University of Arkansas, Fayetteville
ScholarWorks@UARK

Theses and Dissertations

5-2013

Directed Percolation and the Abstract Tile Assembly Model

Tyler Garrett Moore
University of Arkansas, Fayetteville

Follow this and additional works at: http://scholarworks.uark.edu/etd

 Part of the Nanoscience and Nanotechnology Commons, Numerical Analysis and Scientific Computing Commons, and the Statistical, Nonlinear, and Soft Matter Physics Commons

Recommended Citation
Moore, Tyler Garrett, "Directed Percolation and the Abstract Tile Assembly Model" (2013). *Theses and Dissertations*. 765.
http://scholarworks.uark.edu/etd/765

This Thesis is brought to you for free and open access by ScholarWorks@UARK. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of ScholarWorks@UARK. For more information, please contact scholar@uark.edu, ccmiddle@uark.edu.

DIRECTED PERCOLATION AND THE ABSTRACT TILE ASSEMBLY MODEL

DIRECTED PERCOLATION AND THE ABSTRACT TILE ASSEMBLY MODEL

A thesis submitted in partial
fulfillment of the requirements for the degree of
Master of Science in Computer Science

By

Tyler Moore
University of Arkansas
Bachelor of Science in Computer Science, 2011

May 2013
University of Arkansas

ABSTRACT

Self-assembly is a process by which simple components build complex structures through local interactions. Directed percolation is a statistical physical model for describing competitive spreading processes on lattices. The author describes an algorithm which can transform a tile assembly system in the abstract Tile Assembly Model into a directed percolation problem, and then shows simulations of the aTAM which support this algorithm. The author also investigates two new constructs designed for Erik Winfree's abstract Tile Assembly Model called the NULL tile and temperature 1.5. These constructs aid the translation between self-assembly and directed percolation and may assist self-assembly researchers in designing tilesets in the aTAM with non-deterministic local properties, but guaranteed global properties. Temperature 1.5 results indicate the brittleness of the standard temperature 2 tile assembly system, and the NULL tile is shown to assist simulations of large assembly processes while also reinforcing the need for variable temperature models to more closely simulate laboratory self-assembly.

This thesis is approved for recommendation
to the Graduate Council.

Thesis Director:

_____

Dr. Russell Deaton

Thesis Committee:

_____

Dr. Gordon Beavers

_____

Dr. Wing Ning Li

THESIS DUPLICATION RELEASE

I hereby authorize the University of Arkansas Libraries to duplicate this thesis when needed for research and/or scholarship.

Agreed _____
Tyler Moore

Refused _____
Tyler Moore

ACKNOWLEDGEMENTS

I thank Dr. Gordon Beavers, Dr. Russell Deaton, and Dr. Wing Ning Li for their insight and support through the years as well as their participation on this committee. I particularly thank Dr. Russell Deaton for his guidance and direction throughout the preparation of this thesis.

I also thank my family for their love and support and especially Alison Brown.

TABLE OF CONTENTS

LIST OF FIGURES

1. INTRODUCTION

1.1 Problem

Self-assembly is the process by which simple components build complex structures through local interactions. Self-assembling systems include atomic systems and molecular formation [1], macromolecular and crystal formation [2], and cellular composition. There is an entity in each case which interacts with other similar entities such that the complexity of the whole system increases and the free energy is minimized. The building blocks of life, proteins, also interact to gain complexity through self-assembly, and because of this continue to be a mainstay in self-assembly research [3]. Investigations into self-assembly through experiments with DNA have resulted in complex DNA structures like bipedal walkers [4] in the lab and theoretical advancements like Erik Winfree's abstract Tile Assembly Model [5].

One of the simplest and most important characteristics of self-assembly lies in the nature of the construction. The components assemble without supervision by a global agent. The nature of self-assembly, its constrained yet globally undirected construction of the target structure, makes it an exceedingly useful tool for myriad disciplines, particularly nanotechnology. Self-assembly is used to construct nanoscale materials because it provides nanoscientists with a bottom-up design process; the structure assembled is implicitly designed by specifying only the components and their binding domains. By definition it disallows an ordering of components by an intelligent processor into predetermined schema. This circumvents the need for a traditional top-down design workflow, something difficult to implement at the nanoscale. Erik Winfree's abstract Tile Assembly Model, or aTAM, provides an invaluable tool for specifying the self-assembling components in this bottom-up design process. Instead of gold nanoparticles, DNA, or

colloidal crystals, the aTAM manipulates tiles: four-sided and rigid structures that cannot be rotated or reflected. Winfree's abstraction borrows heavily from Wang tiles, first introduced in [28]. There are a number of key changes; for example, binding domains of tiles are described by the glues on their four sides and are articulated by the "temperature" of the system, a construct Wang tiles do not consider.

Self-assembly is well-described under Winfree's aTAM and derived models, such as the kinetic Tile Assembly Model, but these models can be expanded further to address additional constraints and concerns [5]. By intuition it seems that self-assembly is a less constrained process than the traditional aTAM can account for and can be better understood by adapting existing statistical mechanical models to the tile assembly process. This is also evident in the laboratory, as experiments infrequently correspond exactly to the predictions of the aTAM [15]. This thesis specifically considers percolation theory, the study of the number and properties of clusters of occupied sites on graphs [6]. Additional constructs are also considered which can aid the assembly process to more accurately reflect *in vitro* self-assembly. The first is the addition of a static tile type to each tile assembly system called the NULL tile whose purpose is to account for vacancies in the tile assembly. The second is the introduction of an unstable temperature value into the aTAM called here temperature 1.5 which gives a probabilistic component to the aTAM.

1.2 Objective

Using simulations, this thesis will demonstrate the equivalence of directed site percolation on the square lattice and self-assembly. Also, the author will examine the usefulness

of two new constructs which could more accurately represent laboratory self-assembly by simulation within the aTAM.

1.3 Approach

The connection between percolation theory and the aTAM is first made by transforming the configuration of tiles in an assembly into a lattice where sites that contain tiles are considered occupied and where sites that do not contain tiles (or contain incorrect tiles) are considered unoccupied. The percolation system introduces probabilistic behavior seen in laboratory experiments while also reinforcing the power of the aTAM. It is by this transformation of the aTAM and the addition of probabilistic behavior that the model is provided more flexibility without losing its broad application in the study of self-assembly. Percolation theory studies dynamical systems which can be expressed on finite or infinite graphs. There are a suite of equations that express expected behaviors at infinite range which can be used to determine the probability of configuring very large assemblies. Experimental results on finite graphs can also be used to interpret the step-by-step process of current laboratory self-assembly.

A cluster is a percolating cluster and the lattice it is on is said to percolate if that cluster "connects the top line or plane with the bottom line or plane [6]." Imagine that a sponge or other porous material is held under a slow drip of water. If the water permeates the material, that is if it penetrates the material fully, then there is a path from the top of the material to the bottom through which water can travel. That path belongs to a percolating cluster of sites within the material and the material can be modeled by a lattice that percolates. A percolating cluster is outlined in figure 1. Note that not all occupied sites on the lattice are a part of the percolating

cluster, only those sites reachable from the seed site at the top of the lattice. This is a result of the directed edges in the graph.

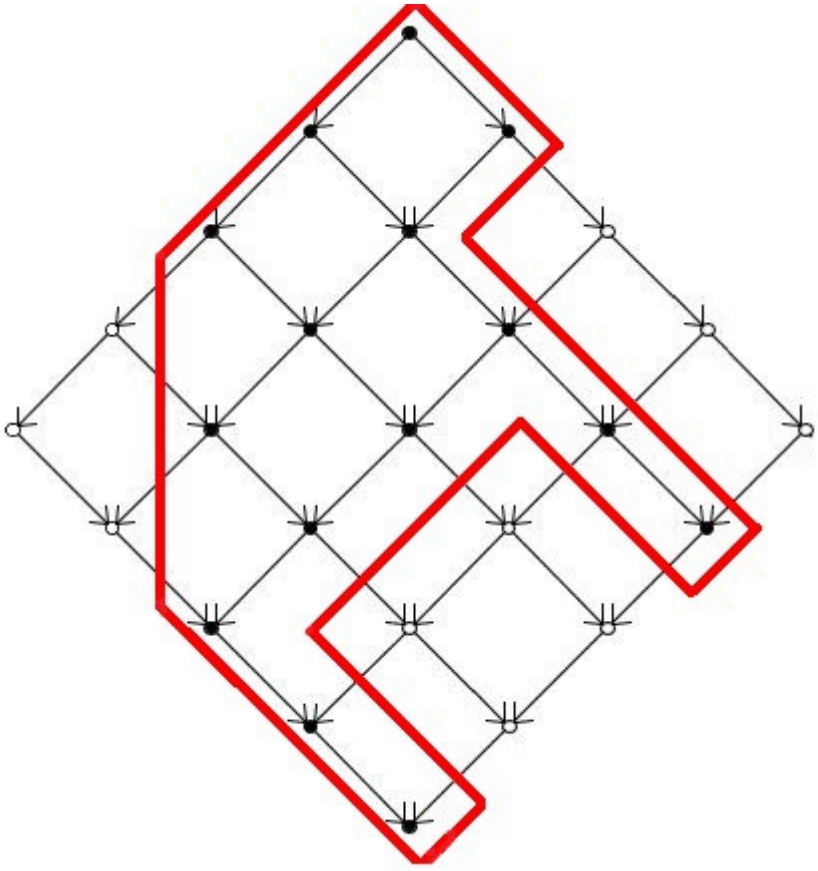

Figure 1: A percolating cluster on a tilted square lattice

By transforming a tile assembly problem into a percolation problem, laboratory results can be examined in a new light. A system that fails to configure appropriately can be considered through terms which are well-described within percolation theory. An assembly under non-ideal conditions can undergo a phase transition as those conditions approach the ideal; these conditions affect attachment attachment probabilities which are instrumental in reaching the final absorbing state. A state is considered **absorbing** if no further action can occur once the system has reached that state [6]. To avoid undesired absorbing states, the aTAM disregards many

interactions among tiles and aggregates which occur in laboratory self-assembly. These undesirable interactions include unexpected tile attachments (tile *a* is expected, but tile *b* attaches for some reason) and uneven attachment rates (assuming a planar construction, tile *a* is expected to attach at site (*i*, *j*), but, before it can, tiles attach at surrounding sites and prevent tile *a* from approaching its binding site). Such interactions short circuit the attachment process or prevent correct assembly, leaving the aggregate such that the system cannot finish the designed assembly. In errorful assemblies, i.e. assemblies with a preponderance of error leading to an absorbing state, the conditions and tile set are such that the state which consists of only the designed assembly can no longer be reached from the current state. Such an aggregate can be said to be in an "absorbing state". No further interactions with tiles in the system can occur, and the aggregate is not in the expected final configuration. This absorbing state is a direct result of tile attachment order and probability. Local, probabilistic choices made during assembly from attachment to attachment can affect the global outcome of the aggregate. Percolation theory shed s new light on self-assembled systems; instead of "Does this tileset assemble the desired configuration or does it form an undesired, intermediate aggregate?", the question becomes "Does this system percolate (assemble the designed aggregate) or does it reach an undesired absorbing state (an errorful assembly)?" This question is at the core of percolation on finite lattices.

Self-assembly can construct both passive (inert materials) and active (walkers and hinged-lid boxes) structures at the nano, micro, meso, and macro scales, but current design principles for self-assembling systems are lacking. Laboratory researchers focus on advancement through trial and error, allowing the process to develop naturally and without constraint, but the

discipline requires a more focused and developed design workflow to continue its relevance. By translating well-described qualities and behaviors inherent in percolating lattices into self-assembly, a whole new wealth of tools are available to researchers. These tools include descriptive equations of a comparable dynamical system and alternate dynamical models of attachment and growth. Through these tools, more flexible work-flows can be established which will allow researchers to construct both more reliably and also more efficiently at any scale. By better describing self-assembly, a wide array of disciplines are positively affected, including nanotechnology, medicine, chemistry and physics.

In this paper the author will describe a number of new avenues of investigation and initial results in applying percolation theory to the aTAM and the process of self-assembly. Multiple tilesets were found or created that exhibited complex but predictable finite and infinite properties. Finite properties were verified by hand and infinite properties were extrapolated from large finite simulations. These tilesets were assembled within a simulation tool that allows temperature 2 attachments, temperature 1 attachments, and a mix of the two called by the author temperature 1.5. Local and global behaviors were studied at different system temperatures on all tilesets. In addition, a NULL tile type was introduced into the assembly process of each tileset and the results compared to equivalent directed percolation systems. Particular attention is paid to temperature 1.5 systems which include the NULL tile. Tile assembly systems operating at temperature 1.5 with NULL tiles are shown to be nearly identical to the directed site percolation problem.

1.4 Organization of this Thesis

Chapter 2 covers key materials necessary for the author's investigations of self-assembly, the aTAM, and directed percolation. Section 2.1 relates core concepts vital to the objective of this thesis, particularly percolation and self-assembly through the aTAM. Section 2.2 investigates relevant recent works in the fields of self-assembly and percolation.

Chapter 3 presents the author's methodology and introduces the key constructs necessary to evaluate this methodology, including the tilesets and new concepts utilized, definitions of measures used to assess these constructs, and a formal algorithm which translates a tile assembly system into a system undergoing directed percolation. A brief summary of the author's investigative method is given in section 3.3.

Chapter 4 relates the results of experimentation on each tileset with the NULL tile and temperature 1.5. Directed percolation simulations are described and approximate critical probabilities discussed for each tileset. In section 4.6, these results are analyzed with respect to each other, and conclusions are reached regarding the properties of the tilesets used and how those properties interacted with the new constructs introduced in this thesis.

Chapter 5 resolves the author's analysis into a conclusive statement regarding the similarity of self-assembly by the aTAM to directed percolation. In addition, the author discusses usefulness of the NULL tile and temperature 1.5 in the aTAM, and also the role that statistical mechanical models like directed percolation can play in the continued development of design workflows and applications of assembled structures. In section 5.2, some open questions are related by the author concerning this thesis, and possible avenues of investigation are discussed.

2. BACKGROUND

2.1 Key Concepts

The three underlying concepts which define the foundations of this research are self-assembly, the abstract Tile Assembly Model, and percolation theory. Self-assembly and the aTAM are described first, after which percolation theory is introduced.

2.1.1 Self-assembly and the aTAM

Winfree describes his abstract Tile Assembly Model as follows [5]:

> The fundamental units in this model are unit square **tiles** (also called monomers) with labeled edges. We have an unlimited supply of each type. **Aggregates** are formed by placing new tiles next to and aligned with existing ones such that sufficiently many of their edges have matching labels. Tiles cannot be rotated or reflected. […] Each edge label σ_i has an associated strength g_i, which must be a non-negative integer. At **temperature** T, an aggregate of tiles can grow by addition of a monomer whenever the summed strength of matching edges exceeds T […] these are called stable additions.

More formally, by Soloveichik and Winfree in [7], let Σ be a set of bond types and let *t*, a tile type, be a 4-tuple $(\sigma_N, \sigma_E, \sigma_S, \sigma_W) \in \Sigma^4$. A *null* bond type represents a null interaction. If *t* has 4 *null* bond types, *t* is called the NULL tile type. A tile type cannot be rotated or reflected because such a transformation disrupts the ordering of bond types which define that tile type. If *T* is a set of tile types, then a tile is an ordered pair $(t, (i, j)) \in T \times Z^2$ which belongs to tile type *t* and is at location (i, j) in the configuration. A configuration *C* is a $Z^+ \times Z^+$ space which contains tiles at each location (i, j). A strength function $g: \Sigma \times \Sigma \rightarrow Z$, where $null \in Z$, defines the interactions between tiles orthogonally adjacent in the configuration. *g* must be symmetric, non-negative, and diagonal. If *g* is symmetric, then $g(t, u) = g(u, t)$ where *t*, *u* are tiles. For *g* to be non-negative, there can be no adverse reactions between tiles; that is, there can be no negative bond strengths

between tiles. If g is diagonal, then only sides with matching bond types interact. Finally, $g(t, null) = 0$, that is, a tile cannot bind with nothing.

In addition to Winfree's terminology[1] there are a number of terms concerning self-assembly which should be introduced less formally. A **configuration**, **structure**, or **assembly** is the resultant of the self-assembly process. A **tile** or **component** can **configure**, **organize**, or **assemble** into a final configuration or assembly.[2] Tiles that interact with other tiles or aggregates such that stable additions are made are said to **attach** or **accrete** to the aggregate. Winfree's aTAM can support any number of dimensions, but two-dimensional tile assembly is most relevant here. Further discussion will assume such planar configurations. Each tile type has a **binding domain**, a unique set of labels, or **glues**, which describe the matching labels (glues) to which its faces can bind. A tile type is identified by its binding domains and these domains must be unique among tile types in the tileset. Unbound tiles, i.e. tiles that are not a part of an aggregate, seek to bind to open binding sites on other tiles and aggregates. If a tile t has a glue g on its north face, another tile u with glue g' on the south face can attempt to bind on that glue. For simplicity, the author assumes the convention that g characterizes it's own binding domain, or $g = g'$. Although usually represented by unique single character names, glues can also be characterized by color or other unique identifiers. Here is an example of tile t binding with an aggregate.

---

[1]The author uses Winfree's term **aggregate** to describe any collection of bound tiles that are not in the expected final arrangement.

[2]Although self-assembly shares commonalities with self-organization, the author does not wish to conflate the two; the sharing of terminology between self-assembly and self-organization is an unfortunate circumstance. To avoid confusion, the author endeavors to use the terms assembly and component when referring to the general process of self-assembly, and the terms configuration and tile when referring to self-assembly by the aTAM.

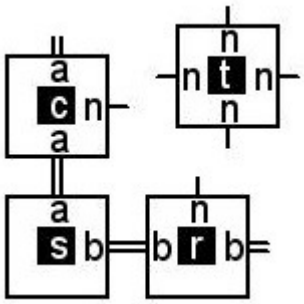

Figure 2: A tile *t* attempts to attach to a configuration in the aTAM in a temperature 2 Tile Assembly System. Tiles have names in white and four glues in black, one to each side of the tile. Strength 1 glues are represented by single connectors and strength 2 glues are represented by double connectors.

Note the binding domain of *t*. It seeks to bind its south face on glue *n* and its west face on glue *n*. The open binding site has such a binding domain. If we assume that the glues involved in forming the bond have a sum strength equal to the temperature of the system, a stable addition to configuration *C* is made. Otherwise, no addition is made. A tile assembly system also requires a seed, an abstract catalyst to which tiles from the tileset bind until some final aggregate or the expected configuration is formed. In figure 2, we can consider the aggregate consisting of tiles *c*, *s*, and *r* the seed of the assembly. Informally, the seed, tileset, and temperature of the system are the input into some black box which outputs the expected final configuration. This black box can be formally represented by the ordered triple *T* called a tile assembly system. $T = (\mathbf{T}, \sigma, \tau)$ where **T** is the tileset (a finite set of tile types), σ is the seed tile or aggregate, and τ is the temperature of the system [5]. Certain assemblies exhibit a property called local determinism. This property is dependent on the glues and glue strengths, tiles, and the temperature of the system. If these conditions are such that for each possible binding domain there exists one or fewer tiles, then the assembly is said to be locally deterministic.

2.1.2 Percolation

Percolation is can model many different spreading processes, including forest fires, dispersion of fluid through rock, and epidemics [6], [7]. Informally, percolation can be exemplified as the study of fluid flow in porous media. In particular, as water is poured on a stone, it will pass into cracks, fissures, and pores on the stones surface. If the stone is very porous or fractured, the water will penetrate the material and flow through the stone. If the material stone is very dense and tightly packed, water will simply flow over and around it. In this way, many percolation problems can be visualized.

Fundamentally, percolation theory is the study of clusters on special graphs called lattices. A lattice is an array of points with even spacing which exhibit translational symmetry. In the case of the square lattice, orthogonally adjacent points in the lattice, called **sites**, are connected by simple edges, called **bonds**. Sites can be either **occupied** or **unoccupied** and bonds can be either **open** or **closed**. Imagine the porous stone from the previous paragraph. It is riddled with nooks and crannies, cracks and holes through which water can flow. A pore in the rock (called a site in percolation theory) is either occupied with water or unoccupied. If that water can flow from one pore into another, deeper in the stone, then their is a bond between these pores that is open. If no such connection between these pores exists, then the bond is closed. A **cluster** is a group of occupied sites connected by open bonds. The very porous stone allows water to pass nearly unimpeded; therefore it has many clusters. An impermeable stone has no such clusters because there are no pores in the stone for water to occupy, and there are no fissures or holes for water to flow through.

The minimum cluster size is one, and the maximum cluster size is the size of the lattice. Although percolation theory can be applied to a wide variety of graphs, the two-dimensional square lattice is most relevant to the aTAM. In the square lattice, each site is adjacent to at most four neighbors, one each to the north, south, east, and west. If the lattice is finite, sites along the border and at the corners of the lattice may have as few as two neighbors. Other lattices considered in percolation theory, but not considered here include the honeycomb, triangular, diamond, simple cubic, body centered cubic, face centered cubic, and high dimensional hypercubic lattices. Only the honeycomb (hexagonal), triangular, and square lattices are two-dimensional.

In the simplest percolation problem, sites and bonds are occupied or open independent of any other site or bond in the lattice. That is, if site A is occupied, the probabilities that sites B, C, and D are occupied remain unaffected. There are three percolation problems which can be considered on a particular lattice, site percolation, bond percolation, and site-bond percolation. In site percolation, all bonds are considered open. This constraint ensures that any adjacent occupied sites form a cluster, since a cluster is defined as a collection of adjacent occupied sites connected by open bonds. Site percolation is **random** if each site is occupied with probability p and unoccupied with probability $(1 - p)$ and is found occupied or unoccupied independently of the status of other sites in the lattice. A cluster on a lattice undergoing random site percolation consists of a group of adjacent occupied sites. In bond percolation, all sites are considered occupied. Bond percolation is random if each bond is open with probability q and closed with probability $(1 - q)$ and if each bond is found open or closed independently of other bonds in the lattice. A cluster on a lattice undergoing random bond percolation consists of a group of sites

adjacent by open bonds. A special hybrid of site percolation and bond percolation, conveniently called site-bond percolation, is also important to consider. In site-bond percolation, sites are occupied with probability p and unoccupied with probability $(1 - p)$ and bonds are open with probability q and closed with probability $(1 - q)$. In site-bond percolation, a cluster is a group of occupied sites adjacent by open bonds. One property of percolation under intense scrutiny is the behavior of a percolation system at some critical probability p_c. For site percolation, if $p < p_c$, then no percolating, also called spanning, cluster can form. If $p > p_c$, then a percolating, or spanning, cluster is guaranteed to form. This critical probability is said to mark a phase transition because the system behaves in a fundamentally different way when the probability of a particular bond or site being open or occupied is below the critical probability compared to when the probability of open bonds or occupied sites is above the critical probability. For some infinite systems with low dimensionality, this p_c can be found exactly within some tolerance, but for high dimensional systems, only inexact estimates gathered from the analysis of mathematical models and finite experimental systems exist. On a finite square lattice, a percolating cluster is a cluster which has at least one site on opposite edges of the lattice. An example of a such a cluster can be found outlined in figure 1.

Self-assembly by the aTAM is inherently dependent on the open binding sites on an aggregate and is described by seeded growth. Directed percolation is a constrained percolation problem which can relate the aTAM to percolation theory. In directed percolation, bonds are given a direction which disallows interactions with “previous” sites on the lattice. This introduces anisotropy into the formerly isotropic percolation problem. Clusters are now defined as being a group of occupied sites connected by open, directed bonds. In addition, directed

percolation is seeded. This means that some initial collection of sites are first marked occupied and then the system is allowed to change state through some probability of growth or expansion.

Directed percolation excels at modeling systems that are characterized by two competing processes. In the epidemic system mentioned earlier, each site can be infected or healthy. Spreading is controlled by some probability p which describes a rate of infection. Healthy sites adjacent by a directed bond to infected sites can become infected with probability p and infected sites can spontaneously recover with probability $(1 - p)$. If $p < p_c$, then the system enters an absorbing state where the infection fails to spread and ultimately dies out. This state is called an absorbing state because the system cannot spontaneously regenerate the infection.

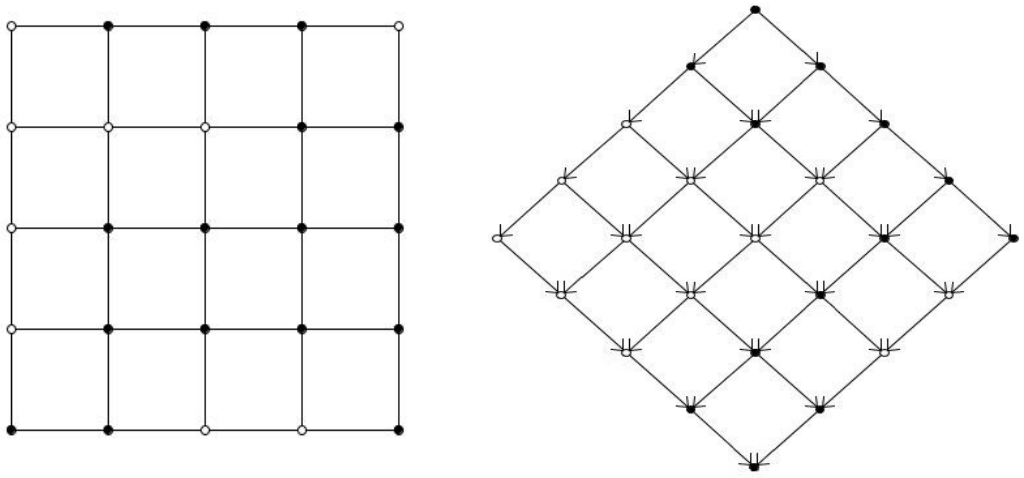

Figure 3: A square lattice undergoing random percolation (left) and a tilted square lattice undergoing directed site percolation (right) where time increases down the lattice

The process of directed percolation can be analyzed layer by layer in a time-dependent formulation where t encodes the current layer, or the boundary of the percolation process, as it advances across the lattice. At $t = 0$, the process begins with a seeded set of sites s_0; at $t = i$ the

leading edge of the system is encoded in the set of sites s_i. To model directed percolation, a tilted two-dimensional square lattice is constructed by taking an undirected square lattice and rotating it 45 degrees before adding directions to all bonds such that no site has a path to any of its ancestors. An example of the transformation from the square lattice to the tilted square lattice is presented in figure 3.

It is interesting to note both self-assembly and percolation can be related through cellular automata. In the case of directed percolation, the Domany-Kinzel cellular automaton is used [8]. The DKCA is a simple stochastic cellular automaton which when composed into a system of automata on a directed tilted square lattice exhibits directed percolation behavior. The probability of spreading is captured by three rules which relate a site and its parents. The child site is occupied with probability 0 if neither parent is occupied, and the child is occupied with probability p_1 if one parent is occupied and the other is unoccupied. The child is occupied with probabilityp_2 if both parents are occupied. These relationships are shown in Table 1, where the child's probability of being marked occupied is in column Child.

Parent 1	Parent 2	Child
Unoccupied	Unoccupied	0
Unoccupied	Occupied	p_1
Occupied	Unoccupied	p_1
Occupied	Occupied	p_2

Table 1: The probabilistic rules of the DKCA

Suppose p represents the probability that a particular site is marked occupied and q represents the probability of a particular bond being open. If $p_1 = p \cdot q$ and $p_2 = p \cdot q(2 - q)$, then a site-bond percolation problem is created [9]. When $p = 1$, then a directed bond percolation problem is

formed and $p_1 = q$ and $p_2 = q(2 - q)$. A directed site percolation problem can be created if $q = 1$, yielding $p_1 = p_2 = p$.

2.2 Related Work

2.2.1 Adaptations to the aTAM

A number of adaptations to the aTAM have been made to better model different aspects of self-assembly. Winfree's kinetic Tile Assembly Model, or kTAM, is one such adaptation [5]. The strict constraints on aggregate growth are done away with in the kTAM and allowances are made which more accurately describe self-assembly in "laboratory conditions," although certain abstractions are introduced to simplify calculations. Winfree's kTAM assumes the following:

1. Tile concentrations are constant and equal.
2. Aggregates cannot interact (bind) with each other.
3. Tiles are equally likely to bind to an aggregate.
4. Tiles can dissociate from an aggregate at a rate exponentially inverse to the strength of its bond to the aggregate.

Winfree also defines 3 system parameters, G_{mc}, G_{se}, and k_f. G_{mc} measures the cost of binding a tile to an aggregate and is set by the experimenter through the concentration of tiles in the system. G_{se} measures the energy cost of breaking a single bond. The final system parameter, k_f, defines the timestep and doesn't affect the behavior of the system. The binding rate of a particular tile to a particular open binding site on a particular aggregate can be written as

$$r_f = k_f * e^{-Gmc}$$

This equation demonstrates that the binding rate is simply the current timestep multiplied by the tile concentration. The equation

$$r_{r,b} = k_f * e^{-b * Gse}$$

describes the likelihood that a particular tile will dissociate from the aggregate at a particular timestep. Although the kTAM is an improvement upon the aTAM for laboratory work, the fundamental tile assembly process remains unchanged and the analysis and conclusions presented here can be specified for either the kTAM or the aTAM.

In addition, work at Duke has yielded a number of generalized models for self-assembly, each with a specific application [10]. Aggarwal et al. have developed a flexible glue model which allows for interactions between non-binding glues, a multiple temperature model which allows the temperature of the system to change during the tile assembly process, a multiple tile model which allows for the accretion of aggregates to other aggregates, and a unique shape model which allows a tileset to assemble any final configuration as long as each configuration has the same shape. Additional work in other research groups has resulted in the probabilistic Tile Assembly Model or pTAM and the Two-handed Tile Assembly Model [11], [12]. In the pTAM, tilesets are designed such that it is likely that the final configuration will be a linear "ruler" of length *N*. Each tile *t* has a companion tile *r* called a reset tile. If *t* can attach to an aggregate at a particular site, so can *r*. Since tile concentrations are held equal, the probability that tile *t* will bind to the aggregate is equivalent to the probability that tile *r* will bind to the aggregate. This models a system which can guarantee that the given tileset will probabilistically assemble rulers of length *N*, even though the number of tiles in the tileset is lower than the

expected lower bound based on the Kolmogorov complexity[3] of the ruler. In order to achieve this curious result, constraints in the aTAM are modified and adapted such that many different rulers are allowed to assemble from a given tileset and the majority are provably of length *N*.

The Two-handed Tile Assembly Model builds on early work on supertiles, or aggregates of tiles, by allowing supertiles to assemble larger configurations without affecting the process by which single tiles assemble. The Two-handed Tile Assembly Model resembles the multiple tile model developed by Aggarwal et al., but Doty et al. in [12] define a formalization which is consistent with other definitions of the aTAM and expands on the simple generalization established at Duke. Doty et al. produce a model which allows for tile assembly systems which produce multiple terminal assemblies without disallowing infinite assemblies and while allowing finite quantities of particular tile types. By allowing attachments between aggregates and abstracting this process such that tile-to-aggregate attachments are still allowed, the Two-handed Tile Assembly Model establishes the usefulness of well-described and stable additions to the existing aTAM.

The pTAM represents a different type of modification to the aTAM. The pTAM seeks to adapt and change existing constraints in the aTAM, while creating processes which capitalize on these adaptations to generate new functionality not present in the simpler aTAM. By allowing

[3] Kolmogorov complexity is the measure of resources needed to describe an object absolutely without unnecessary repetition. In other words, the Kolmogorov complexity of an object represents the absolute minimum number of units needed to represent that object. Rothemund, Winfree, and Soloveichik demonstrated that the Kolmogorov complexity of the shape of an assembled configuration represents a lower bound on the minimal number of distinct tile types necessary to self-assemble it in a series of papers which formalized the constraints on construction of arbitrary shapes via self-assembly [13], [14]. Their conclusions involve the description of a configuration in terms of arbitrarily small units (which could be tiles or supertiles/aggregates), and the resolution that assembly in the aTAM at temperature 2 is Turing Universal. The minimum number of tile types which assemble a single arbitrary shape and nothing else is $\Theta(K/\log K)$ where K is the Kolmogorov complexity of the shape.

the use of tile types which share glues on their operative binding faces (the west face for east-growing assemblies, and the east face for west-growing assemblies), the pTAM seeks to use nondeterministic processes such that conventional tile complexity boundaries can be circumvented. Tile complexity, or the minimum number of tile types needed to assemble a particular configuration, is a costly constraint on assembly. Current DNA assembly techniques require significant investment in synthesizing appropriate materials, with some large samples costing as much as $700 dollars per unit to fabricate uniquely [15]. By reducing tile complexity without introducing unreasonable constraints which would be difficult to replicate in laboratory assembly, the pTAM has adapted the aTAM to fill a different role for self-assembly researchers. Where the aTAM focuses on creating locally deterministic configurations by accreting tiles one by one to an aggregate, the pTAM focuses on creating conditions which yield a particular configuration with a relatively high probability. This flexibility comes at a cost, though. The pTAM is designed to only produce 1-dimensional structures called rulers; no work has been done to expand the results into higher dimensions [11]. Therefore, the pTAM seems primarily useful for researchers performing a **bin** or **staged** assembly. A staged assembly is an assembly which constructs intermediate aggregates in the aTAM by introducing tile types in a measured way [16]. In this fashion, multiple non-interacting rulers can assemble in a single step before introducing connective tiles which would construct rulers that interact with existing rulers in solution. Such adaptations show the worth of investigating the constraints of the aTAM further and adapting new processes and models to describe the aTAM so that future investigations can adapt the aTAM further.

3. METHODOLOGY

3.1 High Level Design

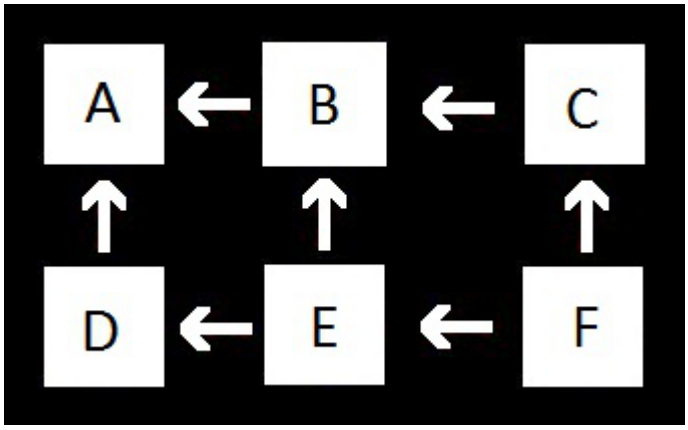

Figure 4: Here is the order of attachment to a seed *A* by a series of tiles. The next tile to be attached occupies the first open binding site in the aggregate where the first open binding site is either directly to the right of the previously attached tile or the first tile on a new row in the aggregate.

In order to model tile assembly, a tile assembly simulation tool designed to mimic the abstract Tile Assembly Model was constructed using Java in the Eclipse IDE. The tool assembles a Configuration, a class that wraps a two-dimensional array of Tiles with additional functionality. Each Tile object has an array of Glues and each Glue has a binding domain and a name which uniquely identifies it. Each Tile is given a north, east, south, and west Glue and is identified in the Configuration by its name and location. A simplified class diagram is shown in figure 5. Assembly begins with a seed row of pre-assembled seed Tiles. This row is prepared outside of the Configuration and then attached explicitly as the seed. Additional seed tiles necessary for assembly, such as seed “column” tiles, are attached explicitly during aggregation.

Each new Tile is attached to the Configuration at the next open site, which is always either the eastern neighbor of the most recent attachment or is a new seed column Tile. Thus, assembly occurs from left to right, top to bottom. See figure 4 for a simple example. The seed Tile is labeled *A* and subsequent tiles *B* through *F* are attached in alphabetic order. Each Tile in the simulation is assumed to be present in the system in an infinite quantity and is selected for attachment greedily. When a Tile attempts to attach to a Configuration, temperature and chance of correct attachment are given as parameters. Since attachment is probabilistic and constrained by the binding domains of the open binding site, situations can arise where no Tile in the tileset is able to attach to the current binding site. In such situations a NULL Tile is attached. The NULL Tile is a Tile that shares no Glues with any Tiles in the tileset and it signifies a vacancy/error in the Configuration. A Tile from the tileset may not be able to attach to the configuration for two reasons: the temperature of the system inhibits the attachment or a NULL Tile is attached with preference over a Tile in the tileset based on some probability determined at attachment time. The Configuration can assemble at temperature $\tau = 1$ or $\tau = 2$. In a tile assembly system where $\tau = 2$, the Configuration can attach a Tile at $\tau = 1$ with some probability provided at attachment time. This behavior, called temperature 1.5, describes a temperature system which may more closely approximate a laboratory environment without losing the aTAM's descriptive yet succinct properties. Note that this does not violate the requirement that the temperature of a tile assembly system must be an integer as the behavior at temperature 1.5 is simply the behavior of a Tile *t* at $\tau = 1$ with some probability *k* and the behavior of a Tile *t* at $\tau = 2$ with some probability $(1 - k)$.

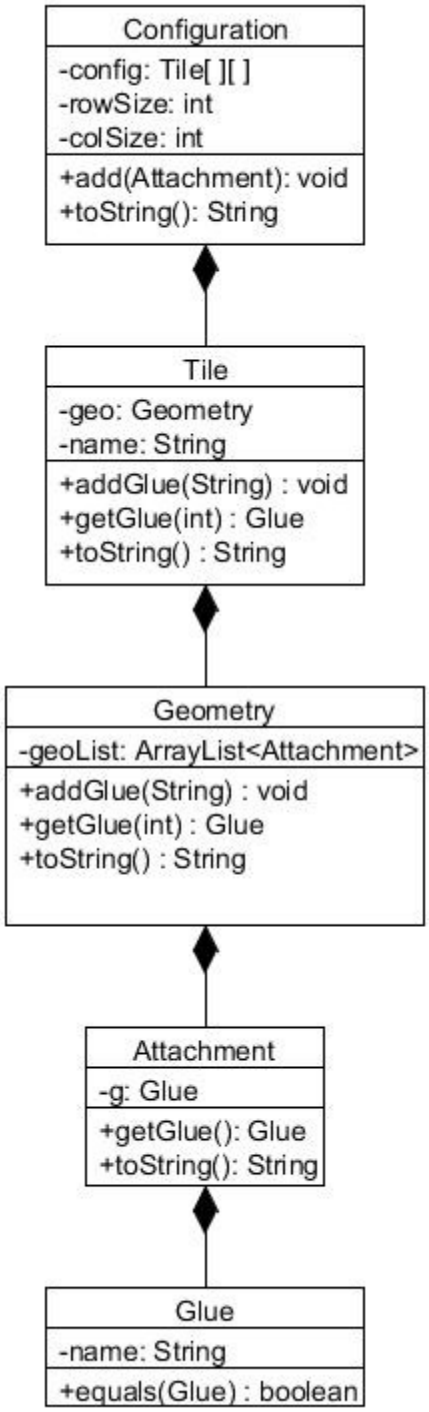

Figure 5: A simplified class diagram of the tile assembly simulation tool

In order to quantify final assemblies, a Correlator was written which takes two n x n Configurations as text files composed of n lines of n tokens each delimited by white space. Each token is a tile name and represents a tile from the tileset which assembled the Configuration. The Correlator parses the files such that for all $0 \leq i, j \leq n$, t_1 at position (i_1, j_1) in Configuration C_1 and t_2 at position (i_2, j_2) in Configuration C_2 are compared by name. If the names are matched, then both a correlation measure and the total number of comparisons are incremented. If the names are not matched, then the total number of comparisons is incremented while the correlation measure remains the same. This creates a correlation percentage which measures the correlation between the tiles in two separate Configurations. Two perfectly correlated Configurations are said to be identical, and two Configurations which have no correlation are said to be distinct. While Configurations which are assembled from distinct tilesets must always be distinct, Configurations which are assembled from the same tileset may not necessarily be identical depending on the parameters given to the tile assembly system. This is partially due to the behavior of the Correlator; Tile neighborhoods are not considered. This means that in temperature 1 or temperature 1.5 tile assembly systems a tile t could attach at location (i_1, j_1) in configuration C_1 and also at location (i_2, j_2) in configuration C_2 even if tiles at locations $(i - 1, j)$ and $(i, j - 1)$ in both C_1 and C_2 do not match. The correlation between a perfect (ideal) temperature 2 assembly and a temperature 1.5 assembly is controlled by the presence of an absorbing phase transition. If the temperature 1.5 assembly sees a preponderance of error during the attachment of new tiles to the aggregate, then it is impossible to continue the assembly of the expected design. In this case, correlation decreases sharply and such an aggregate is said to be in an absorbing phase transition.

In addition, a visualization tool called VisTool was written in Processing, a Java-derivative language created by Casey Reas and Ben Fry. VisTool accepts a text file representing a Configuration, much like the Correlator, and parses it for tile names. Each tile in the tileset is interpreted as a color and that color is rendered to the screen at pixel location (i, j) where (0, 0) is the upper-leftmost pixel. In this way, VisTool translates an n x n Configuration into an n x n image. This image can then be easily analyzed and saved for future comparison or examined for cluster patterns and sizes. VisTool is used to determine whether a Configuration contains clusters of infinite size given that Configuration's tile assembly system. If a Configuration shows a large cluster which spans the image from the upper left corner to the lower right corner, VisTool interprets this cluster as a cluster of infinite size. The primary purpose of VisTool is to relate configurations of tiles assembled with the aTAM to systems undergoing directed percolation processes.

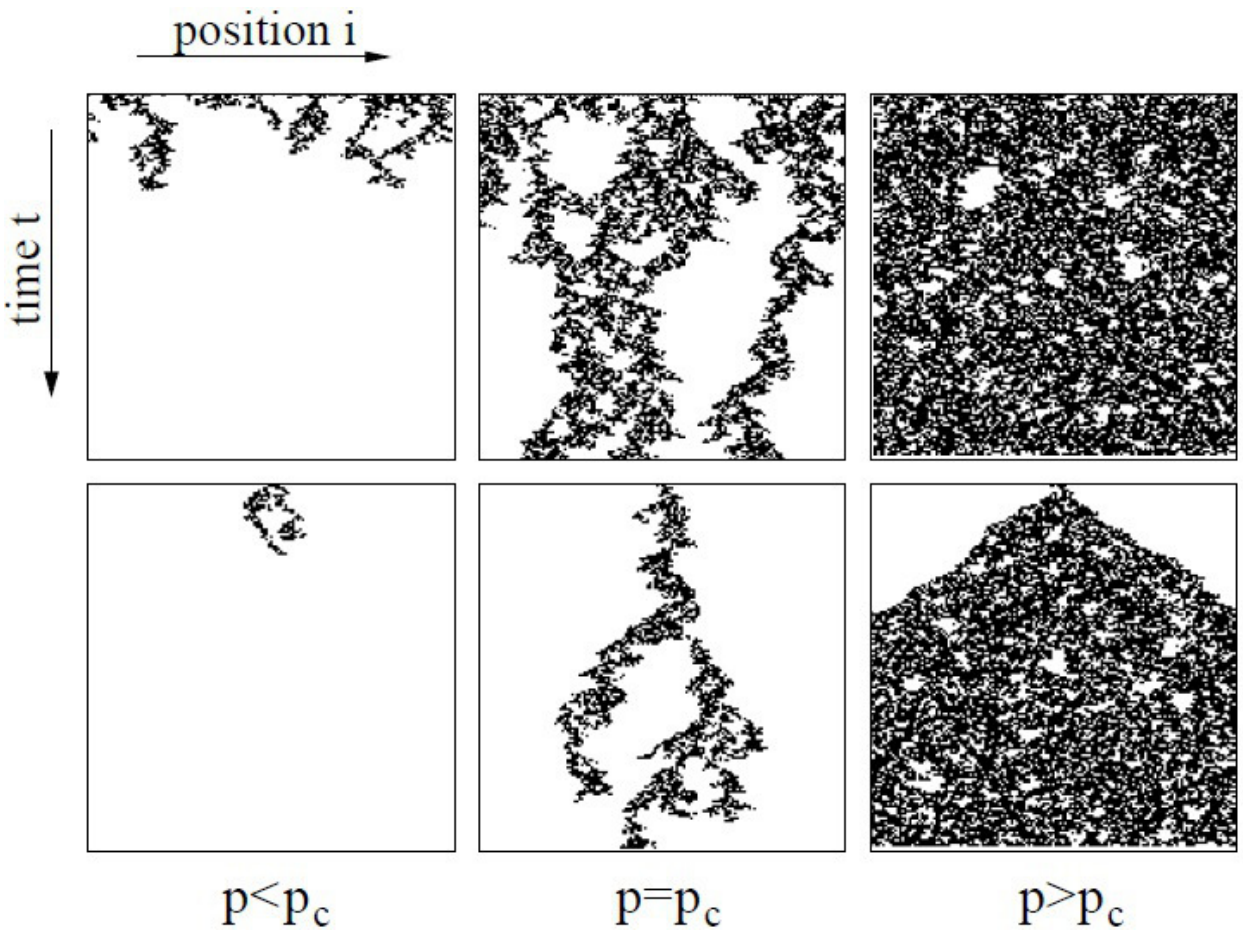

Figure 6: An example of a pure directed percolation problem below, at, and above critical percolation probabilities from a seed row (top) and a single seed (bottom)

As seen in figure 6 from [8], directed percolation produces predictable patterns of clustered sites below, at, and above critical percolation values. Measures like correlation length, or the difference in either time or space (vertical and horizontal distance in figure 6, respectively), can be used to evaluate cluster size objectively. The correlation length of clusters which are the product of a tile assembly process was not measured for this thesis; instead broad cluster behaviors were determined subjectively by observation and through VisTool and those observations were used to find approximate critical percolation values for each tileset. Further

analysis beyond the scope of this thesis will require the use of correlation length, a more standard metric.

3.2 Tilesets and Definitions

A number of steps were taken to ensure the correctness of the simulation tool. Five tilesets were found or created which assembled independently verifiable configurations in the aTAM . These tilesets, discussed in sections 3.2.1 through 3.2.5, were chosen independently for a variety of reasons including historical relevance to the aTAM, their complexity, repetition or predictable long-range interactions, and regularity. Once the simulation tool was verified to correctly assemble final configurations given a particular tile assembly system, these tilesets were used to evaluate two new adaptations to the aTAM: the NULL tile and temperature 1.5. The details concerning the NULL tile and its application are related fully in 3.2.6 and temperature 1.5 is explained in section 3.2.7. Finally, these adaptations to the aTAM were used in tandem with each tileset to study directed site percolation behavior in the aTAM. A formal statement of equivalency is given in section 3.2.8.

3.2.1 Single Glue Tileset

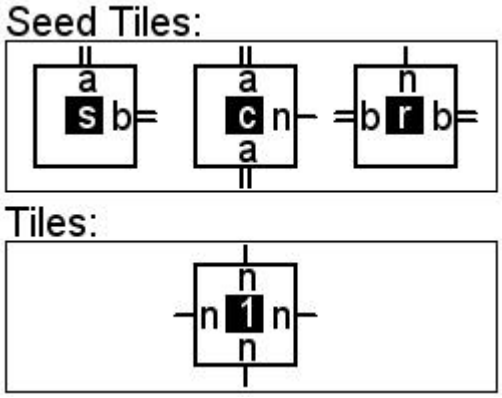

Figure 7: The Single Glue Tileset where tile names are shown in white and glue names are shown in black, one to each tile side. Glue strengths are represented by either a single connector (strength 1) or a double connector (strength 2).

The Single Glue tileset consists of one tile and one glue plus a set of seed tiles. This simple tileset was employed for the following two reasons: to understand experimentally the effect of NULL tile attachments on a trivial tileset, and to verify the correctness of the simulation for arbitrarily large configuration sizes. In addition, this tileset corresponded to directed percolation in a straightforward way, and was used to confirm the equivalence between tile assembly and directed percolation through simulations using the NULL tile at temperature 1.5. Assembly is resilient at temperature 1, that is aggregation is not halted by an error, and single NULL tile attachments (which model such errors) do not interrupt the normal assembly process. This is because the tile set has only a single glue. In the case of a NULL attachment, a subsequent temperature 1 attachment effectively "corrects" for the NULL and aggregation continues. Assemblies which are resilient to errors can be used to make large aggregates of tiles which reliably exhibit an expected behavior. A tile assembly system where **T** is the Single Glue Tileset can be reduced easily to a directed percolation problem using the algorithm described in

section 3.2.8 and directed percolation properties are immediately evident in the final configurations. This behavior is partially due to the tileset's resiliency to attachment errors due to its trivial size.

3.2.2 Double Glue Tileset

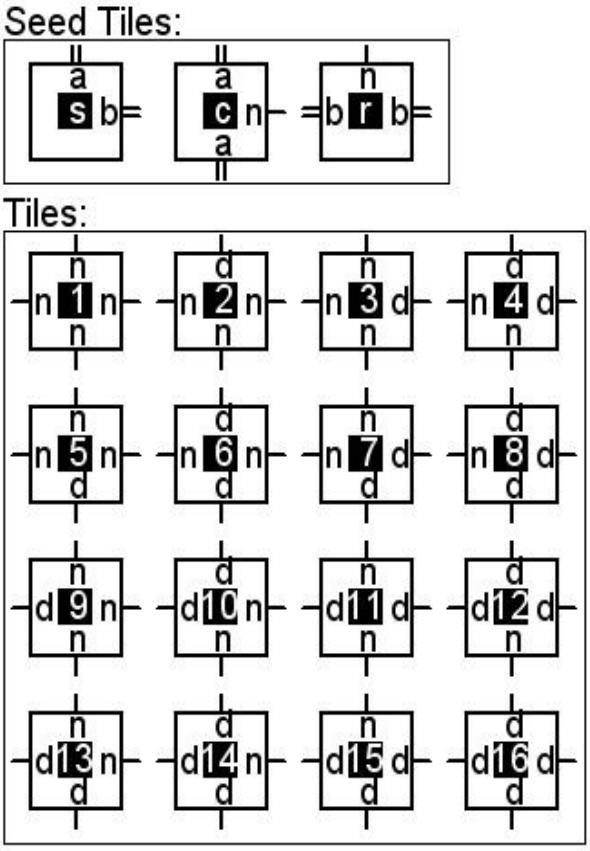

Figure 8: The Double Glue Tileset

The Double Glue Tileset consists of 16 tiles and two glues plus a set of seed tiles. Each of the 16 possible combinations of glues is present on some tile in the tileset. This contributes symmetry to the tileset. The Double Glue Tileset was created in order to reinforce and extend the analysis of the similar but simpler Single Glue Tileset. The tiles present in the tileset can bind

non-uniquely on any open binding site in the configuration, which introduces a non-deterministic element in a previously locally deterministic system. The full potential of this tileset wasn't explored in this thesis due to simulation constraints, but initial experimentation shows an intriguing relationship between the Double Glue Tileset and the Single Glue Tileset when NULL tiles are allowed. Consider a configuration of tiles from the Double Glue Tileset and the NULL tile. Each open binding site has 5 possible tiles which can occupy that site at temperature 2. At temperature 1, each site can bind 13 possible tiles. This highly non-deterministic attachment scheme builds on the trivial properties which are exhibited by the Single Glue Tileset. These properties include resilience to temperature 1 assembly as well as non-programmed growth by non-deterministic attachment. The Double Glue Tileset exhibits these properties but in a more chaotic way due to the number of tiles which can attach at a given open binding site. The sheer number of correct tile attachments on a single open site regardless of binding domain creates less predictable assemblies which are all still technically correct.

Assemblies of the Double Glue Tileset continue to grow, but correlation between a perfect assembly (or intended assembly) and another assembly even at temperature 2 is low. This indicates that even though the aggregate is still undergoing assembly, this assembly is no longer following an intended program.

3.2.3 Counter Quilt Tileset

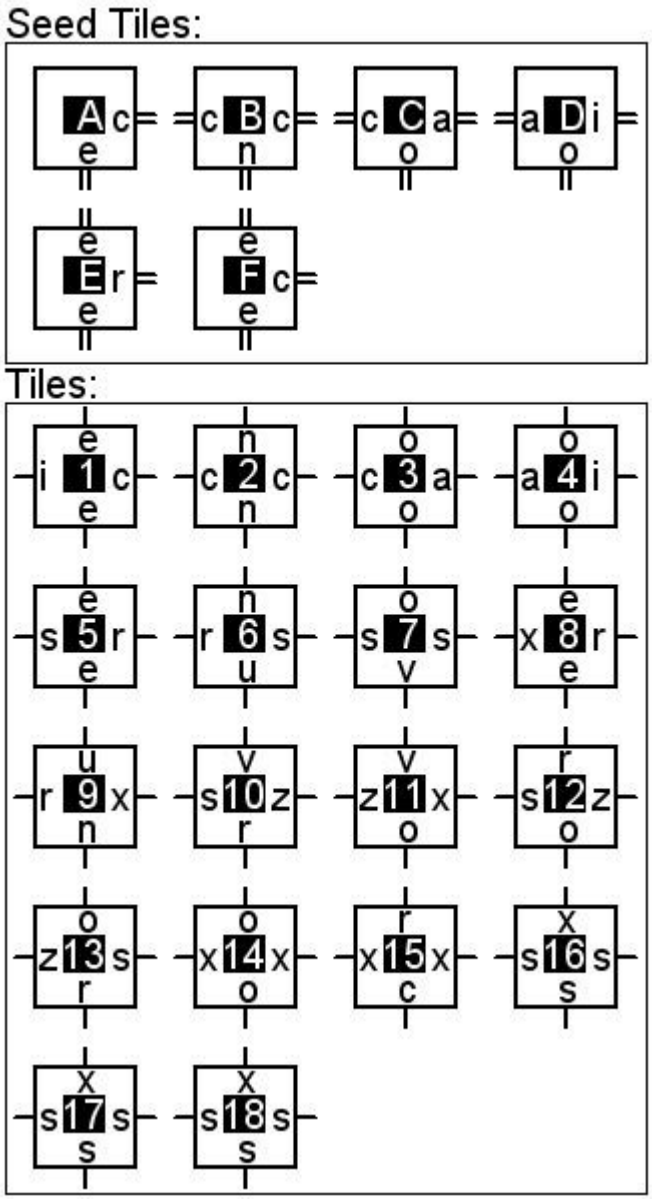

Figure 9: The Counter Quilt Tileset

The Counter Quilt Tileset has 18 tiles plus a set of seed tiles and consists of a tileset which assembles a finite configuration that is embedded within an infinite "quilting" tileset. The finite configuration is a small tileset which "counts" from 0 to 7 and is found in [25]. The purpose of this tileset was to construct small, test-able configurations which had well-understood

finite and infinite tiling properties not represented in the Single or Double Glue Tilesets. Essentially, if the tileset can construct a small quilt with perfect accuracy, then it can construct a large quilt with perfect accuracy. The predictable infinite properties can be used to easily show how errors during assembly can propagate or repair by random chance in a temperature 1.5 tile assembly system. Partial aggregates begin to assemble correctly and then an unfavorable attachment derails the aggregate. This is evidence that highly constrained tilesets (or tilesets with detailed and complex programs which necessarily have a large number of glues and possible binding domains) are more susceptible to entering an errorful state which is characterized by a preponderance of vacancies and incorrect tile attachments. Stated another way, complex tilesets are more susceptible to an absorbing phase transition.

A small configuration of 10000 tiles is shown below in figure 10. Each tile is represented by a small ellipse. Gray ellipses represent '0' tiles, while white ellipses represent '1' tiles. Quilting and seed tiles are shown in black.

Figure 10: A small configuration assembled from the Counter Quilt tileset

3.2.4 Sierpinski Tileset

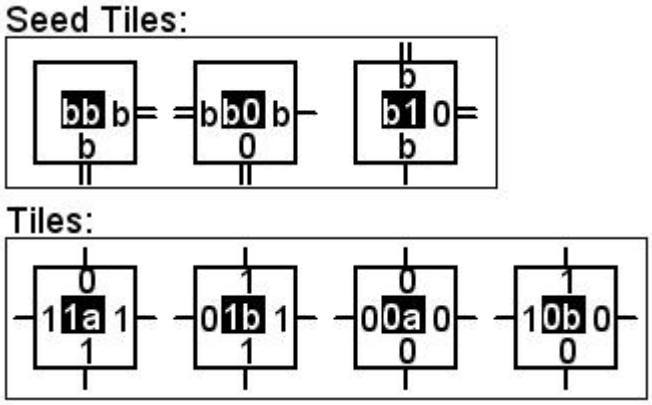

Figure 11: The Sierpinski Tileset

The Sierpinski Tileset was first proposed by Erik Winfree in [5] and later shown to configure within predicted error rates *in vitro* by Rothemund, Papadakis, and Winfree in [26]. The tileset consists of four tiles and a seed aggregate. Tiles are labeled as '1' or '0', and the final configuration is read based on these labels. The growth occurs diagonally in the fashion of a Sierpinski triangle. Evaluation of this tileset in assembly systems outside of the aTAM has generated interesting variations, including the fibered Sierpinski tileset. In [27], Lathrop et al. show that the fibered version of the Sierpinski tileset has the same zeta dimension as the Sierpinski tileset, and can be strictly assembled in a generalized Tile Assembly Model. The fibered Sierpinski tileset is interesting in that it is fundamentally a collection of linear structures of tiles which are composed parallel and perpendicular to one another. These linear structures are the “fibers” of the Sierpinski triangle design.

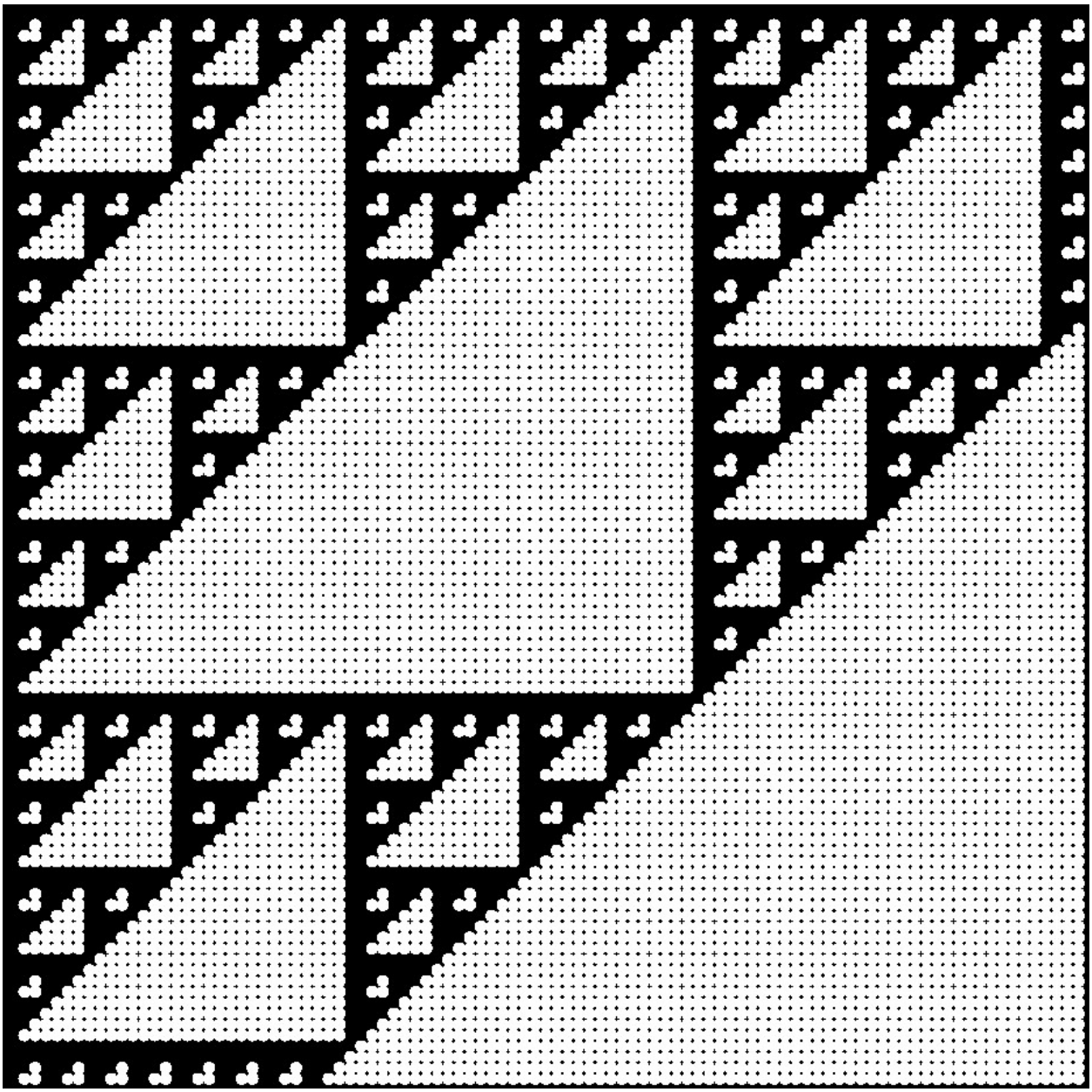

Figure 12: A small Sierpinski triangle configuration

The Sierpinski tileset was chosen because of the diagonalized pattern consisting of regular triangles which grow according to rules encoded into the tileset's design. Like the Counter Quilt Tileset and unlike the Single and Double Glue Tilesets, the Sierpinski Tileset has been designed to yield a particular final configuration, thus, errors in attachment and other

probabilistic behavior can be analyzed based on the aggregates formed by tile assembly systems with known parameters.

3.2.5 Binary Counter Tileset

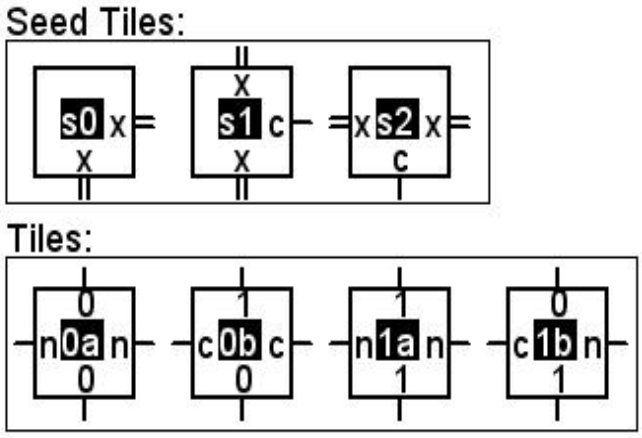

Figure 13: The Binary Counter Tileset

The Binary Counter Tileset from [13] is a simple tileset consisting of four unique tile types plus a seed aggregate. Each tile type is labeled as either '1' or '0' much like the Sierpinski tileset. After the configuration is assembled, correctness is tested by polling the assembly and "reading" the tile labels. In the tileset, two tiles are labeled '1' and two tiles are labeled '0'. The configuration grows infinitely assuming infinite quantities of each tile type. Each row of the assembly is a single number in binary representation and each column encodes the "place" of its contained digits. The underlying logic of the algorithmic assembly is borrowed from the ripple-carry adder. The initial conditions set by the seed determine the first "number" to be assembled. Figure 14 shows a configuration assembled from the Binary Counter Tileset which counts from 0 to 63. '1' tiles are shown in white and '0' tiles are shown in gray. The least significant digit is represented by the leftmost column in the configuration. The Binary Counter Tileset was chosen

to give counterpoint to the Counter Quilt Tileset, a derivative tileset, while confirming assembly behavior seen in the Sierpinski Tileset.

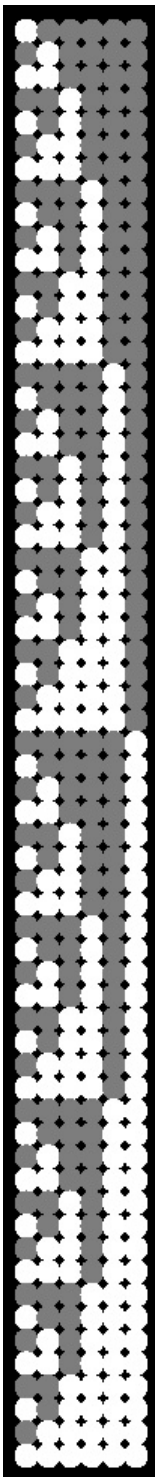

Figure 14: A configuration of the Binary Counter Tileset which "counts" from 1 to 63 where 0 is a seed row for the counter.

3.2.6 The NULL Tile

The NULL tile represents a "hole" in the configuration. In the aTAM, there is no consideration for a binding site which fails to attach the appropriate tile. If no tile can occupy a particular binding site, then that site is left unoccupied and assembly continues. In general, this behavior is acceptable when taken with the other constraints imposed by the aTAM, but laboratory assembly can introduce geometric and spatial limitations which cannot be modeled in the aTAM. If an open binding site is surrounding on all sides by tiles, the tile which wishes to bind to that site cannot due to geometric constraints. This type of vacancy can be referred to as a kinetic trap. The speed of the assembly at some locality of binding sites is greater than the global speed of assembly, which means that the aggregate can "grow around" an open binding site such that no tile can attach to the binding site. In other words, two tiles in solution cannot occupy the same location, so the open binding site becomes "fenced in" by other tiles. In addition, if tiles in solution are represented in different concentrations (and not held at constant and equivalent concentrations as expected in the aTAM or kTAM), then tiles with higher concentrations are more likely to attach to open binding sites than tiles with lower concentrations by simple probability.

In order to model this, the author introduces a NULL tile to the aTAM which attaches probabilistically to an aggregate. In order to work within the attachment constraints set by the aTAM, at each open binding site a tile must attach, if possible. If no such attachment is possible, then a NULL tile is attached. To simulate kinetic traps, it could also be the case that a NULL tile is attached instead of a tile from the tileset. In this case there is some probability p that tile t will attach to aggregate C and some probability $(1 - p)$ that the NULL tile will attach instead, where 0

$\leq p \leq 1$. Therefore there are two conditions under which the NULL tile attaches to the aggregate: either no tile in the tileset can make a stable addition, or the NULL tile has been selected to attach over any other tile in the tileset. Whenever the NULL tile is allowed, either condition for attachment is valid. The NULL tile's behavior is unchanged for temperature 1.5 systems.

3.2.7 Temperature 1.5 Systems

Temperature 1.5 is a hybrid temperature. A tile assembly system at temperature 1.5 assembles at either temperature 1 or temperature 2 probabilistically. More formally, temperature 1.5 describes a system where each attachment is made at temperature 1 with some probability p and temperature 2 with the probability $(1 - p)$. A tile assembly system at temperature 1.5 can simulate temperature 1 assembly if the probability of temperature 2 attachment is 0, and a tile assembly system at temperature 1.5 can simulate a temperature 2 assembly if the probability of temperature 1 attachment is 0. The temperature of each attachment is determined at attachment time for each tile accreting to the aggregate. If a tile is unable to attach to the configuration at a given open binding site, then a NULL tile is attached, regardless of current attachment temperature or other probabilistic considerations.

3.2.8 The aTAM and Directed Percolation

Informally, a tile assembly system can be transformed into a directed percolation problem by turning each tile into a site in a square lattice, and each stable bond between two tiles into a directed edge between the appropriate two sites. If the tile in the aggregate is an error, then the corresponding site is unoccupied, else if the tile is correct, then the corresponding site is occupied.

More formally, consider some tile assembly system $T = (\mathbf{T}, \sigma, \tau)$ that assembles some configuration C. C is an n x n square. Let site $s \in S$, where S is the set of sites, be $s = (i, j)$ where $i, j \in$ Z is a site on a lattice indexed by (i, j) from 0 to n. Let g: (S, S) $\rightarrow$ Z where S is the set of sites in C, and Z is the integers and where $g(s_1, s_2)$ is the function that returns the binding strength between sites 1 and 2 in S. Create a square lattice of size n x n called L where each node is called a site and can be labeled occupied or unoccupied and each edge is called a bond and can be labeled open or closed. For each edge e from l_1 to l_2 in L, make e a directed edge from l_1, the site at index (i, j), to l_2, the site at index $(i+1, j)$, then do the same for the edge from l_1 to l_3, the site at index (i, j+1). Then tilt L such that the first tilted row of L contains one site, the second tilted row contains two sites, the third tilted row contains three sites, and so on. Each tilted row is called a timestep, and sites in a timestep are referenced from left to right such that the leftmost site is 0, the next site is 1, and so on. Each site in L_{TILT} can thus be referenced by a pair (u, v) where $u, v \in Z$ and $u, v \geq 0$ where u is the timestep and v is the index into that timestep. See figure 15 for clarification. At each timestep u, for each index v, mark site (u, v) occupied in this way:

I. If $u = v = 0$, mark site (u, v) occupied and call this the seed for the directed percolation
II. If $0 < v < u - 1$, mark site (u, v) occupied with probability p if both site $(u - 1, v - 1)$ and site $(u - 1, v)$ are occupied where $\tau = 2$, else if either site $(u - 1, v - 1)$ or site $(u - 1, v)$ are occupied mark site (u, v) occupied with probability p where $\tau = 1$, else mark site (u, v) unoccupied with probability $(1 - p)$

- mark directed edges e_1 from site $(u - 1, v - 1)$ to site (u, v) and e_2 from site $(u - 1, v)$ to site (u, v) open with probability q if $g((u - 1, v - 1), (u, v)) + g((u - 1, v), (u, v)) \geq \tau$

III. If $v = 0$, mark site (u, v) occupied with probability p if site $(u - 1, v)$ is occupied, else mark site (u, v) unoccupied with probability $(1 - p)$

- mark directed edge e from $(u - 1, v)$ to (u, v) open with probability q if $g((u - 1, v), (u, v)) \geq \tau$

IV. If $v = u$ and $v > 0$, mark site (u, v) occupied with probability p if site $(u - 1, v - 1)$ is occupied, else mark site (u, v) unoccupied with probability $(1 - p)$

- mark directed edge e from $(u - 1, v - 1)$ to (u, v) open with probability q if $g((u - 1, v - 1), (u, v)) \geq \tau$

In this way, if $T = (\mathbf{T}, \sigma, \tau)$ where $\tau = 2$ and T configures an n x n square, all sites (u, v) in L_{TILT} are occupied and all bonds are open. Thus, where $p = q = 1$, T is a deterministic tile assembly. Where $p < 1$, T is a tile assembly system operating at temperature 1.5. In general, the probability of an assembly is given by p_1 and p_2 from section 2.1.2.

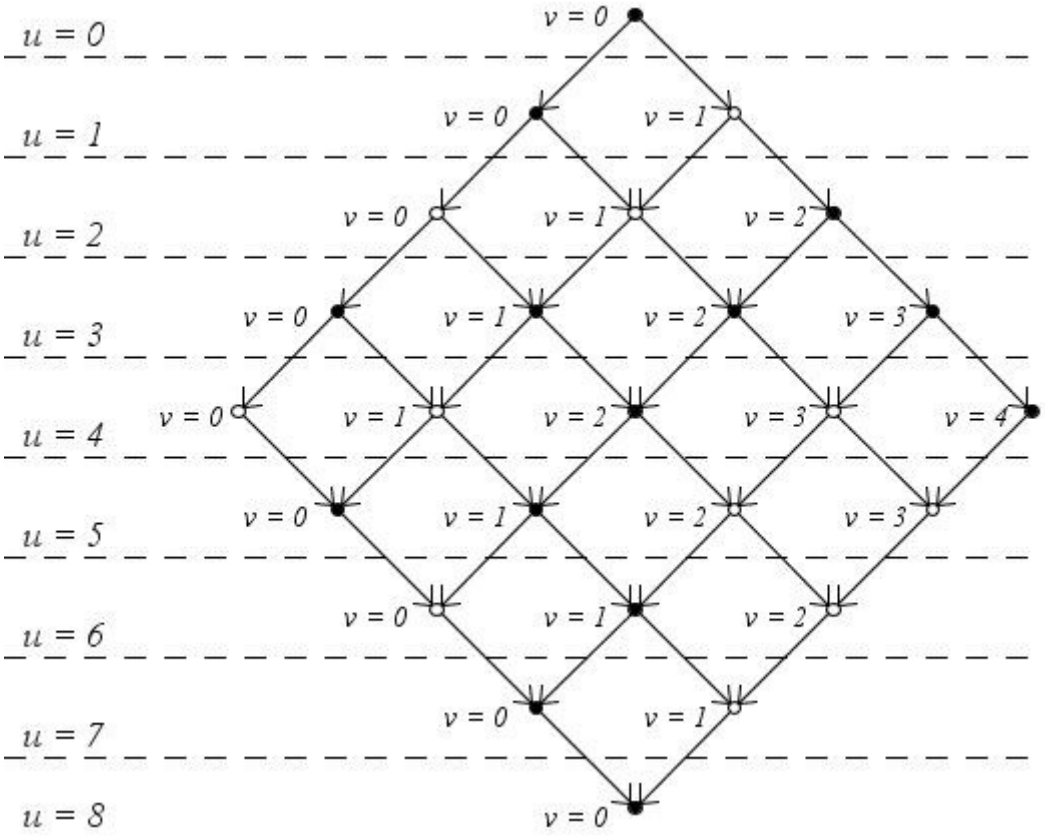

Figure 15: A tilted lattice *L* where each site is referenced by a coordinate pair (*u*, *v*) such that *u* is the timestep and *v* is the index into that timestep. For example, node (2, 3) in configuration *C* is site (5, 1) in *L*.

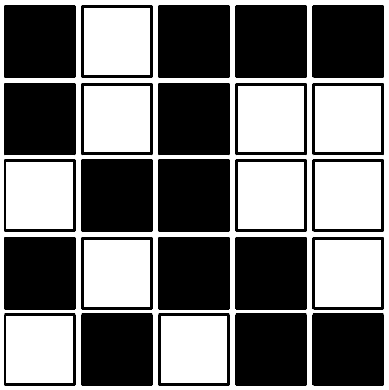

Figure 16: The corresponding tile assembly where black tiles are correct and white tiles are errors at temperature 1.5.

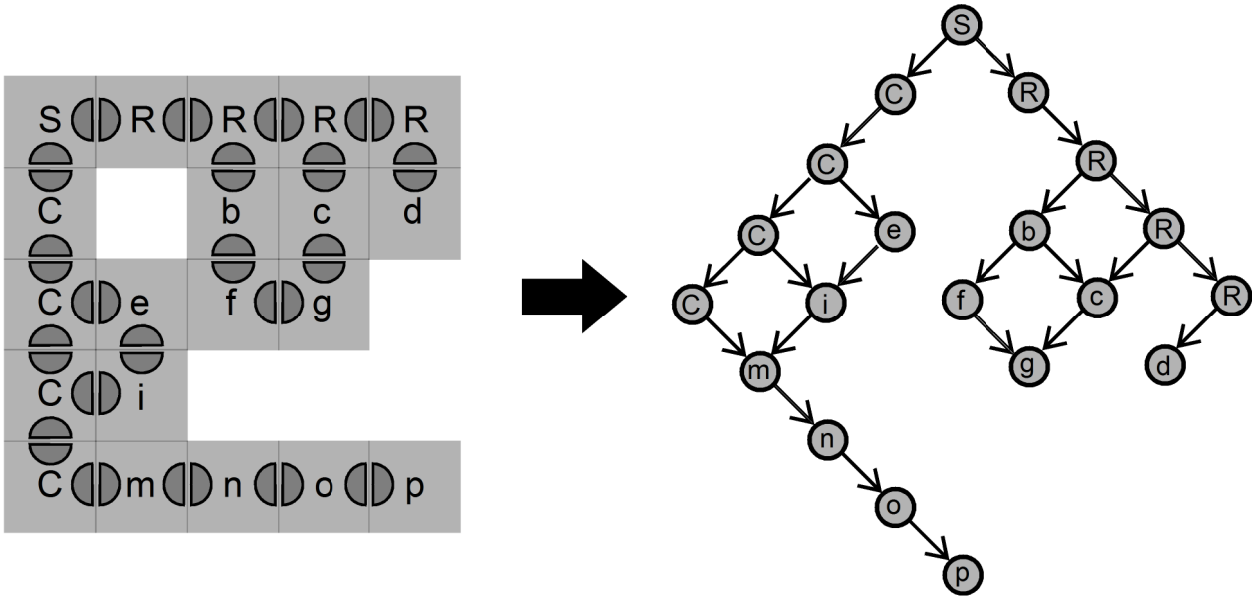

Figure 17: A sample configuration of tiles with errors and both temperature 1 and temperature 2 stable bonds mapped to a square lattice of occupied sites connected by open bonds.

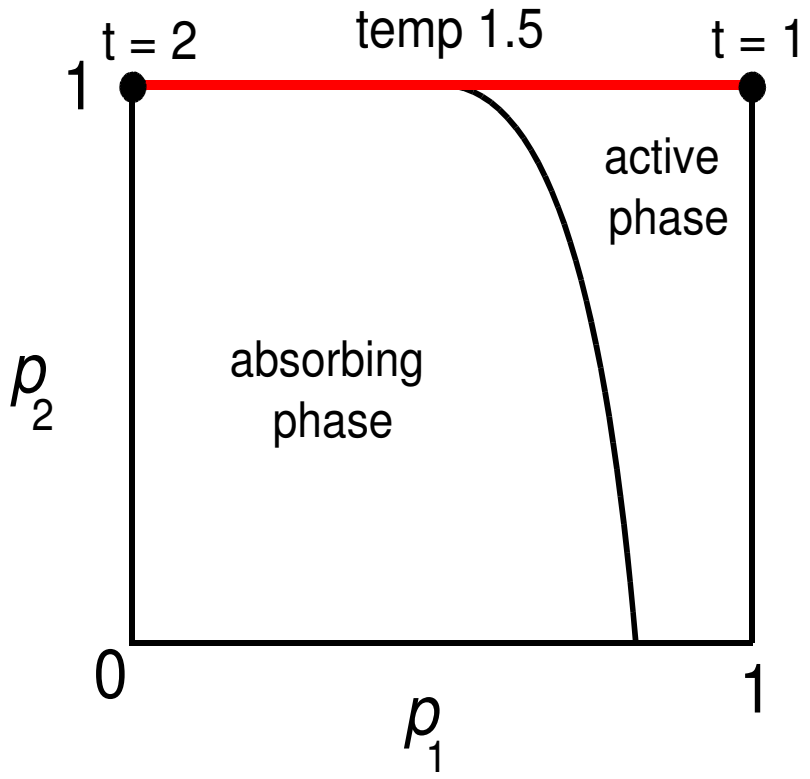

Figure 18: A phase diagram of the DKCA adapted from [8] with additional points of interest. Where $p_2 = 1$ and $p_1 = 0$, this corresponds to a temperature 2 assembly. Where $p_2 = p_1 = 1$, this corresponds to a temperature 1 assembly. Along the red line, temperature 1.5 behavior is observed.

3.2.9 Directed Percolation in the aTAM

Directed percolation can be simulated in the aTAM with a simple tileset which follows the rules of the Domany-Kinzel automata. See table 1 for these rules. Each sequence of "parents" is given 2 child tiles, one which represents a occupied site and one which represents an unoccupied site. The probability of occupation, represented by p_1 and p_2, is controlled by tile concentrations in solution. Figure 19 shows the tileset which corresponds to the DKCA.

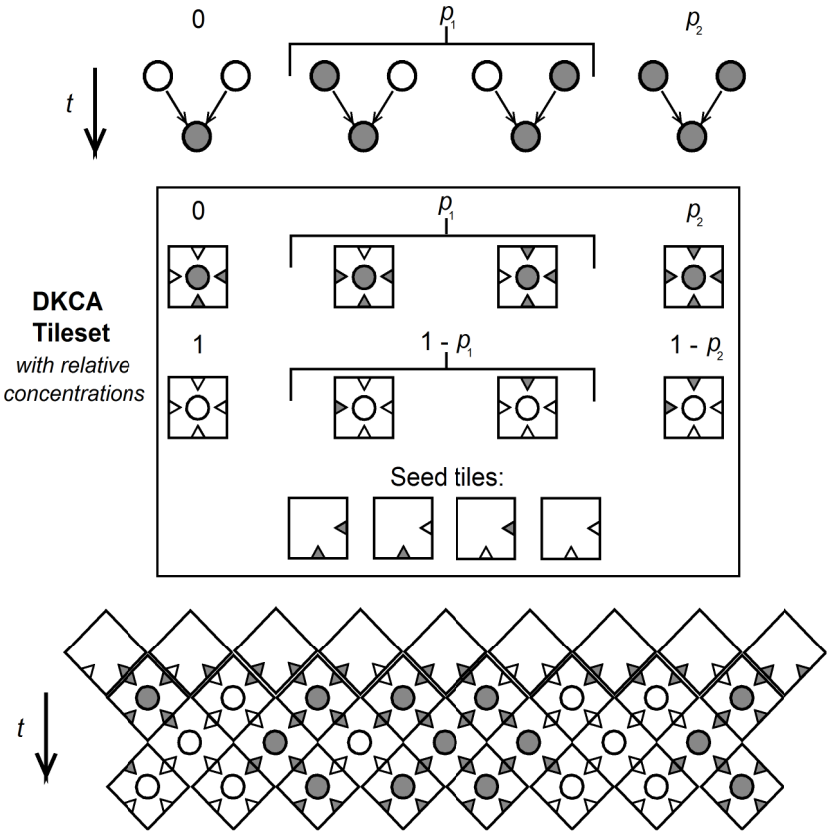

Figure 19: The DKCA tileset with concentrations of tiles. In this way, the aTAM can simulate the DKCA at temperature 2 and thus an arbitrary DP process.

3.3 Implementation and Methods

As mentioned previously, each Configuration can be output to a text file where each Tile is identified by name and separated by white space. Lines in the file correspond to rows in the Configuration. Text files representing small final Configurations of the pre-existing tilesets mentioned above were checked by hand against existing correct examples to verify the accuracy of the assembled Configurations. In order to cross-verify the accuracy of assembled Configurations, VisTool and the Correlator were employed. By constructing a verified target

Configuration at $\tau = 2$ and then comparing it to a Configuration at $\tau = 1$ or a temperature 1.5 system, an analysis of the applicability of temperature 1.5 is made which tests fidelity or the deviation of a temperature 1 or 1.5 Configuration from the target. Correlation data was also found between NULL inclusive tilesets and NULL exclusive tilesets. Although test Configurations were constructed which included 25 million tiles, conclusions are drawn only from Configurations of 1 million Tiles due to memory and computer time constraints. Also, tile concentrations were held equal and infinite quantities of each tile are assumed to keep all probabilities uniform.

A series of simulations using each tileset in the tile assembly tool were performed. The first set of simulations was of each tileset at temperature 2 with no modifications. A second series of simulations was then performed at temperature 1. Correlation between the temperature 2 and temperature 1 assemblies was then recorded. Many simulations of each tileset were performed at temperature 1 since the non-deterministic attachment process can create wildly different final configurations and assembly can fail before a final configuration is reached. A series of simulations was then performed at temperature 1.5 with varying probabilities of temperature 1 attachment, ranging from 0% to 100%. Expected behavior at $p = 0$ was perfect correlation with temperature 2 assemblies and expected behavior at $p = 1$ was approximate temperature 1 correlation. Next a series of simulations was performed at temperature 2 and temperature 1 which included the NULL tile with probability of attachment ranging from 0% to 100%. At $p = 0$, correlation behavior commensurate with the temperature of the system was expected and at $p = 1$ an all NULL configuration was expected. Finally, simulations were performed at temperature 1.5 and with NULL tiles. Directed percolation has clearly defined

behaviors in the qualities of clusters it produces on lattices below, at, and above critical percolation values. This behavior is seen clearly in such systems where the probabilities of temperature 1 attachment and NULL tile attachment control the cluster behavior. After each set of simulations, high contrast images were created with VisTool of selected configurations in order to analyze the behavior of each tileset under each set of conditions.

Configurations were examined for a number of properties, particularly correlation and the existence of spanning clusters. First, these configurations were parsed by hand and those which seemed to exhibit interesting properties were given to VisTool for further analysis. Each configuration processed by VisTool is transformed into a 1,000 x 1,000 pixel image where each pixel represents a single tile. Tiles from the tileset or seed are colored black and NULL tiles are colored red. A configuration is determined to have a percolating cluster if the tileset assembled such that there exists a path from the lower right hand corner of the configuration to the upper left hand corner which consists only of orthogonally adjacent black pixels. These paths are found greedily and such a path is guaranteed to exist if an infinite cluster exists because of the nature of the directed percolation process. The probability of spontaneous generation was held at 0 and the temperature of the system was always at least 1, therefore each tile from the tileset must be adjacent to either one or more seed tiles, one or more tiles from the tileset, or a mix of the two.

4. RESULTS

4.1 Single Glue Tileset

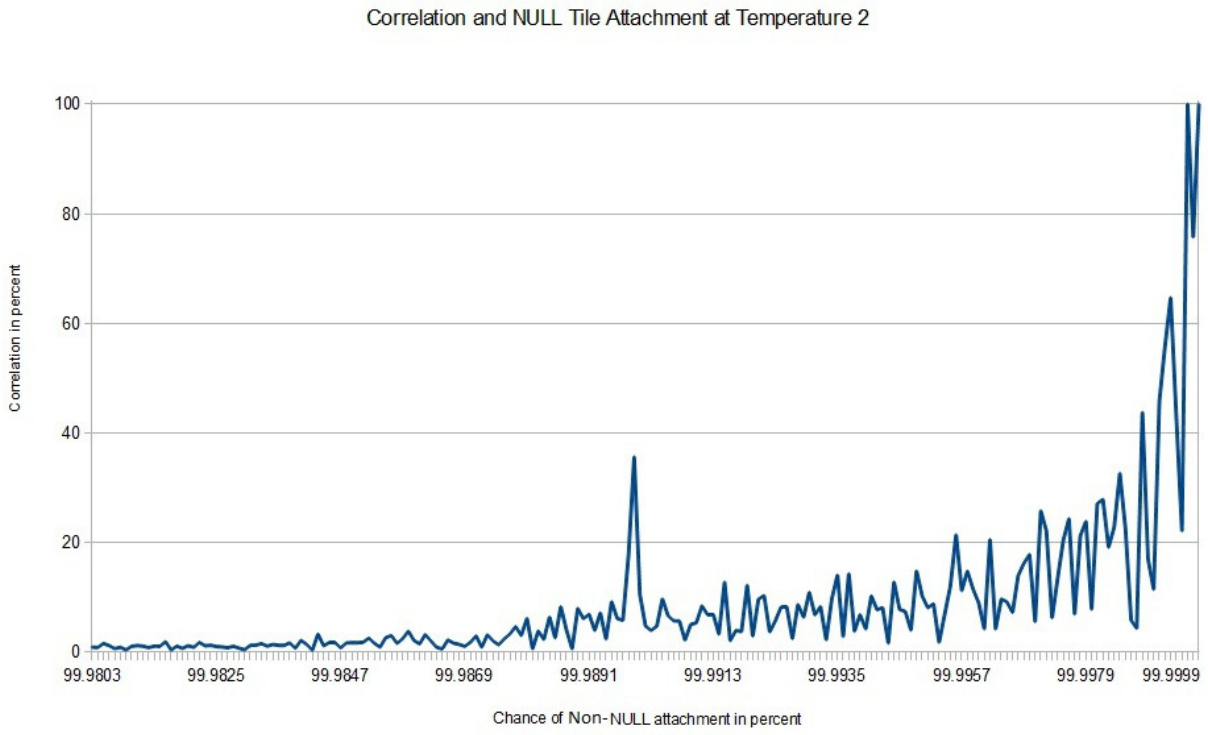

Figure 20: The correlation between a 1000x1000 Single Glue Tileset Configuration in a temperature 2 tile assembly system with NULL Tiles and without. Each datapoint represents a single simulation at the appropriate chance of non-NULL attachment.

As seen in figure 20, the NULL tile at temperature 2 greatly affects the correlation between the configuration assembled with no probability of NULL tile attachment and the configuration with some probability of NULL tile attachment. As long as there is more than a trivial probability of NULL tile attachment, correlation remains very low. In other words, any chance of error in the assembly destroys correlation with an assembly of the same tileset at

temperature 2. As the probability of NULL attachment increases, an absorbing phase transition is seen. In figure 18, the assembly moves off of the temperature 2 point of interest and into the absorbing phase portion of the diagram. Even at a 99.98% chance of non-NULL attachment, correlation hovers around 1%, meaning only 1 in 100 tiles at some (i, j) in the configurations match. As this probability of non-NULL attachment approaches 99.9999%, correlation approaches 100%. At temperature 1, a positive linear relationship between the NULL tile configuration and the non-NULL tile configuration is seen, as shown in figure 21. For this simple tileset, the probability of a correct configuration is given by the probability of non-null attachment. This is illustrated further by assembly at temperature 1.5 where the probability of attaching a NULL tile is 0%. As shown in figure 21, the assembly is not affected by the probability of temperature 1 attachment, whether it is a 0% or 100% chance. The assembly only depends on the probability of NULL attachment. Figure 22 also shows that the same sharp transition seen in figure 20 at temperature 2 is also present at temperature 1.5 where $p_1 = .5$ for attaching at temperature 1 and $p_2 = (1 - p_1) = .5$ for attaching at temperature 2.

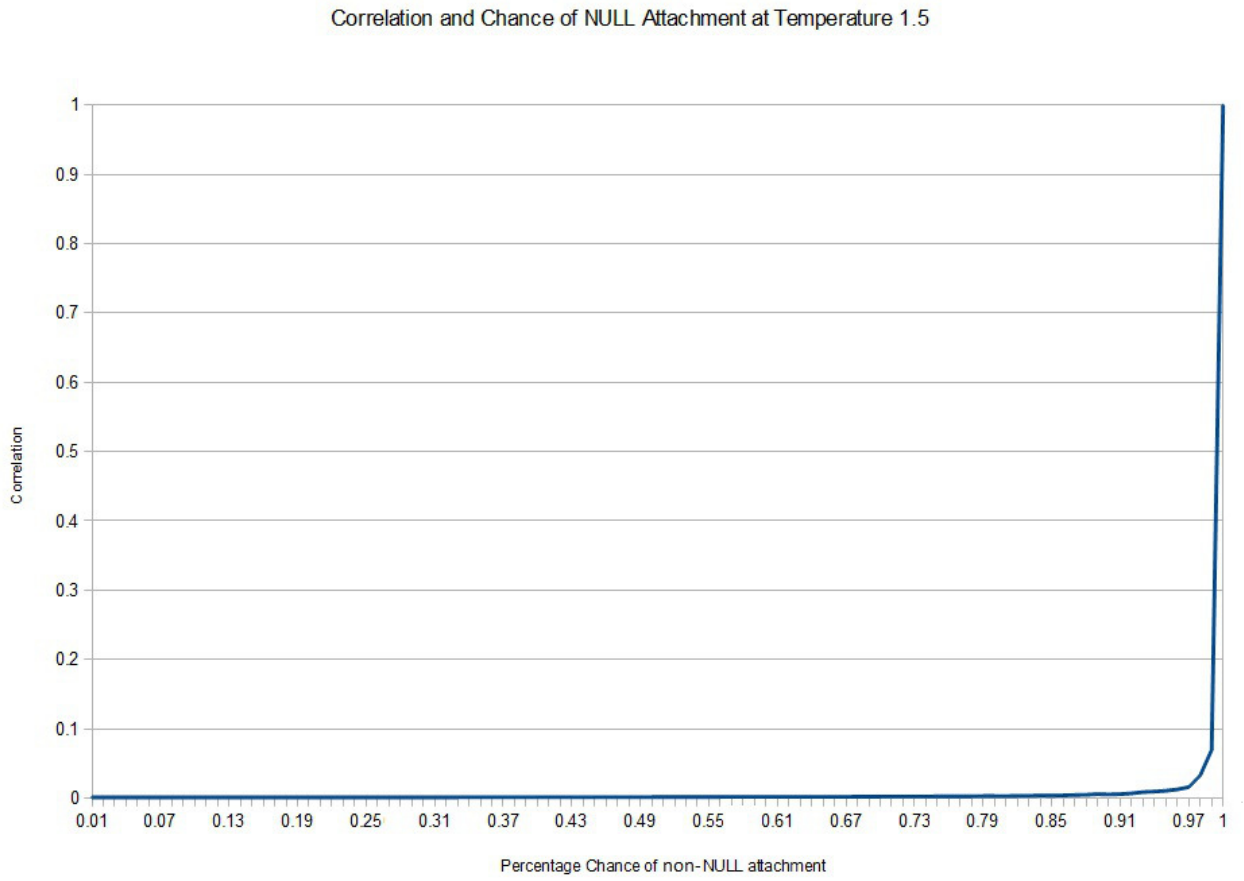

Figure 21: Correlation between a temperature 2 assembly and a temperature 1.5 assembly of Single Glue Tileset where probability of temperature 1 attachment is 50% and probability of non-NULL attachment ranges from 0% to 100%

Directed percolation behavior is seen with a non-NULL attachment probability of 71% which is close to the directed site percolation critical probability seen in the DK model mentioned in section 2.2.2. As mentioned previously, the chance of temperature 1 attachment is superfluous for this tileset since behavior is identical in temperature 1 and temperature 2 tile assembly systems using the Single Glue Tileset. Figures 23, 24, and 25 depict aggregates of the Single Glue Tileset below, at, and above the critical probability for percolation.

Figure 23: Single Glue Tileset where $p < p_c$

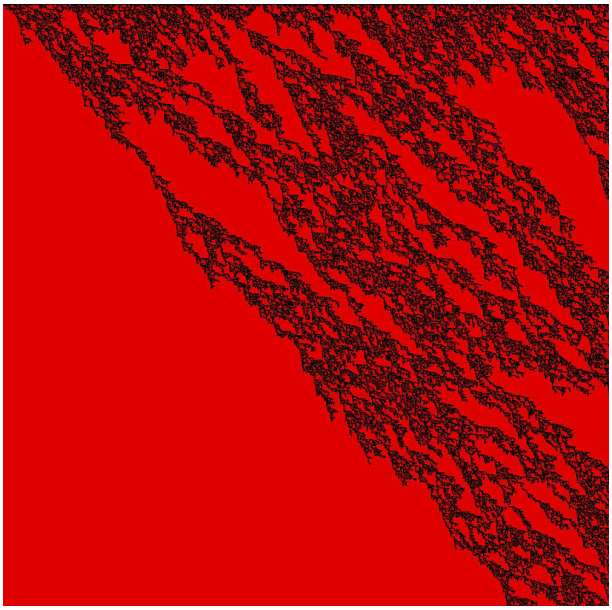

Figure 24: Single Glue Tileset where $p = p_c$

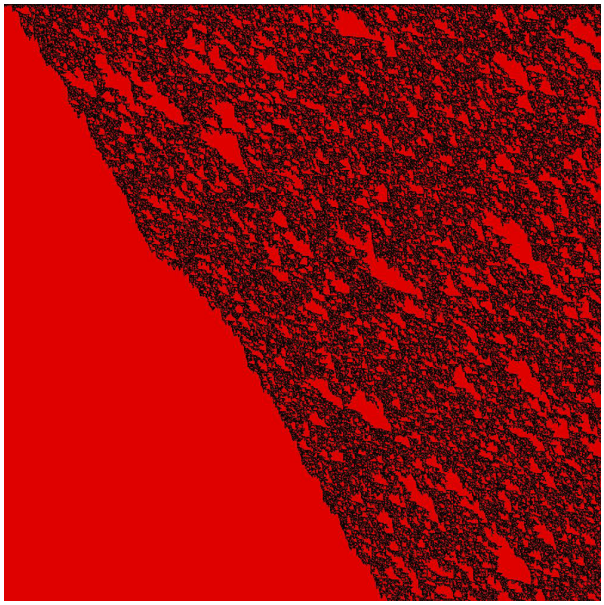

Figure 25: Single Glue Tileset where $p > p_c$

4.2 Double Glue Tileset

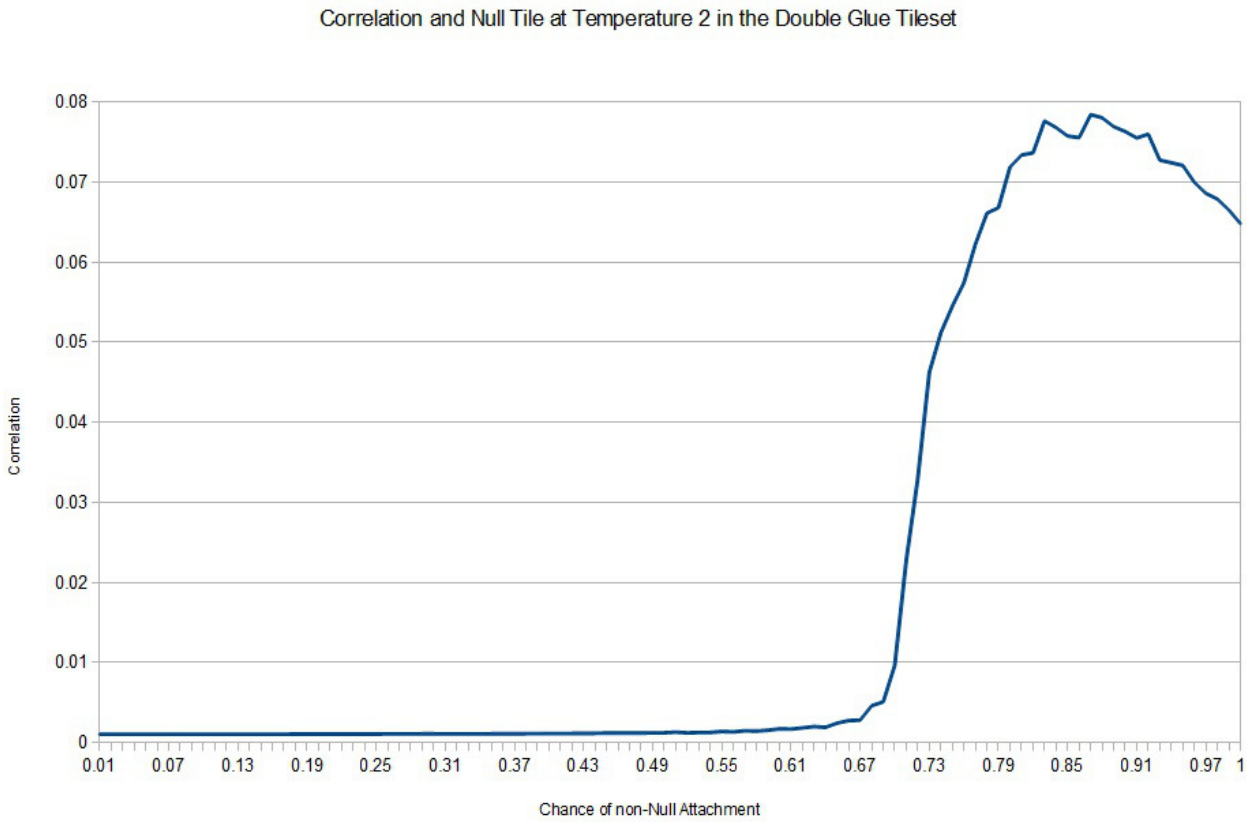

Figure 26: When the Double Glue Tileset with the NULL tile assembles at temperature 1, correlation with assemblies at temperature 2 stabilizes at a little more than 6%. This result is indicative of the effect of local determinism on assembly correlation.

As seen in figure 26, correlation converges to about 6% as the chance of non-NULL tile attachment approaches 100%. It should be noted that correlation is not a strong measure for "correctness" when considering a highly non-deterministic tile assembly system. The Double Glue Tileset has high binding domain symmetry among its constituent tiles and this fouls the

simulation process because attachments are made greedily. Correlation between a temperature 2 Double Glue configuration and a temperature 1.5 Double Glue configuration with NULL tiles is therefore a poor measure of correctness. In order to avoid this issue, any non-NULL attachment is considered a correct attachment in further analysis of the Double Glue tileset.

Allowing NULL tiles has a negligible effect on this correctness unless the chance of NULL attachment is so high that the chance of attaching a NULL row (a row of NULL tiles which begins at site (*i*, *j)* and ends at (*i*-1, *j*+1)) approaches unity. *In vitro* this may not necessarily be a true "row" of components, it is simply the leading edge of the assembly which must be errorful (or NULL). Much like a temperature 2 tile assembly system cannot recover from a single NULL tile, a temperature 1 tile assembly system cannot recover from a NULL row attachment. Effectively, the NULL row and NULL tile operate in the same fashion in their respective tile systems. They disrupt the attachment of new tiles to the assembly to such a degree that the process fails and the returned configuration is incomplete and practically uncorrelated with complete temperature 2 assemblies. This process mirrors directed percolation's absorbing state. In temperature 2 assembly, a single NULL tile is enough to signal an absorbing state, while in temperature 1 assembly, a row of NULL tiles (as defined above) signals an absorbing state. As long as the chance of non-NULL attachment is sufficiently high (as defined by simulation of other percolation systems collected in [8], this is about a 70% chance of non-NULL attachment), the assembly continues and as the size of the configuration approaches infinity, infinite range percolation properties are expected. This phase transition at about 70% is seen coincidentally in figure 26 as well. Figures 27, 28, and 29 show Double Glue Tilesets below, at, and above this critical probability for percolation.

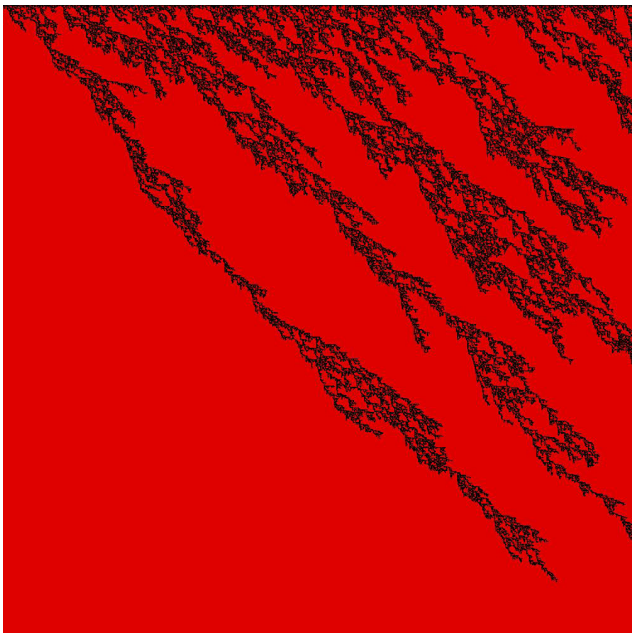

Figure 27: Double Glue Tileset where black represents any tile from the tileset and where $p < p_c$

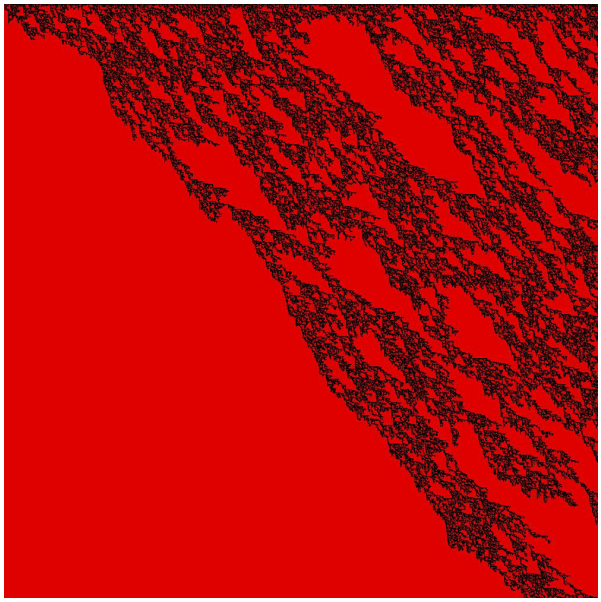

Figure 28: Double Glue Tileset where black represents any tile in the tileset and where $p = p_c$

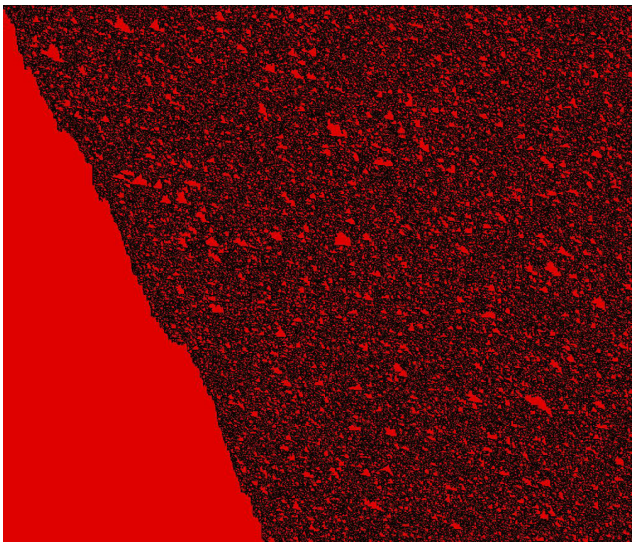

Figure 29: Double Glue Tileset where black represents any tile from the tileset and where $p > p_c$

4.3 Counter Quilt Tileset

The Counter Quilt tileset is not resilient to temperature 1 assembly in a temperature 1.5 system. It is difficult for the system to recover from a single temperature 1 attachment in a predominantly temperature 2 system. Correlation between temperature 2 systems and temperature 1.5 systems exhibits a strong phase transition when the probability of a temperature 1 attachments approaches 0, as seen in figure 30. Despite this interaction, the Counter Quilt tileset does not halt or fail to assemble in a temperature 1 system. The highly symmetrical nature of the quilt and the binary counter create a robust environment for attachment which fuels the assembly as long as the NULL tile is disallowed. When the NULL tile is introduced, directed percolation behavior is seen with critical probabilities of temperature 2 attachment at approximately 86% and non-NULL attachment probability at 87%. Figures 31, 32, and 33 depict the aggregates formed by the Counter Quilt Tileset below, at, and above these critical probabilities.

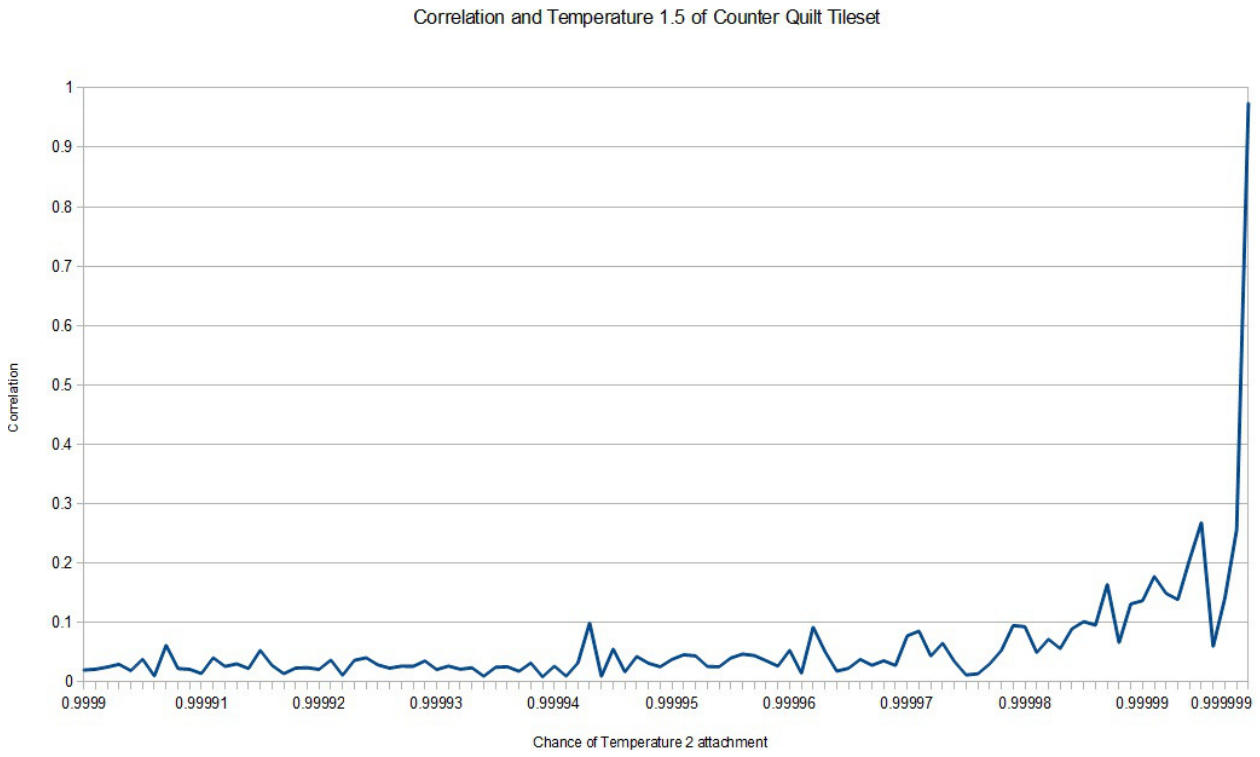

Figure 30: Correlation between temperature 2 systems and temperature 1.5 systems as probability of temperature 2 attachments approaches 100%

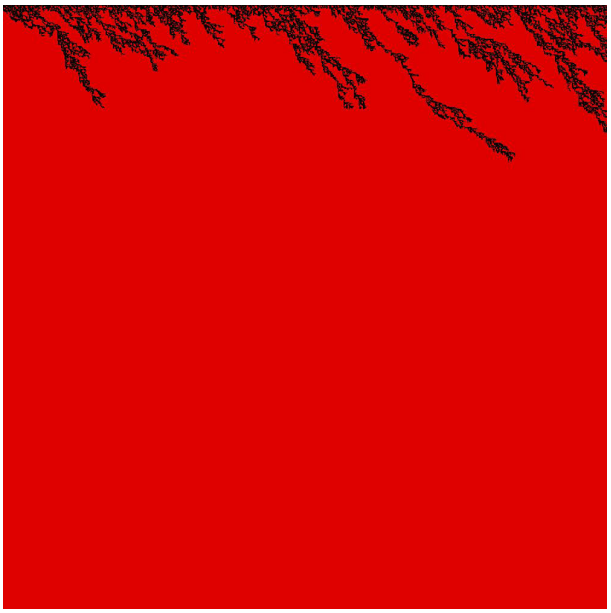

Figure 31: Counterquilt Tileset where black represents any tile from the tileset and where $p < p_c$

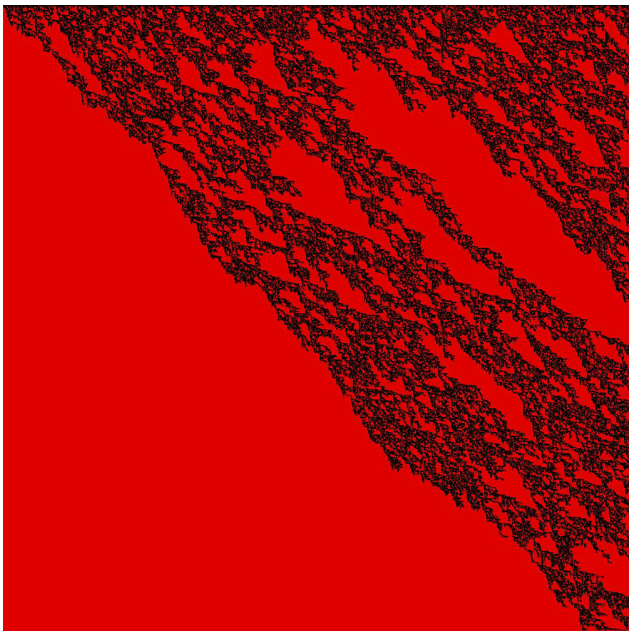

Figure 32: Counterquilt Tileset where black represents any tile from the tileset and where $p = p_c$

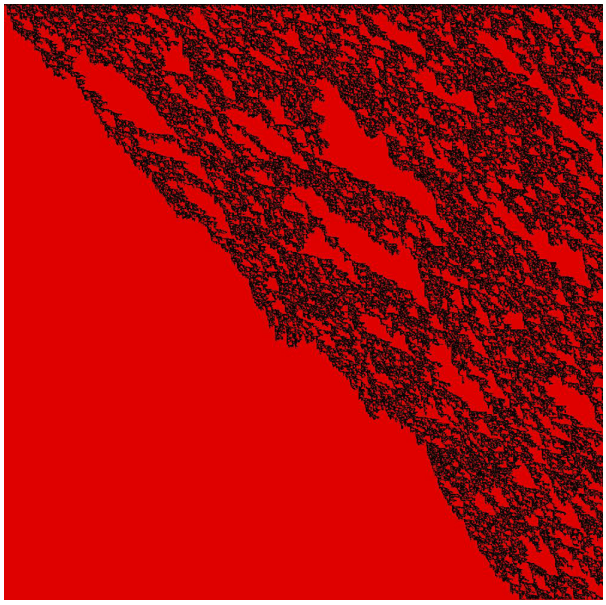

Figure 33: Counterquilt Tileset where black represents any tile from the tileset and where $p > p_c$

4.4 Sierpinski Tileset

The Sierpinski Tileset shows behavior not seen before in other tilesets when considering NULL attachment. As seen in figure 34, there is a smoother phase transition as the system approaches a 100% chance of non-NULL attachment, but it is not as marked as previous tilesets. Assembly at temperature 1.5 yields critical probabilities of temperature 2 attachment at approximately 75% and non-NULL attachment probability at 87%. An assembly constructed at temperature 1 without NULL attachment is illustrated in figure 35. The design of the tileset is still somewhat visible, with some collections of tiles taking on triangle-like formations, but the unconstrained aggregation of tiles leads to restarts and non-triangular clusters sprinkled throughout the configuration.

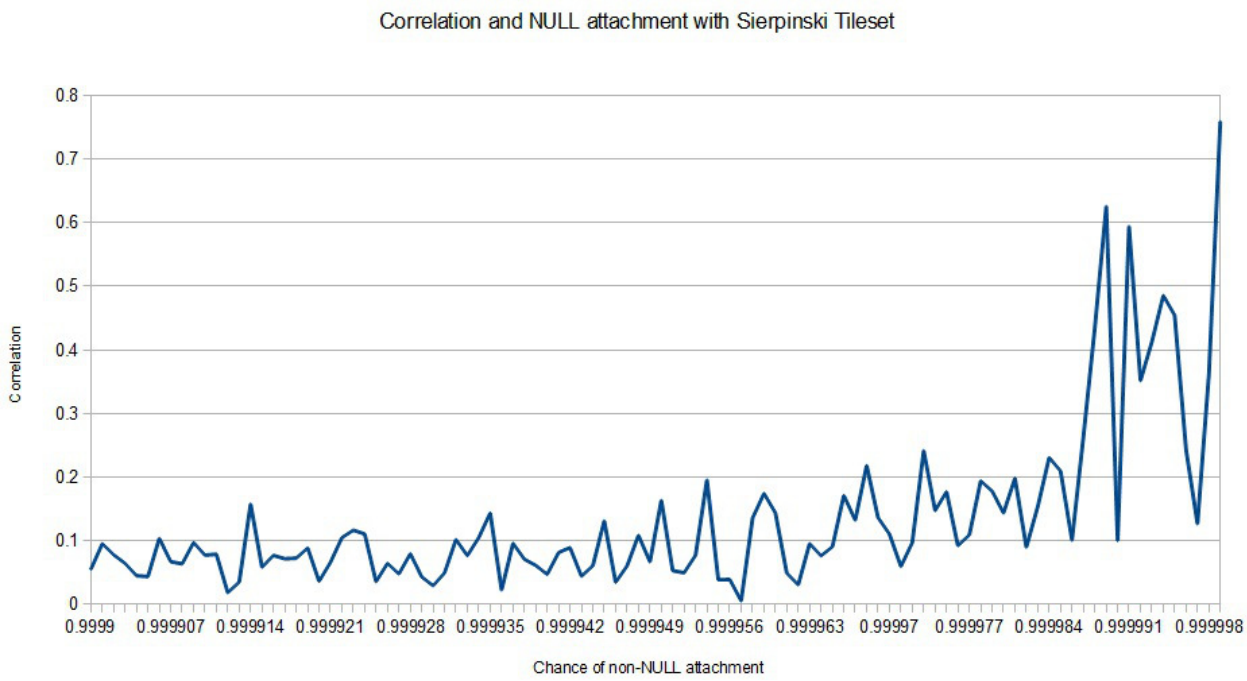

Figure 34: The phase transition evident in other tilesets is difficult to determine from correlation percentage in the Sierpinski Tileset

In addition, if the whole configuration is considered and NULL tiles are allowed with a small probability at temperature 1, the relationship between directed percolation and tile assembly becomes more clear at a high level. In figure 36, elongated clusters of tiles from the tileset come to narrow points before entering an absorbing state, so it is easy to see as NULL attachment probabilities and the probability of temperature 1 attachment both approach the critical probabilities, these clusters will eventually become more attenuated and exhibit infinite length in the system. Conceptually, figure 36 depicts the behavior of the Sierpinski tileset around NULL tile attachments. Examples of the Sierpinski Tileset below, at, and above the critical probabilities given earlier can be seen in figures 37, 38, and 39. As the chance of NULL attachment increases, the connection between figure 36 and 37 becomes more clear. Tiles from the tileset begin to fill in around the NULL attachments, producing clusters of NULL tiles surrounded by tiles from the tileset.

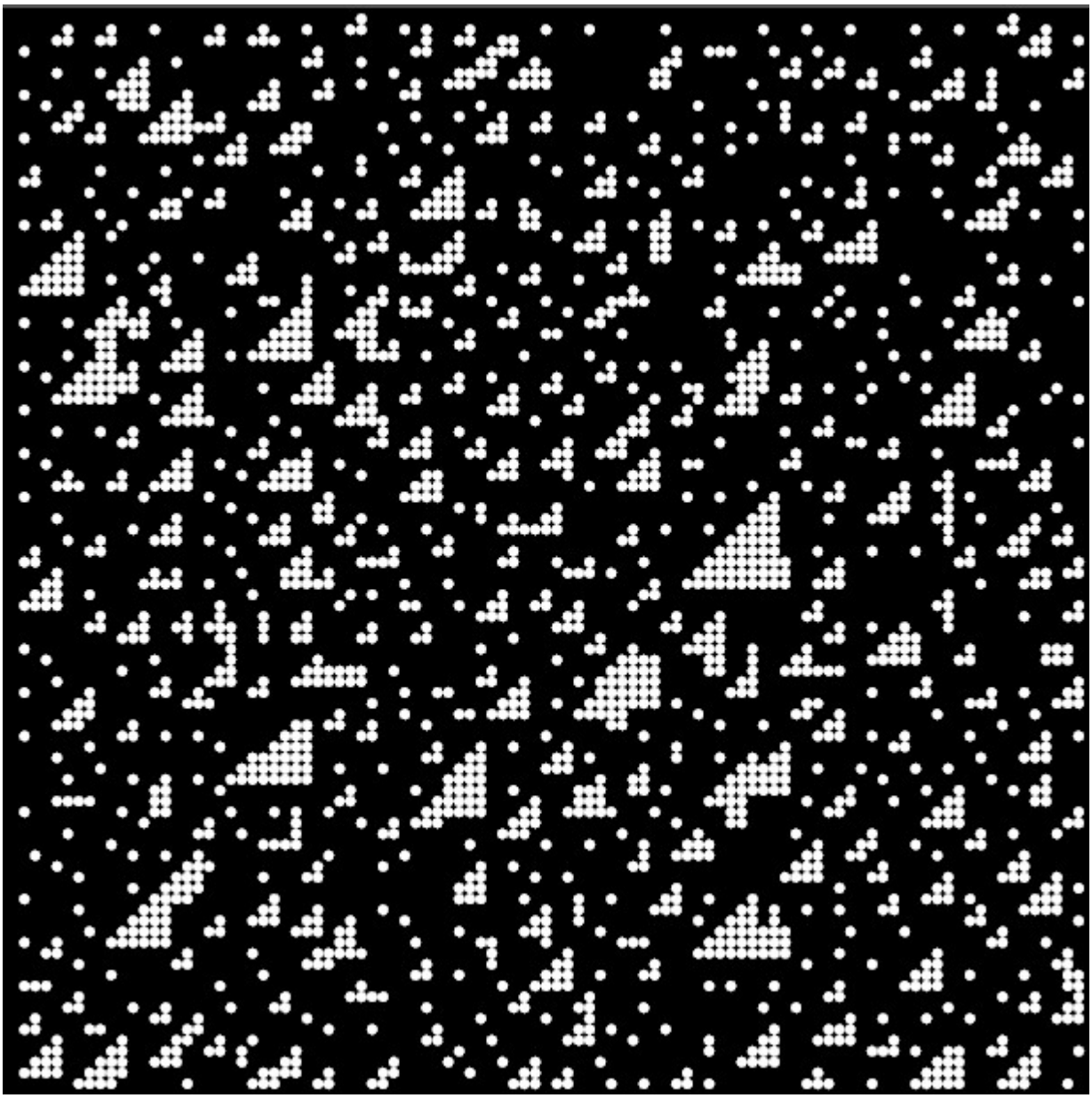

Figure 35: The Sierpinski Tileset at temperature 1 displays a number of interesting properties. Triangle-like shapes are still constructed with some regularity, but a large number of errors and restarts are also seen.

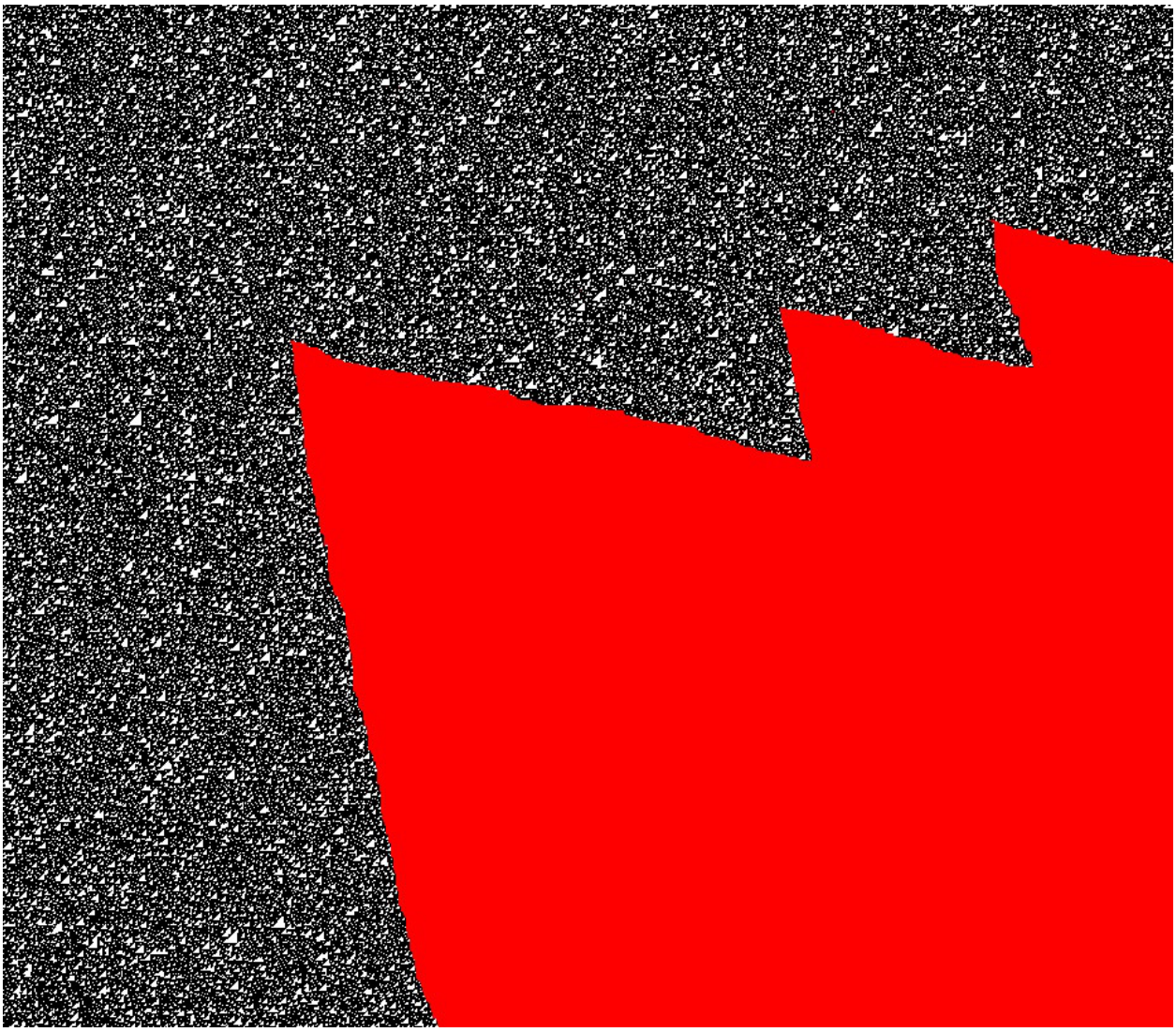

Figure 36: The Sierpinski Tileset with a non-trivial chance of NULL attachment at temperature 1, where black and white tiles are from the tileset and NULL tiles are red

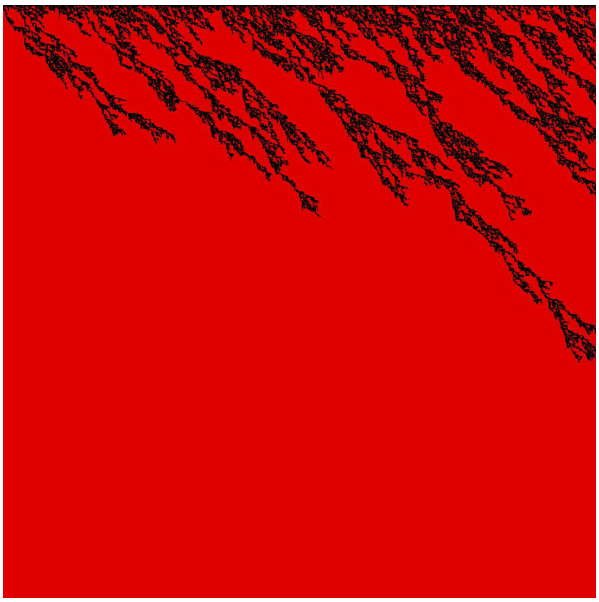

Figure 37: Sierpinski Tileset where black represents any tile from the tileset and where $p < p_c$

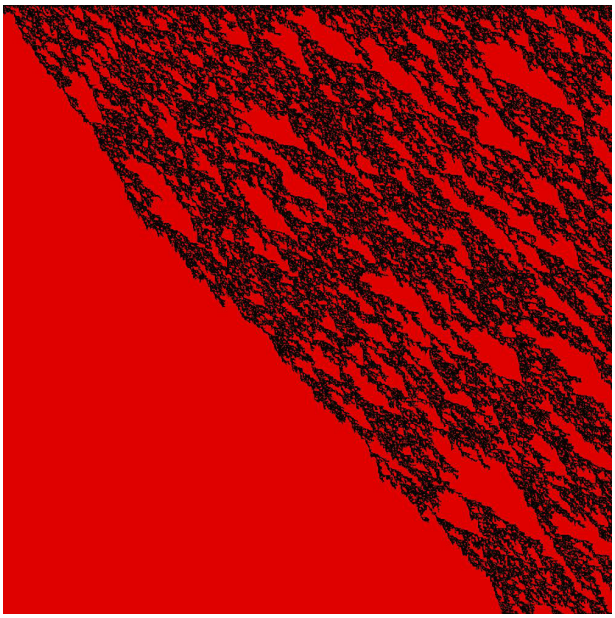

Figure 38: Sierpinski Tileset where black represents any tile from the tileset and where $p = p_c$

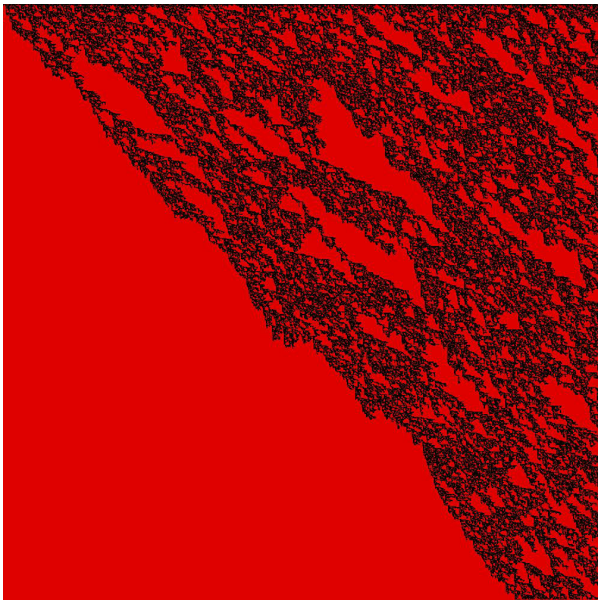

Figure 39: Counterquilt Tileset where black represents any tile from the tileset and where $p > p_c$

4.5 Binary Counter Tileset

The Binary Counter Tileset displays correlation properties similar to the Sierpinski Tileset, as seen in figure 40. As the chance of non-NULL attachment approaches 100%, the correlation between the assemblies approaches 100%, but the steep transition seen in previous tilesets is absent. Assembly at temperature 1.5 yields critical probabilities of temperature 2 attachment at approximately 76% and non-NULL attachment probability at 87%. Figures 41, 42, and 43 show the behavior of the Binary Counter Tileset below, at, and above this critical percolation probability.

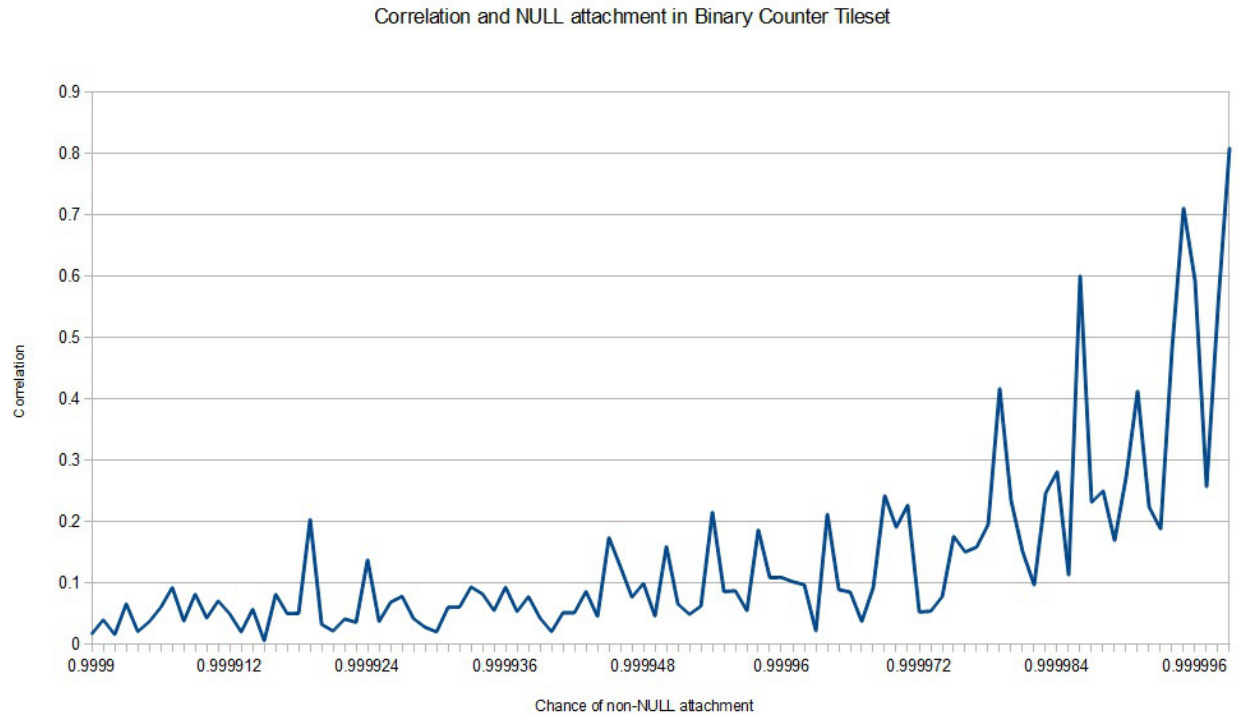

Figure 40: The correlation between a temperature 2 assembly without NULL tiles and with NULL tiles in the Binary Counter Tileset

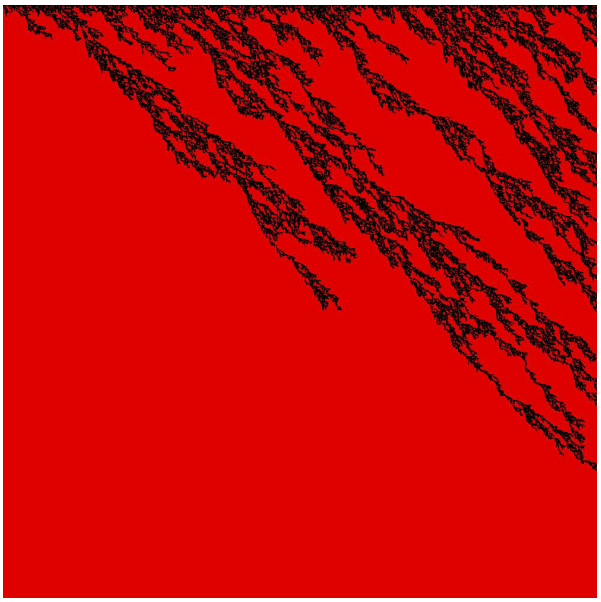

Figure 41: Binary Counter Tileset where black represents any tile from the tileset and where $p < p_c$

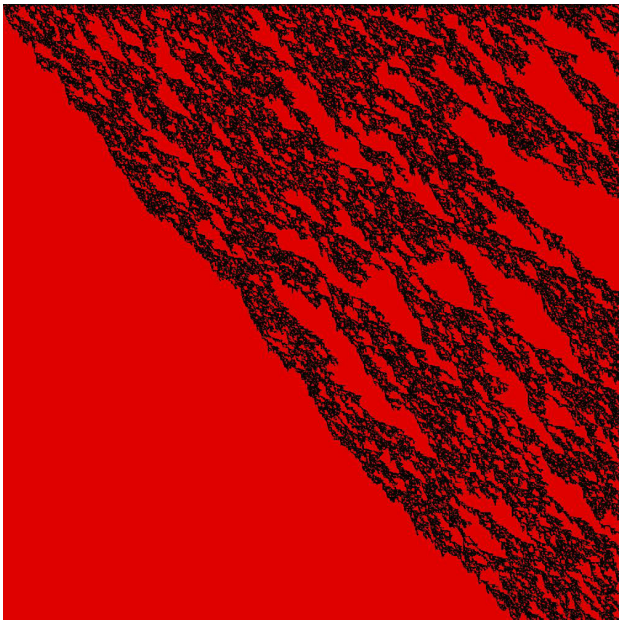

Figure 42: Binary Counter Tileset where black represents any tile from the tileset and where $p = p_c$

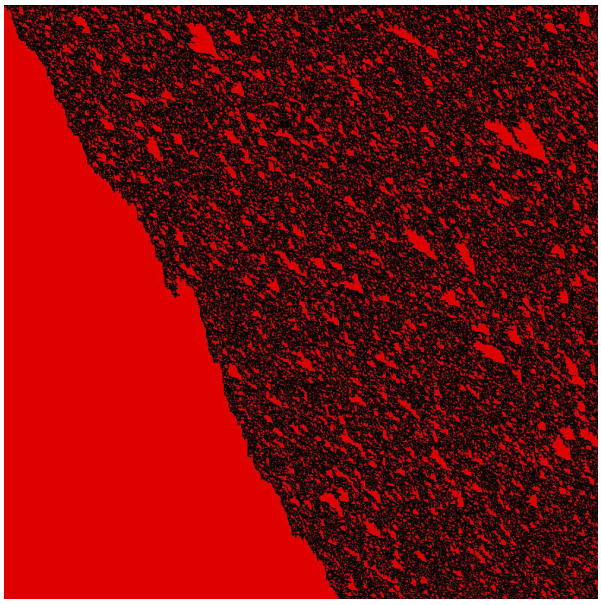

Figure 43: Binary Counter Tileset where black represents any tile from the tileset and where $p > p_c$

4.6 Analysis

Each tileset assembled in very large configurations (1000 tiles by 1000 tiles) shows an interesting behavior at temperature 1.5. At practically any probability of temperature 1 behavior, the system adopts predominantly temperature 1 characteristics (uncontrolled growth which does not follow the intended program for the tileset) and does not strongly correlate with the temperature 2 assembly. As the probability of temperature 2 attachment approaches 1, the behavior of the system at temperature 1.5 approaches temperature 2 suddenly. This behavior happens in such a sharp and predictable way that it seems to represent a phase transition between system temperatures. This result indicates that temperature 2 is difficult to achieve in any real system. Although behavior of the aTAM at temperature 2 is well-described, behaviors in temperature 1 systems are not. The computational power of temperature 1 systems in general is difficult to determine because non-determinism expands the local determinism seen in temperature 2 systems. Instead of a 1-to-1 correspondence between tiles and sites, temperature 1 systems exhibit a k-to-1 correspondence between k tiles with partial matches in their binding domains where $k > 0$. An example of this is seen in temperature 1.5 assemblies of the Double Glue tileset. In a tileset with few glues, the assembly can be so unconstrained that specific configurations are very difficult to achieve. Therefore, the fact that temperature 1 behaviors are seen even when allowing a very large portion of total attachments to be temperature 2 means further analysis of temperature 1 assemblies is necessary. Even a well-designed tileset will suffer losses in total yield if temperature 2 behavior is expected, but not seen in the experimental system.

In addition to the number of glues, the quantities of each tile used in the system is also relevant. For systems where some tiles are used in much higher quantities than others, correlation can be misleading. For example, both the Sierpinski Tileset and the Binary Counter Tileset use a disproportionately large number of tiles from one tile type, seen most explicitly in figure 12. Each white ellipse represents a tile from a single tile type. The correlation measure simply checks location (i, j) in both configurations, if the tile name is the same, the correlation percentage increases. Therefore, if a relatively large number of tiles from a single tile type appear in a configuration, correlation is positively biased. Such bias could contribute to the behavior seen at very high chances of non-NULL attachment shown in figures 34 and 40. The phase transition isn't as sharp as it is in the Single Glue Tileset and Double Glue Tileset.

Similarities to directed percolation are seen at temperature 1.5 for tilesets which would construct infinite configurations given infinite space and infinite quantities of each tile type. When the NULL tile is introduced at temperature 1.5, with critical probabilities consistent with those found through simulation of directed percolation in this and other models, tile assemblies generate probable infinite clusters of non-NULL tiles. As the probability of temperature 2 and non-NULL attachment increases, these clusters begin to saturate the configuration until a perfect assembly is replicated at probability 1 for temperature 2 attachment with no probability for NULL attachment. Site percolation is easily translated into a tile assembly problem. In the Single Glue Tileset with NULL tiles at temperature 1, the DK model site percolation probability is nearly spot on. In configurations of 1,000,000 tiles, clusters spanning the entire lattice begin to form reliably at around $p_c = .705$ for the chance of attaching non-NULL tiles. These early results seem promising in successfully relating directed percolation and tile assembly. Further study is

needed to better articulate these findings for shapes other than the relatively simple n by n square and for tilesets other than the Single Glue Tileset. Experimentally, some loose critical probabilities were found for the Counter Quilt, Sierpinski, and Binary Counter Tilesets, but further trials are needed to evaluate and refine these initial findings.

The fundamental shift inherent in temperature 1.5 requires more investigation and deeper analysis, but these findings indicate some interesting properties. As discussed earlier, a temperature 1.5 system has some probability of attachment at temperature 2 and some probability of attachment at temperature 1. This is simulated by giving a probability t of attachment at temperature 2 and defining the system's probability of temperature 1 attachment as $(1 - p)$. Two key results indicate that temperature 1.5 could help future tile assembly systems more carefully approximate laboratory assembly processes while also creating new avenues for investigation regarding the computational complexity of temperature 1 assembly.

A system operating at temperature 1.5 can represent a system at temperature 2 or 1 depending on the value given for t. As t approaches 1, there is a predictable change in large assemblies which indicates a phase transition. This shift from temperature 1 behavior to temperature 2 behavior is over a relatively miniscule threshold, as seen in figures 20, 30, 34, and 40 and reveals a fundamental difference between temperature 1 and temperature 2 systems which assemble finite structures. Although the aTAM is a great model for describing particular assemblies under set conditions, it does not generalize well to procedures outside of those conditions. It is difficult to describe and assemble a material with particular global properties without articulating those global properties through locally deterministic attachment procedures. Relevant adaptations to the aTAM, such as the pTAM, are a step in the right direction, but

further work in this area is required. Current design principles can yield complex configurations with guaranteed global behaviors, but at the cost of yield. As configurations become more geometrically complex (measured by their Kolmogorov complexity), the chance of error during assembly grows. By developing tilesets that are resilient despite temperature 1.5, researchers can guarantee global properties despite unfavorable local interactions.

Global attributes like conductivity, hydrophobicity or -philicity, and optical qualities like transparency can be described by global behaviors and need not involve universally correct, locally deterministic component attachment. By keeping a hand on the dial, so to speak, a researcher could tune a tileset to a broad field of attachment probabilities and describe a range or class of acceptable configurations (similar to the probabilistic assembly of rulers of length N seen in the pTAM). Since temperature 1.5 is easy to implement in tile assembly simulations, this tuning can occur without the use of costly materials. Once broad changes have been made, finer adjustments can be made *in vitro* which can then be fed back into the simulation in order to refine the model.

The use of NULL tiles can benefit these analyses. By abstracting failed attachments into a catch-all tile type, the desired global properties of an assembly can be iteratively simulated until they fall within acceptable margins. Focusing on directed percolation and the behavior of clusters on the finite lattice, the global concentration of correct tiles can easily be determined at certain probabilities in a particular configuration. Since the critical probability of producing infinite clusters can now be described by directed site percolation, models of finite-size systems in the literature can be employed to analyze similarly sized finite configurations. Conductivity in particular is a global property which has already been addressed in percolation theory literature

[6]. A lattice structure with a certain expectation of conductivity can be more reliably built and tested within the context of previous findings in studies on finite percolating lattices. Also, consider a self-assembling superstructure made via staged self-assembly. By layering the assembly process and constructing stages of supertiles with global properties, researchers can more easily guarantee global properties of the total assembly by guaranteeing the properties of the component supertiles which are less complex and easier to assemble.

Also, the representation of tile assembly as a directed percolation process (particularly directed site percolation) is a helpful first step in relating well-described statistical mechanical models of systems with phase transitions to self-assembly, even outside of the aTAM. Results shown here indicate that assembling n by n squares of arbitrarily large size can be represented by a directed site percolation problem on that same square lattice. The simple translation between the two problems seems to indicate a promising and elegant relationship between the two processes which could prove fruitful in solving self-assembly systems for certain global properties. Investigations into directed lattice animals and their geometric relationships to arbitrarily organized configurations could also prove fruitful now that initial simulations have shown a connection between directed percolation and the aTAM.

In addition, an absorbing state once reached cannot be left (by the nature of directed percolation), which indicates that tiling fidelity is difficult to achieve under non-ideal conditions, where non-ideal conditions are taken to be non-temperature 2 systems which can undergo errorful aggregation. Even specially designed error-correcting tilesets can theoretically enter absorbing states where errors are so numerous that the system cannot recover. When taken with the DKCA tileset presented in figure 19, this result indicates that directed percolation models are

well adapted to describe self-assembly in the aTAM . Also, figure 19 describes a tileset which is capable of sweeping out the phase diagram described by figure 18, which may lead to additional CA which can program tile assemblies for bulk or global statistical properties.

5 CONCLUSIONS

5.1 Findings

Each tileset constructs the expected configuration with perfect accuracy in a temperature 2 system with no NULL tiles. This behavior is expected in the aTAM. Conversely, tilesets which include NULL tiles very rarely assemble the expected configuration. The NULL tile halts the assembly process, and the temperature of the system prevents further attachments from occurring. As seen in figure 44, at temperature 2, a NULL tile attachment can derail an entire assembly. Note that although *t* has a partially correct binding domain, it cannot attach, even if it is the next expected tile in the configuration, ignoring the NULL attachment.

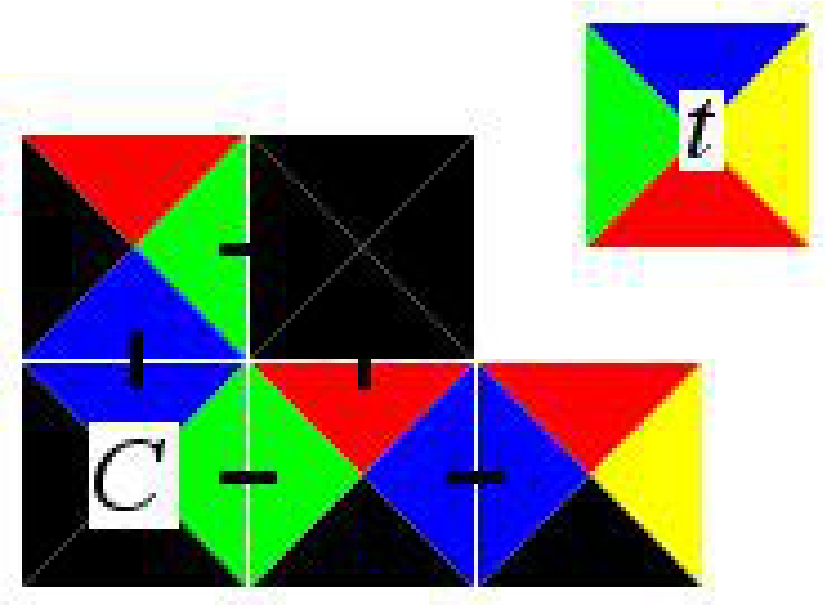

Figure 44: Tile *t* attempts to attach to configuration *C*

As the probability of a NULL tile attachment approaches zero, the correlation between the NULL allowance configuration and the regular temperature 2 configuration approaches unity. At temperature 1, NULL attachment isn't as detrimental to the continued assembly, but correlation

remains consistently low. This is due to the less constrained attachment process which both allows the aggregate to correct a NULL attachment while also allowing unexpected tile attachment on open binding sites.

Results also indicate that percolation and self-assembly by the aTAM are related. Simulations of the Single Glue Tileset and the Double Glue Tileset show that infinite clusters develop at approximately the same critical percolation values in each, although the non-deterministic nature of attachment generated variant clusters in the Double Glue Tileset which were impossible to relate via correlation at any temperature. Furthermore, these critical percolation values are similar to critical percolation values for the DK model of directed percolation. The relationship between percolation and self-assembly by the aTAM is clearly shown by these results, but further simulation of designed tilesets with particular final configurations, like the Sierpinski and Counter Quilt tilesets, indicates that when a tileset is crafted such that a specific configuration is expected in a temperature 2 tile assembly system, infinite clusters are developed at critical percolation values higher than those seen with the Single Glue and Double Glue Tilesets. This behavior is a direct result of the asymmetry in binding domains seen in these tilesets; as the set of possible binding domains becomes larger and less regular, opportunities for new tiles to attach become more rare. This is due to the dependence on the variable temperature system for a particular temperature during a particular attachment attempt.

Temperature 1.5 allows for variable system temperature determined probabilistically at attachment time for each new attachment and can be used to more accurately predict how a tileset might assemble in a laboratory environment. This is due to the pervasive and unavoidable

nature of attachment errors which are ignored in temperature 2 aTAM tile assembly systems. Simulations of temperature 1.5 behavior indicate that non-trivial probabilities of temperature 1 attachment in the tile assembly system tend to cause the entire system to behave as though it were temperature 1 when correctness is measured by correlation to a temperature 2 configuration. When considering tilesets like the Sierpinski or the Counter Quilt, this behavior becomes more clear, as seen in figure 36.

The NULL tile has an interesting effect on assembly and is necessary to distinguish directed percolation behavior in the chosen tilesets. Without a NULL tile, assembly will halt when no tile can attach, but allowing NULL attachments makes the tile assembly process more robust for temperature 1 or temperature 1.5 tile assembly systems. The addition of the NULL tile creates another phase transition in correlation percentage for non-trivial NULL attachment rates which decays as the tileset integrates more glues and tile types (such as the Sierpinski and Binary Counter Tilesets). This decay can be seen by comparing the sharp transition seen in the Single Glue Tileset in figure 20 to the shallower and more scattered transition in the Sierpinski tileset and Binary Counter Tileset (figures 34 and 40). The nature of these tilesets also contributes to the decay, since both the Sierpinski and Binary Counter do not use all tile types equally.

When the NULL tile is allowed and the temperature of the system is set to temperature 1.5, strong directed percolation behavior is seen. This is most clear when examining configurations of the Single and Double Glue Tilesets below, at, and above the critical probabilities found through simulation. Essentially the NULL tile and temperature 1.5 increase the flexibility of the aTAM by reducing or removing constraints on attachment which would cause the assembly to stop attaching new tiles.

In conclusion, simulation results show the NULL tile to be a necessary addition to the aTAM where designed tilesets or temperature 1.5 are parameters of the tile assembly system and the constraints imposed by the aTAM on assembly are too binding. When the NULL tile and temperature 1.5 are allowed with some probability, strong directed percolation behavior becomes evident. These simulation results support the formal equivalence of self-assembly and directed percolation given in section 3.2.8. Additionally, well-known critical probabilities for percolation in the DK model for cellular automata match critical probabilities for percolation seen in simulations of two distinct tilesets.

5.2 Future Work

As mentioned previously in section 4, there are a number of avenues of investigation which need further development. With regards to the experimental applicability of the author's conclusions concerning the usefulness of temperature 1.5, the first step is to articulate minimal tilesets which are inclined to construct configurations with predictable global properties, a feat in its own right. Perhaps the Single Glue and Double Glue Tilesets hold the key to such investigations. Both tilesets can be used to represent random media where some components or elements of the medium have certain desirable properties that other components do not have. If this property is binary, then such random media can be represented by the Single Glue tileset with a NULL tile. Properties with up to 16 categorical values can be represented with the Double Glue Tilesets, although the directed bond percolation problem would need to be addressed probabilistically with such a large number of values. Potentially, if a tileset can arbitrarily represent a medium that follows some functional distribution of its components, and that tileset

is minimal or close to minimal, then a whole new class of system can be represented with the aTAM which does not rely on local determinism except as a constraint which aids assembly.

Also, the connection between directed percolation and the aTAM must be formalized. The mathematical connection between the aTAM and directed percolation or other statistical mechanical models must be developed and articulated for researchers looking to simulate such systems. Since self-assembly has a strong connection to DK automata and directed percolation shares that connection, perhaps the beginnings of this association could be made there. Additionally, directed percolation gives researchers access to a number of potentially helpful resources, including the concept of directed lattice animals. A lattice animal is a collection of occupied sites adjacent by open bonds and is the result of a percolation process. Lattice animals are defined by a perimeter and area which are calculated from the statistical and geometric properties of the animal. In the aTAM with NULL tiles, configurations of non-NULL tiles can be thought of as lattice animals. By relating these aggregates or clusters in a configuration to lattice animals, new metrics and workflows are introduced to the process of self-assembly which can be used to improve design and assembly in simulations and in the laboratory by giving researchers more tools to describe and investigate self-assembled products.

REFERENCES

[1] M. Gómez-López *et al.* “The art and science of self-assembling molecular machines.” *Nanotechnology*, vol. 7, pp. 183-192, Sept. 1996.

[2] F. Li, D. P. Josephson and A. Stein, "Colloidal assembly: The road from particles to colloidal molecules and crystals,” *Angewandte Chemie Int. Edition*, vol. 50, no. 2, pp. 360-388, Jan. 2011.

[3] L. Adleman *et al.*, “The undecidability of the infinite ribbon problem: Implications for computing by self-assembly,” *SIAM J. Comput.*, vol. 38, no. 6, pp. 2358-2381, 2009.

[4] W. B. Sherman and N. C. Seeman, "A precisely controlled DNA biped walking device,” *Nano Lett.*, vol. 4, no. 7, pp. 1203-1207, July 2004.

[5] E. Winfree, “Algorithmic Self-Assembly of DNA,” Ph.D. dissertation, Dept. Computation and Neural Systems and Bioeng. California Inst. of Technology, Pasadena, CA, 2008.

[6] D. Stauffer and A. Aharony, *Introduction to Percolation Theory*. Bristol, PA: Taylor & Francis Inc., 1992, pp. 1 – 171.

[7] D. Soloveichik and E. Winfree, "Complexity of self-assembled shapes,"*SIAM J. Comput.*, vol. 36, no. 6, pp. 1544-1569, 2007.

[8] H. Hinrichsen, "Non-equilibrium critical phenomena and phase transitions into absorbing states," *Advances in Phys.*, vol. 49, no. 7, pp. 815-958, 2000.

[9] E. Domany and W. Kinzel, “Equivalence of Cellular Automata to Ising Models and Directed Percolation,” *Phys. Rev. Lett.*, vol. 53, no. 4, pp 311-314, July 1984.

[10] G. Aggarwal *et al*, "Complexities for generalized models of self-assembly," *SIAM J. Comput.*, vol. 34, no. 6, pp. 1493-1515, 2005.

[11] H. Chandran, “Tile complexity of linear assemblies,” *SIAM J. Comput.*, vol. 42, no. 4, pp. 1051-1073, Aug. 2012.

[12] D. Doty *et al*, “Strong fault-tolerance for self-assembly with fuzzy temperature,” *FOCS 2010: Proc. of the 51st Annu. IEEE Symp. on Found. of Comput. Sci.*, Nevada, USA, pp 417–426, 2010.

[13] P. W. K. Rothemund and E. Winfree, "The program-size complexity of self-assembled squares," *Proc. of the Thirty-Second Annu. ACM Symp. on Theory of Computing*, pp. 459-468. ACM, 2000.

[14] D. Soloveichik and E. Winfree, "Complexity of self-assembled shapes,"*SIAM J. Comput.*, vol. 36, no. 6, pp. 1544-1569, 2007.

[15] A. V. Pinheiro *et al*, "Challenges and opportunities for structural DNA nanotechnology," *Nature nanotechnology*, vol. 6, pp. 763-772, Nov. 2011.

[16] E. D. Demaine *et al*, "Staged self-assembly: Nanomanufacture of arbitrary shapes with *O*(1) glues," *Natural Comput.* vol. 7, no. 3, pp. 347-370, Sept. 2008.

[17] J. W. Essam, "Percolation theory," *Rep. Prog. Phys.,* vol. 43, no. 7, pp. 833-913, July 1980.

[18] J. L. Cardy and R. L. Sugar, "Directed percolation and Reggeon field theory," *J. Phys. A: Math. Gen.*, vol. 13, no. 12, pp. L423, Nov. 1980.

[19] R. C. Brower *et al*, "Reggeon field theory on a lattice: A formulation," *Nucl. Phys. B*, vol. 94, no. 3, pp. 460-476, Aug. 1975.

[20] S. Redner, "Flory theory for directed lattice animals and directed percolation," *J. Phys. A: Math. Gen*, vol. 15, no. 6, pp. L273-L278, June 1982.

[21] I. Jensen and H. C. Fogedby, “Critical exponents for an irreversible surface reaction model,” *Phys. Review A*, vol. 41, no. 6, pp. 3411-3414, March 1990.

[22] J.-F. Markert, "Directed animals, quadratic and rewriting systems,"*arXiv preprint arXiv:1112.0910* (2011).

[23] M. Z. Bazant, "Largest cluster in subcritical percolation," *Physical Review E* vol. 62, no. 2, Aug. 2000, pp. 1660-1669.

[24] J. Machta, C. M. Newman and D. L. Stein, "A Percolation-theoretic Approach to Spin Glass Phase Transitions," in *Spin Glasses: Statics and Dynamics*, A. Boutet de Monvel and A. Bovier, Eds.Basel, Switzerland: Birkhäser Verlag AG, 2009, pp. 205-223.

[25] Q. Cheng *et al*., “Optimal self-assembly of counters at temperature two.” in *Proceedings of the first conference on Foundations of Nanoscience: Self-assembled architectures and devices*, 2004.

[26] P. W. K. Rothemund, N. Papadakis and E. Winfree, "Algorithmic self-assembly of DNA Sierpinski triangles," *PLoS biology*, vol. 2, no. 12, pp. 2041-2053, Dec. 2005.

[27] J. I. Lathrop, J. H. Lutz and S. M. Summers, "Strict self-assembly of discrete Sierpinski triangles," *Theoretical Comput. Sci.*, vol. 410, no. 4, pp. 384-405, Feb. 2009.

[28] H. Wang, “Proving theorems by pattern recognition—II,” *Bell Systems Technical Journal*, vol. 40, pp. 1–41, 1961.