DNA Computing by Self Assembly

Presented by
Mohammed Ashraf Ali

Outline

- Self- Assembly
- Examples and behaviour of self-assembly
- DNA Computing
- Tiling Theory
- Nanotechnology
- DNA Self-assembly
- Visualization
- Experiments with self- assembly
- Further research and conclusion

Self- Assembly

Self assembly is the process in which disordered objects come together without any external activity to form a complex structure. First demonstrated by biological and inorganic physical system. Afterwards, a great number of scientists from different disciplines like engineering, biology, mathematics, chemistry, physics stated to research and investigate to design and control the behaviour of self assembling systems

Examples of self assembly

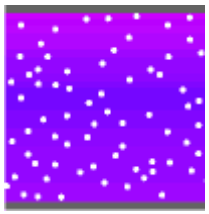
- By mutual attraction between atoms in a specific spatial manner, atoms bind to form *molecules*
- While molecules start mutual sharing of electrons, atoms start to ionize in water. And in absence of water, molecules bind to form *crystal*
- Very small structural units, Cells, are interacting to form *organism*.
- Even rarely self assemble of heavenly bodies form *astronomical system*

Behaviour of self assembly

Self assembly display different behaviour in different simple setting. According to the forces that cause assembly to happen are classified into three categories that are given below

Assembly by capillary forces:

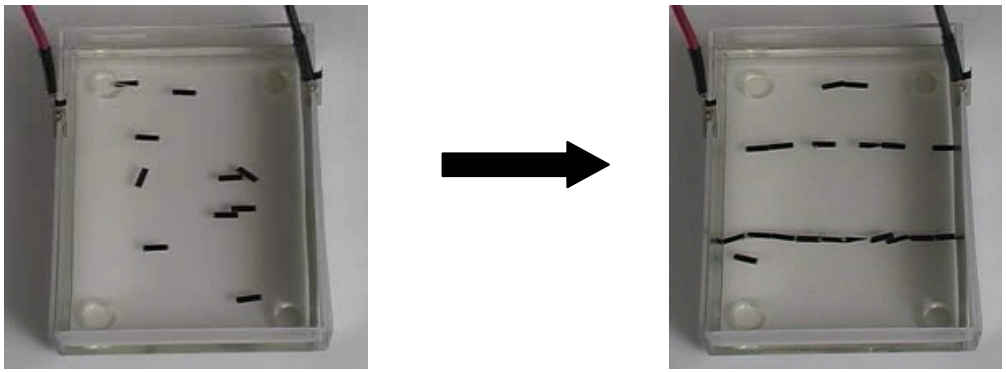
Different types of objects live in water. These small objects feel a force of attraction with other objects that are resting on the surface of water and come in contacts. When two or more objects contact, they remain in contact. This is called capillary forces.



Behaviour of self assembly cont..

Assembly by electrostatic forces:

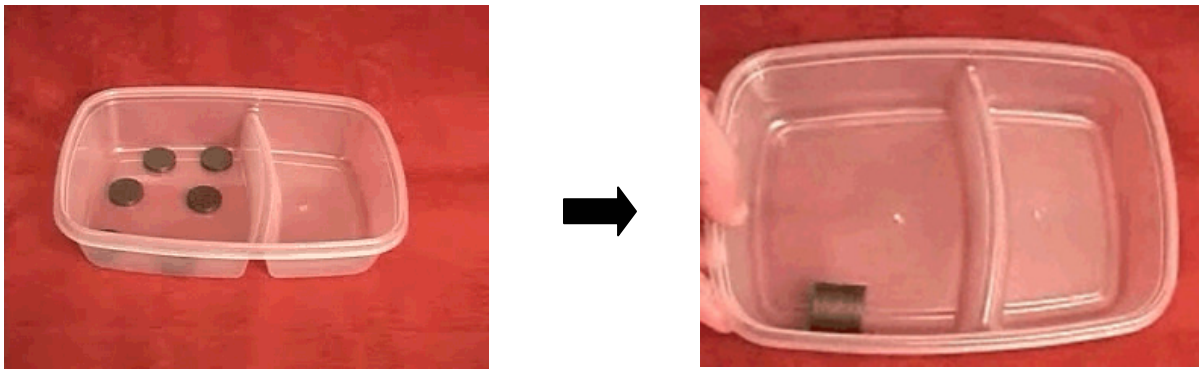
A particle can freely floating in fluid like oil. As the particle and oil has different dielectric properties , their voltage also difference. As a result, particles start to move backwards and forwards to transfer charge and after a time period, they self assemble like a chain. But the precondition is oil depth should be greater than particles length so that particles can move freely.



Behaviour of self assembly cont..

Assembly by magnetic forces:

Have some disk shape magnets close to each other in a container. Shaking the container a little bit is enough to cause formation of contact of particles means magnets in a highly ordered manner. Highly ordered manned applied here as they attach chain like shape.



Molecular Computing

A biochemical algorithm named molecular computation has some characteristics of programming in biochemical systems. Molecular computation can be done by two complementary perspectives:

- Mathematical problems, such as combinatorial search problems can be solved using the astounding parallelism of chemistry; and
- Control molecular processes, such as complex fabrication tasks using biochemical algorithms.

The second one currently appears to be more helpful compare to first one.

Molecular Computing cont…

This two approaches has some common major theoretical issues

- How algorithms with programmable binding interactions can be encoded efficiently in molecules

and

- How algorithms can be make robust to asynchronous and unreliable molecular processes.

Algorithmic Self-Assembly

By combining the following three topics, algorithmic self-assembly can be define

- DNA computing (Adleman, 1994)
- The theory of tiling (Grunbaum and Sheppard, 1986)

  and
- DNA nanotechnology (Seeman, 2003)

DNA Computing

The nervous system can perform sophisticated computations by storage, production and reproduction of genetic information. In this biological organization and processes, genetic information and algorithms lie at the core.

As electromechanical devices are controlled by using electronic microprocessors, molecular and chemical events are controlled by biological organisms.

DNA Computing cont…

According to Leonard Adleman, the direct DNA computing self-assembly can be performed by reactions of programmability of DNA crossing.

- In the first step, all possible executable paths through the target graph were gathered by DNA crossing and represented by DNA molecules.
- In the second step, each possible sequence spontaneously occurred in any order produces a double-stranded DNA molecule whose sequence encodes a valid path through the graph.

DNA Computing cont…

But this Leonard Adleman's technique has one problem

- It's linear self-assembly only support to perform simple computations, not any sophisticated computations.

DNA Computing cont…

According to on Wang's (1961, 1962) embedding of computation in geometrical tiling, Winfree in 1996 builds 2D self-assembly

- Turing universal Computation like an algorithm embedded and guided by potentially a periodic crystallization process can be perform by this two dimensional (2D) self-assembly of DNA.

- For this type of computation or molecular fabrication task, consider the set of molecular Wang tiles as program.

DNA Computing cont…

Comparative study of 1D & 2D

1D algorithmic self-assembly	*2D algorithmic self-assembly*
It offers only limited computational power, not construction power	It offers new capabilities for computation and construction
It does not offers physical phenomena or experimental challenges	It offers a new range of physical phenomena and experimental challenges

Tilling Theory

A tiling is an arrangement of a finite set of unique tiles that fit together perfectly in the infinite plane. Here the precondition is, shape of all tiles in a tiling must be unique like all may be octagon or square with unique length

In three dimensions (3D), tiling contains 230 symmetries and in two dimensions (2D), tiling contains 17 symmetries. It allows one to determine whether more symmetry can be added on the plane or not.

DNA Tilling

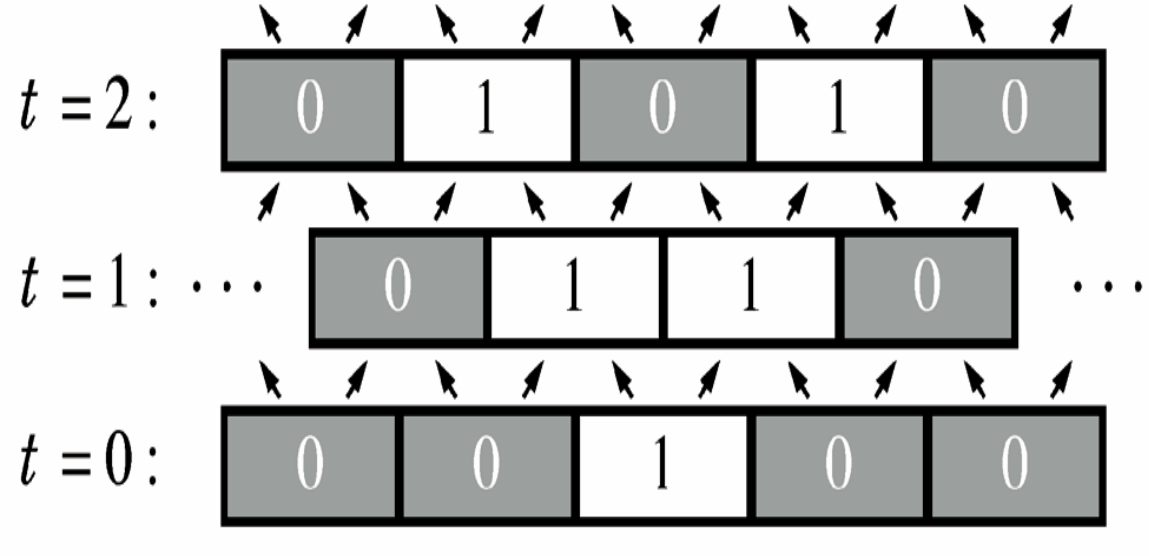
The most interesting and simple cellular automaton rule is the exclusive–or (XOR) function:

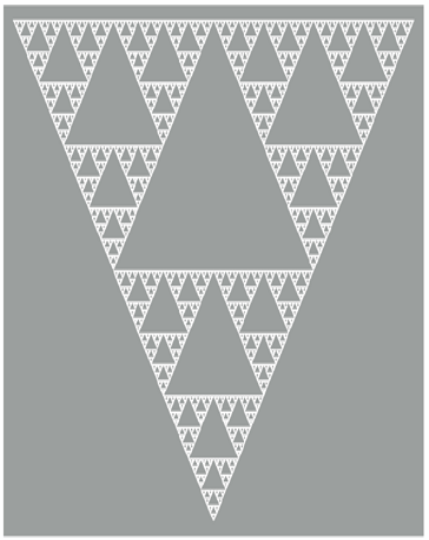
At the beginning stage at time t=0, the row has all cells of '0's except one cell of '1' in center position. Two neighbour tiles are compare according to XOR truth table and the output are picked out. So, at next time slot of space-time history, cells of '1' increase and cells of '0' decrease. With the time slot go on, cells of '1' continue to increase and cells of '0' continue to decrease to have Pascal's triangle (Bondarenko 1993). When assemble is completed, it looks like 'V' shape that is the form of Sierpinski's fractal triangle.

DNA Tilling cont…

XOR truth table

Tile	Ix	Iy	O
1	0	0	0
2	0	1	1
3	1	0	1
4	1	1	0

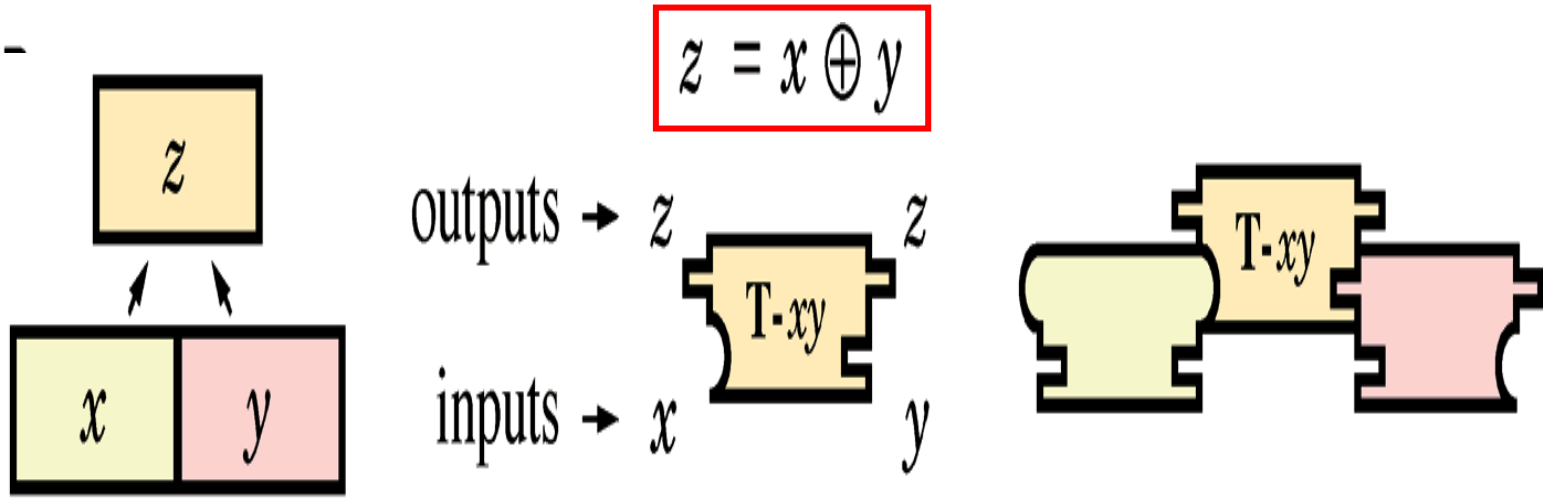




Sierpinski's fractal triangle

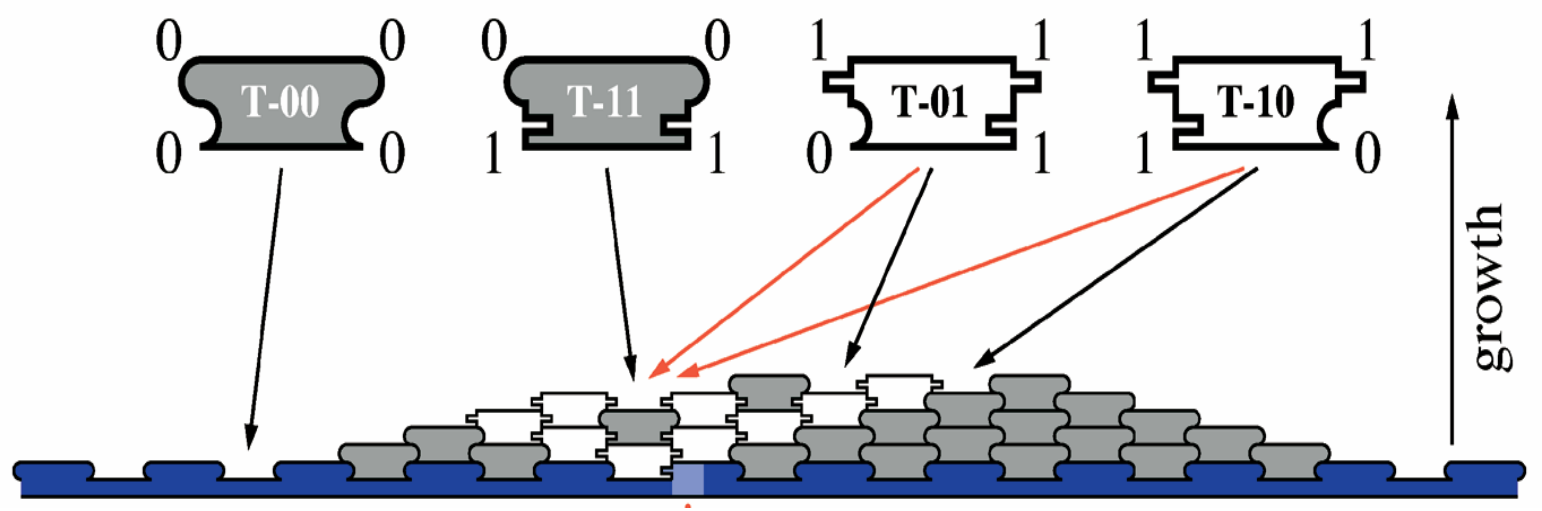
DNA Tilling cont…

By using the cellular automaton rule of exclusive–or (XOR) function, the tiles can be arranged to have tilling. The space–time also related with tilling. The lower portion of tile (i.e. *x* and *y*) represents the input and the upper portion (i.e. *z*) of tile represents the output. The shape of each tile at input side and at output side is different according to value.



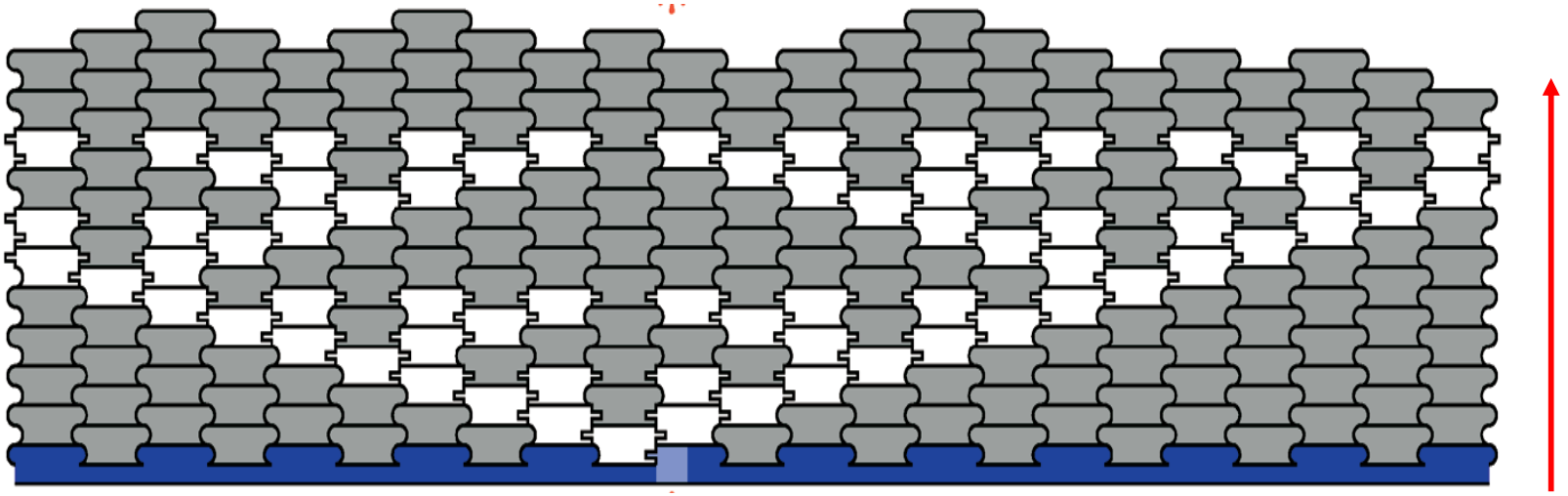
DNA Tilling cont…

The initial state of the cellular automaton means the base line of tilling is called linear input row marked by Blue color. ‘0’ tiles means tiles whose output ‘0’ (T-00 and T-11) marked as grey color and ‘1’ tiles means tiles that output ‘1’ (T-01 and T-10) marked as white color. At t=0, baseline contains all grey tiles except one white in center. Thus, following the above mentioned rules, tiles start to tide with time. Due to unique shape of tiles, wrong tiles can not place marked by red arrow.

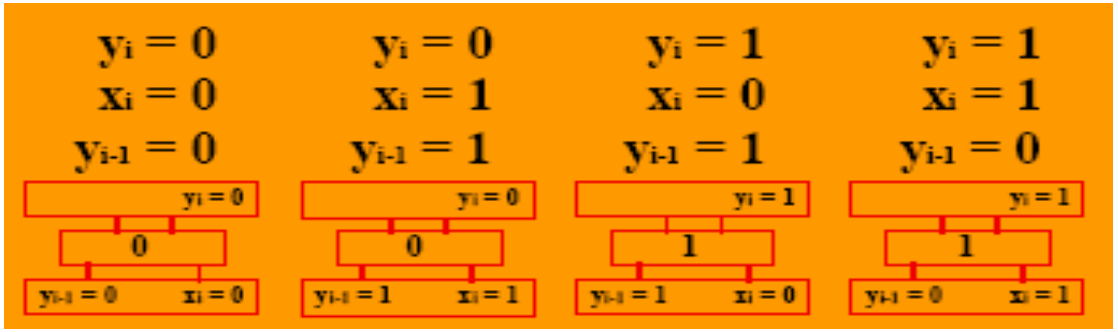


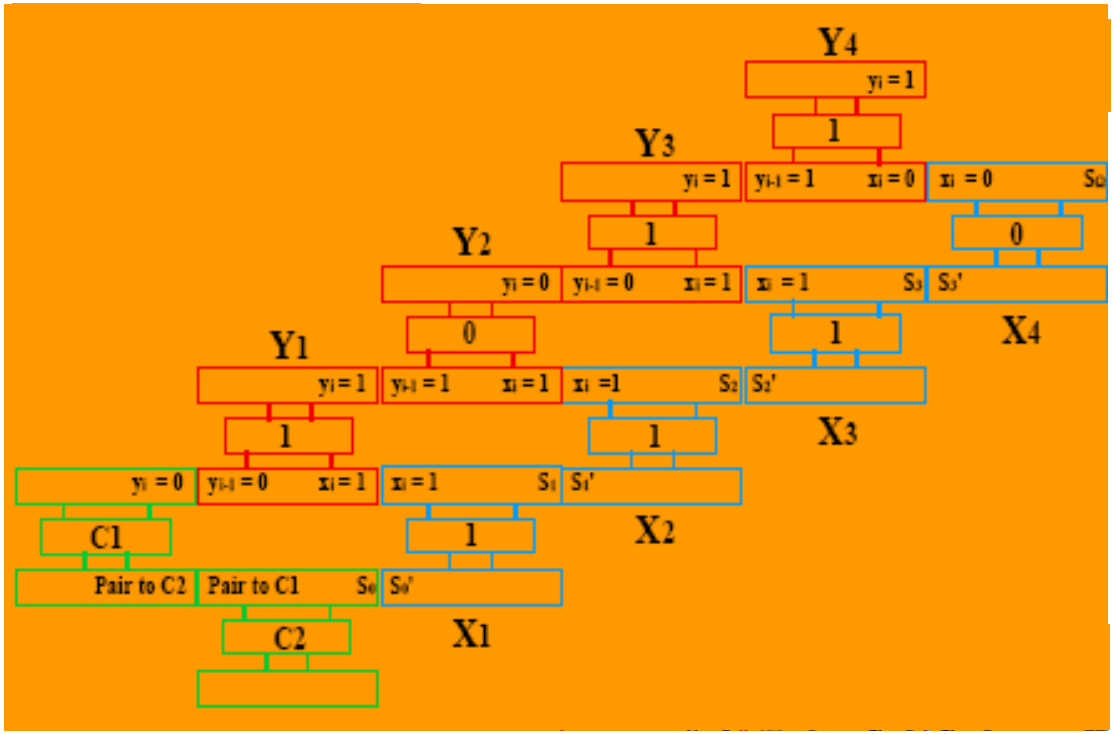
DNA Tilling cont…

Following the cellular automaton rule of XOR function, tilled the tiles one after another with time.
Subsequently, a complete tilling can be model without mismatches or missing tiles



DNA Tilling cont…





According to Thomas H. LaBean, Erik Winfree, and John H. Reif

tiles are tilling diagonally instead of straight up by following the cellular automaton (XOR) function

the input is a sequence of Booleans $X_{1=}$ $X_{2=}$ $X_{3=}1$ and $X_{4=}0$.

the output is a sequence of Booleans $Y_{1=}Y_{3=}Y_{4=}1$ and Y2=0,

To properly bind,
$Y_1 = X_1$ and for i > 1,

$$Y_i = Y_{i-1} \text{ XOR } X_i$$

Thus the ith output is the cumulative XOR of the 1st through the ith inputs

DNA Tilling cont…

The operation of modeling till-base self-assembly is perfect and synchronous in a cellular automaton. But the operation of molecular self-assembly is asynchronous and may have many types of errors. To overcome and avoid creation these error, the following four challenges must be consider [2]:

- To form 2D crystals, abstract tiles have to translate into

molecules means molecular tiles as soon as possible.

- To cope with the logic of chosen abstract tiles,

molecular tiles has to program with specific binding domains.

DNA Tilling cont…

- To assemble followed by correct order and to protect

errors, the molecular tiles binding must be cooperative. And

- These molecular tiles should be assemble on a specified nucleating structure and inhibited spurious nucleation.

These properties are compulsory to implement XOR cellular automaton as well as 1D cellular automaton

DNA Tilling cont…

In 1960s, Hao Wang discovered that the tiling problem is provably unsolvable.

To prove his comments, Wang start tilling a set of tiles and put together uniquely to reproduce the space-time. And this method are developed by Turing machine, in such a way that,

If the Turing machine become stops with an output, then the tiling also becomes stuck;

If the Turing machine continuously compute and does not stop, then an ordered tiling will come out.

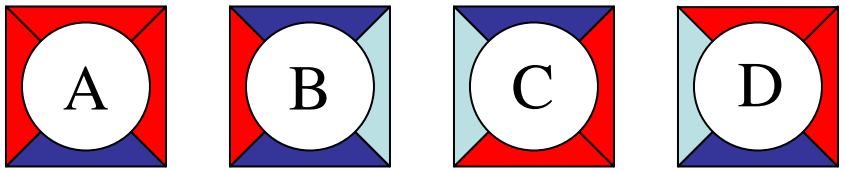
DNA Tilling cont…

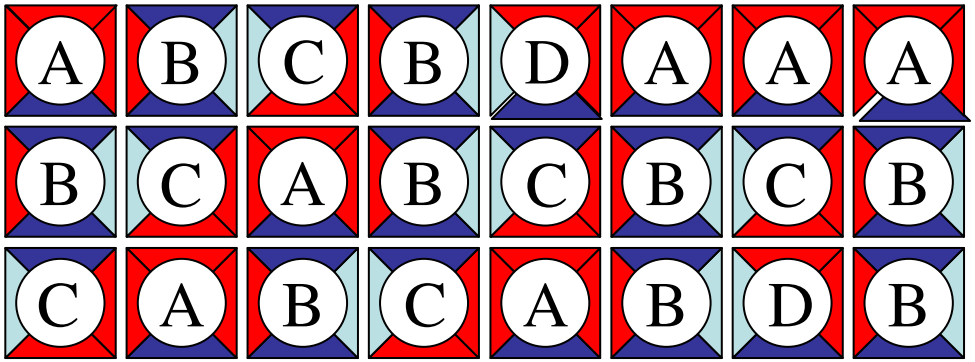
But Wang define some precondition for tiling:

- All the tiles should be same size
- All the tiles should be square shape
- Each tile has four side- top, bottoms, left and right.
- Only the glue (color) of each side can be different
- One tile match with another tile while their touching edge has unique glue (color)
- The tiles can not be rotate

DNA Tilling cont…

Finite set of tiles





Nanotechnology

The nanotechnology are used to figure out the architecture of DNA

- In biological perspective, DNA can be structured as hairpins and three or four way branch points
- In program perspective, DNA can be structured as hinges and joints, bolts and braces that is bind to each other to have a DNA sequence

Nanotechnology cont…

Some more nanostructure by Seeman and his students

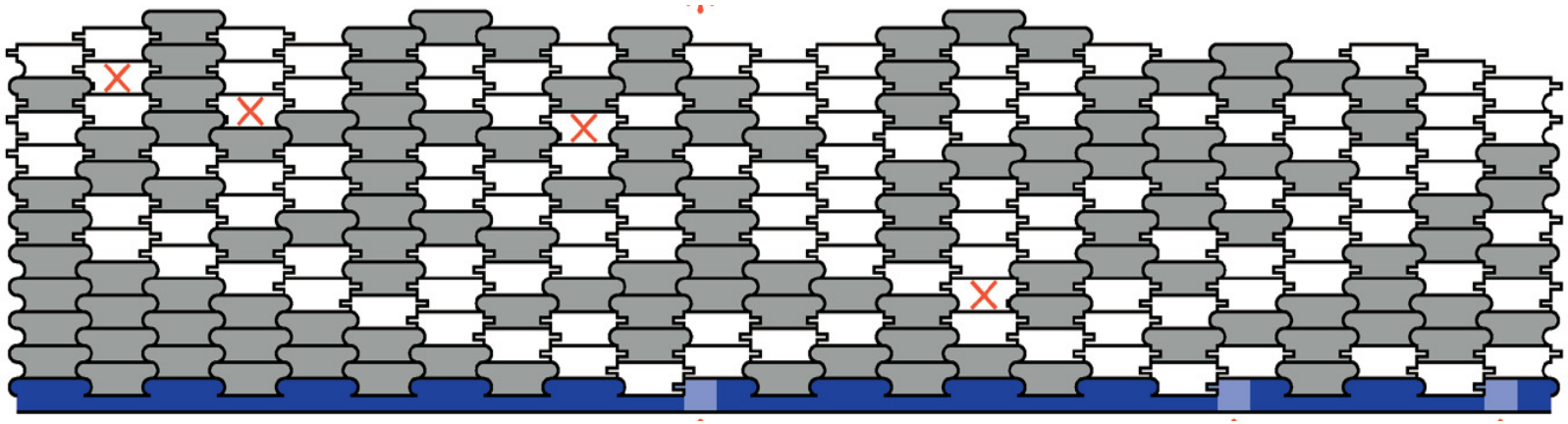
- Wire-frame cube and truncated octahedron
- Single-stranded DNA and RNA knots, including the trefoil, the figure-eight and Borromean rings
- Rigid building-block structures, such as triangles and four-armed "bricks" known as doublecrossover (DX) molecules
- More

Nanotechnology cont...

The problem of Wang tiles method that we seen earlier,

Tiles match with only the neighbour tiles according to glue. After certain times, the configuration come like no way to proceed

- Without creating mismatch or
- Removing tiles that are not permitted



Nanotechnology cont…

To **solve** the problem, Wang proposed tile assembly model

- One tile binds to a growing assembly with a certain strength (for example 0, 1, or 2) ;
- If this tile bind to the growing assembly with a total strength greater than some threshold (typically 1 or 2) will remain stick;
- If this tile bind to the growing assembly with a total strength less than some threshold means weaker strength will fall off immediately;

Following this process, Turing machines and cellular automata can be simulated

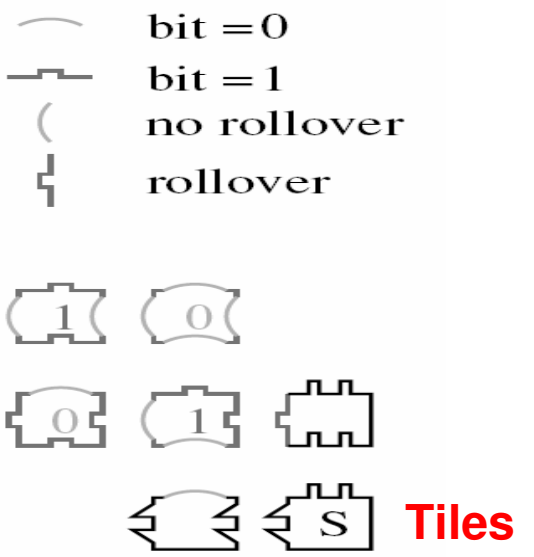
Nanotechnology cont…

Example:

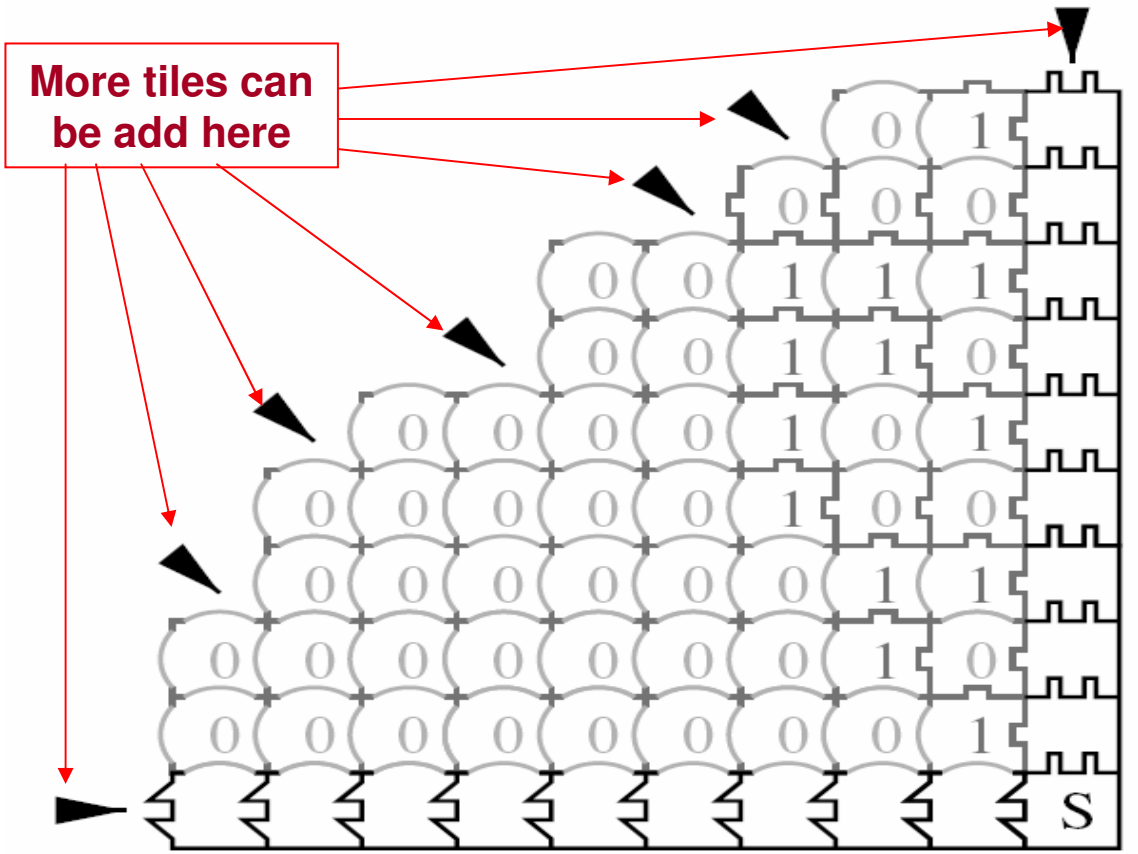
- Consider seven tiles with threshold value 2 with unique shape
- S is the seed tile whose strength value is 2, polymerized to V-shaped to add further
- Select the unique tile that can fit into the corner of V as well as strength value 2. So two tiles of strength value 1 can be added.
- Again new two corners created on the growing assembly. Find two unique tile that can be added on this two corner and so on. The assembly will grow up continuously.

Nanotechnology cont…

tile	c_i	I_{A_i}	I_{B_i}	O_i	c_{i+1}
1	0	0	0	0	0
2	0	0	1	1	0
3	0	1	0	1	0
4	0	1	1	0	1
5	1	0	0	1	0
6	1	0	1	0	1
7	1	1	0	0	1
8	1	1	1	1	1



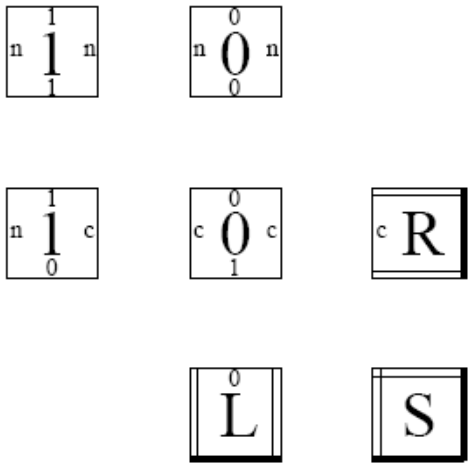
Tiles

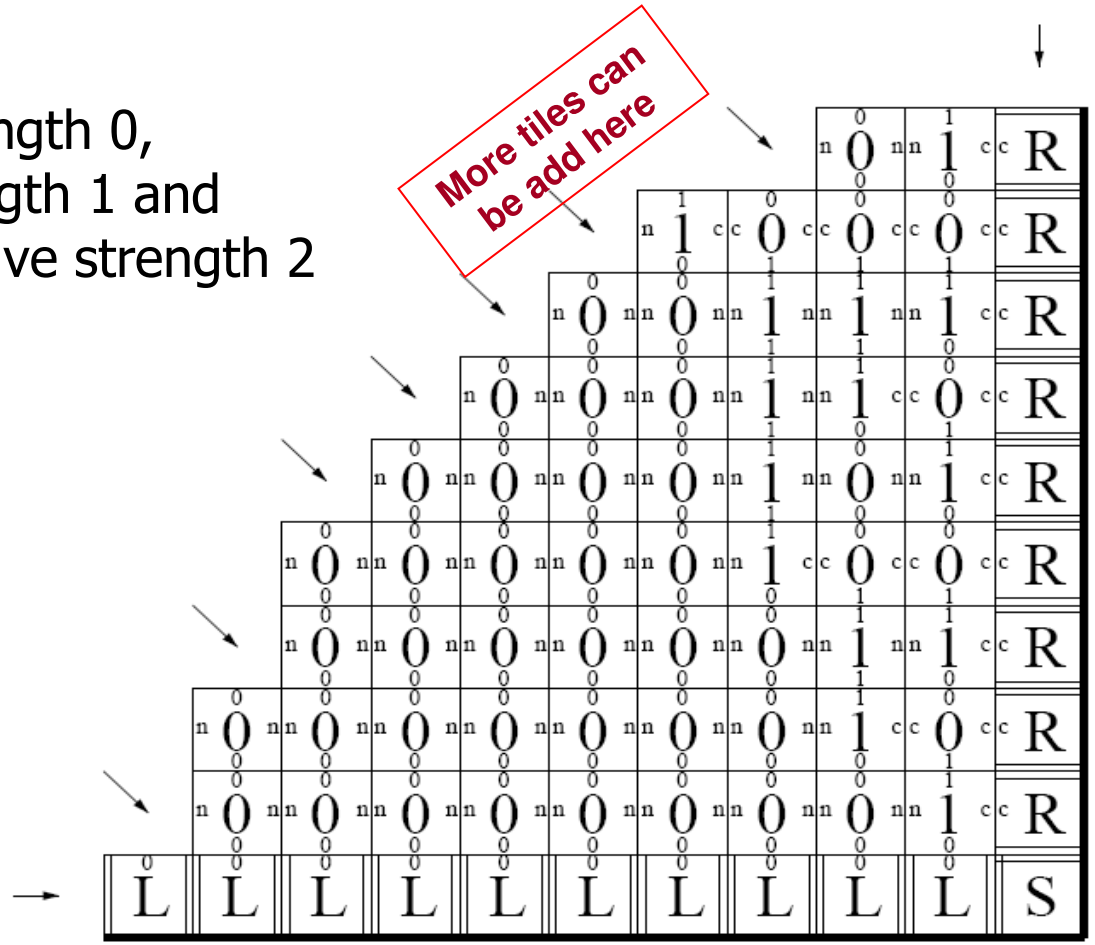


Nanotechnology cont...

According to Paul W. K. Rothemund and Erik Winfree:

Tiles with
thick sides have strength 0,
thin sides have strength 1 and
double-lined sides have strength 2





DNA Self-Assembly

One important concern is:
what classes of DNA produce what classes of languages

According to Chomsky hierarchy

- Regular classes of languages can be produced by self-assembly of duplex DNA while single sticky end bond with the formation of linear DNA complexes.
- 2-metalinear languages can be produced by self-assembly of hairpin and duplex DNA while single sticky end bond with the formation of linear DNA complexes.

DNA Self-Assembly cont…

- Context-free languages can be produced by self-assembly of hairpin, duplex and 3-arm while single sticky end bond with the formation of DNA complexes. and
- Recursively enumerable languages can be produced by self-assembly of Double Crossover (DX) units while double sticky end bond and the temperature is critical with the formation of 2D DNA complexes.

DNA Self-Assembly cont…

Two more classes were defined by Erik Winfree, Tony Eng and Grzegorz Rozenberg

- ✓ $ETOL_{fin}$ languages can be produced by self-assembly of DNA multi-crossover units while multiple simultaneous sticky end bond with the formation of linear multi-helix DNA complexes
- ✓ $ETOL_{ml}$ languages can be produced by restricting the internal tiles coding strands from one side to other side.
- ✓ Another class of language named context-sensitive is not develop yet by DNA self-assembly.

Experiments with Self-Assembly

In the beginning, the simulations proposed the following two presumptions to experiment with self-assembly:

(1) Figure out two-dimensional lattices independently so that the computation can be embedded into them and

(2) Double-crossover molecule contains four binding domains whose activity placed into growing two-dimensional lattice.

But the problem of these two presumption is that they do not support direct experiments with self-assembly

Experiments with Self-Assembly cont…

To overwhelm the problem, Erik Winfree proposed experimental tests of these two presumptions. Each experiment has to go through the following three stages:

- The desired building blocks are framed by following the sequences of design for DNA oligonucleotides
- Synthesis and self-assembly of DNA oligonucleotides passed to building blocks and the subsequent self-assembly of this building blocks passed to create a large molecular structures and
- the resulting molecular structures are then characterized and analyzed for experimental use.

Experiments with Self-Assembly cont…

To complete the experiments, a design sequence has to establish. The reason is multiple of ten number oligonucleotides and thousands of nucleotide is involved.

For evaluate the design sequence, Erik Winfree developed some software tools following some criteria.

The other facility that this software offered are to improve the sequences automatically according to different criteria.

Experiments with Self-Assembly cont…

The first property of the design sequence is a 150-K Dalton molecular system was developed for a growing two-dimensional lattice. In the lattice, two DNA single-stranded are bind to each other for hybridization and the gap between these two DNA is 20 nm. Then the capability of cooperatively of two molecule were tested.

The second property is, a system is designed that can self-assemble two double-crossover molecules into a two-dimensional lattice. Then Gel electrophoresis characterize these two molecules and the resulting lattices are visualized using atomic force microscopy.

Experiments with Self-Assembly cont…

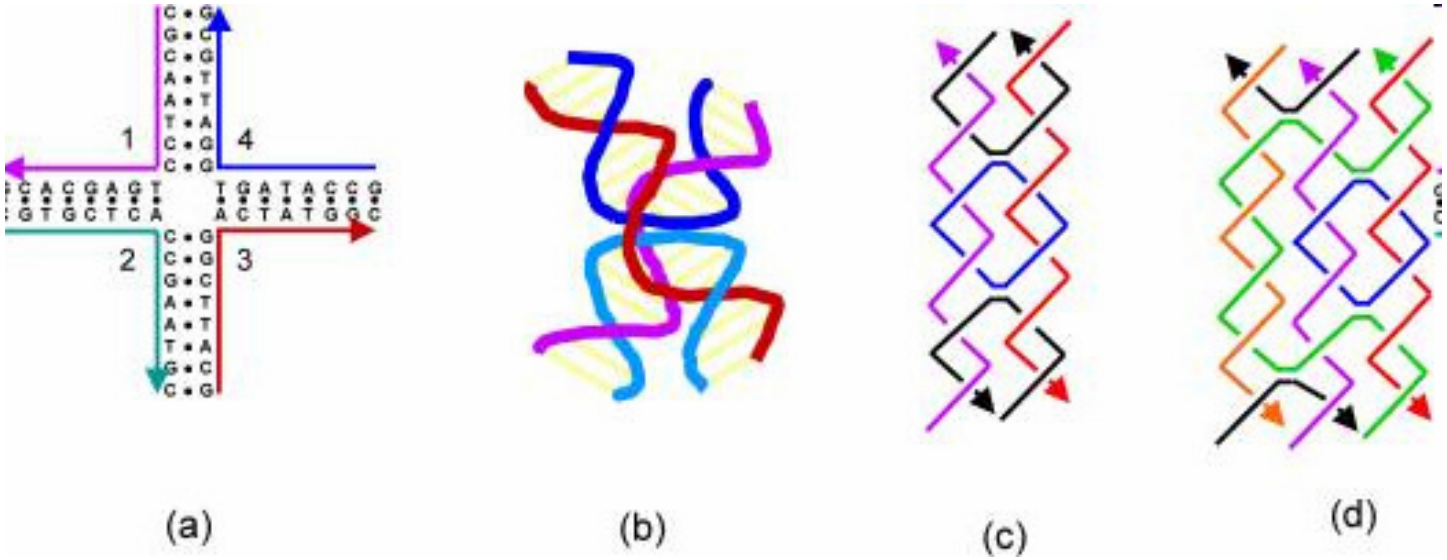
The perfect lattice structures of the self-assembled crystal are obtained by binding one, out of these two double-crossover molecules with a bulky DNA. Name of this bulky DNA is `arm` for a defined period in the atomic force microscope images.

The target molecule had perfect complementarities that are proved by testing the capacity of cooperatively of the two molecules. The perfect lattice structure of the self-assembled crystal is obtained by binding one double-crossover molecules with a DNA that is the core part to implement the Turing-universal model of computation by self-assembly of DNA.

Visualization of DNA tile

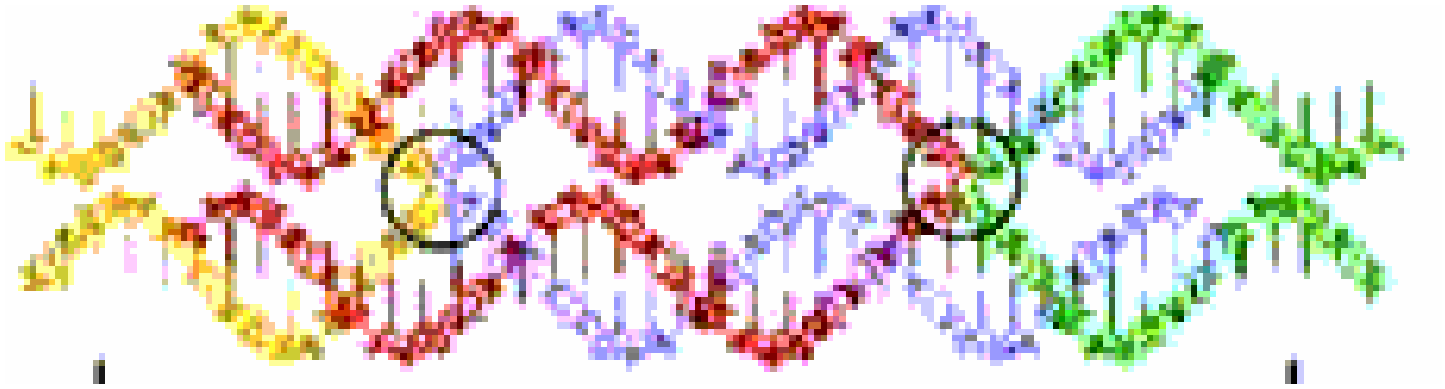
Basic DNA Structures for Self-Assembly

(A) A four-arm junction
(B) Three-dimensional structure
(C) Double crossover (DX) DNA and
(D) Triple crossover (TX) DNA



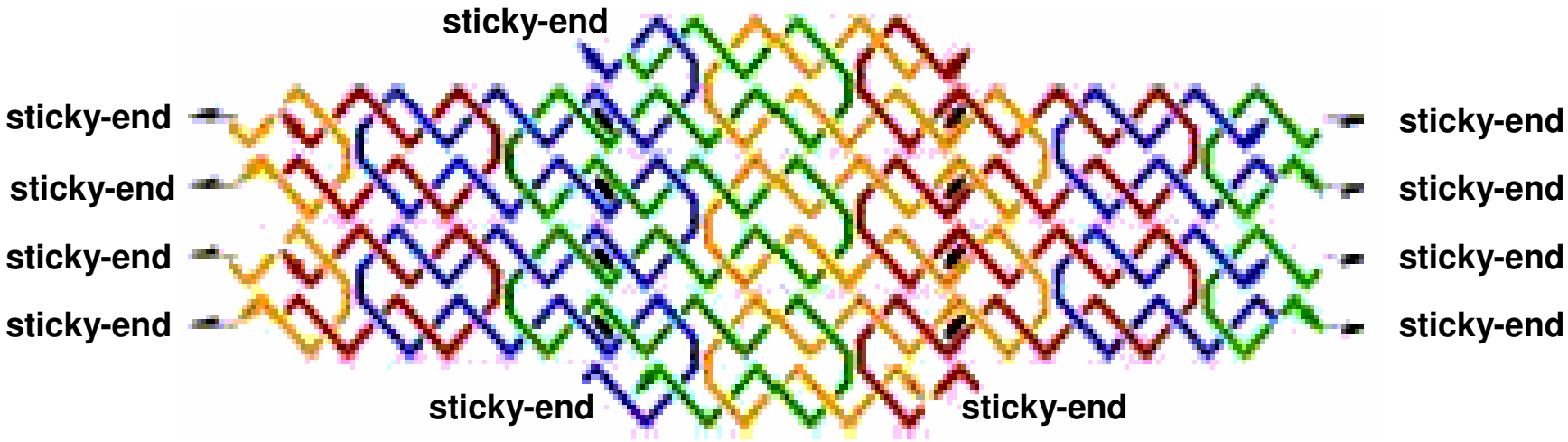
Visualization of DNA tile cont...

The circular regions between red and green; yellow and purple strands represent the crossover regions. The strands (yellow and green), that stick out on both sides, are known as sticky ends. Virtual reality (VR) systems are used to create, display and interact with virtual objects in space using head mounted 3D displays technology.

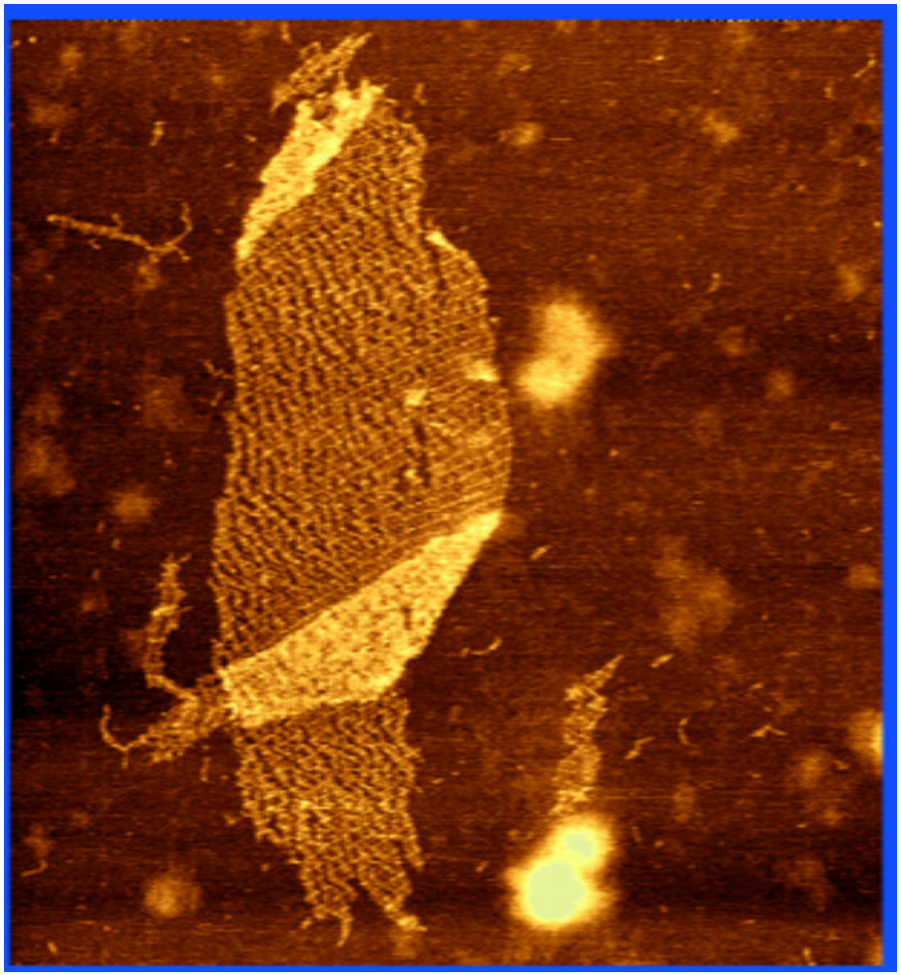


Visualization of DNA tile cont..

DNA tiled network made from individual DNA tiles. The sticky ends come from different tiles put together to form a network of DNA tiles.



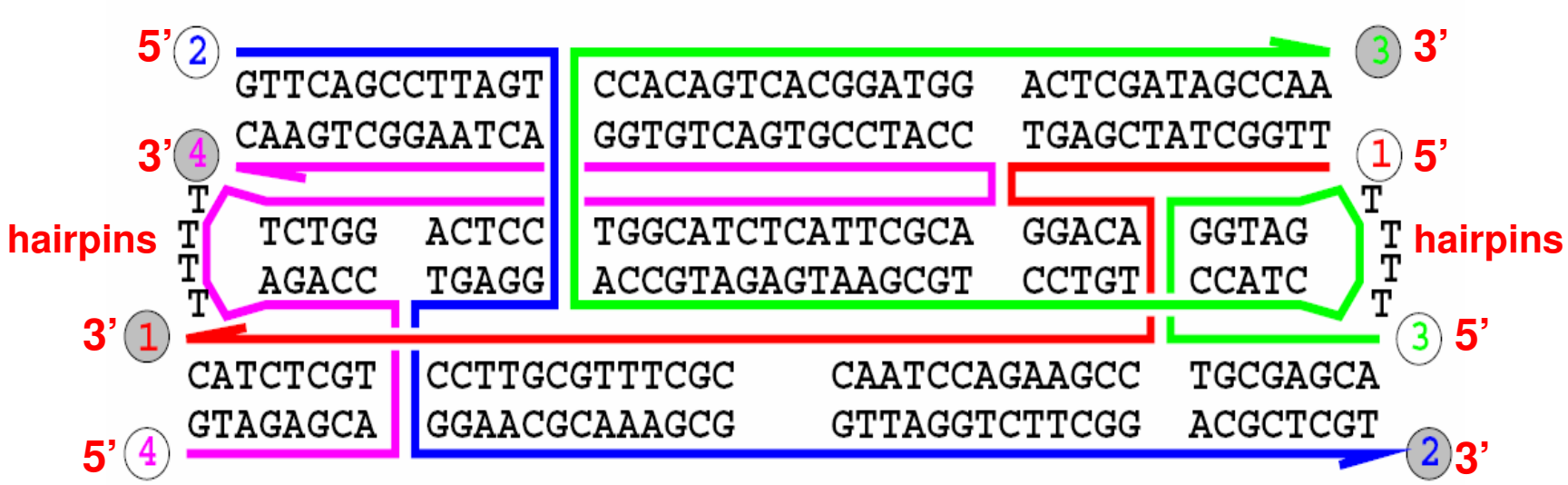
Visualization of DNA tile cont..



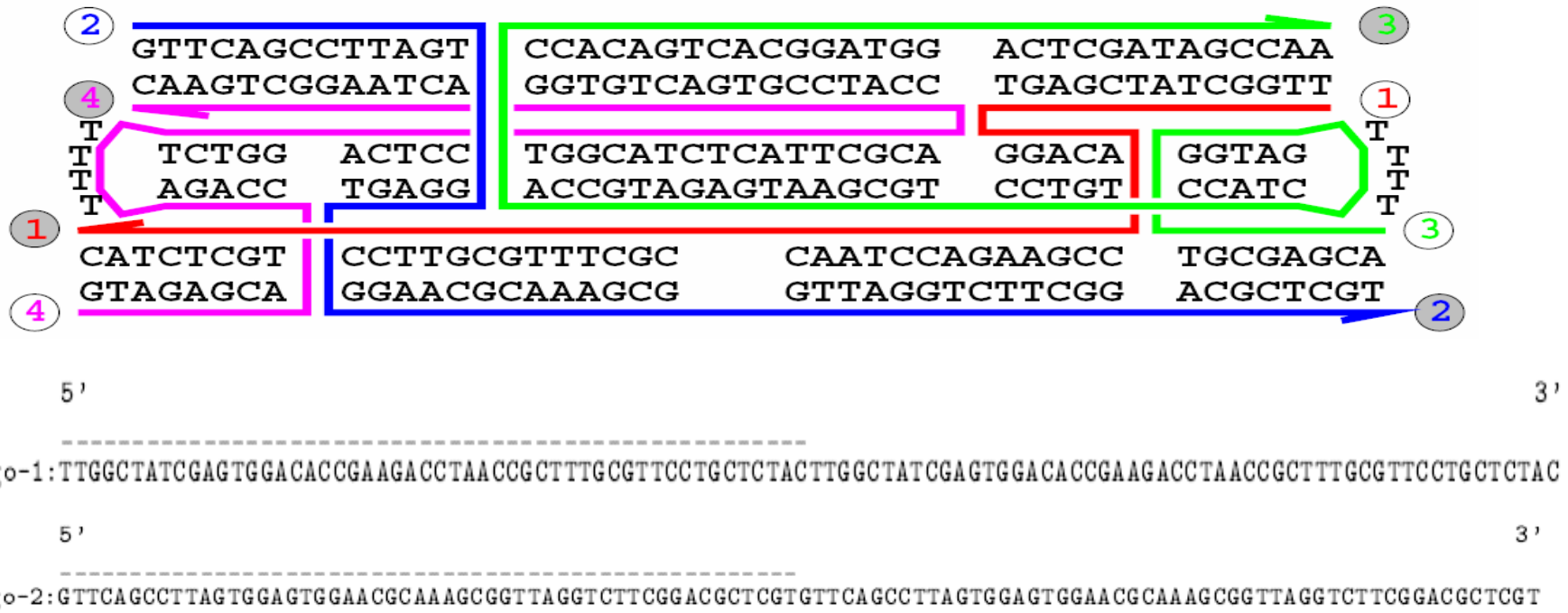
An atomic force microscope takes an image of DNA by two tiles one of which contains an extra loop directed out of the plane. These loops figure out the visible stripe features with approximately 28 nm springs.

Strand and Sequence Trace Through TAO Tile (TTTT)

- The four oligonucleotides are numbered and labelled with white circles on their 5' ends and dark circles on the 3' ends.
- The two TTTT segments are hairpins on the ends of the middle helix.
- Horizontal lines show chain direction and Vertical lines represent the strand crossover points.
- Oligonucleotides 1 (Red) and 2 (Blue) are 52 bases in length; oligonucleotides 3 (Green) and 4 (Purple) are 72 bases long.



Strand and Sequence Trace through TAO Tile cont..



The over lined sequence represents a single copy of tile oligonucleotides which is then repeated (without over lining) to show how two copies of a tile oligonucleotides are held together by a bridge oligonucleotides functioning as a splint.

Application of Self-Assembly

The first application of algorithmic self-assembly is to solve combinatorial optimization problems. It follows the NP-complete problems that are very hard to compute. It used exponential time or parallelism to check properties of generate-and-test form.

So, algorithmic self-assembly tried first to figure out all possible solutions. Then filter all the possible solutions based on predefined information to find one best solution that contains all the desired properties. Then test the best solution whether it can satisfy a small number of simple properties within a short time. If it meets all the criteria as well as it correspond properly, Then It can be say that this is the best final solution.

Application of Self-Assembly cont...

The single self-assembly also used to solve the following problems:

- The Hamiltonian path problem (HPP)

  (Winfree et al., 1998b),
- The Boolean formula satisfiability problem (SAT) (Lagoudakis and LaBean, 2000), and
- Other mathmatical calculations (Reif, 1997)

Application of Self-Assembly cont...

But the question is how much DNA computing can be performing using this self-assembly?

If the SAT has 40-variable problems, about 30 millilitres of DNA can be compute that required few hours

The best possible solution can be achieved while the operation corresponds 1012 bit per second.

Though this definition is fit with chemistry but still some question for electronic computers.

Future work

About 20 years ago, Seeman proposed a conceptual formula to create a three-dimensional DNA tiles with periodic arrays that is still undiscovered.

It may allow some techniques to solve more complicated information-processing in algorithmic self-assembly as well as increase power to step up from one-dimensional to two-dimensional cellular automata i.e. Turing machines

Summary

- Self-assembly algorithm are defined by DNA computing, theory of tiling and DNA nanotechnology
- XOR function used to design simple cellular automaton and tilling
- DNA computing by self-assembly appears to be a robust readily programmable phenomenon.
- Self assembly is not a concrete method for all types of DNA computing
- Comparison between logical structure of self-assembly programs and structure of existing models of computation will proof that algorithmic self-assembly of DNA are reliable.

References

- Erik Winfree, “DNA computing by self-assembly”
- Paul W. K. Rothemund, Nick Papadakis, Erik Winfree, “Algorithmic Self-Assemblyof DNA Sierpinski Triangles”, December 2004 Rothemund et al.
- Erik Winfree, Tony Eng and Grzegorz Rozenberg, “ String tile model for DNA computing by self-assembly”
- Chengde Mao, Thomas H. Labean, John H. Reif & Nadrian C. Seeman, “Logical computation using algorithmic self-assembly of DNA triple-crossover molecules”, *Nature* 407, 493 - 496 (2000) © Macmillan Publishers Ltd.
- Robert D. Barish, Paul W. K. Rothemund, and Erik Winfree,” Two Computational Primitives for Algorithmic Self-Assembly: Copying and Counting”, *Nano Lett.*, 23 November 2005

Thank you