Error Detection/Correction in DNA Algorithmic Self-Assembly

Stephen Frechette

Riverside Research Institute
Lexington, MA 02421
sfrechette@rri-usa.org

Fabrizio Lombardi

Dept of Electrical and Computer Engineering
Northeastern University
Boston, MA 02115
lombardi@ece.neu.edu

Abstract

A novel error detection/correction technique for algorithmic self-assembly is presented in this paper. Through the use of a tile set that allows errors to be isolated and propagated to the boundary edge of 2D (two-dimensional) assemblies, the proposed technique permits growth errors to be detected and corrected. For assemblies in which each four-sided tile is a party to only one tile mismatch, all growth errors in the assembly can be detected and corrected using the proposed method with only two additional tiles. This technique relies on the attachment of so-called isolation tiles at set periods, thus implementing a checkpoint for error detection/correction. The physical environment and related features for the removal of the erroneous sections of an assembly are presented.

Index Terms: error detection and correction, checkpointing, error tolerance, DNA self-assembly, tiling.

1. Introduction

Structures for IC manufacturing can be programmed to self-assemble through the use of binding properties of DNA strands, which are composed of unpaired bases binding with reciprocal bases. This process uses complexes commonly referred to as *tiles*; for example, in rectangular structures a tile has four types of DNA strands. Previous works have reported that the assembly of incorrect tiles occurs with error rates between 1 to 10 percent [4]; as millions of molecules are usually involved, the presence of these errors represents a serious challenge for efficient manufacturing in the nano ranges. Therefore *error detection and correction techniques* are needed.

Error correction techniques that require additional matching ends have been proposed; in [7] pads between tiles decrease the error rate by requiring more bonds to match for each assembled tile. The use of checkpoints for DNA-self assembly has been investigated in previous works; [6] has proposed, simulated, and evaluated a checkpoint scheme based on *temperature pulsing* during crystalline array growth. Periodic pulses remove defective, and adversely remove some of the correct (error-free) sections of the crystalline arrays. In [6] bonds between incorrectly attached DNA tiles are broken, so that defective (erroneous) substructures can separate from the lattice, thus restarting growth at an earlier error-free structure. In the process outlined in [6], it is assumed that a temperature pulse (i.e., a temporary increase in temperature) removes all defective regions of the structure. Although a portion of the removed tiles are correct, the incorrect tiles are removed at a higher rate. This paper presents a tile set and a novel error correction technique with checkpointing for a rectangular pattern, that is amenable to DNA self-assembly.

2. Overview of Proposed Method

In algorithmic self-assembly, the growth of a DNA crystal is used for information processing. A set of DNA tiles is used to execute an algorithm. The *abstract Tile Assembly Model (aTAM)* [2] provides the basis for analysis of algorithmic self-assembly in an ideal case. A *tile set* consists of a finite set of unique *tiles* that is used to self-assemble into a DNA *crystal*. A tile is assumed to be square and tiles can not rotate by assumption. Each of the four sides of a tile has a *bond type*. The bond types of a tile determines the uniqueness of the tile. Each bond type has an associated *bond strength*. Bond types can be null (strength of 0), single (strength of 1), or double (strength of 2). Two bonds of the same type can glue (i.e. bond) together, with a corresponding strength. It is assumed that the strength between different bond types is always 0. In an ideal process, self-assembly always begins with a *seed tile*. A tile can be added to the existing crystal when its total bond strength to the crystal is greater than or equal to 2. The crystal generated from the seed tile via a series of legal (correct) tile additions is referred to as the *produced assembly*. The *integer* in the tile denotes the *bond type*. In practice, a *mismatch* may occur during this process, such that a tile with a total bound less than 2 is attached. Additionally, a tile can also fall off from a crystal.

The *kinetic Tile Assembly Model (kTAM)* [2] provides a framework to analyze and simulate the non-ideal self-assembly. The kTAM model includes rates for both *association and dis-association* of tiles from the crystal. In this model, it is assumed that the on-rate (association) r_{on} is determined only by the tile concentration, that is determined by the parameter G_{mc}. The off-rate r_{off} (dis-association) is determined by the total bound strength b that holds the tile to the crystal and the parameter G_{se}. These rates are given as follows: $r_{on} = k \times e^{-G_{mc}}$ and $r_{off,b} = k \times e^{-bG_{se}}$, where k is a constant, G_{mc} is the physical parameter measuring the tile concentration, while G_{se} is the physical parameter measuring the unit bond strength.

This paper specifically addresses the detection and correction of *growth errors*, which are defined as weakly-bonded tile attachments at a location where another tile could and should attach [3]. Growth errors occur if there exist at least two incorrect (erroneous) attachments upon the same four-sided tile. The error correction technique presented in this paper is referred to as the *Error Isolation Tile method*. *Border tiles* are defined as any tile on an edge of an assembly, in which growth has ceased. In the proposed model, border tiles assemble at the same rate as the entire assembly, thus a tile mismatch is allowed to *propagate to an edge*, and causes a *disjointed line* to occur in the border tiles, i.e., the initial error results in a disjointed line in the border tiles, that should otherwise be straight. At the break in the border line using the proposed method, a so-called *Error Isolation Tile* attaches and thus, it effectively *tags* that section of the assembly for removal. A growth error is illustrated in Figure 1; the disjoint line (as effect of this error) is evident.

Consider the error-free assembly shown in Figure 2a, the subsequent attachment along the southern border of an Error Isolation Tile that detects the error is shown in Figure 2b. Although a set of Error Isolation Tiles that attach to every possible defective structure of two adjacent tiles could be constructed, the number of tiles in such an Error Isolation Tile set would be prohibitively large, thus very inefficient in terms of overhead in the final assembly. In addition to a tile

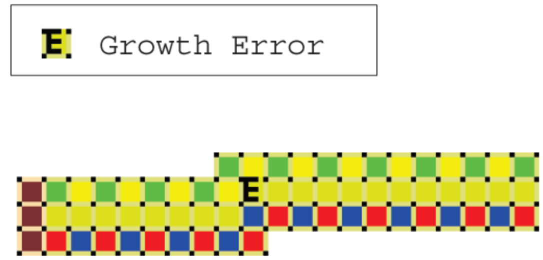


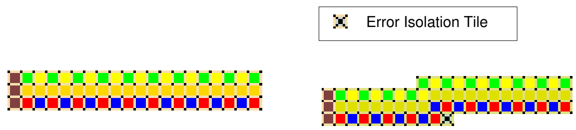
Figure 1. Initial Growth Error.

mismatch, it is possible that a growth error (as denoted in Figure 1 by an E) is caused by weak bonds. The bond strength must not allow for shifting a subsequent column up or down by one row; if this occurs, then it is considered a growth error. This is classified as a growth error (even though no mismatch has occurred), because the rules of the bond strength must cause the tiles to fall off prior to other tiles to subsequently attach and increase the adjacent bonds for a permanent attachment to the assembly.

The proposed approach requires only two additional tiles, which are referred to as the Error Isolation Tile set. The rectilinear growth of the assembly results in the need for two tiles in the tile set, as opposed to just one tile. The two types of Error Isolation Tiles are illustrated as error tags in Figure 3a. (1) The first Error Isolation Tile attaches to the two disjoint border tiles from the southeast. (2) The second Error Isolation Tile attaches from the southwest. Errors are detected when an Error Isolation Tile attaches to the assembly at the location of two disjoint and adjacent border tiles, which can only be adjacent in error. Any growth error results in a disjoint line at the southern border tiles, so only two Error Isolation Tiles are required to detect a large number of possible growth errors.

3. Error Detection and Correction

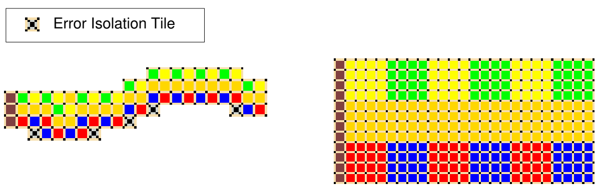
Border tiles are employed in most tile sets for algorithmic self-assembly [5]. In most tile sets these tiles assembly prior to the interior tiles. Thus, propagation of an error within the interior tiles will be blocked by the border tiles and prevented from reaching an edge of the assembly. In the proposed model, border tiles in the tile set may only assemble if an interior tile next to the border is assembled. The growth error in the interior will cause a break in the straight line of the border tiles, as shown in the middle of Figure 1 by the disjoint line due to E as growth error. The border tiles make up the disjoint bottom row. The set employed to demonstrate the Error Isolation Tile correction technique is a modified version of the linear tile set pre-



a) Error-Free Assembly

b) Assembly with One Error Isolation Tile

Figure 2. Example of Two Assemblies.



a) Assembly with Five Error Isolation Tiles

b) Assembly via 4×4 Snake Proofreading Tile Set

Figure 3. Example of Two Assemblies.

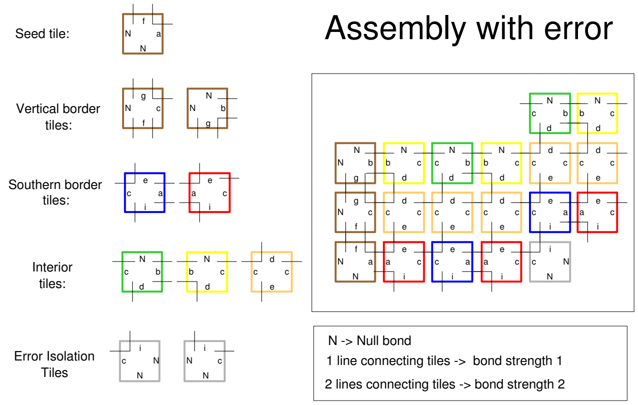


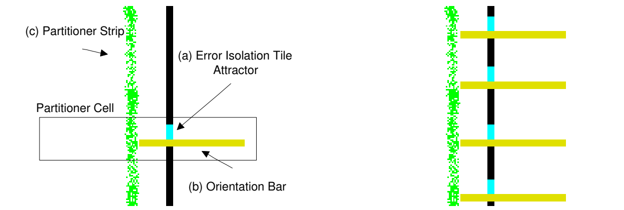
Figure 4. $3 \times Y$ **Rectangular Tile Set, with Error Isolation Tiles.**

sented in [1]. The tile set employed through the remainder of this paper is the $3 \times Y$ rectangular tile set detailed in Figure 4, and illustrated in an error free assembly in Figure 2a. This tile set has a potential application as a data bus with one data line, and could be modified to have any number of data lines. The tile set defined in Figure 4 does not contain an input or output, but it self-assembles into a defined pattern. The error correction method proposed in this paper can be applied to a modified version of any tile set that exhibits algorithmic self-assembly, such as the Sierpinski triangle or the binary counter tile sets of [5]. Hence, all growth errors in which no four-sided tile is part of two or more errors, will cause a disjoint line in the southern border tiles, and an Error Isolation Tile will attach to those two disjoint border tiles.

4. Removal of Defective Sections

Previous works have presented theoretical models and tiles sets in which a damaged section of an assembly can heal perfectly, i.e., the damaged section of the assembly is removed. The use of a particular tile (with conductive cargo or a tag) for detecting an error, has been proposed in [9]; in that work, the use of magnetic attraction has been proposed for removing a defective section of an assembly [8]. The previous section has shown that the Error Isolation Tile correction method detects most growth errors in the $3 \times Y$ rectangular tile set (as defined in Figure 4). The method to remove only a specific portion of the assembly (that is tagged for removal by an Error Isolation Tile) is detailed in this section. The proposed error correction technique can be employed by any tile set that exhibits rectilinear growth, i.e., growth that is limited to only two specified directions. Most existing tile sets exhibit rectilinear growth, and the proposed technique can also be extended to these sets. The proposed Error Isolation Tile method uses *checkpoint intervals* in which all growth activity ceases, and defective sections of the assembly are removed.

In the proposed method, all Error Isolation Tiles hold a metal, i.e., *a conductive cargo*. Thus, the assemblies with an attached Error Isolation Tile can be attracted to a charged metal and a section of the assembly (within a given radius of the Error Isolation Tile) can be raised in temperature, thus causing bonds to break and a partition of the assembly to occur. This partition divides the assembly into error-free and erroneous (defective) sections. This is accomplished as follows: (1) The device that removes the defective section of the assembly is referred to as the *Partitioner Cell*, and is shown in Figure 5a. An array of partitioner cells is illustrated in Figure 5b. The Error Isolation Tiles are attached to the charged metal strip, known as the Error Isolation Tile *Attractor*, this is shown as (a) in Figure 5a. The portion of the assembly to be removed contains the initial error and all tiles to the right of the error, so the assembly attaches to the Partitioner Cell with the Error Isolation Tile on the south side. (2) To ensure a proper orientation of the assembly, a 3-D *Orientation Bar* (shown as (b) in Figure 5a) is em-



a) Single Empty Partitioner Cell

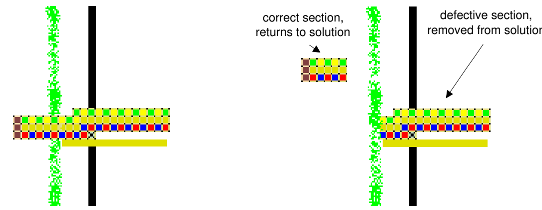
b) Array of Partitioner Cells

Figure 5. Partition Cells.

ployed to prevent the assembly from attaching to the Partitioner Cell at an incorrect orientation. A proper orientation is required, such that only tiles near and to the east of the error are removed. (3) The *Partitioner Strip* (shown as (c) in Figure 5a) is a 2-D strip of metal that heats the local portion of the assembly directly above it, thus causing the bonds (also directly above it) to break. When bonds of the assembly located above the Partitioner Strip are broken, due to the west to east rectilinear growth that the $3 \times Y$ rectangular tile set exhibits, every tile to the east of the error will be removed. Therefore after the checkpointing stage, all tiles that subsequently attach after the initial growth error are removed from the assembly.

Figure 6a illustrates an occupied Partitioner Cell. After an assembly attaches to a Partitioner Cell and the portion of the assembly (that is located above the Partitioner Strip) breaks, the assembly is partitioned into error-free and erroneous sections, as illustrated in Figure 6b. The erroneous section of the assembly remains connected to the charged strip, and the error-free sections of the assembly return to the solution. The array of Partitioner Cells continues to fill with defective portions of the assembly, so there must be sufficient arrays of Partition Cells for all assemblies with Error Isolation Tiles attached. Arrays of Partitioner Cells are always immersed in the solution; they attract only erroneous assemblies when the self-assembly is in a checkpointing state. The proposed checkpoint process is therefore complete, and it may be repeated.

The use of error detection/correction via checkpointing and Error Isolation Tiles can be extended for use with assemblies other than rectangles. Through an alteration of the orientation of the 3-D and 2-D metal lines (i.e., (b) and (c) in Figure 5a), the Partitioner Cell can partition triangular or circular assemblies, therefore the Partitioner Cell is not limited to the partitioning of exclusively rectangular assemblies. Every tile within a specified radius and to the east of the defective section can be removed (e.g., Figure 6).



a) Occupied Partitioner Cell

b) A Tile Assembly is Successfully Partitioned into Error-free and Erroneous Sections

Figure 6. Example of an Occupied Partitioner Cell Before and After Partitioning an Erroneous Tile Assembly.

References

[1] E. Winfree et al.. The xgrow simulator. [Online]. Available: www.dna.caltech.edu/Xgrow/

[2] E. Winfree, "Simulations of Computing by Self-Assembly," Tech. Report CS-TR:1998.22, Caltech, 1998.

[3] E. Winfree, *Self-Healing Tile Sets*, Foundations of Nanoscience: Self-Assembled Architectures and Devices, 2005, pp. 21-22.

[4] P. Rothemund, N. Papadakis, and E. Winfree "Algorithmic Self-Assembly of DNA Sierpinski Triangles," *PLoS Biology*, 2 (12). pp. 2041-2053. ISSN 1544-9173 2004.

[5] M. Cook, PWK Rothemund, and E. Winfree, "Self-Assembled Circuit Patterns," *DNA Based Computers 9*, 2004, pp. 91-107.

[6] Y. Baryshnikov, E.G. Coffman, N. Seeman and B. Yimwadsana, "Self-Correcting Self-Assembly: Growth Models and the Hammersley Process," *Proc. of the 11th Int'l Meeting on DNA Computing*, London, Ontario, 2005.

[7] J.H Reif, S. Sahu, and P. Yin, "Compact Error-Resilient Computational DNA Tiling Assemblies," *10th Int'l Meeting on DNA Based Computers*, Lecture Notes in Computer Science, Springer-Verlag: New York, 2004.

[8] H. Yan, X. Zhang, Z. Shen, and N.C. Seeman, "A Robust DNA Mechanical Device Controlled by Hybridization Topology," *Nature*, vol. 415, pp. 62-65, 2002.

[9] A. Carbone and N.C. Seeman, "Circuits and Programmable Self-Assembling DNA Structures," *Proc. of National Academy of Science, USA* 99:12577-12582, Sept. 13, 2002. [Online]. Available: www.pnas.org/cgi/content/full/99/20/12577