CSCI 2570 Introduction to Nanocomputing

DNA Tiling

John E Savage

Computing with DNA

- Prepare oligonucleotides (“program them”)
- Prepare solution with multiple strings.
- Only complementary substrings q and $\underline{q}$ combine, e.g. q = CAG and $\underline{q}$ = GTC
- E.g. GCTCAG + GTCTAT =

```
GCTCAG
   | | |
   GTCTAT
```

- 1D & 2D crystalline structures **self-assemble**

Generating Random Paths Through the Graph

- Edge strings $q'_u p'_v$ combine with vertex strings $p_v q_v$ to form **duplexes**, shown below.

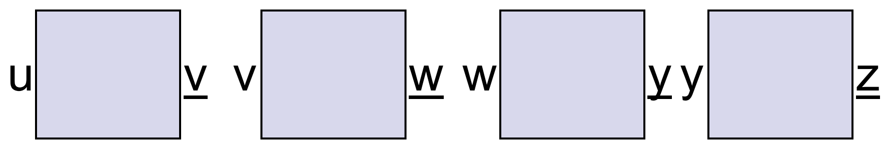
```
                 q'u p'v                                q'v p'w
GTATATCCGAGCTATTCGAG CTTAAAGCTAGGCTAGGTAC
           | | | | | | | | | | | | | | | | | | | | | | | |   | | | | | | | | |
           CGATAAGCTCGAATTTCGAT CCGATCCATGTTAGCACCGT
                           pv qv                             pw qw
```

- Colored pairs of coupled strings act as a unit.

- Each duplex has two sticky ends that can combine with another duplex or strand.

1D Tiling Model

- Modeled by non-rotating tiles with binding sites on E & W sides.



- All paths in a graph G can be produced with such tiles.
- Minimal bonding strength needed for adhesion

2D Tiling Model

- Square tiles with labels on each side.
  - Tiles do not rotate.
- A tile "sticks" only if the sum of the strengths of all bonds ≥ t, **threshold** of tiling system.
- Goal: build a pattern from a seed tile.
- **Note**: This is a random process!

Emulation of a Binary Counter

- Non-rotating tiles have binding sites on all 4 sides. Tile bounding strength: red = 2, other = 1
- Threshold = 2 (arrows where tiles can add).

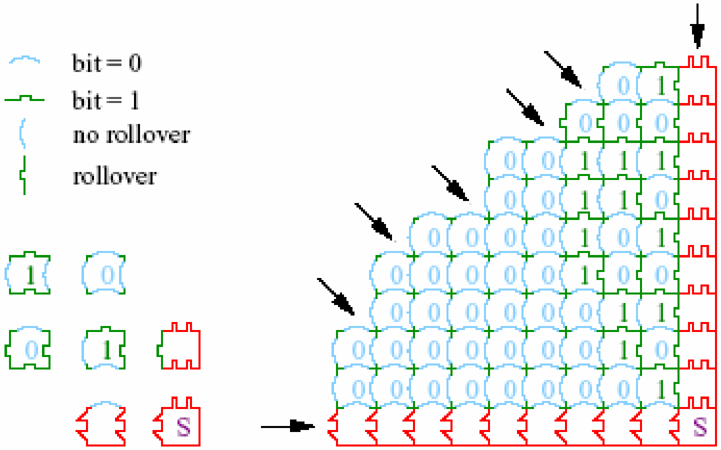
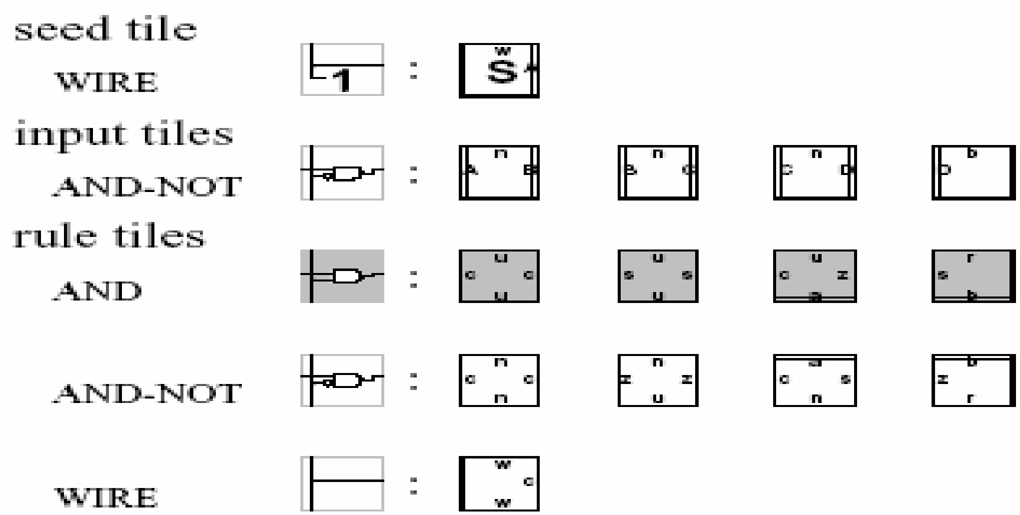


Fig: @NAE Bridge, vol 31, p34, Winfree

- Tiling starts at seed tile S.

Tiles Emulating a Decoder

Can a CPU be self-assembled?



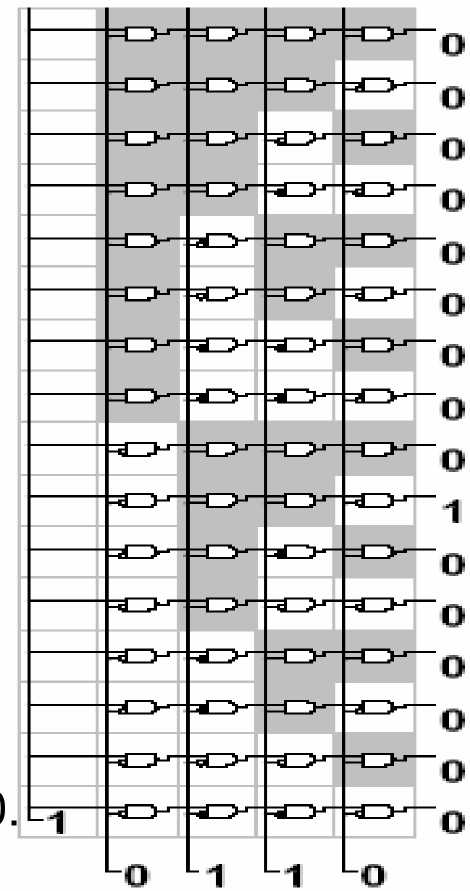


Fig: @ DNA9 2004 p91 Cook et al.

Double edges have strength 2. Thick edges have strength 0.
Others have strength 1. Threshold t = 2.

Addressable Memory Constructed from Tiling System

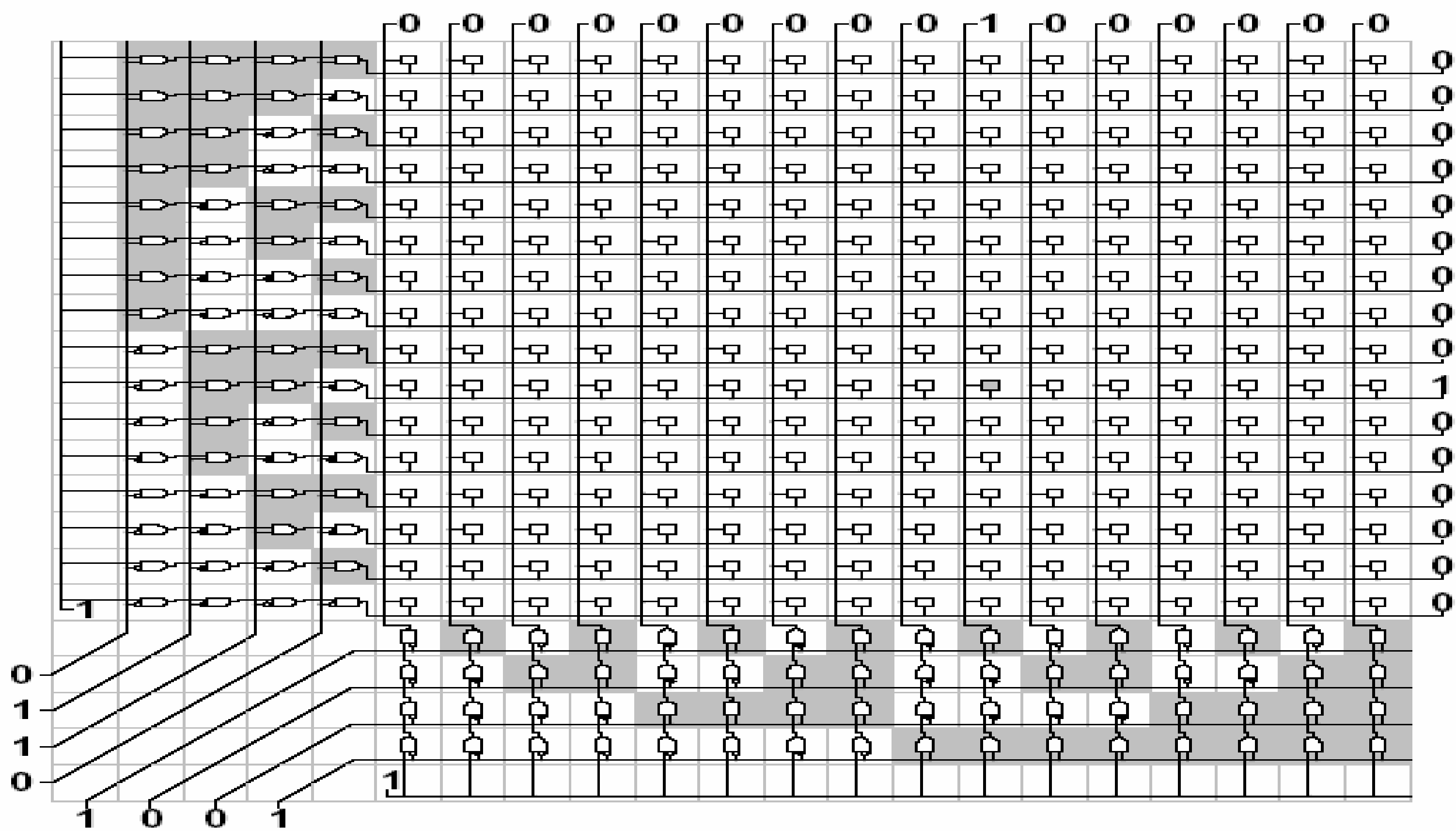


Fig: @ DNA9 2004 p91 Cook et al.

Languages and Tiling Systems

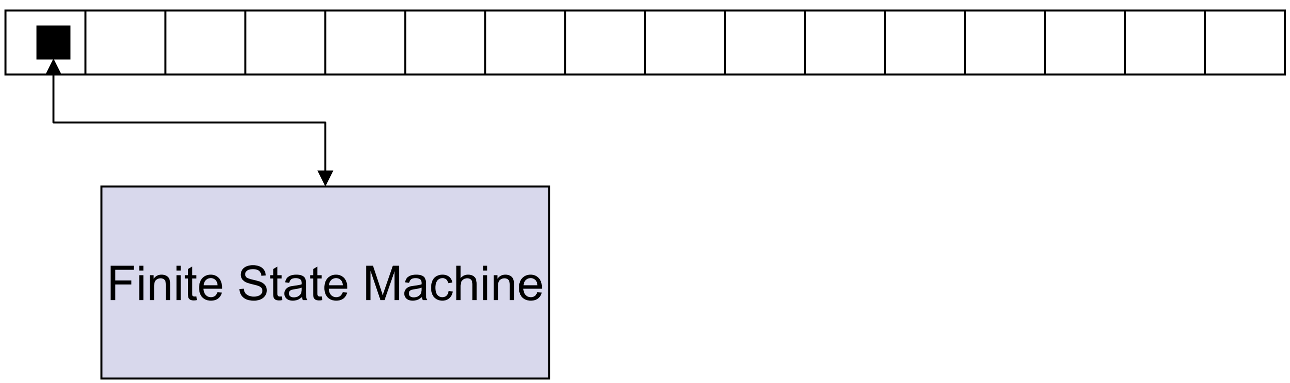
- Regular, context-free and recursively enumerable languages correspond to tiling systems with various restrictions
  - See "Universal Computation via Self-assembly of DNA: Some Theory and Experiments" by Winfree Yang and Seeman

Questions About Tile Systems

- Can a tile system fill the plane?
- What's the smallest tile system that generates a pattern?
- How hard is it to determine if a tile system uniquely assembles to a shape?

Universality of Tile Systems

- The Turing machine (TM) is "universal."



- We show that a tile system can simulate TM by computing **TM configurations**.

TM Configurations

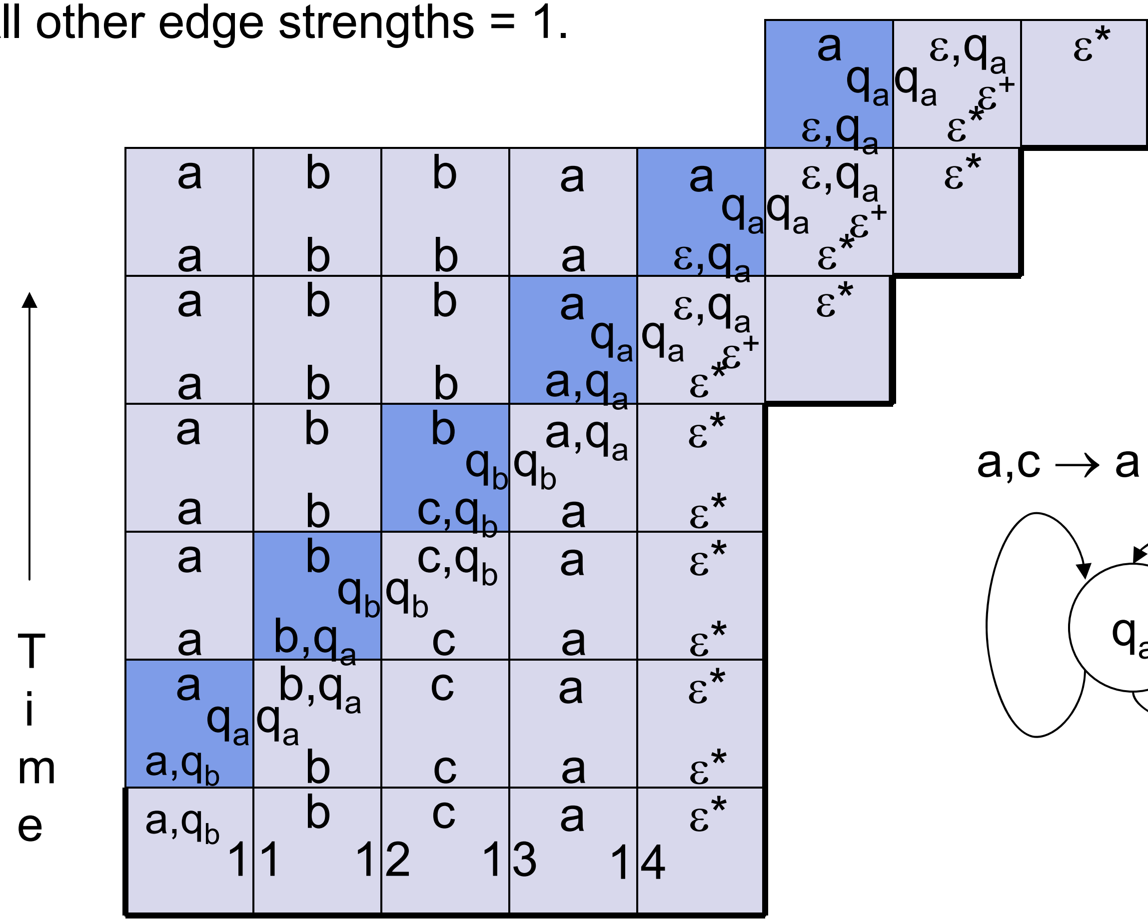
- Cell contains (q_i,x) if head over it or (-,x) if not.
- Get next config. from current & FSM state table
- Shows exist universal cellular automata.

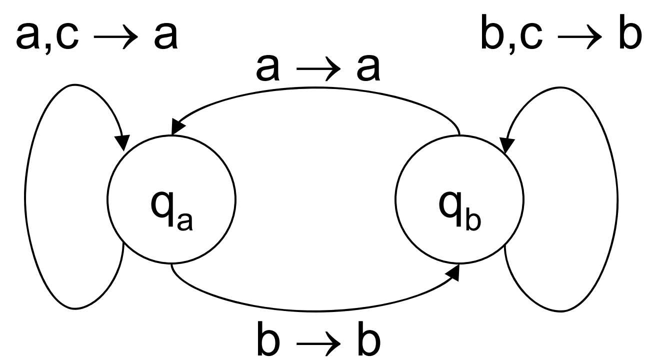
Time ↓

q_0 x_1	- x_2	- x_3	- x_4	- x_5	- β	- β	- β	- β	- β
- y_1	q_1 x_2	- x_3	- x_4	- x_5	- β	- β	- β	- β	- β
- y_1	- y_2	q_2 x_3	- x_4	- x_5	- β	- β	- β	- β	- β
- y_1	- y_2	q_4 y_3	- x_4	- x_5	- β	- β	- β	- β	- β

Tiling Emulation of a TM

Colored tile binds to edge with strength = 2.
All other edge strengths = 1.





Tiling Emulation of TM

- Example illustrates the writing of a new symbol and moving the head.

- Must also handle writing over a blank cell and creating a new one on the right (or left), if necessary.
  - What tiles would handle this case?

Answers to Questions

- Can a tile system fill the plane?
  - Yes, if TM doesn't halt.
  - How hard is it to determine if this is possible?
- What is smallest tile system that generates a pattern?
  - Can the "busy beaver problem" be applied?
    - On empty tape, what's longest string written by halting TM?
  - Related to the Kolmogorov complexity of the pattern?
    - Shortest input string generating given string on universal TM.
- How hard is it to determine if a tile system uniquely assembles to a shape?
  - NP-complete

Self Assembly

- DNA tile systems illustrate self assembly
- Errors occur in practice.
  - Tiles adhere where they shouldn't and get locked into place by subsequent attachments
  - They can also nucleate without using a seed.
- Methods to control errors:
  - Proofreading tile sets
  - Zig-zag tile set and control of concentrations

Sierpinski Triangle

- Double-edge strength = 2, others = 1, t = 2

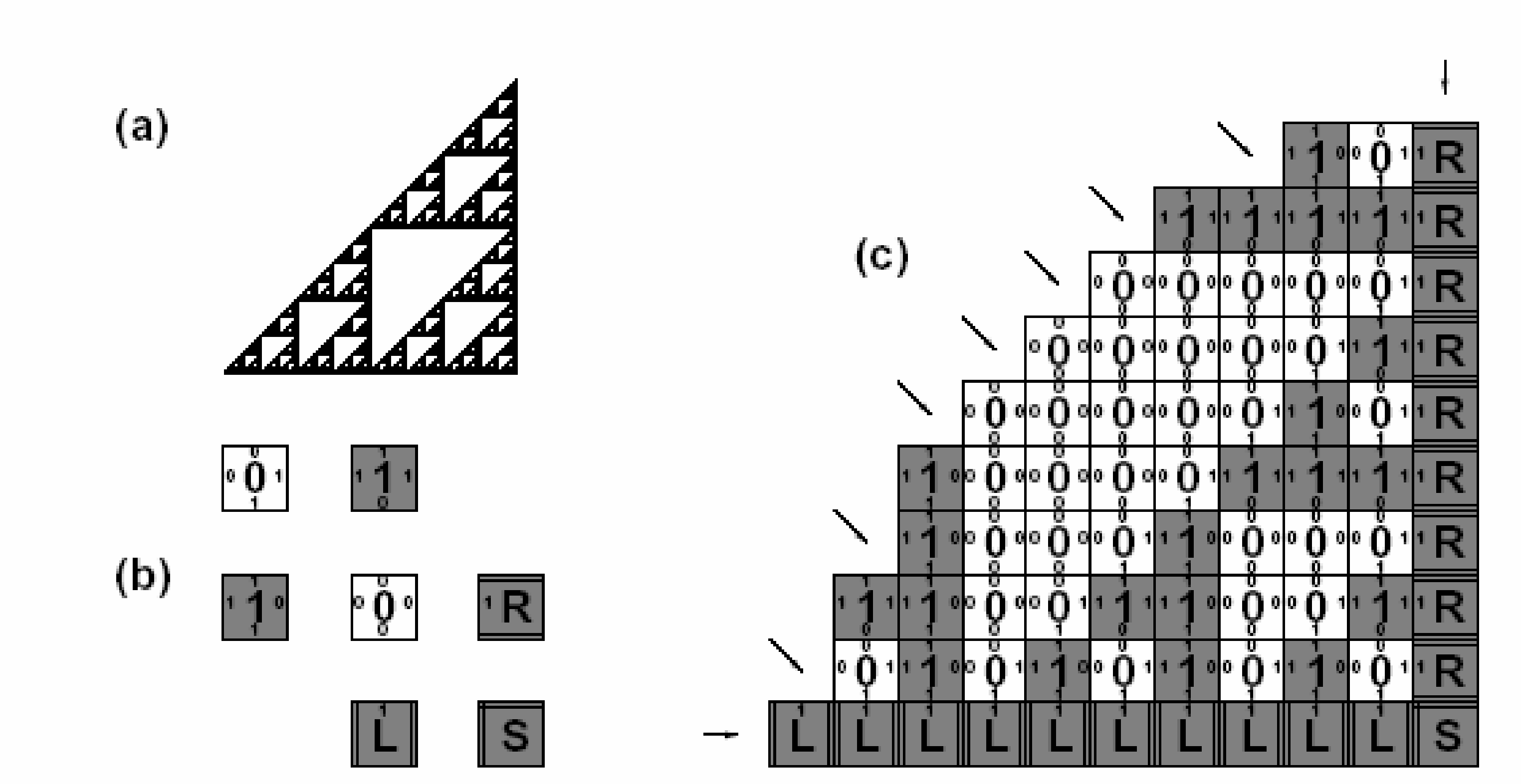


Fig: @ DNA9, vol 2943, p.91, Cook et al.

Error in Self Assembly of Sierpinski Triangle

- A single error will propagate
  - Error rates in a DNA tiling experiment were 1-10%.
  - Spurious nucleation dominated outcomes.

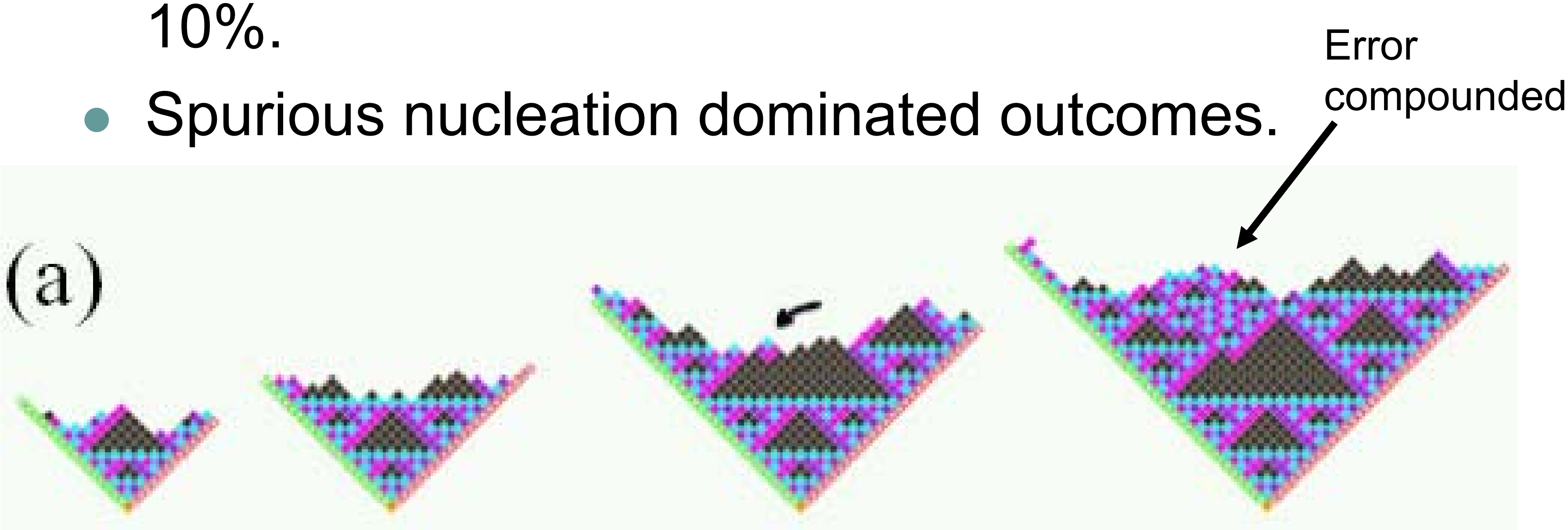


Fig: @ Procs. DNA9, 2003, p126

How to Control Errors in DNA Self-Assembly?

- Error correction?
  - Fault tolerant cellular automata are known.
  - But challenging.
- Optimizing conditions for assembly?
  - A 10-fold reduction in mismatch rates in standard DNA tiling requires 100-fold increase in assembly time by cooling down the process.
- Redesigning the tile set to reduce error rate?

Self Assembly/Disassembly

- Rate of assembly is determined by the concentration of free tiles.
- Rate of disassembly is a function of binding energies and temperature of the environment
- Winfree has modeled this process.

Proofreading Tile Sets† Reduces Spurious Nucleation

- Each original tile replaced by 4 tiles

(x,y) ✹ (z,z), z = x ⇖ y

- When a mismatch occurs, there is no way to continue without making an additional error.

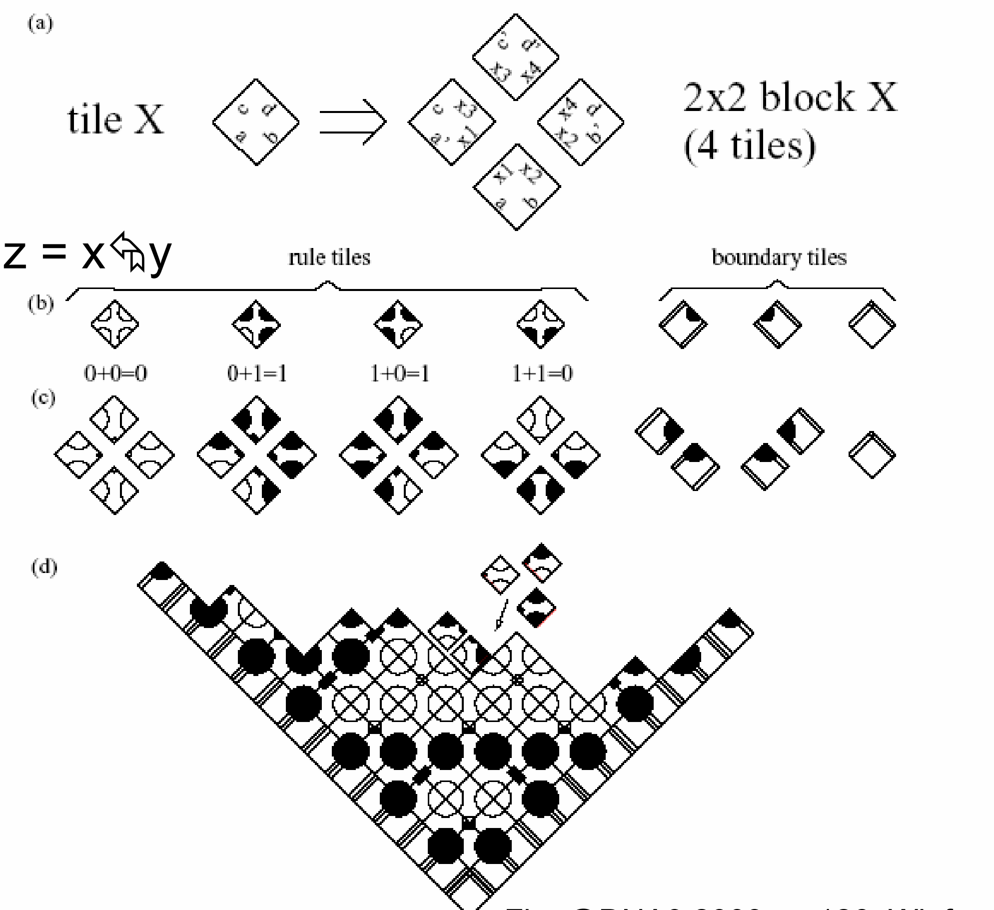


Fig: @DNA9 2003, p. 126, Winfree

† Winfree, Procs. DNA9, 2003

Simulation with 2x2 Proofreading Tiles

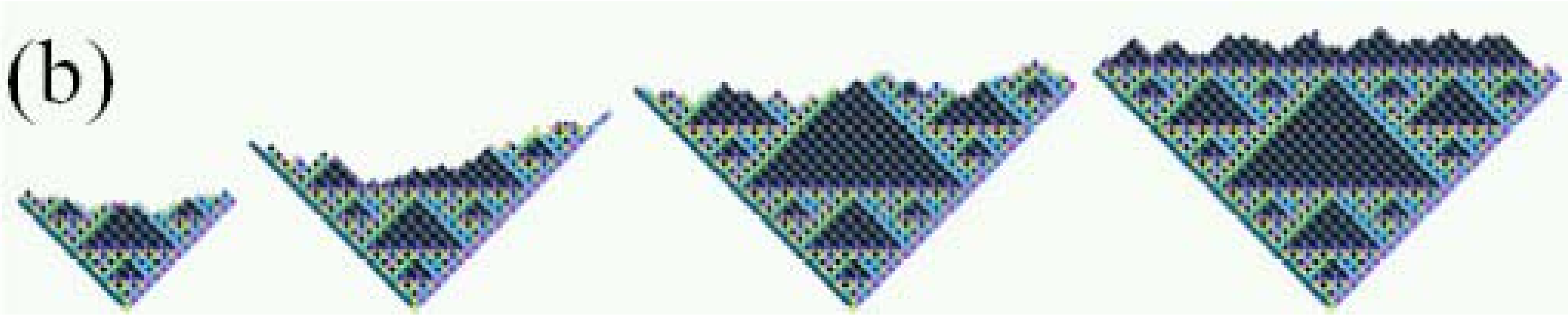
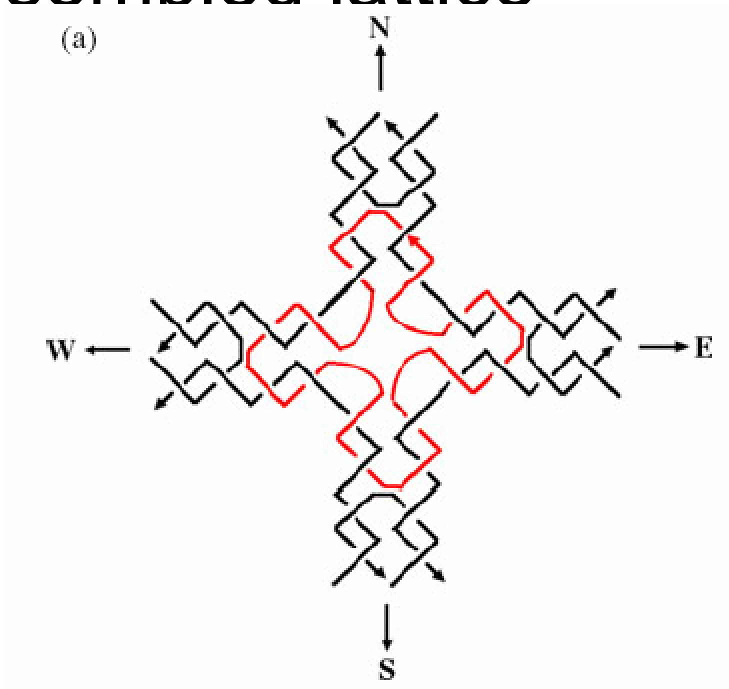
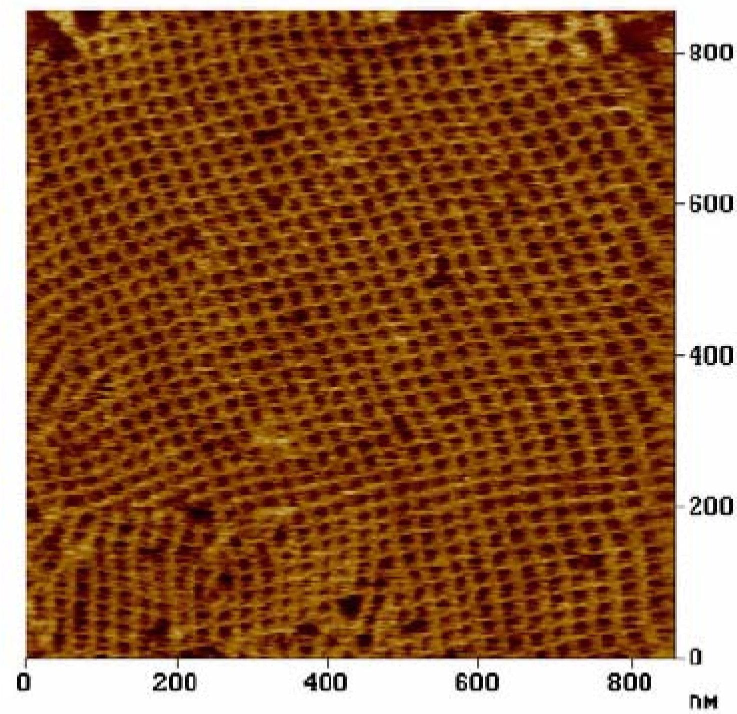


Fig: @ Procs. DNA9, 2003, p126

DNA Scaffolds

- DNA tile (a Holliday junction) and self-assembled lattice





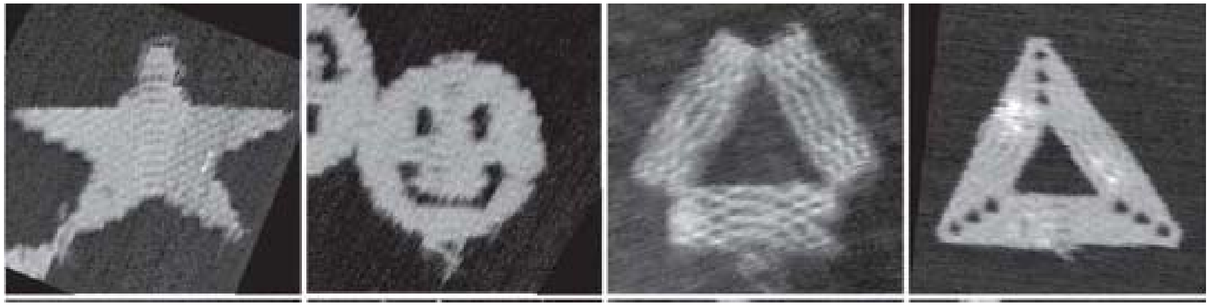
Figs: @Nanotechnology, v 15, (2004) p S525

Prospects for DNA-Based Algorithmic Self Assembly

- Combinatorial problems: at best 10^{12} ops/sec
  - Can be done faster on conventional computers.
  - Not very promising.

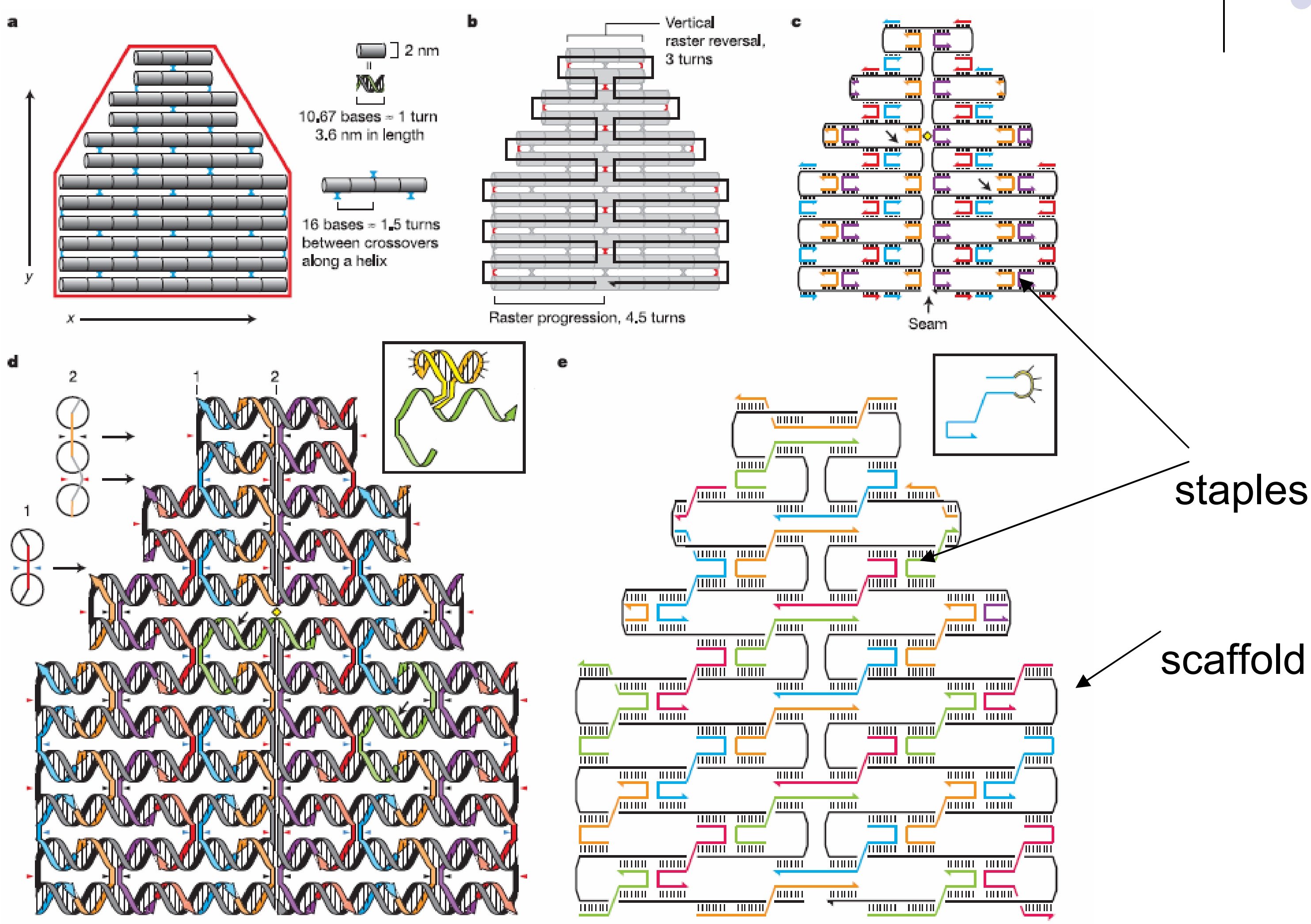
Patterning & Templating DNA

- Rothemund[+] has presented a remarkably effective method for generating shapes from DNA which he can decorate with molecules to produce patterns. (See his website.)



[+]Folding DNA to Create Nanoscale Shapes and Patterns, Nature, March 2006.

Rothemund's Approach



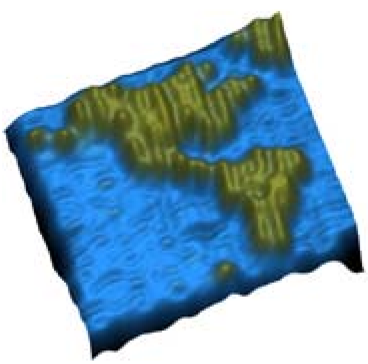
Rothemund's Commentary[+] on Self-Assembly of DNA Strands

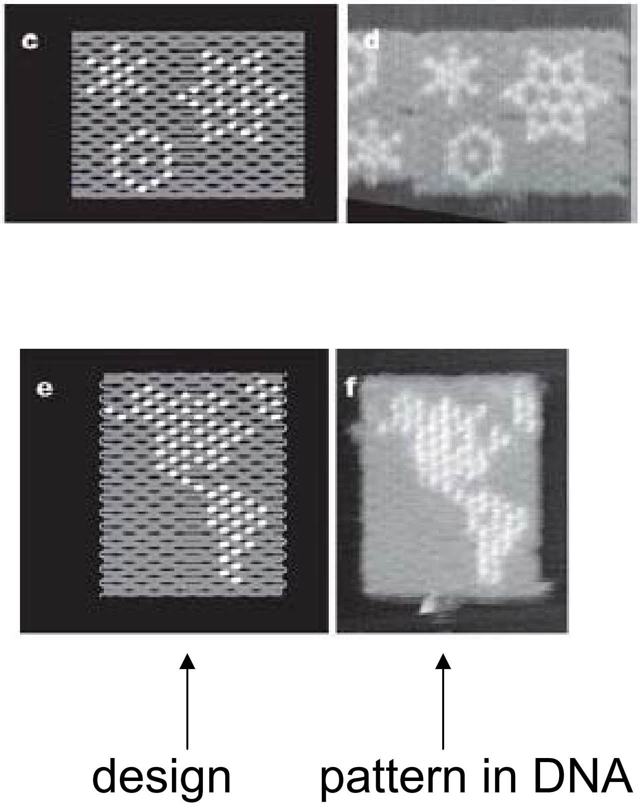
- The widespread use of scaffolded self-assembly … of long DNA scaffolds in combination with hundreds of short strands, has been inhibited by several (assumptions):
  - Sequences must be optimized to avoid secondary structure or undesired binding interactions,
  - Strands must be highly purified, and
  - Strand concentrations must be precisely equimolar …
- All three are ignored in the present method.

[+]Folding DNA to Create Nanoscale Shapes and Patterns, Nature, March 2006.

Rothemund's Patterns

- Staples were decorated with molecules visible under an atomic force microcroscope.





Conclusion

- DNA-based computing offers interesting possibilities
- Most likely to be useful for nano fabrication
  - However, high error rates may preclude its use