Minimum Tile and Edge-bond Types for Self-Assembled DNA Nanostructures

Laura Beaudin, Jo Ellis-Monaghan*, Natasha Jonoska, David Miller, and Greta Pangborn

VGN NIH INBRE

NATIONAL SECURITY AGENCY

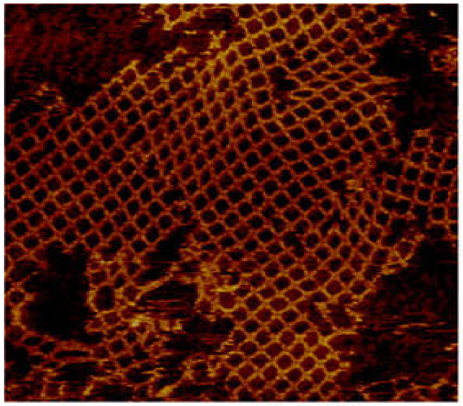
SAINT MICHAEL'S COLLEGE

An Application-Driven Problem

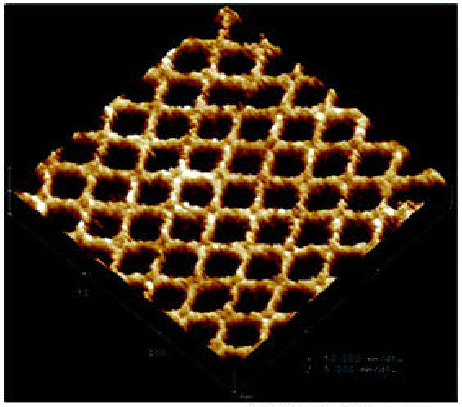
Determine the minimum number of branched junction molecules, and their combinatorial structures, for self-assembling DNA nanostructures.

Why Self-Assembling Nanostructures?

- Biomolecular computing (Hamilton Cycle/3-Sat)
- Nanoelectronics
- Fine screen filters (lattices) at the nano-size scale
- Biosensors and drug delivery mechanisms



500x500 nm

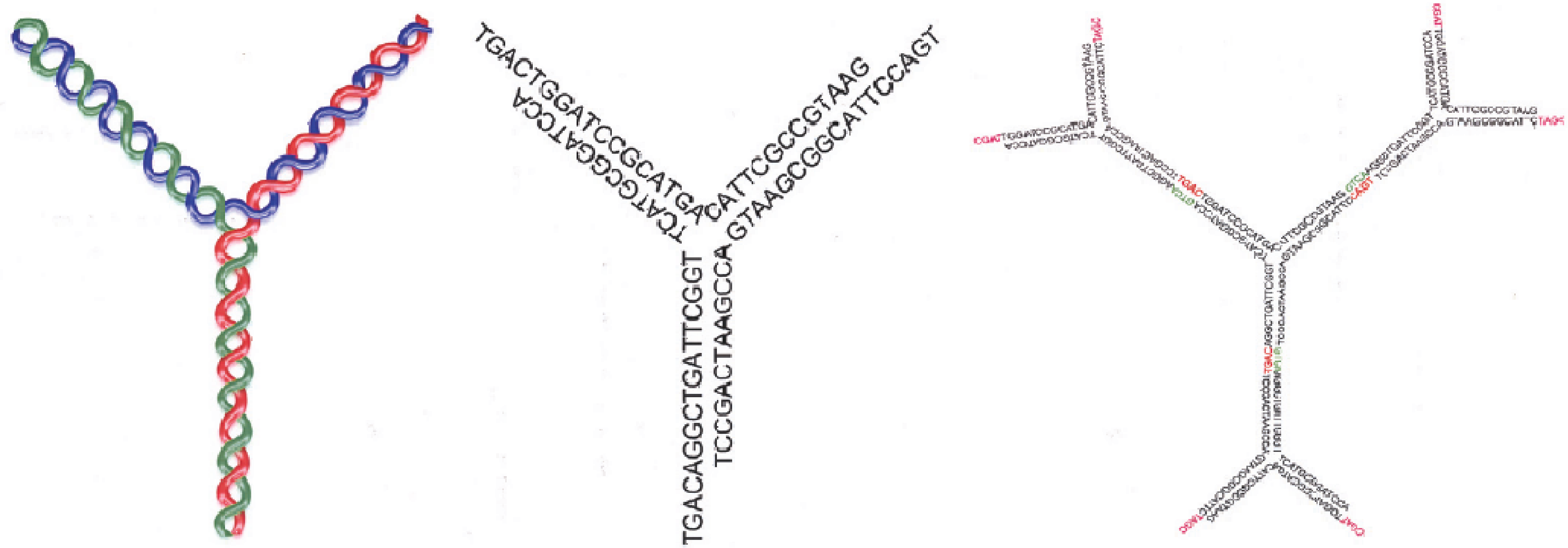


150x150 nm

http://www.nanopicoftheday.org/2004Pics/April2004/DNAmesh.htm

The molecular building blocks

- *K*-armed branched junction molecules

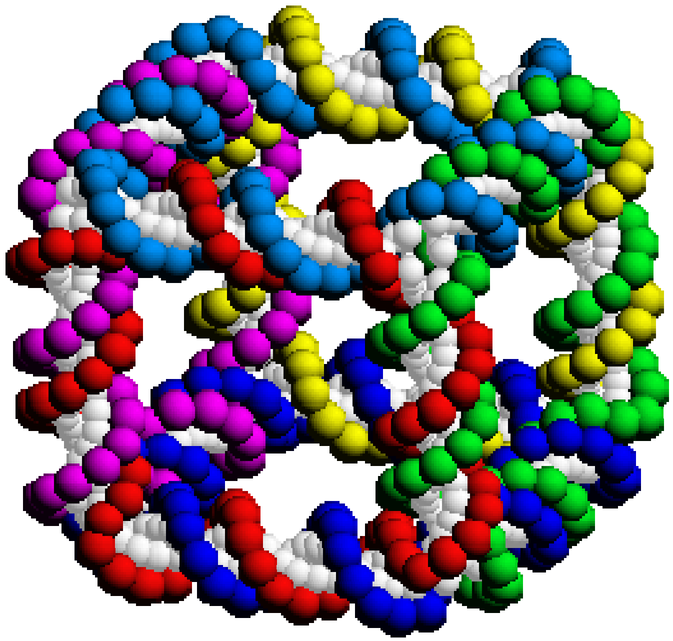


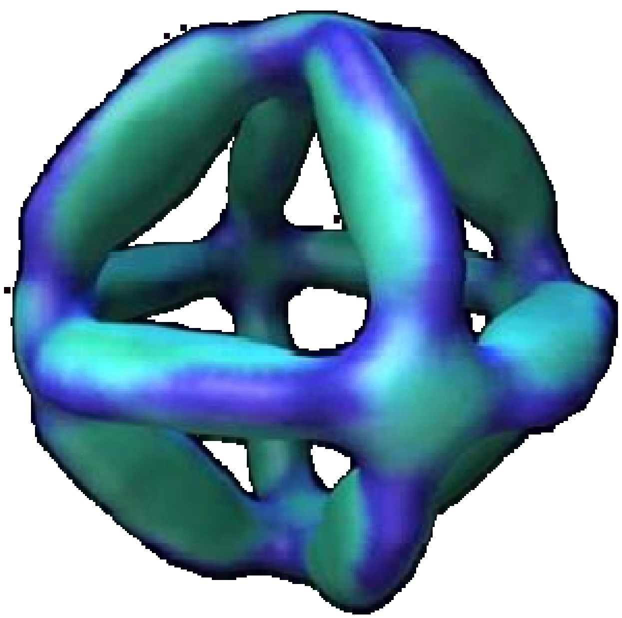
Y-shaped DNA. Schematic diagrams of the structure (left) and sequence (middle) of Y-DNA, and dendrimer-like DNA (right).

D. Luo, "The road from biology to materials," Materials Today, **6** (2003), 38-43

A Complete Complex

- A self-assembled DNA cube and Octahedron





http://seemanlab4.chem.nyu.edu/nanotech.html

The Fundamental Questions

Given a target graph,

- what is the minimum number of k-armed branched junction molecules that must be designed to create the graph?
- What is the minimum number of bond types needed?
- What is the combinatorial structure of the molecules in a minimal set?

Three Different Laboratory Constraints

In solving this problem, we consider the following three scenarios:

1) The incidental construction of a graph smaller than G is acceptable

2) The incidental construction of a graph smaller than G is not acceptable but a graph with the same size as G (same number of edges and vertices) is acceptable

3) The incidental construction of any graph other than G is not acceptable (that is, no smaller graphs nor same-size graphs, where a same-size graph uses all the intended tiles but has the wrong adjacencies)

In all cases, we assume flexible armed molecules (abstract, not embedded, graphs).

Definitions

a $\hat{a}$

ATTCG GGTAACATTCG
TAAGCCCATTG TAAGC

Sticky end types a, b, c, $\hat{c}$, $\hat{a}$, etc. are different types of unadjoined arms sticking off of molecules

Two sticky ends which are able to adjoin such as those of type a and $\hat{a}$, are known as complementary sticky ends

A tile type t represents a *flexible-armed* branched junction molecule specified by a set of sticky end branches

An edge formed by joining two complementary sticky ends is called a bond-edge type

A pot P is a set of tiles such that for any sticky end type a on a tile in P, a complementary sticky end of type $\hat{a}$ also exists on some tile in P

A complex C is an arrangement of tiles from a pot type P with as many adjoined complementary sticky ends as possible with the given tiles

A complete complex is a complex which has no unadjoined sticky ends hanging off of itself

a $\hat{a}$

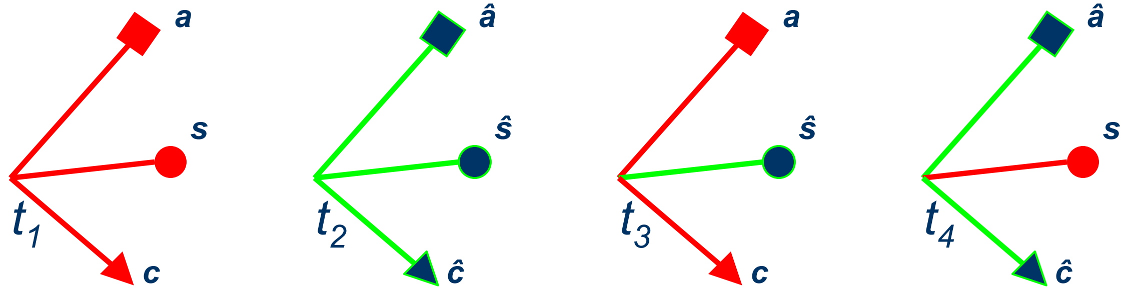
ATTCG GGTAACATTCG

TAAGCCCATTG TAAGC

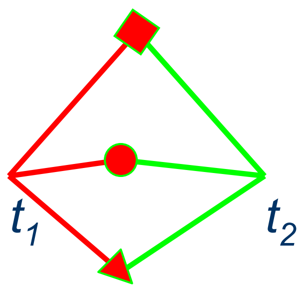
Visual Example

Both complete complexes (1) and incomplete complexes (2) can be constructed by the following pot type P with 4 tiles:

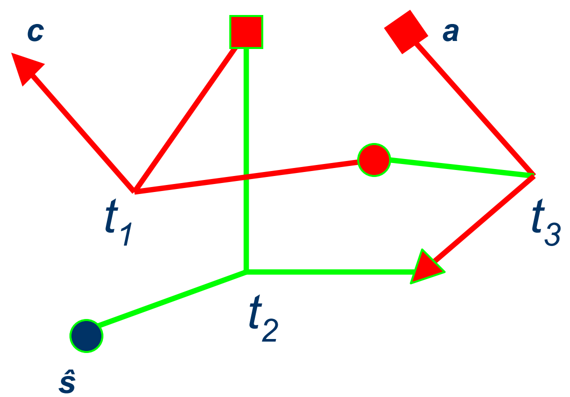
P:



1)



2)



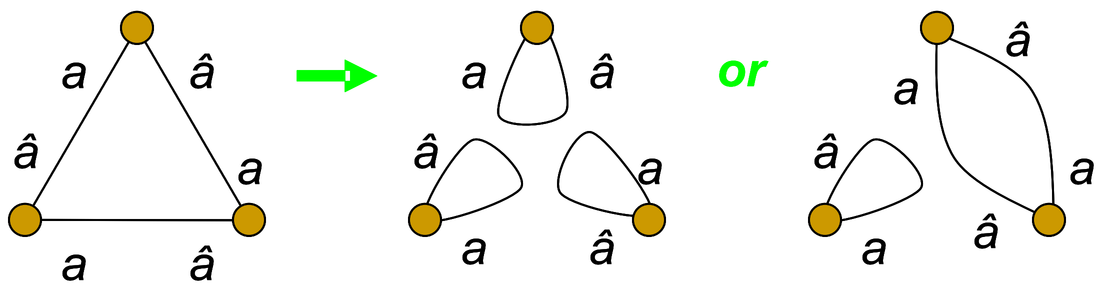
Simple Constraints

- A graph G with m vertices may be constructed from the pot P if and only if the number of hatted sticky ends of each type used in the construction of G equals the number of unhatted sticky ends of the same type that appear in the construction.
- The total number of hatted sticky end types must equal the total number of unhatted sticky end types in a complete complex.

These constraints drive parity arguments.

Some Things to Consider

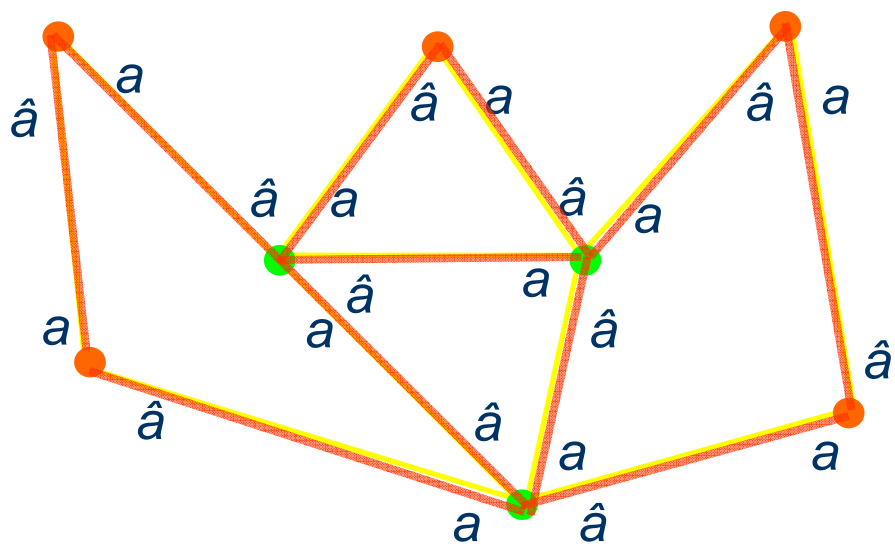
- Since every edge in a graph G represents the connection of two complementary sticky ends, a complete complex will be required to construct G.
- Since no two vertices of different degree can represent the same tile type, the minimum number of tile types need for the construction of G will be greater than or equal to the number of different vertex degrees in G.
- Under the restrictions given by scenario 3, no two adjacent vertices can represent the same tile type because multi-edges and loops could be formed by swapping sticky ends



Scenario 1 Example

The vertex sequence of a graph G be the list of vertex degrees in G.

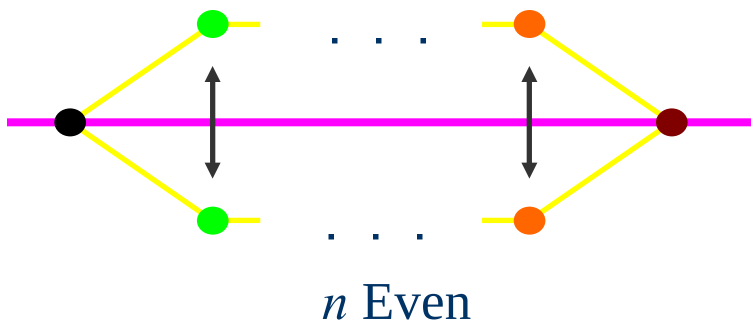
For Eulerian graphs (graphs containing an Euler circuit), the minimum number of tile types required for construction is equal to the number of different digits that appear in the vertex sequence. This can be shown by labeling sticky end types as we follow a graphs Euler circuit (labeling sticky end type *a* for outgoing sticky ends and *â* for incoming sticky ends).

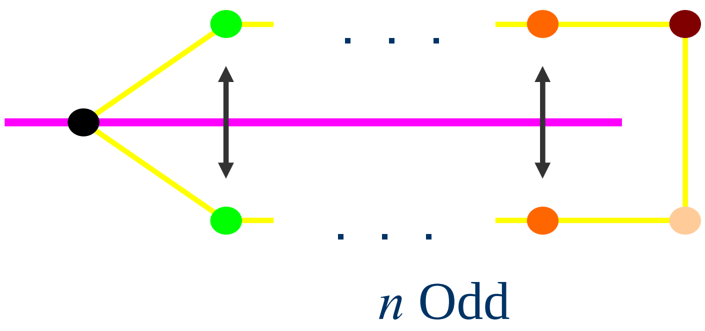


Only 1 bond-edge type is required for Eulerian graphs, and only as many tile types as valencies!

Scenario 2 Example

The minimum number of tile types required to construct a cycle such that no smaller graphs can be constructed out of the tiles is $\left\lceil \frac{n}{2} \right\rceil + 1$ where n is the number of vertices in the cycle C_n.

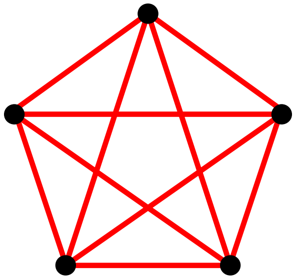




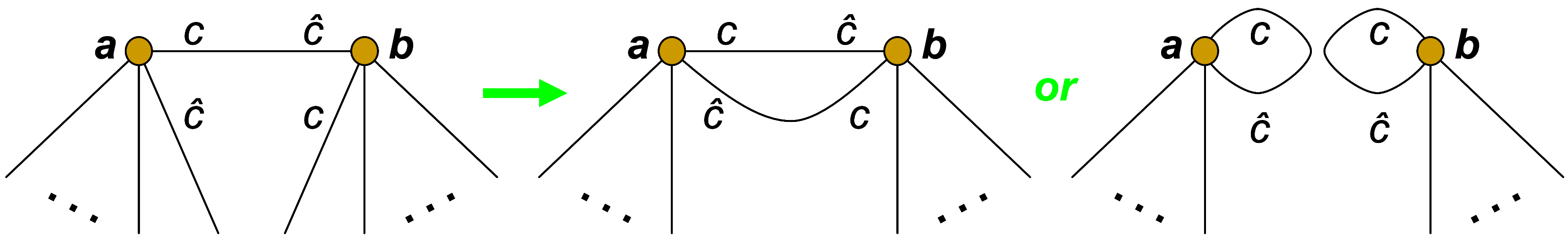
The bisecting line reflects identical tile types

The minimum number of bond-edge types in this case is $\left\lceil \frac{n}{2} \right\rceil$.

Scenario 3 Example



Complete graphs K_n can only be constructed using n tile types and n-1 bond-edge types. Since every vertex in a complete graph is adjacent to every other vertex, no two vertices can represent the same tile type under the constraints of scenario 3. The image below shows the result of two tiles (***a*** and ***b***) of the same type appearing in K_n.



A complex other than K_n is formed!

Table A: Minimum Tile Types

Scenario 1	$T_1(G)$ = minimum number of tile types required if complexes of smaller size than the target graph are allowed.
General graph *G*	The number of different vertex degrees ≤ $T_1(G)$ ≤ the number of different even vertex degrees + 2*(the number of different odd vertex degrees).
Trees	The number of different vertex degrees ≤ $T_1(G)$ ≤ the number of different vertex degrees + 1.
C_n	$T_1(C_n) = 1$.
K_n	$T_1(K_n) = 1$ if n is even, and $T_1(K_n) = 2$ if n is odd.
$K_{n,m}$	$T_1(K_{n,m}) = 1$ if $n=m$ and even, and $T_1(K_{n,m}) = 2$ otherwise.
K-regular graphs	$T_1(G) = 1$ if n is even, and $T_1(G) = 2$ if n is odd.
Scenario 2	$T_2(G)$= minimum number of tile types required if complexes of the same size as the target graph, but not smaller, are allowed.
Trees	$T_2(T)$ = the number of different lesser size subtree sequences.
C_n	$T_2(C_n) = \lceil n/2 \rceil + 1$.
K_n	$T_2(K_n) = 2$ if n is even, and $T_2(K_n) = 3$ if n is odd.
$K_{n,m}$ *with n ≠ m*	$T_2(K_{n,m}) = 2$ if $gcd(m,n)=1$, and $T_2(K_{n,m}) = 3$ if $gcd(m,n)>1$.
$K_{n,n}$	$2 \leq T_2(K_{n,n}) \leq 3$.
Scenario 3	$T_3(G)$ = minimum number of tile types required if complexes of the same size as (or smaller than) the target graph are not allowed.
Trees	$T_3(T)$ = the number of induced subtree isomorphisms.
C_n	$T_3(C_n)= \lceil n/2 \rceil + 1$.
K_n	$T_3(K_n) = n$.
$K_{n,m}$	$T_3(K_{n,m}) = min(n,m)+1$.

Table B: Minimum Bond-Edge Types	
Scenario 1	$B_1(G)$= minimum number of bond-edge types required if complexes of smaller size than the target graph are allowed.
General graph *G*	$B_1(G)$= 1 for all graphs.
Scenario 2	$B_2(G)$ = minimum number of bond edge types required if complexes of the same size as the target graph, but not smaller, are allowed.
Trees	$B_2(T)$ = the number of different sizes of lesser size subtrees.
C_n	$B_2(C_n) = \lceil n/2 \rceil$.
K_n	$B_2(K_n)$ = 1 if n is even, and $B_2(K_n)$ = 2 if n is odd.
$K_{n,m}$	$B_2(K_{n,m})$ = 1 if $gcd(m,n)$=1 , and $B_2(K_{n,m})$ = 2 if $gcd(m,n)$>1.
Scenario 3	$B_3(G)$ = Minimum number of bond edge types required if complexes of the same size as (or smaller than) the target graph are not allowed.
C_n	$B_3(C_n) = \lceil n/2 \rceil$.
K_n	$B_3(K_n) = n - 1$.

Thus far, the same pots have achieved both minimum tile types and minimum bond-edge types, but we don't know if this is always possible.

Pending…

- Various lattices, both 2 and 3 dimensional (as incomplete complexes?)
- Tubes (C_m x P_n) (ditto)
- C_m x C_n
- Various Platonic and Archimedean solids

And a whole other kettle of fish…

- Same set up and questions, but now assume rigid armed molecules—i.e. a fixed rotation (or location) of the sticky end types about a tile vertex.
- Edge-length constraints—because the helixes have to twist, if we call a twist a unit, each edge is of integer length.

Recap of project criteria

- Engaging Application (often collaborative team—networking is important)
- Don’t need a lot of sophisticated theory to at least get started
- Can construct small examples to build intuition
- Development of new mathematical theory

Bibliographical References

N. Jonoska, G. L. McColm, A. Staninska. The Spectrum of a Pot with DNA molecules. University of South Florida Department of Mathematics.

N. Jonoska, G. L. McColm, A. Staninska. The Graph of a Pot with DNA molecules. University of South Florida Department of Mathematics. 2006.

D. Luo, “The road from biology to materials,” Materials Today, **6** (2003), 38-43.

http://www.nanopicoftheday.org/2004Pics/April2004/DNAmesh.htm

http://seemanlab4.chem.nyu.edu/nanotech.html

Acknowledgement:

The project described was supported in by the Vermont Genetics Network through NIH Grant Number 1 P20 RR16462 from the INBRE program of the National Center for Research Resources, and by a National Security Agency Standard Grant.