



**HAL**  
open science

## Self-assembling (classes of) shapes with a constant number of tiles.

Ivan Rapaport, Eric Rémila

► **To cite this version:**

Ivan Rapaport, Eric Rémila. Self-assembling (classes of) shapes with a constant number of tiles.. [Research Report] LIP RR-2004-38, Laboratoire de l'informatique du parallélisme. 2004, 2+12p. hal-02101777

**HAL Id: hal-02101777**

**<https://hal-lara.archives-ouvertes.fr/hal-02101777>**

Submitted on 17 Apr 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



*Laboratoire de l'Informatique du Parallélisme*

École Normale Supérieure de Lyon  
Unité Mixte de Recherche CNRS-INRIA-ENS LYON-UCBL n° 5668

*Self-assembling (classes of) shapes  
with a constant number of tiles*

Ivan Rapaport  
Eric Rémila

September 2004

Research Report N° 2004-38

**École Normale Supérieure de Lyon**

46 Allée d'Italie, 69364 Lyon Cedex 07, France

Téléphone : +33(0)4.72.72.80.37

Télécopieur : +33(0)4.72.72.80.80

Adresse électronique : [lip@ens-lyon.fr](mailto:lip@ens-lyon.fr)



**INRIA**



# Self-assembling (classes of) shapes with a constant number of tiles

Ivan Rapaport  
Eric Rémila

September 2004

## Abstract

Squares are the most studied shapes in the tile assembly model. Adleman and al. proved that the program size complexity of an  $n \times n$  square is  $\Theta(\frac{\log n}{\log \log n})$ . In other words, for each  $n$  we need a different set of tiles. Also, the size of the set increases with  $n$ .

Our approach is to fix *a priori* the set of tiles in such a way that it always self-assembles into an  $N \times N$  square. If  $n$  is an arbitrary positive integer, we show how to calculate the tile concentrations in order to ensure that  $\mathbb{E}(N) = n$ . To our knowledge, the tile concentrations, a parameter of the original model, has never been seriously considered before. We claim that it is therefore not necessary to construct the tiles (which are in fact tiny molecules) for each shape we are asked to assemble. These tiny molecules are, for us, fixed primitives. In order to obtain (in expectation) the required shape, we only need to “play” with the concentrations. This is, of course, a standard procedure in Chemistry and Biology. Let  $d$  be a distance in  $\mathbb{Z}^2$ . In the present work we tackle the specific problem of constructing tile systems that assemble into shapes of the form  $\{x \in \mathbb{Z}^2 \mid d(x, 0) \leq N\}$  and such that  $\mathbb{E}(N) = n$ . We solve the problem for the max distance  $d_\infty$  and for the  $d_1$  distance. In the first case the induced shapes are squares while in the second case the induced shapes are diamonds. Diamonds are much more difficult to produce than squares. We leave as an open problem the Euclidean distance  $d_2$ , which induces the class of “discrete circles”.

**Keywords:** Wang tile, self-assembly

## Résumé

Les carrés sont les formes les plus étudiées dans le modèle d'auto-assemblage. Adleman et al ont prouvé que complexité en taille d'un carré  $n \times n$  est  $\Theta(\frac{\log n}{\log \log n})$ . En d'autres termes, le nombre de tuiles nécessaires pour réaliser le carré  $n \times n$  augmente avec  $n$ .

Notre approche consiste à fixer *a priori* l'ensemble de tuiles de telle façon que les seuls auto-assemblages terminaux possibles soient les carrés. Pour tout entier positif  $n$ , nous montrons comment choisir les concentrations de tuiles, pour garantir que  $\mathbb{E}(N) = n$ . À notre connaissance, les concentrations de tuiles, qui sont un paramètre du modèle original, n'ont jamais été sérieusement considérées auparavant. Les tuiles représentent des petites molécules pouvant s'assembler. Ces molécules de base sont, pour nous, des primitives fixées. Dans le but d'obtenir (en moyenne) la forme de la taille requise, nous avons seulement besoin, de “jouer” sur les concentrations. Ceci est, évidemment, une procédure standard en chimie et en biologie. Soit  $d$  une distance dans  $\mathbb{Z}^2$ . Dans cet article, nous travaillons sur le problème spécifique de construire des systèmes de tuiles qui s'assemblent en formes du type :  $\{x \in \mathbb{Z}^2 \mid d(x, 0) \leq N\}$  et telles que  $\mathbb{E}(N) = n$ .

Nous résolvons le problème pour la distance  $d_\infty$  et pour la distance  $d_1$ . Dans le premier cas, les formes induites sont des carrés tandis que dans le deuxième cas, ce sont des “carreaux”. Les carreaux sont beaucoup plus difficiles à produire que les carrés. Nous laissons ouvert le problème pour la distance Euclidienne  $d_2$ , qui induit la classes des “cercles discrets”.

**Mots-clés:** tuiles de Wang, auto-assemblage

# Self-assembling (classes of) shapes with a constant number of tiles \*

Ivan Rapaport <sup>†</sup>      Éric Rémila <sup>‡</sup>

## Abstract

Squares are the most studied shapes in the tile assembly model. Adleman et al. proved that the program size complexity of an  $n \times n$  square is  $\Theta(\frac{\log n}{\log \log n})$ . In other words, for each  $n$  we need a different set of tiles. Also, the size of the set increases with  $n$ . Our approach is to fix *a priori* the set of tiles in such a way that it always self-assembles into an  $N \times N$  square. If  $n$  is an arbitrary positive integer, we show how to calculate the tile concentrations in order to ensure that  $\mathbb{E}(N) = n$ . To our knowledge, the tile concentrations, a parameter of the original model, has never been seriously considered before. We claim that it is therefore not necessary to construct the tiles (which are in fact tiny molecules) for each shape we are asked to assemble. These tiny molecules are, for us, fixed primitives. In order to obtain (in expectation) the required shape, we only need to “play” with the concentrations. This is, of course, a standard procedure in Chemistry and Biology. Let  $d$  be a distance in  $\mathbb{Z}^2$ . In the present work we tackle the specific problem of constructing tile systems that assemble into shapes of the form  $\{x \in \mathbb{Z}^2 \mid d(x, 0) \leq N\}$  and such that  $\mathbb{E}(N) = n$ . We solve the problem for the max distance  $d_\infty$  and for the  $d_1$  distance. In the first case the induced shapes are squares while in the second case the induced shapes are diamonds. Diamonds are much more difficult to produce than squares. We leave as an open problem the Euclidean distance  $d_2$ , which induces the class of “discrete circles”.

## 1 Introduction

The tile assembly model was introduced by Rothmund and Winfree [4, 6]. This model, based on the classical one of Wang [5], includes a mechanism of growth (a dynamics) which takes into account global parameters such as the temperature and the tile concentrations.

The individual components are square tiles. These tiles “float” on the two dimensional plane. They can not be rotated. Each side of a tile has a specific “glue”. When two tiles collide they

---

\*Partially supported by projects Fondap on Applied Mathematics and Fondecyt 1020611 (I.R.)

<sup>†</sup>Departamento de Ingeniería Matemática and Centro de Modelamiento Matemático, umr 2071 CNRS-Universidad de Chile, Blanco Encalada 2120 Piso 5 Casilla 170-3 Correo 3 Santiago-Chile [irapapor@dim.uchile.cl](mailto:irapapor@dim.uchile.cl)

<sup>‡</sup>Laboratoire de l’Informatique du Parallélisme, umr 5668 CNRS-INRIA-UCB Lyon 1-ENS Lyon, 46 allée d’Italie, 69364 Lyon cedex 07, France and GRIMA, IUT Roanne, 20 avenue de Paris, 42334 Roanne cedex, France, [Eric.Remila@ens-lyon.fr](mailto:Eric.Remila@ens-lyon.fr)

stick if their abutting sides have the same glue and, crucially, if the strength of the glue is “high enough” with respect to the temperature.

The dynamics of such a tile system is modeled as a Markov process. The precise process we consider here was introduced by Adleman et al. [1]. It is, however, a simplification of the reversible version proposed by Winfree [6]. Roughly speaking, the higher the concentration of a particular tile the higher the rate at which it is encountered. And, when encountered, the particular tile can eventually be incorporated into the growing structure. At the end of the process, which begins with a “seed” tile placed at the origin of the plane, a given shape  $S$  will be produced.

Notice that the produced shapes are the sink states of the Markov chain. If there is a unique sink state  $S$ , and if also the state space turns out to be finite, we say that the tile system uniquely produces the shape  $S$ . Researchers, until now, have almost exclusively considered such tile systems. And they have mainly been concerned with the question of finding the minimal number of tiles needed in order to produce particular shapes. A typical result is the one obtained by Aggarwal et al. [2], a generalization of [1], saying that the minimal number of tiles that uniquely produces the  $m \times n$  rectangle is  $\Omega(\frac{n}{m})$  if  $m \ll n$ , and  $\Theta(\frac{\log n}{\log \log n})$  otherwise.

The time  $t$  passed from the beginning of the process until the emergence of  $S$  is a random variable. Even if a tile system  $\mathbf{T}$  uniquely produces  $S$ , the value of  $t$  could change in different “repetitions” of the Markov process. The expected value of  $t$  corresponds to the time it takes to  $\mathbf{T}$  to produce  $S$ . In [1] it is shown that the time complexity for producing an  $n \times n$  square is  $\Theta(n)$ . Of course, some other random variables are also relevant. And, in fact, in this work we focus our attention on the random variable that corresponds to the “size” of the produced shape.

Let us consider, for instance, the class of all squares. In our construction we fix the tile system in such a way that each time we run the Markov chain a (different) square is produced. Let us call  $N$  the random variable corresponding to the size of the produced square. If  $n$  is a fixed positive integer, then we will show how to calculate the tile concentrations in order to ensure that  $\mathbb{E}(N) = n$ . To our knowledge, the tile concentrations, a parameter of the original model, has never been seriously considered before.

We claim that it is therefore not necessary to construct the tiles (which are in fact tiny molecules) for each shape we are asked to assemble. These tiny molecules are, for us, fixed primitives. In order to obtain (in expectation) the required shape, we only need to “play” with the concentrations. This is, of course, a standard procedure in Chemistry and Biology.

Let  $d$  be a distance in  $\mathbb{Z}^2$ . In the present work we tackle the problem of constructing tile systems that assemble into shapes of the form  $\{x \in \mathbb{Z}^2 \mid d(x, 0) \leq N\}$ . More precisely, if an arbitrary positive integer  $n$  is given, our aim is to calculate the tile concentrations in order to get  $\mathbb{E}(N) = n$ . We solve the problem for the max distance  $d_\infty$  and for the  $d_1$  distance. In the first case the induced shapes are squares while in the second case the induced shapes are

diamonds. We leave as an open problem the Euclidean  $d_2$  distance, which induces the class of “discrete circles”.

## 2 Tile systems

A tile system is a 5-tuple  $\mathbf{T} = \langle T, t_0, \tau, g, P \rangle$ . Each of these variables is defined in the following.

**The set of tiles.**  $T$  is a finite set of tiles. Each of these tiles is an oriented unit square with the north, east, south and west edges labeled from some alphabet  $\Sigma$  of glues (or colors). For each  $t \in T$ , the labels of its four edges are canonically denoted  $\sigma_N(t)$ ,  $\sigma_E(t)$ ,  $\sigma_S(t)$  and  $\sigma_W(t)$ .

**The seed.**  $t_0 \in T$  is a particular tile known as the seed.

**The temperature.**  $\tau$  is a positive integer called the temperature.

**The strength function.** The (glue) strength function  $g$  goes from  $\Sigma \times \Sigma$  to  $\mathbb{N}$ . We assume that  $g(\alpha, \beta) = 0$  for all  $\alpha, \beta$  in  $\Sigma$  such that  $\alpha \neq \beta$ . The value  $g(\alpha, \alpha)$  is called the strength of  $\alpha$ . We also assume that the set of glues  $\Sigma$  contains a special one, denoted by *null*, such that for all  $\alpha$  in  $\Sigma$ ,  $g(\text{null}, \alpha) = 0$ . The tiles are represented as in Figure 1: the number of lines in front of the glue corresponds to the strength of it. There is one exception to that convention: no lines mean strength 1.

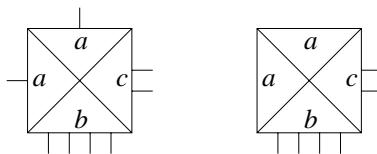


Figure 1: Two ways of representing the same tile.

**The concentration.** The concentration  $P$  associates to each tile  $t \in T$  a positive value  $P(t)$ . The concentration function  $P$  satisfies  $\sum_{t \in T} P(t) = 1$ .

## 3 The dynamics

The dynamics of a tile system is modeled as a Markov process. At the end of the process a particular shape  $S$  is produced. The time  $t$  passed from the beginning of the process until the emergence of  $S$  is an important random variable. More precisely, the expected value of  $t$  is known as the time complexity for producing  $S$ . But some other random variables are also relevant. In fact, in this work we focus our attention on the size of  $S$ .

**T-transitions.** A configuration is a map from  $\mathbb{Z}^2$  to  $(T \cup \{\text{empty}\})$ , where the tile *empty* is the one having in its four sides the glue *null*. Let  $A$  and  $B$  be two configurations. Suppose

that there exist  $t \in T$  and  $(x, y) \in \mathbb{Z}^2$  such that  $A = B$  except for  $(x, y)$  with  $A(x, y) = \text{empty}$  and  $B(x, y) = t$ . If also

$$g(\sigma_E(t), \sigma_W(A(x+1, y))) + g(\sigma_W(t), \sigma_E(A(x-1, y))) + \\ g(\sigma_N(t), \sigma_S(A(x, y+1))) + g(\sigma_S(t), \sigma_N(A(x, y-1))) \geq \tau$$

then we say that the position  $(x, y)$  is attachable in  $A$ , and we write  $A \rightarrow_{\mathbf{T}} B$ . Informally, this means that  $B$  can be obtained from  $A$  by adding a tile  $t$  in such a way that the total strength of the interaction between  $A$  and  $t$  is at least  $\tau$ . Let  $\rightarrow_{\mathbf{T}}^*$  denote the transitive closure of  $\rightarrow_{\mathbf{T}}$ .

**Derived supertiles.** The seed configuration,  $\Gamma_{t_0}$ , is the one that satisfies  $\Gamma_{t_0}(0, 0) = t_0$  and, for all  $(x, y) \neq (0, 0)$ ,  $\Gamma_{t_0}(x, y) = \text{empty}$ . Informally, the seed is reduced to the tile  $t_0$  placed at the origin  $(0, 0)$  of  $\mathbb{Z}^2$ . The derived supertiles of the tile system  $\mathbf{T}$  are those configurations  $X$  such that  $\Gamma_{t_0} \rightarrow_{\mathbf{T}}^* X$ .

**Continuous time Markov process.** The dynamics of the tile system  $\mathbf{T}$  is modeled as a continuous time Markov process where the states are in one-to-one correspondence with the derived supertiles and the initial state corresponds to the seed configuration  $\Gamma_{t_0}$ . There is a transition from state  $A$  to state  $B$  if  $A \rightarrow_{\mathbf{T}} B$ . If  $B$  is obtained from  $A$  by adding the tile  $t$  then the rate of the transition is  $P(t)$ . More precisely, the time for the occurrence of such a transition follows an exponential law of parameter  $P(t)$ . In this paper all we need to know about continuous time Markov chains is the following. Suppose that in state  $A$  there are  $k$  possible transitions to states  $B_1, \dots, B_k$ . And suppose that the transition rates are  $P_1, \dots, P_k$ . Then, the probability to jump to state  $B_i$  equals  $\frac{P_i}{P_1 + \dots + P_k}$ . Finally, the time spent in state  $A$  follows an exponential law of parameter  $P_1 + \dots + P_k$ .

**Production of shapes.** A shape is a 4-connected finite subset of  $\mathbb{Z}^2$ . The shape of a derived supertile  $A$  will be denoted by  $[A]$  and corresponds to  $\{(x, y) \in \mathbb{Z}^2 : A(x, y) \neq \text{empty}\}$ . A derived supertile  $A$  is called terminal if it is a sink state of the Markov process. In other words, if there is no supertile  $B$  such that  $A \rightarrow_{\mathbf{T}} B$ . The set of shapes produced by the tile system is

$$S(\mathbf{T}) = \{[A] : A \text{ is terminal}\}.$$

Let  $\mathcal{C}$  be a set of shapes. We say that the tile system  $\mathbf{T}$  uniquely produces the set  $\mathcal{C}$  if on one hand  $S(\mathbf{T}) = \mathcal{C}$  and, on the other hand, the event “the structure grows indefinitely” has probability zero of occurrence. This notion is the natural generalization of the one of Winfree where the set  $\mathcal{C}$  was a singleton.

## 4 A one dimensional example

Let  $n$  be a positive fixed integer. The one dimensional segment of size  $n$  is the shape

$$L_n = \{(x, 0) \mid 0 \leq x < n\}.$$

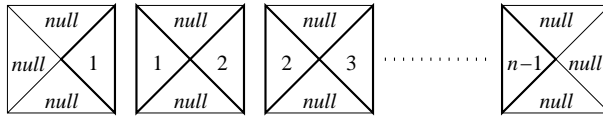


Figure 2: The  $n$  tiles needed in order to uniquely produce an  $n$ -segment.

Let us fix the temperature  $\tau = 1$ . If we want our tile system to uniquely produce  $L_n$  then, with respect to the set of tiles  $T$ , we do not have much choice. In fact, the solution is depicted in Figure 2.

Suppose now that, instead, we have the three tiles of Figure 3. The seed is  $t_0$  and the concentrations are  $P(t_0) = 0$ ,  $P(t_1) = 1 - \epsilon$ ,  $P(t_2) = \epsilon$ . This tile system produces the set of all segments. The main question is the following: what is the expected length of the produced segment? The answer is rather direct. Let  $N$  be the random variable corresponding to the length of the segment. Obviously:

$$\Pr\{N = k\} = (1 - \epsilon)^{k-2}\epsilon.$$

The random variable  $N$  follows a geometric law. So  $\mathbb{E}(N) = \epsilon^{-1} + 1$ . In other words, if we repeat this experiment many times with the concentration of tile  $t_2$  being  $(n - 1)^{-1}$ , then the average size of the produced segments will be exactly  $n$ . And we do this with a constant number of tiles. Moreover: with a fixed set of tiles. We just need to prepare the right concentrations of each tile. This is an elementary procedure in Biology and Chemistry.

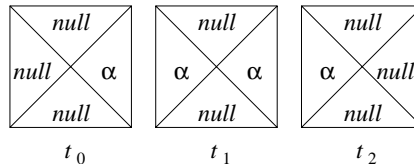


Figure 3: Three tiles producing a segment of expected length  $n$ .

## 5 Rectangles and squares

Let  $m, n$  be positive integers. The rectangle of width  $m$  and height  $n$  is the shape

$$R_{m,n} = \{(x, y) \mid 0 \leq x < m, 0 \leq y < n\}.$$

Let us fix the temperature  $\tau = 2$ . With this temperature, by generalizing the result of [1], it has been proved in [2] that the minimal number of tiles that uniquely produce the rectangle  $R_{m,n}$  is

$$\Omega\left(\frac{n}{m}\right) \quad \text{if } m \ll n,$$

$$\Theta\left(\frac{\log n}{\log \log n}\right) \quad \text{otherwise.}$$



Following our approach we fix the set of tiles that appears in Figure 4. Let us consider  $t_{SW}$  as the seed. It is easy to notice that this set of tiles uniquely produces rectangles. If  $A, B, C$  are arbitrary positive values satisfying  $A + B + C = 1$ , then we fix the concentrations as follows:

$$\begin{aligned} P(t_S) &= A(1 - (m - 1)^{-1}), & P(t_{SE}) &= A(m - 1)^{-1}, \\ P(t_W) &= B(1 - (n - 1)^{-1}), & P(t_{NW}) &= B(n - 1)^{-1}, \\ P(t_\beta) &= C, & P(t_{SW}) &= 0. \end{aligned}$$

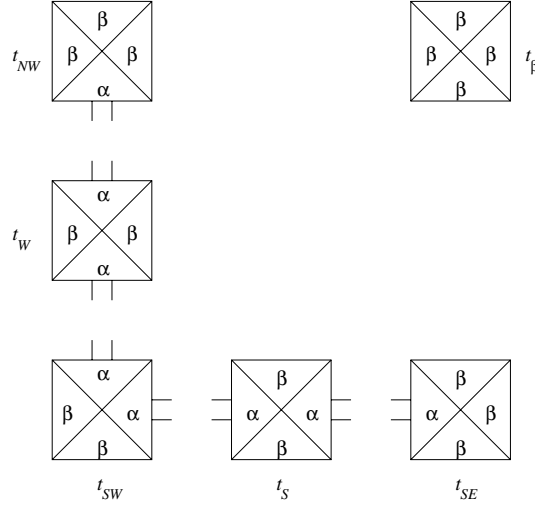


Figure 4: The set of tiles used for producing rectangles.

The same reasoning we applied for the one dimensional segment leads us to the following result.

**Proposition 1** *The tile system defined above uniquely produces rectangles. If  $M$  is the random variable corresponding to the width and  $N$  is the random variable corresponding to the height, then  $\mathbb{E}(M) = m$  and  $\mathbb{E}(N) = n$ .*

The height and the width of the rectangles produced by the previous tile system were independent. Is it possible to make these two values absolutely dependant in order to uniquely produce, for instance, squares? Squares are the most studied objects in the tile assembly model.

If we want to produce squares it is rather natural to create diagonals. Informally, we use four tiles in order to construct the diagonal. These four tiles  $-t_D, t_{D_{right}}, t_{D_{up}}$  and  $t_{stop}$  appear in Figure 5. We need two more tiles in order to fill the square:  $t_\beta$  for the northwest half and  $t_\gamma$  for the southeast half. The seed is  $t_D$  and the temperature is  $\tau = 2$ .

For arbitrary positive values  $A, B, C, D$  such that  $A + B + C + D + E = 1$ , we fix the concentrations as follows:

$$P(t_D) = A(1 - (n - 1)^{-1}), \quad P(t_{stop}) = A(n - 1)^{-1},$$

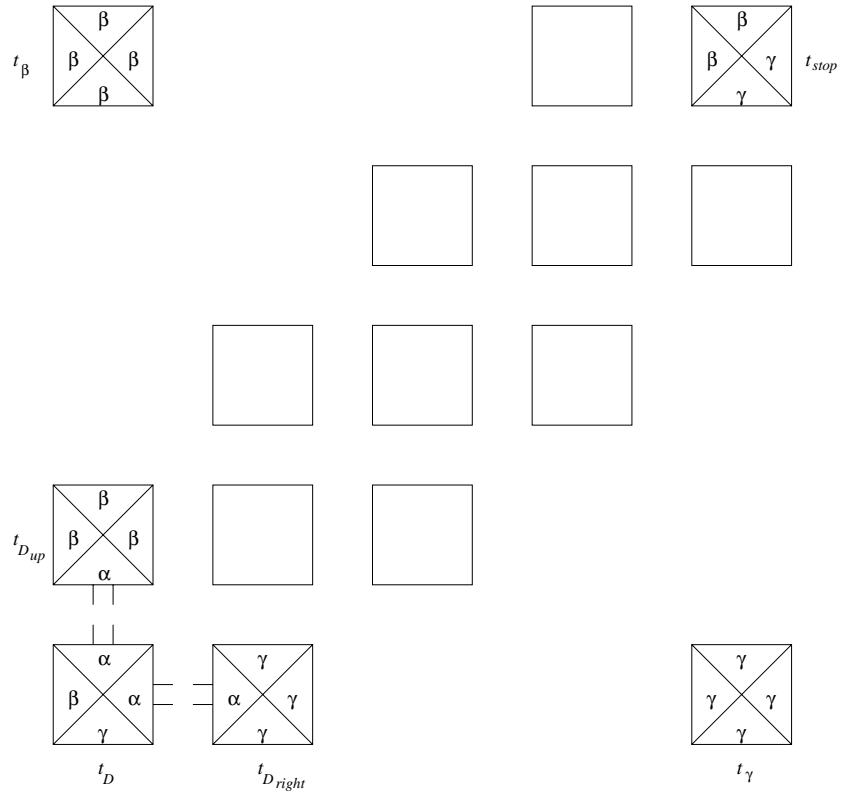


Figure 5: The set of tiles used for producing squares.

$$P(t_{D_{right}}) = B, P(t_{D_{up}}) = C, P(t_{\beta}) = D, P(t_{\gamma}) = E.$$

**Proposition 2** *The tile system defined above uniquely produces squares. If  $N$  is the random variable corresponding to the length of the sides, then  $\mathbb{E}(N) = n$ .*

## 6 Diamonds

Notice that a square corresponds to the set  $\{(x, y) \mid d_{\infty}[(x, y), (0, 0)] \leq n\}$ , with

$$d_{\infty}[(x, y), (x', y')] = \max\{|x - x'|, |y - y'|\}.$$

If we change the metrics towards the more “natural”

$$d_1[(x, y), (x', y')] = |x - x'| + |y - y'|,$$

then the induced shape is the diamond  $D_n$  that appears in Figure 6. The problem of producing diamonds is much more complicated than those we tackled before. Any naive approach seem not to work. We are going to construct diamonds with the help of a very particular and non-trivial cellular automaton called “the firing squad”.

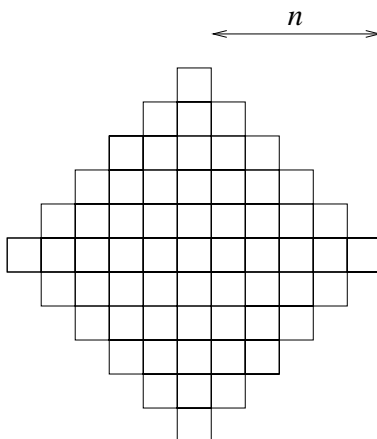


Figure 6: An  $n$ -diamond.

### 6.1 The firing squad cellular automaton

A one dimensional cellular automaton (CA) is a couple  $(Q, \delta)$  where  $Q$  is a finite set of states and  $\delta : Q^3 \rightarrow Q$  is a transition function. We always assume the existence of a state  $\rho \in Q$  such that  $\delta(x, \rho, y) = \rho$  for all  $x, y \in Q$ . A configuration of a CA  $(Q, \delta)$  is a bi-infinite sequence  $\mathcal{C} \in Q^{\mathbb{Z}}$ , and its global transition function  $G_{\delta} : Q^{\mathbb{Z}} \rightarrow Q^{\mathbb{Z}}$  is such that

$$(G_{\delta}(\mathcal{C}))_i = \delta(\mathcal{C}_{i-1}, \mathcal{C}_i, \mathcal{C}_{i+1}).$$

There is a particular CA, defined in [3] and related to the classical “firing-squad problem”, capable of synchronizing an array of cells of arbitrary size “as soon as possible”. The main properties of this CA, that we denote by  $(Q_{FS}, \delta_{FS})$ , are the following.

- $\{G_l, G_r, s, F\} \subseteq Q_{FS}$  (left general, right general, soldier, fire).
- For every  $n \in \mathbb{N}$ , when beginning from the initial configuration

$$\dots \rho\rho\rho \underbrace{G_l s s s \dots s s s G_r}_{n} \rho\rho\rho \dots$$

the CA evolves, after  $n - 1$  steps, to the configuration

$$\dots \rho\rho\rho \underbrace{F F F \dots F F F}_{n} \rho\rho\rho \dots$$

in such a way tht the state  $F$  never appears before the  $(n - 1)$ -th step.

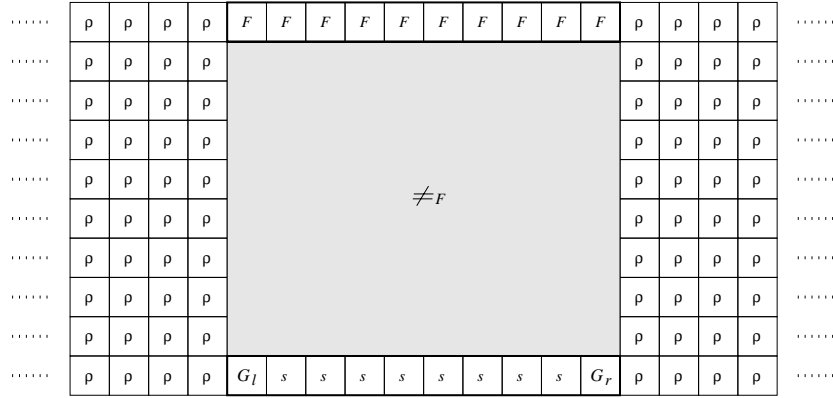


Figure 7: Two-generals firing squad.

## 6.2 Simulating the firing squad CA by self-assembly

Let us fix the temperature  $\tau = 2$ . We will represent, for simplicity, the tiles rotated in  $45^\circ$ . Let us first consider the set of six tiles of Figure 8. The tiles are, from left to right,  $t_\alpha$  (the seed),  $t_{G_l}$ ,  $t_\beta$ ,  $t_s$ ,  $t_{G_r}$  and  $t_\gamma$ . The colors  $\alpha, \beta, \gamma \notin Q_{FS}$ . The color *null* is omitted. For instance,

$$\sigma_N(t_\alpha) = \sigma_S(t_\alpha) = \sigma_W(t_\alpha) = \text{null}.$$

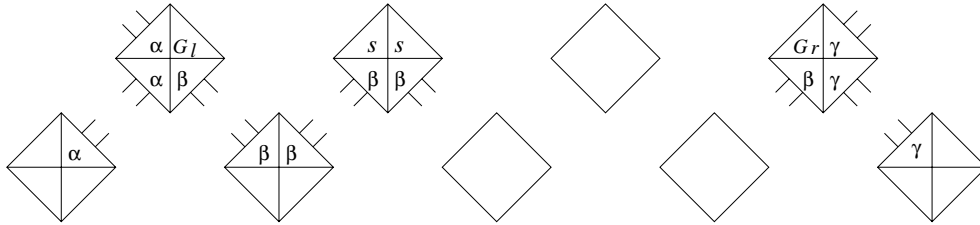


Figure 8: The set of six tiles that codifies the initial configuration.

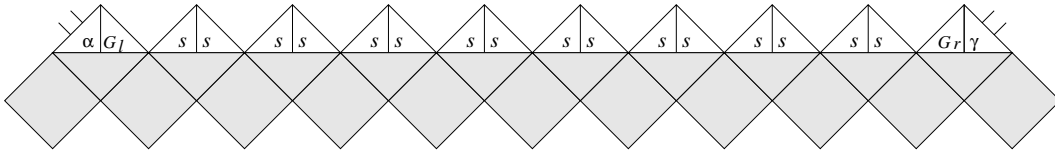


Figure 9: The way the initial configuration  $\dots \rho\rho\rho G_l s s s \dots s s s G_r \rho\rho\rho \dots$  is assembled.

As it is schematically explained in Figure 9, the assembling process of these tiles is such that the structure they produce represents the initial configuration

$$\dots \rho\rho\rho G_l s s s \dots s s s G_r \rho\rho\rho \dots$$

The size of the initial structure of Figure 9 determines the size of the diamond we are going to produce. This part of the self-assembly process is, in fact, the only nondeterministic one. Therefore, the expected size of the diamond can be calculated as a function of the concentrations of the previously introduced tiles. Moreover, the only concentrations relevant for the process are those of  $t_s$  and  $t_{G_r}$ . Let  $0 < A < 1$ . Let us define the concentrations as follows:

$$P(t_s) = A(1 - (n - 1)^{-1}), P(t_{G_r}) = A(n - 1)^{-1}.$$

The only requirement for the concentrations of the other tiles (the four already introduced and those to come) is that they must be positives with their sum being  $1 - A$ .

There are two other classes of tiles: transmission tiles and transition tiles. The transmission tiles are divided into six subclasses: left-border, internal, right-border, upper-left-border, upper-border, upper-right-border. More precisely, for all  $a, b \in Q_{FS} \setminus \{F\}$ , the transmission tiles are constructed in Figure 10.

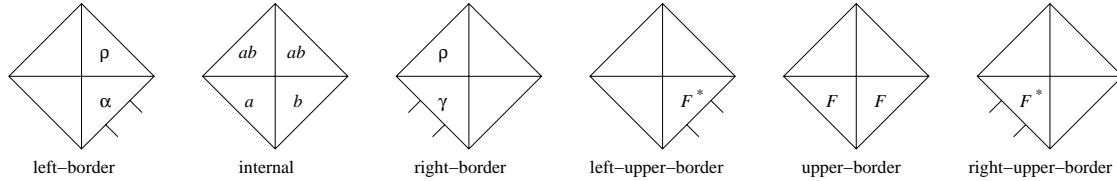


Figure 10: Transmission tiles.

The transition tiles are divided into five subclasses: left-border, internal, right-border, upper-left-border and upper-right-border. More precisely: let  $a, b, c, d, e, f, g \in Q_{FS}$  be such that

$$\delta_{FS}(\rho, a, b) \neq F \text{ and } \delta_{FS}(f, g, \rho) \neq F.$$

The  $(a, b)$ -left-border,  $(c, d, e)$ -internal and  $(f, g)$ -right-border tiles are constructed in Figure 11.

Finally, let  $a, b, c, d \in Q_{FS}$  be such that  $\delta_{FS}(\rho, a, b) = \delta_{FS}(c, d, \rho) = F$ . The upper-left-border and upper-right-border tiles are constructed in Figure 12.

From the previously defined construction follows the last proposition.

**Proposition 3** *The tile system defined above uniquely produces diamonds. If  $N$  is the random variable corresponding to the length of the diagonal, then  $\mathbb{E}(N) = 2n + 1$ .*

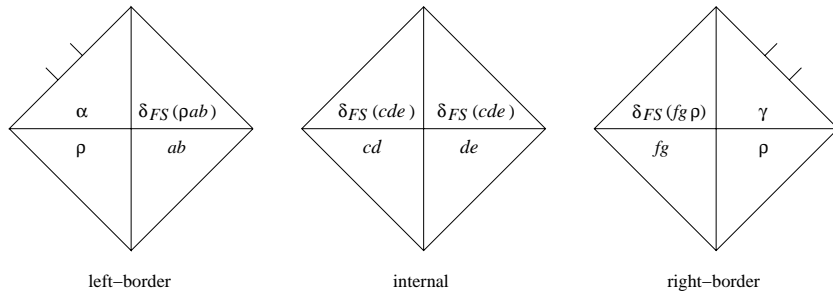


Figure 11: Transition tiles: left-border, internal, right-border.

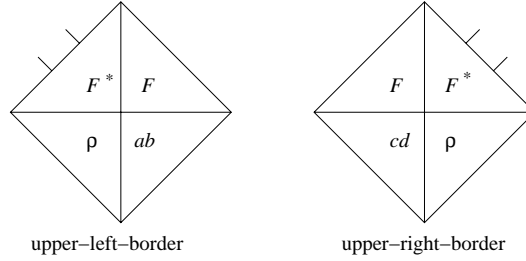


Figure 12: Transition tiles: upper-left-border, upper-right-border.

## 7 Perspectives

We have considered the  $d_\infty$  and the  $d_1$  distances. It is a natural step forward to consider the Euclidean distance  $d_2$ . In other words, we should try to answer whether it is possible or not to produce the class of “discrete circles”:

$$\{(x, y) \mid \sqrt{x^2 + y^2} \leq n\}.$$

We should take into account the multiple temperature model introduced in [2]. In that model, roughly speaking, the temperature  $\tau$  is adjusted during assembly. We can exhibit a class of shapes that, in our standard model, can not be produced with a constant number of tiles: the  $L$ -shape of Figure 13(a). Nevertheless, as it is shown in Figure 13(b), we can construct a square with  $\tau = 2$  and then increase the temperature to  $\tau = 3$ . The only portion of the square that will “resist” this higher temperature is the desired  $L$ -shape.

Given a class of shapes  $\mathcal{C}$  and a positive value  $n$ , our aim was to find a tile system  $\mathbf{T}$  producing  $\mathcal{C}$  and such that  $\mathbb{E}(N) = n$  (where  $N$  is the random variable corresponding to the size of the produced shape). An extremely natural complexity problem arises: we should find, among all those feasible tile systems, the one that minimizes the variance of  $N$ .

## References

- [1] L. Adleman, Q. Cheng, A. Goel and M. Huang. “Running time and program size for self-assembled squares”. In Proceedings of the 33rd Annual ACM Symposium on Theory of Computing (STOC 2001), 740-748, 2001.

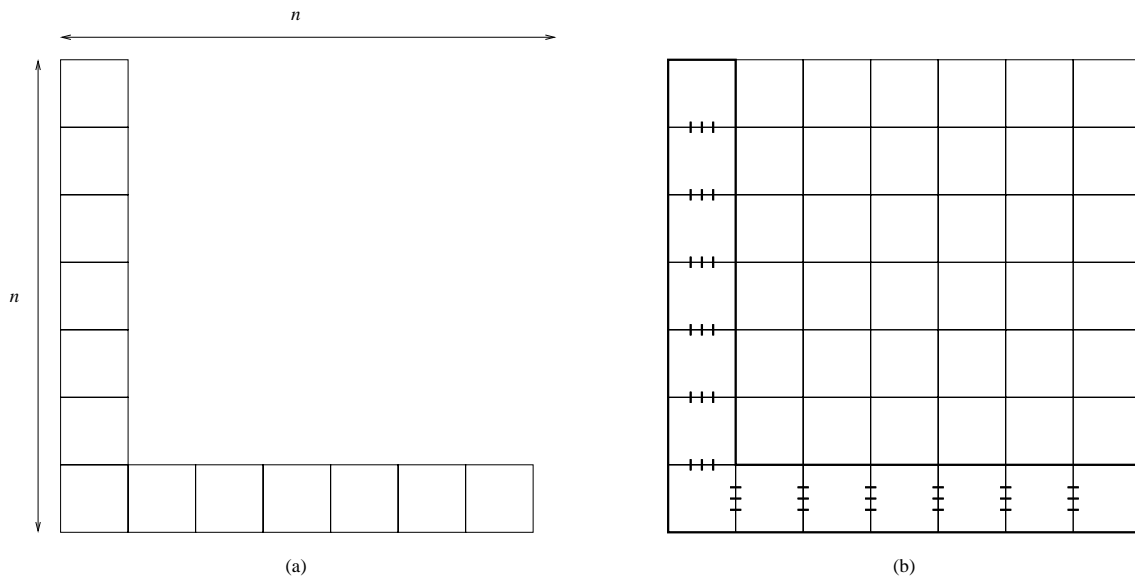


Figure 13: (a) An  $L$ -shape. (b) A square having an  $L$ -shape as a more stable substructure.

- [2] G. Aggarwal, M. H. Goldwasser, M-Y. Kao and R. T. Schweller. “Complexities for generalized models of self-assembly”. Proceedings of the 15th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2004), 880–889, 2004.
- [3] J. Mazoyer and N. Reimen. “A linear speed-up theorem for cellular automata”. Theoretical Computer Science 101 (1992), 59-98.
- [4] P. Rothmund and E. Winfree. “The program size complexity of self-assembled squares”. In Proceedings of the 32nd Annual ACM Symposium on Theory of Computing (STOC 2000), 459-468, 2000.
- [5] H. Wang. “Proving theorems by pattern recognition”. H. Bell System Technical Journal, 40:1-41, 1961.
- [6] E. Winfree. “Algorithmic self-assembly of DNA”. Ph.D. thesis, California Institute of Technology, Pasadena, 1998.