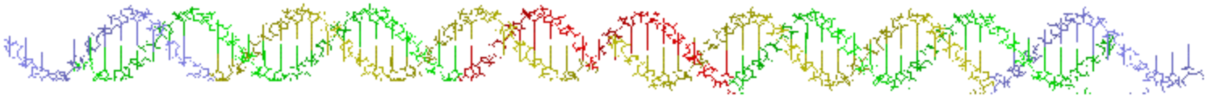
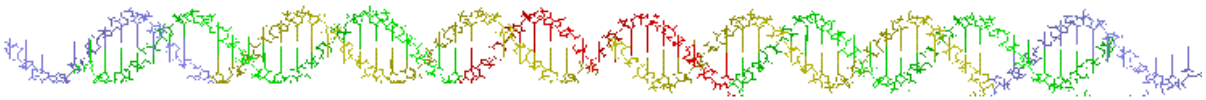


Intro to Tiling Assembly



John Reif

From PPT by Ho-Lin Chen, Ashish Goel, Tianqi Song

What is Molecular Self-assembly?

Self-Assembly is the process by which *simple* objects *autonomously* assemble into complexes.

- **In Nanoscience, Self-assembly is the spontaneous formation of a complex by small (molecular) components under simple combination rules**
- **Geometry, dynamics, combinatorics are all important**
- **Inorganic: Crystals, supramolecules**
- **Organic: Proteins, DNA**

Goals:

- **Understand self-assembly,**
- **Design self-assembling systems**
- **Applications to nano-technology, molecular robotics, molecular computation**

A Matter of Scale

- **Question:** Why a mathematical study of "molecular" self-assembly specifically?
- **Answer:** The scale changes everything
  - Consider assembling micro-level (or larger) structures such as robotic swarms. We can attach rudimentary computers, motors, and radios to these structures.
    - Can now implement an intelligent distributed algorithm.
  - In molecular self-assembly, we have nano-scale components. No computers. No radios. No antennas. We **need self-assembly** to make computers, radios, antennas, motors.
    - Local rules such as "attach to another component if it has a complementary DNA strand"
  - Self-assembly at larger scales is interesting, but is more a sub-discipline of distributed algorithms, artificial intelligence etc.

Applications of Self-Assembly

- **Building blocks of nano-machines.**

- **DNA computing.**

- **Small electrical devices such as FLASH memory. [Black et. al. 2003]**

- **Nanostructures which "steer" light in the same way computer chips steer electrons. [Percec et. al. 2003]**

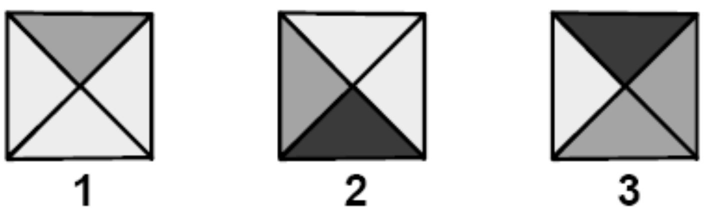
Wang Tiling

- **Proposed in** [Hao Wang, Proving theorems by Pattern Recognition II, 1961].
- Class of formal tiling systems
- Tiles:
  - Given a finite set of square tiles with a glue on each side.
  - Tiles are modeled visually by squares with a color (glue type) on each side
  - The tiles cannot be rotated or reflected and
  - You can use infinite number of copies of each tile.
- Tiling Question: whether they can tile the plane with same abutting glue.

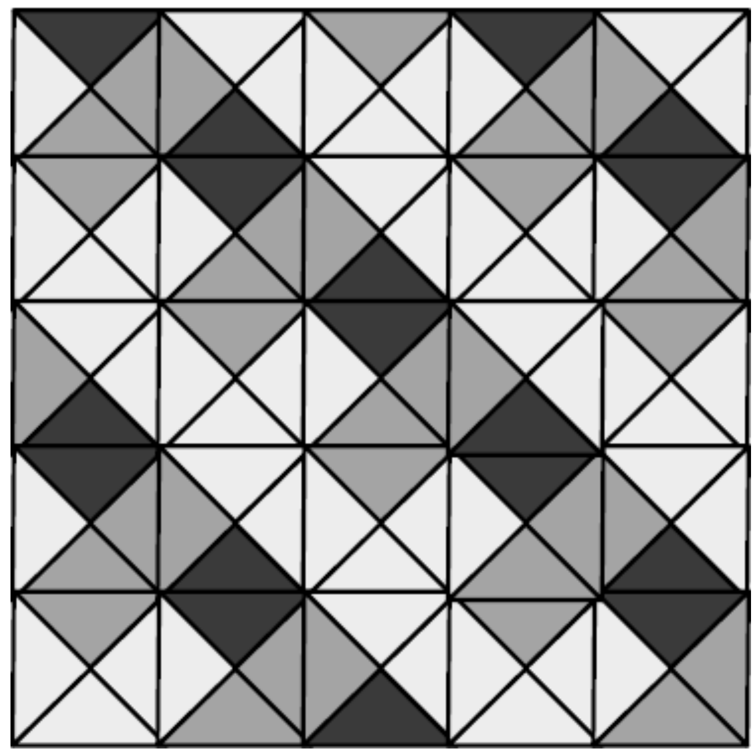
Wang Tiling Problem:

Given a set of tiles, can copies of the tiles be arranged one by one to fill an infinite plane such that adjacent edges of abutting tiles share the same color ?

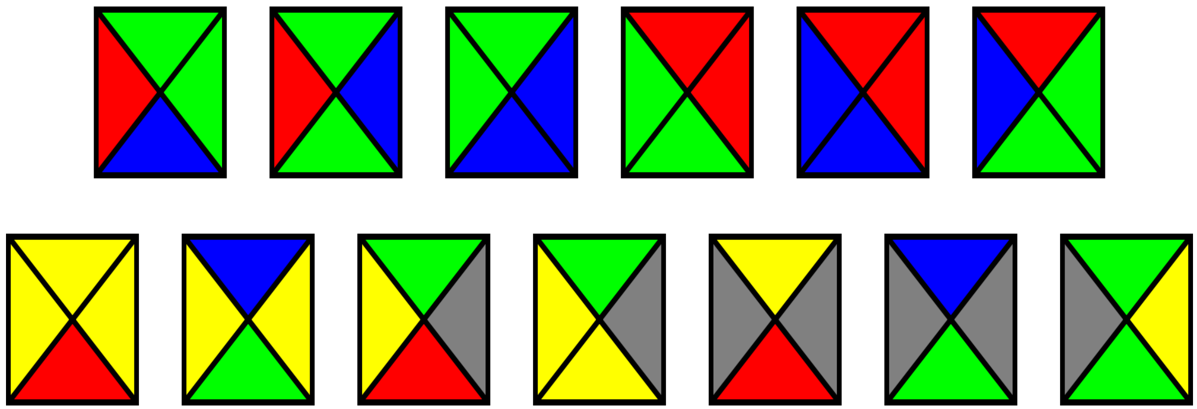
3 tile example:



Source: Savi Maharaj



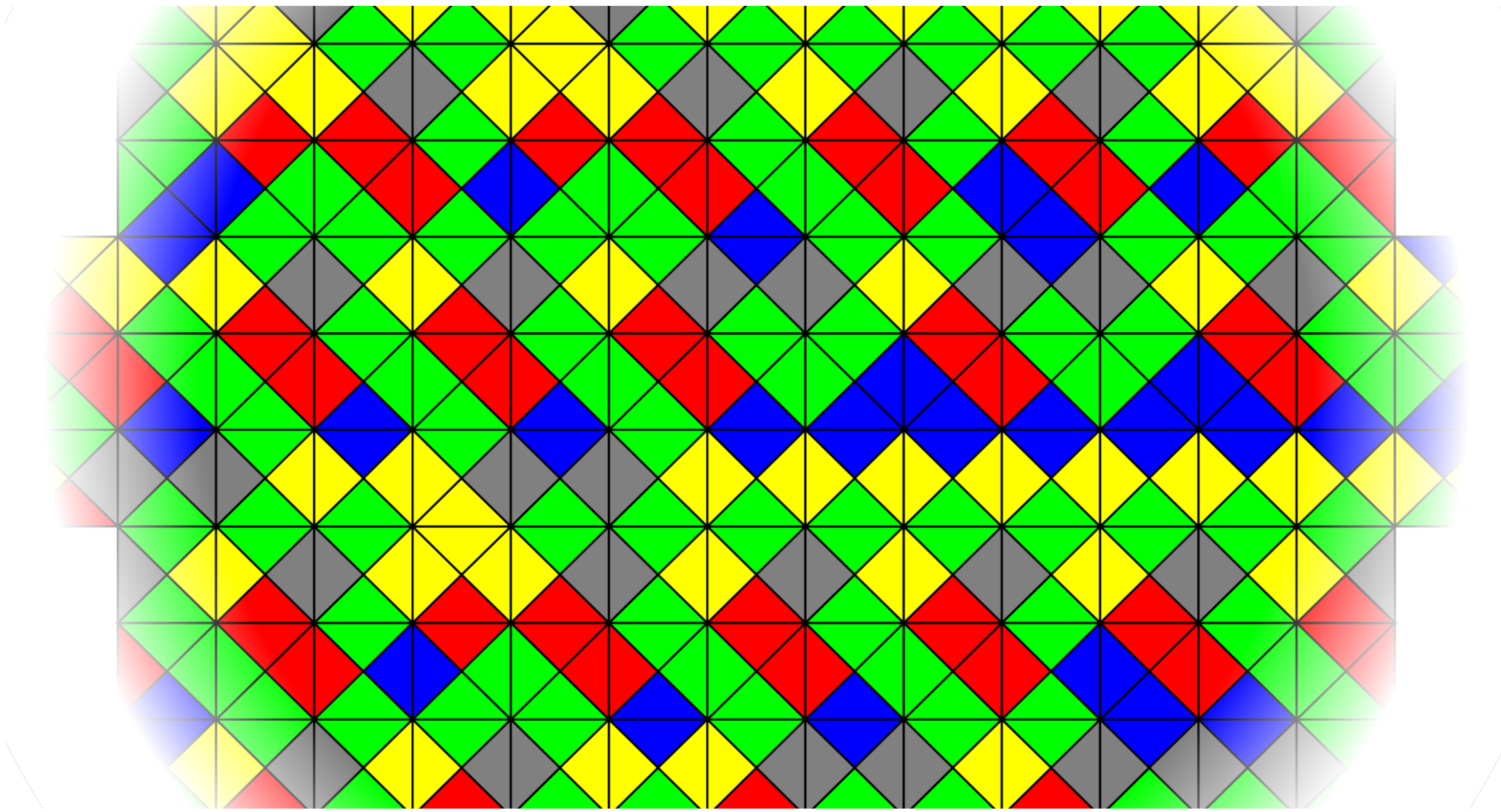
Example of Wang Tiling



Designed by ***[Karel Culik, 1996].***
[Wikipedia]

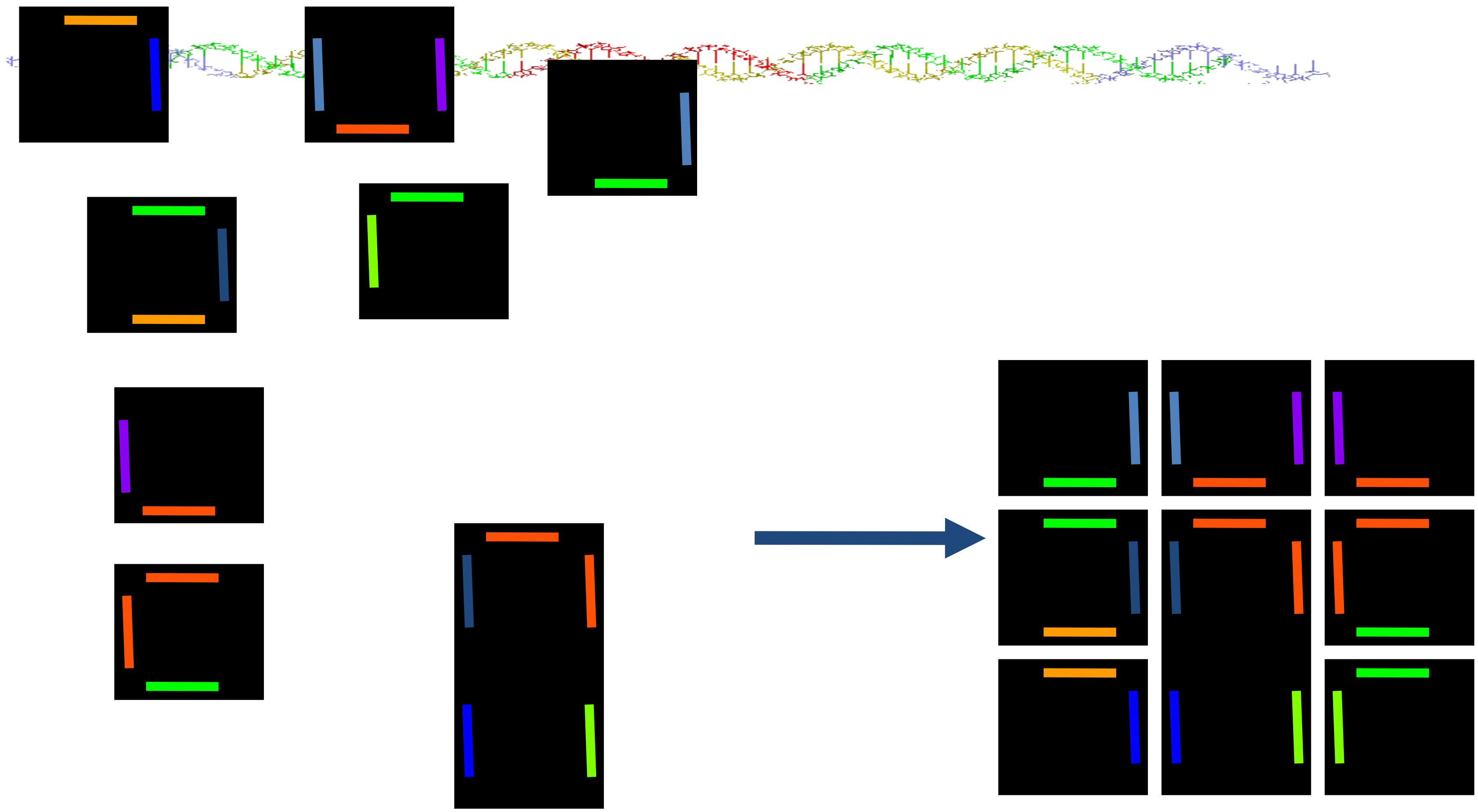
- **This tile set contain 13 tiles.**
- **They can tile the plane aperiodically as shown in next page.**

Example of Wang Tiling



[Wikipedia]

Wang Tiling Construction with "Smart Bricks"



A tiling assembly using `Smart Bricks' with affinity between colored pads.

Is Self-Assembly Just Crystallization?

- **No. Crystals do not grow into unique terminal structures!**
  - **A sugar crystal does not grow to precisely 20nm**
- **Crystals are typically made up of a small number of different types of components**
  - **Two types of proteins; a single Carbon molecule**
- **Crystals have regular patterns**
  - **Computer circuits, which we would like to self-assemble, don't**
- **Self-assembly = combinatorics + crystallization**
  - **Can count, make interesting patterns**
  - **Nature doesn't count too well, so molecular self-assembly is a genuinely new engineering paradigm.**
  - **Think engines.**
  - **Think semiconductors.**

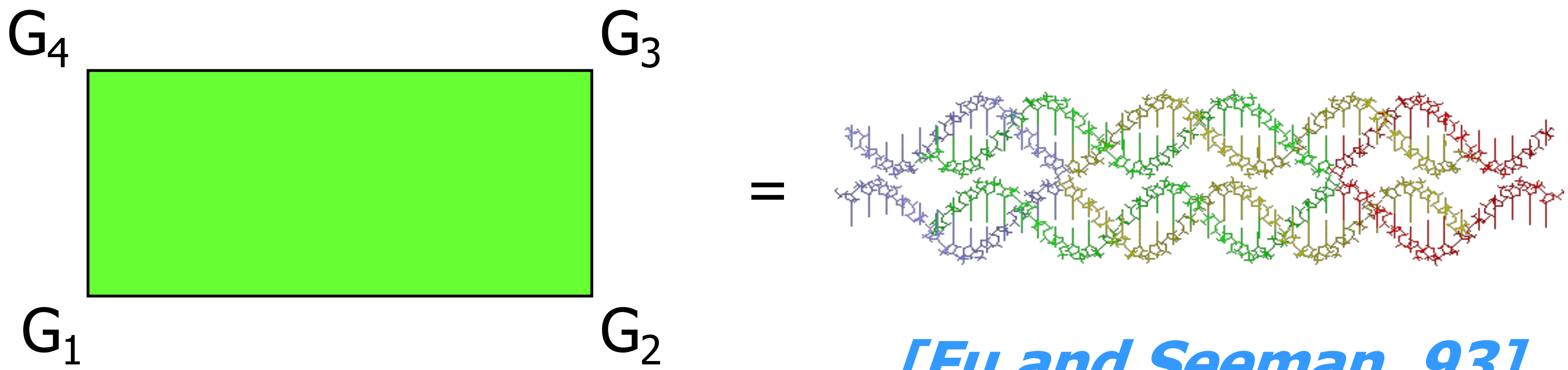
Undecidability of tiling problem

Proof of Undecidability and Nonperiodicity for Tilings of the Plane: ***[Robinson, Undecidability and Nonperiodicity for Tilings of the Plane, 1971]:***

- **Encoded the moves of a Turing machine into a set of Wang tiles, such that:**
- **The Wang tiles can tile the plane if and only if the Turing machine will never halt.**

=> a tiling system can compute any computable function!

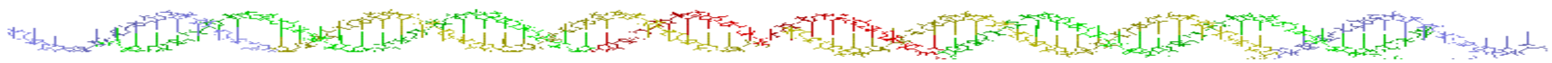
DNA Tiles

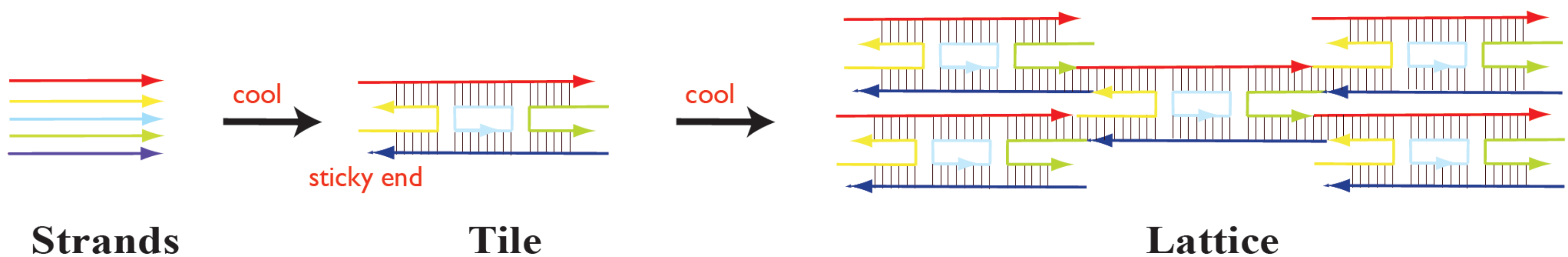


[Fu and Seeman, 93]
[Winfree and Seeman]

Glues = sticky ends
Tiles = molecules

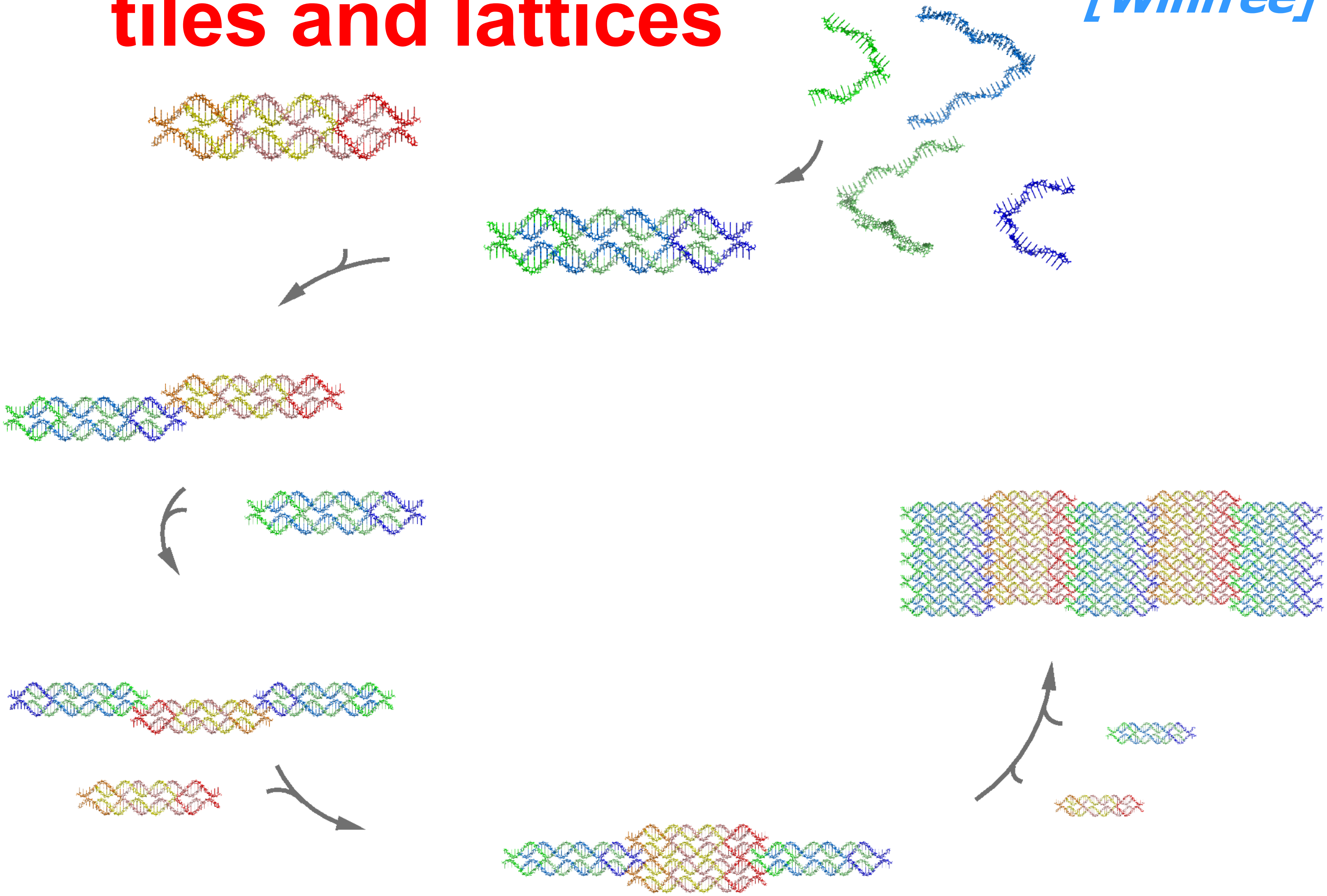
Self-assembly of DNA tiles into Lattices





Self-Assembly of DNA into tiles and lattices

[Winfree]



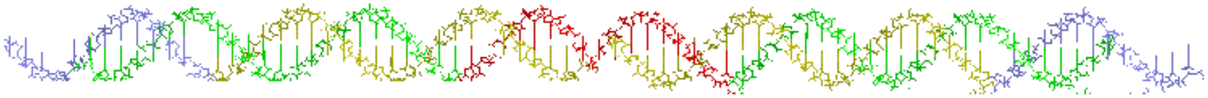
DNA Self-Assembly

Trying to "compute" using protein tiles would be challenging:

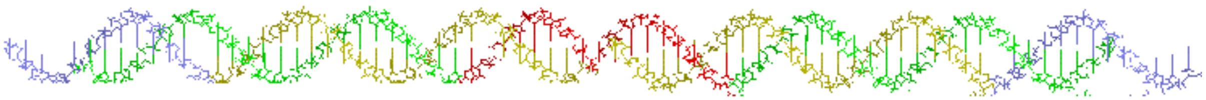
- Proteins have a complicated geometry and
- it is hard to predict what shape a single protein will take.

Instead, we assume that tiles are made of DNA strands hybridized together, and that the glues are really single-stranded DNA strands

- **DNA is combinatorial: the functionality of DNA is determined largely by the sequence of ACTG bases.**
- **Proof-of-concept from nature: A DNA strands can hybridize to another complementary DNA sequences**
- **DNA tiles have been constructed in the lab, and DNA computation has been demonstrated**



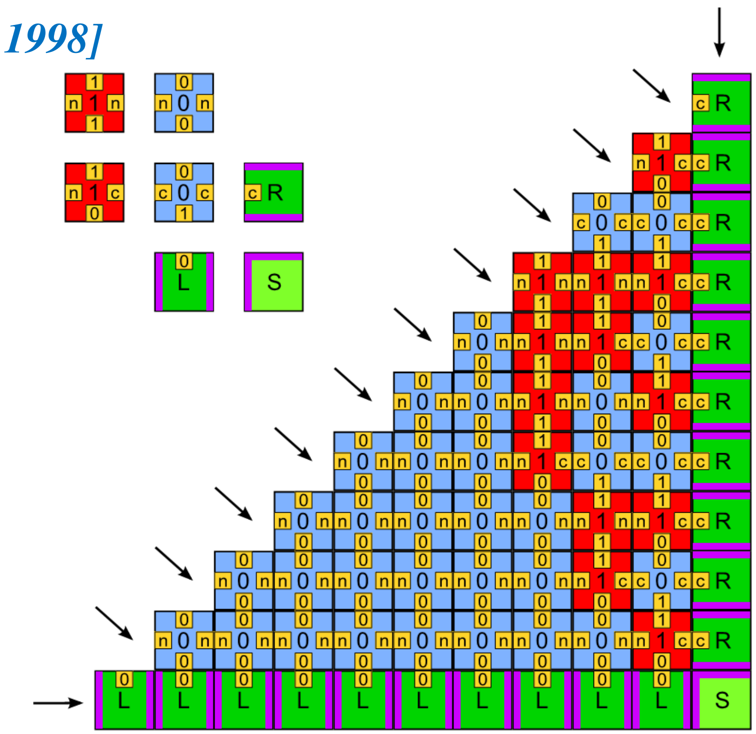
Abstract Tiling Model (ATAM)



Tile assembly model (TAM)

- **Proposed by Erik Winfree developing on Wang tilings**

  - *[Winfree: Simulations of Computing by Self-Assembly, 1998]*

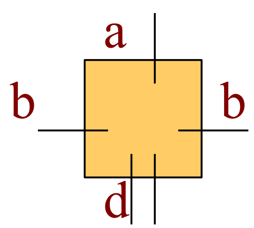


- **Simple, yet powerful model**
- **Refines Wang tiling**
- **Models crystal growth**
- **Also, Turing-complete**
- **Can be implemented using DNA molecules**

The Tile Assembly Model (ATAM)

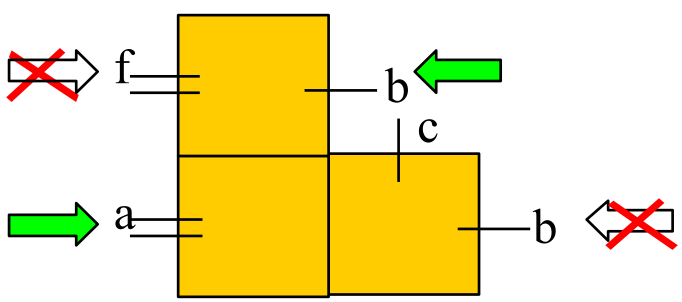
Oriented Tiles with a glue on each side

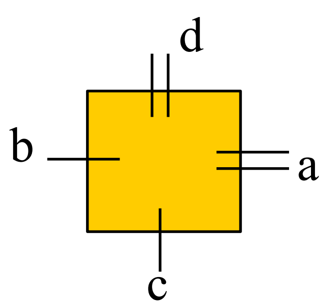
- Each glue is labeled by a strength



Single bar: strength 1 glue
Double bar: strength 2 glue

- Tiles floating on an infinite grid; Temperature τ
- A tile can add to an existing assembly if
  total strength of matching glues $\geq \tau$





$\tau = 2$

Abstract Tile Assembly Model ATAM

- Proposed by ***[Erik Winfree and Paul W.K. Rothemund, 2000].***
- **An ATAM tiling system is a quadruple**

  <T, s, τ, g>.
- T is a set of tiles.
- A **tile** is a square with a glue on each side:

Abstract Tile Assembly Model

- **s is seed tile:** The initial configuration has only seed tile.
- **τ is temperature:** the minimum accumulative strength that can fix a tile in a configuration.
- **g is glue strength function:** $G \times G \text{->} N^+$ where G is glue set.

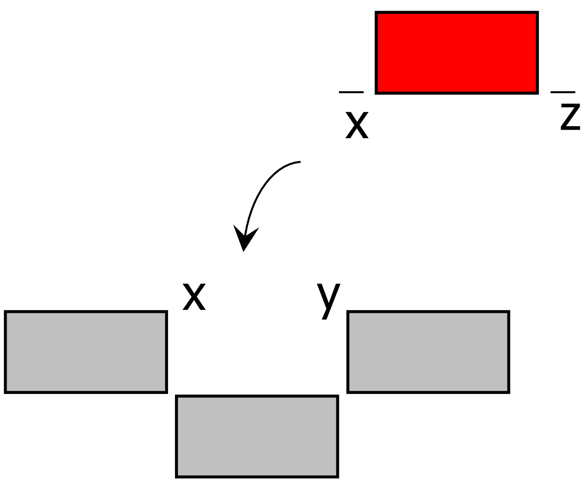
  For any x, y in G, g(x, y)=g(y, x).
- **A configuration is a function:**

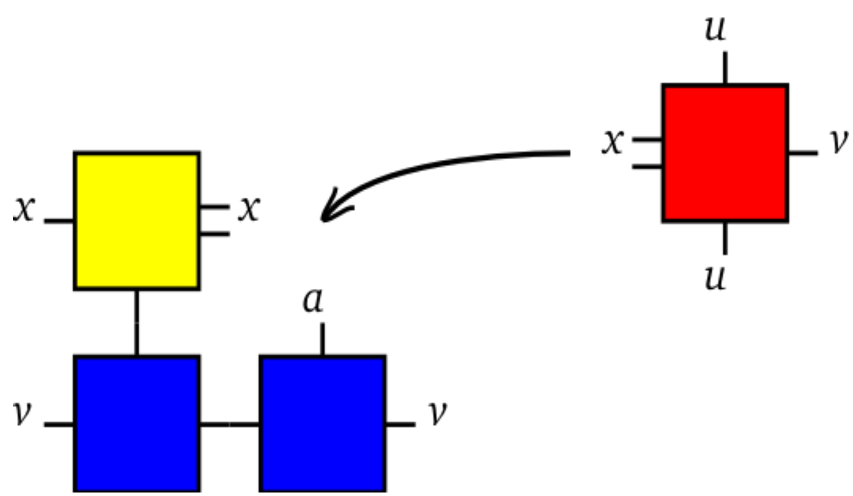
Z $\times$ **Z ->T** $\cup$ **{empty},** where **Z** is the integers.

Abstract Tile Assembly Model(ATAM):

[Rothemund, Winfree, ’ 2000]

- **Temperature T:** A positive integer. (Usually 1 or 2)
- **A set of tile types:** Each tile is an oriented rectangle with glues on its corners. Each glue has a non-negative strength (0, 1 or 2).
- **The initial assembly (seed).**





A tile can attach to an assembly iff the combined strength of the "matched glues" is greater or equal than the temperature.

Abstract Tile Systems

- Tile: the four glues and their strengths
- Tile System:
  - κ tiles
  - Infinitely many copies available of each tile
  - Temperature τ
  - A seed tile s
- Accretion Model:
  - Assembly starts with a single seed tile, and proceeds by repeated addition of single tiles e.g. Crystal growth
  - Are interested primarily in tile systems that assemble into a unique terminal structure

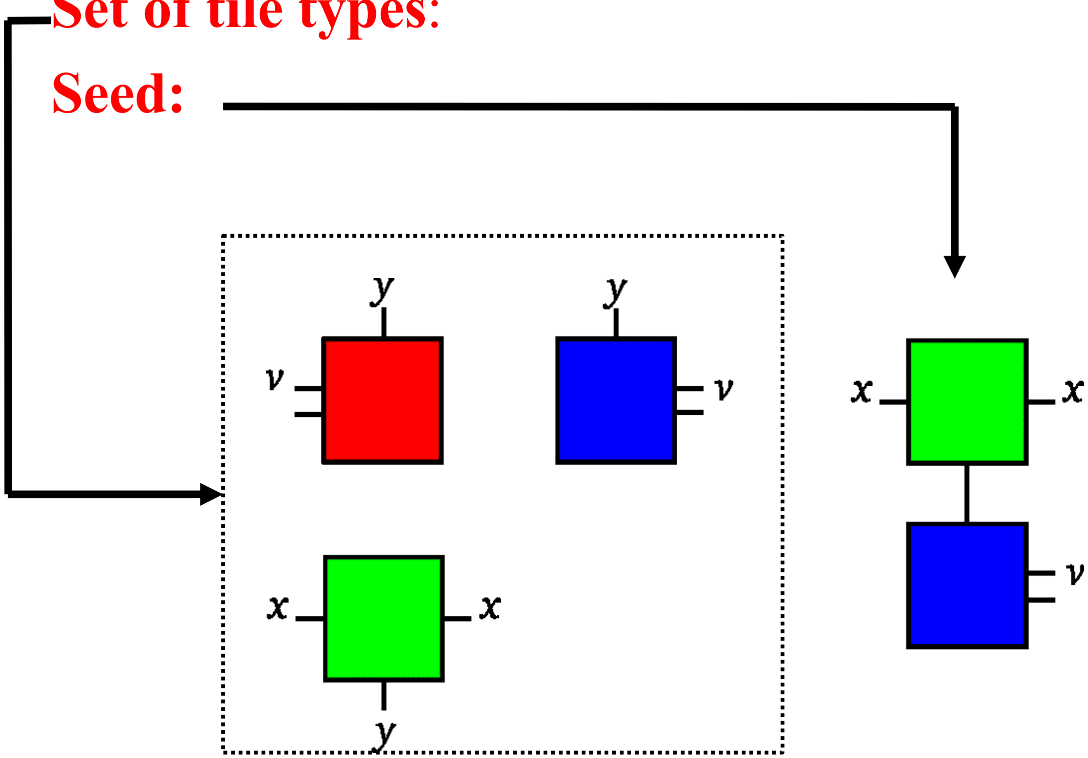
[Rothemund and Winfree '00] [Wang '61]

Example of a Tiling Self Assembly Process

Temperature: 2

Set of tile types:

Seed:



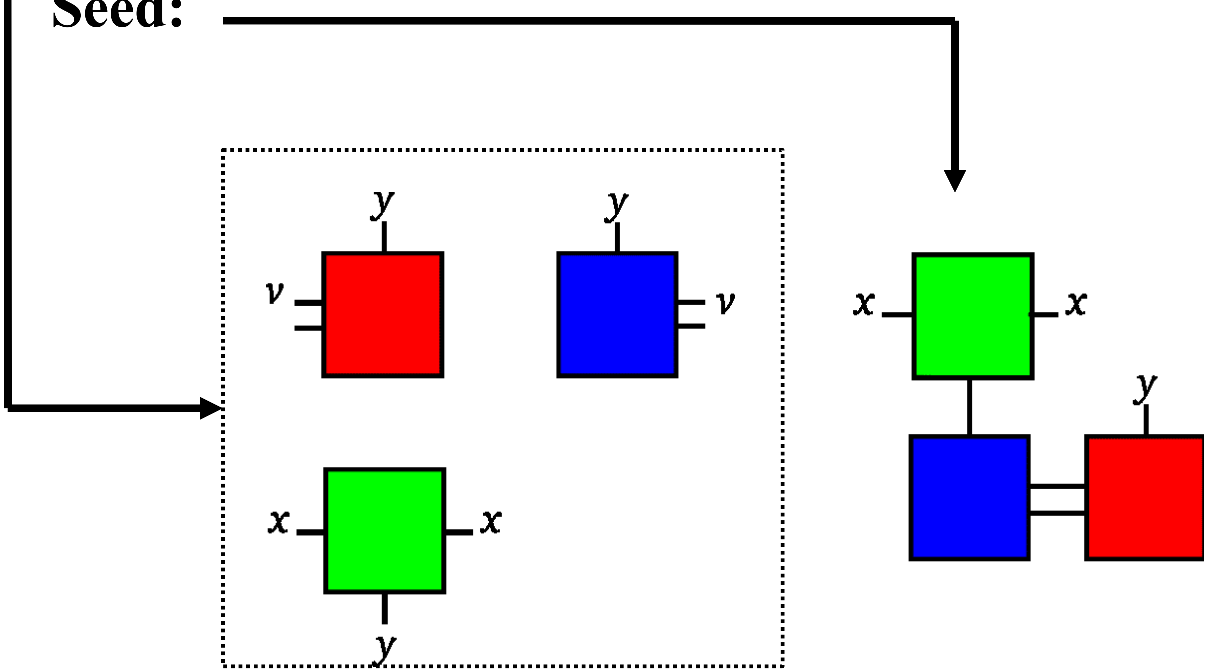
[Cheng, Goel, Cheng]

Example of a Tiling Self Assembly Process, cont

Temperature: 2

Set of tile types:

Seed:



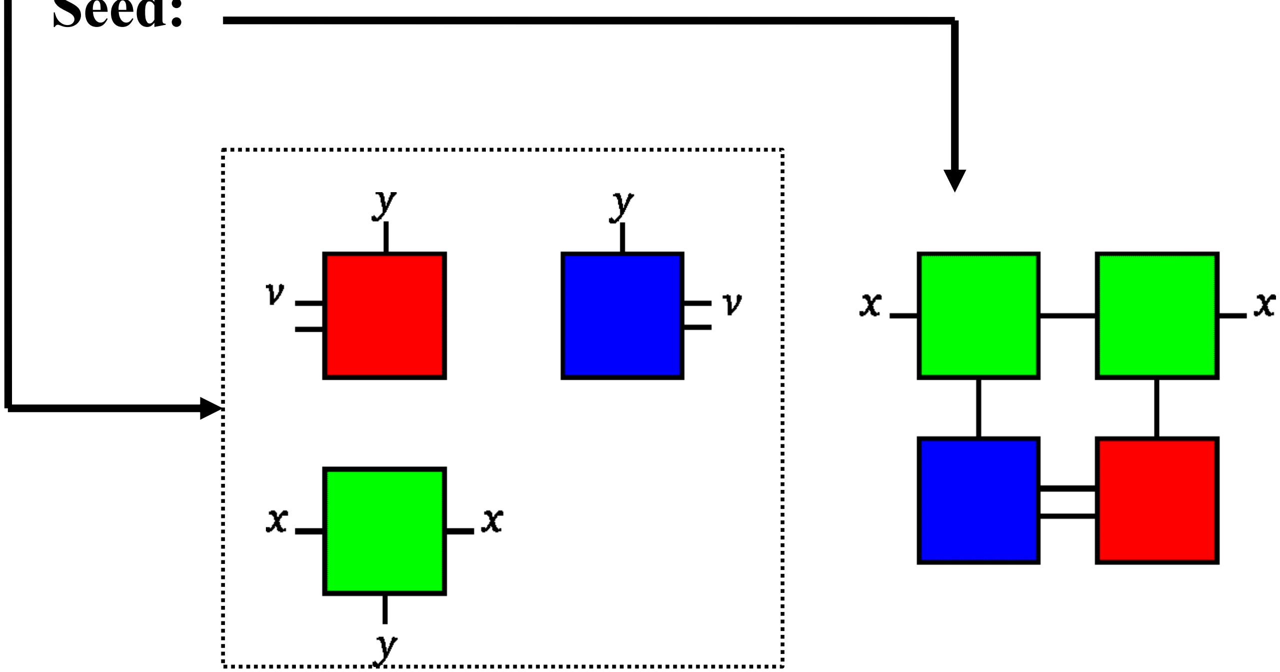
[Cheng, Goel, Cheng]

Example of a Tiling Self Assembly Process, cont

Temperature: 2

Set of tile types:

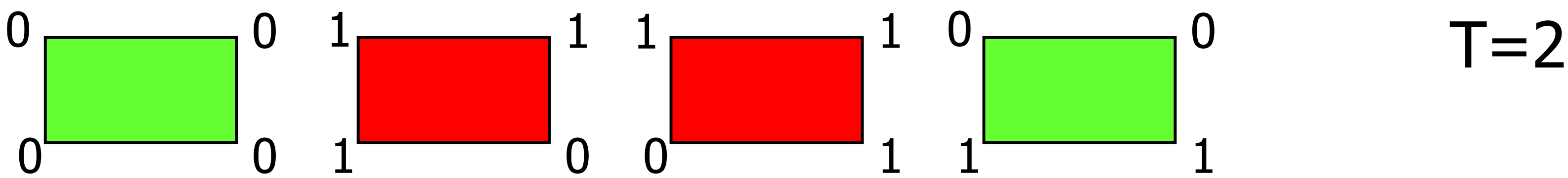
Seed:



[Cheng, Goel, Cheng]

Example:
Tiling Assembly of Sierpinski System

[Winfree, 2096]

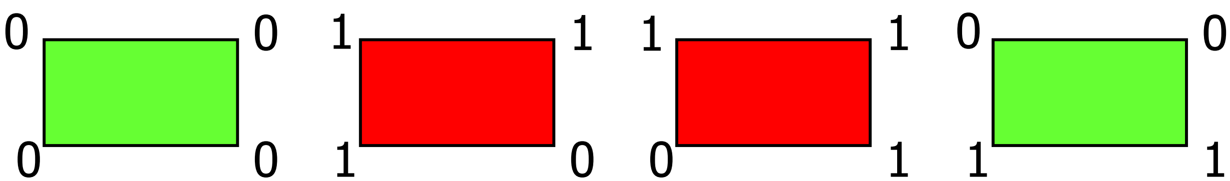


Seed Tiles forming Base Assembly

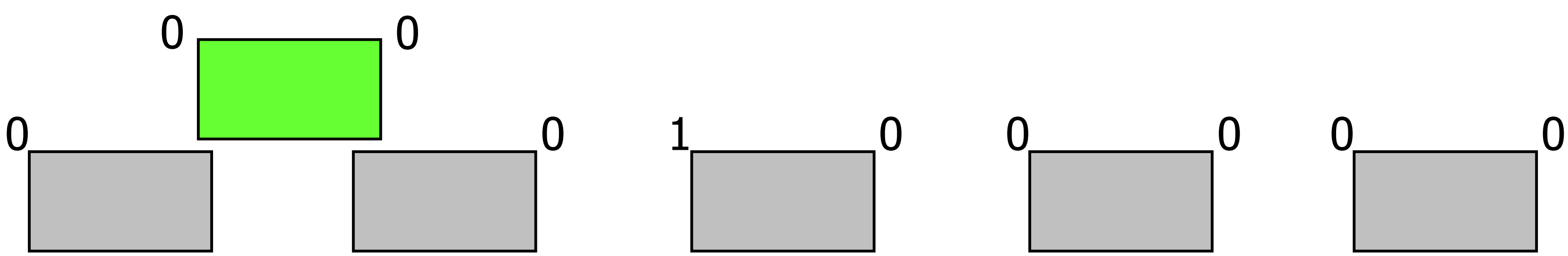
[Cheng, Goel, Cheng]

Example:
Tiling Assembly of Sierpinski System, cont

Tiles to be used in assembly:



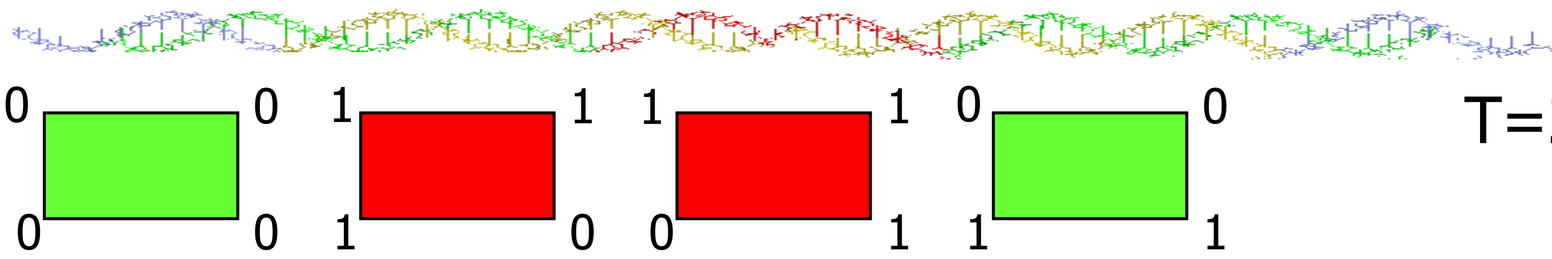
T=2



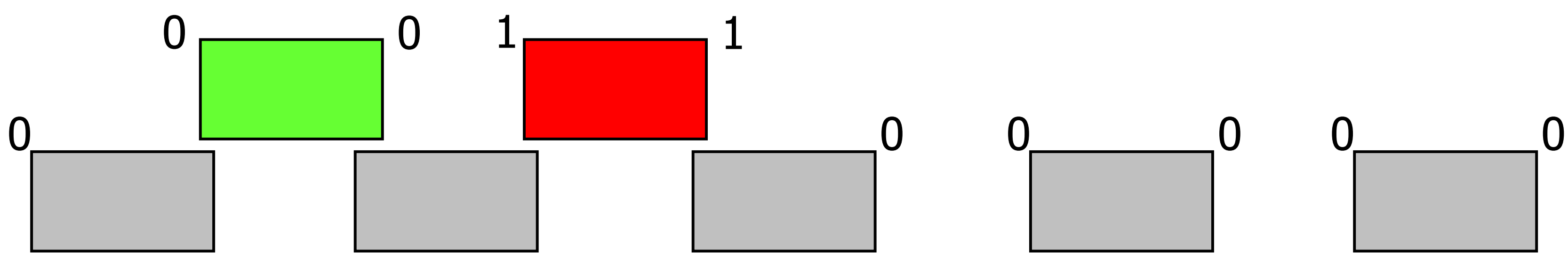
Seed Tiles forming Base Assembly

[Cheng, Goel, Cheng]

Example:
Tiling Assembly of Sierpinski System, cont



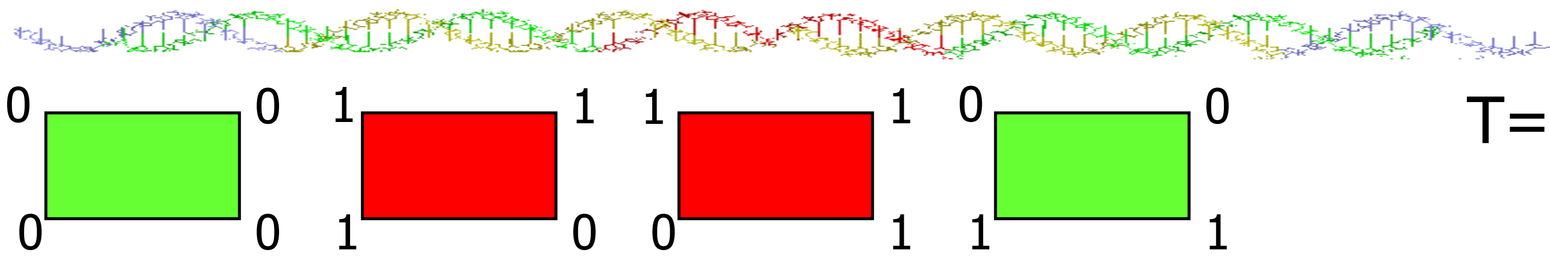
T=2



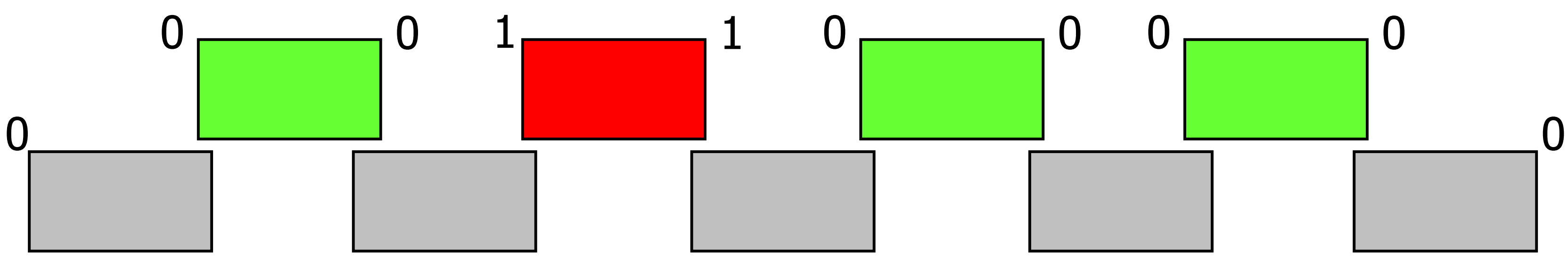
Seed Tiles forming Base Assembly

[Cheng, Goel, Cheng]

Example:
Tiling Assembly of Sierpinski System, cont



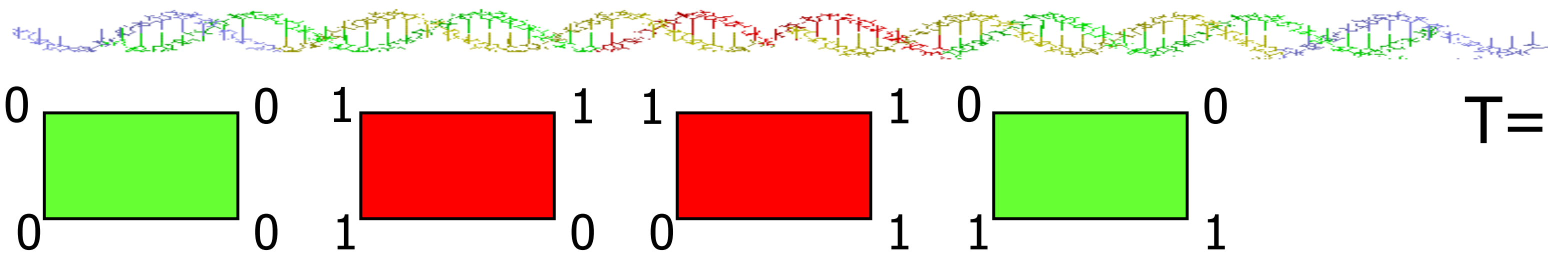
T=2



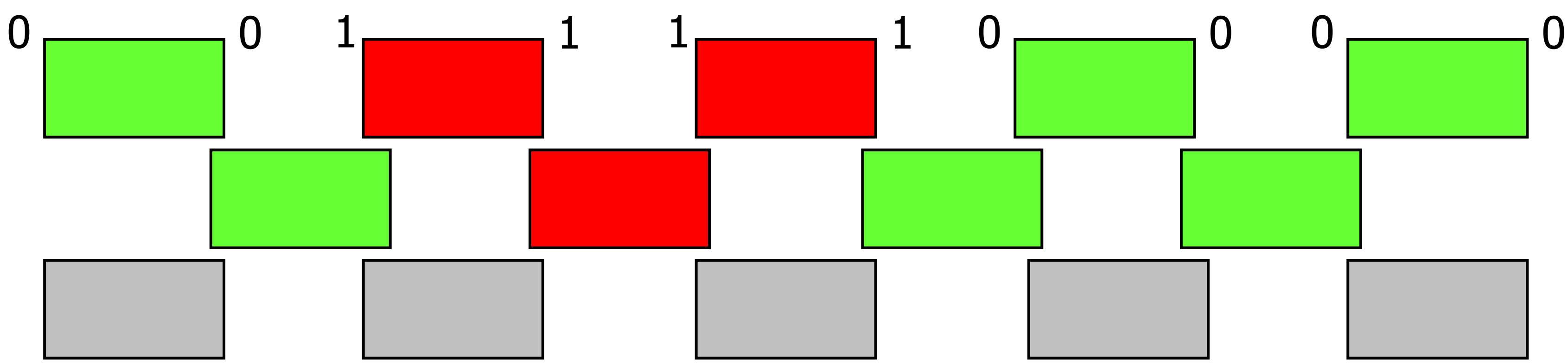
Seed Tiles forming Base Assembly

[Cheng, Goel, Cheng]

Example:
Tiling Assembly of Sierpinski System, cont



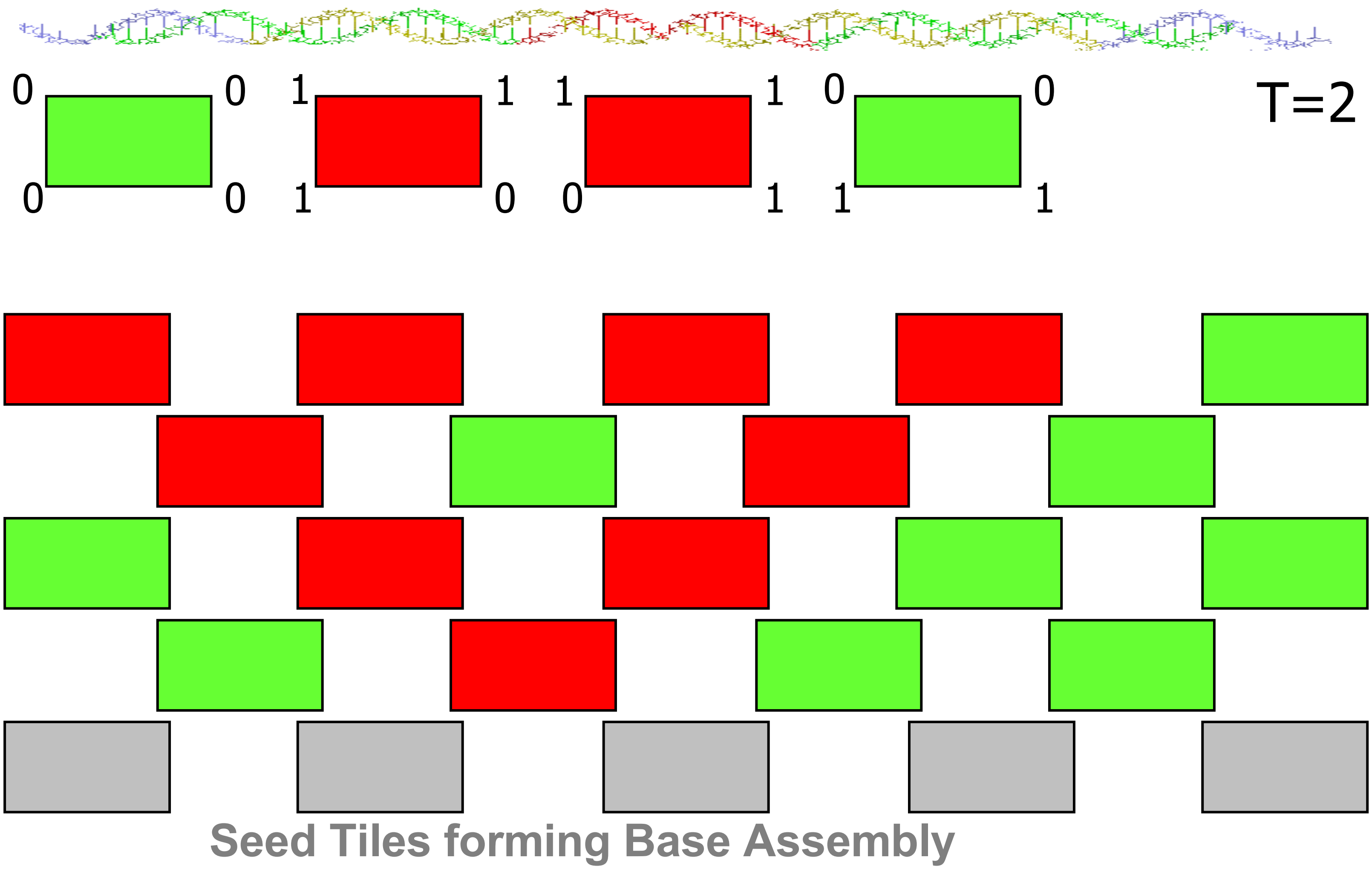
T=2



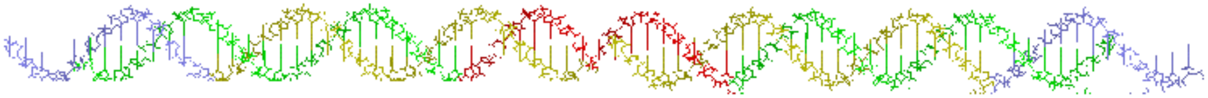
Seed Tiles forming Base Assembly

[Cheng, Goel, Cheng]

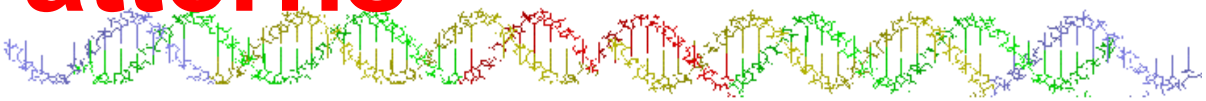
Example:
Tiling Assembly of Sierpinski System, cont



[Cheng, Goel, Cheng]

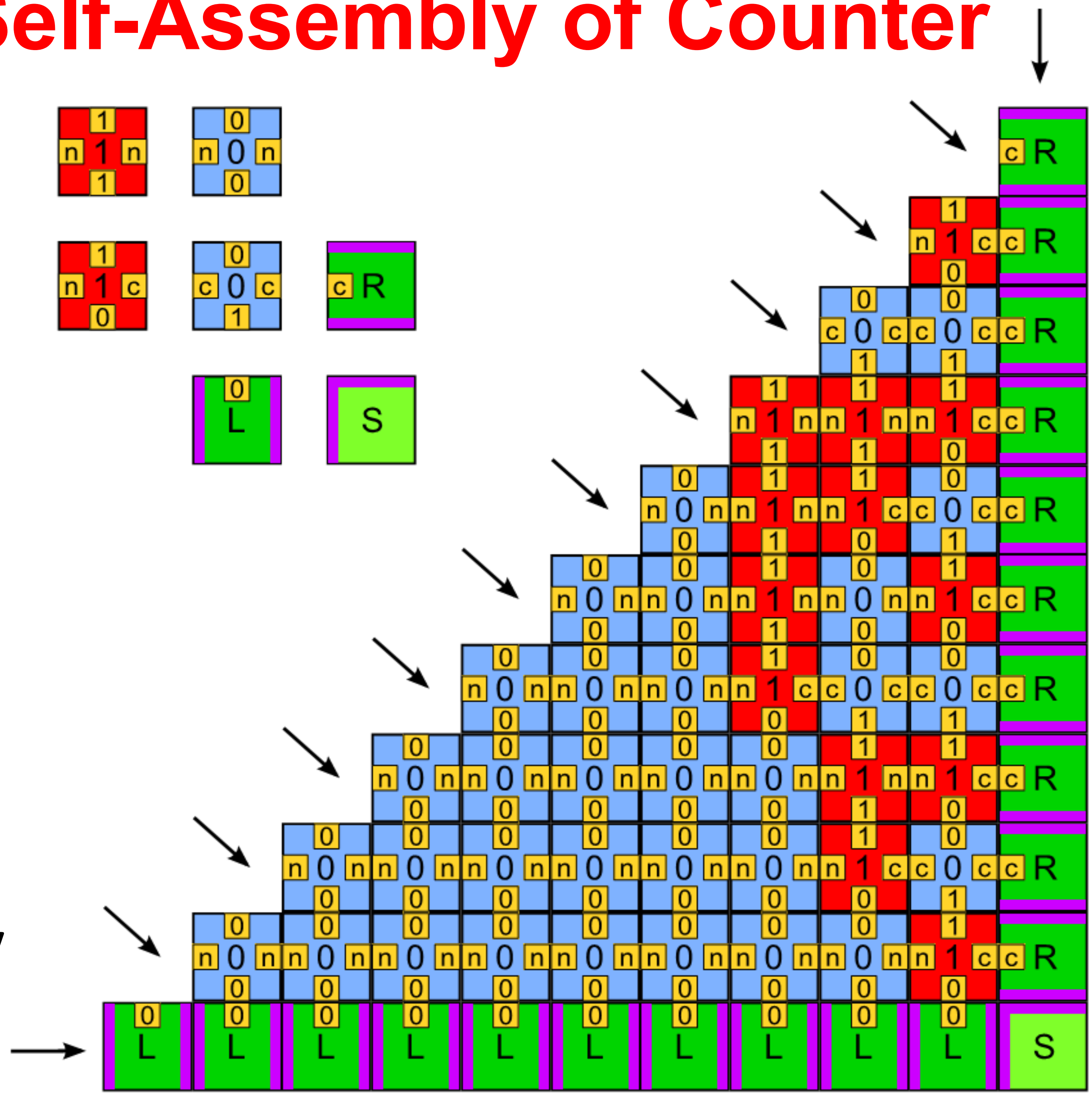


Self-Assembled Circuit Patterns



[Cook, Rothmund, Winfree, DNA9]

Tiling Self-Assembly of Counter

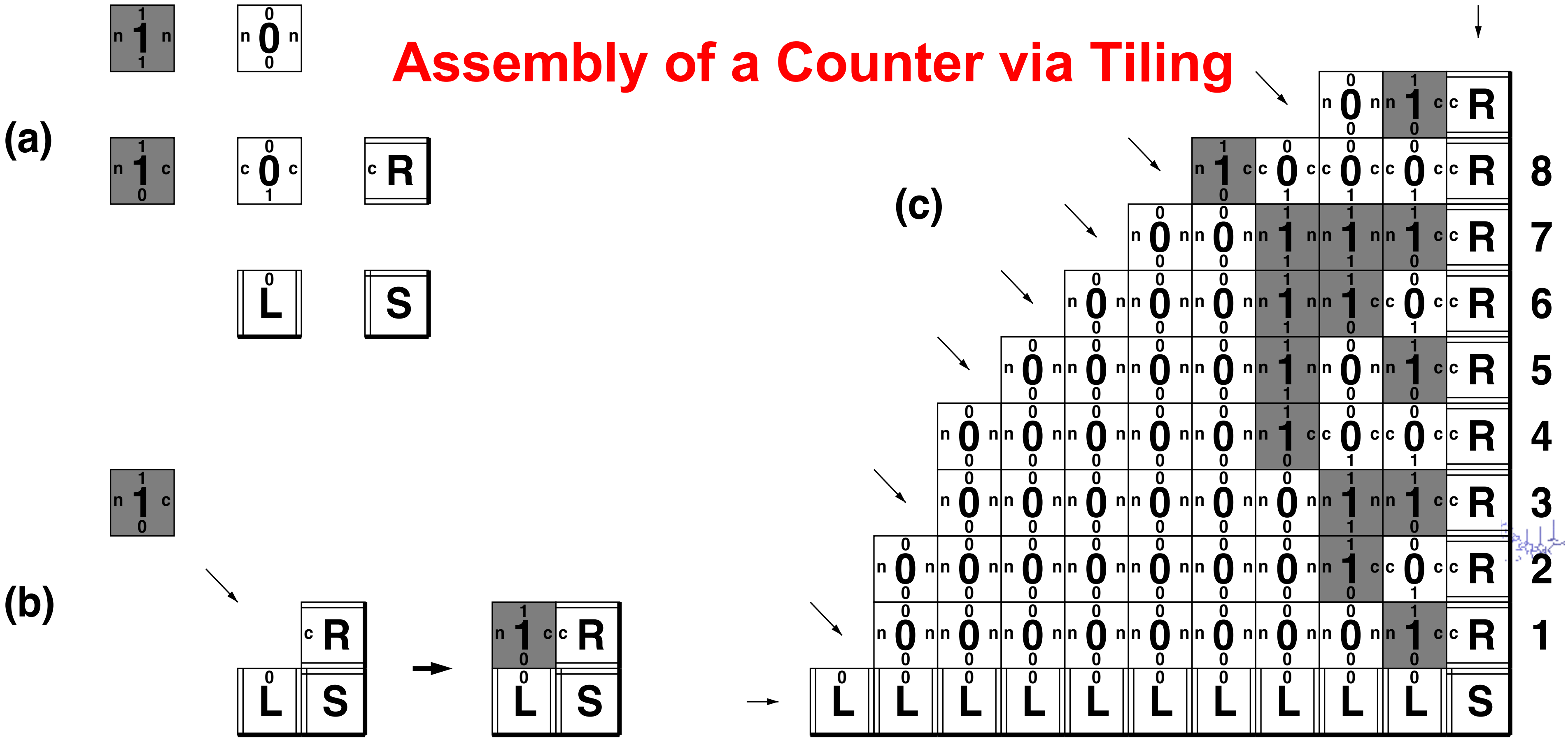


Counter made by self-assembly *[Adleman, Cheng, Goel, Huang 2001]*

Execute Counting Row by Row:

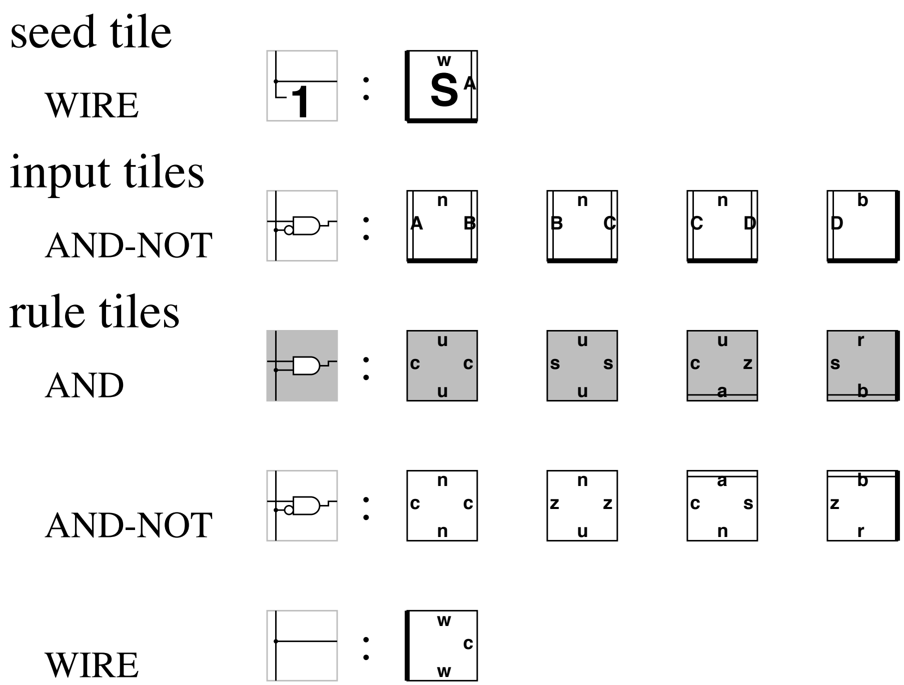
- **c is the carry bit**
- **n is the no-carry bit**

Assembly of a Counter via Tiling

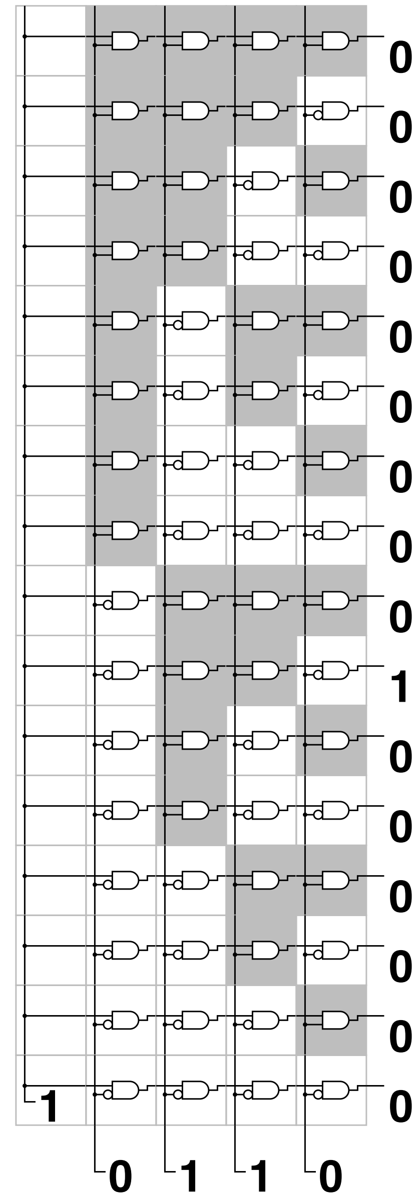


Assembly Model program for counting in binary. The tiles labeled "1" are colored gray to make it easier to see the resulting pattern, visible in (c). The self-assembly progresses by individual tiles accreting to the assembly as shown in (b). Edges marked with a small letter or number have bond strengths of 1, while edges with a double line have bond strengths of 2 (and do not require a further label here, since there is only one vertical and one horizontal kind). A later stage of self-assembly is shown in (c), with arrows indicating all the places that a new tile could accrete.

Assembly of a Demultiplexer via Tiling



Demultiplexer: **transforms a binary number to a unary number position**



Using a binary counter to self-assemble a demultiplexer. Logic levels for an example input-output pair are shown: only the row that exactly matches the input pattern is set to “1”. To make a pattern with N rows, $10 + \log N$ tiles are used.

Assembly of two Demultiplexers via Tiling

The Two Demultiplexers allow addressing in 2D memory:

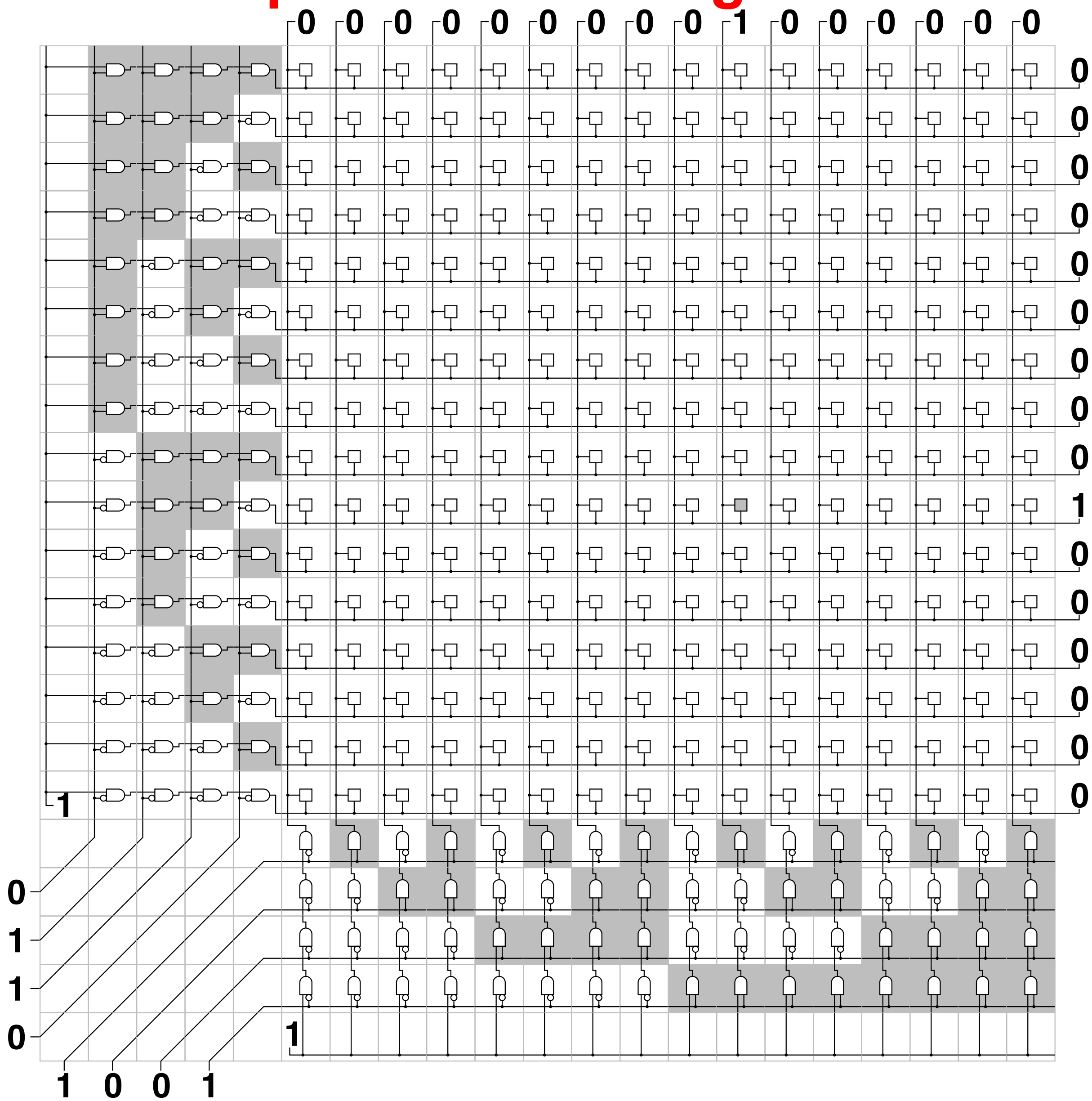
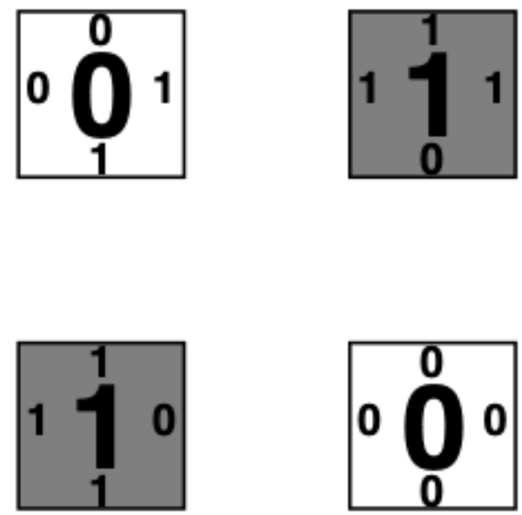


Fig. 3. Two self-assembled demultiplexers at right angles can address a memory. The gray memory cell is being addressed in this figure.

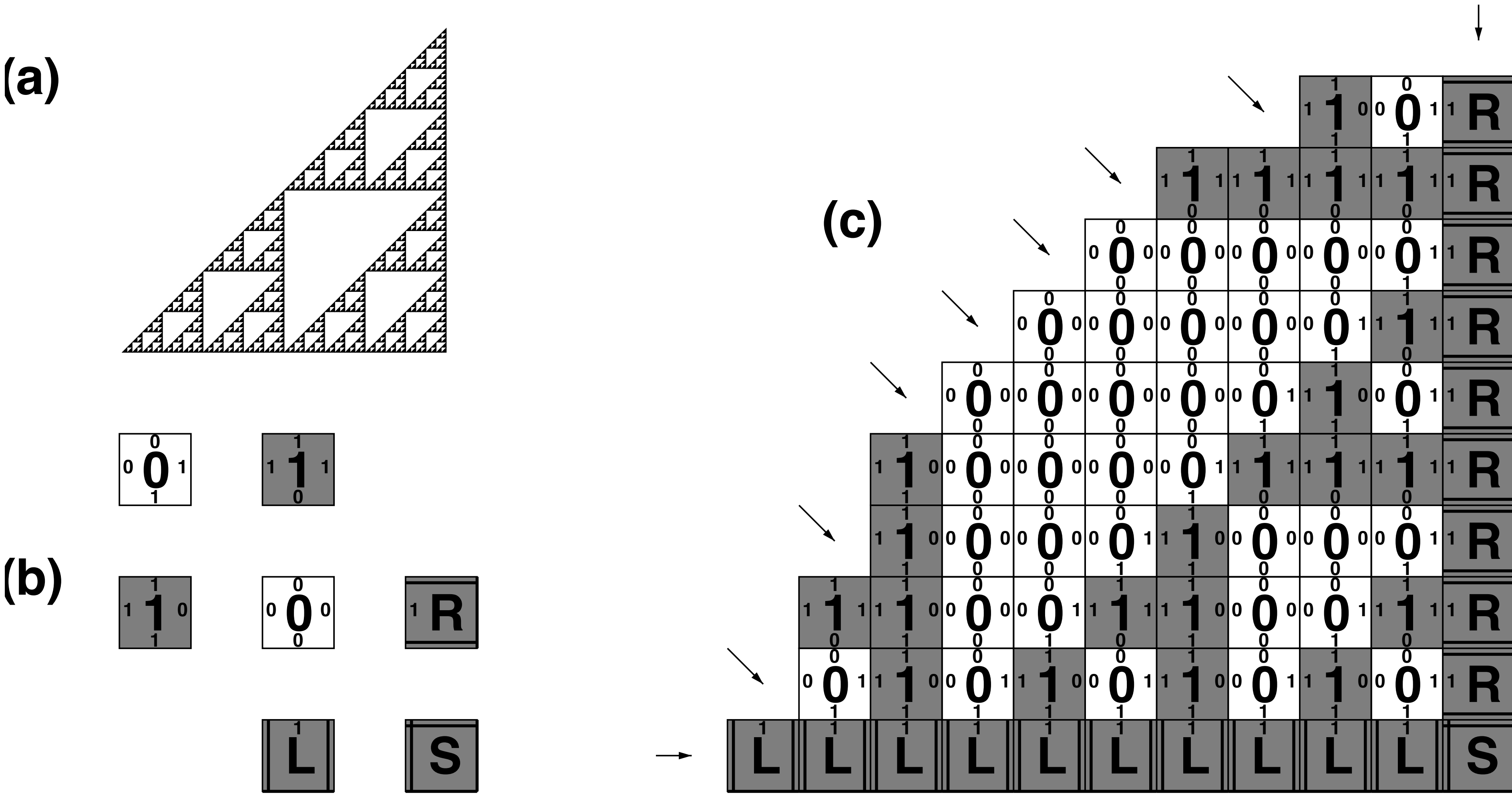
Assembly of Sierpinski Triangle (mod 2 Pascal Triangle) via Tiling Self Assembly

- **Assume τ = 2, and all glue strengths are 1**
- **Uses four "XOR" tiles, where:**
  - **the bottom side is labeled either 0 or 1, and**
  - **the right side is labeled either 0 or 1, and**
  - **the left and upper sides are labeled e XOR s where e and s are the bits on the east and south sides**



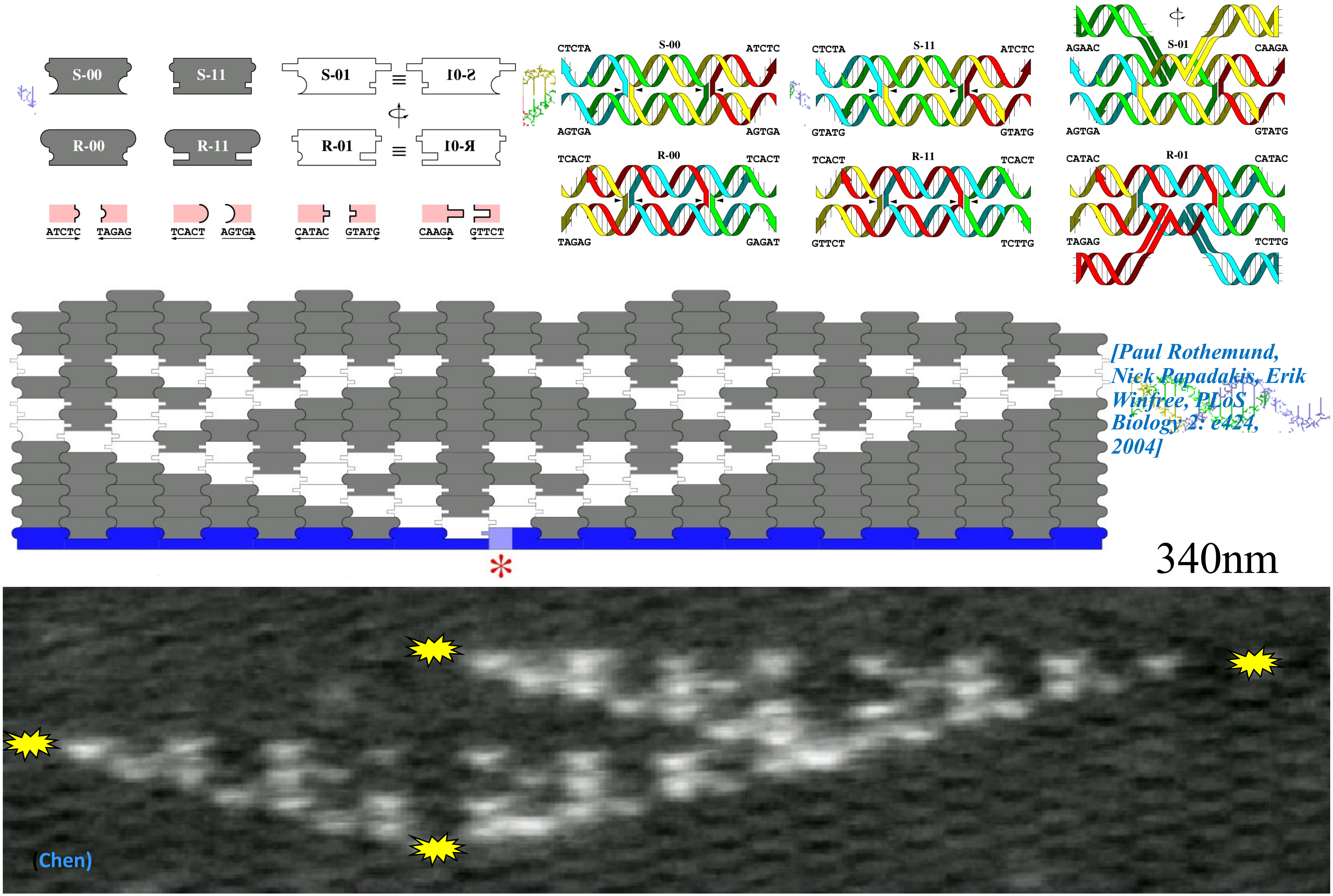
- **Given the appropriate inputs (i.e. seeds), these tiles can do some very interesting things like**
  - **Computing the parity**
  - **Assembling into a fractal pattern**

Assembly of Sierpinski Triangle (mod 2 Pascal Triangle) via Tiling Self Assembly



The Sierpiński triangle and a set of tiles that construct it in the limit.

Sierpinski triangle experiments using DAO-E Tiles



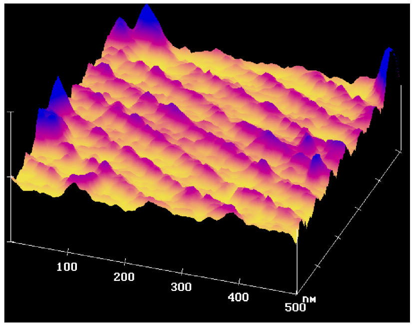
[Paul Rothemund, Nick Papadakis, Erik Winfree, PLoS Biology 2: e424, 2004]

Applications of Tiling Self-Assembly

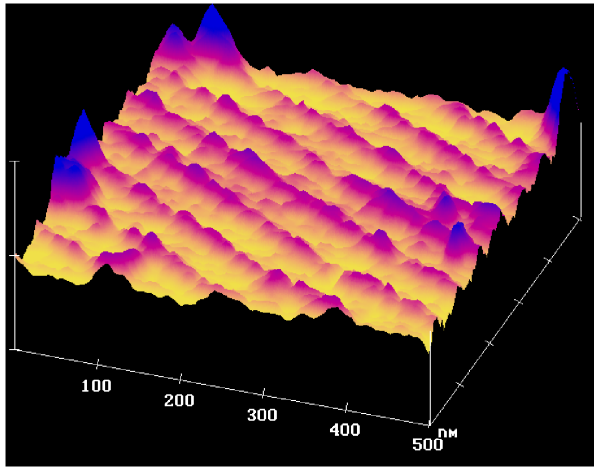
- Self-assembly can be used to create small electrical devices such as FLASH memory. [Black et. Al.’ 03]

- Self-assembly can create nanostructures which “steer” light in the same way computer chips steer electrons. [Percec et. Al.’ 03]

- DNA strands can self-assemble into tiles and those tiles can further self-assemble into larger structures. This has many potential applications. *[Winfree 96]*



DNA “rug” by Winfree

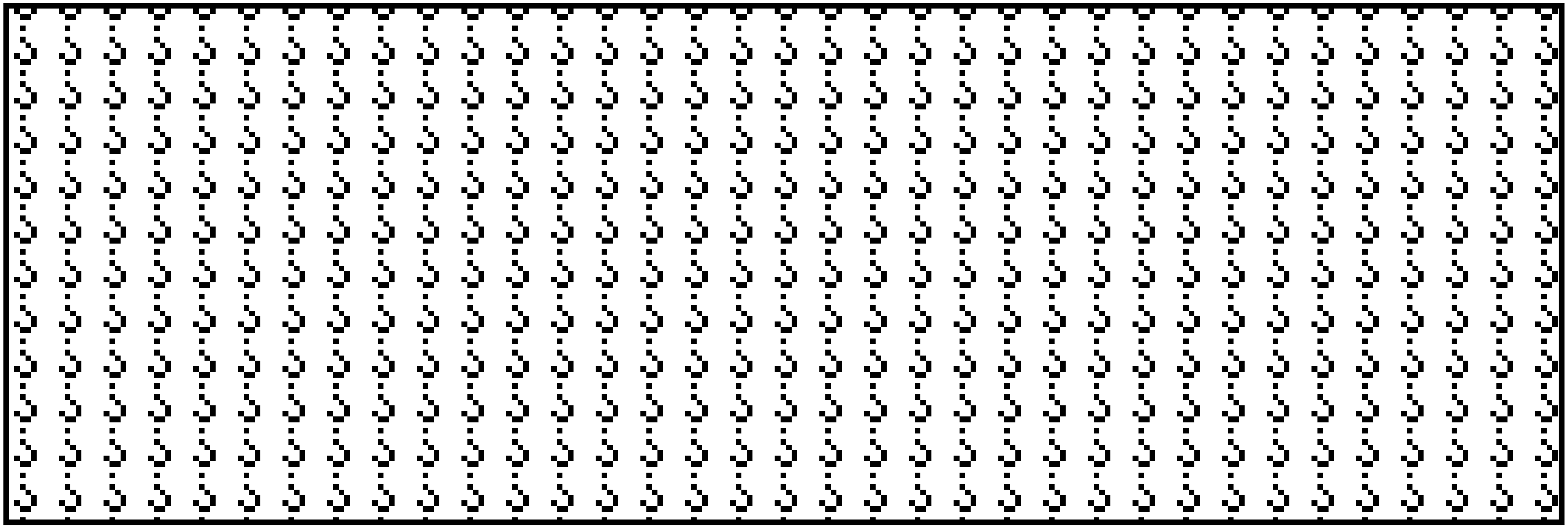


A DNA "rug" assembled using DNA tiles

- The rug is roughly 500 nm wide, and
- is assembled using DNA tiles 12nm by 4nm (false colored)

[Erik Winfree, Caltech]

Mesoscale Tile Systems



Performing XOR using self-assembling mesoscale (1 cm) tiles

(Square tiles floating on oil with "glues" on each side)

[Paul Rothemund, USC]

Theory of 1-D, 2-D, and 3-D Tiling Self-assemblies

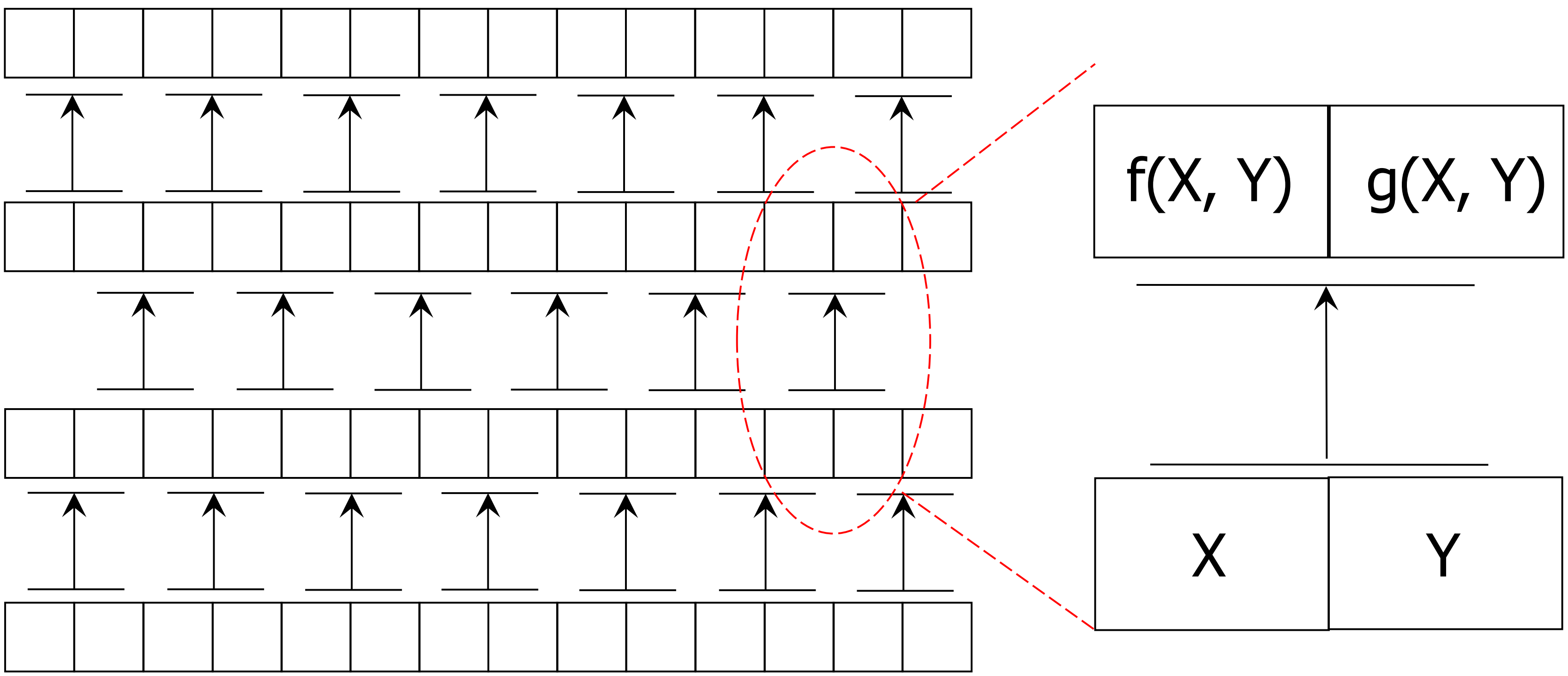
- **The theory of Linear 1-D tiling self-assemblies are very different from 2-D tiling self-assemblies**

- **2-D and 3-D tiling self-assemblies are very similar in terms of the theory**

2D Tiling Simulation of Block Cellular Automata (BCA)

- Giving a BCA, Winfree showed that a 2D Tiling system can be designed to simulate the BCA's computation.
- Each row of tiles given the BCA configuration at a given time.
- BCAs can simulate Turing Machines, so this implies that
  - **Tiling assemblies can execute any computation.**
  - **various properties of Tiling assemblies are undecidable.**

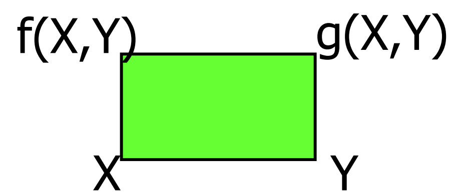
2D Tiling Simulation of Block Cellular Automata (BCA)

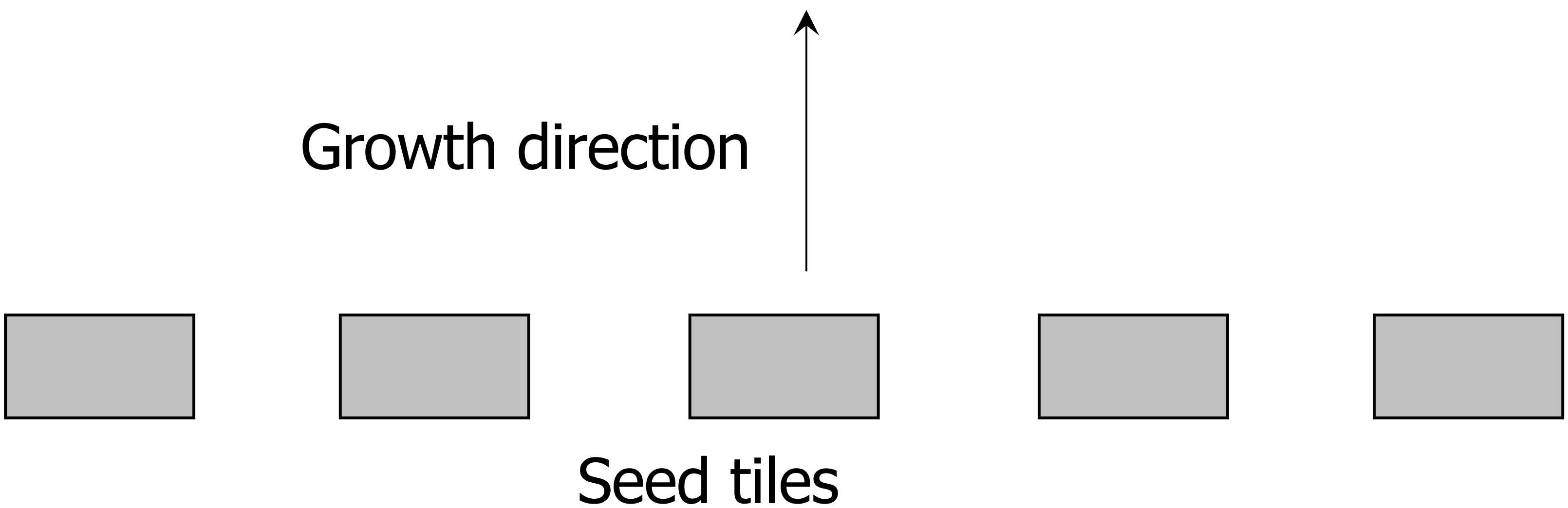


2D Tiling Simulation of Block Cellular Automata (BCA)

T=2

A series of tiles with format:

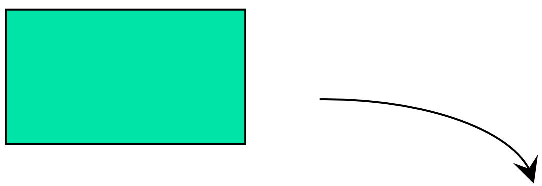




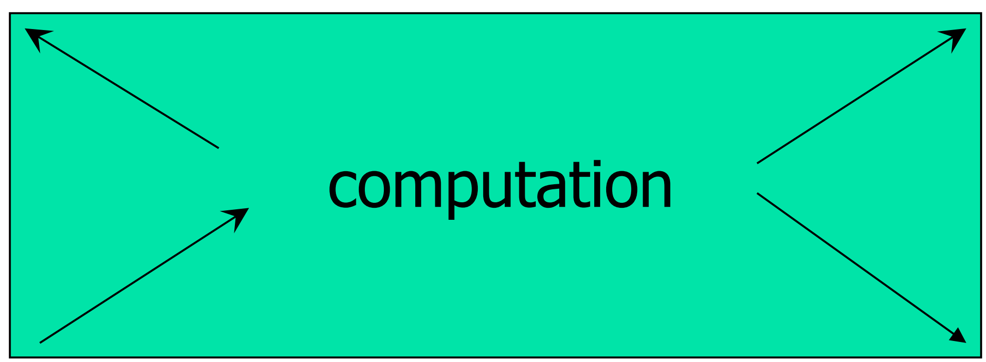
Tile Complexity

- Tile complexity or program-size complexity: the minimum number of tile types required to assemble a shape. ***[Erik Winfree and Paul W.K. Rothemund STOC 2000]***
- A special case: in the linear version proposed by ***[Chandran et al]*** the tile set is MULTISET.

Assemble Arbitrary Shapes



Replace each tile by a block.
Size of block = O(computation time)



Theoretical and Algorithmic Results for Tiling Assemblies

- **Efficiently assembling basic shapes with precisely controlled size and pattern**
  - **Constructing N X N squares with Ω(log n/log log n) tiles** [Adleman, Cheng, Goel, Huang, ’01]
  - **Perform universal computation by simulating BCA** [Winfree ’99]
- **Library of primitives to use in designing nano-scale structures** [Adleman, Cheng, Goel, Huang, ’01]
- **Automate the design process** [Adleman, Cheng, Goel, Huang, Kempe, Moisset de espanes, Rothemund ’01]
- **Robustness**

Design Complexity of a Tile System

- Suppose the tile system has k different tiles
  - Assume that we have θ(k) different glues
  - Assume for simplicity that all glues are represented by DNA strands of the same length, L
  - **How many different glues can we have?**
    - Each position can be one of the bases A, C, T, or G. An A pairs with T on the complementary strand, and a C pairs with G. So we can think of each position as corresponding to an AC base pair or a CG base pair.
    - Two choices) Binary encoding
    - Number of glues · 2^L
  - **Total length of all the DNA strands in all the distinct tiles = θ(k log k). We will refer to k log k as the design complexity**
    - Does not depend on the size of the final structure. This is how much time and expertise and resources you would have to invest into designing the components (like program complexity)

Notes

- We will sometimes use the quantity k to refer to the design complexity.

- The rules of assembly are easy to code up in a simulator:
  - If a tile system of k tiles can assemble into a certain shape, then
  - There is a computer program of size $\theta(k \log k)$ which generates that shape.

Three Exercises in Tile Assembly

1. **Suppose we did not assume that the number of glues is $\theta(k)$. Can you still prove that the total length of all the strands in all the tiles is $\theta(k \log k)$?**

2. **We assumed that our tiles are oriented, i.e., east is always east and north is always north, and the tiles are not allowed to rotate. Show how an oriented tile system can be simulated using a system where tiles can rotate by +/- 90 degrees.**

3. **Now use the fact that the glues are really DNA strands to show how an oriented tile system can be simulated using a system where tiles can rotate by +/- 90 as well as 180 degrees.**

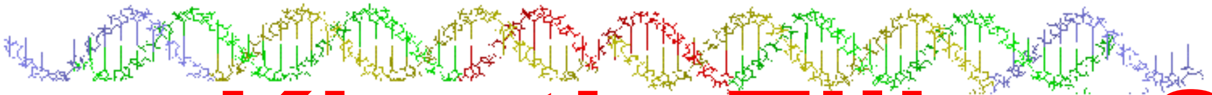
Tile System Design

- **Library of primitives to use in designing nano-scale structures**

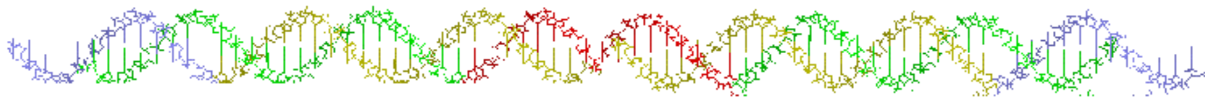
  [Adleman, Cheng, Goel, Huang, 2001]

- **Automate the design process**

  [Adleman, Cheng, Goel, Huang, Kempe, Moisset de espanes, Rothemund 2001]



Kinetic Tiling System Model (KTAM)

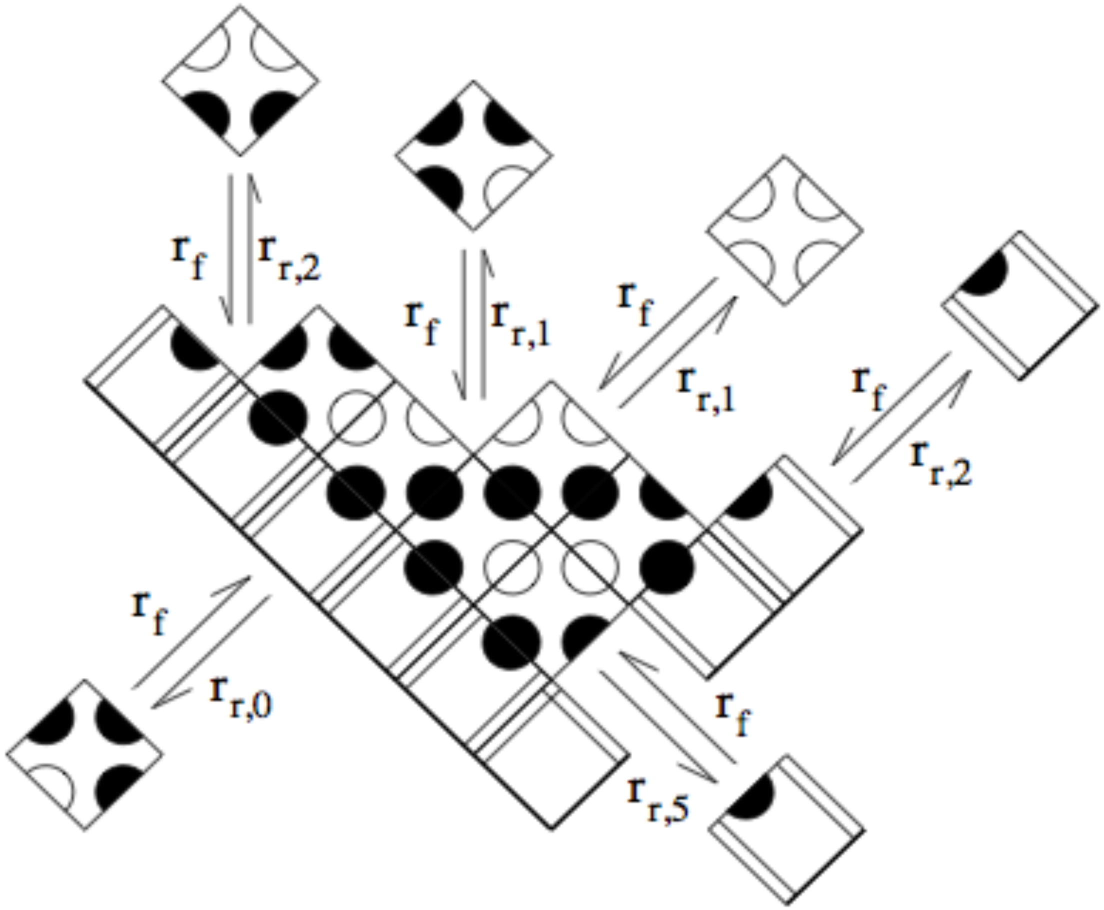


Kinetic Tile Assembly (KTAM) Model

[Erik Winfree,1988] **Assumptions:**

- All monomers hold the same constant concentration.
- There are not interactions among aggregates.
- All monomers have the same forward assembly rate constant.
- The reverse tile assembly rate depends exponentially on the number of pairs need to be broken.

Kinetic Tile Assembly Model



[Erik Winfree, 1998]

Kinetic Tile Assembly Model

- [monomer] is the concentration of of a monomer (single isolated tile).
- k_f is the forward tile assembly rate constant.
- $r_f = k_f\,[monomer] = k_f e^{-G_{mc}}$
- $r_{r,b} = k_{r,b} = k_f e^{-bG_{se}}$
- G_{mc} is a measure of entropic cost of fixing a monomer at a particular position. It is determined by the concentration of monomer.

Kinetic Tile Assembly Model

- G_{se} is a measure of free energy cost to break a double helix of length s, where

$$G_{se} = (4000K/T - 11)s.$$

- b is the number of s-length double helix that need to be broken.

Kinetic Tile Assembly Model:

[Winfree, 1998]

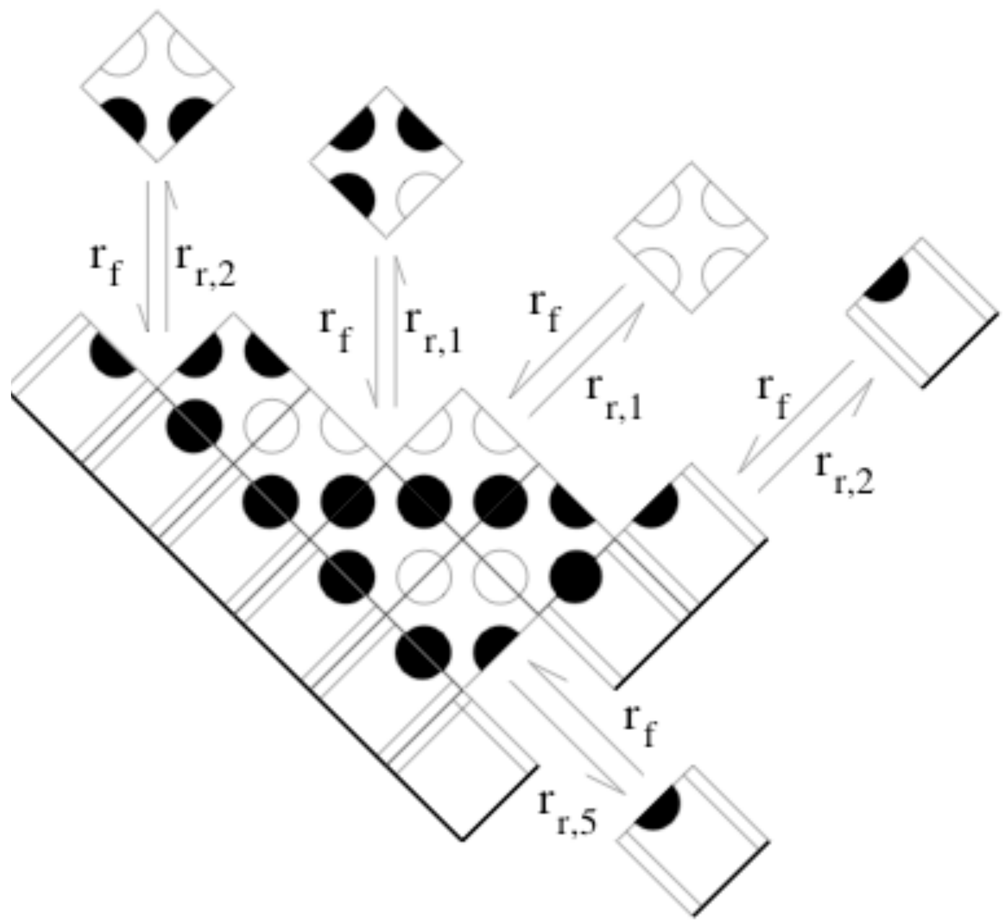
A tile can attach at any location.

The rate of attachment

r_f = constant.

The rate of detachment

$r_{r,b} = c\, e^{-bG}$



Kinetic Tile Assembly Model => Abstract Tile Assembly Model

- Set the temperature and concentration to

$$r_{r,T+1} \ll r_{r,T} < r_f \ll r_{r,T-1}$$

  - If a tile attaches with strength < T-1, it is likely to fall off very fast.
  - If a tile is held by strength at least T+1, it is unlikely to fall off

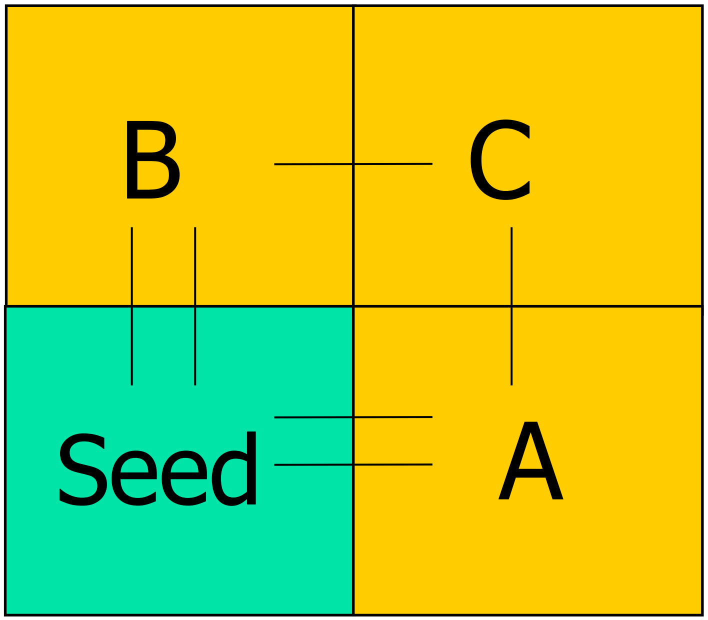
Time Complexity of Kinetic Tile Assembly

- **States:** possible assembly configurations.
- **Rate between state S1 and S2:** if S2 is got by attaching a tile x to S1. The rate is the concentration of tile x. Otherwise, no edge between S1 and S2.

- **The time from initial configuration to the terminal configuration**
  - Is a random variable.
  - **Assembly Time complexity** is the expected value of the random variable

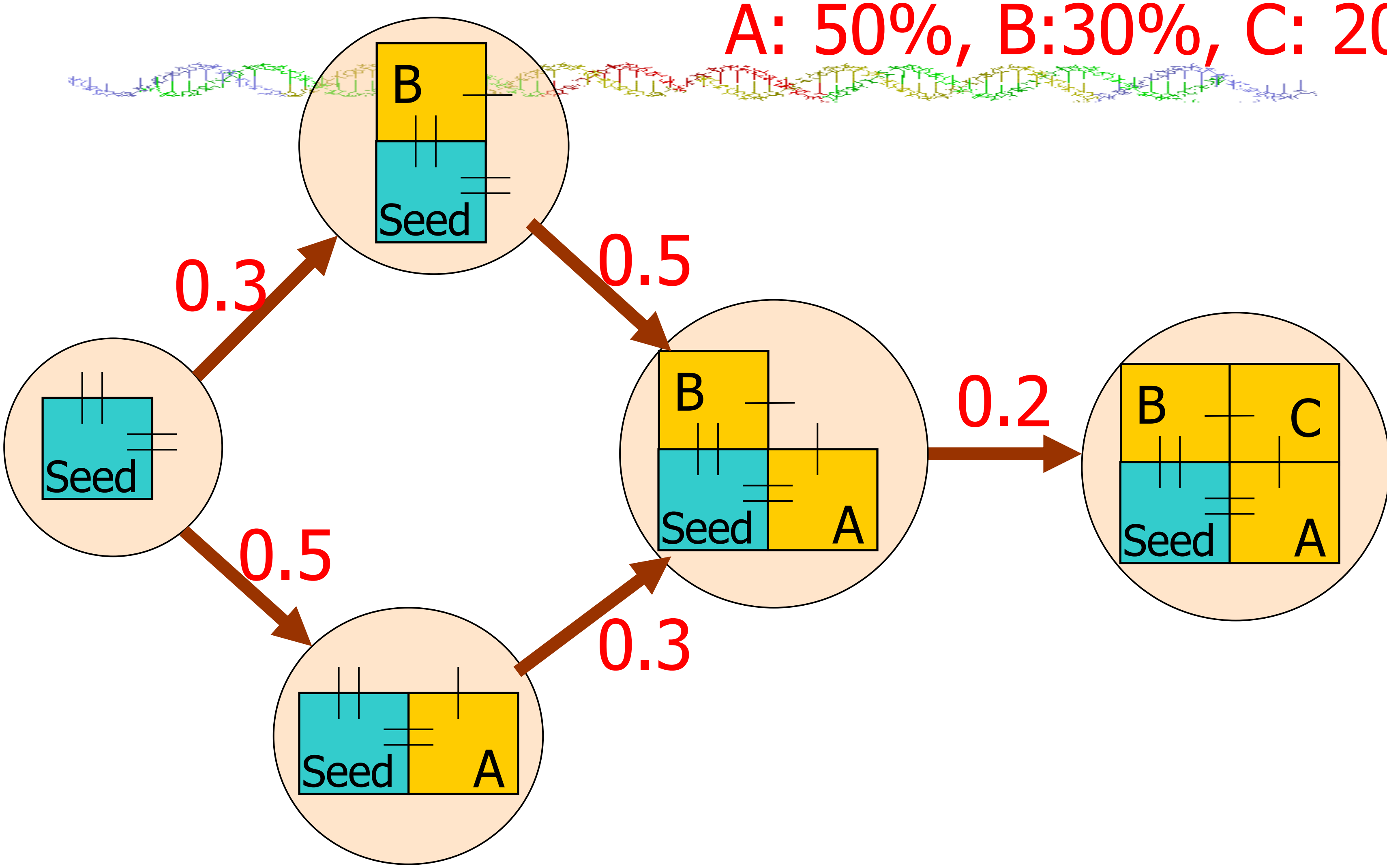
Example: A Tile System and its Running Time

$\tau = 2$



Example: A Tile System and its Running Time

A: 50%, B:30%, C: 20%



Reversible Tile Assembly Model

- Proposed by Leonard M. Adleman in 1999 to study linear assembly.
- Define two functions:
  - (1) σ: $G \times G$->[0,1]
  - (2) v: $G \times G$->[0,1] where G is the glue set.
- **σ(g1,g2) is the probability of a tile sticking between glue g1 and g2.**
- **v(g1,g2) is the probability of a tile unsticking between glue g1 and g2.**

Tile Systems and Running Time

Define a continuous time Markov chain M:

- State space S: set of all structures that a tile system can assemble into
- Tiles of type T_i have concentration c_i
  - $\Sigma_i\, c_i = 1$
- Unique terminal structure => unique sink in M
- Seed tile is the unique source state

- **Tiling Assembly time:**
  - the hitting time to reach the sink state from the source state in the Markov Chain

Simple Example:
1-D Tile Assemblies

- **Resulting Markov Chain:**
  - Average time for the i^{th} tile to attach is $1/c_i$
  - Assembly time = $\sum_i 1/c_i$
- For fastest assembly, all tiles must have the same concentration of $1/(n-1)$
  - Expected assembly time is $\frac{1}{4} n^2$
  - Can assemble thicker rectangles much faster and with much fewer different tiles

Multiple Temperature Tiling Assembly Model

- Proposed by ***[Aggarwal et al. 2004].***
- Replace temperature t in tiling model by sequence of temperatures $\{\tau_i\}_{i=1}^{k}$.
- A system that has k temperatures is a k-temperature system.
- **Assembly mechanism:** the assembly process of a k-temperature system has k phases.

Multiple Temperature Tiling Assembly Model

Assemble and disassemble under temperature τ_1 for long enough time.

Phase 1

⋮

Assemble and disassemble under temperature τ_k for long enough time.

Phase k

The terminal product of phase k is the terminal product of this k-temperature system.

Flexible Glue Tiling Assembly Model

- Proposed by ***[Aggarwal et al. 2004].***
- **In standard model:** $g(x, y) = 0$ if $x \neq y$.
- **In flexible glue model:** $g(x, y)$ may not be 0 when $x \neq y$.

q-Tile Tiling Assembly Model

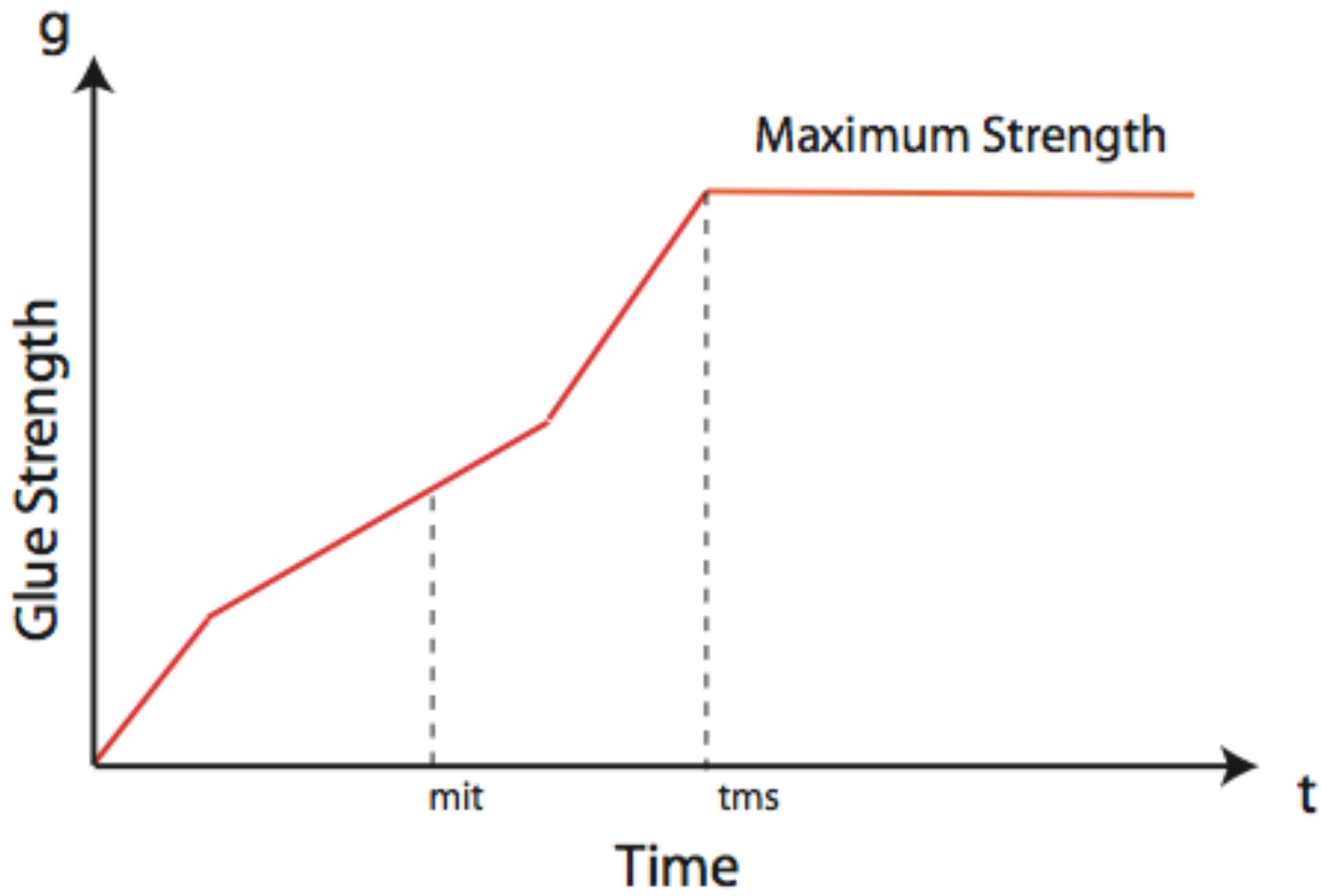
- Proposed by *[Aggarwal et al. 2004].*
- **In standard model:** only single tile can attach to the growing supertile containing seed.
- **In q-tile model:** supertiles of size not larger than q can form and attach to the seed supertile if the accumulative glue strength between two supertiles is not less than temperature.
  - Standard model is exactly 1-tile model.

Time-dependent Glue Tile Assembly Model

- **Proposed by** ***[Sahu et al. 2005].***
- **The glue strength function g is defined a a function of time t:**

  $g: G \times G \times R \rightarrow R$. $g(x, y, t)$ is the glue strength between glue x and glue y when they have been juxtaposed for t time.

Time-dependent Glue Tiling Assembly Model



[Sahu et al. 2005]

Time-dependent Glue Tiling Assembly Model

- Define $r: G \times G \rightarrow R$ as ***time for maximum glue strength***.
- $g(x, y, t)$ is growing when $t < r(x, y)$.
- When $t \geq r(x, y)$,
  - $g(x, y, t) = g(x, y, r(x, y))$.

- Define $u: G \times G \rightarrow R$ as ***minimum interaction time*** **between two glues**.

Time-dependent Glue Tiling Assembly Model

Next, we illustrate the time-dependent model with an example of the addition of a single tile to an aggregate. In a configuration C, when a position (i, j) becomes available for the addition of a tile A, it will stay at (i, j) for a time interval t_0, where $t_0 = \max\{\mu(e(A), w(C(i, j+1))), \mu(n(A), s(C(i+1, j))), \mu(w(A), e(C(i, j-1))), \mu(s(A), n(C(i-1, j)))\}$. Recall that our model requires that if two tiles ever come in contact, they will stay together till the minimum interaction time of the corresponding glues.

After this time interval t_0, if $g(e(A), w(C(i, j+1)), t_0) + g(n(A), s(C(i+1, j)), t_0) + g(w(A), e(C(i, j-1)), t_0) + g(s(A), n(C(i-1, j)), t_0) < \tau$, tile A will detach; otherwise, A will continue to stay at position (i, j).

Step-wise Tiling Assembly Model

- Proposed by Reif in 1999.
- A tiling system under step-wise assembly model is a quadruple

$$S_{step} = <\{T_i\}_{i=1}^{k}, s, g, \{\tau_i\}_{i=1}^{k}>,$$ where

- **k is the number of steps,**
- **T_i is the tile set at step i, and**
- **τ_i is the temperature at step i.**

Step-wise Tiling Assembly Model

Put in T_1 including s. Assemble under temperature τ_1 for long enough time.

step 1 (in tube 1)

↓ terminal product of step 1

⋮

↓ terminal product of step k-1

Put in T_k in. Assemble under temperature τ_k for long enough time.

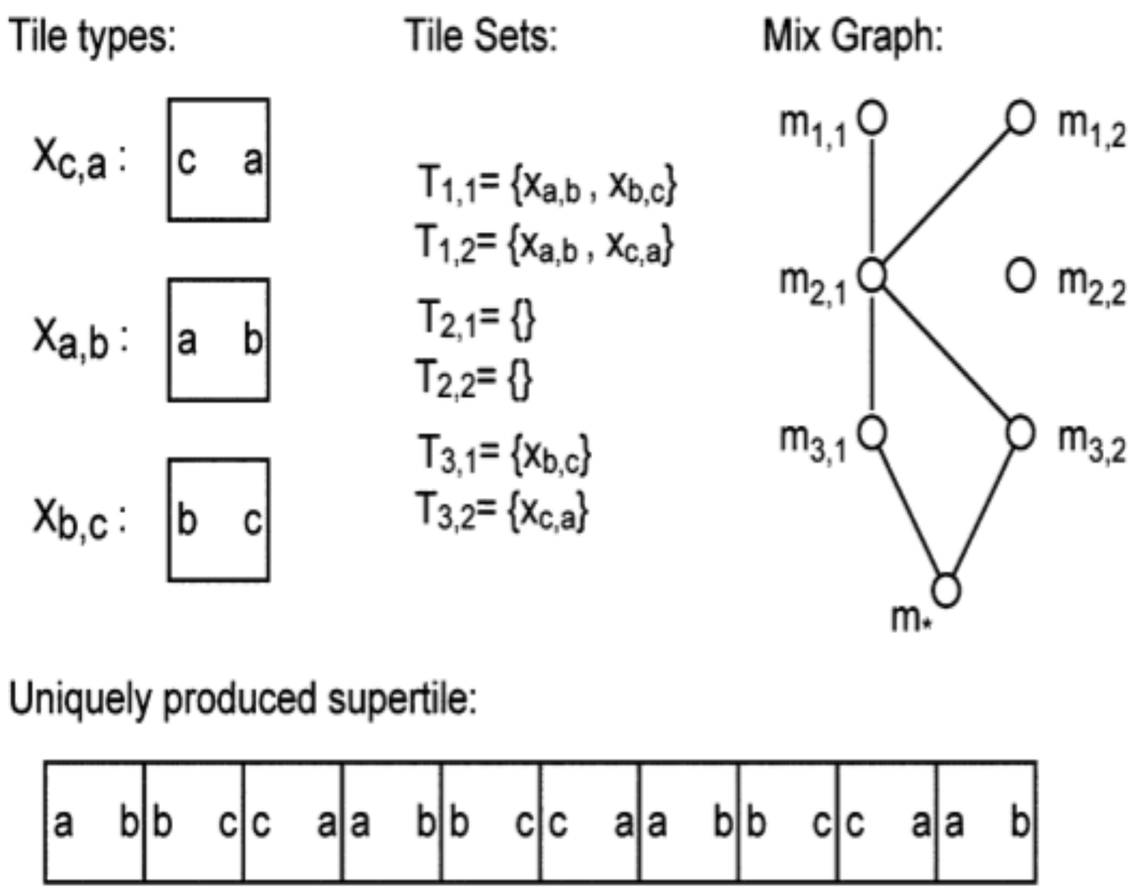
step k (in tube k)

The terminal product of step k is the terminal product of this k-step system.

Staged Tile Assembly Model

- Proposed by ***[Demaine et al. 2007].***
- It is a generalized version of step-wise assembly model first proposed by Reif.
- The assembly process under staged assembly model is shown in the Figure (next page).

Staged Tiling Assembly Model



A sample staged assembly system that uniquely assembles a 1×10 line. The temperature is $\tau = 1$, and each glue a, b, c has strength 1. The tile, stage, and bin complexities are 3, 3, and 2, respectively

- **Vertices of mix graph represent bins for separated assembly reactions.**
- **Each bin has its only tile set and temperature.**
- **Only terminal product of one bin is delivered to bins in next stage.**