of comma-free codes, providing thus synchronization of messages without delay. In addition, they possess the following remarkable property: if a message over a solid code undergoes any type of errors, then the codewords containing no errors can be identified and decoded correctly without delay. We refer the reader to [4] for a relevant discussion and further references on these codes, as well as to [3, 8, 1, 6, 10, 13, 2] for more recent results on these objects.

The general problem of constructing "good" solid codes is of central importance in our context. Some of the criteria used in the literature for evaluating the "goodness" of codes are the following: maximality, information ratio (or average word length), encoding/decoding complexity, and error-detectability. In addition, there have been systematic efforts to characterize explicitly, or generate algorithmically, all possible solid codes of certain types [3, 8].

[1] Research supported by a Discovery Research Grant of NSERC, Canada.