3-Color Bounded Patterned Self-assembly*
(extended abstract)

Lila Kari[1], Steffen Kopecki[1], and Shinnosuke Seki[2]

[1] Department of Computer Science, The University of Western Ontario, London Ontario N6A 5B7, Canada
`lila@csd.uwo.ca, steffen@csd.uwo.ca`

[2] Helsinki Institute of Information Technology (HIIT), Department of Computer Science, Aalto University, P. O. Box 15400, FI-00076, Aalto, Finland
`shinnosuke.seki@aalto.fi`

Abstract. Patterned self-assembly tile set synthesis (PATS) is the problem of finding a minimal tile set which uniquely self-assembles into a given pattern. Czeizler and Popa proved the NP-completeness of PATS and Seki showed that the PATS problem is already NP-complete for patterns with 60 colors. In search for the minimal number of colors such that PATS remains NP-complete, we introduce multiple bound PATS (MBPATS) where we allow bounds for the numbers of tile types of each color. We show that MBPATS is NP-complete for patterns with just three colors and, as a byproduct of this result, we also obtain a novel proof for the NP-completeness of PATS which is more concise than the previous proofs.

1 Introduction

Tile self-assembly is the autonomous formation of a structure from individual *tiles* controlled by local attachment rules. One application of self-assembly is the implementation of nanoscopic tiles by DNA strands forming double crossover tiles with four unbounded single strands [10]. The unbounded single strands control the assembly of the structure as two, or more, tiles can attach to each other only if the bonding strength between these single strands is big enough. The general concept is to have many copies of the same tile types in a solution which then form a large crystal-like structure over time; often an initial structure, the *seed*, is present in the solution from which the assembly process starts.

A mathematical model describing self-assembly systems is the *abstract tile self-assembly model* (aTAM), introduced by Winfree [9]. Many variants of aTAMs have been studied: a main distinction between the variants is whether the *shape* or the *pattern* of a self-assembled structure is studied. In this paper we focus on the self-assembly of patterns, where a property, modeled as color, is assigned to each tile; see for example [6] where fluorescently labeled DNA tiles self-assemble

* The research of L. K. and S. K. was supported by the NSERC Discovery Grant R2824A01 and UWO Faculty of Science grant to L. K. The research of S. S. was supported by the HIIT Pump Priming Project Grant 902184/T30606.

into Sierpinski triangles. Formally, a pattern is a rectilinear grid where each vertex has a color: a k-colored $m \times n$-pattern P can be seen as a function $P\colon [m] \times [n] \to [k]$, where $[i] = \{1, 2, \ldots, i\}$. The optimization problem of *patterned self-assembly tile set synthesis* (PATS), introduced by Ma and Lombardi [4], is to determine the minimal number of tile types needed to uniquely self-assemble a given pattern starting from an L-shaped seed. In this paper, we consider the decision variant of PATS, defined as follows:

Problem (k-PATS)
GIVEN: A k-colored pattern P and an integer m;
OUTPUT: "Yes" if P can uniquely be self-assembled by using m tile types.

Czeizler and Popa proved that PATS, where the number of colors on an input pattern is not bounded, is NP-hard [1], but the practical interest lies in k-PATS. Seki proved 60-PATS is NP-hard [8]. By the nature of the biological implementations, the number of distinct colors in a pattern can be considered small. In search for the minimal number k for which k-PATS remains NP-hard, we investigate a modification of PATS: *multiple bound* PATS (MBPATS) uses individual bounds for the number of tile types of each color.

Problem (k-MBPATS)
GIVEN: A pattern P with colors from $[k]$ and $m_1, \ldots, m_k \in \mathbb{N}$;
OUTPUT: "Yes" if P can uniquely be self-assembled by using m_i tile types of color i, for $i \in [k]$.

The main contribution of this paper is a polynomial-time reduction from PATS to 3-MBPATS which proves the NP-hardness of 3-MBPATS. However, our reduction does not take every pattern as input, we only consider a restricted subset of patterns for which PATS is known to remain NP-hard. The patterns we use as input are exactly those patterns that are generated by a polynomial-time reduction from 3-SAT to PATS. Using one of the reductions which were presented in [1,8] as a foundation for our main result turned out to be unfeasible. Therefore, we present a novel proof for the NP-hardness of PATS which serves well as foundation for our main result. Furthermore, our reduction from 3-SAT to PATS is more concise compared to previous reductions in the sense that in order to self-assemble a pattern P we only allow three more tile types than colors in P. In Czeizler and Popa's approach the number of additional tile types is linear in the size of the input formula and Seki uses 84 tile types with 60 colors.

Let us note first that the decision variants of PATS and MBPATS can be solved in NP by simple "guess and check" algorithms. Before we prove NP-hardness of k-PATS, in Sect. 3, and 3-MBPATS, in Sect. 4, we introduce the formal concepts of patterned tile assembly systems, in Sect. 2. We only present some shortened proofs for our lemmas. Full proofs for all lemmas as well as additional figures, depicting our patter designs, can be found in the arXiv version [3].

2 Rectilinear Tile Assembly Systems

In this section we formally introduce patterns and rectilinear tile assembly systems. An excellent introduction to the fundamental model aTAM is given in [7].

Let C be a finite *alphabet of colors*. An $m \times n$*-pattern* P, for $m, n \in \mathbb{N}$, with colors from C is a mapping $P\colon [m] \times [n] \to C$. By $C(P) \subseteq C$ we denote the colors in the pattern P, i. e., the codomain or range of the function P. The pattern P is called k*-colored* if $|C(P)| \leq k$. The width and height of P are denoted by $w(P) = m$ and $h(P) = n$, respectively. The pattern is arranged such that position $(1, 1)$ is on the bottom left and position $(m, 1)$ is on the bottom right.

Let Σ be a finite *alphabet of glues*. A *colored Wang tile*, or simply *tile*, $t \in C \times \Sigma^4$ is a unit square with a color from C and four glues from Σ, one on each of its edges. $\chi(t) \in C$ denotes the color of t and $t(N)$, $t(E)$, $t(W)$, and $t(S)$ denote the glues on the north, east, west, and south edges of t, respectively. We also call the south and west glues the *inputs* of t while the north and east glues are called *outputs* of t.

A *rectilinear tile assembly system* (RTAS) (T, σ) over C and Σ consists of a set of colored Wang tiles $T \subseteq C \times \Sigma^4$ and an L-shaped seed σ. The seed σ covers positions $(0, 0)$ to $(m, 0)$ and $(0, 1)$ to $(0, n)$ of a two-dimensional Cartesian grid and it has north glues from Σ on the positions $(1, 0)$ to $(m, 0)$ and east glues from Σ on positions $(0, 1)$ to $(0, n)$. We will frequently call T an RTAS without explicitly mentioning the seed. The RTAS T describes the self-assembly of a structure: starting with the seed, a tile t from T can attach to the structure at position $(x, y) \in [m] \times [n]$, if its west neighbor at position $(x - 1, y)$ and south neighbor at position $(x, y - 1)$ are present and the inputs of t match the adjacent outputs of its south and west neighbors; the self-assembly stops when no more tiles in T can be attached by this rule. Arbitrarily many copies of a each tile type in T are considered to be present while the structure is self-assembled, thus, one tile type can appear in multiple positions. A *tile assignment* in T is a function $f\colon [m] \times [n] \to T$ such that $f(x, y)(W) = f(x - 1, y)(E)$ and $f(x, y)(S) = f(x, y - 1)(N)$ for $(x, y) \in [m] \times [n]$. The RTAS self-assembles a pattern P if there is a tile assignment f in T such that the color of each tile in the assignment f is the color of the corresponding position in P, i. e., $\chi \circ f = P$. A terminological convention is to call the elements in T *tile types* while the elements in a tile assignment are called *tiles*.

A *directed RTAS* (DRTAS) T is an RTAS where any two distinct tile types $t_1, t_2 \in T$ have different inputs, i. e., $t_1(S) \neq t_2(S)$ or $t_1(W) \neq t_2(W)$. A DRTAS has at most one tile assignment and can self-assemble at most one pattern. If T self-assembles an $m \times n$-pattern P, it defines the function $P_T\colon [m] \times [n] \to T$ such that $P_T(x, y)$ is the tile in position (x, y) of the tile assignment given by T. In this paper, we investigate minimal RTASs which uniquely self-assemble one given pattern P. As observed in [2], if P can be uniquely self-assembled by an RTAS with m tile types, then P can also be (uniquely) self-assembled by a DRTAS with m tile types.

3 NP-hardness of Pats

In this section, we prove the NP-hardness of Pats. The proof we present uses many techniques that have already been employed in [1,8]. Let us also point out that we do not intend to minimize the number of colors used in our patterns or the size of our patterns. Our motivation is to give a proof that is easy to understand and serves well as a foundation for the results in Sect. 4.

A boolean formula F over variables V in *conjunctive normal form with three literals per clause*, 3-CNF for short, is a boolean formula such that

$$F = (c_{1,1} \vee c_{1,2} \vee c_{1,3}) \wedge (c_{2,1} \vee c_{2,2} \vee c_{2,3}) \wedge \cdots \wedge (c_{\ell,1} \vee c_{\ell,2} \vee c_{\ell,3})$$

where $c_{i,j} \in \{v, \neg v \mid v \in V\}$ for $i \in [\ell]$ and $j = 1, 2, 3$. It is well known that the problem 3-Sat, to decide whether or not a given formula F in 3-CNF is satisfiable, is NP-complete; see e. g., [5]. The NP-hardness of Pats follows by the polynomial-time reduction from 3-Sat to Pats, stated in Theorem 1.

Theorem 1. *For every formula F in 3-CNF there exists a pattern P_F such that F is satisfiable if and only if P_F can be self-assembled by a DRTAS with at most $|C(P_F)|+3$ tile types. Moreover, P_F can be computed from F in polynomial time.*

Theorem 1 follows by Lemmas 3 and 5, which are presented in the following.

The pattern P_F consists of several rectangular *subpatterns* which we will describe in the following. None of the subpatterns will be adjacent to another subpattern. The remainder of the pattern P_F is filled with *unique colors*; a color c is unique in a pattern P if it appears only in one position in P, i. e., $\left|P^{-1}(c)\right| = 1$. As a technicality that will become useful only in the proof of Theorem 2, we require that each position adjacent to the L-shaped seed or to the north or east border of pattern P_F has a unique color. Clearly, for each unique color in P_F we require exactly one tile in any DRTAS which self-assembles P_F. Since each subpattern is surrounded by a frame of unique colors, the subpatterns can be treated as if each of them would be adjacent to an L-shaped seed and we do not have to care about the glues on the north border or east border of a subpattern.

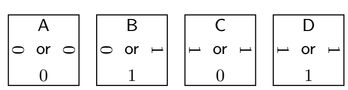

Fig. 1. The four tile types used to implement the or-gate.

As stated earlier, the number of tile types m that is required to self-assemble P_F, if F is satisfiable, is $m = |C(P_F)| + 3$. Actually, every color in $C(P_F)$ will require one tile type only except for one color which is meant to implement an or-gate; see Fig. 1. Each of the tile types with color or is supposed to have west input $w \in \{0, 1\}$, south input $s \in \{0, 1\}$, east output $w \vee s$, and an independent north output.

Our first subpattern p, shown in Fig. 2, ensures that every DRTAS which self-assembles the subpattern p contains at least three tile types with color or. For the upcoming proof of Theorem 2 we need a more precise observation which draws a connection between the number of distinct output glues and the number of distinct tile types with color or.

Lemma 1. *A DRTAS T which self-assembles a pattern including the subpattern p contains either*

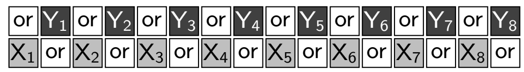

Fig. 2. The subpattern p.

i.) *three distinct tile types $o_1, o_2, o_3 \in T$ with color* or *all having distinct north and east glues,*

ii.) *four distinct tile types $o_1, o_2, o_3, o_4 \in T$ with color* or *all having distinct north glues and together having at least two distinct east glues,*

iii.) *four distinct tile types $o_1, o_2, o_3, o_4 \in T$ with color* or *all having distinct east glues and together having at least two north glues, or*

iv.) *eight distinct tile types $o_1, \ldots, o_8 \in T$ with color* or *all having distinct east or north glues.*

Lemma 1 follows by the fact that each of the tiles with colors Y_1 to Y_8 has the or-gate as west and south neighbors, hence, the number of east glues times the number of north glues of all tile types with color or has to be at least eight.

We aim to have statement *ii.)* of Lemma 1 satisfied, but so far all four statements are possible. The subpatterns q_1 to q_5 in Fig. 3 will enforce the functionality of the or-gate tile types.

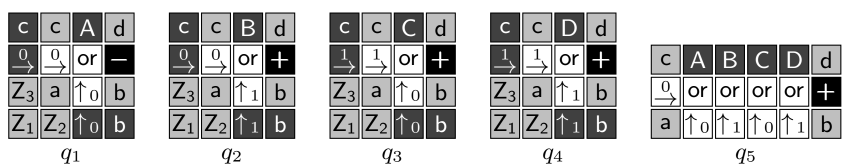

Fig. 3. The subpatterns q_1 to q_5.

Lemma 2. *Let P be a pattern that contains the subpatterns p and q_1 to q_5, and let $m = |C(P)| + 3$. A DRTAS T with at most m tile types which self-assembles pattern P contains four tile types with color* or *of the forms shown in Fig. 1. For every other color in $C(P)$ there exists exactly one tile type in T. Moreover, the tile type with color* $\xrightarrow{0}$ *has east output* 0 *and the tile type with color* + *has west input* 1*.*

There are at least three or-gate tile types, thus, only the color of one tile type in T is not determined yet. The clue of patterns q_1 to q_4 is that if, e. g., the two tiles with colors c in q_1 and q_2 were of different types, there would be only one tile type of the other colors, and in particular, their west neighbors would be of the same type as well as their south neighbors. Thus, these two tile types would have the same inputs, which is prohibited for DRTAS by definition. This implies that the tiles with colors A and B have the same west input and can only be placed because their south neighbors, the or-gate tiles in q_1 and q_2, are of different types. By analogous arguments the four or-gate tiles in q_1 to q_4 are of four different types. Subpattern q_5 ensures that the east and west glues of the or-gates match in the way shown in Fig. 1.

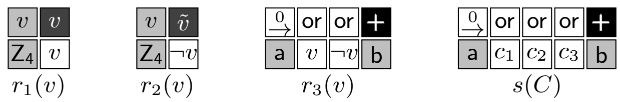

Fig. 4. The subpatterns $r_1(v)$ to $r_3(v)$ for a variable $v \in V$ and the subpattern $s(C)$ for a clause $C = (c_1 \vee c_2 \vee c_3)$ in F where $c_i = v$ or $c_i = \neg v$ for some variable $v \in V$ and $i = 1, 2, 3$.

The subpatterns that we defined so far did not depend on the formula F. Now, for each variable $v \in V$ we define three subpatterns $r_1(v)$, $r_2(v)$, $r_3(v)$ and for a clause C from F we define one more subpattern $s(C)$; these patterns are given by Fig. 4. For a formula F in 3-CNF we let P_F be the pattern that contains all the subpatterns p, q_1 to q_5, $r_1(v)$ to $r_3(v)$ for each variable $v \in V$, and $s(C)$ for each clause C from F, where each subpattern is placed to the right of the previous subpattern with one column of unique colors in between. Then, P_F has height 6, because the top and bottom rows contain unique colors only, and P_F has width $45 + 11 \cdot |V| + 6 \cdot \ell$. The next lemma follows from this observation.

Lemma 3. *Given a formula F in 3-CNF, the pattern P_F can be computed from F in polynomial time.*

The subpatterns $r_1(v)$ and $r_2(v)$ ensure that the two tile types with colors $\boxed{v}$ and $\boxed{\neg v}$ have distinct north outputs. The subpattern $r_3(v)$ then implies that one of the north glues is 0 and the other one is 1.

Lemma 4. *Let P_F be the pattern for a formula F over variables V in 3-CNF and let T be a DRTAS with at most $m = |C(P_F)| + 3$ tile types which self-assembles pattern P_F. For all variables $v \in V$, there is a unique tile type $t_v^{\oplus} \in T$ with color $\boxed{v}$ and a unique tile type $t_v^{\ominus} \in T$ with color $\boxed{\neg v}$ such that either $t_v^{\oplus}(N) = 1$ and $t_v^{\ominus}(N) = 0$ or $t_v^{\oplus}(N) = 0$ and $t_v^{\ominus}(N) = 1$.*

Now, these glues serve as input for the or-gates in the subpatterns $s(C)$. The following lemma concludes the proof of Theorem 1.

Lemma 5. *Let P_F be the pattern for a formula F over variables V in 3-CNF and let $m = |C(P_F)| + 3$. The formula F is satisfiable if and only if P_F can be self-assembled by a DRTAS T with at most m tile types.*

The formula F is satisfiable if and only if there is a variable assignment $f \colon V \to \{0, 1\}$ which satisfies every clause in F. In order for $s(C)$ with $C = (c_1 \vee c_2 \vee c_3)$ to self-assemble, one of the north glues of the tiles for c_1, c_2, or c_3 has to be 1. Let $t_v^{\oplus}$ and $t_v^{\ominus}$ for $v \in V$ as before. Since $t_v^{\oplus}(N)$ and $t_v^{\ominus}(N)$ represent opposite truth values, the pattern P can be self-assembled using m tile types if and only if $f(v) = t_v^{\oplus}(N)$ satisfies every clause in F. How the remaining tile types and glues in T can be chosen is shown in the arXiv version [3].

4 NP-hardness of 3-mbPats

The purpose of this section is to prove the NP-hardness of 3-mbPats. Let us define a set of restricted input pairs $\mathcal{I}$ for Pats. The set $\mathcal{I}$ contains all pairs

(P, m) where $P = P_F$ is the pattern for a formula F in 3-CNF as defined in Sect. 3 and $m = |C(P)| + 3$. Consider the following restriction of PATS.

Problem (MODIFIED PATS)
GIVEN: A pair (P, m) from $\mathcal{I}$;
OUTPUT: "Yes" if P can uniquely be self-assembled by using m tile types.

As we choose exactly those pairs (P, m) as input for the problem that are generated by the reduction, stated in Theorem 1, we obtain the following corollary which forms the foundation for the result in this section.

Corollary 1. MODIFIED PATS *is* NP*-hard.*

The NP-hardness of 3-MBPATS follows by the polynomial-time reduction from MODIFIED PATS to 3-MBPATS, stated in Theorem 2.

Theorem 2. *For every input pair $(P, m) \in \mathcal{I}$ there exist a black/white/gray-colored pattern Q and integers m_b, m_w, m_g such that: P can be self-assembled by a DRTAS with at most m tile types if and only if Q can be self-assembled by a DRTAS with at most m_b black tile types, m_w white tile types, and m_g gray tile types. Moreover, the tuple (Q, m_b, m_w, m_g) can be computed from P in polynomial time.*

Lemma 12 states the "if part" and Lemma 8 states the "only if part" of Theorem 2. Lemma 6 states that (Q, m_b, m_w, m_g) can be computed from P in polynomial time.

For the remainder of this section, let $(P, m) \in \mathcal{I}$ be one fixed pair, let $C = C(P)$ and $k = |C|$. We may assume that $C = [k]$ is a subset of the positive integers. The tile bounds are $m_b = 1$ for black tile types, $m_w = 5k - 3(w(P) + h(P)) + 14$ for white tile types, and $m_g = 2k + 3$ for gray tile types. Note that, due to the pattern design in Sect. 3, $h(P) = 6$ is constant.

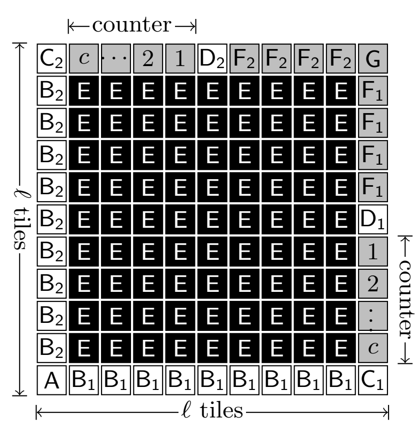

Fig. 5. Black/white/gray supertile which portrays a color $c \in C$.

Let $\ell = 5k + 8$. For a color $c \in C$, we define an $\ell \times \ell$ square pattern as shown in Fig. 5. We refer to this pattern as well as to its underlying tile assignment as *supertile*. In contrast to the previous section, the positions in the supertile are labeled which does not mean that the colors or the tiles used to self-assemble the pattern are labeled; the colors are black, white, or gray. The horizontal and vertical *color counters* are the c gray positions in the top row, respectively right column, which are succeeded by a white tile in position D_2, respectively D_1. The color counters illustrate the color c that is *portrayed* by the supertile. The patterns of two supertiles which portray two distinct colors differ only in the place the white tile is positioned in its top row and right column.

For colors in the bottom row and left column of the pattern P we use *incomplete supertiles*: a supertile portraying a color c in the bottom row of pattern P lacks the white row with positions A, B_1, and C_1; a supertile representing a color c in the left column of pattern P lacks the white column with positions A, B_2, and C_2. In particular, the supertile portraying color $P(1,1)$ does not contain any of the positions A, B_1, B_2, C_1, and C_2. Recall that all incomplete supertiles portray a color c that is unique in P.

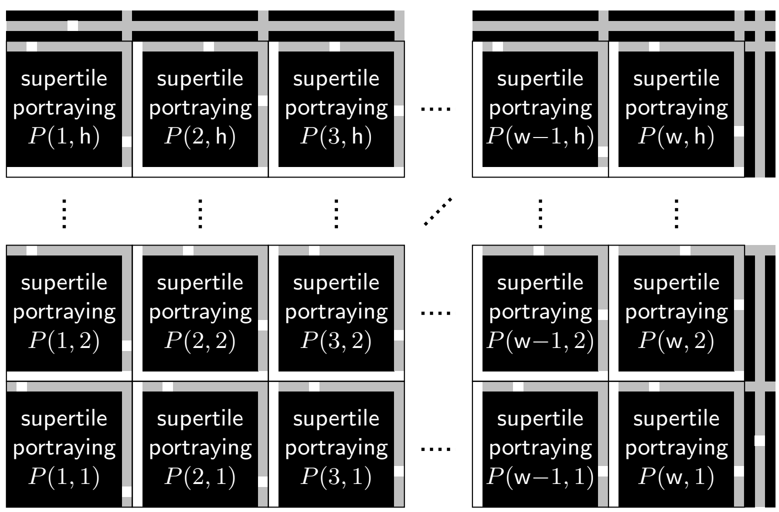

Fig. 6. Black/white/gray pattern Q defined by the k-color pattern P with $\mathsf{w} = w(P)$ and $\mathsf{h} = h(P)$.

The pattern Q is shown in Fig. 6. By $Q\langle x, y\rangle$ we denote the pattern of the supertile covering the square area spanned by positions $((x-1)\cdot \ell, (y-1)\cdot \ell)$ and $(x\cdot \ell - 1, y \cdot \ell - 1)$ in Q; the incomplete supertiles cover one row and/or column less. The pattern is designed such that supertile $Q\langle x, y\rangle$ portrays the color $P(x, y)$ for all $x \in [w(P)]$ and $y \in [h(P)]$. Additionally, Q contains three *gadget rows* and three *gadget columns* which are explained in Fig. 7. The purpose of these gadget rows and columns is to ensure that the color counters can only be implemented in one way when using no more than m_g gray tile types. All together Q is of dimensions $w(Q) = \ell \cdot w(P) + 2$ times $h(Q) = \ell \cdot h(P) + 2$. Obviously, the pattern Q can be computed from P in polynomial time.

Lemma 6. *(Q, m_b, m_w, m_g) can be computed from P in polynomial time.*

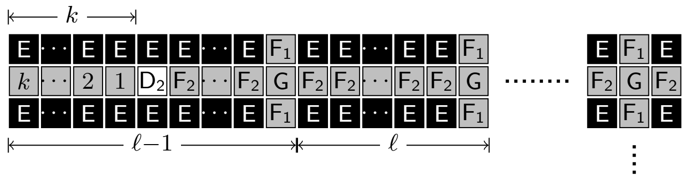

Fig. 7. The gadget rows on the north border of the pattern Q, the gadget columns are symmetrical: the middle row (resp., column) contains gray tiles except for one white tile in position $k+1$; the upper and lower rows (resp., left and right columns) contain gray tiles in positions above the gray column (resp., right of the gray row) of a supertile, the other tiles are black.

For a DRTAS Θ which self-assembles Q, we extend our previous notion such that $Q_\Theta\langle x, y\rangle$ denotes the tile assignment of supertile $Q\langle x, y\rangle$ given by Θ. In the following, we will prove properties of such a DRTAS Θ. Our first observation is about the black and gray tile types plus two of the white tile types.

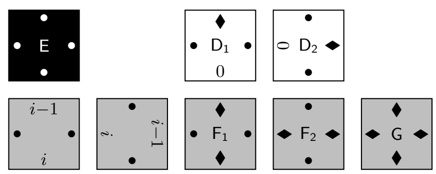

Fig. 8. The black tile type, two of the white tile types, and all gray tile types: the labeled tile types are used in the corresponding positions of each supertile and the gadget pattern; the unlabeled tile types, called *counter tiles* for $i \in [k]$, implement the vertical and horizontal color counters.

Lemma 7. *Let Θ be a DRTAS which self-assembles the pattern Q using at most $m_b = 1$ black tile types and $m_g = 2k+3$ gray tile types. The black and gray tile types in Θ are of the form shown in Fig. 8 and Θ contains two white tiles of the form shown in the figure. In every supertile, the horizontal and vertical color counters are implemented by a subset of the counter tile types and for a position* E, D_1, D_2, F_1, F_2, *or* G *the correspondingly labeled tile type is used. Furthermore, the glues $\bullet, \blacklozenge, 0, 1, \ldots, k$ are all distinct.*

Since there is only one black tile type which can tile the black square area in each supertile, the black tile type has to be of the given form. In particular, no kind of information can be passed through the black square areas in the supertiles. The k gray tiles, followed by one white tile in the gadget rows and columns, ensure that some kind of horizontal and vertical counter tile types are present in Θ. The three remaining gray tile types have to be used for positions F_1, F_2, and G; it is easy to see that they are of the given forms.

Remark 1. Consider a DRTAS Θ that self-assembles the pattern Q using most m_b black tile types and m_g gray tile types. If we have a look at the tile assignment of the black square plus the gray column and row in a supertile, we see that this block has inputs $\bullet$ on all edges except for edges where the color counters are initialized and it has outputs • on all edges, except for its right-most and top-most output edges which are ♦. This means that all information on how to initialize the color counters has to be carried through the white lines and rows, that are, the tiles in positions A, B_1, B_2, C_1, C_2. Moreover, the tile in position A is the only one with non-generic input from other supertiles. This tile fully determines the tile assignment of the supertile and can be seen as the *control tile* or *seed* of the supertile. Henceforth, for a supertile $s = Q_\Theta\langle x, y\rangle$ we extend our notion of glues such that $s(S)$ and $s(W)$ denote the south and west input of the tile in position A, respectively, $s(N)$ and $s(E)$ denote the north and east output of the tiles in positions C_2 and C_1, respectively. For incomplete supertiles only one of $s(N)$ or $s(E)$ is defined.

Two supertiles in Q_Θ are considered distinct if their tile assignment differs in at least one position. By the observations above, two complete supertiles are distinct if and only if their control tiles are of distinct types; this is equivalent to require that the inputs of the two supertiles differ. Since incomplete supertiles portray unique colors in P, they are distinct from any supertile in Q_Θ but itself.

There is some flexibility in how the white tile types are implemented in a DRTAS Θ which self-assembles Q. Let us present one possibility which proves the "only if part" of Theorem 2.

Lemma 8. *If P can be self-assembled by a DRTAS T with m tile types, then Q can be self-assembled by a DRTAS Θ using m_b black tile types, m_w white tile types, and m_g gray tile types.*

Proof. Let Θ contain the tile types given in Fig. 8. For a supertile portraying a color $c \in C \setminus \{\boxed{\mathsf{or}}\}$ we use the five tile types given in Fig. 9. Note that we need less tile types for incomplete supertiles which leads to $5\cdot(k-1)-3\cdot(h(P)+w(P))+1$ white tile types in total. Thus, we have 16 white tile types left for the or-gate.

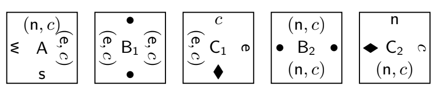

Fig. 9. White tile types for the supertile portraying a color $c \in C$, except for the or-gate, where $t \in T$ with $c = \chi(t)$, $\mathsf{n} = t(N)$, $\mathsf{e} = t(E)$, $\mathsf{s} = t(S)$, and $\mathsf{w} = t(W)$.

Since three of the or-gates have the same east output, see Fig. 1, they can share tile types in positions B_1 and C_1. The 16 white tile types in Fig. 10 are used to self-assemble the supertiles representing the or-gates. The tile types are designed such that they can self-assemble pattern Q. □

For the converse implication of Theorem 2, let us show how to obtain a DRTAS that self-assembles P from the supertiles in Q_Θ. The following result follows from the bijection between supertiles in Q_Θ and tiles in P_T.

Lemma 9. *Let Θ be a DRTAS which self-assembles Q using at most m_b black tile types and m_g gray tile types, and let*

$$S = \{Q_\Theta\langle x, y\rangle \mid x \in [w(P)], y \in [h(P)]\}$$

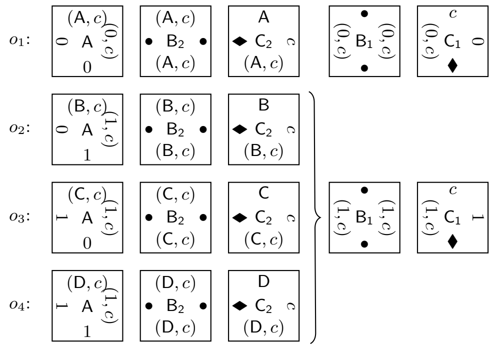

Fig. 10. White tile types for supertiles portraying the or-gate where $o_1, o_2, o_3, o_4 \in T$ are defined in Fig. 1.

be the set of all distinct supertiles in Q_Θ. There exists a DRTAS T with $|S|$ tile types which self-assembles P such that for each supertile $s \in S$ there exists a tile type $t_s \in T$ with the same glues on the respective edges and s portrays the color of t_s. For an incomplete supertile the statement holds for the defined glue.

We continue with the investigation of the white tile types that are used to self-assemble the pattern Q. The next lemma follows by a case study of what would go wrong if one tile type were used in two of the positions.

Lemma 10. *Let Θ be a DRTAS which self-assembles the pattern Q using at most m_b black tile types and m_g gray tile types. A white tile type from Θ which is used in one of the positions A, B_1, B_2, C_1, C_2, D_1, or D_2 cannot be used in another position in any supertile.*

Let B_1^* be the right-most position B_1 in a supertile, adjacent to position C_1, and let B_2^* be the top-most position B_2 in a supertile, adjacent to position C_2. The following argument is about tiles in the five positions $K = \{\mathsf{A}, \mathsf{B}_1^*, \mathsf{B}_2^*, \mathsf{C}_1, \mathsf{C}_2\}$ of each supertile. Following Remark 1 it is clear that a tile in position A fully determines the supertile, tiles in positions B_1^* and C_1 carry the color and the east glue of a supertile, whereas tiles in positions B_2^* and C_2 carry the color and the north glue.

Lemma 11. *Let Θ be a DRTAS which self-assembles Q using at most m_b black tile types and m_g gray tile types. Let s_1 and s_2 be supertiles in Q_Θ.*
i.) If s_1 and s_2 portray different colors, they cannot share any tile types in positions from K.
ii.) If $s_1(E) \neq s_2(E)$, they cannot share any tile types in A, B_1^*, or C_1.
iii.) If $s_1(N) \neq s_2(N)$, they cannot share any tile types in A, B_2^*, or C_2.
The three statements hold for all available positions in incomplete supertiles.

Let us conclude the proof of Theorem 2.

Lemma 12. *The pattern P can be self-assembled by a DRTAS T with m tile types if Q can be self-assembled by a DRTAS Θ with m_b black tile types, m_w white tile types, and m_g gray tile types.*

Proof. We show that Q_Θ cannot contain more than m distinct supertiles, then, the claim follows from Lemma 9. The black, gray, and two white tile types in Θ are defined by Lemma 7. The number of distinct tile types in Θ that can be used as control tiles, equals to the number of distinct complete supertiles of Q_Θ. By Lemma 11 we need five white tile types for each complete supertile portraying a color in $C \setminus \{\boxed{\textsf{or}}\}$; of these five tile types one can be used as control tile. For incomplete supertiles we need just two white tile types, and none for the one supertile portraying $P(1,1)$. There are 16 white tile types left for the or-gate supertiles. From Lemma 1 and Lemma 9 we infer that among these 16 white tile types we can have at most four control tiles. Therefore, the number of distinct supertiles in Q_Θ is $k + 3 = m$ — concluding the proof. □

Conclusions

We prove that k-mbPats, a natural variant of k-Pats, is NP-complete for $k = 3$. Furthermore, we present a novel proof for the NP-completeness of Pats and our proof is more concise than previous proofs. We introduce several new techniques for pattern design in our proofs, in particular in Sect. 4, and we anticipate that these techniques can ultimately be used to prove that 2-mbPats and also 2-Pats are NP-hard.

References

1. E. Czeizler and A. Popa. Synthesizing minimal tile sets for complex patterns in the framework of patterned DNA self-assembly. In *DNA computing and molecular programming*, pages 58–72. Springer, 2012.
2. M. Göös and P. Orponen. Synthesizing minimal tile sets for patterned DNA self-assembly. In *DNA computing and molecular programming*, pages 71–82. Springer, 2011.
3. L. Kari, S. Kopecki, and S. Seki. 3-color bounded patterned self-assembly. *arXiv preprint arXiv:1306.3257*, 2013.
4. X. Ma and F. Lombardi. Synthesis of tile sets for DNA self-assembly. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 27(5):963–967, 2008.
5. C. H. Papadimitriou. *Computational complexity*. John Wiley and Sons Ltd., 2003.
6. P. W. Rothemund, N. Papadakis, and E. Winfree. Algorithmic self-assembly of DNA Sierpinski triangles. *PLoS Biology*, 2(12):e424, 2004.
7. P. W. Rothemund and E. Winfree. The program-size complexity of self-assembled squares. In *Proceedings of the thirty-second annual ACM Symposium on Theory of Computing*, pages 459–468. ACM, 2000.
8. S. Seki. Combinatorial optimization in pattern assembly (extended abstract). In *Prodeedings of Unconventional Computation and Natural Computation, 12th International Conference*, volume 7956 of *LNCS*, pages 220–231, 2013.
9. E. Winfree. *Algorithmic self-assembly of DNA*. PhD thesis, California Institute of Technology, 1998.
10. E. Winfree, F. Liu, L. A. Wenzler, and N. C. Seeman. Design and self-assembly of two-dimensional DNA crystals. *Nature*, 394(6693):539–544, 1998.