Aalto University
School of Science
and Technology

Combinatorial optimization in pattern assembly (extended abstract)

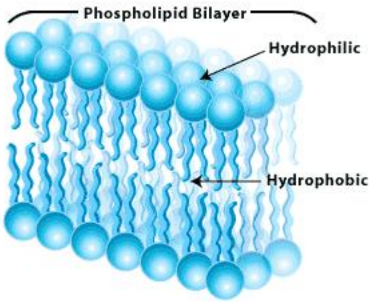
Shinnosuke Seki

Helsinki Institute for Information Technology (HIIT)
Department of Information and Computer Science, Aalto University
shinnosuke.seki@aalto.fi

Talks at UCNC 2013 (Milan, Italy, July 1-5th, 2013)

An example of self-assembly

Lipid bilayer



The water, external environment, affects components (lipids), but does not intend to lead them to the membrane structure.

Other examples of self-assembly

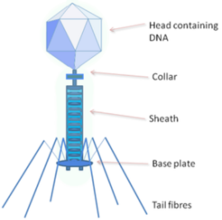
Self-assembly is an omnipresent phenomenon:

snow crystal

garaxy

pattern

city



virus capsid

Molecular assembly

Engineering Goal

DNA self-assembly

Why DNA?

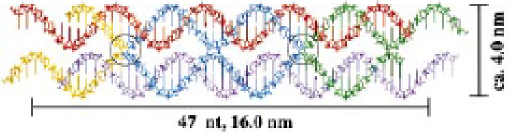
DNA is the molecule of choice in many labs [Doty 2012]:

- it is easy to synthesize
- its physical properties are well-understood
- due to its information-bearing properties

DNA self-assembly

DNA tile implementation

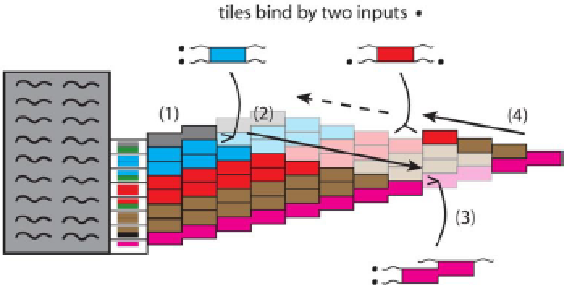
Winfree, Liu, Wenzler, and Seeman implemented **interactive DNA tiles** *in vitro* [Winfree et al. 1998] as a DNA double-crossover molecule:



4 single strands (red, yellow, purple, green), called **sticky ends** provide this “tile” with the capability of interacting with other “tiles”.

DNA self-assembly

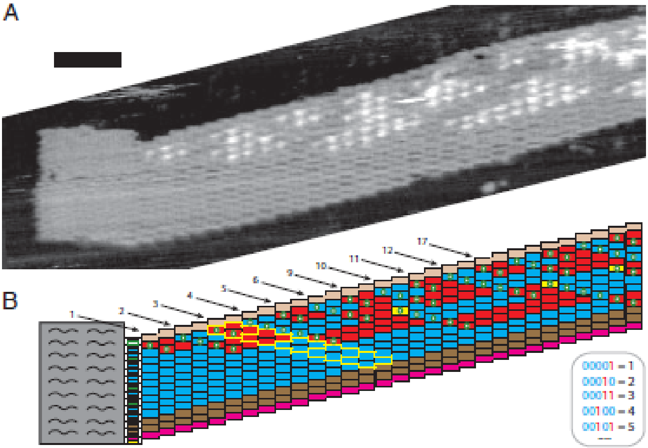
Binary counter assembles from DNA tiles [Barish et al. 2009]



The gray box to the left is the *seed* (starting point of assembly process) made of **DNA origami** [Rothemund 2006].

DNA self-assembly

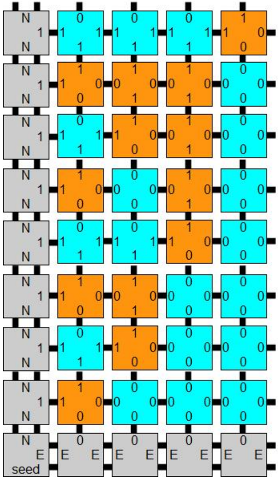
Binary counter assembles from DNA tiles [Barish et al. 2009]



Pattern assembly

What is a pattern?

For $h, w, k \geq 1$, a **k-colored $(w \times h)$-pattern** P is a function from the region $\{(x, y) \mid 0 \leq x < w, 0 \leq y < h\}$ to the color set $\{1, 2, \ldots, k\}$.



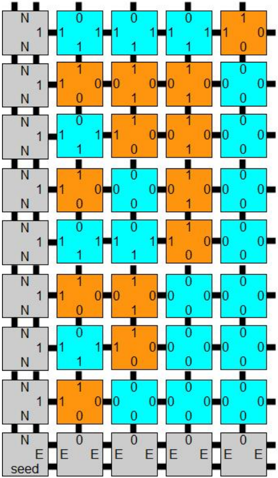
Example

The right is a 3-colored (gray, blue, orange) 5×9-pattern, which is the well-known **binary counter pattern**.

Pattern assembly

What is a pattern?

For $h, w, k \geq 1$, a **k-colored $(w \times h)$-pattern** P is a function from the region $\{(x, y) \mid 0 \leq x < w, 0 \leq y < h\}$ to the color set $\{1, 2, \ldots, k\}$.



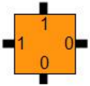
Example

The right is a 3-colored (gray, blue, orange) 5×9-pattern, which is the well-known **binary counter pattern**.

Pattern assembly

Rectilinear TAS

A **tile type** $t \in T$ is a DNA tile abstracted to be a square of specific color:



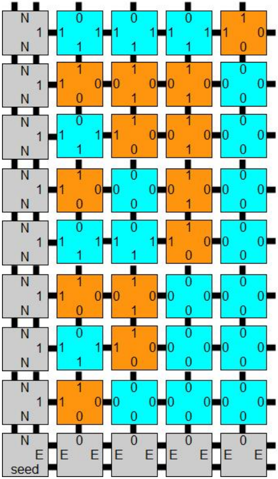
Each side has a label $(0, 1, \dots)$ that models a kind of sticky ends.

Rectilinear Tile-Assembly System (RTAS)

An RTAS is a pair (T, σ_L), where

T a finite set of tile types;

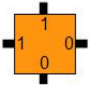
σ_L an L-shape seed.



Pattern assembly

Rectilinear TAS

A **tile type** $t \in T$ is a DNA tile abstracted to be a square of specific color:



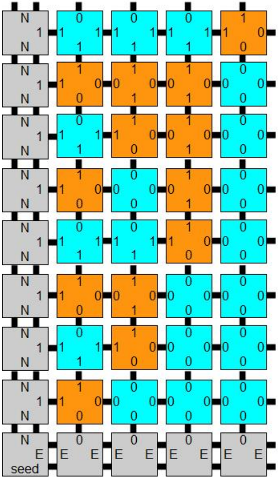
Each side has a label $(0, 1, \dots)$ that models a kind of sticky ends.

Rectilinear Tile-Assembly System (RTAS)

An RTAS is a pair (T, σ_L), where

T a finite set of tile types;

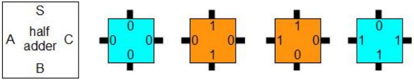
σ_L an L-shape seed.



Pattern assembly

Rectilinear TAS

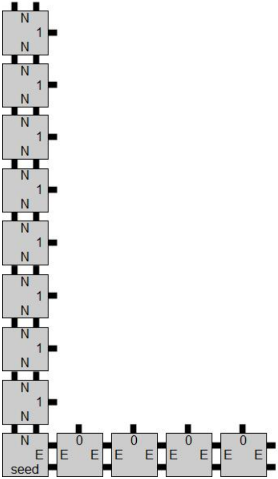
An RTAS with the following 4 tile types



assembles the binary counter pattern.

RTAS' attachment rule

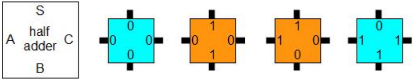
- Assembly begins with the L-shape seed;
- A tile attaches if both of its west and south glues match.
- Assembly proceeds rectilinearly, from south-west to north-east.



Pattern assembly

Rectilinear TAS

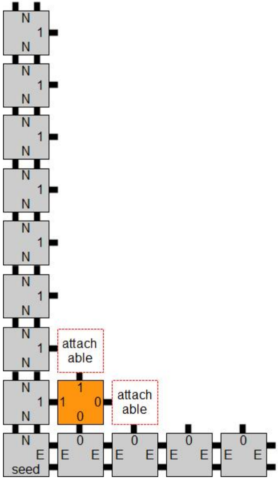
An RTAS with the following 4 tile types



assembles the binary counter pattern.

RTAS' attachment rule

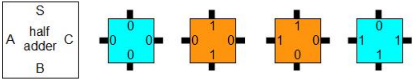
- Assembly begins with the L-shape seed;
- A tile attaches if both of its west and south glues match.
- Assembly proceeds rectilinearly, from south-west to north-east.



Pattern assembly

Rectilinear TAS

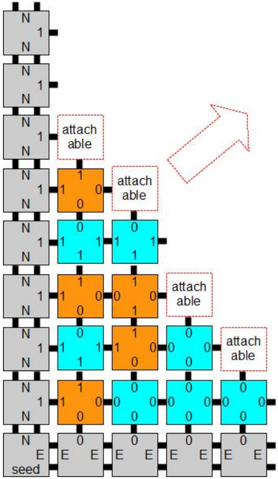
An RTAS with the following 4 tile types



assembles the binary counter pattern.

RTAS' attachment rule

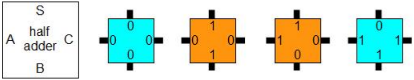
- Assembly begins with the L-shape seed;
- A tile attaches if both of its west and south glues match.
- Assembly proceeds rectilinearly, from south-west to north-east.



Pattern assembly

Rectilinear TAS

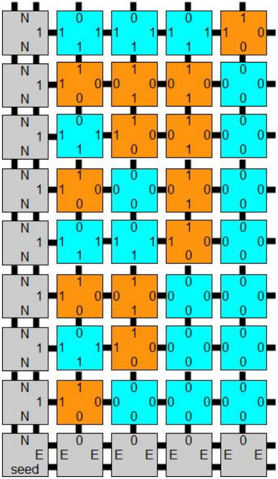
An RTAS with the following 4 tile types



assembles the binary counter pattern.

RTAS' attachment rule

- Assembly begins with the L-shape seed;
- A tile attaches if both of its west and south glues match.
- Assembly proceeds rectilinearly, from south-west to north-east.



Pattern Assembly

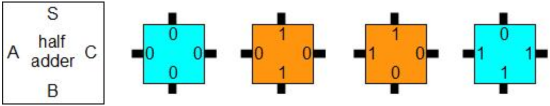
Rectilinear TAS

2-in 2-out logic gates

RTAS can have 4 tile types implement various 2-in 2-out logic gates such as `AND`, `OR`, `XOR`, and half-adder.

Example

The 4 tile types for counter pattern assembly implement **half-adder**.



Directed RTAS

Definition

An RTAS (T, σ_L) is **directed** if for any distinct $t_1, t_2 \in T$,

- their west glues are different, **or**
- their south glues are different.

Unique pattern assemblability

A directed RTAS assembles a unique pattern.

Directed RTAS

Definition

An RTAS (T, σ_L) is **directed** if for any distinct $t_1, t_2 \in T$,

- their west glues are different, **or**
- their south glues are different.

Unique pattern assemblability

A directed RTAS assembles a unique pattern.

PATTERN SELF-ASSEMBLY TILE SET SYNTHESIS (PATS)
[Ma & Lombardi 2008]

GIVEN: a pattern P
FIND: a directed RTAS with min # of tile types
that assembles P.

Theorem ([Czeizler & Popa 2012])

PATS *is* NP*-hard.*

Proof. It is due to a polynomial-time reduction from 3SAT. Its concise proof can be found in [Kari, Kopecki, S. 2013].

Constant colored PATS

Definition

"Any given logic circuit can be formulated as a colored rectangular pattern with tiles, using only a constant number of colors [Czeizler & Popa 2012]".

k-colored PATS(k-PATS)

GIVEN: a ***k*-colored** pattern P

FIND: a directed RTAS with min # of tile types that assembles P.

Constant colored PATS

Our contribution

Main Theorem

60-PATS is NP-hard.

Proof.

A given 3SAT instance ϕ is reduced to a 60-colored pattern $P(\phi)$ in a polynomial-time such that:

ϕ is satisfiable $\iff$ a directed RTAS assembles $P(\phi)$ using at most 84 tile types.

□

Proof idea

$P(\phi)$ consists of the following sub-patterns:

- MAIN CIRCUIT (evaluating 3SAT)
- GADGETS

Proof idea

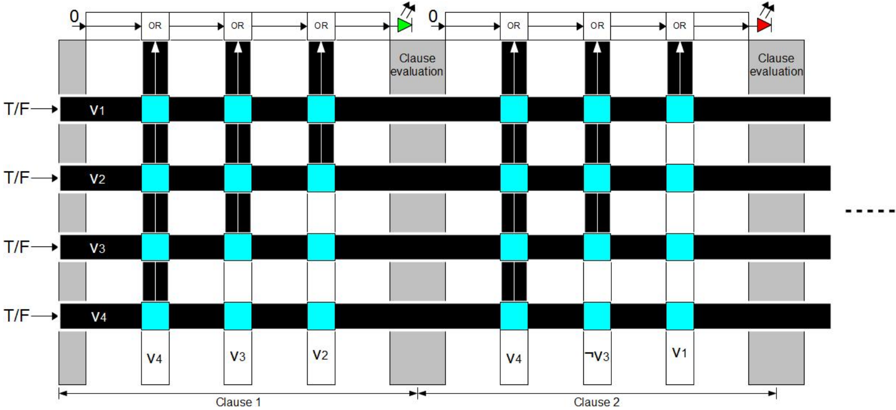
$P(\phi)$ consists of the following sub-patterns:

- MAIN CIRCUIT (evaluating 3SAT)
- GADGETS

Proof idea

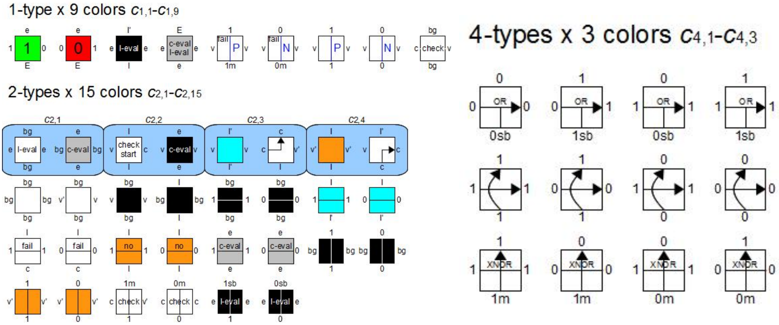
Circuit for evaluating 3SAT instance

The 3SAT instance ϕ can be evaluated using **2-in 2-out** logic gates (OR, XNOR, and intersection).



Proof idea

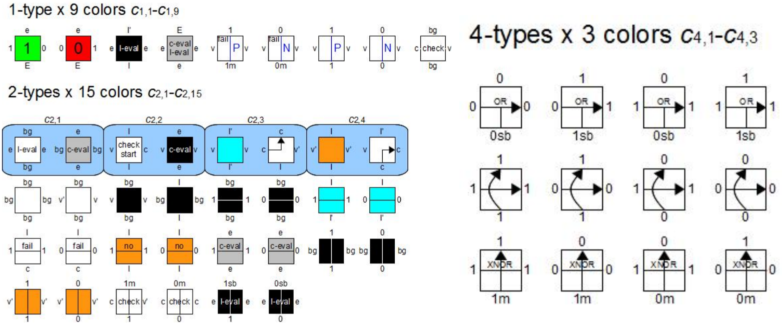
Set T_{3SAT} of 51 tile types for MAIN CIRCUIT



Tiles of these types evaluate ϕ according to the assignment, which are encoded on the L-shape seed σ_L.

Proof idea

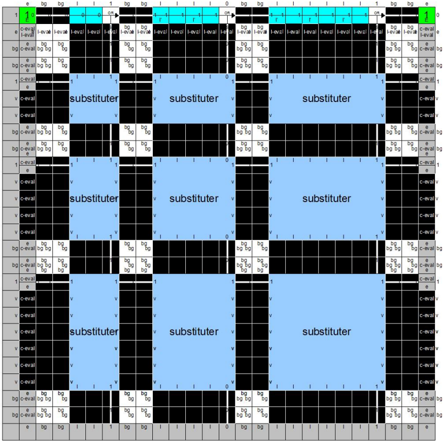
Set T_{3SAT} of 51 tile types for MAIN CIRCUIT



Tiles of these types evaluate ϕ according to the assignment, which are encoded on the L-shape seed σ_L.

Proof idea

MAIN CIRCUIT



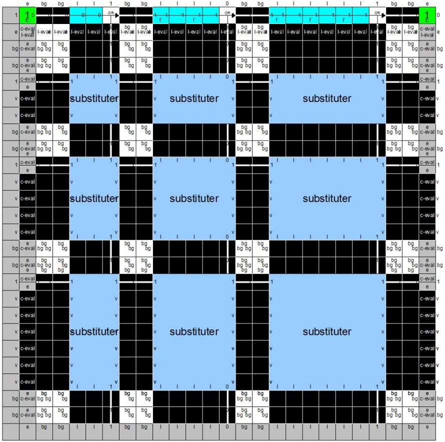
MAIN CIRCUIT is intuitivevly a snapshot of the evaluator circuit when all LEDs flush **green**, that is, ϕ **is satisfied**.

Satisfiability $\rightarrow$ assemblability

If ϕ is satisfiable, then encode a satisfying assignment on the L-shape seed and tiles in T_{3SAT} self-assemble MAIN CIRCUIT.

Proof idea

Main Circuit



Main Circuit is intuitivevly a snapshot of the evaluator circuit when all LEDs flush **green**, that is, ϕ **is satisfied**.

Satisfiability $\rightarrow$ assemblability

If ϕ is satisfiable, then encode a satisfying assignment on the L-shape seed and tiles in T_{3SAT} self-assemble Main Circuit.

Proof idea

Reduction

Satisfiability $\rightarrow$ 84 tile types are enough

84 tile types are enough:

- T_{3SAT} (of 51 tile types) for MAIN CIRCUIT;
- 33 tile types for 33 auxiliary colors (1 type per color) in GADGET.

We need to prove:

Unsatisfiability $\rightarrow$ need for 85 tile types

With T_{3SAT}, some LED position flushes red no matter what seed is, that is, MAIN CIRCUIT would not self-assemble. If 84 were enough, we must find something but T_{3SAT}.

Proof idea

Reduction

Satisfiability $\rightarrow$ 84 tile types are enough

84 tile types are enough:

- T_{3SAT} (of 51 tile types) for MAIN CIRCUIT;
- 33 tile types for 33 auxiliary colors (1 type per color) in GADGET.

We need to prove:

Unsatisfiability $\rightarrow$ need for 85 tile types

With T_{3SAT}, some LED position flushes red **no matter what seed is**, that is, MAIN CIRCUIT would not self-assemble. If 84 were enough, we must find something but T_{3SAT}.

Proof idea

Role of GADGET

GADGET endows $P(\phi)$ with the following property:

Lemma

In order for a directed RTAS to self-assemble $P(\phi)$ using 84 tile types, the set T_{3SAT} must be used.

Corollary

If ϕ is not satisfiable, then no directed RTAS with 84 tile types can self-assemble $P(\phi)$.

Corollary

60-PATS is NP-hard.

Proof idea

Role of GADGET

GADGET endows $P(\phi)$ with the following property:

Lemma

In order for a directed RTAS to self-assemble $P(\phi)$ using 84 tile types, the set T_{3SAT} must be used.

Corollary

If ϕ is not satisfiable, then no directed RTAS with 84 tile types can self-assemble $P(\phi)$.

Corollary

60-PATS is NP-hard.

Proof idea

Role of Gadget

Gadget endows $P(\phi)$ with the following property:

Lemma

In order for a directed RTAS to self-assemble $P(\phi)$ using 84 tile types, the set T_{3SAT} must be used.

Corollary

If ϕ is not satisfiable, then no directed RTAS with 84 tile types can self-assemble $P(\phi)$.

Corollary

60-Pats is NP-hard.

Approximability and PTAS

Corollary

For any $c \geq 60$*, there is no polynomial-time algorithm for* c-PATS *with approximation ratio* $\frac{85}{84}$*.*

Corollary

For any $c \geq 60$*,* c-PATS *does admit any polynomial-time approximation scheme (PTAS).*

Recent Developments

Theorem ([Johnsen, Kao, S. 2013])

29-PATS is NP-hard.

Proof.

This is due to a polynomial-time reduction from SUBSET SUM ψ to a 29-colored pattern $P(\psi)$ as: ψ is summable iff there is a directed RTAS with 46 tile types that self-assembles $P(\psi)$. □

Corollary

For any $c \geq 29$, there is no polynomial-time algorithm for c-PATS with approximation ratio $\frac{47}{46}$.

Corollary

For any $c \geq 29$, c-PATS does admit any PTAS.

Recent Developments

Theorem ([Johnsen, Kao, S. 2013])

*29-*PATS *is* NP*-hard.*

Proof.

This is due to a polynomial-time reduction from SUBSET SUM ψ to a 29-colored pattern $P(\psi)$ as: ψ is summable iff there is a directed RTAS with 46 tile types that self-assembles $P(\psi)$. □

Corollary

For any $c \geq 29$*, there is no polynomial-time algorithm for* c-PATS *with approximation ratio* $\frac{47}{46}$*.*

Corollary

For any $c \geq 29$*,* c-PATS *does admit any PTAS.*

Future Works

- finding the minimum c such that c-PATS remains NP-hard (conjectured to be 2).
- the design of good approximation algorithms for constant colored PATS.

Acknowledgements

Thank you!!

My special thanks are due to

- Ho-Lin Chen (National Taiwan Univ., Taiwan), Da-Jung Cho (Yonsei University, Republic of Korea), David Doty (CalTech, USA), Oscar Ibarra (Univ. California, Santa Barbara, USA), Aleck Johnsen (Northwestern Univ., USA), Nataša Jonoska (Univ. South Florida, USA), Ming-Yang Kao (NU), Jarkko Kari (Univ. Turku, Finland), Lila Kari (Univ. Western Ontario, Canada), Satoshi Kobayashi (Univ. Electro-Communications, Japan), Steffen Kopecki (UWO), Florence Linez (Aalto Univ., Finland), Amirhossein Simjour (UWO), Damien Woods (CalTech), and anonymous referees for valuable discussions, comments, and encouragements;
- HIIT (Helsinki Institute for Information Technology) Pump Priming Project Grant No. 902184/T30606
- Department of Information and Computer Science, Aalto University

References

- R. D. Barish, R. Schulman, P. W. K. Rothemund, and E. Winfree.
  An information-bearing seed for nucleating algorithmic self-assembly.
  PNAS 106(15): 6054-6059, 2009.
- E. Czeizler and A. Popa.
  Synthesizing minimal tile sets for complex patterns in the framework of patterned DNA self-assembly.
  DNA 18, LNCS 7433, pp. 58-72, Springer, 2012.
- D. Doty.
  Theory of algorithmic self-assembly.
  Communications of the ACM 55(12): 78-88, 2012.

References (cont.)

- A. C. Johnsen, M-Y. Kao, and S. Seki.
  Patterned self-assembly tile set synthesis in 29 colors is NP-hard.
  submitted.
- L. Kari, S. Kopecki, and S. Seki.
  3-color bounded patterned self-assembly (extended abstract).
  DNA 19, to appear, 2013.
- X. Ma and F. Lombardi.
  Synthesis of tile sets for DNA self-assembly.
  IEEE T. Comput. Aid. D. 27(5): 963-967, 2008.

References (cont.)

- P. W. K. Rothemund.
  Folding DNA to create nanoscale shapes and patterns.
  Nature 440: 297-302, 2006.
- E. Winfree, F. Liu, L. A. Wenzler, and N. C. Seeman.
  Design and self-assembly of two-dimensional DNA crystals.
  Nature 394: 539-544, 1998.