Concentration Programming:
Randomized Self-Assembly for Approximate Shapes
Outline

• Assembly Model
• Basic Constructions
• Probabilistic Assembly Model
• Main Result
Tile Assembly Model
(Rothemund, Winfree, Adleman)

Tile Set:

Seed Tile:

Glue Function:

Temperature: \( t = 2 \)

\[
T = \begin{align*}
\text{S} & = 2 \\
\text{a} & = 2 \\
\text{b} & = 2 \\
\text{x} & = 2 \\
\text{c} & = 2 \\
\text{d} & = 2 \\
\text{p} & = 1 \\
\text{w} & = 1
\end{align*}
\]
Tile Assembly Model
(Rothemund, Winfree, Adleman)

\[ T = \begin{align*}
& S \quad a \quad b \\
& x \quad c \quad d
\end{align*} \]

\[ \begin{align*}
G(y) & = 2 \\
G(g) & = 2 \\
G(r) & = 2 \\
G(b) & = 2 \\
G(p) & = 1 \\
G(w) & = 1
\end{align*} \]

\[ t = 2 \]
Tile Assembly Model
(Rothemund, Winfree, Adleman)

\[ \begin{align*}
T &= S \quad a \quad b \\
    &\quad x \quad c \quad d
\end{align*} \]

\[
\begin{align*}
G(y) &= 2 \\
G(g) &= 2 \\
G(r) &= 2 \\
G(b) &= 2 \\
G(p) &= 1 \\
G(w) &= 1
\end{align*}
\]

\[ t = 2 \]
Tile Assembly Model
(Rothemund, Winfree, Adleman)

$T = \begin{array}{ccc}
S & a & b \\
x & c & d
\end{array}$

$G(y) = 2$
$G(g) = 2$
$G(r) = 2$
$G(b) = 2$
$G(p) = 1$
$G(w) = 1$

t = 2$
Tile Assembly Model
(Rothemund, Winfree, Adleman)

\[ T = \]

\[ G(y) = 2 \]
\[ G(g) = 2 \]
\[ G(r) = 2 \]
\[ G(b) = 2 \]
\[ G(p) = 1 \]
\[ G(w) = 1 \]

\[ t = 2 \]
Tile Assembly Model
(Rothemund, Winfree, Adleman)

\[ T = \begin{array}{c}
S \\
\text{a} \\
\text{b} \\
\text{x} \\
\text{c} \\
\text{d}
\end{array} \]

\[
\begin{align*}
G(y) &= 2 \\
G(g) &= 2 \\
G(r) &= 2 \\
G(b) &= 2 \\
G(p) &= 1 \\
G(w) &= 1
\end{align*}
\]

\[ t = 2 \]
Tile Assembly Model
(Rothemund, Winfree, Adleman)

\[ T = \begin{align*}
S \quad & a \quad & b \\
\text{x} \quad & c \quad & d
\end{align*} \]

\[
G(y) = 2 \\
G(g) = 2 \\
G(r) = 2 \\
G(b) = 2 \\
G(p) = 1 \\
G(w) = 1
\]

\[ t = 2 \]
Tile Assembly Model
(Rothemund, Winfree, Adleman)

\[ T = \begin{bmatrix}
S & a & b \\
x & c & d
\end{bmatrix} \]

\[
\begin{align*}
G(y) &= 2 \\
G(g) &= 2 \\
G(r) &= 2 \\
G(b) &= 2 \\
G(p) &= 1 \\
G(w) &= 1
\end{align*}
\]

t = 2
Tile Assembly Model
(Rothemund, Winfree, Adleman)

\[
T = \begin{array}{ccc}
\text{S} & \text{a} & \text{b} \\
\text{x} & \text{c} & \text{d}
\end{array}
\]

\[
T = \begin{array}{ccc}
\text{d} & \text{x} \\
\text{c} & \text{x} & \text{x} \\
\text{S} & \text{a} & \text{b}
\end{array}
\]

\[
G(\text{y}) = 2 \\
G(\text{g}) = 2 \\
G(\text{r}) = 2 \\
G(\text{b}) = 2 \\
G(\text{p}) = 1 \\
G(\text{w}) = 1
\] 

\[
t = 2
\]
Tile Assembly Model
(Rothemund, Winfree, Adleman)

\[ T = \]

\[
\begin{array}{ccc}
S & a & b \\
x & c & d \\
x & x & x \\
\end{array}
\]

\[ G(y) = 2 \]
\[ G(g) = 2 \]
\[ G(r) = 2 \]
\[ G(b) = 2 \]
\[ G(p) = 1 \]
\[ G(w) = 1 \]

\[ t = 2 \]
How efficiently can you build an n x n square?
How efficiently can you build an $n \times n$ square?
How efficiently can you build an $n \times n$ square?

Tile Complexity: $2n$
How efficiently can you build an n x n square?
How efficiently can you build an $n \times n$ square?

- Use $\log n$ tile types to seed counter:

```
0 0 0 0 0
```

$\log n$
How efficiently can you build an $n \times n$ square?

- Use $\log n$ tile types capable of Binary counting:

- Use 8 additional tile types capable of binary counting:

```
0 0 0 0 0
```

$\log n$
How efficiently can you build an $n \times n$ square?

- Use $\log n$ tile types capable of Binary counting:

- Use 8 additional tile types capable of binary counting:
How efficiently can you build an n x n square?

- Use $\log n$ tile types capable of binary counting:
- Use 8 additional tile types capable of binary counting:
How efficiently can you build an n x n square?

- Use log n tile types capable of Binary counting:

- Use 8 additional tile types capable of binary counting:
How efficiently can you build an \( n \times n \) square?

**Tile Complexity:** \( O(\log n) \)
(Rothemund, Winfree 2000)

With optimal counter:
**Tile Complexity:** \( O(\log n / \log \log n) \)
(Adleman, Cheng, Goel, Huang 2001)

Meets lower bound:
\( \Omega(\log n / \log \log n) \)
(Rothemund, Winfree 2000)
Outline

• Assembly Model
• Basic Constructions
• **Probabilistic Assembly Model**
• Main Result
Probabilistic Assembly Model
(Becker, Remila, Rapaport)

Assign Relative Concentrations:

d %5

x %5

s %5

a %60

b %20

c %5
Probabilistic Assembly Model
(Becker, Remila, Rapaport)

Tileset =

\[
\begin{align*}
G(y) &= 2 \\
G(g) &= 2 \\
G(r) &= 2 \\
G(p) &= 1 \\
G(w) &= 1 \\
t &= 2
\end{align*}
\]
Probabilistic Assembly Model
(Becker, Remila, Rapaport)

Tileset =

\[ \begin{align*}
S & \quad 5
d & \quad 5 \\
S & \quad 5 \\
b & \quad 20 \\
c & \quad 5 \\
\end{align*} \]

\[ \begin{align*}
G(y) & = 2 \\
G(g) & = 2 \\
G(r) & = 2 \\
G(p) & = 1 \\
G(w) & = 1 \\
t & = 2
\end{align*} \]

(Becker, Remila, Rapaport)
Probabilistic Assembly Model
(Becker, Remila, Rapaport)

Tileset =

\[ \begin{align*}
\text{d} & \quad \%5 \\
\text{x} & \quad \%5 \\
\text{S} & \quad \%5 \\
\text{a} & \quad \%60 \\
\text{b} & \quad \%20 \\
\text{c} & \quad \%5 \\
\end{align*} \]

\[ G(y) = 2 \]
\[ G(g) = 2 \]
\[ G(r) = 2 \]
\[ G(p) = 1 \]
\[ G(w) = 1 \]

\[ t = 2 \]
Probabilistic Assembly Model
(Becker, Remila, Rapaport)

Tileset =

\[
\begin{align*}
& d \quad \%5 \\
& x \quad \%5 \\
& S \quad \%5 \\
& a \quad \%60 \\
& b \quad \%20 \\
& c \quad \%5 \\
\end{align*}
\]

\[
G(y) = 2 \\
G(g) = 2 \\
G(r) = 2 \\
G(p) = 1 \\
G(w) = 1 \\
t = 2
\]
Probabilistic Assembly Model

(Becker, Remila, Rapaport)

Tileset =

d %5  x %5
S %5  a %60
b %20  c %5

G(y) = 2
G(g) = 2
G(r) = 2
G(p) = 1
G(w) = 1

t = 2
Probabilistic Assembly Model  
(Becker, Remila, Rapaport)

Tileset =

\[
\begin{align*}
\text{d} & \quad \%5 \\
\text{x} & \quad \%5 \\
\text{S} & \quad \%5 \\
\text{a} & \quad \%60 \\
\text{b} & \quad \%20 \\
\text{c} & \quad \%5 \\
\end{align*}
\]

\[
\begin{align*}
G(y) & = 2 \\
G(g) & = 2 \\
G(r) & = 2 \\
G(p) & = 1 \\
G(w) & = 1 \\
\end{align*}
\]

t = 2
Probabilistic Assembly Model
(Becker, Remila, Rapaport)

Tileset =

\[
\begin{align*}
&d & \text{\%5} \\
&S & \text{\%5} \\
&a & \text{\%60} \\
&b & \text{\%20} \\
&c & \text{\%5}
\end{align*}
\]

Two Terminal Shapes Produced

\[
\begin{align*}
&d \parallel x \\
&S \parallel a \\
&d \parallel b \parallel c
\end{align*}
\]

\[
\begin{align*}
G(y) &= 2 \\
G(g) &= 2 \\
G(r) &= 2 \\
G(p) &= 1 \\
G(w) &= 1 \\
t &= 2
\end{align*}
\]
Probabilistic Assembly Model
(Becker, Remila, Rapaport)

Tileset =

\[
\begin{align*}
\text{d} & \quad \%5 \\
\text{x} & \quad \%5 \\
\text{S} & \quad \%5 \\
\text{a} & \quad \%60 \\
\text{b} & \quad \%20 \\
\text{c} & \quad \%5 \\
\text{S} & \quad \%5 \\
\text{a} & \\
\text{S} & \\
\text{b} & \\
\text{S} & \\
\text{S} & \\
\text{S} & \\
\text{S} & \\
\text{S} & \\
\end{align*}
\]

\[
\begin{align*}
.60/.85 &= \%70.6 \\
.20/.85 &= \%23.5 \\
.05/.85 &= \%5.9
\end{align*}
\]

\[
\begin{align*}
G(y) &= 2 \\
G(g) &= 2 \\
G(r) &= 2 \\
G(p) &= 1 \\
G(w) &= 1 \\
t &= 2
\end{align*}
\]
Probabilistic Assembly Model
(Becker, Remila, Rapaport)

Tileset =

- d: 5%
- x: 5%
- S: 5%
- a: 60%
- b: 20%
- c: 5%

G(y) = 2
G(g) = 2
G(r) = 2
G(p) = 1
G(w) = 1

t = 2

%70.6 %5.9 %23.5

%75 %25

(Becker, Remila, Rapaport)
Generic Tileset for all Squares
Generic Tileset for Approximate Squares

\[(1-\varepsilon)n \quad \text{n} \quad (1+\varepsilon)n\]
Generic Tileset for Approximate Squares

$(\varepsilon, \delta) –$ Approximate Square Assembly

Given:
- $\varepsilon, \delta < 0$
- integer $n$

Design:
A probabilistic tile system that will assemble an $n' \times n'$ square with:

$$(1-\varepsilon)n < n' < (1+\varepsilon)n$$

With probability at least:

$$1 - \delta$$
High Level Idea

- Build random structure:
  - Dimensions are random
  - Internal pattern is random
High Level Idea

- Build random structure:
  - Dimensions are random
  - Internal pattern is random

- Incorporate arithmetic tiles to extract a binary number from the random pattern

10110110
Finish off Square

Output n approximation

Binary Counter

10110110

n

n

n

n
Line Estimation of $n$
Line Estimation of $n$
Line Estimation of n
Line Estimation of $n$
Line Estimation of $n$
Line Estimation of n
Line Estimation of $n$
Line Estimation of n
Line Estimation of $n$
Line Estimation of $n$
Line Estimation of $n$
Line Estimation of \( n \)
Line Estimation of n

\[ s \]

\[ x \% \frac{1}{n} \]

\[ x \% \frac{(n-1)}{n} \]
Line Estimation of $n$

Length has Geometric distribution with $p = \frac{1}{n}$

$E[\text{Length}] = n$
Line Estimation of $n$

[Becker, Rapaport Remila, 2006]

$E[\text{Length}] = n$
Line Estimation of $n$

[Becker, Rapaport Remila, 2006]

- Assembles all $n \times n$ squares.
- Expected dimension specified by percentages.
- Geometric distribution: Large variance

$$E[\text{Length}] = n$$
Improved Estimation of $n$: Binomial Distribution

Key Idea
Improved Estimation of n: Binomial Distribution

Probability of placing a red tile given either a red or green tile is placed:

\[
1/n
\]
Improved Estimation of n: Binomial Distribution

Probability of placing a red tile given either a red or green tile is placed:

\[ \frac{1}{n} \]

To compute estimation of n:
Compute LENGTH / REDS
Improved Estimation of $n$: Binomial Distribution

Binary Counter

Length: 10000
Reds: 100
Improved Estimation of n: Binomial Distribution

Length: 10000
Reds: 100

Compute Length / Reds: 100

Estimate for n:

Division tiles
Problem: Estimation Line too Long

Chernoff Bounds only yield high accuracy for $\text{Length} >> n$

$\text{Length} >> n$:
Too long for an $n \times n$ square…
Solution: Estimation Frame

Phase 1: Build dimensions of frame.

Multiple lines = HEIGHT: Determined by Geometric Distribution

WIDTH: Determined by Geometric Distribution
Solution: Estimation Frame

Phase 1:
Build dimensions of frame.

Phase 2:
Build Sampling Lines

Multiple lines = HEIGHT: Determined by Geometric Distribution

WIDTH: Determined by Geometric Distribution
Solution: Estimation Frame

Phase 1:
Build dimensions of frame.

Phase 2:
Build Sampling Lines

Phase 3: Sum Reds and Length for each Line

Multiple lines = HEIGHT: Determined by Geometric Distribution

WIDTH: Determined by Geometric Distribution
Solution: Estimation Frame

Phase 1: Build dimensions of frame.

Phase 2: Build Sampling Lines

Phase 3: Sum Reds and Length for each Line

Phase 4: Sum subtotals

Multiple lines = HEIGHT: Determined by Geometric Distribution

WIDTH: Determined by Geometric Distribution
Solution: Estimation Frame

Phase 1: Build dimensions of frame.

Phase 2: Build Sampling Lines

Phase 3: Sum Reds and Length for each Line

Phase 4: Sum subtotals

Phase 5: Compute Length to Reds ratio

Multiple lines = HEIGHT: Determined by Geometric Distribution

WIDTH: Determined by Geometric Distribution

Output Estimation
Solution: Estimation Frame

With high probability:
- \( \text{HEIGHT} < n \)
- \( \text{WIDTH} < n \)

\[ \rightarrow \text{Frame fits within } n \times n \text{ square} \]

With high probability:
- \( \text{HEIGHT} \times \text{WIDTH} \gg n \)

Chernoff Bounds imply:

\[ \rightarrow \text{Estimation is accurate with high probability:} \]

\[ (1 - \epsilon)n' < n < (1 + \epsilon)n' \]
Finish off Square

Output n approximation
Finish off Square

Output $n$ approximation

Binary Counter

$n$
We have a fixed size $O(1)$ tileset that:

- For any given $\varepsilon, \delta$
- $n > C(\varepsilon, \delta)$

We can assign percentages such that:

With probability at least $1-\delta$, a size $n' \times n'$ square is assembled with

$$(1-\varepsilon)n < n' < (1+\varepsilon)n$$
We have a fixed size $O(1)$ tileset that:

- For any given $\varepsilon, \delta$
- $n > C(\varepsilon, \delta)$

We can assign percentages such that:

With probability at least $1 - \delta$, a size $n' \times n'$ square is assembled with

$$(1 - \varepsilon)n < n' < (1 + \varepsilon)n$$

Question: **Exact** squares with high probability? (yes, see the paper in your reading)