

On the decidability of self-assembly of infinite ribbons

Leonard Adleman

Laboratory for Molecular Science
University of Southern California
adleman@usc.edu

Lila Kari

Department of Computer Science
University of Western Ontario
lila@csd.uwo.ca

Jarkko Kari

Department of Computer Science
University of Iowa
jjkari@cs.uiowa.edu

Dustin Reishus

Laboratory for Molecular Science
University of Southern California
reishus@usc.edu

Abstract

Self-assembly, the process by which objects autonomously come together to form complex structures, is omnipresent in the physical world. A systematic study of self-assembly as a mathematical process has been initiated. The individual components are modelled as square tiles on the infinite two-dimensional plane. Each side of a tile is covered by a specific “glue”, and two adjacent tiles will stick iff they have matching glues on their abutting edges. Tiles that stick to each other may form various two-dimensional “structures” such as squares, rectangles, or may cover the entire plane. In this paper we focus on a special type of structure, called ribbon: A non-self-crossing sequence of tiles on the plane, in which successive tiles are adjacent along an edge, and abutting edges of consecutive tiles have matching glues. We prove that it is undecidable whether an arbitrary finite set of tiles with glues (infinite supply of each tile type available) can be used to assemble an infinite ribbon. The proof is based on a construction, due to Robinson, of a special set of tiles that allow only aperiodic tilings of the plane. This construction is used to create a special set of directed tiles (tiles with arrows painted on the top) with the “strong plane-filling property” - a variation of the “plane-filling property” previously defined by J. Kari. A construction of “sandwich” tiles is then used in conjunction with this special tile set, to reduce the well-known undecidable Tiling Problem to the problem of the existence of an infinite directed zipper (a special kind of ribbon). A “motif” construction is then introduced that allows one tile system to simulate another by using geometry to represent glues. Using motifs, the infinite directed zipper problem is reduced to the infinite ribbon problem, proving the latter undecidable.

The result settles an open problem formerly known as the

“unlimited infinite snake problem”. Moreover, an immediate consequence is the undecidability of the existence of arbitrarily large structures self-assembled using tiles from a given tile set.

1 Introduction

Self-assembly, the process by which objects autonomously come together to form complex structures, is omnipresent in the physical world. Atoms bind to each other by chemical bonds to form molecules, molecules may form crystals or macromolecules, cells interact to form biological organisms. Recently it has been suggested that complex self-assembly processes will ultimately be used in circuit fabrication, nano-robotics, DNA computation and amorphous computing. Indeed, in electronics, engineering, medicine, material science, manufacturing and other disciplines, there is a continuous drive toward miniaturization. Unfortunately, current “top-down” techniques such as lithography may not be capable of efficiently creating structures with features the size of molecules or atoms. Self-assembly provides a “bottom-up” approach by which such fine structures might be created.

Experimental research on self-assembly includes simulation of one-dimensional cellular automata by using macroscopic plastic tiles that assemble at an oil/water interface [25], studies of self-directed growth of hybrid organic molecules on a silicon substrate [18], self-assembly of regular lipid hollow icosahedra [8], self-assembly of patterns of lead on a copper surface as templates for fabricating nanostructures [22], formation of electrical networks by self-assembly of small plastic and metal objects [10], [33], self-assembly of molecular machines [9], and the construction by self-assembly of the world’s first molecular-thickness tran-

sistor [28].

Investigations into DNA Computing [2] and amorphous computing [1] have pointed out a strong connection between self-assembly and computation. While [30], [21], [23], [32] have shown the potential of DNA self-assembly for computation, [31] has experimentally demonstrated the self-assembly of periodic two-dimensional arrays from DNA “tiles”, in which “binding” is regulated by the “sticky” ends of the DNA tiles. Moreover [19] has demonstrated the execution of logical operations (cumulative XOR) by self-assembly of DNA triple cross-over molecules.

A systematic study of self-assembly as a computational process has been initiated in [3]. The individual components are therein modelled as square tiles on the infinite two-dimensional plane. The tiles cannot be rotated. Each side of a tile is covered by a specific “glue” and two adjacent tiles will stick iff they have matching glues on their abutting edges. As such, tiles have been previously studied in the context of classic questions about tilings of the plane [29]. The Tiling Problem, for example (proved undecidable in [7], [24]), asks whether a given set of tiles (supply of each tile type unlimited) can be used to correctly tile the entire plane. However, self-assembly of tiles into required shapes poses new types of questions. What is the minimal number of tiles required for the self-assembly of a desired shape and no other objects [5]? What is the running time of such a self-assembly [5], [6]? Do the results still hold if one generalizes the model by making the “sticking” reversible, or by defining different bond strengths and requiring more than one bond for sticking to occur [27], [6], [4], [5]? Do reversible self-assemblies achieve equilibrium [4]? What is the optimal initial concentration of tile types that guarantees the fastest assembly of a required shape [5]?

One of the interesting problems of self-assembly is whether or not a given set of tiles allows uncontrollable growth, that is, whether arbitrarily large structures can be produced. It turns out that this question is equivalent to asking whether or not, given a set of tiles, there exists an infinite ribbon that can be formed with tiles from this set. A *ribbon* is a non-self-crossing sequence of tiles on the plane, in which successive tiles are adjacent along an edge, and abutting edges of consecutive tiles have matching glues. If a given “seed” tile is specified, the problem of existence of an infinite ribbon starting at that tile has been proved undecidable [12], [11], [13]. However, when no such seed is specified, existing proof techniques failed to produce an answer and the problem was declared open in [13]. The unseeded problem is even more relevant given that, in physical simulations of self-assembly, the growth mechanism was deemed incompatible with computations that use a chosen input [25]. Here we prove that the unseeded problem is also undecidable. This result settles the decade-old problem known hitherto as the “unlimited infinite snake problem”,

[13].

The paper is organized as follows. Section 2 introduces the notions of tile, glue, tile system, sticking, ribbon, zipper, valid tiling of the plane, directed tiles. In particular, a zipper is a ribbon with the additional requirement that even non-consecutive tiles that touch must have matching glues at their abutting edges.

Section 3 starts with defining a directed tile system with the strong plane-filling property: in such a system, any infinite directed zipper is forced to follow a specific plane-filling self-similar path and, moreover, an infinite directed zipper is always guaranteed to exist. The construction of a directed tile system with the above property uses a 3×3 block-construction starting from directed tiles defined in [16], [15] which, in turn, resemble tiles devised by Robinson in [24] to only produce aperiodic tilings of the plane. These tiles are augmented in [15], [16] with directions that force any directed tiled path to form a self-similar plane-filling Hilbert curve that recursively fills arbitrarily large squares by filling each of their quadrants. The original construction used the condition that no glue mismatch be present in the 3×3 neighbourhood of any tile on the path, to force the path to follow the desired curve. Here, the additional 3×3 block construction is needed to ensure that the weaker condition of no glue mismatch between any two tiles on the directed tiled path is enough to force its course.

Section 3 then proves the undecidability of the existence of an infinite directed zipper by reducing the Tiling Problem to it. The reduction is based on a construction of sandwich tiles, that have directed tiles from the directed tile system with the strong plane-filling property on top and undirected tiles from a given tile system on the bottom.

The next result of the section proves the undecidability of the existence of an infinite ribbon by reducing the undecidable infinite directed-zipper problem to it. The proof uses a “ribbon-motif” construction that simulates each directed zipper tile by a ribbon of undirected tiles following its contours, and uses geometry (bump and dents) to simulate zipper-tile glues.

Section 4 and Section 5 point to the implications of the undecidability of existence of infinite ribbons for the problem of self-assembly of arbitrarily large supertiles, and summarize our results.

2 Notation and definitions

A tile is an oriented unit square. The north, east, south and west edges of the tile are each labelled with a symbol called a *glue* from a finite alphabet X . Tiles can be placed on the plane but not rotated. The positions of the tiles on the plane are indexed by \mathbb{Z}^2 , the set of pairs of integers.

Formally, a *tile* t is a quadruple $t = (t_N, t_E, t_S, t_W) \in X^4$ where X is a finite set. The components t_N, t_E, t_S, t_W

will be called the glues on the north, east, south, and west edges of the tile, respectively. A *tile system* T is a finite subset of X^4 . For all tiles $t = (t_N, t_E, t_S, t_W)$ and $t' = (t'_N, t'_E, t'_S, t'_W)$, t sticks on the north to t' iff $t_N = t'_S$. Sticking on the east, south, and west is defined similarly.

The directions $\mathcal{D} = \{N, E, S, W\}$ are functions from \mathbb{Z}^2 to \mathbb{Z}^2 : $N(x, y) = (x, y + 1)$, $E(x, y) = (x + 1, y)$, $S(x, y) = (x, y - 1)$, $W(x, y) = (x - 1, y)$. We say that the positions (x, y) and (x', y') are *adjacent* iff $(x', y') \in \{N(x, y), E(x, y), S(x, y), W(x, y)\}$. In addition, (x, y) *abuts* (x', y') on the north iff $(x', y') = N(x, y)$, and similarly for the other directions.

Given a tile system T , a *T -tiling of the plane* is a total function $f : \mathbb{Z}^2 \rightarrow T$. The tiling f is *valid at position* (x, y) iff $f(x, y)$ sticks on the north to $f(N(x, y))$, $f(x, y)$ sticks on the east to $f(E(x, y))$, $f(x, y)$ sticks on the south to $f(S(x, y))$, and $f(x, y)$ sticks on the west to $f(W(x, y))$. That is to say, f is valid at (x, y) iff the tile at position (x, y) sticks on the appropriate sides to all the tiles that are at positions adjacent to it. The tiling is *valid* iff it is valid at every position $(x, y) \in \mathbb{Z}^2$. The well-known *Tiling Problem* posed by Wang asks: Given a tile system T , does there exist a valid T -tiling of the plane? The Tiling Problem has been first proven undecidable by Berger [7], and the result improved by Robinson [24]. One can generalize the notion of T -tiling of the entire plane by defining a *partial T -tiling* as a partial function $g : \mathbb{Z}^2 \rightarrow T$. A partial tiling g is *valid on a region* $R \subseteq \text{dom}(g)$ iff, for all $(x_1, y_1), (x_2, y_2) \in R$, if (x_1, y_1) abuts (x_2, y_2) on the north (east, south, west), then $g(x_1, y_1)$ sticks to $g(x_2, y_2)$ on the north (east, south, west). The partial tiling g is called *valid* iff it is valid on the entire $\text{dom}(g)$.

A *path* is a function $P : I \rightarrow \mathbb{Z}^2$ where I is a set of consecutive integers and $\forall(i, i + 1 \in I)$, $P(i)$ and $P(i + 1)$ are adjacent. That is, a path is a sequence of adjacent positions on the plane. For all $i \in I$, we denote $P(i)$ as (x_i, y_i) . We say that (x, y) is on P iff $(x, y) \in \text{range}(P)$. For all tile systems T , a *T -tiled path* is a pair (P, r) where P is a path and r is a function $r : \text{range}(P) \rightarrow T$.

A T -tiled path (P, r) is a *T -ribbon* iff

(a) P is one-to-one and

(b) $\forall(i, i + 1) \in \text{dom}(P)$, if (x_{i+1}, y_{i+1}) abuts (x_i, y_i) on the north (south, east, west) then $r(x_{i+1}, y_{i+1})$ sticks to $r(x_i, y_i)$ on the north (south, east, west).

Informally, a tiled path is a ribbon iff it does not cross itself and the glue between tiles at consecutive positions along the path match.

A T -ribbon (P, r) is a *T -zipper* iff for all $(x, y), (x', y')$ on P if (x', y') abuts (x, y) on the north (south, east, west) then $r(x', y')$ sticks to $r(x, y)$ on the north (south, east, west). Note that the notion of a zipper is more restrictive than the notion of a ribbon in that a zipper requires all of its tiles (consecutive or not) in adjacent positions to stick.

A *directed tile system* is a pair (T, d) where T is a tile system and $d : T \rightarrow \{N, E, S, W\}$ is a function from the tile system to the direction functions. A directed tile system can be thought of as a tile system in which each tile has an arrow painted on it pointing north, east, south, or west.

A (T, d) -*directed tiled path* is a pair (P, r) , where (P, r) is a T -tiled path and $\forall(i, i + 1 \in \text{dom}(P))[(x_{i+1}, y_{i+1}) = d(r(x_i, y_i))(x_i, y_i)]$. A directed tiled path is a tiled path in which each tile points to its successor on the path. If (P, r) is a (T, d) -directed tiled path and (P, r) is a T -ribbon (zipper) then (P, r) is a (T, d) -directed ribbon (zipper).

A path, (directed) tiled path, (directed) ribbon, (directed) zipper is finite, semi-infinite, infinite, iff the domain of the associated path is finite, has a least or greatest element and is infinite, or is \mathbb{Z} respectively.

3 Undecidability of existence of infinite ribbons

To prove the undecidability of existence of infinite ribbons, we shall first prove the undecidability of existence of infinite directed zippers. Afterwards, a “ribbon-motif” construction will be employed to prove the result for ribbons: the construction simulates a directed zipper by a ribbon-motif of smaller tiles that goes around the contours of the zipper tiles, and uses geometry to simulate zipper-tile glues.

In order to prove the undecidability of existence of infinite directed zippers, we shall use the undecidability of the Tiling Problem in conjunction with a sandwich construction that makes use of the existence of a directed tile system with the so-called *strong plane-filling property*.

Definition 3.1 A directed tile system (T, d) has the strong plane-filling property iff:

- (a) There exists an infinite (T, d) -directed zipper and
- (b) For all infinite (T, d) -directed zippers (P, r) , $\forall n \exists(x, y)$ such that $(x + i, y + j)$ is on P for $i = 0, 1, 2, \dots, n$ and $j = 0, 1, 2, \dots, n$.

Definition 3.1 states that in a directed tile system with the strong plane-filling property any infinite directed zipper is forced to be plane-filling (by filling arbitrarily large squares) and, moreover, that such an infinite directed zipper always exists. An important element of our subsequent proofs will be Theorem 3.1, namely the construction of a directed tile system that satisfies the strong plane-filling property of Definition 3.1. In fact, our directed tiles satisfy an even stronger requirement than (b): any infinite directed tiled path is either plane-filling or it has a glue mismatch between two of its tiles. In particular, using our tiles, an infinite directed tiled path may not even form a loop without encountering a glue mismatch along the way. Consequently, every non-plane-filling infinite directed tiled path

must necessarily contain two neighbouring tiles with a glue mismatch. The constructed tiles satisfy therefore a stronger version of Definition 3.1, satisfying thus a stronger property than the original plane-filling property defined in [15, 16].

In the original version, any infinite directed tiled path was forced to be plane-filling unless there was a mismatch of glues inside the 3×3 neighbourhood of some tile of the path. In [16] directed tiles with this weaker plane-filling property were explicitly constructed by augmenting Robinson's aperiodic tiles [24] with directions. With these tiles, all directed tiled paths consisting of tiles without errors in their 3×3 neighbourhood formed self-similar plane-filling curves of the kind used by Hilbert [14] to show that the unit square is a continuous image of the unit segment. The constructed tiles filled arbitrarily large squares recursively, by first filling the individual quadrants of a square. The process was guided by the direction associated to each tile, which forced the construction to proceed along the self-similar Hilbert curve that traversed each quadrant of a square and linked the quadrants to each other.

In the present application we need the stronger variant of the plane-filling property introduced in this paper. In our case the infinite directed tiled path must be forced to be plane-filling even under the weak assumption that there is no glue mismatch between any two tiles belonging to it. It turns out that such a directed tile system is obtained by taking 3×3 blocks of the tiles in [16]. For this particular set of tiles, our 3×3 block-construction ensures that the weak requirement of absence of glue mismatches between any two tiles belonging to the directed tiled path is sufficient to determine its uniqueness. Namely, we can prove that the constructed directed tile system has the property that, starting with any arbitrary tile, the only way to build an infinite error-free directed zipper is by forming a Hilbert curve that covers arbitrarily large squares.

The exact construction is rather complex and is not included here. We have:

Theorem 3.1 *There exists a directed tile system (T_0, d_0) with the strong plane-filling property.*

We shall now use Theorem 3.1 to prove the undecidability of the existence of an infinite directed zipper.

Theorem 3.2 *The set $\{(T, d) \mid (T, d) \text{ is a directed tile system and there exists an infinite } (T, d)\text{-directed zipper}\}$ is undecidable.*

Proof. Reduce the undecidable Tiling Problem to our problem.

By Theorem 3.1 there exists a directed tile system (T_0, d_0) with the strong plane-filling property.

Let T_1 be a tile system. Consider the following “sandwich tiles”, each consisting of a tile from T_0 placed on top

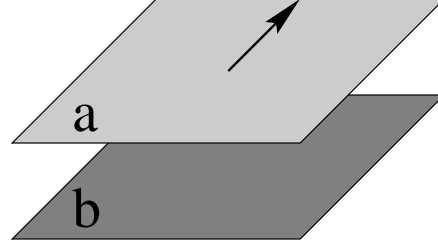


Figure 1. A sandwich tile $\sigma(a, b)$ consists of a directed tile $a \in T_0$ placed on top of a tile $b \in T_1$. The direction of $\sigma(a, b)$ is $d(\sigma(a, b)) = d_0(a)$, the direction of its top tile.

of a tile from T_1 . That is, create a new directed tile system (T, d) of sandwich tiles such that $T = \{\sigma(a, b) \mid a \in T_0 \text{ and } b \in T_1\}$ where for all $a = (a_N, a_E, a_S, a_W) \in T_0$ and $b = (b_N, b_E, b_S, b_W) \in T_1$, the sandwich tile $\sigma(a, b) = ((a_N, b_N), (a_E, b_E), (a_S, b_S), (a_W, b_W))$ and the direction of the sandwich tile is $d(\sigma(a, b)) = d_0(a)$ (see Figure 1).

Hence, a sandwich tile $\sigma(a, b)$ sticks on the north (east, south, west) to $\sigma(a', b')$ iff a sticks on the north (east, south, west) to a' and b sticks on the north (east, south, west) to b' .

We will show now that there exists a valid T_1 -tiling of the plane iff there exists an infinite (T, d) -directed zipper.

“ \Rightarrow ” Assume that there exists a valid T_1 -tiling of the plane $f : \mathbb{Z}^2 \rightarrow T_1$. Consider an infinite (T_0, d_0) -directed zipper (P, r) (such exists, since (T_0, d_0) has the strong plane-filling property). Consider the T -tiled path (P, q) of sandwich tiles such that, for all (x, y) on P , $q(x, y) = \sigma(r(x, y), f(x, y))$. Informally, the directed tiled path of sandwich tiles consists of the infinite (T_0, d_0) -directed zipper on its top, and the corresponding partial valid T_1 -tiling $f|_{\text{range}(P)}$ on the bottom. It is clear that in fact (P, q) is an infinite (T, d) -directed zipper of sandwich tiles.

“ \Leftarrow ” Assume that there exists an infinite (T, d) -directed zipper (P, q) of sandwich tiles. Then, let (P, r) be its top layer, i.e., for all (x, y) on P , $r(x, y) = a$ iff $q(x, y) = \sigma(a, b)$ for some $b \in T_1$. Then clearly, (P, r) is an infinite (T_0, d_0) -directed zipper. Hence, as (T_0, d_0) has the strong plane-filling property, P contains “arbitrarily large squares”: $\forall n \exists (x, y)$ such that $(x + i, y + j)$ is on P for $i = 0, 1, 2, \dots, n$ and $j = 0, 1, 2, \dots, n$. Let (P, z) be the bottom layer of (P, q) , i.e. the T_1 -tiled path such that for all (x, y) on P , $z(x, y) = b$ iff $q(x, y) = \sigma(a, b)$. Then clearly, (P, z) is in fact an infinite T_1 -zipper. It follows that $\text{range}(P) = \text{dom}(z)$ contains arbitrarily large squares and moreover, for all n , $z : \text{range}(P) \rightarrow T_1$ is a partial tiling valid on a square of size n . It now follows from the König infinity lemma [17] that there exists a valid T_1 -tiling of the plane. \square

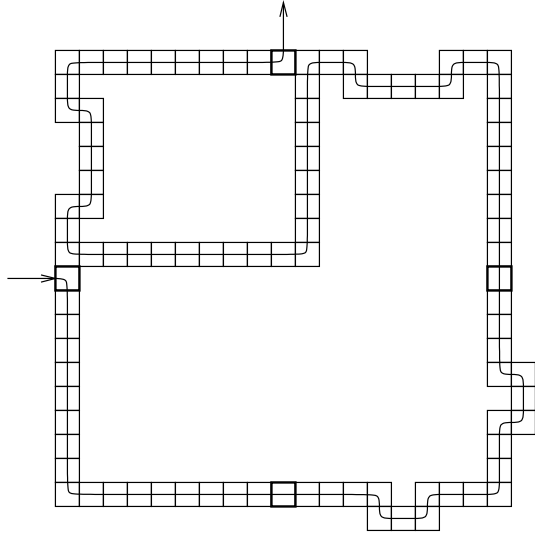


Figure 2. A ribbon motif with input tile on the west and output tile on the north.

Having proved the undecidability of existence of an infinite directed zipper, we shall use this result to show that the existence of an infinite ribbon is also undecidable.

Theorem 3.3 *The set $\{T \mid T \text{ is a tile system and there exists an infinite } T\text{-ribbon}\}$ is undecidable.*

Proof. Reduce the set of Theorem 3.2 to this set.

Given a directed tile system (T, d) , we will construct a tile system T' such that there exists an infinite (T, d) -directed zipper iff there exists an infinite T' -ribbon. Recall that a zipper differs from a ribbon in that in a ribbon only consecutive tiles are required to have matching glues on abutting edges, while on a zipper even non-consecutive neighbouring tiles have to have matching glues on their abutting edges.

The construction is as follows. For each directed tile $t \in T$, construct three T' -motifs. A motif is a finite T' -ribbon (P, r) of special form. We construct these ribbon motifs as follows (see Figure 2).

Each motif is essentially a tiled path that outlines the contours of a square. There are dents and bumps on the north, east, south and west sides of the motif. In addition, if (P, r) is a motif, then the first position on P is midway along one side of the motif. This side is called the input side of the motif and the tile at the first position is called the input tile of the motif. The last position on P is midway along a side of the motif different than the input side. This side is called the output side of the motif and the tile at the last position is called the output tile of the motif.

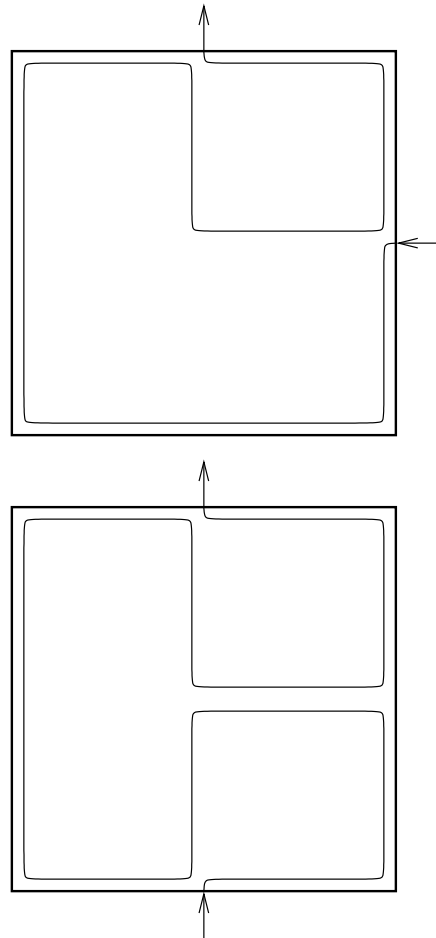


Figure 3. The two remaining variant motifs corresponding to a directed zipper-tile with direction north (the first one is depicted in Figure 2). The bumps and dents are omitted from the variant motifs for clarity.

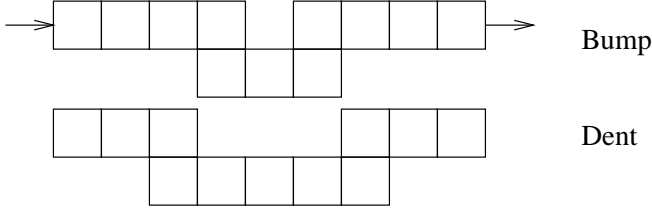


Figure 4. The bump-and-dent pair that simulate a glue. Each glue is assigned a unique position on the side of a motif. If the glues on two directed zipper tiles situated at adjacent positions match, then the bump on the first motif will be placed exactly at the necessary position to fit in the dent of the second motif.

Given a directed tile $t \in T$, we construct the three possible motifs with output side $d(t)$. We call these three motifs the “variants” of t (see Figure 3). On each variant, we put a bump on the east (south) side, to encode the glue on the east (south) side of t . We also put a dent on the west (north) side, to encode the glue on the west (north) side of t . The dents and bumps are designed so that if, for example, t_1 sticks on the north to t_2 , then the dents on the north of t_1 variant motifs fit the bumps on the south of t_2 variants (see Figure 4). If however, t_1 does not stick on the north to t_2 , then the north sides of t_1 variant motifs will overlap the bumps on the south of t_2 variants. Since overlaps are not allowed in ribbons, if t_1 does not stick on the north to t_2 , no T' -ribbon can have a t_2 variant motif directly north of a t_1 variant motif.

The glues on tiles in T' are chosen so that each tile can occur in exactly one variant motif, and in each variant motif it occurs exactly once. To this aim, the two edges of each ribbon tile connecting it with the preceding respectively succeeding tile on a variant motif are labelled so that this is the only sticking that can form along those edges. The only exception to this rule are the input (output) tiles of motifs, which have the edges connecting them to the preceding (succeeding) tiles labelled differently, as detailed below.

So far, the bumps and dents were used to simulate the glues of zipper-tiles. We now simulate the direction of a zipper-tile by incorporating it in the glues of the input and output tiles of its variant motifs. We namely use four new glues, *West-to-East*, *East-to-West*, *North-to-South* and *South-to-North* as follows. If the direction of a directed zipper-tile t_1 is north, i.e. $d(t_1) = N$, then:

- (a) the output ribbon-tiles of all three variant motifs of t_1 will have their north edge labelled *South-to-North*, and
- (b) the input tile of the variant motif of t_1 with a west (east, south) input side will have its west (east, south) edge labelled *West-to-East* (*East-to-West*, *South-to-North*).

We label in a similar way the appropriate edges of the input and output tiles of other motifs, namely those edges that are meant to connect the motifs to each other. This labelling ensures that the variant motifs will be connected to each other in the proper order dictated by the direction of the originating directed zipper-tiles.

Lastly, we want to ensure that only motifs of the kind described above can form, and that motifs can connect to each other only through their input and output tiles. To this aim, we have two new different “null” glues: $null(1)$ and $null(2)$. We label, for all the above constructed ribbon tiles, the so-far-unlabelled north and west edges with $null(1)$ and the unlabelled east and south edges with $null(2)$. Because $null(1)$ only matches $null(1)$, and $null(1)$ is only used on the west and north edges, the edges of tiles labelled with these glues will not stick to any other tiles along those edges. The same is true for $null(2)$. These “non-stick” glues ensure that the ribbon will only follow the intended motif and will not fill it in, or stick to anything outside it, except at the input and output tiles.

To summarize, given the directed tile system (T, d) we can construct the tile system T' consisting of the ribbon tiles used in all variant motifs of tiles in T , as described above.

Assume that there exists an infinite directed (T, d) -zipper (P, r) . One can easily construct an infinite T' -ribbon (P', r') . The idea is that, for each position (x_i, y_i) on P the tile $r(x_i, y_i)$ is replaced by one of its variant motifs. The variant chosen is one with input side opposite $d(r(x_{i-1}, y_{i-1}))$.

Conversely, assume that there exists an infinite T' -ribbon, (P', r') . It is clear from our choice of ribbon tiles and glues that (P', r') must consist of an infinite sequence of motifs. Hence we can construct an infinite (T, d) -directed zipper by replacing each motif with the tile t of which it is a variant. \square

Theorem 3.3 proves the undecidability of existence of an infinite ribbon. This settles the open problem [13], stating: “Problem 4.1. Given a tiling system T , is there an infinite T -snake within the infinite grid $G = \mathbb{Z} \times \mathbb{Z}$?”

In the terminology of [13], a tiling system T is exactly as defined in Section 2, a finite set of tiles, i.e. of squares with coloured edges that cannot be rotated and with infinitely many copies of each tile available; The grid G is the integer grid of positions in the plane; An infinite T -snake is a sequence of tiles on the plane in which successive tiles are adjacent along an edge and touching edges have the same colour, i.e., infinite T -snakes are (possibly self-crossing) 2-way infinite ribbons where identical tiles must be present at the crossing sites. In general, an infinite snake problem asks, given a tiling system T , and some portion P of the plane, whether there is an infinite T -snake that lies entirely within P . [13] proves that, given T and a strip of width $k \in \mathbb{N}$, the existence of an infinite T -snake that lies entirely within

the strip is decidable. Given a tile system T , and a specific tile $t_0 \in T$, the problem of whether there exists an infinite T -snake that contains t_0 is proved undecidable in [13] using methods in [12] (the case of a 1-way infinite snake starting at t_0 was proven undecidable in [11]). If the special “seed” tile is not specified, like in Problem 4.1, [13] conjectures the problem undecidable but states that “[...] it seems that this would be difficult to prove. We have not been able to adjust the proof techniques of other undecidability results for this purpose”.

Theorem 3.3, by showing the undecidability of existence of infinite non-self-crossing snakes, proves Problem 4.1 undecidable.

4 Undecidability of self-assembly of arbitrarily large supertiles

Let us return to the discussion of self-assembly. Supertiles are constructed by an incremental process starting from a single tile and proceeding by addition of single tiles that “stick” to the hitherto built structure. The problem we are addressing is whether or not, given a tile system, an infinite supertile can self-assemble with tiles from that system. To formalize our notions,

Definition 4.1 A shape is a function $f : I \rightarrow \mathbb{Z}^2$ such that I is a set of consecutive natural numbers, $0 \in I$ and $(\forall i \in I)[i > 0 \Rightarrow (\exists j \in I)[j < i \text{ and } f(i) \text{ and } f(j) \text{ are adjacent}]]$.

A shape describes thus a connected region of the plane. The size of a shape f is the cardinality of its range, i.e. the number of positions it contains, regardless of how many times they are visited. If we fill in each position of a shape with tiles from a given tile system T , we obtain the notion of a T -supertile.

Definition 4.2 For all tile systems T , a T -supertile is a pair (f, g) where f is a shape and $g : \text{range}(f) \rightarrow T$ is a function such that $(\forall i \in \text{dom}(f))[i > 0 \Rightarrow (\exists j \in \text{dom}(f))[j < i \text{ and } f(i) \text{ abuts } f(j) \text{ on the north (east, south, west) and } g(f(i)) \text{ sticks on the north (east, south, west) to } g(f(j))]]$.

The size of a T -supertile (f, g) is the size of its corresponding shape f . Two T -supertiles (f, g) and (f', g') are equivalent iff there exists $i, j \in \mathbb{Z}$ such that for all $(x, y) \in \mathbb{Z}^2$, $(x, y) \in \text{range}(f)$ iff $(x + i, y + j) \in \text{range}(f')$ and for all $(x, y) \in \text{range}(f)$, $g(x, y) = g'(x + i, y + j)$. That is, the T -supertile (f', g') is equivalent to the T -supertile (f, g) if and only if (f', g') can be obtained by translating (f, g) on the plane.

Note that the definitions of shape and supertile may differ in other papers. The ideas are similar, but the details may not be the same. The problem stated at the beginning of this section is settled by the following result.

Theorem 4.1 The following sets are undecidable:

$S_1 = \{T \mid T \text{ is a tile system and there exists an infinite } T\text{-supertile}\}$, and

$S_2 = \{T \mid T \text{ is a tile system and there exists infinitely many non-equivalent finite } T\text{-supertiles}\}$.

Proof. It is easily shown that sets S_1 and S_2 are identical to the set $\{T \mid T \text{ is a tile system and there exists an infinite } T\text{-ribbon}\}$. Hence Theorem 4.1 follows from Theorem 3.3. \square

5 Conclusion

In this paper we prove the undecidability of the problem of distinguishing tile systems that allow infinite ribbons to self-assemble from those that do not.

This result settles an open problem formerly known as the “unlimited infinite snake problem”.

We also prove the undecidability of the problem of distinguishing tile systems that allow the self-assembly of infinite supertiles from those that do not.

We introduce a “motif” construction that allows one tile system to simulate another by using geometry to represent glues. This construction may be useful in other contexts.

Acknowledgements We thank David Harel, Dale Myers, Yael Petruschka, and Qi Cheng, Ashish Goel, Ming-Deh Huang, Pablo Moisset des Espanes, Paul Rothmund for discussion and comments. This research was supported by NASA/JPL, NSF, ONR, DARPA grants to L.A., NSF Grant CCR 97-33101 to J.K., and Grant R2824A01 of the Natural Sciences and Engineering Research Council of Canada to L.K.

References

- [1] H. Abelson, D. Allen, D. Coore, C. Hanson, G. Homsy, T. Knight, R. Nagpal, E. Rauch, G. Sussman and R. Weiss. Amorphous computing. *Communications of the ACM* 43(2000), 74-82.
- [2] L. Adleman. Molecular computation of solutions to combinatorial problems. *Science* 266(1994), 1021–1024.
- [3] L. Adleman. Towards a mathematical theory of self-assembly. *Technical Report* 00-722(2000), Department of Computer Science, University of Southern California.
- [4] L. Adleman, Q. Cheng, A. Goel, M. Huang and H. Wasserman. Linear self-assemblies: equilibria, entropy, and convergence rates. Unpublished, 2000.
- [5] L. Adleman, Q. Cheng, A. Goel, M. Huang, D. Kempe, P. Moisset des Espanes and P. Rothmund. Combinatorial optimization problems in self-assembly. *Proc. ACM Symposium on Theory of Computing (STOC) 2002*, 23-32.

- [6] L.Adleman, Q.Cheng, A.Goel and M.Huang. Running time and program size for self-assembled squares. *Proc. ACM Symposium on Theory of Computing (STOC) 2001*, 740-748.
- [7] R.Berger. The undecidability of the domino problem. *Mem.Amer.Math.Soc.* 66(1966).
- [8] M. Dubois et al. Self-assembly of regular hollow icosahedra in salt-free cationic solutions. *Nature*, 411(2001), 672-675.
- [9] M. Gomez-Lopez, J. Preece, and J. Stoddart. The art and science of self-assembling molecular machines. *Nanotechnology*, Vol. 7, No. 3 (1996), 183-192.
- [10] D. Gracias, J. Tien, T. Breen, C. Hsu and G. Whitesides. Forming electrical networks in three dimensions by self-assembly. *Science* 289(2000), 1170-1173.
- [11] H.D.Ebbinghaus. Domino threads and complexity. In E.Borger, ed. *Computation Theory and Logic*. Lecture Notes in Computer Science, 270(1987), Springer, 131-142.
- [12] Y.Etzion. On the solvability of domino snake problems. M.Sc. Thesis. Dept. of Applied Math. and Computer Science, The Weizmann Institute of Science, Rehovot, Israel, 1991.
- [13] Y.Etzion-Petruschka, D.Harel, D.Myers. On the solvability of domino snake problems. *Theoretical Computer Science*, 131(1994), 243-269. (Received July 1992.)
- [14] D.Hilbert. Über die stetige Abbildung einer Linie auf ein Flächenstück. *Math. Ann.* 38(1891), 459-460.
- [15] J.Kari. Reversibility of 2D cellular automata is undecidable. *Physica D*, 45(1990), 379-385.
- [16] J.Kari. Reversibility and surjectivity problems of cellular automata. *Journal of Computer and System Sciences*, vol.48, no.1(1994), 149-182.
- [17] D.König. Über eine Schlussweise aus dem Endlichen ins Unendliche. *Acta Litt.Sci.*, Szeged, 3(1927), 121-130.
- [18] G. Lopinski, D. Wayner and R. Wolkow. Self-directed growth of molecular nano-structures on silicon. *Nature* 406(2000), 48.
- [19] C. Mao, T. LaBean, J. Reif and N. Seeman. Logical computation using algorithmic self-assembly of DNA triple-crossover molecules. *Nature* 407(2000), 493-496.
- [20] D.Myers. Decidability of the tiling connectivity problem. Abstract 79T-E42, *Notices Amer. Math. Soc.*, 195(26)(1979), A-441.
- [21] Lagoudakis and T. LaBean. 2D DNA self-assembly for satisfiability. In *DIMACS Series*, 54(2000), E. Winfree and D.K. Gifford Eds., *DNA-Based Computers V*, American Mathematical Society, 141-154.
- [22] R. Plass, J.A. Last, N.A. Bartelt, G.L. Kellogg. Self-assembled domain patterns. *Nature*, 412(2001), 875.
- [23] J. Reif. Local parallel biomolecular computation. In *DIMACS Series*, 48 (1999), H. Rubin, D.H.Wood Eds., *DNA Based Computers III*, American Mathematical Society, 217-254.
- [24] R.M.Robinson. Undecidability and nonperiodicity for tilings of the plane. *Inven.Math.* 12(1971), 177-209.
- [25] P. Rothmund. Using lateral capillary forces to compute by self-assembly. *Proceedings of the National Academy of Sciences*, vol. 9(2000), 984-989.
- [26] P. Rothmund. Theory and experiments in algorithmic self-assembly, University of Southern California Ph.D. Thesis, 2001.
- [27] P. Rothmund and E. Winfree. The program-size complexity of self-assembled squares. *ACM Symposium on Theory of Computing (STOC) 2001*, 459-468.
- [28] J.H.Schön, H.Meng and Z.Bao. Self-assembled monolayer organic field-effect transistors. *Nature*, 413(2000), 713-715.
- [29] H. Wang. Proving theorems by pattern recognition. II. *BellSystems Technical Journal*, 40(1961), 1-42.
- [30] E. Winfree, X. Yang and N. Seeman, Universal computation via self-assembly of DNA: Some theory and experiments. In *DIMACS Series*, 44(1998), L.F.Landweber, E.Baum, eds., *DNA Based Computers II*, American Mathematical Society, 191-213.
- [31] E. Winfree, F. Liu, L. Wenzler, N. Seeman. Design and self-assembly of two-dimensional DNA crystals. *Nature* 394(1998), 539-544.
- [32] E. Winfree. Algorithmic Self-Assembly of DNA, Ph.D. thesis. California Institute of Technology, Pasadena, 1998.
- [33] G. Whitesides, J. Mathias and Christopher T. Seto. Molecular self-assembly and nanochemistry: a chemical strategy for the synthesis of nanostructures, *Science*, 254(1991), 1312-1319.