Journal of Automata, Languages and Combinatorics **10** (2005) 1, 3–24

A LANGUAGE-THEORETIC APPROACH TO COVERING PROBLEMS

MARCELLA ANSELMO[1]
Dip. di Informatica ed Applicazioni, Università di Salerno
I-84081 Baronissi (SA), Italy
e-mail: anselmo@dia.unisa.it

and

MARIA MADONIA[2]
Dip. di Matematica ed Informatica, Università di Catania
Viale Doria 6/A, Catania, Italy
e-mail: madonia@dmi.unict.it

ABSTRACT

The formal language of all words 'covered' by words in a given language is investigated. A language is a covering code when any word has at most one minimal covering over it; in this case the language generated by the covering operation is said cov-free. We study the notion of cov-freeness, in analogy to the theory developed on classical freeness. In particular cov-freeness is characterized by the notion of cov-stability introduced here. Further, cov-maximality of a regular covering code is characterized by its cov-completeness. Some more properties are obtained using these characterizations. We also show that the series counting the minimal coverings of a word over a regular language is rational. All along the paper we compare new definitions and results to their counterpart in classical monoids and in monoids of zig-zag factorizations.

Keywords: Formal languages, variable-length codes, covering, formal power-series

1. Introduction

Recently several problems related to covers were examined, especially motivated by molecular biology, but also by data compression and computer-assisted music analysis. A wide literature exists; see for example [4, 7, 11, 13, 14, 19, 21] and all other papers there cited. Roughly speaking, a covering of a word w over a language X is a way to obtain w by concatenations and overlaps of elements of X. The above cited literature affords covering problems from an algorithmic point of view: given a word w find an

[1]Partially supported by "Progetto Cofinanziato MIUR: Linguaggi formali e automi: teoria ed applicazioni".

[2]Partially supported by "Progetto Cofinanziato MIUR: Ragionamento su aggregati e numeri a supporto della programmazione e relative verifiche"