Journal of Automata, Languages and Combinatorics **28** (2023) 1–3, 287 – 338

FINITE MAXIMAL CODES AND FACTORIZATIONS OF CYCLIC GROUPS

Clelia De Felice

Dipartimento di Informatica, Università degli Studi di Salerno
via Giovanni Paolo II 132, 84084 Fisciano (SA), Italy
cdefelice@unisa.it

ABSTRACT

Variable-length codes are the *bases* of the free submonoids of a free monoid. There are some important longstanding open questions about the structure of finite *maximal codes*, namely the *factorization conjecture* and the *triangle conjecture*, proposed by Perrin and Schützenberger. The latter concerns finite codes Y which are subsets of a^*Ba^*, where a is a letter and B is an alphabet not containing a. A structural property of finite maximal codes has recently been shown by Zhang and Shum. It exhibits a relationship between finite maximal codes and factorizations of cyclic groups. With the aim of highlighting the links between this result and other older ones on maximal and factorizing codes, we give a simpler and a new proof of this result. As a consequence, we prove that for any finite maximal code $X \subseteq (B \cup \{a\})^*$ containing the word a^{pq}, where p, q are prime numbers, $X \cap a^*Ba^*$ satisfies the triangle conjecture. Let n be a positive integer that is a product of at most two prime numbers. We also prove that it is decidable whether a finite code $Y \cup a^n \subseteq a^*Ba^* \cup a^*$ is included in a finite maximal code and that, if this holds, $Y \cup a^n$ is included in a code that also satisfies the factorization conjecture.

Keywords: formal languages, variable-length codes, finite maximal codes, factorizations of cyclic groups

1. Introduction

The theory of *variable-length codes* takes its origin in the framework of the theory of information, since Shannon's early works in the 1950's. An algebraic theory of codes was subsequently initiated by Schützenberger, who proposed in [38] the semigroup theory as a mathematical setting for the study of these objects. In this context the theory of codes has been extensively developed, showing strong relations with automata theory, combinatorics on words, formal languages and the theory of semigroups (see [2] for a complete treatment of this topic). In this paper we follow this algebraic approach and codes are defined as the *bases* of the free submonoids of a free monoid.

0000-0002-1789-1706