

THE CODE PROBLEM FOR DIRECTED FIGURES

MICHAŁ KOLARZ¹

Abstract. We consider directed figures defined as labelled polyominoes with designated start and end points, with two types of catenation operations. We are especially interested in codicity verification for sets of figures, and we show that depending on the catenation type the question whether a given set of directed figures is a code is decidable or not. In the former case we give a constructive proof which leads to a straightforward algorithm.

Mathematics Subject Classification. 68R15, 68R99.

INTRODUCTION

Variable-length word codes, *i.e.*, subsets X of a monoid such that every product of the elements decomposes uniquely over X , have been studied by many authors and extensive literature exists in this subject, including *e.g.* the well-known monograph [3]. Much less is known about the codicity in two dimensions. Some authors have extended word codes to trees (see [6]) or polyominoes (*e.g.* [1,2]); decidability of the codicity testing problem varies in these cases. However, it is undecidable in the case of both standard and labelled polyominoes (*cf.* [8,9]).

The interest in two-dimensional codes is natural in the context of various disciplines that use picture encodings, *e.g.* [4]. In the present paper we study pictures composed of labelled polyominoes, equipped with start and end points; we call them directed figures. Catenation of directed figures is defined with an optional merging function. We characterize decidability of the question whether a given set of directed figures is a code in both situations (with or without the merging function). In the decidable cases we give proofs that lead to simple algorithms.

Keywords and phrases. Directed figures, variable-length codes, codicity verification, Sardinias-Patterson algorithm.

¹ Institute of Computer Science, Jagiellonian University, Łojasiewicza 6, 30-348 Kraków, Poland; mko1@smp.if.uj.edu.pl

In Section 1 we define directed figures and related operations. Then, in Section 2, the notion of a code is introduced in the context of catenation with and without the merging function. The decidability question for some specific cases is also solved here. Section 3 is concerned with the codicity verification for both kinds of codes, giving algorithms in typical cases. We end with some plans for future study.

1. PRELIMINARIES

Let Σ be a finite, nonempty alphabet. For $u = (u_x, u_y) \in \mathbb{Z}^2$, a *translation* in \mathbb{Z}^2 by vector u is denoted by τ_u ,

$$\tau_u : \mathbb{Z}^2 \ni (x, y) \mapsto (x + u_x, y + u_y) \in \mathbb{Z}^2.$$

For a set $V \subseteq \mathbb{Z}^2$ and an arbitrary function $f : V \rightarrow X$ it obviously induces

$$\begin{aligned} \tau_u : P(\mathbb{Z}^2) \ni V &\mapsto \{\tau_u(v) \mid v \in V\} \in P(\mathbb{Z}^2), \\ \tau_u : X^V \ni f &\mapsto f \circ \tau_{-u} \in X^{\tau_u(V)}. \end{aligned}$$

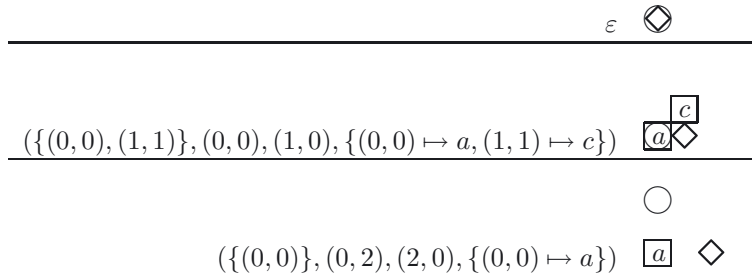
Definition 1.1 (Directed figure). Let $D \subseteq \mathbb{Z}^2$ be finite and non-empty, $b, e \in \mathbb{Z}^2$ and $l : D \rightarrow \Sigma$. A quadruple $f = (D, b, e, l)$ is called a *directed figure* (over Σ) with

<i>domain</i>	$\text{dom}(f) = D,$
<i>start point</i>	$\text{begin}(f) = b,$
<i>end point</i>	$\text{end}(f) = e,$
<i>labelling function</i>	$\text{label}(f) = l,$
<i>translation vector</i>	$\text{tran}(f) = e - b.$

In addition we define the *empty directed figure* ε as $(\emptyset, (0, 0), (0, 0), \emptyset)$.

Note that we make no additional assumptions about the relative placement of domain, start point and end point of the figure and connectedness of the domain.

Example 1.2. Directed figures with graphical representation (a circle marks the start point and a diamond marks the end point of the figure).



The set of all directed figures over Σ is denoted by Σ^\diamond . Two directed figures x, y are equal if there exists $u \in \mathbb{Z}^2$ such that

$$y = (\tau_u(\text{dom}(x)), \tau_u(\text{begin}(x)), \tau_u(\text{end}(x)), \tau_u(\text{label}(x))).$$

Thus, we actually consider figures up to a translation.

Definition 1.3 (Catenation). Let $x = (D_x, b_x, e_x, l_x)$ and $y = (D_y, b_y, e_y, l_y)$ be directed figures. If $D_x \cap \tau_{x_e - y_b}(D_y) = \emptyset$, *catenation* of x and y is defined as

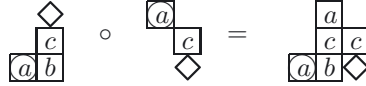
$$x \circ y = (D_x \cup \tau_{x_e - y_b}(D_y), b_x, \tau_{x_e - y_b}(e_y), l),$$

where

$$l(z) = \begin{cases} l_x(z) & \text{for } z \in D_x, \\ \tau_{x_e - y_b}(l_y)(z) & \text{for } z \in \tau_{x_e - y_b}(D_y). \end{cases}$$

If $D_x \cap \tau_{x_e - y_b}(D_y) \neq \emptyset$, catenation of x and y is not defined.

Example 1.4. Catenation of two figures.



Example 1.5. Catenation of the following figures is not defined (point labeled by c in the first figure overlaps with point labeled by b in the second figure).



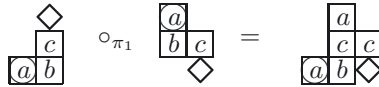
Definition 1.6 (m -catenation). Let $x = (D_x, b_x, e_x, l_x)$ and $y = (D_y, b_y, e_y, l_y)$ be directed figures. m -catenation of x and y w.r.t. a merging function $m : \Sigma \times \Sigma \rightarrow \Sigma$ is defined as

$$x \circ_m y = (D_x \cup \tau_{x_e - y_b}(D_y), b_x, \tau_{x_e - y_b}(e_y), l),$$

where

$$l(z) = \begin{cases} l_x(z) & \text{for } z \in D_x \setminus \tau_{x_e - y_b}(D_y), \\ \tau_{x_e - y_b}(l_y)(z) & \text{for } z \in \tau_{x_e - y_b}(D_y) \setminus D_x, \\ m(l_x(z), \tau_{x_e - y_b}(l_y)(z)) & \text{for } z \in D_x \cap \tau_{x_e - y_b}(D_y). \end{cases}$$

Example 1.7. π_1 -catenation of two figures (π_1 denotes projection on the first argument).



Observe that \circ is associative. \circ_m is associative if and only if m is associative. So for associative m , $\Sigma_m^\circ = (\Sigma^\circ, \circ_m)$ is a monoid (which is never free).

Abusing this notation, we also write X° (resp. X_m°) to denote the set of all figures that can be composed by \circ catenation (resp. \circ_m m -catenation) from figures in $X \subseteq \Sigma^\circ$. When some results hold for \circ catenation as well as for \circ_m m -catenation, we use \bullet catenation symbol, and “ $x \bullet y$ ” should then be read as “ $x \circ y$ (resp. $x \circ_m y$)”. In the same way we use “ $x \in X_\bullet^\circ$ ” to denote “ $x \in X^\circ$ (resp. $x \in X_m^\circ$)”.

From now on let m be an arbitrary associative merging function.

2. CODICITY

2.1. CODES

In this subsection we show that in general it is not decidable whether a given set is a code w.r.t. \circ catenation.

Definition 2.1 (Code). $X \subseteq \Sigma^\circ$ is a *code* if for any $x \in X^\circ$ there exists only one sequence $x_1, \dots, x_k \in X$ such that $x = x_1 \circ \dots \circ x_k$.

Example 2.2. $\{\boxed{a}\diamond, \diamond\boxed{a}\} \subseteq \{a\}^\circ$ is a code.

Example 2.3. $\{w = \begin{array}{c} \diamond \\ \boxed{a} \ \boxed{a} \end{array}, x = \boxed{a} \ \boxed{a} \diamond, y = \begin{array}{c} \boxed{a} \\ \diamond \end{array}, z = \begin{array}{c} \boxed{a} \\ \boxed{a} \end{array} \diamond\} \subseteq \{a\}^\circ$ is not a code since $w \circ x = y \circ z = \begin{array}{c} \boxed{a} \ \boxed{a} \\ \boxed{a} \ \boxed{a} \end{array} \diamond$

The following consideration is based on ideas presented in [2] (Sects. 3 and 4).

Let $\Sigma = \{a\}$. For $h, h_N, h_E, h_S, h_W \in \mathbb{Z}_+$ such that $h_N, h_E, h_S, h_W \leq h$ and $b, e \in \{N, E, S, W\}$ (N as north, E as east, S as south and W as west) we define a *directed hooked square* $\text{DHS}_h(h_N, h_E, h_S, h_W)_e^b$ to be a directed figure $f \in \Sigma^\circ$ with:

$$\begin{aligned} \text{dom}(f) &= B \setminus (H_N^- \cup H_E^- \cup H_S^- \cup H_W^-) \cup (H_N^+ \cup H_E^+ \cup H_S^+ \cup H_W^+), \\ \text{begin}(f) &= \begin{cases} (0, h+2) & \text{if } b = N, \\ (h+2, 0) & \text{if } b = E, \\ (0, -h-2) & \text{if } b = S, \\ (-h-2, 0) & \text{if } b = W, \end{cases} \\ \text{end}(f) &= \begin{cases} (0, h+3) & \text{if } e = N, \\ (h+3, 0) & \text{if } e = E, \\ (0, -h-3) & \text{if } e = S, \\ (-h-3, 0) & \text{if } e = W, \end{cases} \end{aligned}$$

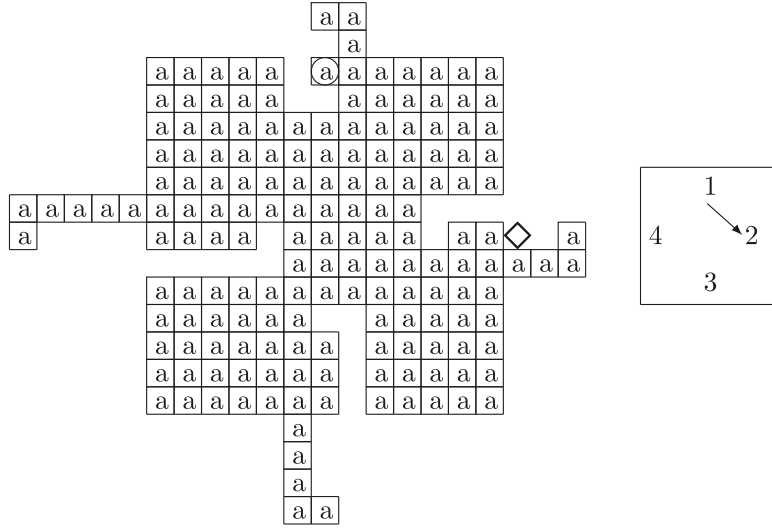


FIGURE 1. $DHS_4(1, 2, 3, 4)_E^N$ - full and reduced graphical representation.

where

$$\begin{aligned}
 B &= \{(x, y) \mid x, y \in \{-h - 2, \dots, h + 2\}\}, \\
 H_N^- &= \{(-1, y) \mid y \in \{h + 2 - h_N, \dots, h + 2\}\} \cup \{(0, h + 2 - h_N)\}, \\
 H_E^- &= \{(x, 1) \mid x \in \{h + 2 - h_E, \dots, h + 2\}\} \cup \{(h + 2 - h_E, 0)\}, \\
 H_S^- &= \{(1, y) \mid y \in \{-h - 2, \dots, -h - 2 + h_S\}\} \cup \{(0, -h - 2 + h_S)\}, \\
 H_W^- &= \{(x, -1) \mid x \in \{-h - 2, \dots, -h - 2 + h_W\}\} \cup \{(-h - 2 + h_W, 0)\}, \\
 H_N^+ &= \{(1, y) \mid y \in \{h + 3, \dots, h + 3 + h_N\}\} \cup \{(0, h + 3 + h_N)\}, \\
 H_E^+ &= \{(x, -1) \mid x \in \{h + 3, \dots, h + 3 + h_E\}\} \cup \{(h + 3 + h_E, 0)\}, \\
 H_S^+ &= \{(-1, y) \mid y \in \{-h - 3 - h_S, \dots, -h - 3\}\} \cup \{(0, -h - 3 - h_S)\}, \\
 H_W^+ &= \{(x, 1) \mid x \in \{-h - 3 - h_W, \dots, -h - 3\}\} \cup \{(-h - 3 - h_W, 0)\},
 \end{aligned}$$

i.e. f is a square with hooks on each side (see *e.g.* Fig. 1).

Observe that for $x = DHS_h(h_N, h_E, h_S, h_W)_e^b$ and $x' = DHS_h(h'_N, h'_E, h'_S, h'_W)_{e'}^{b'}$ catenation $x \circ x'$ is defined if and only if e matches to b' , *i.e.*:

$$\begin{aligned}
 e = N &\quad \text{and} \quad b' = S &\quad \text{or} \\
 e = E &\quad \text{and} \quad b' = W &\quad \text{or} \\
 e = S &\quad \text{and} \quad b' = N &\quad \text{or} \\
 e = W &\quad \text{and} \quad b' = E
 \end{aligned}$$

and $h_e = h'_{b'}$.

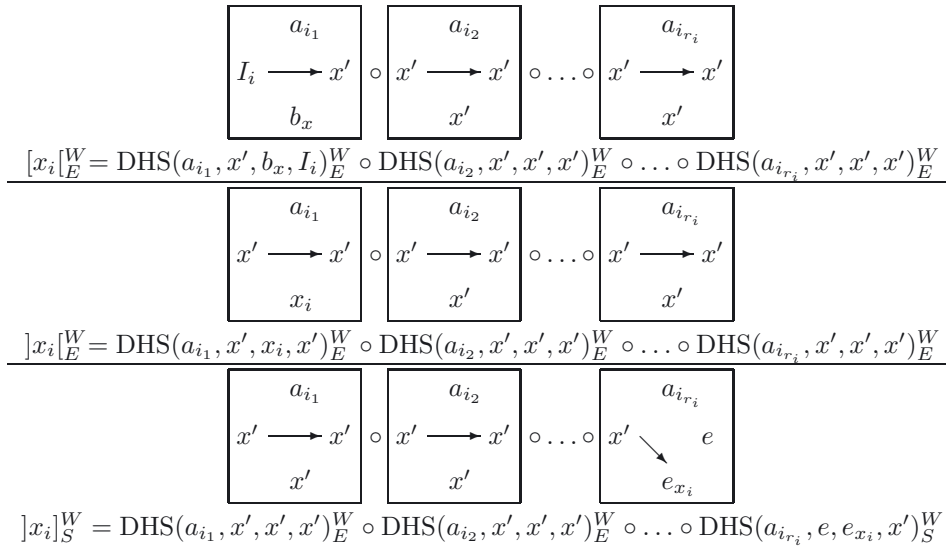


FIGURE 2. Basic-figures for $x_i = a_{i_1} \cdots a_{i_{r_i}}$.

Now we encode the Post problem in a set of directed figures over $\Sigma = \{a\}$. The Post problem can be stated as follows:

Let $A = \{a_1, \dots, a_p\}$ be a finite alphabet, $x_1, \dots, x_k, y_1, \dots, y_k \in A^+$ such that $x_i \neq y_i$ (for $i \in \{1, \dots, k\}$). One cannot decide if there exists $i_1, \dots, i_n \in \{1, \dots, k\}$, $n \geq 2$, such that $x_{i_1} \cdots x_{i_n} = y_{i_1} \cdots y_{i_n}$.

We describe a set of directed figures X such that the Post problem has a solution if and only if X is a code.

Consider the following set:

$$H = \bigcup_{i \in \{1, \dots, k\}} \{x_i, y_i, e_{x_i}, y_{y_i}, I_i\} \cup \{a_i \mid i \in \{1, \dots, p\}\} \cup \{x, y, x', y', b_x, b_y, e\},$$

where I_i are additional elements related to each pair (x_i, y_i) from corresponding Post Problem. Set $h = |H| = 5k + p + 7$. We can define a bijection between H and $\{1, \dots, h\}$, so from now on, each element of H is identified with its image by this bijection. Since h is fixed in our consideration we write $\text{DHS}(h_N, h_E, h_S, h_W)_e^b$ instead of $\text{DHS}_h(h_N, h_E, h_S, h_W)_e^b$.

For each $x_i \in X$ we define *basic-figures* $[x_i[$, $]x_i]$ and $]x_i]$ (Fig. 2); these figures will be used to encode the word x_i standing at the beginning (we call it *begin solution figure*), in the middle (*middle solution figure*) and at the end (*end solution figure*) of a solution of the Post problem, respective.

In addition we define *annex-figures* (Figs. 3–6); first three types of these figures will be used to convey information about solution of the Post problem from north to south, from north to west and from east to west respective, last one will convey no information.

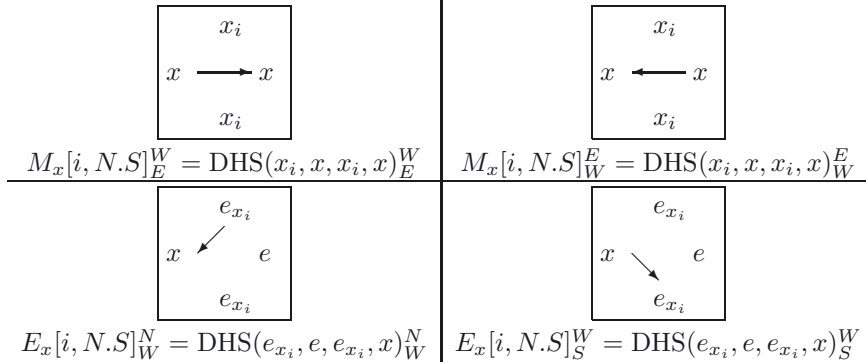


FIGURE 3. Annex-figures for passing information from north to south.

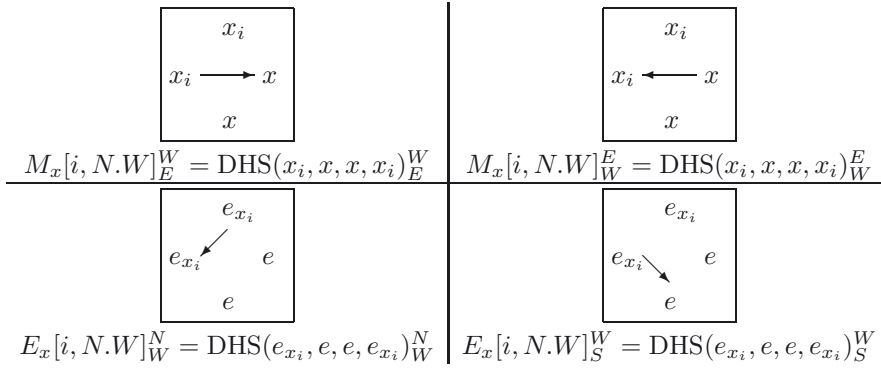


FIGURE 4. Annex-figures for passing information from north to west.

In the same way we define figures for “ y -part” of the Post problem, replacing the letter x by the letter y .

Let \mathbf{X} be the set of all defined figures ($6k$ basic-figures and $32k + 2$ annex-figures, $16k$ for each parts: “ x -part” and “ y -part”). Observe that there exists no half-plane of integer values anchored in $(0, 0)$ (i.e. $\{v \in \mathbb{Z}^2 \mid u \cdot v > 0\}$ for some $u \in \mathbb{Z}^2$, where \cdot denotes the usual dot product) containing all translation vectors of the figures we have defined.

Proposition 2.4. *If the Post problem has a solution then \mathbf{X} is not a code.*

Proof. Let $i_1, \dots, i_n \in \{1, \dots, k\}$ be a solution of the Post problem, i.e.

$$x_{i_1} \cdots x_{i_n} = y_{i_1} \cdots y_{i_n}.$$

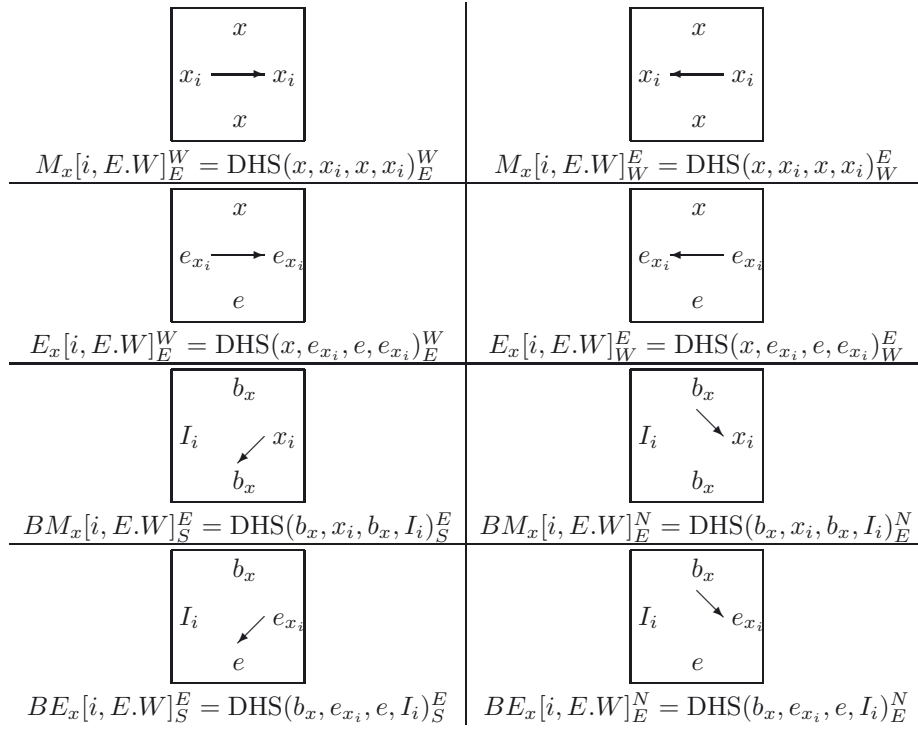


FIGURE 5. Annex-figures for passing information from east to west. Consider following directed figures:

$$wx_1 = [x_{i_1} [{}^W_E \circ] x_{i_2} [{}^W_E \circ] \dots \circ] x_{i_{n-1}} [{}^W_E \circ] x_{i_n}]_S^W,$$

$$wx_j = E_x[i_n, N.S]_W^N \circ \underbrace{N_x [{}^E_W \circ] \dots \circ N_x [{}^E_W \circ]}_{|x_n|-2 \text{ times}} \circ$$

$$\underbrace{N_x [{}^E_W \circ] \dots \circ N_x [{}^E_W \circ]}_{|x_{n-1}|-1 \text{ times}} \circ M_x[i_{n-1}, N.E]_W^E \circ$$

$$\dots$$

$$\underbrace{N_x [{}^E_W \circ] \dots \circ N_x [{}^E_W \circ]}_{|x_{j+1}|-1 \text{ times}} \circ M_x[i_{j+1}, N.E]_W^E \circ$$

$$\underbrace{N_x [{}^E_W \circ] \dots \circ N_x [{}^E_W \circ]}_{|x_j|-1 \text{ times}} \circ M_x[i_j, N.E]_W^E \circ$$

$$\underbrace{M_x[i_j, E.W]_W^E \circ \dots \circ M_x[i_j, E.W]_W^E}_{|x_{i_1} \dots x_{i_{j-1}}|-1 \text{ times}} \circ BM_x[i_j, E.W]_S^E$$

(for even $j < n$),

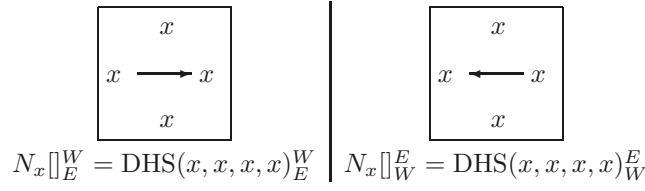


FIGURE 6. Annex-figure which pass no information.

$$\begin{aligned}
 wx_j &= BM_x[i_j, E.W]_E^N \circ \underbrace{M_x[i_j, E.W]_E^W \circ \dots \circ M_x[i_j, E.W]_E^W}_{|x_{i_1} \dots x_{i_{j-1}}| - 1 \text{ times}} \circ \\
 &M_x[i_j, N.E]_E^W \circ \underbrace{N_x \square_E^W \circ \dots \circ N_x \square_E^W}_{|x_j| - 1 \text{ times}} \circ \\
 &M_x[i_{j+1}, N.E]_E^W \circ \underbrace{N_x \square_E^W \circ \dots \circ N_x \square_E^W}_{|x_{j+1}| - 1 \text{ times}} \circ \\
 &\dots \\
 &M_x[i_{n-1}, N.E]_E^W \circ \underbrace{N_x \square_E^W \circ \dots \circ N_x \square_E^W}_{|x_{n-1}| - 1 \text{ times}} \circ \\
 &\underbrace{N_x \square_E^W \circ \dots \circ N_x \square_E^W}_{|x_n| - 2 \text{ times}} \circ E_x[i_n, N.S]_S^W \\
 &\text{(for odd } j < n),
 \end{aligned}$$

$$\begin{aligned}
 wx_n &= E_x[i_n, N.W]_W^N \circ \underbrace{E_x[i_n, E.W]_W^E \circ \dots \circ E_x[i_n, E.W]_W^E}_{|x_{i_1} \dots x_{i_n}| - 2 \text{ times}} \circ BE_x[i_n, E.W]_S^E \\
 &\text{(if } n \text{ is even),} \\
 wx_n &= BE_x[i_n, E.W]_E^N \circ \underbrace{E_x[i_n, E.W]_E^W \circ \dots \circ E_x[i_n, E.W]_E^W}_{|x_{i_1} \dots x_{i_n}| - 2 \text{ times}} \circ E_x[i_n, N.W]_S^W \\
 &\text{(if } n \text{ is odd).}
 \end{aligned}$$

In the same way we define figures wy_1, \dots, wy_n .

It is easy to see that $wx_1 \circ \dots \circ wx_n = wy_1 \circ \dots \circ wy_n \subseteq X^\diamond$. Hence \mathbf{X} is not a code. \square

Example 2.5. Consider

$$\begin{aligned}
 \Sigma &= \{a, b\}, \\
 X &= (x_1, x_2, x_3) = (b, ab, bab), \\
 Y &= (y_1, y_2, y_3) = (ba, abb, b).
 \end{aligned}$$

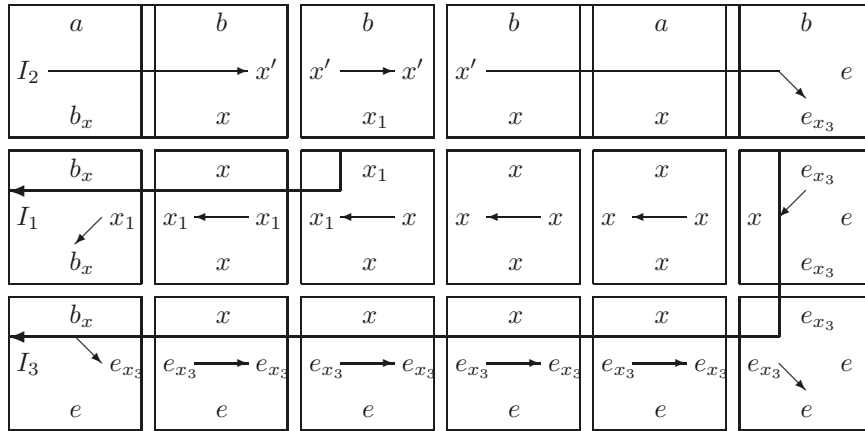


FIGURE 7. “X”-tiling of f .

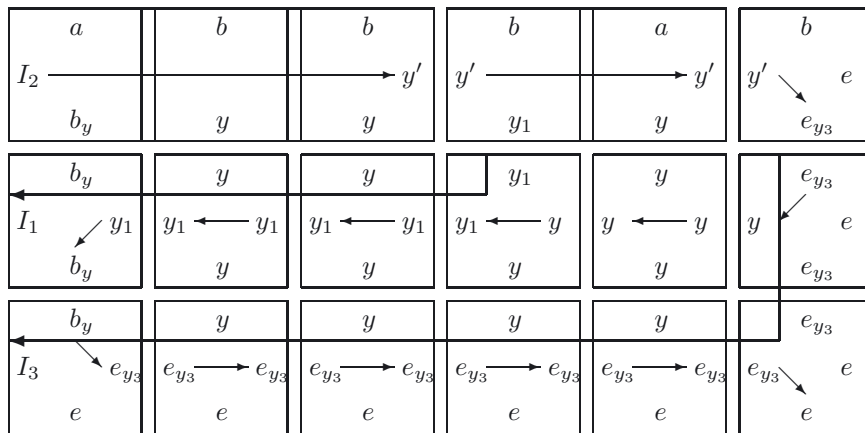


FIGURE 8. “Y”-tiling of f .

We have $x_2x_1x_3 = y_2y_1y_3$. Figure f with two different tiling with elements of \mathbf{X} is presented in Figures 7 and 8 (where thick arrows show the flow of information through annex-figures).

Proposition 2.6. *If \mathbf{X} is not a code then the related Post problem has a solution.*

Proof. Let f be a figure of minimal size (w.r.t. to the size of its domain) which admits two tilings with elements of \mathbf{X} , i.e. there exist $f_1, \dots, f_p, g_1, \dots, g_q \in \mathbf{X}$ such that $f_1 \neq g_1$ and $f_1 \circ \dots \circ f_p = g_1 \circ \dots \circ g_q$.

Consider directed hooked squares tiling f (these are annex-figures and squares of which basic-figures are built). Let d be the westmost among the northmost of them. We have following possibilities:

Case 1: $d \in \bigcup_{z \in \{x,y\}} \bigcup_{j \in \{1, \dots, k\}} \{E_z[j, N.S]_W^N, E_z[j, N.W]_W^N, E_z[j, E.W]_W^N\}$.

Since d is the westmost among the northmost of all squares tiling f , it cannot have north and west neighbour squares, *i.e.* squares hooked to it at north side and west side, respective. This implies that $f = d$, which contradicts the definition of the double tiling of f .

Case 2: $d \in \bigcup_{z \in \{x,y\}} \bigcup_{j \in \{1, \dots, k\}} \{E_z[j, N.S]_S^W, E_z[j, N.W]_S^W\}$.

Since d has no north and west neighbours, north and west hooks of d are uniquely determined by f . Each of listed figures is uniquely determined by its north and west hooks. This implies that d is also uniquely determined by f . d has no west neighbour and it has the start point at its west side, which implies that it must be the first one in sequence of figures whose catenation gives f , *i.e.* $d = f_1 = g_2$. Then either $f = d$ (contradiction as previously), or $f' = f_2 \circ \dots \circ f_p = g_2 \circ \dots \circ g_q$ is a smaller figure with two tilings, which contradicts the minimality of f .

Case 3: $d \in \bigcup_{z \in \{x,y\}} \bigcup_{j \in \{1, \dots, k\}} \{M_z[j, N.S]_W^E, M_z[j, N.W]_W^E, M_z[j, E.W]_W^E\} \cup \{N_x \square_W^E, N_y \square_W^E\}$.

As in *Case 1*, d is uniquely determined by f . d has no west neighbour and it has the end point at its west side, which implies that it must be the last one in sequence of figures whose catenation gives f , *i.e.* $d = f_p = g_q$. Then either $f = d$ (contradiction as previously) or $f' = f_1 \circ \dots \circ f_{p-1} = g_1 \circ \dots \circ g_{q-1}$ is smaller figure with two tilings, which contradicts the minimality of f .

Case 4: $d \in \bigcup_{z \in \{x,y\}} \bigcup_{j \in \{1, \dots, k\}} \{BM_z[j, E.W]_E^N, BE_z[j, E.W]_E^N\}$.

Now d must be the first one in tiling since it has the start point at its north side and it is the northmost in tiling. Observe that there exists no square with e -hook at the north side. This implies that $BM_z[j, E.W]_E^N$ and $BE_z[j, E.W]_E^N$ ($z \in \{x, y\}$) cannot be the first elements of two different tiling sequences of f . This implies that d is uniquely determined by f and $d = f_1 = g_1$. Contradiction as in *Case 2*.

Case 5: $d \in \bigcup_{z \in \{x,y\}} \bigcup_{j \in \{1, \dots, k\}} \{BM_z[j, E.W]_S^E, BE_z[j, E.W]_S^E\}$

As in *Case 4*, d is uniquely determined by f . If $d = BE_z[j, E.W]_S^E$ (for $z \in \{x, y\}$) then d is the last element of a tiling sequence. Contradiction as in *Case 3*. If $d = BM_z[j, E.W]_S^E$ (for $z \in \{x, y\}$) then (for some $i \in \{1, \dots, p\}$) $f = f_1 \circ f_2 \circ \dots \circ f_{i-1} \circ d \circ f_{i+1} \circ \dots \circ f_p$, where $f_1 \in \{M_z[j, E.W]_W^E, M_z[j, N.W]_W^E\}$ and $f_2 = \dots = f_{i-1} = M_z[j, E.W]_W^E$. Contradiction as in *Case 1*.

This leads us to a conclusion that:

Case 6: Directed hooked square d is a part of a basic-figure. In particular d is a “first part” of f_1 and g_1 .

Now it is easy to observe the following properties of f 's tiling:

- (1) If f_1 is a figure that encodes one of the words from X , then all f_i ($i \in \{1, \dots, p\}$) are figures encoding “ x -part” of the related Post problem (since

- there is no figure that links a figure from “ x -part” with a figure from “ y -part”). In the same way, if f_1 encodes a word from Y , then f_i encodes “ y -part” of the Post problem. Similar statement is true for g_i ($i \in \{1, \dots, q\}$).
- (2) First “row” of figures in tiling is a sequence of middle solution figures (may be empty) which is ended by an end solution figure (that ends row) and may be started with a begin solution figure.
 - (3) Sequence of middle solution figures from the first row implies that in the tiling, leftmost column’s hooks (I_j hooks of some BM and BE annex-figures) correspond to the sequence of indices of words encoded by those figures.

This leads us to simple observation, that the only possible two tilings of f are tilings of the form defined in the proof of Proposition 2.4. This implies that the related Post problem has a solution. \square

Theorem 2.7. *It is not decidable whether a given set $X \subseteq \Sigma^\diamond$ is a code.*

Proof. Follows directly from Proposition 2.4 and Proposition 2.6 and the undecidability of the Post problem. \square

2.2. m -CODES

Definition 2.8 (m -code). $X \subseteq \Sigma^\diamond$ is an m -code if for any $x \in X_m^\diamond$ there exists only one sequence $x_1, \dots, x_k \in X$ such that $x = x_1 \circ_m \dots \circ_m x_k$.

Example 2.9. $\{\boxed{a}\diamond, \boxed{a}\} \subseteq \{a\}^\diamond$ is a π_1 -code.

Example 2.10. $\{x = \boxed{a}\diamond, y = \diamond\boxed{a}\} \subseteq \{a\}^\diamond$ is not a π_1 -code since $x \circ_{\pi_1} y = x \circ_{\pi_1} y \circ_{\pi_1} x \circ_{\pi_1} y = x = \boxed{\diamond a}$.

The results for the m -catenation are quite different from those for catenation (see [5]).

Theorem 2.11. *Let $X \subseteq \Sigma^\diamond$ be such that for any $u \in \mathbb{Z}^2$*

$$\bigcup_{x \in X} \{\text{tran}(x)\} \cap \{v \in \mathbb{Z}^2 \mid u \cdot v \leq 0\} \neq \emptyset. \tag{1}$$

Then X is not an m -code.

3. CODICITY VERIFICATION

Results on codicity verification in this section are based on our previous work [5], where they were obtained in the context of m -catenation with an additional restriction for the figures. Considerations presented here work for both types of catenation, with no constraints on the relative position of domain and both begin and end points of figures.

Let us consider a situation, when for a given $X = \{x_1, \dots, x_n\} \subseteq \Sigma^\diamond$ there exists a vector τ such

$$\forall x \in X : \tau \cdot \text{tran}(x) > 0.$$

Without loss of generality we can assume that

$$\forall x \in X : \text{begin}(x) = (0, 0)$$

and

$$\angle(R_{-\frac{\pi}{2}}(\tau), \text{tran}(x_1)) \leq \angle(R_{-\frac{\pi}{2}}(\tau), \text{tran}(x_2)) \leq \dots \leq \angle(R_{-\frac{\pi}{2}}(\tau), \text{tran}(x_n)),$$

where \angle denotes an angle between two vectors, and R_ϕ denotes a rotation by ϕ . Now choose constants $r_E, r_N, r_W, r_S > 0$ such that the vectors

$$\begin{aligned} \tau_E &= r_E \tau, \\ \tau_N &= r_N R_{\frac{\pi}{2}}(\text{tran}(x_n)), \\ \tau_W &= -r_W \tau, \\ \tau_S &= r_S R_{-\frac{\pi}{2}}(\text{tran}(x_1)) \end{aligned}$$

define a “bounding area” for figures in X , *i.e.*,

$$\forall x \in X : \text{dom}(x) \cup \{\text{end}(x)\} \subseteq \bigcap_{u \in \{\tau_E, \tau_N, \tau_W, \tau_S\}} \{\text{HP}(u, \text{begin}(x))\},$$

where for $u, v \in \mathbb{Z}^2$ $\text{HP}(u, v)$ denotes half-plane $\{w \in \mathbb{Z}^2 \mid u \cdot (w - (v + u)) \leq 0\}$ (see Fig. 9).

For $x \in X_\bullet^\diamond$ define

$$\begin{aligned} \text{CE}^+(x) &= \text{HP}(\tau_s, \text{end}(x)) \cap \text{HP}(\tau_n, \text{end}(x)) \cap \text{HP}(\tau_w, \text{end}(x)), \\ \text{CE}^-(x) &= \mathbb{Z}^2 \setminus \text{CE}^+(\text{end}(x)), \\ \text{CW}^+(x) &= \bigcup_v \{v + (\text{CE}^+(\text{end}(x)) \cap \text{HP}(\tau_e, \text{end}(x)))\}, \\ \text{CW}^-(x) &= \mathbb{Z}^2 \setminus \text{CW}^+(\text{end}(x)), \end{aligned}$$

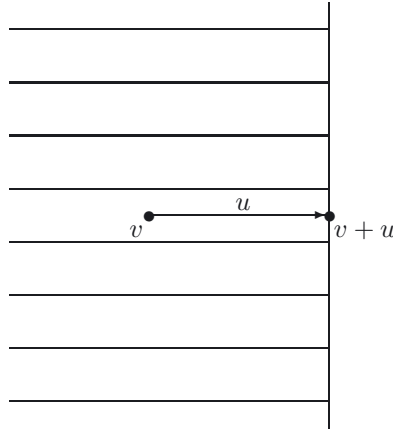


FIGURE 9. $HP(u, v)$. The half-plane contains integer grid points lying on vertical line and to the left side of that line (the region marked by horizontal lines).

where the union in the definition of $CW^+(x)$ is taken over $v \in \mathbb{Z}^2$ lying within an angle spanned by vectors $-\tau(x_1)$ and $-\tau(x_n)$ (see Figs. 10 and 11).

Immediately from definition we have following properties:

Proposition 3.1. *Let $x, y \in X_\bullet^\diamond$. Labels of $CE^-(x)$ cannot be changed or defined in $x \bullet y$:*

$$\begin{aligned} u \in CE^-(x) \cap \text{dom}(x) &\Rightarrow \text{label}(x)(u) = \text{label}(x \bullet y)(u), \\ u \in CE^-(x) \setminus \text{dom}(x) &\Rightarrow u \notin \text{dom}(x \bullet y). \end{aligned}$$

Proposition 3.2. *Let $x \in X_\bullet^\diamond$. Labels of $CW^-(x)$ are not defined in x :*

$$u \in CW^-(x) \Rightarrow u \notin \text{dom}(x).$$

Proposition 3.3.

$$\forall x, y \in X_\bullet^\diamond : CE^+(x \bullet y) \subseteq CE^+(x).$$

Proposition 3.4.

$$\forall x, y \in X_\bullet^\diamond : CW^+(x) \subseteq CW^+(x \bullet y).$$

For $x, y \in X_\bullet^\diamond$ we define a *configuration* as a pair (x, y) . We say that for $x', y' \in X_\bullet^\diamond$ configuration (x', y') is a *successor* of (x, y) and write $(x, y) \prec (x', y')$ if for some $z \in X$ either $x' = x \bullet z$ or $y' = y \bullet z$. By \prec^* we denote the transitive closure of \prec . Obviously we have

$$X \text{ is not a code (resp. } m\text{-code)} \Leftrightarrow \exists x, y \in X, z \in X_\bullet^\diamond : x \neq y, (x, y) \prec^* (z, z).$$

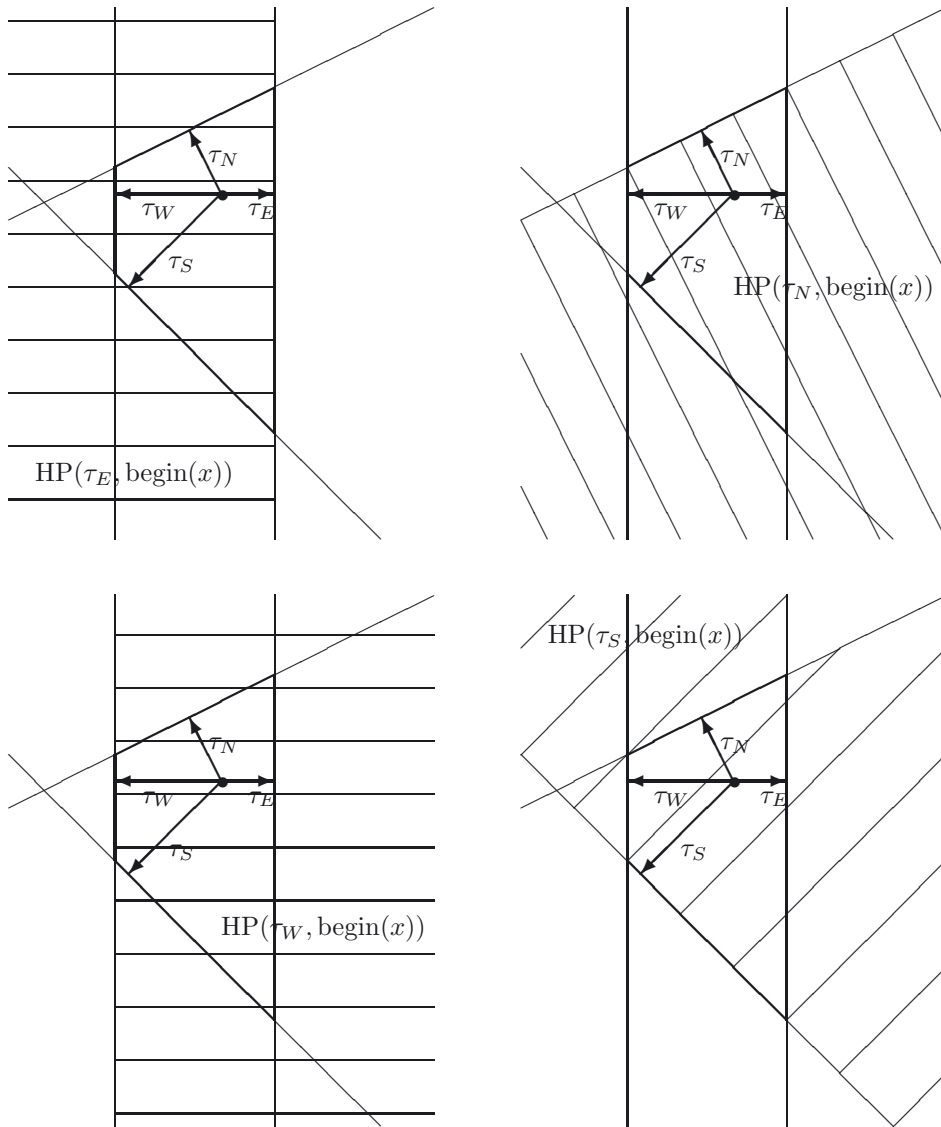


FIGURE 10. Half-planes $HP(\tau, \text{begin}(x))$ ($\tau \in \{\tau_E, \tau_N, \tau_W, \tau_S\}$); the black dot denotes the start point of x .

Our goal is either to find a configuration (x, y) such that

$$(x, y) \prec \dots \prec (z, z) \tag{2}$$

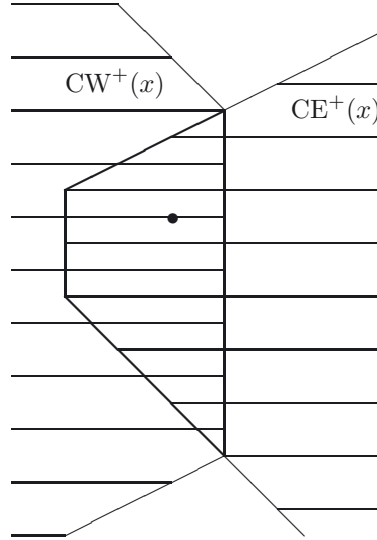


FIGURE 11. $CW^+(x)$ and $CE^+(x)$ regions; the black dot denotes the end point of x .

– then X is not a code (resp. m -code), or to prove that such a configuration does not exist – then X is a code (resp. m -code). A configuration satisfying (2) is called a *proper configuration*.

Proposition 3.5. *If (x, y) is proper and $(x', y') \prec (x, y)$, then (x', y') is proper.*

Proof. Obvious. □

Let

$$\rho = \max_{x \in X} \{ \text{minimal natural number } n \text{ such that } B(\text{begin}(x), n) \cap \text{dom}(x) \neq \emptyset \},$$

where for $u = (u_x, u_y) \in \mathbb{Z}^2$ and $n \in \mathbb{N}$ $B(u, n)$ denotes a ball on integer grid with center u and radius n , i.e.,

$$B(u, n) = \{ (v_x, v_y) \in \mathbb{Z}^2 \mid |u_x - v_x| + |u_y - v_y| \leq n \}$$

(see Fig. 12).

Proposition 3.6. *If (x, y) is proper, then*

$$\begin{aligned} B(\text{end}(x), \rho) \cap (CW^+(y) \cup CE^+(y)) &\neq \emptyset, \\ B(\text{end}(y), \rho) \cap (CW^+(x) \cup CE^+(x)) &\neq \emptyset. \end{aligned}$$

Proof. See definitions of $CW^+(\cdot)$ and $CE^+(\cdot)$. □

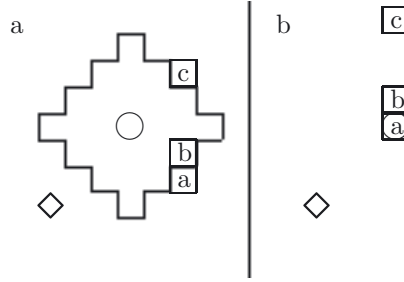


FIGURE 12. Directed figures and minimal balls (marked with thick lines).

Proposition 3.7. *If (x, y) is proper, then*

$$\text{label}(x) \upharpoonright_{\text{CE}^-(x) \cap \text{CE}^-(y)} \equiv \text{label}(y) \upharpoonright_{\text{CE}^-(x) \cap \text{CE}^-(y)}.$$

Proof. See definition of $\text{CE}^-()$ and Proposition 3.1. □

Notice that we do not need all of the information contained in configurations, just those labellings that can be changed by future catenations. By Proposition 3.7, instead of (x, y) we can consider a *reduced configuration* defined as a pair $(\pi_{RC}(x, y), \pi_{RC}(y, x))$ where

$$\pi_{RC}(z, z') = (\text{end}(z), \text{label}(z) \upharpoonright_{\text{dom}(z) \setminus (\text{CE}^-(z) \cap \text{CE}^-(z'))}).$$

Now Proposition 3.5 implies that we need only consider configurations where the span along τ_e is bounded by $|\tau_e|$, i.e., $|\tau_e \cdot (\text{end}(x) - \text{end}(y))| \leq |\tau_e|^2$, since no single figure advances $\text{end}(x)$ or $\text{end}(y)$ by more than $|\tau_e|$. Moreover, Proposition 3.6 restricts the perpendicular span (in the direction of $R_{-\frac{\pi}{2}}(\tau_e)$). Hence the number of reduced configurations, up to translation, is finite.

This leads us to the following result:

Theorem 3.8. *Let $X \subseteq \Sigma^\diamond$ be such that there exists $u \in \mathbb{Z}^2$*

$$\bigcup_{x \in X} \{\text{tran}(x)\} \subseteq \{v \in \mathbb{Z}^2 \mid u \cdot v > 0\}. \tag{3}$$

It is decidable whether X is a code (resp. m -code).

In particular, for m -catenation we have:

Corollary 3.9. *It is decidable whether a given finite set $X \subseteq \Sigma^\diamond$ is a m -code.*

Proof. If for any $u \in \mathbb{Z}^2$ condition (1) holds then, from Theorem 2.11, X is not an m -code. On the other hand, if there exists $u \in \mathbb{Z}^2$ such that condition (3) holds, from Theorem 3.8, we can verify if X is an m -code. □

Observe that from what we have presented we can obtain an algorithm that verifies whether a given set is a code by either finding identity configuration $(x_{i_1} \bullet \dots \bullet x_{i_k}, x_{j_1} \bullet \dots \bullet x_{j_l})$ or checking that there is no such configuration.

4. FINAL REMARKS

We are now interested in the complexity of the codicity verification algorithms. Apparently it depends on the angle spanned by the translation vectors of figures. Bigger angles obviously give higher complexity, and when the angle tends to zero, the algorithm becomes similar to the well-known Sardinas-Patterson algorithm. We also plan to study the specific case of figures with parallel translation vectors, resembling partial words. We hope that in this case, with possible additional constraints, the defect effect could be saved, *cf.* [5,7].

Acknowledgements. The author is thankful to Włodzimierz Moczurad for fruitful discussions and useful comments.

REFERENCES

- [1] P. Aigrain and D. Beauquier, Polyomino tilings, cellular automata and codicity. *Theoret. Comput. Sci.* **147** (1995) 165–180.
- [2] D. Beauquier and M. Nivat, A codicity undecidable problem in the plane. *Theoret. Comput. Sci.* **303** (2003) 417–430.
- [3] J. Berstel and D. Perrin, *Theory of Codes*. Academic Press (1985).
- [4] G. Costagliola, F. Ferrucci and C. Gravino, Adding symbolic information to picture models: definitions and properties. *Theoret. Comput. Sci.* **337** (2005) 51–104.
- [5] M. Kolarz and W. Moczurad, Directed figure codes are decidable. *Discrete Mathematics and Theoretical Computer Science* **11** (2009) 1–14.
- [6] S. Mantaci and A. Restivo, Codes and equations on trees. *Theoret. Comput. Sci.* **255** (2001) 483–509.
- [7] W. Moczurad, Defect theorem in the plane. *RAIRO-Theor. Inf. Appl.* **41** (2007) 403–409.
- [8] M. Moczurad and W. Moczurad, Decidability of simple brick codes, in *Mathematics and Computer Science, Vol. III (Algorithms, Trees, Combinatorics and Probabilities)*. Trends in Mathematics, Birkhäuser (2004), 541–542.
- [9] M. Moczurad and W. Moczurad, Some open problems in decidability of brick (labelled polyomino) codes, in *Cocoon 2004 Proceedings. Lect. Notes Comput. Sci.* **3106** (2004) 72–81.

Communicated by A. Restivo.

Received March 18, 2009. Accepted November 5, 2010.