Closure Operations on Picture Languages and Floor Plans[1]

Extended abstract of a full paper appeared in DLT 2026

Stefano Crespi Reghizzi[1], Antonio Restivo[2] and Pierluigi San Pietro[1,*]

[1]*Politecnico di Milano - DEIB and CNR - IEIIT, Milan, Italy*
[2]*Dipartimento di Matematica e Informatica, Università di Palermo, Palermo, Italy*

Abstract

We study two natural two-dimensional analogues of the Kleene star: tiling closure and closure under horizontal and vertical concatenation. For recognizable (REC) picture languages, several questions that are elementary for word languages become undecidable: in particular, whether a local language is already closed under either operation, and whether a recognizable language contains only indecomposable pictures. On the positive side, every language closed under tiling closure or concatenation closure has a unique minimal basis. However, unlike the one-dimensional regular case, the basis of a recognizable picture language need not be recognizable.
We connect these closure operations with the widely investigated geometric model of floor plan, which represents rectangular partitions: we encode rectangular floor plans as ordinary pictures by materializing block contours with a finite alphabet of edge and corner symbols. The languages of basic floor plans, all floor plans, and triangulated floor plans are local. Guillotine floor plans coincide with the concatenation closure of basic floor plans; this language is not local, and its recognizability remains open. We show that a positive answer would imply closure of REC under concatenation closure. We also prove that strong equivalence classes of floor plans are recognizable, as well as the complement of the triangulated guillotine language.

Keywords

picture languages, recognizable picture languages, closure operations, floor plans, guillotine decompositions

1. Introduction

Recognizable word languages have a robust algebraic theory. In particular, the Kleene star of a regular language is regular, fixed-point questions for star-closed regular languages are decidable, and the minimal set of generators of a regular submonoid is again regular. The purpose of the full paper is to examine how much of this framework survives when words are replaced by finite rectangular arrays.

Two closure operations are considered. The first one is *tiling closure*: a picture belongs to the closure of a language L if it can be partitioned into rectangular blocks, each of which is a picture of L. The second one is *concatenation closure*: a picture belongs to the closure of L if it can be obtained from pictures of L by a finite expression using horizontal and vertical concatenation. The second operation is more restrictive, because each concatenation step cuts the current rectangle by one full horizontal or vertical line, whence the *guillotine* terminology comes. Thus every concatenation decomposition is a tiling, but not every tiling is a concatenation decomposition; the standard obstruction is a windmill pattern (e.g., in [2]). Some variants of the concatenation closure operation have been introduced in [3]. The T-closure is the essential operation of tile-rewriting picture grammars [4, 5].

The first contribution is an algebraic and decision-theoretic analysis of these two closures over the family REC, defined by the tiling systems of [6]. Basic indecomposability questions are undecidable already for simple cases. We then introduce the natural notion of basis (roughly a minimal generating set) —which for word languages was studied already in [7] and more recently in [8]—and prove that it is well-behaved: the basis exists and is unique for closed languages. Nevertheless, recognizability of the language does not imply recognizability of its basis.

[1]The full version of this summary was published in [1].

ICTCS 2026: Italian Conference on Theoretical Computer Science, September 7–9, 2026, Udine, Italy

*Corresponding author.
stefano.crespireghizzi@polimi.it (S. Crespi Reghizzi); antonio.restivo@unipa.it (A. Restivo); pierluigi.sanpietro@polimi.it (P. San Pietro)
0000-0001-5061-7402 (S. Crespi Reghizzi); 0000-0002-1972-6931 (A. Restivo); 0000-0002-2437-8716 (P. San Pietro)

The second contribution is a connection with rectangular partitions, also called floor plans or rectangulations. Floor plans are familiar in discrete geometry and in layout applications such as VLSI design. We encode them as picture languages by replacing geometric line segments with symbols occupying cells. This makes it possible to ask whether natural classes of floor plans are local or recognizable, and to relate guillotine decomposability to concatenation closure.

2. Pictures, closures, and recognizability

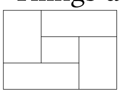
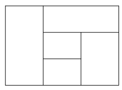
A *picture* over a finite alphabet Σ is a nonempty rectangular array of letters of Σ. If p has m rows and n columns, its size is written (m, n). A picture language is a subset of Σ^{++}, the set of all finite nonempty pictures over Σ.

For pictures p and q, the horizontal concatenation $p \obar q$ is defined when p and q have the same number of rows; it places q to the right of p. The vertical concatenation $p \ominus q$ is defined when p and q have the same number of columns; it places q below p. Both operations extend to languages in the usual way.

A *rectangular partition* of a picture domain is a finite family of pairwise disjoint rectangular subdomains (blocks) whose union is the whole domain. A partition is a *tiling of* p *with* L if every block, read as a subpicture of p, belongs to L. The tiling closure of L, denoted here by L^{++}, is the set of all pictures that admit such a tiling. The *concatenation* closure of L, denoted by $L^{\ominus\obar+}$, is the least language containing L and closed under both $\obar$ and $\ominus$. Equivalently, its elements admit a recursive guillotine decomposition into blocks from L. For every L,

$$L \subseteq L^{\ominus\obar+} \subseteq L^{++},$$

and the second inclusion may be proper, e.g., if the only tiling with L of a picture p is the windmill below, then $p \in L^{++} \setminus L^{\ominus\obar+}$.

Tilings are often depicted as rectangular partitions. For instance, this tiling is called a windmill  and this one is a guillotine decomposition: .

We use the standard tiling-system model for recognizable picture languages [9]. A language is *local* if membership is determined by the set of allowed overlapping 2×2 tiles in the picture surrounded by a reserved border symbol. A language is *recognizable*, or in REC, if it is the (alphabetic) projection of a local language. The family REC contains all local languages and is closed under union, intersection, and horizontal and vertical concatenation. It is also closed under tiling closure by Simplot's theorem (Prop. 5.1 of [10]). Whether REC is closed under concatenation closure is an open problem which motivates the floor-plan part of the paper.

3. Closure properties and bases

A language L is T-closed if $L = L^{++}$ and C-closed if $L = L^{\ominus\obar+}$. For word languages, the analogous question is whether $L = L^*$, and it is decidable for regular languages. For picture languages the situation changes sharply.

Theorem 1. *It is undecidable whether a given local picture language is T-closed. It is also undecidable whether it is C-closed.*

The proof reduces from the emptiness problem for local picture languages. Given a local language K over Σ, add a fresh symbol \$ and consider $L = K \cup \{\$\}^{++}$ over the disjoint union of alphabets. If K is empty, then L is just the full unary language over \$ and is closed under both operations. If K is nonempty, a picture p from K can be concatenated or tiled with a unary strip over \$, producing a picture outside L (since $\{\$\}^{++}$ contains pictures of every size, one can always choose a \$-picture matching p dimensions).

The next issue is whether a closed picture language has a canonical minimal generating set. A picture $p \in L$ is T-indecomposable if the only tiling of p with pictures of L is the trivial tiling with p itself. It is C-indecomposable if it is not a horizontal or vertical concatenation of two pictures of L. Denote by $\mathrm{IND}_T(L) \subseteq L$ and $\mathrm{IND}_C(L) \subseteq L$ the corresponding sets.

A set B is a T-basis of a T-closed language L if $L = B^{++}$ and no element of B is generated by the other elements of B under tiling closure. The definition of C-basis is the analogous one with concatenation closure.

Theorem 2. *It is undecidable whether a recognizable language L satisfies $L = \mathrm{IND}_T(L)$. It is also undecidable whether $L = \mathrm{IND}_C(L)$.*

For the tiling case, start from a recognizable language R and add one isolated fresh picture p to R^{++} (e.g., a $(1,1)$ picture with a fresh symbol, which cannot be tiled by smaller blocks). If R is empty, the resulting language contains only p; otherwise it contains a nontrivially tiled picture. The concatenation case is similar, using $R \cup (R \oplus R) \cup \{p\}$.

Theorem 3. *If L is T-closed, then $\mathrm{IND}_T(L)$ is the unique T-basis of L. If L is C-closed, then $\mathrm{IND}_C(L)$ is the unique C-basis of L.*

The proof is by induction on picture area. In the tiling case, every decomposable picture in a closed language can be tiled by strictly smaller pictures from the same language; by induction those smaller pictures are generated by indecomposable ones. The concatenation case is the same argument with the first horizontal or vertical cut. Minimality follows because an indecomposable picture cannot be generated from other elements of the language.

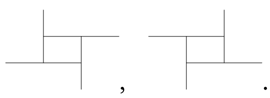
The existence and uniqueness of the basis mirrors the one-dimensional theory. The crucial difference is that the basis need not preserve recognizability.

Theorem 4. *There exists a recognizable language $L \in \mathrm{REC}$ that is both T-closed and C-closed, but whose T-basis and C-basis are not in* REC.

The construction encodes a known nonrecognizable condition: equality of two square subpictures over a non-unary alphabet. One defines a recognizable closed language L in which the intended basis consists of pictures containing exactly one complete marked occurrence of such a subpicture. If the basis were recognizable, closure of REC under intersection would make the double square language recognizable, contradicting the standard syntactic-equivalence argument for tiling systems (Syntactic Equivalence Lemma 9.2 of [9]). The same construction is arranged so that the tiling and concatenation bases coincide.

4. Floor plans as picture languages

We now interpret the two closure operations geometrically via rectangular partitions called floor plans. A *floor plan* (see e.g., [11, 12, 13] is a rectangular region subdivided into smaller rectangles. A triangulated floor plan is one in which no four blocks meet at a common point. A guillotine floor plan is one that can be obtained by recursively cutting a rectangle by full horizontal or vertical cuts. Equivalently, within triangulated floor plans, guillotine floor plans are exactly those avoiding the so-called *windmill*: a quadruple of segments forming one of the two following patterns:



To study floor plans as picture languages, we replace geometric contours with linear arrays of symbols. The alphabet Ξ contains one blank interior symbol, four directed edge symbols, and four directed corner symbols. Each block is represented by a clockwise rectangular contour, and its interior is filled with blanks. This yields the following languages of floor plan pictures, indicated by fpp:

- L_{basic}: pictures consisting of one single rectangular block;
- L_{fp}: all floor-plan pictures;
- L_{tfp}: triangulated floor-plan pictures;
- L_{guil}: guillotine floor-plan pictures.
- L_{tguil}: triangulated guillotine floor-plan pictures.

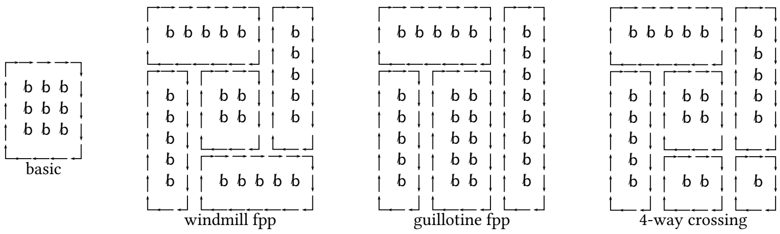


Figure 1: Floor plans represented as pictures. From left to right, an fpp in L_{basic}, a windmill pattern, a guillotine fpp in L_{guil}, and a fpp in $L_{\text{fp}} \setminus L_{\text{tfp}}$.

Fig. 1, from left to right, shows a basic floor plan picture, a windmill fpp, a guillotine one, and a floor plan picture with a four-way junction. Each block is delimited by a clockwise-oriented rectangular contour.

The contour encoding is deliberately local. It allows a 2×2 tile to check whether an edge continues in the correct direction, whether a corner has the right adjacent edges, and whether interiors remain inside the corresponding rectangle. The only additional local distinction needed for triangulation is the exclusion of the tile in which four corners meet.

Theorem 5. *The languages L_{basic}, L_{fp}, and L_{tfp} are local.*

The proof lists the finite tile set enforcing well-formed clockwise contours. Conversely, any picture satisfying these tiles is forced to be a rectangular partition: contour symbols occur only on block boundaries, rectangles cannot overlap, and one rectangle cannot be nested inside another.

The closure operations have an immediate interpretation in this setting:

$$L_{\text{fp}} = L_{\text{basic}}^{++}, \qquad L_{\text{guil}} = L_{\text{basic}}^{\ominus\oplus+}.$$

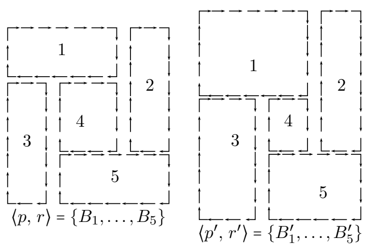
The first identity says that arbitrary rectangular partitions are exactly tilings by basic blocks. The second says that guillotine floor plans are exactly those obtained by recursive horizontal and vertical concatenation of basic blocks. Moreover, $L_{\text{basic}}^{\ominus\oplus+}$ is not local: it has the same set of 2×2 tiles as L_{fp}, but it is a proper subset of L_{fp}.

The recognizability of L_{guil} remains open, but the following conditional result shows this is equivalent to the general case.

Theorem 6. *If $L_{\text{basic}}^{\ominus\oplus+}$ is in* REC, *then $R^{\ominus\oplus+}$ is in* REC *for every recognizable picture language R.*

The proof uses a product tiling system. One component recognizes the guillotine shape, and the other recognizes the content placed in each block. Local constraints synchronize the two components: the contour component identifies the blocks of the guillotine partition, while the content component checks that the interior of each block carries a valid witness for membership in R. Projecting away the auxiliary information gives exactly $R^{\ominus\oplus+}$.

Two positive recognizability results are obtained (without resolving the open problem). First, *strong equivalence classes* of floor plans are recognizable. Strong equivalence preserves the horizontal and vertical contact relations among blocks, as shown in the two strongly-equivalent pictures at the right, where $B_1, \ldots, B_5$ are the 5 blocks in the fpp at the left.



For a fixed floor plan with n blocks, one uses a local language over the product of the contour alphabet with the finite label set $\{1, \ldots, n\}$; the labels enforce the prescribed contacts, and a projection removes them.

Theorem 7. *For every floor-plan picture p, its strong equivalence class $E_p \subseteq L_{\mathrm{fp}}$ is in* REC.

If p is guillotine, then every floor plan strongly equivalent to p is guillotine as well. Hence each strong equivalence class of a guillotine floor plan gives a recognizable guillotine sublanguage.

Second, the complement of the triangulated guillotine language L_{tguil} is recognizable.

Theorem 8. *The language $\Xi^{++} \setminus L_{\mathrm{tguil}}$ is in* REC.

Indeed, within triangulated floor plans, the non-guillotine ones are exactly those containing a windmill. A tiling system can mark and verify one occurrence of such a pattern. The other pictures outside L_{tguil} are either malformed floor-plan encodings or non-triangulated floor plans. These cases are recognizable because the local descriptions above allow a tiling system to guess and verify a violated local condition. Closure of REC under union then gives the result. This does not settle recognizability of L_{guil}, nor of L_{tguil}, because REC is not closed under complement.

5. Conclusion

The results show that two-dimensional closure operations retain part of the algebraic structure of the Kleene star but lose several of its effective and regularity-preserving properties. Tiling-closed and concatenation-closed languages have unique minimal bases, yet recognizable closed languages may have nonrecognizable bases, and natural indecomposability questions are undecidable.

A floor-plan encoding provides a concrete geometric view of these phenomena. Tiling closure corresponds to arbitrary rectangular partitions, while concatenation closure corresponds to guillotine partitions. Thus the open question whether REC is closed under concatenation closure can be reformulated through the recognizability of guillotine floor-plan pictures. The recognizable complement and the recognizable strong equivalence classes give partial positive evidence, but the central closure problem remains open.

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT in order to: Grammar and spelling check, Paraphrase and reword. After using this tool, the author(s) reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] S. Crespi Reghizzi , A. Restivo, P. San Pietro, Closure operations on picture languages and their relation to floor plans, in: M.-P. Béal, P. Caron (Eds.), DLT 2026, Developments in Language Theory, volume 16578 of *Lecture Notes in Computer Science*, Springer, 2026, pp. 168–180. doi:10.1007/978-3-032-28404-4_13.

[2] E. Ackerman, G. Barequet, R. Y. Pinter, A bijection between permutations and floorplans, and its applications, Discrete Appl. Math. 154 (2006) 1674–1684. doi:10.1016/j.dam.2006.03.018.

[3] O. Matz, Regular expressions and context-free grammars for picture languages, in: R. Reischuk, M. Morvan (Eds.), STACS 97, 14th Annual Symposium on Theoretical Aspects of Computer Science, Lübeck, Germany, February 27–March 1, 1997, Proceedings, volume 1200 of *Lecture Notes in Computer Science*, Springer, 1997, pp. 283–294. doi:10.1007/BFB0023466.

[4] A. Cherubini, S. Crespi-Reghizzi, M. Pradella, P. San Pietro, Picture languages: Tiling systems versus tile rewriting grammars, Theor. Comput. Sci. 356 (2006) 90–103. doi:10.1016/j.tcs.2006.01.038.

[5] M. Pradella, A. Cherubini, S. Crespi-Reghizzi, A unifying approach to picture grammars, Inform. and Comput. 209 (2011) 1246–1267. doi:10.1016/J.IC.2011.07.001.

[6] D. Giammarresi, A. Restivo, Recognizable picture languages, Int. J. Pattern Recogn. 6 (1992) 241–256.

[7] J. A. Brzozowski, Roots of star events, J. ACM 14 (1967) 466–477. doi:10.1145/321406.321409.

[8] J. Berstel, D. Perrin, C. Reutenauer, Codes and Automata, Encyclopedia of Mathematics and its Applications, Cambridge University Press, 2009.

[9] D. Giammarresi, A. Restivo, Two-dimensional languages, in: G. Rozenberg, A. Salomaa (Eds.), Handbook of formal languages, vol. 3, Springer, Berlin, 1997, pp. 215–267. doi:10.1007/978-3-642-59126-6_4.

[10] D. Simplot, A characterization of recognizable picture languages by tilings by finite sets, Theor. Comput. Sci. 218 (1999) 297–323. doi:10.1016/S0304-3975(98)00328-4.

[11] A. Asinowski, J. Cardinal, S. Felsner, É. Fusy, Combinatorics of rectangulations: old and new bijections, Comb. Theory 5 (2025). doi:10.5070/C65165025.

[12] Y. Lai, S. M. Leinwand, A theory of rectangular dual graphs, Algorithmica 5 (1990) 467–483. doi:10.1007/BF01840399.

[13] B. D. He, A simple optimal binary representation of mosaic floorplans and baxter permutations, Theor. Comput. Sci. 532 (2014) 40–50. doi:10.1016/J.TCS.2013.05.007.