ARTICLE IN PRESS
YINCO:4180

Information and Computation ••• (••••) •••–•••

Contents lists available at ScienceDirect

Information and Computation

www.elsevier.com/locate/yinco

Picture codes and deciphering delay ☆

Marcella Anselmo [a,*], Dora Giammarresi [b], Maria Madonia [c]

[a] *Dipartimento di Informatica, Università di Salerno, Via Giovanni Paolo II, 132, 84084 Fisciano (SA), Italy*
[b] *Dipartimento di Matematica, Università Roma "Tor Vergata", via della Ricerca Scientifica, 00133 Roma, Italy*
[c] *Dipartimento di Matematica e Informatica, Università di Catania, Viale Andrea Doria 6/a, 95125 Catania, Italy*

ARTICLE INFO

Article history:
Received 27 June 2014
Available online xxxx

Keywords:
Two-dimensional languages
Codes
Deciphering delay

ABSTRACT

A set X of pictures over an alphabet Σ is a code if any picture over Σ is tilable in at most one way with pictures in X. The codicity problem is in general undecidable. Recently, the prefix picture codes were introduced as a decidable subclass of codes that generalize the prefix string codes. In the string theory, the finite deciphering delay sets are some interesting codes which coincide with the prefix codes when the delay is equal to 0. An analogous notion is introduced for the picture codes and it is proved that the codes with deciphering delay k form a decidable class of picture codes which includes interesting examples and special cases.

1. Introduction

Since the beginning of the formal language theory, string codes have been a very important subject of research, for their theoretical interest and also because of their immediate applications to practical problems. The theory of (variable length) codes takes its origin in the theory of information devised by Shannon in the 1950s. A code is a set of strings such that any coded message can be uniquely decomposed as a concatenation of strings in the code. M.-P. Schützenberger discovered that the coding theory is closely related to the classical algebra and he initiated an investigation on codes based on combinatorics and algebraic methods (refer to [11,12] for a comprehensive study).

Extensions of the classical strings to two dimensions can be done in several ways; they bring to the definition of polyominoes, labeled polyominoes, as well as rectangular labeled polyominoes we will refer to as pictures. On the other hand, the notion of code can be intuitively and naturally transposed to the two dimensional objects by exploiting the notion of unique tiling decomposition. A set C of polyominoes is a *code* if every polyomino which is tilable with (copies of) elements of C, it is tilable in a unique way. Unfortunately, most of the published results show that in the two dimensional (2D for short, as opposed to 1D) context most important properties are lost. In [10] D. Beauquier and M. Nivat proved that the problem whether a finite set of polyominoes is a code is undecidable, and that the same result holds also for simple dominoes. Codes involving other variants of polyominoes including bricks (i.e. labeled polyominoes) and pictures are studied in [1,13,19,21,24] and further undecidability results are proved. It is worthwhile to remark that all the mentioned results consider 2D codes independently from a 2D formal language theory.

☆ Partially supported by MIUR Project 2010LYA9RH *"Automata and Formal Languages: Mathematical and Applicative Aspects"*, MIUR Project *"PRISMA PON04a2 A/F"*, FARB Project ORSA138754 of University of Salerno, PRA Project 2012 of University of Catania.
* Corresponding author.
E-mail addresses: anselmo@dia.unisa.it (M. Anselmo), giammarr@mat.uniroma2.it (D. Giammarresi), madonia@dmi.unict.it (M. Madonia).

http://dx.doi.org/10.1016/j.ic.2016.06.003

Very recently, a new definition for picture codes was introduced in [5,7], in connection with the family *REC* of the picture languages recognized by *finite tiling systems*. Observe that the finite tiling systems generalize to two dimensions the finite state automata for strings, and that the family REC is considered as the two-dimensional counterpart of the regular string languages (see [17]). In [5,7] the picture codes are defined by using the formal operation of *tiling star* as defined in [25]; the tiling star of a set X is the set X^{**} of all the pictures that are tilable (in the polyominoes style) by elements of X. Then, X is a code if any picture in X^{**} is tilable in a unique way. Note that if $X \in$ REC then X^{**} is in REC, too. By analogy to the string case, it holds that if X is a finite picture code then one can construct an unambiguous tiling system for X^{**} (see [3] for the definition), starting from the pictures in X. Unfortunately, despite this nice connection to the string code theory, it is proved that it is undecidable whether a given set of pictures is a code. This is actually coherent with the known result of undecidability for the unambiguity inside the family REC ([3]).

Coming back to the string theory, observe that the aim of the theory of codes is to give a structural description of codes in a way that allows their construction. This is easily accomplished for prefix codes. A set S of words is called prefix if no word in S is (left-)prefix of another one. It holds that any prefix set of words is also a code, referred to as a prefix code. Moreover, any string in a prefix code can be decoded *on-time*; while reading the string from left-to-right, at most one element in the code can match the input.

Looking for decidable subclasses of picture codes, a definition of *prefix code* for pictures is proposed in [5,7]. The pictures are considered with a preferred scanning direction – from the top-left corner towards the bottom-right one. Intuitively, the property to be maintained in a prefix code X is that during the decoding process of a picture the next element in X is unique. Nevertheless, the formal definition cannot be merely translated from the strings to two dimensions. The main difficulty comes from the fact that the initial part of a string is still a string, while the initial part of a picture is not in general a picture (it is not rectangular). The formal definition of the prefix sets of pictures involves some special kind of polyominoes. The main results in [5,7] include the proof that it is decidable whether a finite set of pictures is a prefix set and that, as in the one dimensional case, every prefix set of pictures is a code. Furthermore, it is proved that several other properties can be extended from the string prefix codes to the prefix codes of pictures. Very interestingly, exploiting the richness of the two-dimensional structures, a particular subclass of prefix codes, named *strong prefix codes*, is introduced and studied in [6,9]; strong prefix codes have a simpler definition based on the overlapping of pairs of pictures and preserve all the desired prefix properties.

Another important notion that appears at the very beginning of the theory of codes is the one of deciphering delay. In 1959, E.N. Gilbert and E.F. Moore [18] introduced the delay of an encoding, in order to investigate the encodings of a message without the prefix property. The idea is that the delay of a set is d, if scanning a coded message, it is necessary to wait for the receipt of the first d code-words, before the first one can be fixed. In this setting, the prefix codes are the codes with delay zero; that is why they are called *instantaneous* codes (see [11,12] for all formal definitions). Hence, the codes with finite deciphering delay form an intermediate family between prefix codes and general codes. They share many useful properties with the prefix codes. Furthermore, these classes of codes seem to be important still nowadays, for applications as for example the computation of the Hausdorff dimension of language-defined fractals (see [15]).

In this paper we extend the concept of finite deciphering delay code to two dimensions. Again, while the intuitive definition seems quite intuitive and natural, the formalization of such ideas is extremely involved and requires to cleverly invent the right definition of polyomino prefix of another polyomino, as well as to carefully count the k code-pictures that play the role of the delay. The main difficulties arise, as in the prefix case, from the fact that in the decoding process of the initial part of a picture is a polyomino and not a picture. Moreover we have the further inconvenience that it is necessary to compare two different polyominoes, one composed by at least k pictures of the code.

We give the definition of set of pictures with finite deciphering delay. As in the string case, a code with delay equal to 0 is assumed to be a prefix code. We prove that if a set of pictures has finite deciphering delay then it is a code. This result holds both for finite and infinite picture languages. Hence it provides a useful tool to prove that a picture language is a code (recall that the property is in general undecidable). Further we exhibit, for each integer k, a set with deciphering delay equal to k, and point out a set that has not finite deciphering delay. Moreover, since it is decidable whether a finite set of pictures has deciphering delay equal to k, this notion contributes to enlarge the family of known decidable picture codes. Infinite codes are also considered in relation to their deciphering delay and some interesting examples are inductively constructed to show the richness of this new family of picture codes.

The paper is organized as follows. Section 2 reports all needed definitions and known results. Section 3 introduces the definition of picture code with finite delay, while Sections 4 and 5 include all new results in the case of finite and infinite codes, respectively. Some conclusions and further perspectives are discussed in Section 6.

Part of the presented results appeared as a preliminary version in [8].

2. Preliminaries

In this section we collect all the preliminary notions needed to present our results. We first recall the notion of string code with finite deciphering delay, then we introduce all the necessary notations on pictures, picture languages and polyominoes. Finally, we mention some previous results on picture codes and prefix picture codes.