Reducing local alphabet size in recognizable picture languages

Stefano Crespi Reghizzi Antonio Restivo **Pierluigi San Pietro**[1]

[1]DEIB, Politecnico di Milano, Italy
[2]Dipartimento di Matematica e Informatica, Università di Palermo, Italy

DLT 2021, Porto

Recognizable Picture languages

- A *picture* is a rectangular array of letters over an alphabet Σ.
  A *picture language* over Σ is a set of pictures.
- *Recognizable* picture languages, REC, can be defined as *projections of local languages* [Giammarresi and Restivo, 1997].
  The class REC has many important properties that make this definition a good candidate for being the 2D "equivalent" of regular languages.
- A language is *local* if membership of a picture p can be decided by checking if all its subpictures of size 2×2 (called *tiles*) are included in a given set.
- More precisely, local languages are defined using tiles over *bordered pictures* (pictures surrounded with a border of a new symbol $\#$).
  The border is ignored in the projection, but it is necessary to define all local languages.

Example: unary language: #columns = 2·#rows.

A picture over $\Sigma = \{a\}$

a	a	a	a	a	a
a	a	a	a	a	a
a	a	a	a	a	a

Its preimage over Γ

$\searrow$	b	b	b	b	$\nearrow$
b	$\searrow$	b	b	$\nearrow$	b
b	b	$\searrow$	$\nearrow$	b	b

(the projection $\pi : \Gamma \to \Sigma$ is the obvious one:$\pi(\searrow) = \pi(b) = \pi(\nearrow) = a$)

A bordered picture

Its tile set over alphabet Γ

Example: unary language: #columns = 2·#rows.

A picture over $\Sigma = \{a\}$

a	a	a	a	a	a
a	a	a	a	a	a
a	a	a	a	a	a

Its preimage over Γ

$\searrow$	b	b	b	b	$\nearrow$
b	$\searrow$	b	b	$\nearrow$	b
b	b	$\searrow$	$\nearrow$	b	b

(the projection $\pi : \Gamma \to \Sigma$ is the obvious one:$\pi(\searrow) = \pi(b) = \pi(\nearrow) = a$)

A bordered picture

#	#	#	#	#	#	#	#
#	$\searrow$	b	b	b	b	$\nearrow$	#
#	b	$\searrow$	b	b	$\nearrow$	b	#
#	b	b	$\searrow$	$\nearrow$	b	b	#
#	#	#	#	#	#	#	#

Its tile set over alphabet Γ

$$\left\{ \begin{array}{|c|c|}\hline \# & \# \\ \hline \# & \searrow \\ \hline \end{array}, \begin{array}{|c|c|}\hline \# & b \\ \hline \# & \# \\ \hline \end{array}, \begin{array}{|c|c|}\hline \# & \# \\ \hline \# & \nearrow \\ \hline \end{array}, \begin{array}{|c|c|}\hline \# & \# \\ \hline b & \# \\ \hline \end{array}, \begin{array}{|c|c|}\hline \searrow & b \\ \hline b & \searrow \\ \hline \end{array}, \begin{array}{|c|c|}\hline \searrow & \nearrow \\ \hline \# & \# \\ \hline \end{array}, \ldots \right\}$$

Local and regular word languages

- The definition of REC is a generalization of the classical result (Y. Medvedev 1964, Eilenberg 1974):
  every regular language $R \subseteq \Sigma^*$ is the projection of a local language $L \subseteq \Gamma^*$.
- The local alphabet Γ is much larger than Σ.
- The *alphabetic ratio* $|\Gamma|/|\Sigma|$ is $O(n^2)$ where n is the number of states of a NFA recognizing R–it can grow unboundedly with the language.
- How small can the ratio be?
- no significant improvement is possible using local languages

Extended Medveded Theorem for word languages

Local languages are a member of McNaughton and Papert's infinite hierarchy of *k-strictly locally testable* (k-slt), languages, where the one-dimensional "tiles" have size k, $k \geq 2$ (where local = 2-slt, and slt= $\bigcup_{k \geq 2} k$-slt).

A natural question: what happens to the alphabetic ratio when using slt instead of local languages?

Theorem 8 of [Crespi-Reghizzi and San Pietro, 2012]. Also implicitly in a proof of [Thomas, 1982]

Every regular word language is the projection of a k-slt language, with constant (minimal) alphabetic ratio 2.

Hence, *the alphabetic ratio is language-independent* (but k is not).

Extended Medveded Theorem for word languages

Local languages are a member of McNaughton and Papert's infinite hierarchy of *k-strictly locally testable* (k-slt), languages, where the one-dimensional "tiles" have size k, $k \geq 2$ (where local = 2-slt, and slt= $\bigcup_{k \geq 2} k$-slt).

A natural question: what happens to the alphabetic ratio when using slt instead of local languages?

Theorem 8 of [Crespi-Reghizzi and San Pietro, 2012]. Also implicitly in a proof of [Thomas, 1982]

Every regular word language is the projection of a k-slt language, with constant (minimal) alphabetic ratio 2.

Hence, *the alphabetic ratio is language-independent* (but k is not).

Problem: Alphabetic ratio for picture languages in REC?

Can we also define the class REC using projections of slt *picture* languages, with a language-independent (i.e., *constant*) alphabetic ratio?

Theorem 4: constant value of alphabetic ratio

Every language in REC is the image under projection of some slt language L with constant alphabetic ratio **2**.

The alphabetic ratio 2 is also the smallest value possible (following immediately by the same property for words):

Theorem 3: minimality of the alphabetic ratio

There exists a REC picture language R such that for every slt language L, if R is the image of L under a projection, then the alphabetic ratio is at least 2.

Problem: Alphabetic ratio for picture languages in REC?

Can we also define the class REC using projections of slt *picture* languages, with a language-independent (i.e., *constant*) alphabetic ratio?

Theorem 4: constant value of alphabetic ratio

Every language in REC is the image under projection of some slt language L with constant alphabetic ratio **2**.

The alphabetic ratio 2 is also the smallest value possible (following immediately by the same property for words):

Theorem 3: minimality of the alphabetic ratio

There exists a REC picture language R such that for every slt language L, if R is the image of L under a projection, then the alphabetic ratio is at least 2.

Problem: Alphabetic ratio for picture languages in REC?

Can we also define the class REC using projections of slt *picture* languages, with a language-independent (i.e., *constant*) alphabetic ratio?

Theorem 4: constant value of alphabetic ratio

Every language in REC is the image under projection of some slt language L with constant alphabetic ratio **2**.

The alphabetic ratio 2 is also the smallest value possible (following immediately by the same property for words):

Theorem 3: minimality of the alphabetic ratio

There exists a REC picture language R such that for every slt language L, if R is the image of L under a projection, then the alphabetic ratio is at least 2.

Strictly Locally Testable Picture Languages

- A k-tile is a picture of size $k \times k$; the *k-tiling* of p is the set of k-tiles that are subpictures of the bordered picture $\hat{p}$.

k-slt picture languages

A picture language L is *k-strictly locally testable* (k-slt) if there is a set of k-tiles such that all the pictures of L have a k-tiling in the set.

- For $k = 2$ we obtain local languages.

Corollary 1: Expressive power of REC does not change

The family of languages obtained by projections of slt languages coincides with the family REC of recognizable picture languages.

Example: trade-off of alphabetic ratio vs. size k of tiles

R: unary pictures such that the number of columns is double the number of rows.
R defined by the 2-tiling of the picture on the left, over $\Gamma_3 = \{b, \searrow, \nearrow\}$.

Pre-image with 3 symbols

$\searrow$	b	b	b	b	$\nearrow$
b	$\searrow$	b	b	$\nearrow$	b
b	b	$\searrow$	$\nearrow$	b	b

Pre-image with 2 symbols

$\rightarrow$	b	b	b	b	$\rightarrow$
b	$\rightarrow$	b	b	$\rightarrow$	b
b	b	$\rightarrow$	$\rightarrow$	b	b

Pre-image of illegal picture

$\rightarrow$	b	b	b	b	b	$\rightarrow$
b	$\rightarrow$	b	b	b	$\rightarrow$	b
b	b	$\rightarrow$	$\rightarrow$	$\rightarrow$	b	b

- Merging letters $\searrow$ and $\nearrow$ into $\rightarrow$, the 2-tiling of the corresponding pre-image also allows illegal pictures, e.g. the one having the pre-image on the right.
- The smallest alphabet Γ_k for which a 2-tiling recognizes the illegal picture becomes 3 if, instead of a 2-tile, we use a 3-tile with the 2-tiling of the middle picture.

Example: trade-off of alphabetic ratio vs. size k of tiles

R: unary pictures such that the number of columns is double the number of rows.
R defined by the 2-tiling of the picture on the left, over $\Gamma_3 = \{b, \searrow, \nearrow\}$.

Pre-image with 3 symbols

$\searrow$	b	b	b	b	$\nearrow$
b	$\searrow$	b	b	$\nearrow$	b
b	b	$\searrow$	$\nearrow$	b	b

Pre-image with 2 symbols

$\rightarrow$	b	b	b	b	$\rightarrow$
b	$\rightarrow$	b	b	$\rightarrow$	b
b	b	$\rightarrow$	$\rightarrow$	b	b

Pre-image of illegal picture

$\rightarrow$	b	b	b	b	b	$\rightarrow$
b	$\rightarrow$	b	b	b	$\rightarrow$	b
b	b	$\rightarrow$	$\rightarrow$	$\rightarrow$	b	b

- Merging letters $\searrow$ and $\nearrow$ into $\rightarrow$, the 2-tiling of the corresponding pre-image also allows illegal pictures, e.g. the one having the pre-image on the right.
- The smaller alphabet $\Gamma_2 = \{b, \rightarrow\}$ suffices to eliminate the illegal picture in column 3 if, instead of a 2-slt, we use a 3-slt with the 3-tiling of the middle picture.

Example: trade-off of alphabetic ratio vs. size k of tiles

R: unary pictures such that the number of columns is double the number of rows.
R defined by the 2-tiling of the picture on the left, over $\Gamma_3 = \{b, \searrow, \nearrow\}$.

Pre-image with 3 symbols

$\searrow$	b	b	b	b	$\nearrow$
b	$\searrow$	b	b	$\nearrow$	b
b	b	$\searrow$	$\nearrow$	b	b

Pre-image with 2 symbols

$\rightarrow$	b	b	b	b	$\rightarrow$
b	$\rightarrow$	b	b	$\rightarrow$	b
b	b	$\rightarrow$	$\rightarrow$	b	b

Pre-image of illegal picture

$\rightarrow$	b	b	b	b	b	$\rightarrow$
b	$\rightarrow$	b	b	b	$\rightarrow$	b
b	b	$\rightarrow$	$\rightarrow$	$\rightarrow$	b	b

- Merging letters $\searrow$ and $\nearrow$ into $\rightarrow$, the 2-tiling of the corresponding pre-image also allows illegal pictures, e.g. the one having the pre-image on the right.
- The smaller alphabet $\Gamma_2 = \{b, \rightarrow\}$ suffices to eliminate the illegal picture in column 3 if, instead of a 2-slt, we use a 3-slt with the 3-tiling of the middle picture.

Example: trade-off of alphabetic ratio vs. size k of tiles

R: unary pictures such that the number of columns is double the number of rows.
R defined by the 2-tiling of the picture on the left, over $\Gamma_3 = \{b, \searrow, \nearrow\}$.

Pre-image with 3 symbols

$\searrow$	b	b	b	b	$\nearrow$
b	$\searrow$	b	b	$\nearrow$	b
b	b	$\searrow$	$\nearrow$	b	b

Pre-image with 2 symbols

$\rightarrow$	b	b	b	b	$\rightarrow$
b	$\rightarrow$	b	b	$\rightarrow$	b
b	b	$\rightarrow$	$\rightarrow$	b	b

Pre-image of illegal picture

$\rightarrow$	b	b	b	b	b	$\rightarrow$
b	$\rightarrow$	b	b	b	$\rightarrow$	b
b	b	$\rightarrow$	$\rightarrow$	$\rightarrow$	b	b

- Merging letters $\searrow$ and $\nearrow$ into $\rightarrow$, the 2-tiling of the corresponding pre-image also allows illegal pictures, e.g. the one having the pre-image on the right.
- The smaller alphabet $\Gamma_2 = \{b, \rightarrow\}$ suffices to eliminate the illegal picture in column 3 if, instead of a 2-slt, we use a 3-slt with the 3-tiling of the middle picture.

Idea of proof of Th. 4: collect k-tiles of the k-partition

Given $R \in REC$ over Σ, consider a *local* language L over a local alphabet Γ s.t. its projection is R.
In general, in this case the alphabetic ratio is greater than 2.

For simplicity we assume in the presentation that all pictures have both the*number of columns and the number of rows multiple of* k.
In this case, we can define a *k-partition of a preimage* (on the local alphabet Γ): the set of non-overlapping k-tiles covering the preimage.

Example of k-tiles of the k-partiton of preimages

Suppose that $\Sigma = \{a, a\}$ and R is again s.t #columns=2#rows, but both a and a allowed in any position.
The local alphabet is $\Gamma = \{b, \searrow, \nearrow, b, \searrow, \nearrow\}$.

Pictures in R

a	a	a	a	a	a	a	a
a	a	a	a	a	a	a	a
a	a	a	a	a	a	a	a
a	a	a	a	a	a	a	a

$\Longrightarrow$

Pre-images over Γ

$\searrow$	b	b	b	b	b	b	$\nearrow$
b	$\searrow$	b	b	b	b	$\nearrow$	b
b	b	$\searrow$	b	b	$\nearrow$	b	b
b	b	b	$\searrow$	$\nearrow$	b	b	b

Leftmost 4-tile w_1 *Rightmost 4-tile* w_2

a	a	a	a	a	a	a	a
a	a	a	a	a	a	a	a
a	a	a	a	a	a	a	a
a	a	a	a	a	a	a	a

$\Longrightarrow$

$\searrow$	b	b	b	b	b	b	$\nearrow$
b	$\searrow$	b	b	b	b	$\nearrow$	b
b	b	$\searrow$	b	b	$\nearrow$	b	b
b	b	b	$\searrow$	$\nearrow$	b	b	b

Leftmost 4-tile w_3 *Rightmost 4-tile* w_4

Example of k-tiles of the k-partiton of preimages

Suppose that $\Sigma = \{a, a\}$ and R is again s.t #columns=2#rows, but both a and a allowed in any position.
The local alphabet is $\Gamma = \{b, \searrow, \nearrow, b, \searrow, \nearrow\}$.

Pictures in R

a	a	a	a	a	a	a	a
a	a	a	a	a	a	a	a
a	a	a	a	a	a	a	a
a	a	a	a	a	a	a	a

$\Longrightarrow$

Pre-images over Γ

$\searrow$	b	b	b	b	b	b	$\nearrow$
b	$\searrow$	b	b	b	b	$\nearrow$	b
b	b	$\searrow$	b	b	$\nearrow$	b	b
b	b	b	$\searrow$	$\nearrow$	b	b	b

Leftmost 4-tile w_1 *Rightmost 4-tile* w_2

a	a	a	a	a	a	a	a
a	a	a	a	a	a	a	a
a	a	a	a	a	a	a	a
a	a	a	a	a	a	a	a

$\Longrightarrow$

$\searrow$	b	b	b	b	b	b	$\nearrow$
b	$\searrow$	b	b	b	b	$\nearrow$	b
b	b	$\searrow$	b	b	$\nearrow$	b	b
b	b	b	$\searrow$	$\nearrow$	b	b	b

Leftmost 4-tile w_3 *Rightmost 4-tile* w_4

Example of k-tiles of the k-partiton of preimages

Suppose that $\Sigma = \{a, a\}$ and R is again s.t #columns=2#rows, but both a and a allowed in any position.
The local alphabet is $\Gamma = \{b, \searrow, \nearrow, b, \searrow, \nearrow\}$.

Pictures in R

a	a	a	a	a	a	a	a
a	a	a	a	a	a	a	a
a	a	a	a	a	a	a	a
a	a	a	a	a	a	a	a

$\Longrightarrow$

Pre-images over Γ

$\searrow$	b	b	b	b	b	b	$\nearrow$
b	$\searrow$	b	b	b	b	$\nearrow$	b
b	b	$\searrow$	b	b	$\nearrow$	b	b
b	b	b	$\searrow$	$\nearrow$	b	b	b

Leftmost 4-tile w_1 *Rightmost 4-tile* w_2

a	a	a	a	a	a	a	a
a	a	a	a	a	a	a	a
a	a	a	a	a	a	a	a
a	a	a	a	a	a	a	a

$\Longrightarrow$

$\searrow$	b	b	b	b	b	b	$\nearrow$
b	$\searrow$	b	b	b	b	$\nearrow$	b
b	b	$\searrow$	b	b	$\nearrow$	b	b
b	b	b	$\searrow$	$\nearrow$	b	b	b

Leftmost 4-tile w_3 *Rightmost 4-tile* w_4

Idea of proof: define new k-tiles on different alphabet

Each k-tile w_i over Γ of the k-partitions is associated with a k-tile x_i over a new local alphabet $\Lambda = \{0,1\} \times \Sigma$. The binary component is a unique encoding of the *frame* of w_i, i.e., its border. The Σ component of x_i is defined so to follow the original projection to Σ.

$w_1 =$

k-tile over Γ

↘	b	b	b
b	↘	b	b
b	b	↘	b
b	b	b	↘

Frame

↘	b	b	b
b			b
b			b
b	b	b	↘

Binary encoding

1	1	1	1
1	0	1	0
0	1	1	1
0	0	0	0

k-tile over $\Lambda = \{0,1\} \times \Sigma$

$(1,a)$	$(1,a)$	$(1,a)$	$(1,a)$
$(1,a)$	$(0,a)$	$(1,a)$	$(0,a)$
$(0,a)$	$(1,a)$	$(1,a)$	$(1,a)$
$(0,a)$	$(0,a)$	$(0,a)$	$(0,a)$

Idea of proof: define new k-tiles on different alphabet

Each k-tile w_i over Γ of the k-partitions is associated with a k-tile x_i over a new local alphabet $\Lambda = \{0,1\} \times \Sigma$. The binary component is a unique encoding of the *frame* of w_i, i.e., its border. The Σ component of x_i is defined so to follow the original projection to Σ.

$w_1 =$

k-tile over Γ

↘	b	b	b
b	↘	b	b
b	b	↘	b
b	b	b	↘

Frame

↘	b	b	b
b			b
b			b
b	b	b	↘

Binary encoding

1	1	1	1
1	0	1	0
0	1	1	1
0	0	0	0

k-tile over $\Lambda = \{0,1\} \times \Sigma$

$(1,a)$	$(1,a)$	$(1,a)$	$(1,a)$
$(1,a)$	$(0,a)$	$(1,a)$	$(0,a)$
$(0,a)$	$(1,a)$	$(1,a)$	$(1,a)$
$(0,a)$	$(0,a)$	$(0,a)$	$(0,a)$

Idea of proof: define new k-tiles on different alphabet

Each k-tile w_i over Γ of the k-partitions is associated with a k-tile x_i over a new local alphabet $\Lambda = \{0,1\} \times \Sigma$. The binary component is a unique encoding of the *frame* of w_i, i.e., its border. The Σ component of x_i is defined so to follow the original projection to Σ.

$w_1 =$

k-tile over Γ

↘	b	b	b
b	↘	b	b
b	b	↘	b
b	b	b	↘

Frame

↘	b	b	b
b			b
b			b
b	b	b	↘

Binary encoding

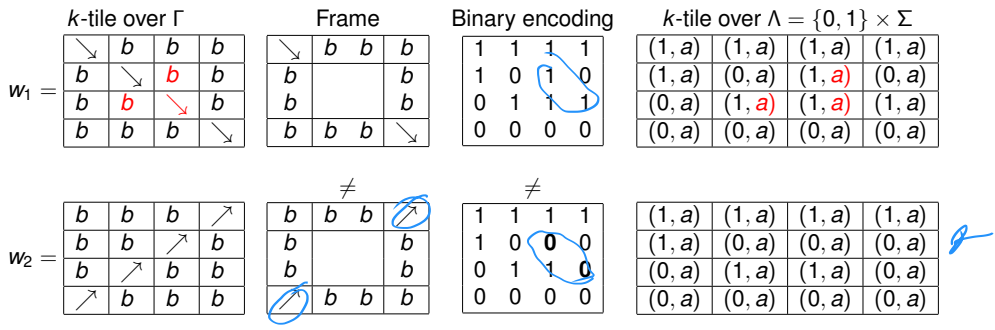
1	1	1	1
1	0	1	0
0	1	1	1
0	0	0	0

k-tile over $\Lambda = \{0,1\} \times \Sigma$

$(1,a)$	$(1,a)$	$(1,a)$	$(1,a)$
$(1,a)$	$(0,a)$	$(1,a)$	$(0,a)$
$(0,a)$	$(1,a)$	$(1,a)$	$(1,a)$
$(0,a)$	$(0,a)$	$(0,a)$	$(0,a)$

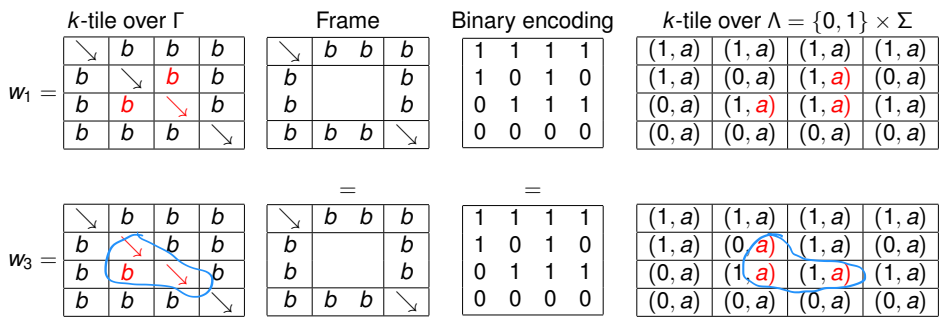
Idea of proof: define new k-tiles on different alphabet

Each k-tile w_i over Γ of the k-partitions is associated with a k-tile x_i over a new local alphabet $\Lambda = \{0,1\} \times \Sigma$. The binary component is a unique encoding of the *frame* of w_i, i.e., its border. The Σ component of x_i is defined so to follow the original projection to Σ.



Idea of proof: define new k-tiles on different alphabet

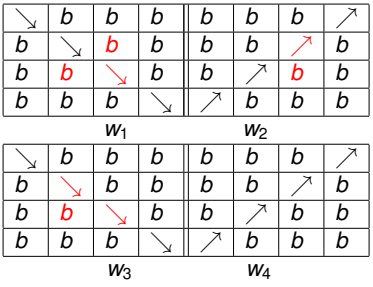
Each k-tile w_i over Γ of the k-partitions is associated with a k-tile x_i over a new local alphabet $\Lambda = \{0,1\} \times \Sigma$. The binary component is a unique encoding of the *frame* of w_i, i.e., its border. The Σ component of x_i is defined so to follow the original projection to Σ.



Interchangeability of k-tiles with same frame

A typical local property: Two k-tiles (verifying a correct 2-tiling) having the same frame can be interchanged (since a local test may only look at 2-tiles)

Previous preimages

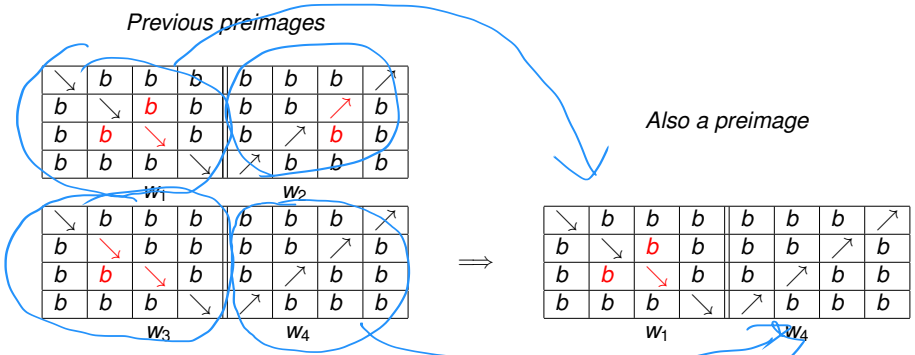


Also a preimage

This explains why we can only encode the frame of a k-tile, rather than the whole k-tile. This will be is crucial to find enough unique encodings.

Interchangeability of k-tiles with same frame

A typical local property: Two k-tiles (verifying a correct 2-tiling) having the same frame can be interchanged (since a local test may only look at 2-tiles)



This explains why we can only encode the frame of a k-tile, rather than the whole k-tile. This will be is crucial to find enough unique encodings.

In summary: define set Z_k of all preimages on alphabet Λ

a	a	a	a	a	a	a	a
a	a	a	a	a	a	a	a
a	a	a	a	a	a	a	a
a	a	a	a	a	a	a	a

$p \in R$ over Σ

$\searrow$	b	b	b	b	b	b	$\nearrow$
b	$\searrow$	b	b	b	b	$\nearrow$	b
b	b	$\searrow$	b	b	$\nearrow$	b	b
b	b	b	$\searrow$	$\nearrow$	b	b	b

preimage of p over Γ

$(1,a)$	$(1,a)$	$(1,a)$	$(1,a)$	$(1,a)$	$(1,a)$	$(1,a)$	$(1,a)$
$(1,a)$	$(0,a)$	$(0,a)$	$(0,a)$	$(1,a)$	$(0,a)$	$(0,a)$	$(0,a)$
$(0,a)$	$(1,a)$	$(1,a)$	$(0,a)$	$(0,a)$	$(1,a)$	$(1,a)$	$(0,a)$
$(0,a)$	$(0,a)$	$(0,a)$	$(0,a)$	$(0,a)$	$(0,a)$	$(0,a)$	$(0,a)$

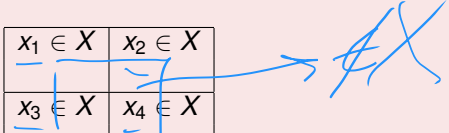
preimage of p over Λ (in Z_k)

Comma-free picture codes

- The image of Z_k (the set of all preimages over $\Lambda = \{0,1\} \times \Sigma$), under the projection from Λ to its component Σ, is the original language R.
- To be slt, the k-tiling of Z_k must not allow erroneous pictures.
  The problem can be solved using *comma-free picture codes* for the binary encodings, that do not allow ambigous overlappings.

Comma-free picture codes

A finite set X of $k \times k$ pictures (called code-pictures) such that for all $x_1, x_2, x_3, x_4 \in X$ the following $2k \times 2k$ pictures do not include *internal pictures* in X:



Example of comma-free picture

Let X be a comma-free picture code.

Code-pictures in X

1	1	1	1
1	0	1	0
0	1	1	1
0	0	0	0

1	1	1	1
1	0	0	0
0	1	1	0
0	0	0	0

A picture using X

1	1	1	1	1	1	1	1
1	0	1	0	1	0	0	0
0	1	1	1	0	1	1	0
0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1
1	0	1	0	1	0	0	0
0	1	1	1	0	1	1	0
0	0	0	0	0	0	0	0

$\notin X$

Every internal 4-tile is not in X.

As a consequence, the set of picture composed of (horizontal/vertical) concatenations of code-pictures in X is a $2k$-slt language.

Application of comma-free picture codes

To complete the proof, we need to find k such that the number of comma-free picture codes with k bits is large enough to encode all k-tiles in the the k-partition.
The frame to be encoded grows as Γ^{4k}, i.e, $2^{O(k)}$.

- From recent results on *non-overlapping* picture codes [Anselmo et al., 2017], the cardinality of a binary comma-free picture code is $2^{\Theta(k^2)}$
- Therefore, $\exists k$ such that the number of codes is larger than the number of frames.

Theorem (comma-free picture codes and recognizability)

There exists $k \geq 2$ such that if the binary encodings used in the set Z_k of preimages on alphabet A is a binary comma-free picture code with k bits, then Z_k is $2k$-slt.

Application of comma-free picture codes

To complete the proof, we need to find k such that the number of comma-free picture codes with k bits is large enough to encode all k-tiles in the the k-partition.
The frame to be encoded grows as Γ^{4k}, i.e, $2^{O(k)}$.

- From recent results on *non-overlapping* picture codes [Anselmo et al., 2017], the cardinality of a binary comma-free picture code is $2^{\Theta(k^2)}$
- Therefore, $\exists k$ such that the number of codes is larger than the number of frames.

Binary comma-free picture codes are adequate

There exists $k \geq 2$ such that if the binary encodings used in the set Z_k of preimages on alphabet Λ is a binary comma-free picture code with k bits, then Z_k is $2k$-slt.

Application of comma-free picture codes

To complete the proof, we need to find k such that the number of comma-free picture codes with k bits is large enough to encode all k-tiles in the the k-partition.
The frame to be encoded grows as Γ^{4k}, i.e, $2^{O(k)}$.

- From recent results on *non-overlapping* picture codes [Anselmo et al., 2017], the cardinality of a binary comma-free picture code is $2^{\Theta(k^2)}$
- Therefore, $\exists k$ such that the number of codes is larger than the number of frames.

Binary comma-free picture codes are adequate

There exists $k \geq 2$ such that if the binary encodings used in the set Z_k of preimages on alphabet Λ is a binary comma-free picture code with k bits, then Z_k is $2k$-slt.

Extending the solutions to pictures of any size

In the presentation we only considered pictures with both number of columns and number of rows being multiple of k.

The paper shows how to extend the solution to the case of pictures of any size.

This is based on *padding* the pictures to make them of size multiple of k, applying the previous procedure and then removing the padding from the k-tiles so obtained.

The padding makes the local alphabet Γ of size $O(k^2)$ rather than a constant, but this only requires a larger value of k to obtain the final result.

Conclusions

- We showed that every recognizable picture language can be defined as the image under projection of a strictly locally testable language, using comma-free picture codes, with minimal alphabetic ratio 2.
- This generalizes the same results obtained for regular word languages and more recently for regular tree languages [Crespi-Reghizzi and San Pietro, 2021].
- An open question is how to compute the required value of k as a function of the local alphabet Γ and of a given alphabetic ratio.

Essential bibliography

M. Anselmo, D. Giammarresi, and M. Madonia. Non-expandable non-overlapping sets of pictures. *Theor. Comput. Sci.*, 657:127–136, 2017. doi: 10.1016/j.tcs.2016.09.025. URL https://doi.org/10.1016/j.tcs.2016.09.025.

S. Crespi-Reghizzi and P. San Pietro. From regular to strictly locally testable languages. *Int. J. Found. Comput. Sci.*, 23(8):1711–1728, 2012. doi: 10.1142/S0129054112400710. URL https://doi.org/10.1142/S0129054112400710.

S. Crespi-Reghizzi and P. San Pietro. Homomorphic characterization of tree languages based on comma-free encoding. In *LATA 2021*, LNCS 12638, pages 241–254. Springer, 2021. doi: 10.1007/978-3-030-68195-1_19. URL https://doi.org/10.1007/978-3-030-68195-1_19.

D. Giammarresi and A. Restivo. Two-dimensional languages. In G. Rozenberg and A. Salomaa, editors, *Handbook of formal languages, vol. 3*, pages 215–267. Springer, 1997. ISBN 3-540-60649-1.

W. Thomas. Classifying regular events in symbolic logic. *J. Comput. Syst. Sci.*, 25(3):360–376, 1982.