Journal of Automata, Languages and Combinatorics **6** (2001) 3, 239–251

SOME RESULTS ON CODES FOR GENERALIZED FACTORIZATIONS

Mouhamed Kadrie Ahmad and Xavier Augros
Laboratoire I3S bât. ESSI, Université de Nice-Sophia Antipolis
930, Route des Colles, B.P. 145, F-06903 Sophia-Antipolis Cedex, France
e-mail: {ahmad,augros}@essi.fr

ABSTRACT

This paper is about unique decipherability, according to a control language, for the concatenation of words chosen in different languages. This control language and this set of languages generate an other language. We give a generalization of the notions of code and stability, and, in the rational case, we also give a polynomial time algorithm to decide if a such generator is a code.

Keywords: rational language, code, factorization, stability, shuffle.

1. Introduction

The *shuffle of two words u and v according to a word t*, called *trajectory*, is the letter by letter shuffle of u and v such that one takes the current letter of u if the current letter of t is 0, and one takes the current letter of v if the current letter of t is 1. As an example, with $u = ab$, $v = cab$, $t = 01101$ one obtains the word $acabb$. The *shuffle of two languages X and Y according to a language T of words over a binary alphabet* {0,1} is the set of shuffles of words in X and Y according to words in T. So one obtains a subset of $X \sqcup\!\sqcup Y$. Here we consider a shuffle of concatenations of words in two languages X and Y according to a set of trajectories T, so one obtains a subset of $(X \cup Y)^*$. The 3-tuple (T, X, Y) generate the language $X \sqcup\!\sqcup_T^* Y$. As an example, with $X = \{ab, abb\}$, $Y = \{ca\}$, $T = \{0110, 11\}$ one obtains the set $X \sqcup\!\sqcup_T^* Y = \{abcacaab, abcacaabb, abbcacaab, abbcacaab, caca\}$. For $\mathcal{F} = (\{0110, 11\}, \{ab, abb\}, \{ca\})$, a tuple such that (ab, ca, ca, abb) is called a factorization (or $\mathcal{F}$-factorization in [5]) of the word $abcacaabb$.

When we consider words of L^*, for a given language L, a factorization of a word $u \in L^*$ is a sequence of words of L: $(u_1, u_2, \ldots, u_n)$, such that $u = u_1 u_2 \ldots u_n$. This operation is introduced in [5] in the case of any number n of languages $X_1, \ldots, X_n$. Here we state results only for two languages, however in the same way they can be proved for any number of languages. Here we are interested in uniquely deciphering property for these $\mathcal{F}$-factorizations. A tuple $\mathcal{F} = (T, X, Y)$ is called a $\mathcal{F}$-code if any word w on the alphabet Σ has at most one $\mathcal{F}$-factorization on (T, X, Y). This