Media Technology

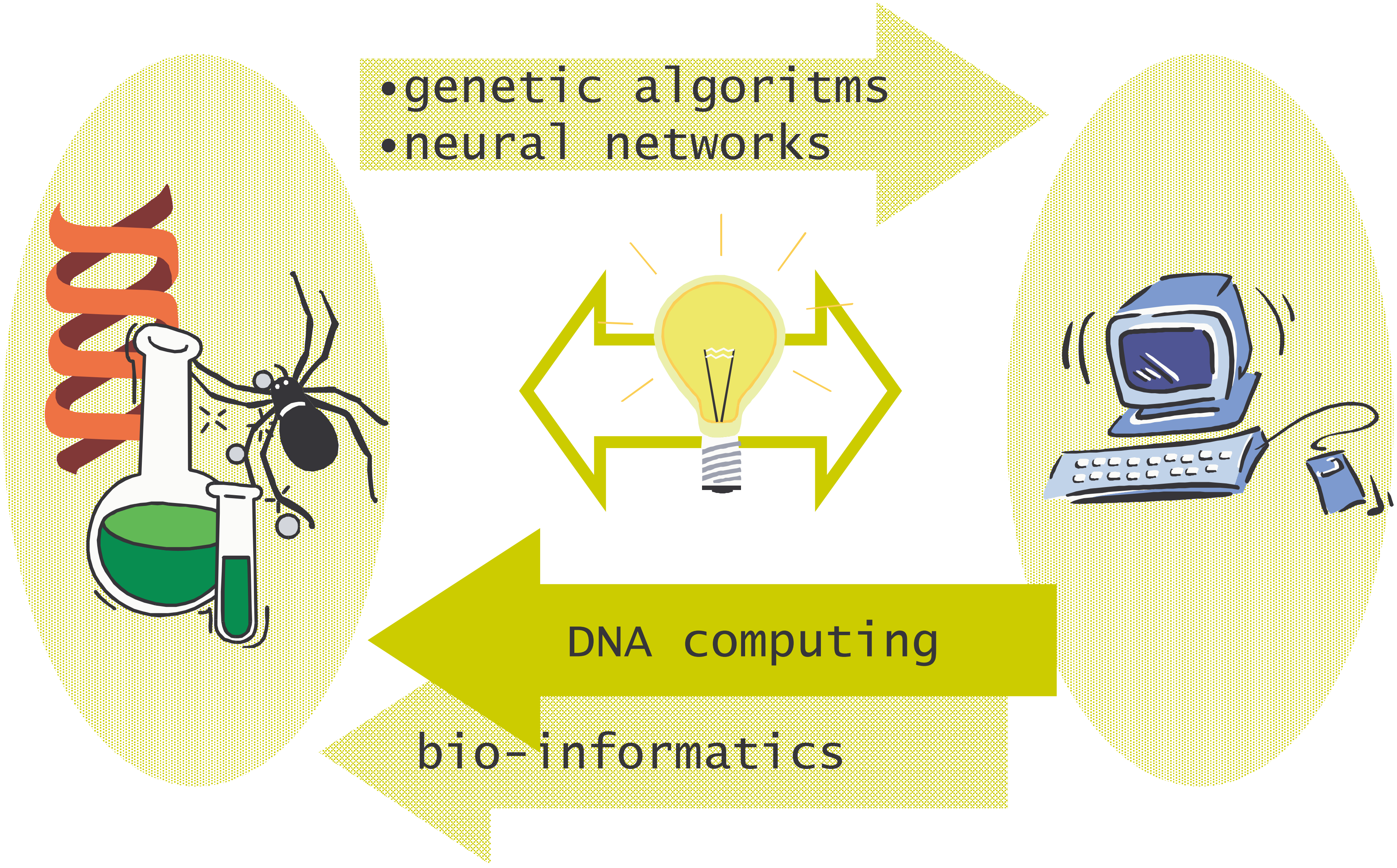
Artificial Intelligence
for Cocktail Parties

Computer in a TestTube

DNA computing

Hendrik Jan Hoogeboom
Computer Science (LIACS)
14 April 2008

natural computation



Len Adleman

Molecular Computation of Solutions to Combinatorial Problem, Science, 266: 1021-1024, (Nov. 11) 1994.

http://www.usc.edu/dept/molecular-science/fm-adleman.htm

Physicists plunder life's tool chest

If we look inside the cell, we see extraordinary machines that we couldn't make ourselves, says Len Adleman. "It's a great tool chest - and we want to see what can we build with it."

Adleman created the first computer to use DNA to solve a problem. He was struck by the parallels between DNA, with its long ribbon of information, and the theoretical computer known as the Turing Machine.

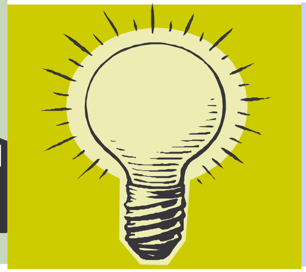
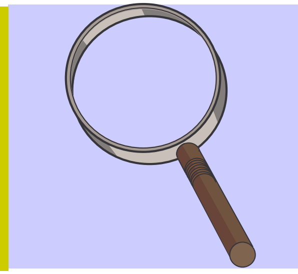
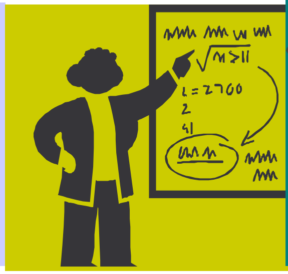
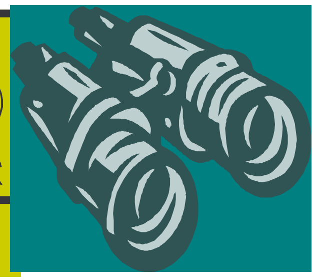
Nature News Service april 2003

Adleman tackled the famous ‘travelling salesman’ problem - finding the shortest route between cities. Such problems rapidly become mind-boggling. The only way is to examine every possible option. With many cities, this number is astronomical.

DNA excels at getting an astronomical amount of data into a tiny space. “One gram of DNA can store as much information as a trillion compact discs,” says Adleman. Myriad DNA molecules can examine every possible route at once, rather than one at a time, as in a conventional computer.

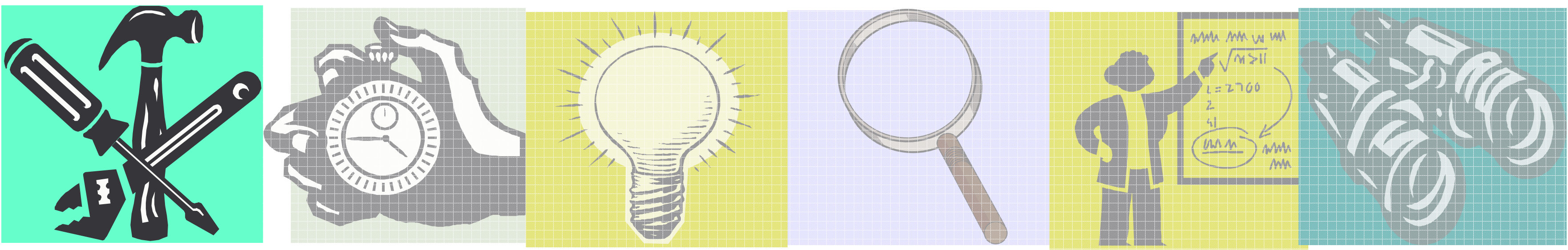
contents

- DNA … the tool chest
- problem complexity … P & NP
  Hamilton Path Problem
- Adleman’s algorithm
- comments
- theory … Turing machine
- recent work + future

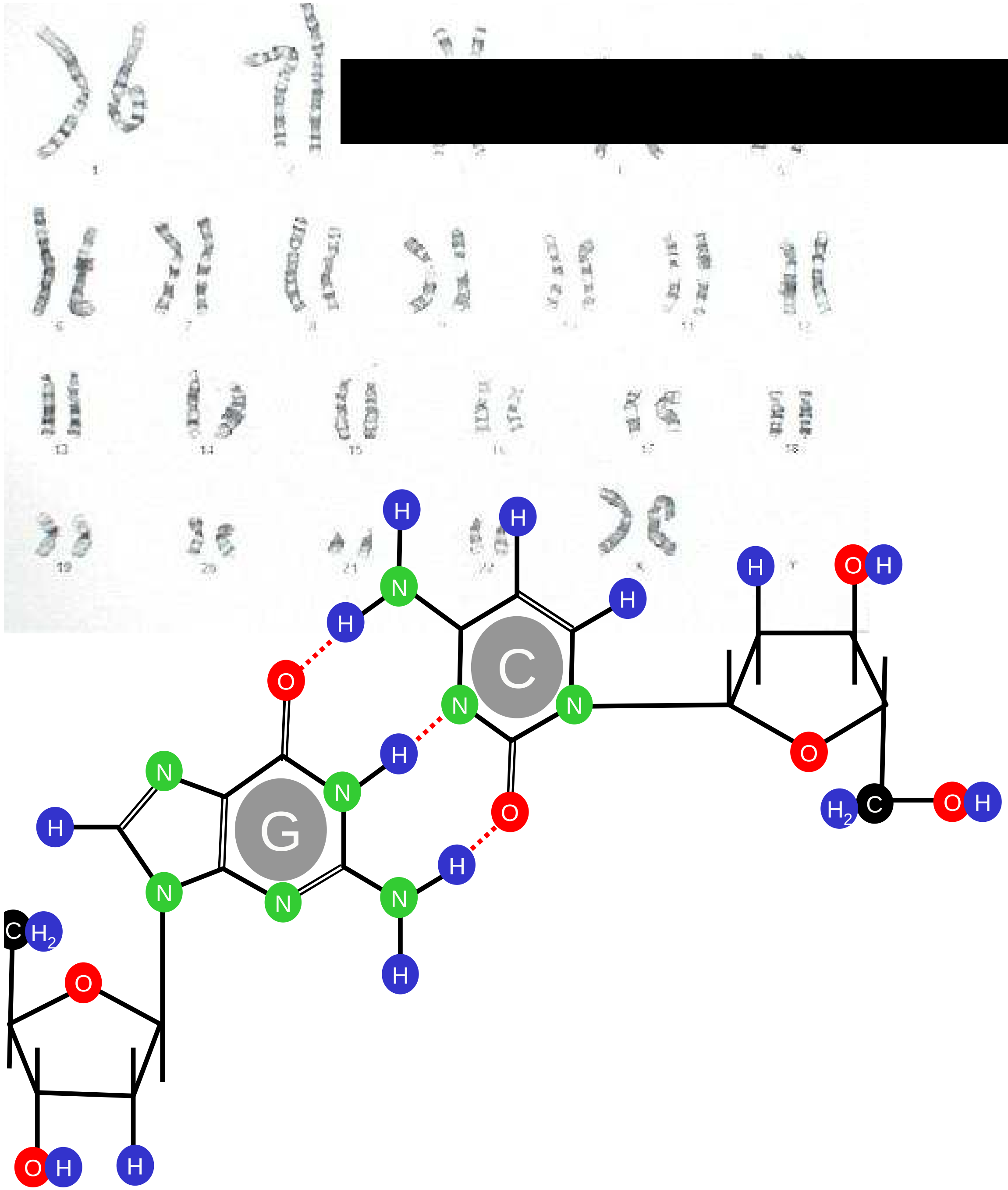
     

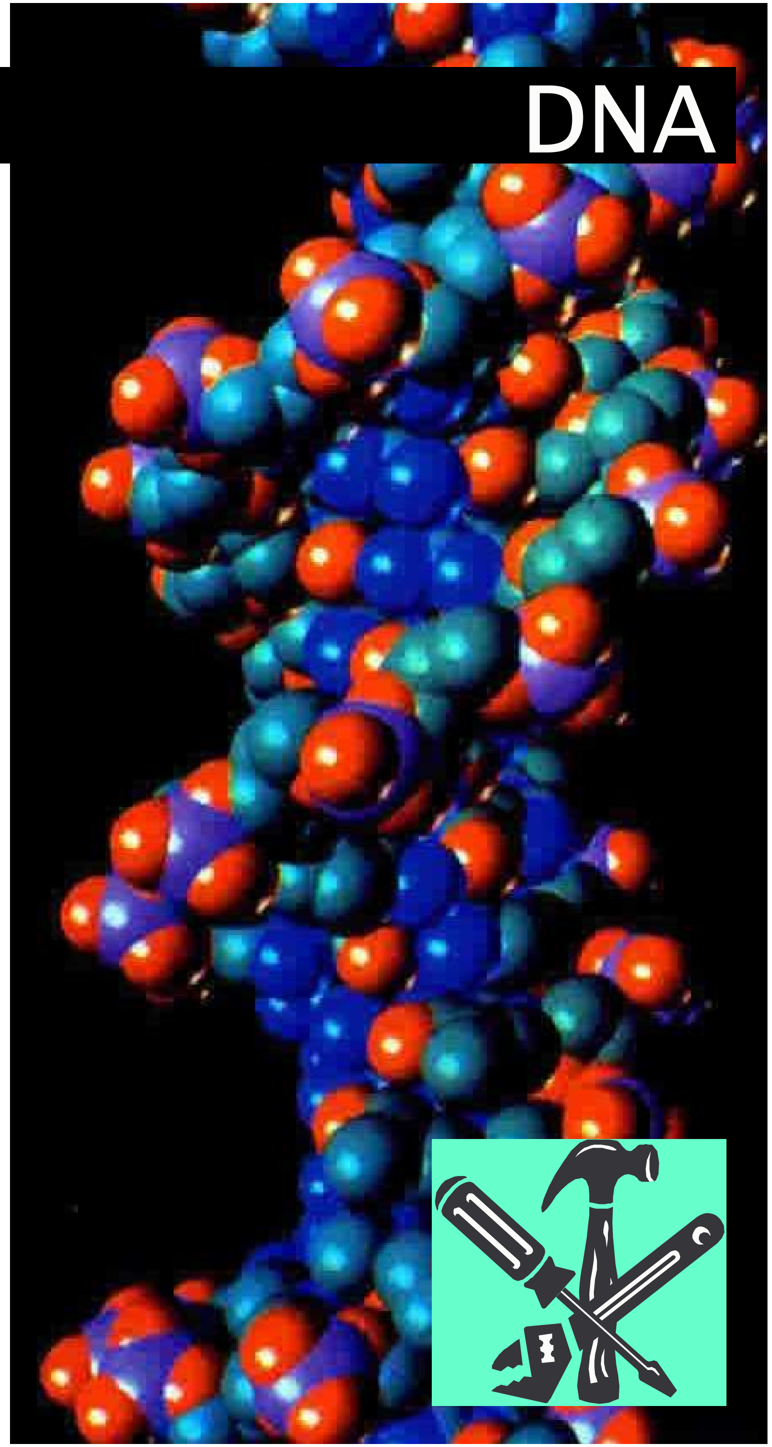
contents

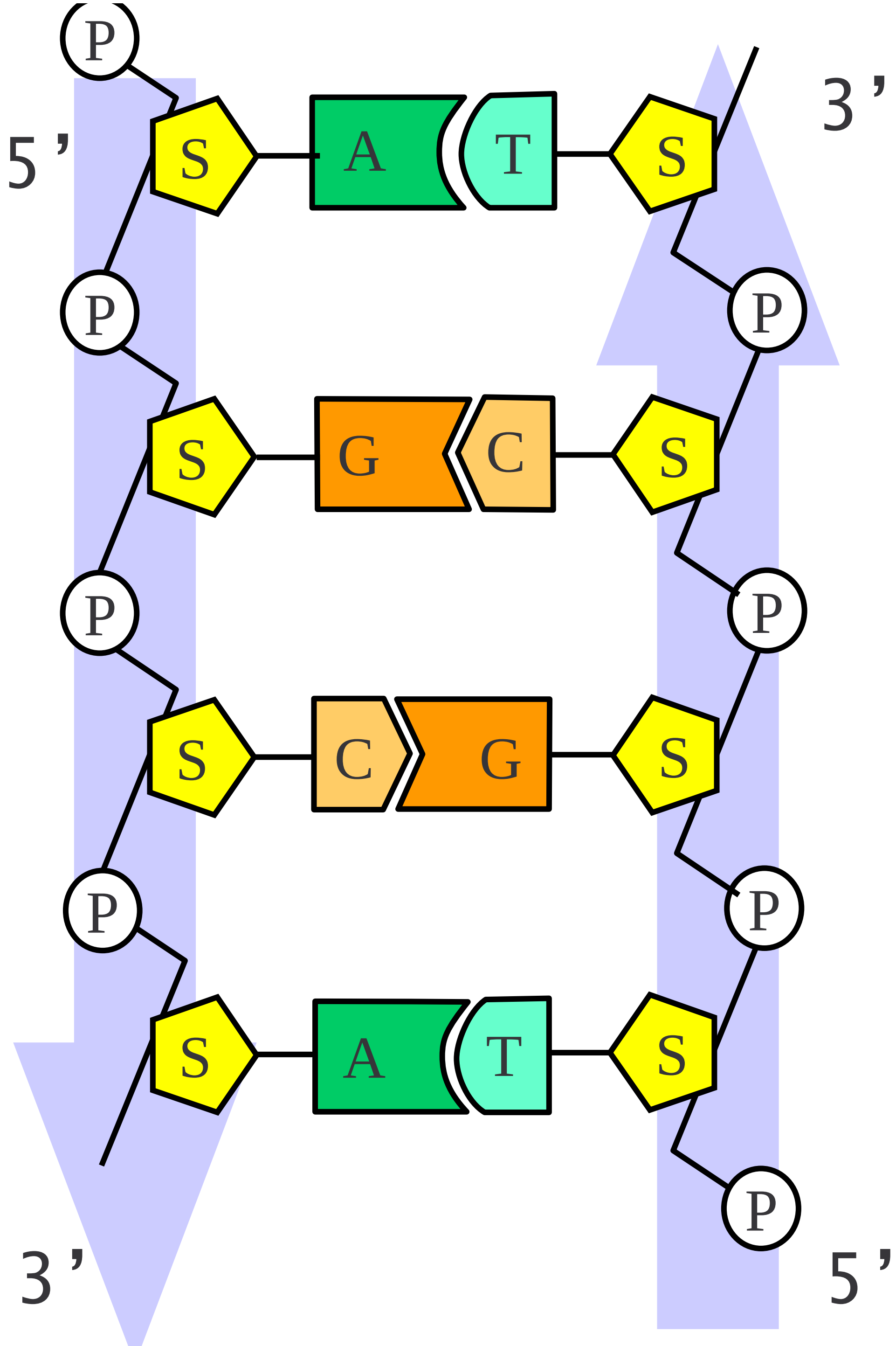
- DNA … the tool chest
- problem complexity … P & NP
  - Hamilton Path Problem
- Adleman's algorithm
- comments
- theory … Turing machine
- recent work + future



DNA







DNA

basepairs

Watson & Crick

[& Rosalind Franklin]

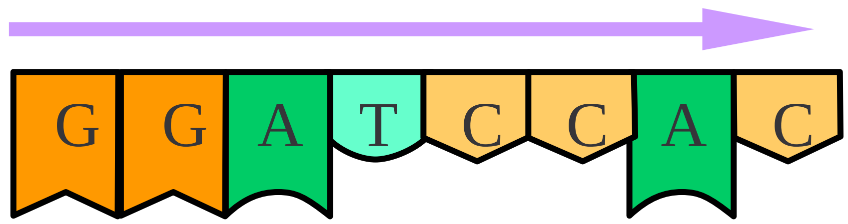
A=T

adenine - thymine

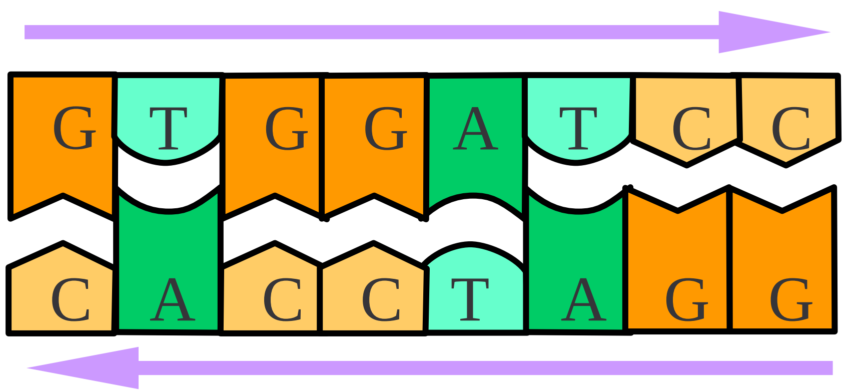
C≡G

guanine - cytosine

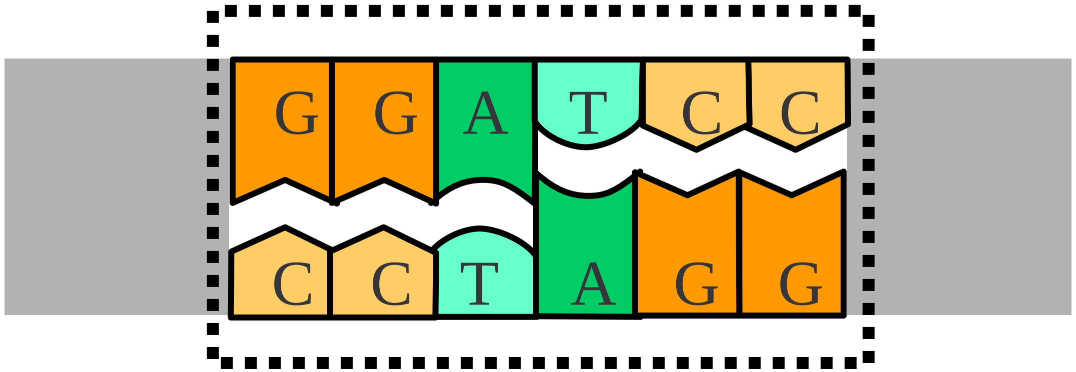
annealing & denaturing



complementary



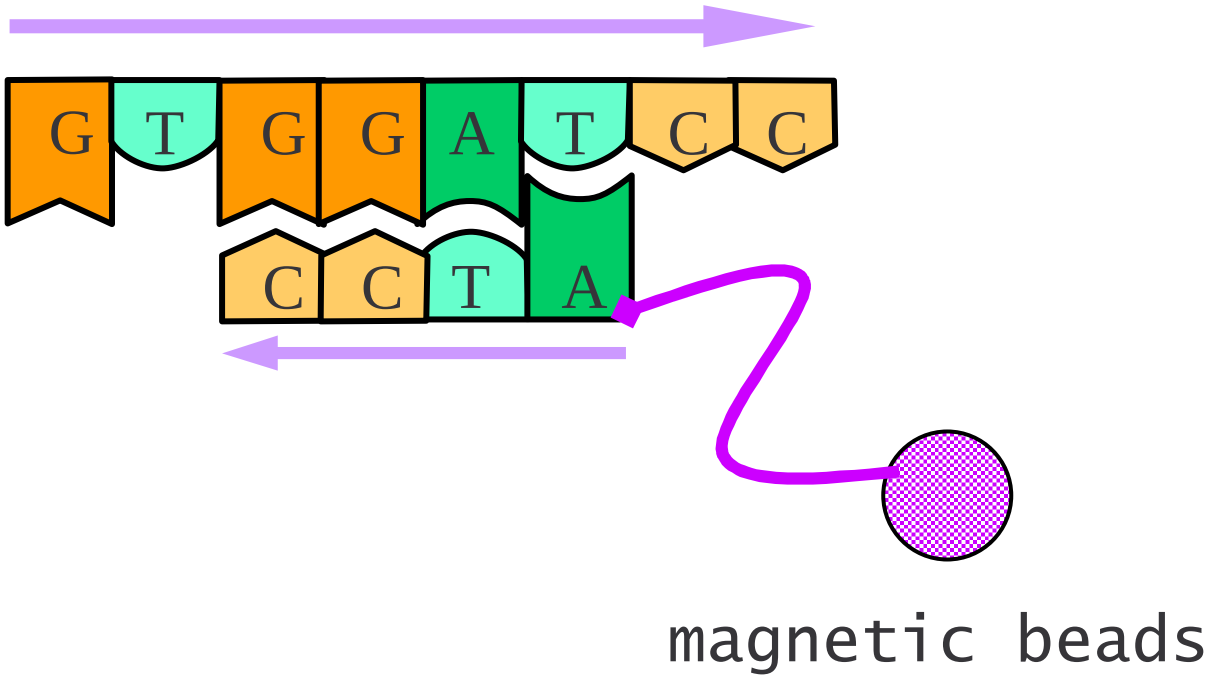
restriction enzymes



BamHI

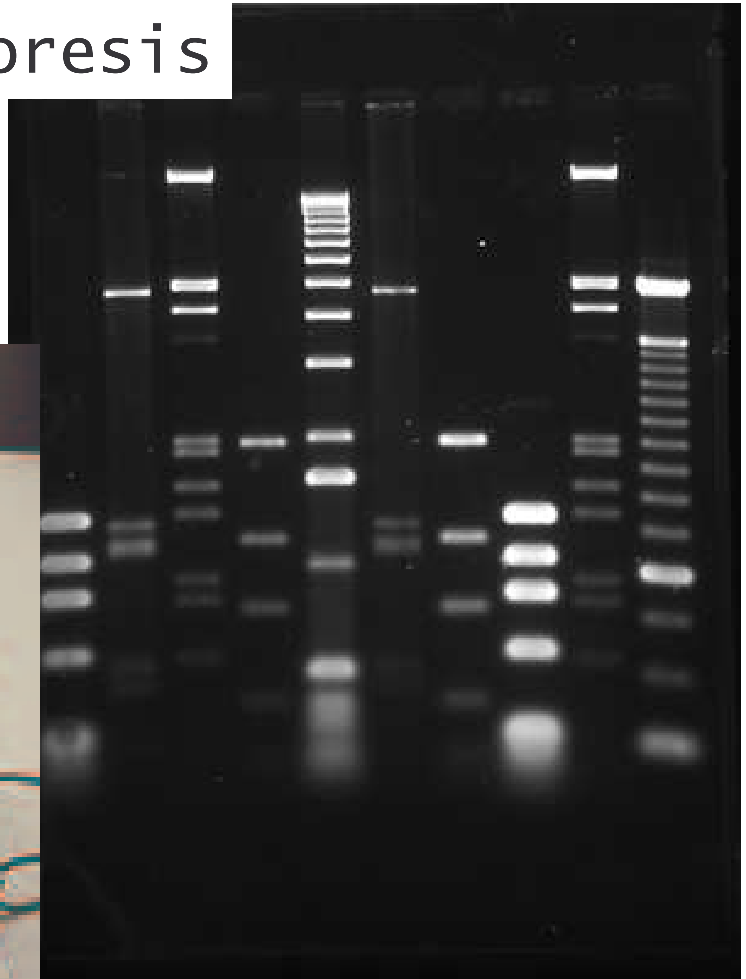
sticky ends

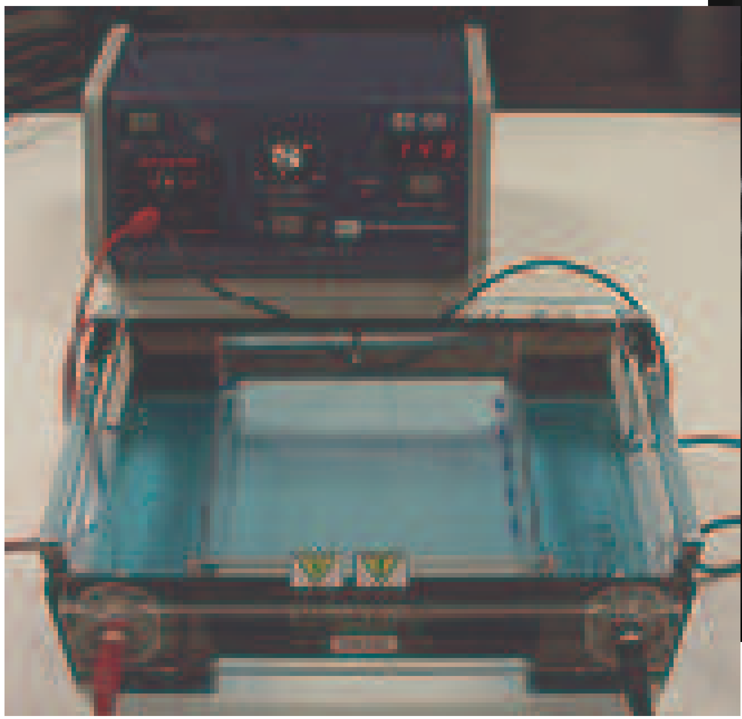
selection



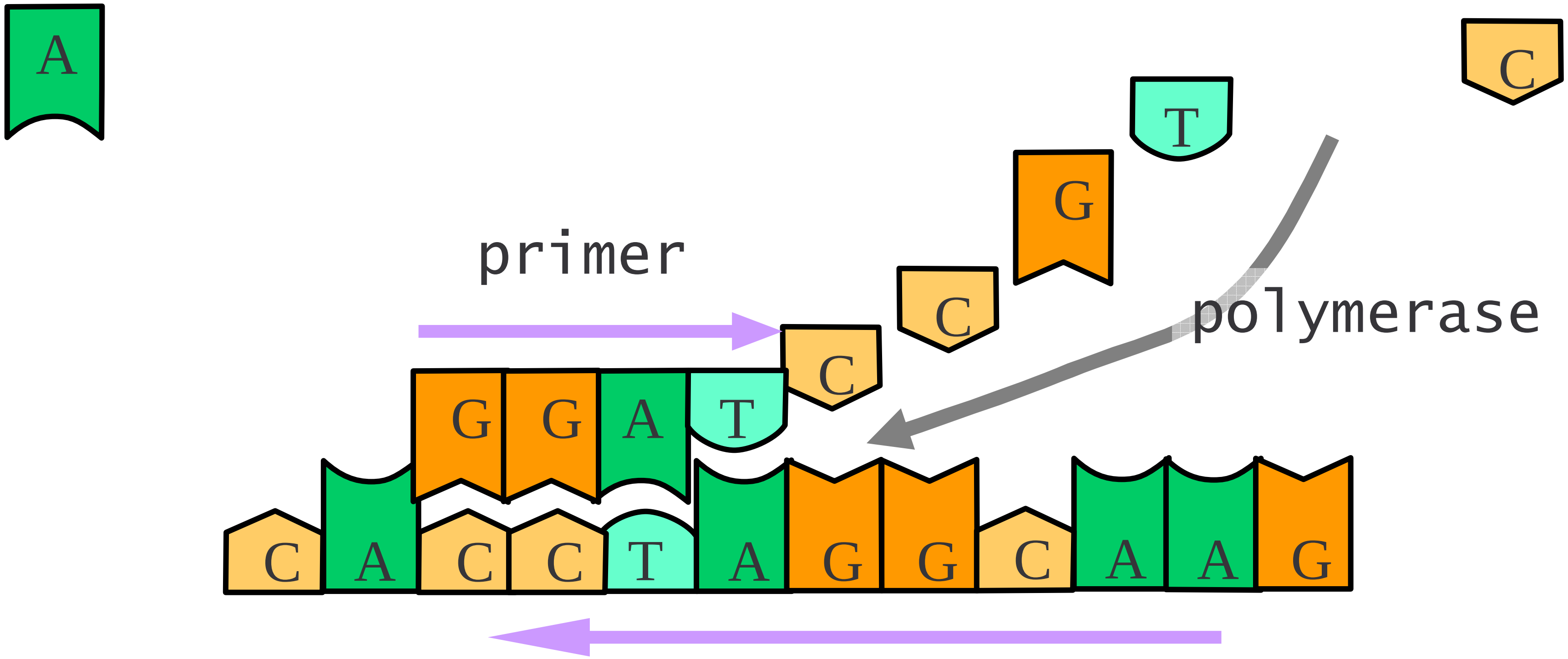
separation on length

DNA gel electrophoresis



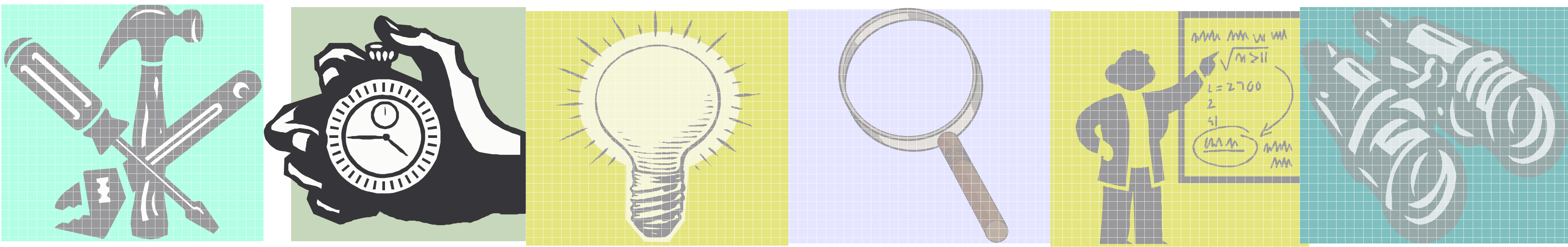


multiplication / amplification



contents

- DNA … the tool chest
- problem complexity … P & NP
  Hamilton Path Problem
- Adleman's algorithm
- comments
- theory … Turing machine
- recent work + future



complexity

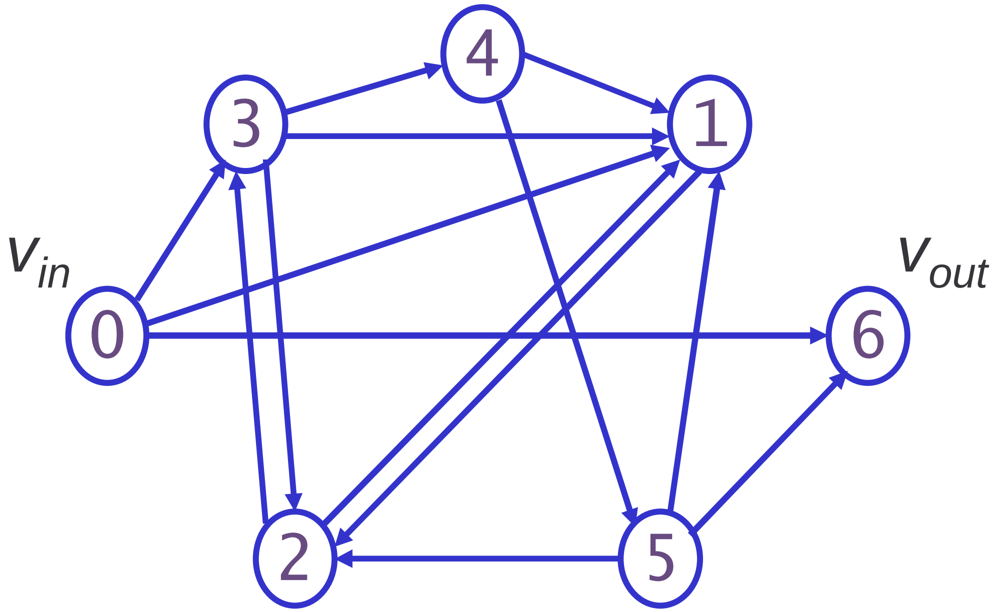
	n=10	30	50	60
n	10^{-5}s	3×10^{-5}s	5×10^{-5}s	6×10^{-5}s
n^2	10^{-4}s	9×10^{-4}s	2×10^{-3}s	4×10^{-3}s
n^5	10^{-1}s	24 s	1.7 m	13 m
2^n	10^{-3}s	18 m	13 d	366 c
3^n	6×10^{-2}s	6.5 y	3855 c	10^{13} c

second
minute
day
year
century

polynomial vs.
exponential

	now	100x	1000x
n	N	100N	1000N
n^2	N	10N	32N
n^5	N	2.5N	4N
2^n	N	N+6.6	N+10
3^n	N	N+4.2	N+6.3

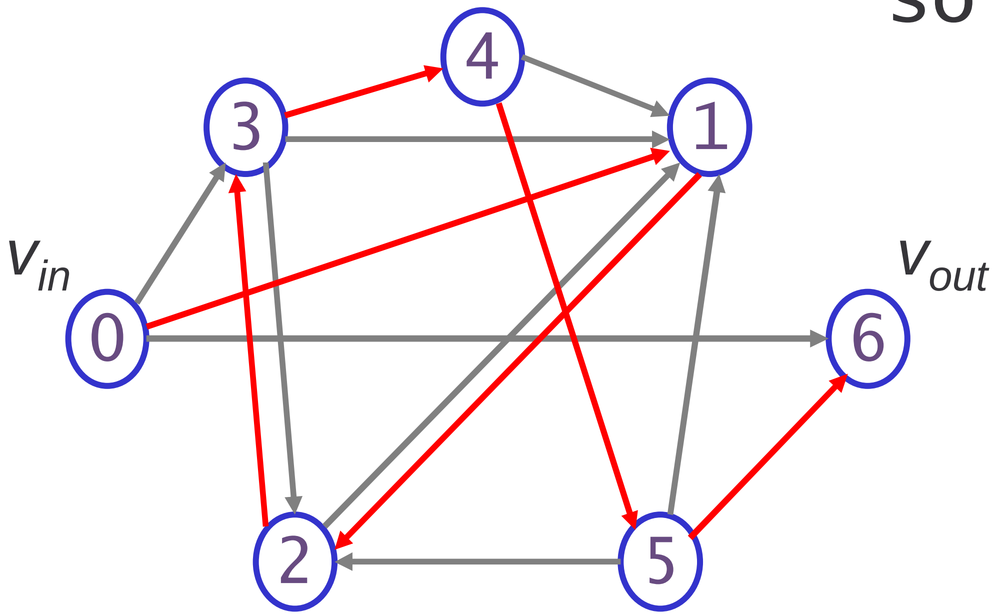
HPP: Hamilton Path Problem



'travelling salesman'

given: points and connections
question: is there a path that visits each point exactly once ?

HPP: Hamilton Path Problem

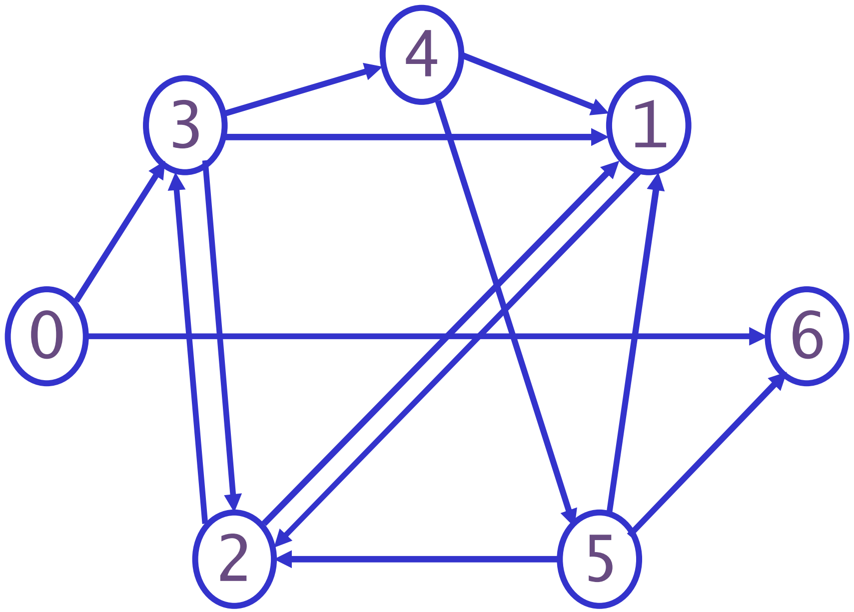


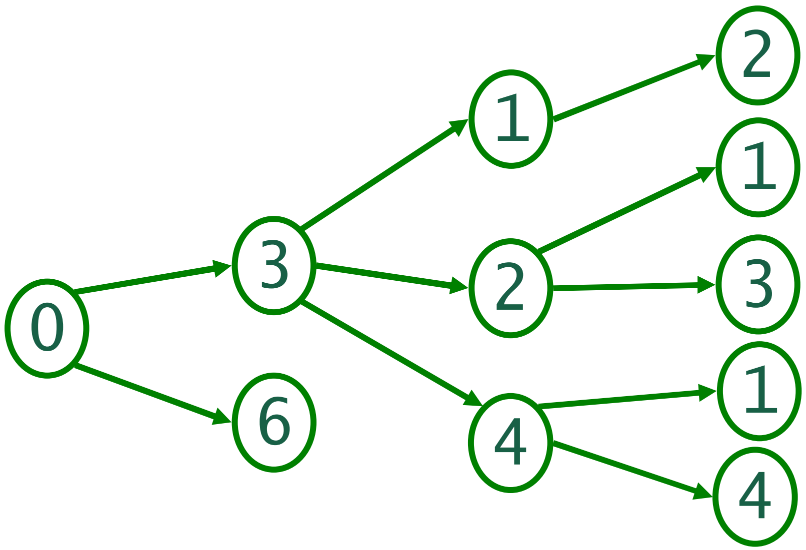
solution

‘travelling salesman’

given: points and connections
question: is there a path that visits each point exactly once ?

HPP: Hamilton Path Problem





no solution?

exponential time:
try all
possibilities

representative
'NP complete'

heuristics

complexity (theory) - P vs. NP

P

polynomial algorithm
to find a solution

NP

polynomial algorithm
to verify a solution

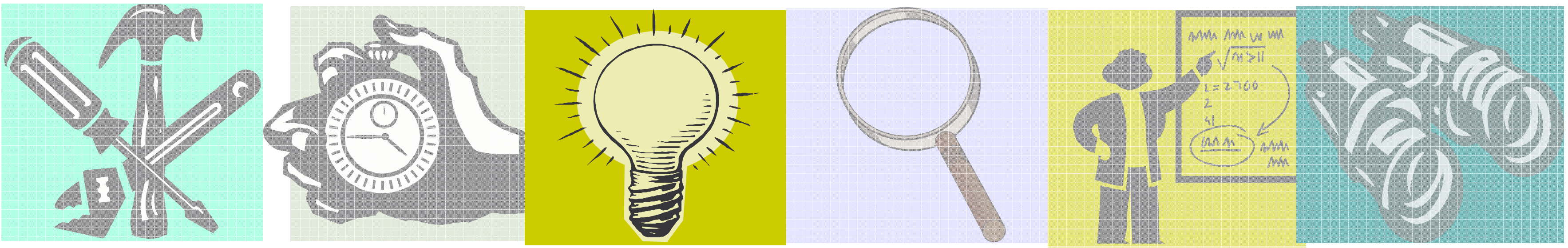
NP-complete

millenium prize problem $P \overset{?}{=} NP$

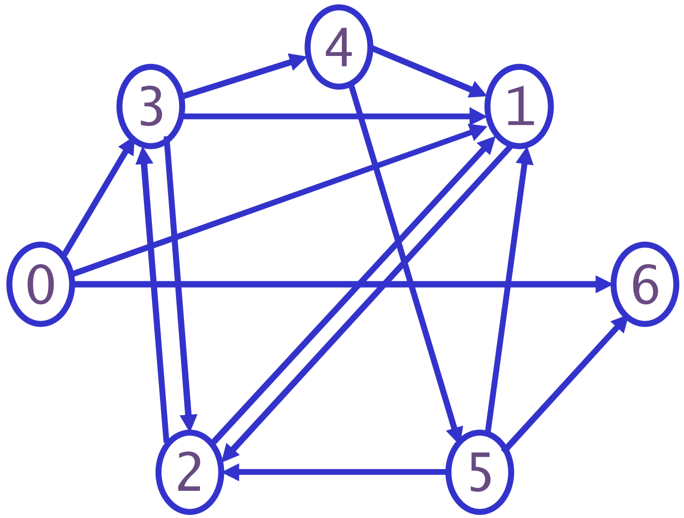
www.claymath.org/Millennium_Prize_Problems/

contents

- DNA … the tool chest
- problem complexity … P & NP
  Hamilton Path Problem
- Adleman's algorithm
- comments
- theory … Turing machine
- recent work + future



Adleman's algorithm

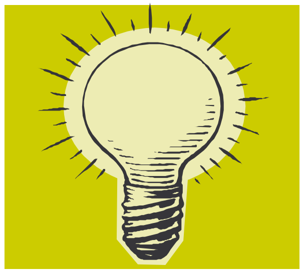


1. generate 'all' paths

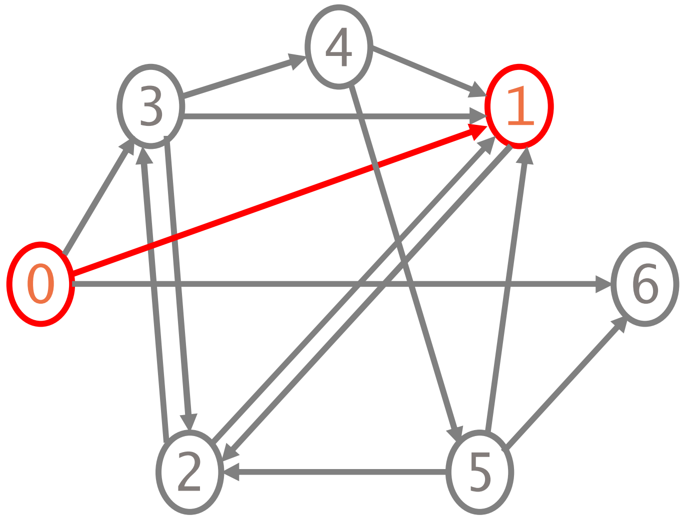
   keep only paths
2. … from v_{in} to v_{out}
3. … that enter n vertices
4. … that enter all vertices

5. if any path remains OK

'massive parallellism'

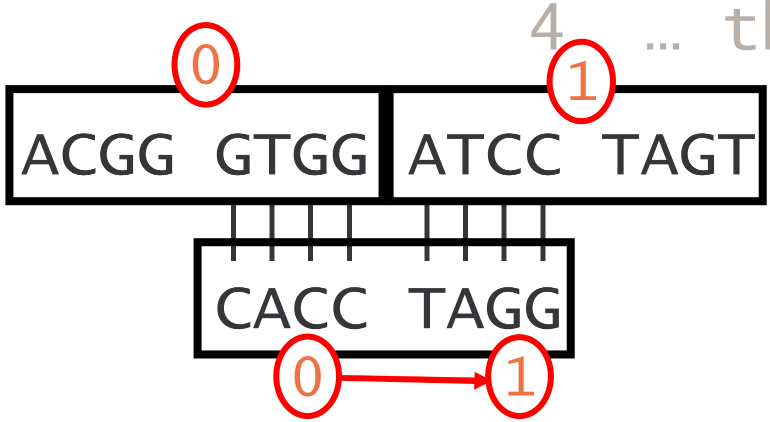


Adleman's algorithm

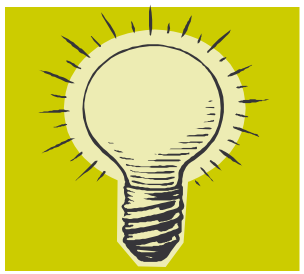


0. coding the graph
1. generate 'all' paths

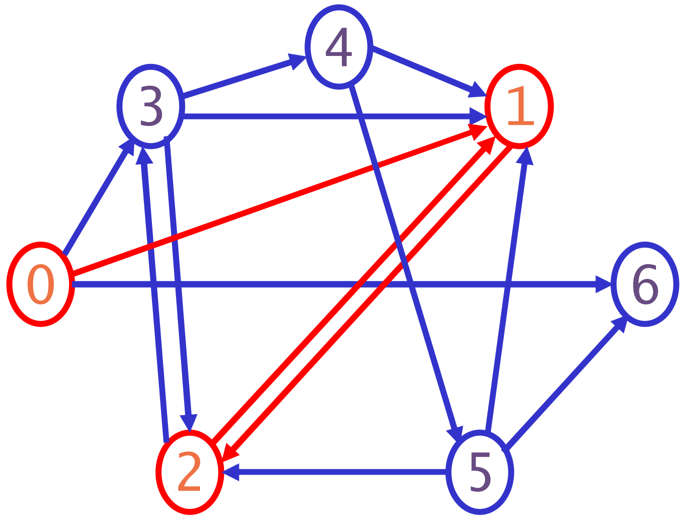
   keep only paths
2. … from v_{in} to v_{out}
3. … that enter n vertices
4. … that enter all vertices



any path remains OK



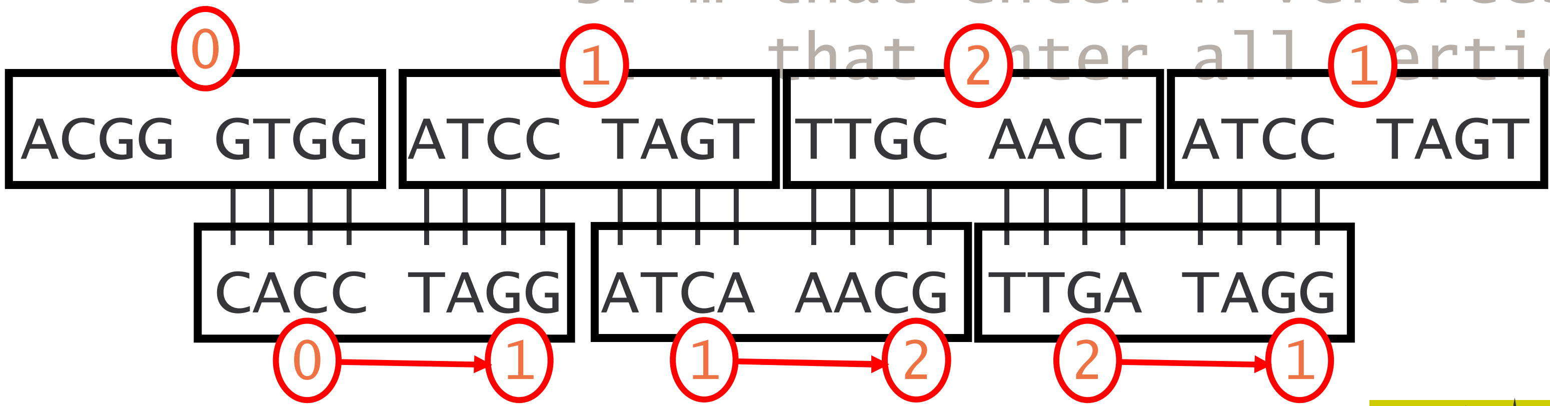
Adleman's algorithm

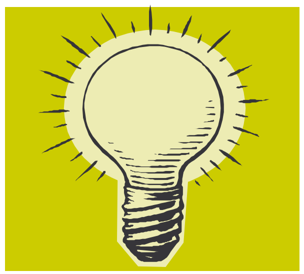


0. coding the graph
1. **generate ‘all’ paths**

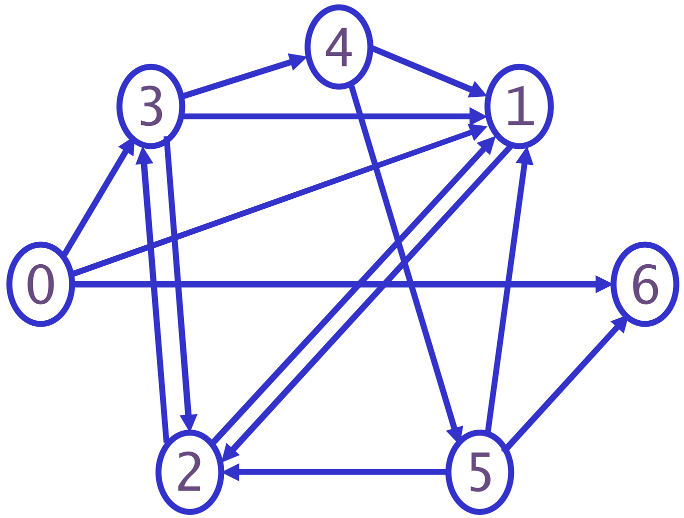
keep only paths

2. … from v_{in} to v_{out}
3. … that enter n vertices
4. … that enter all vertices





Adleman's algorithm



0. coding the graph
1. generate 'all' paths

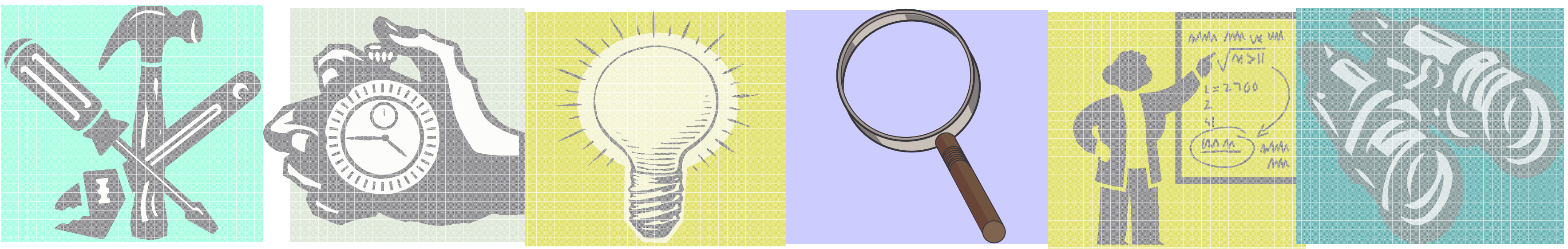
   keep only paths
2. … from v_{in} to v_{out}
3. … that enter n vertices
4. … that enter all vertices

5. if any path remains OK

- PCR with v_{in} and v_{out} primers
- gel: separate on length, amplify & purify
- magnetic beads: select strands
- PCR amplification & gel

contents

- DNA … the tool chest
- problem complexity … P & NP
    - Hamilton Path Problem
- Adleman's algorithm
- comments
- theory … Turing machine
- recent work + future



comments …

- “clear that the methods could be scaled up to … larger graphs”

    + bath tub of DNA ?
    + suitable algorithms

- approximately 7 days of lab work

    + automation
    + alternative molecular algorithms

- possibility of errors

    + pseudopaths: accidental ligation
    + PCR, separation procedures
    + hairpin loops
    + stability when scaled

comments …

- "power of this method of computation"
  - 10^{14} operations 10^{20} plausable
  - exceed supercomputers by thousandfold

  :)

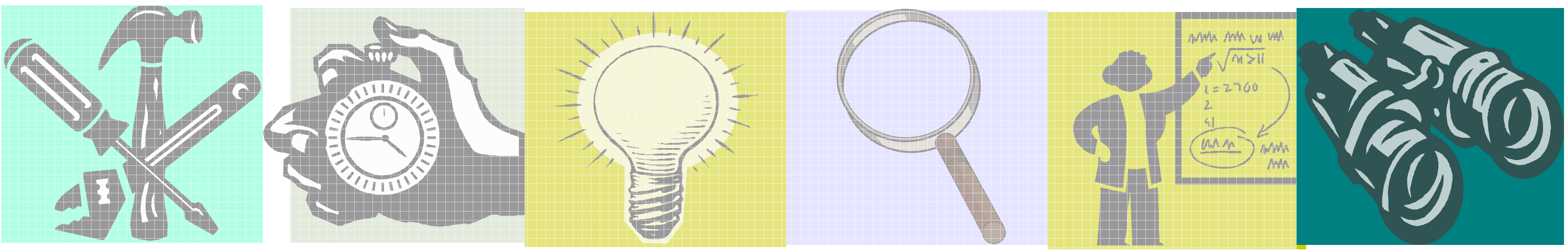
- "not clear whether … used to solve real computational problems"

  . multiplying 100 digit numbers

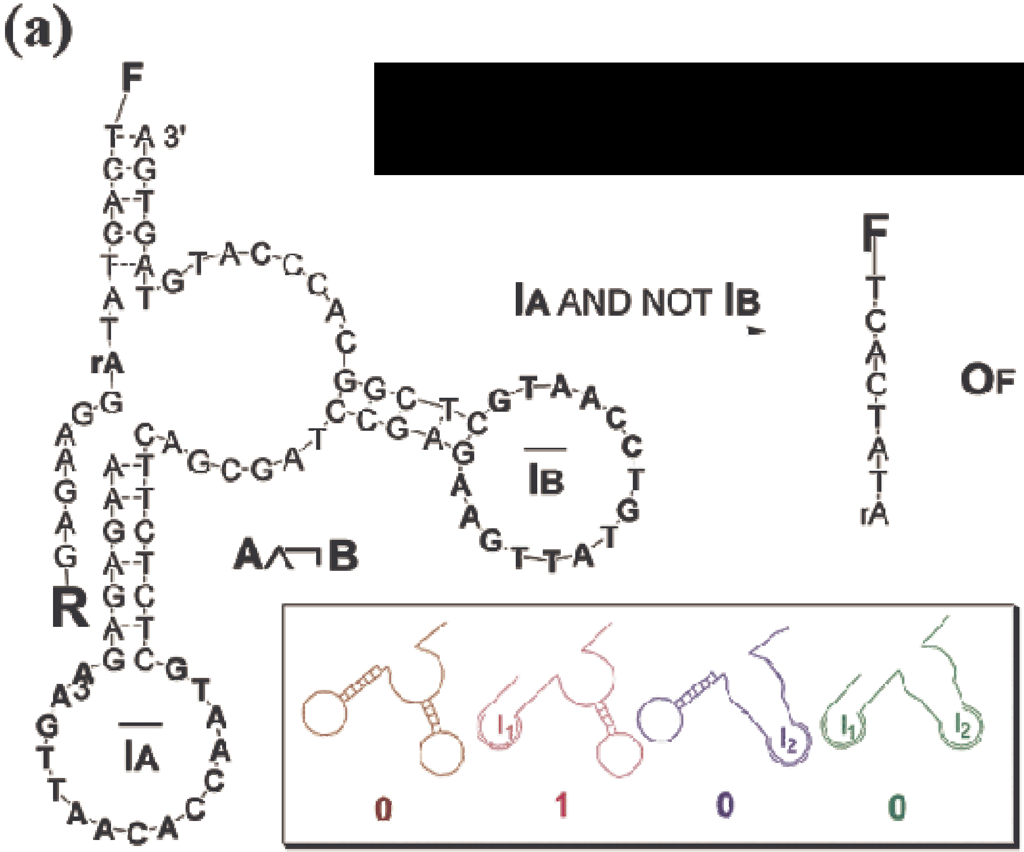
- potential: massively parallel searches

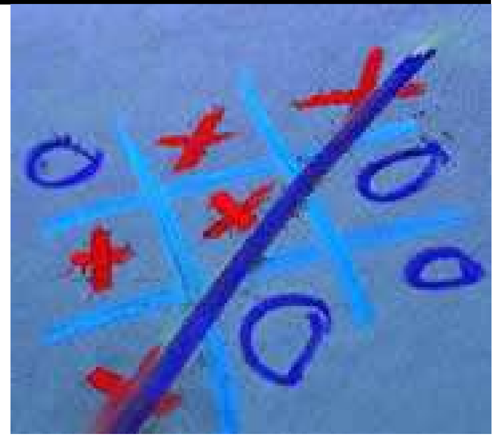
contents

- DNA … the tool chest
- problem complexity … P & NP
  Hamilton Path Problem
- Adleman's algorithm
- comments
- theory … Turing machine
- recent work + future

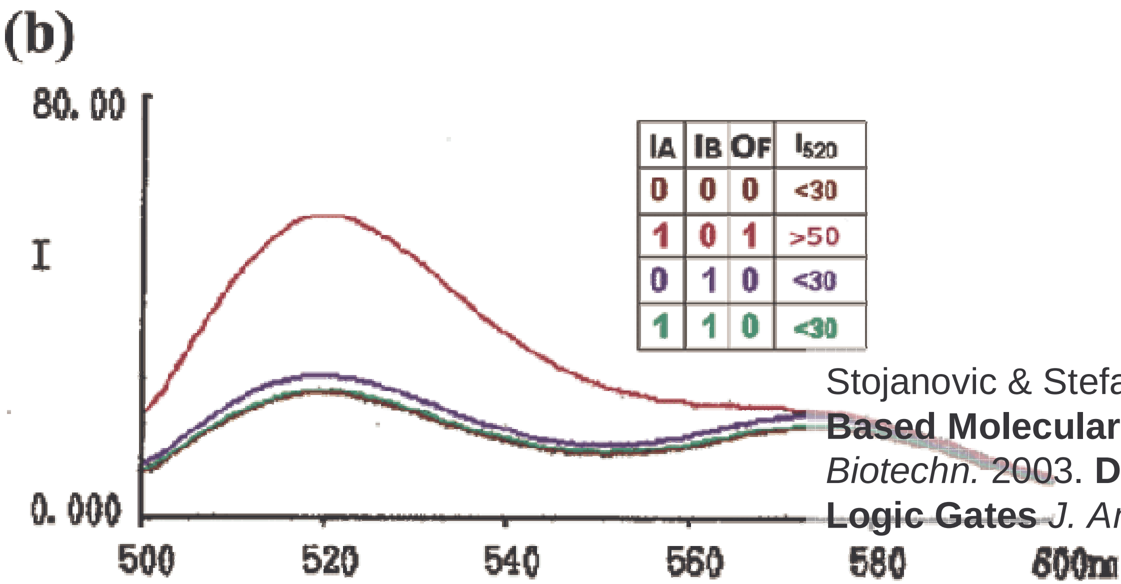


in Nature



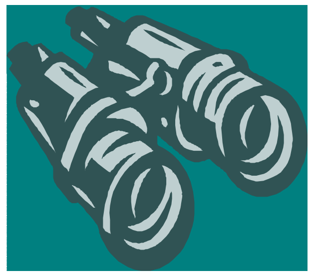


logic gates
fluorescence

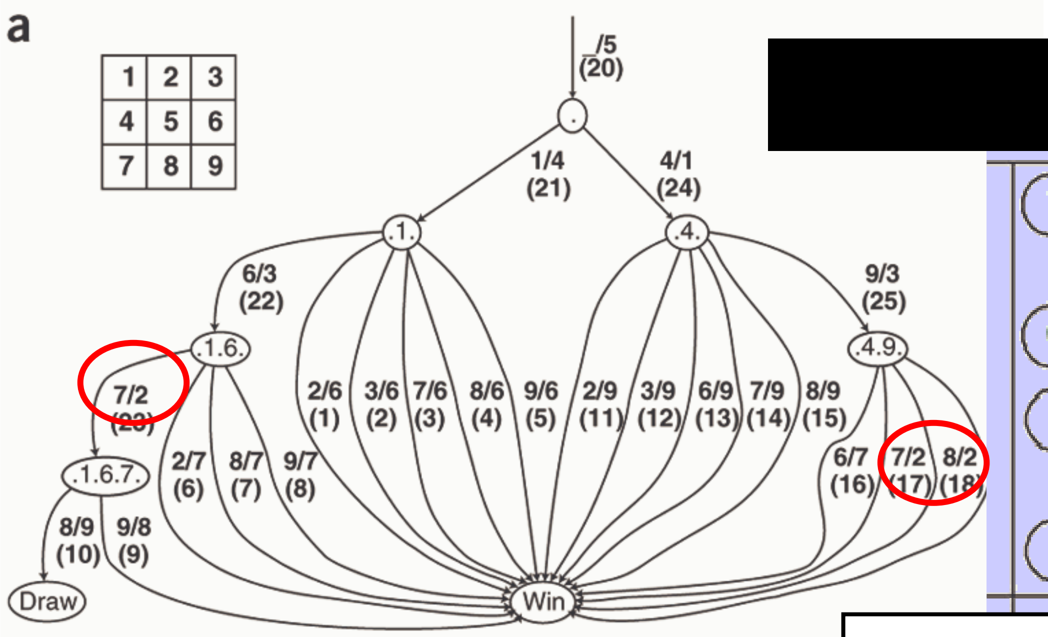


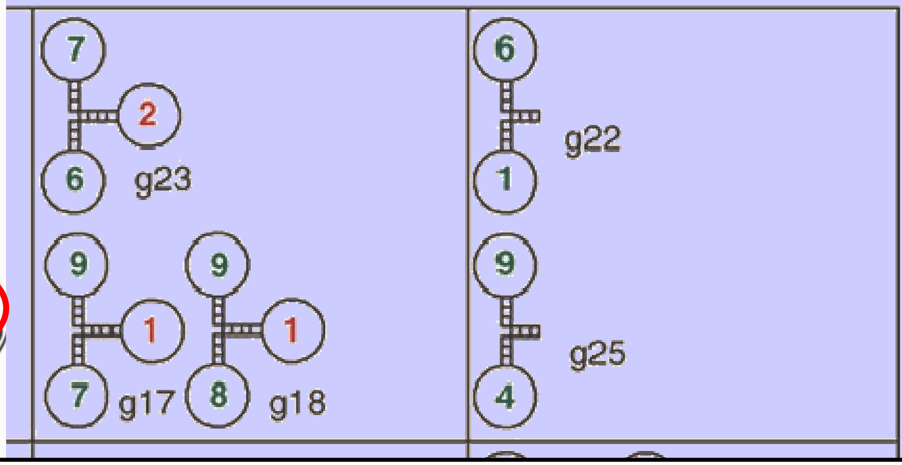
IA	IB	OF	I_{520}
0	0	0	<30
1	0	1	>50
0	1	0	<30
1	1	0	<30

Stojanovic & Stefanovic, **Deoxyribozyme-Based Molecular Automaton**. *Nature Biotechn.* 2003. **Deoxyribozyme-Based Logic Gates** *J. Am. Chem. Soc***.** 2002



in Nature





$$o_2 = (i_6 \wedge i_7 \wedge \neg i_2) \vee (i_7 \wedge i_9 \wedge \neg i_1) \vee (i_8 \wedge i_9 \wedge \neg i_1)$$

b

$$o_1 = \underbrace{i_4}_{\text{edge (24)}}$$

$$o_2 = \underbrace{(i_6 \wedge i_7 \wedge \neg i_2)}_{\text{edge (23)}} \vee \underbrace{(i_7 \wedge i_9 \wedge \neg i_1)}_{\text{edge (17)}} \vee \underbrace{(i_8 \wedge i_9 \wedge \neg i_1)}_{\text{edge (18)}}$$

$$o_3 = \underbrace{(i_1 \wedge i_6)}_{\text{edge (22)}} \vee \underbrace{(i_4 \wedge i_9)}_{\text{edge (25)}}$$

$$o_4 = \underbrace{i_1}_{\text{edge (21)}}$$

$$o_5 = \underbrace{1}_{\text{edge (20)}}$$

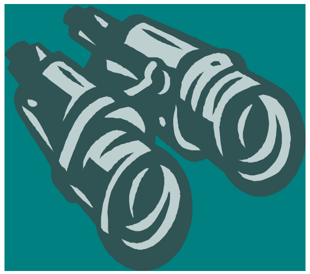
$$o_6 = \underbrace{(i_1 \wedge i_2 \wedge \neg i_6)}_{\text{edge (1)}} \vee \underbrace{(i_1 \wedge i_3 \wedge \neg i_6)}_{\text{edge (2)}} \vee \underbrace{(i_1 \wedge i_7 \wedge \neg i_6)}_{\text{edge (3)}} \vee \underbrace{(i_1 \wedge i_8 \wedge \neg i_6)}_{\text{edge (4)}} \vee ($$

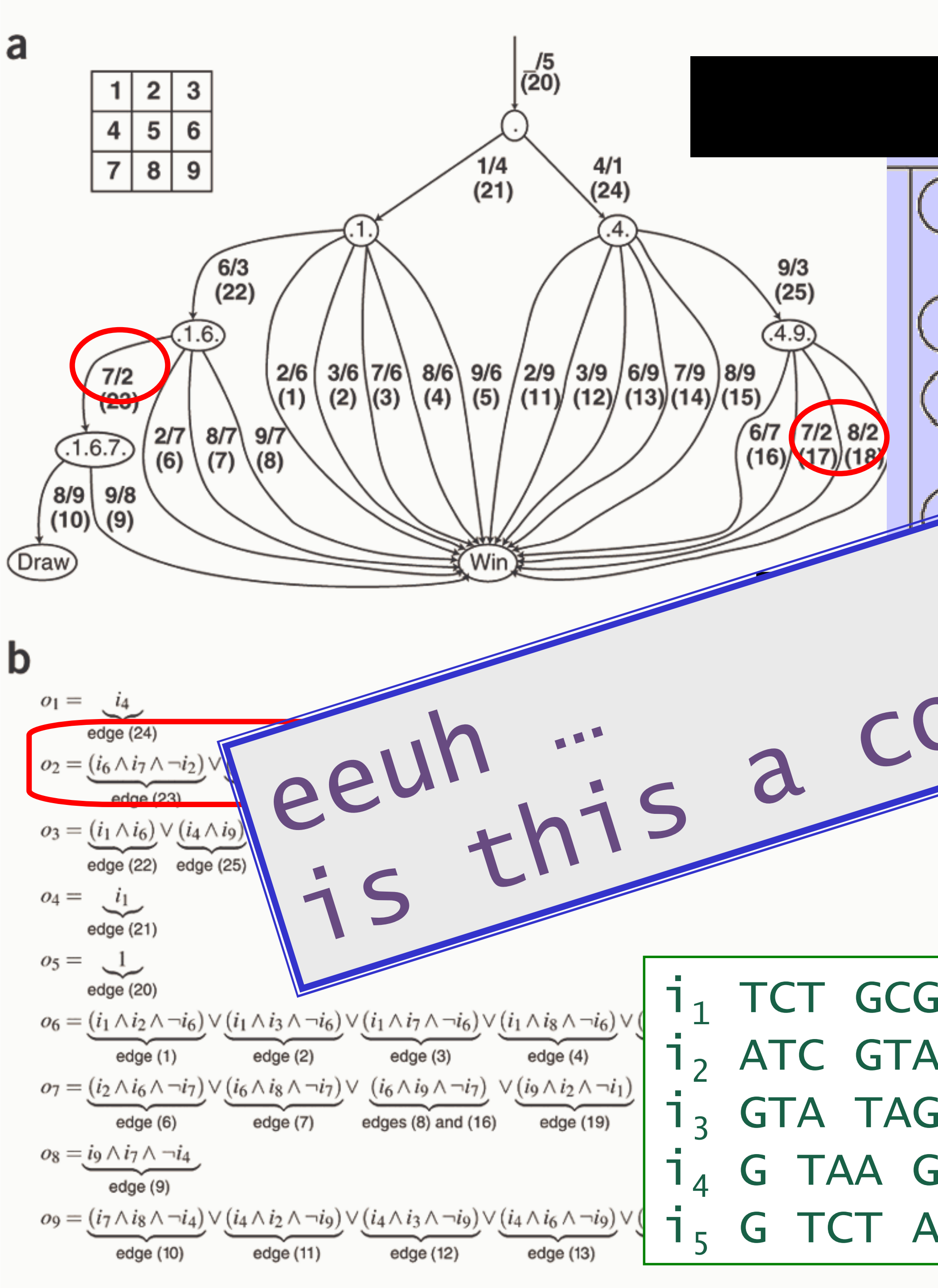
$$o_7 = \underbrace{(i_2 \wedge i_6 \wedge \neg i_7)}_{\text{edge (6)}} \vee \underbrace{(i_6 \wedge i_8 \wedge \neg i_7)}_{\text{edge (7)}} \vee \underbrace{(i_6 \wedge i_9 \wedge \neg i_7)}_{\text{edges (8) and (16)}} \vee \underbrace{(i_9 \wedge i_2 \wedge \neg i_1)}_{\text{edge (19)}}$$

$$o_8 = \underbrace{i_9 \wedge i_7 \wedge \neg i_4}_{\text{edge (9)}}$$

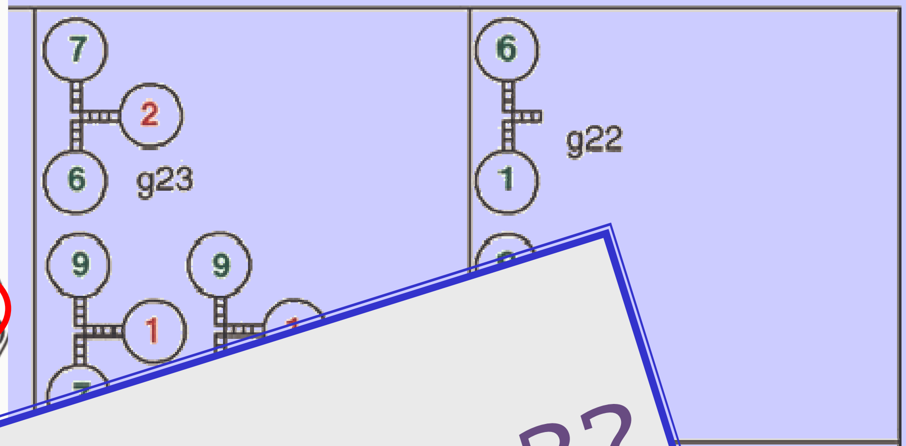
$$o_9 = \underbrace{(i_7 \wedge i_8 \wedge \neg i_4)}_{\text{edge (10)}} \vee \underbrace{(i_4 \wedge i_2 \wedge \neg i_9)}_{\text{edge (11)}} \vee \underbrace{(i_4 \wedge i_3 \wedge \neg i_9)}_{\text{edge (12)}} \vee \underbrace{(i_4 \wedge i_6 \wedge \neg i_9)}_{\text{edge (13)}} \vee ($$

i_1 TCT GCG TCT ATA AAT
i_2 ATC GTA TGT TGT TCA
i_3 GTA TAG TCT GTT TGT
i_4 G TAA GTG CTC AAA TGT C
i_5 G TCT AAT TCT CAC GGT C





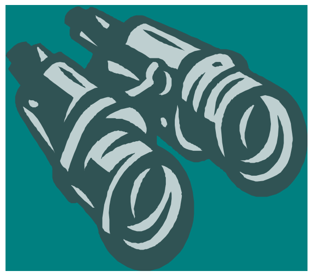
in Nature



$\ldots i_1) \vee (i_8 \wedge i_9 \wedge \neg i_1)$

eeuh …
is this a computer ??

i_1	TCT	GCG	TCT	ATA	AAT		
i_2	ATC	GTA	TGT	TGT	TCA		
i_3	GTA	TAG	TCT	GTT	TGT		
i_4	G	TAA	GTG	CTC	AAA	TGT	C
i_5	G	TCT	AAT	TCT	CAC	GGT	C



DNA computing today

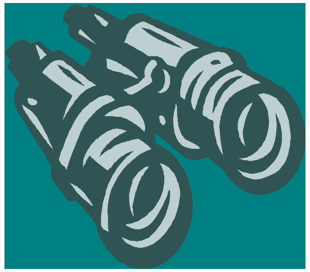
“There are many practical hurdles. Even with the best techniques of today, DNA still lags behind silicon computers,” says Ehud Shapiro. Instead, he advocates creating DNA devices that can do things, and go to places, that silicon can't - such as inside our cells, to make and control drugs.

...

Ultimately, Seeman hopes to build DNA scaffolding for electrical circuits, or for other molecular machines.

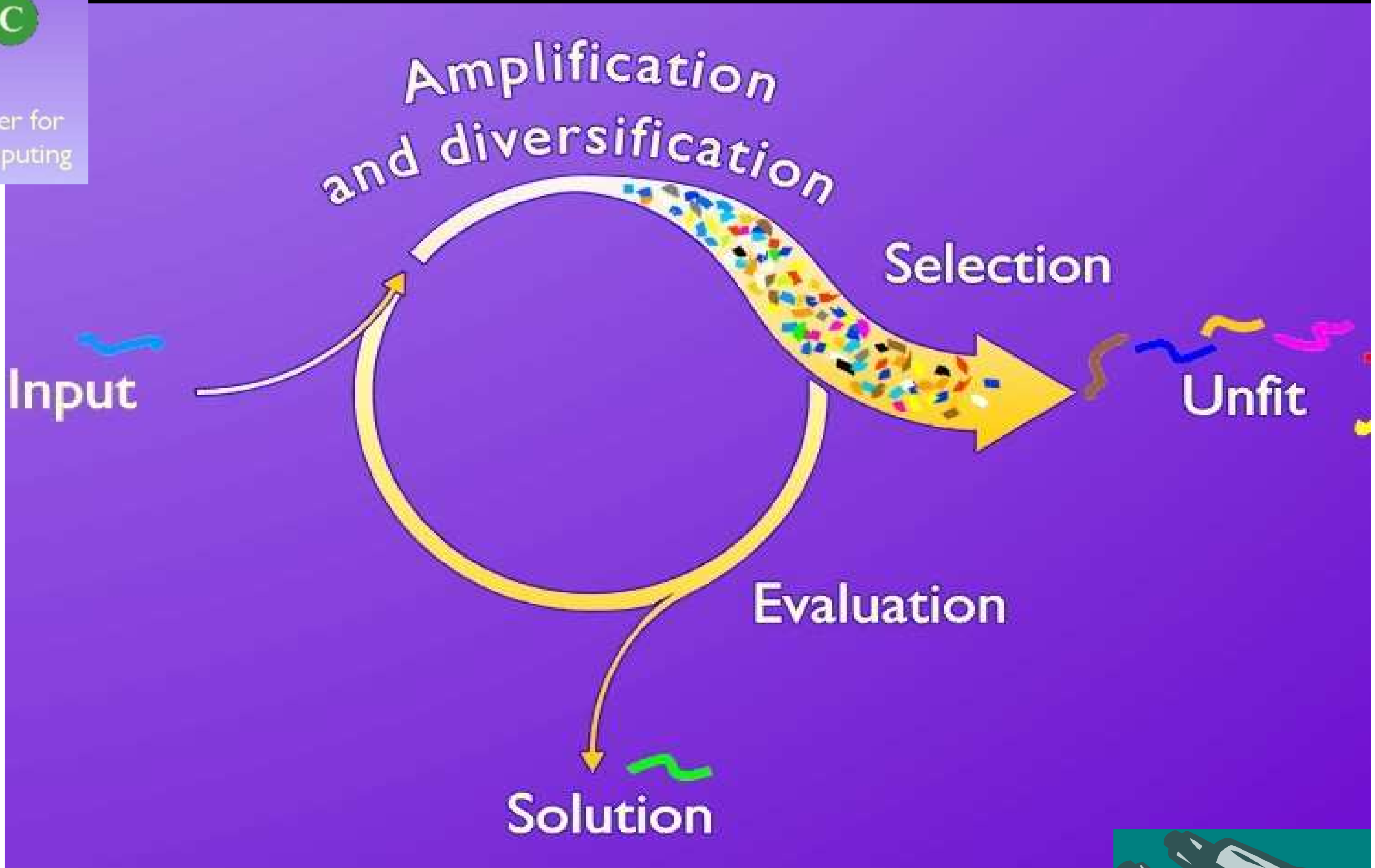
...

Yurke is focusing on DNA machines with moving parts. In 2000, he and his colleagues devised a set of DNA tweezers ....



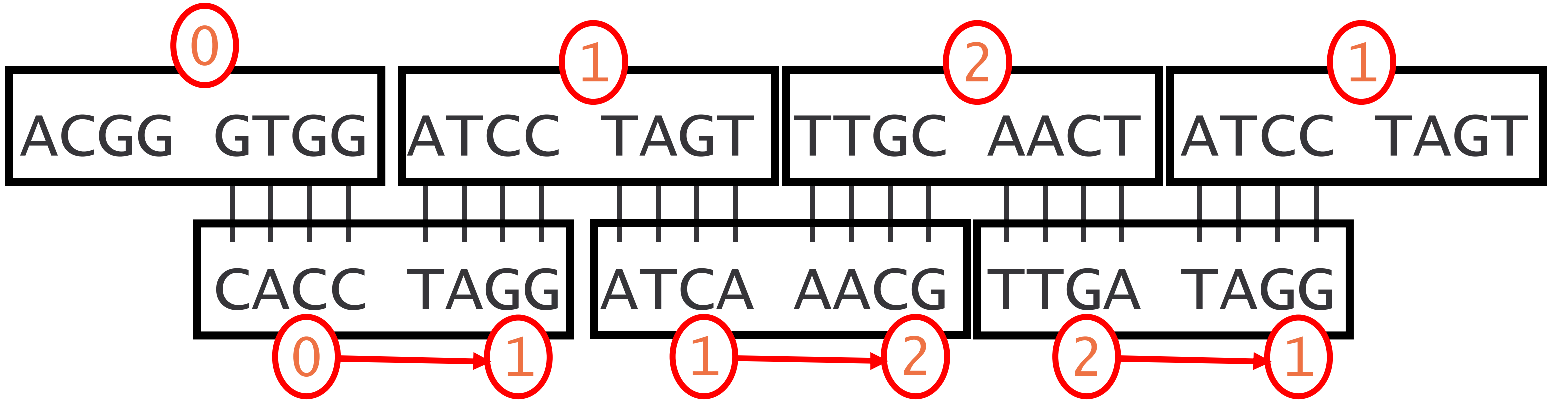
Nature News Service april 2003

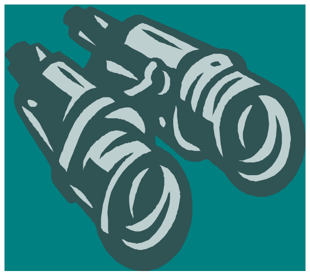
evolutionary DNA computing



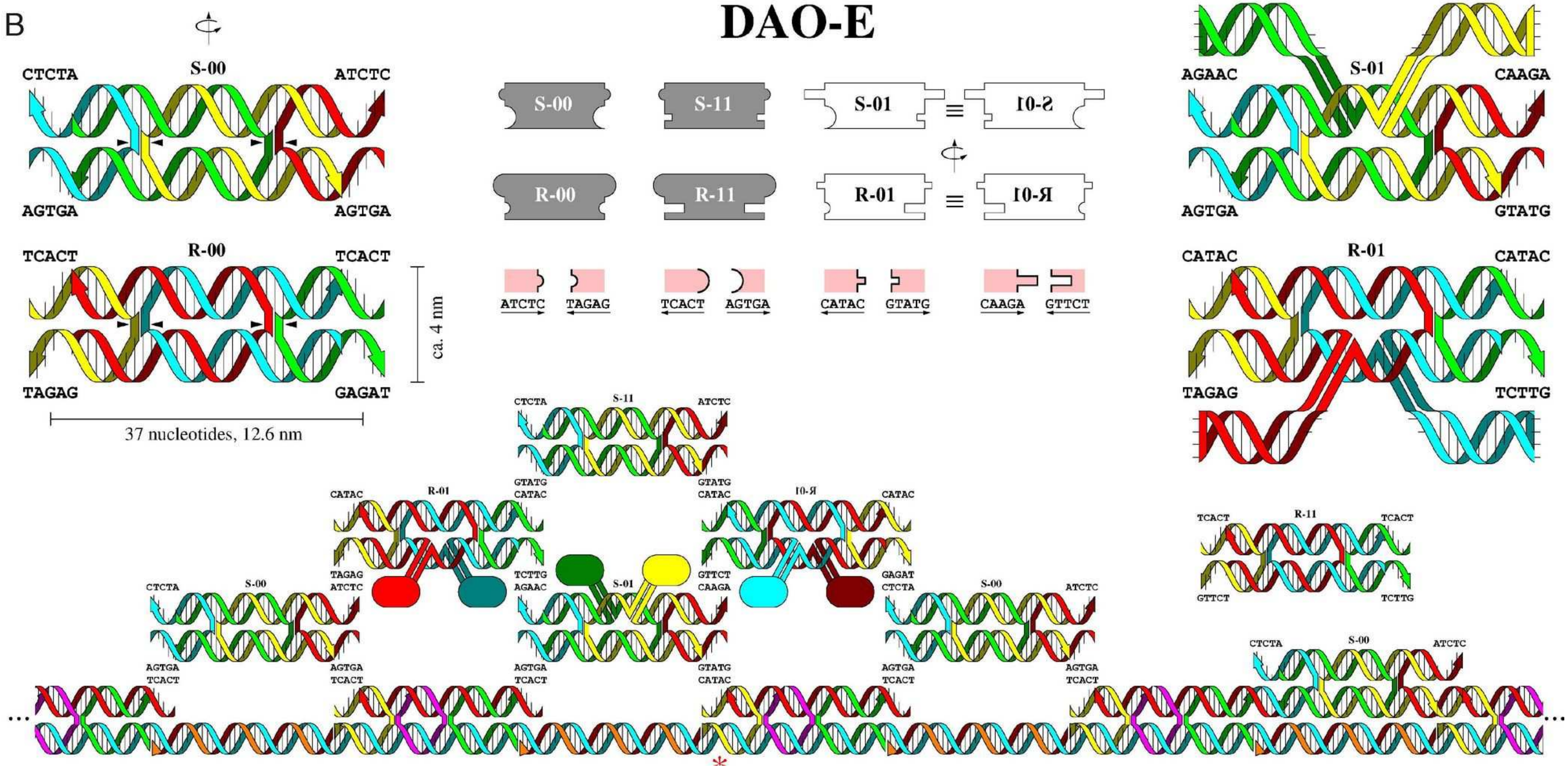
Cross-fertilization between evolutionary computation and DNA-based computing T.Back; J.N. Kok; G. Rozenberg Proceedings 1999 Evolutionary Computation.

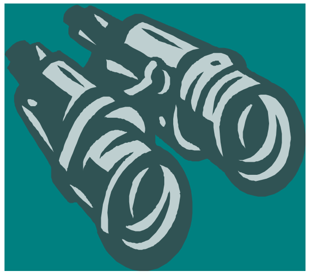
self assembly





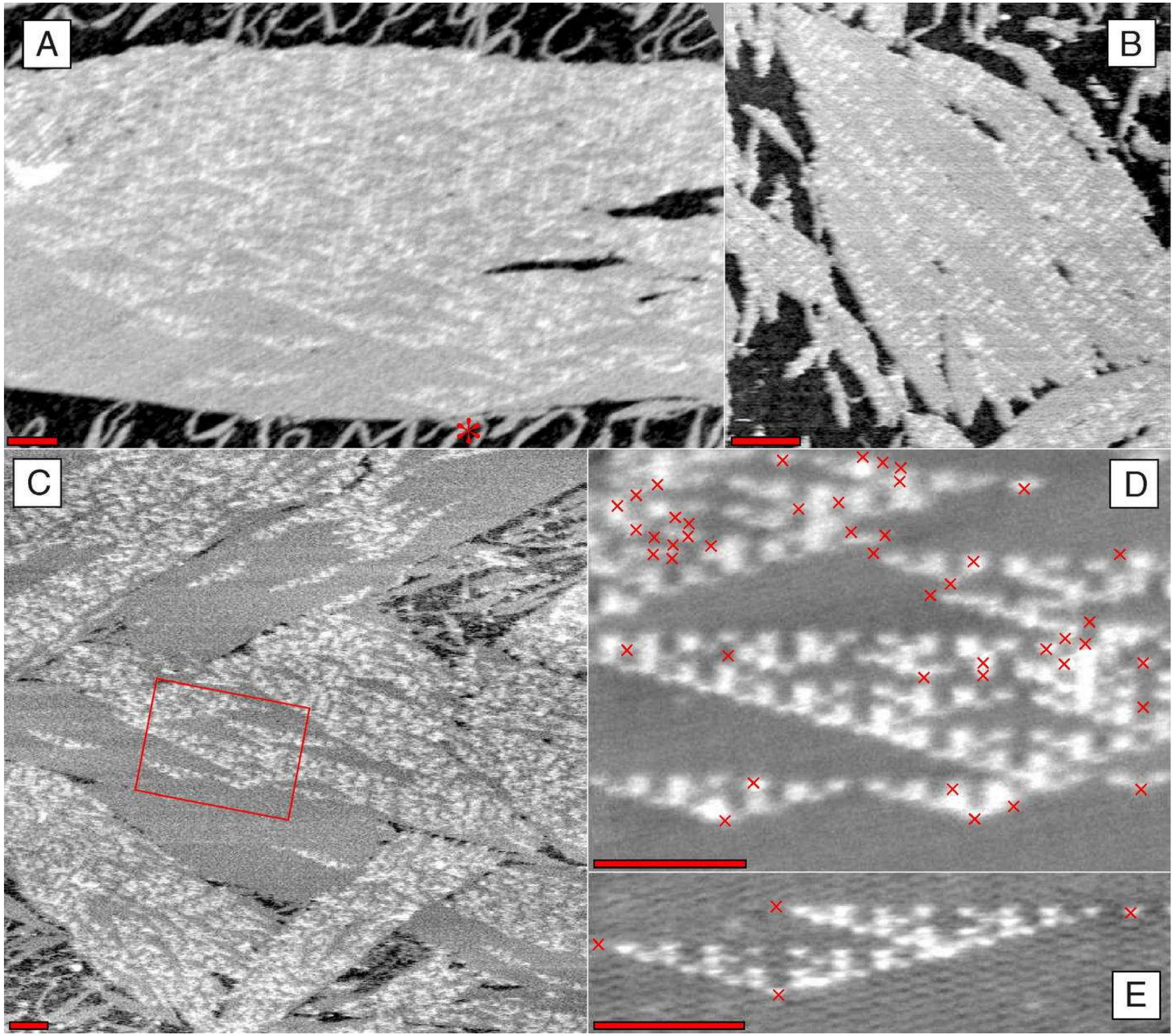
self assembly



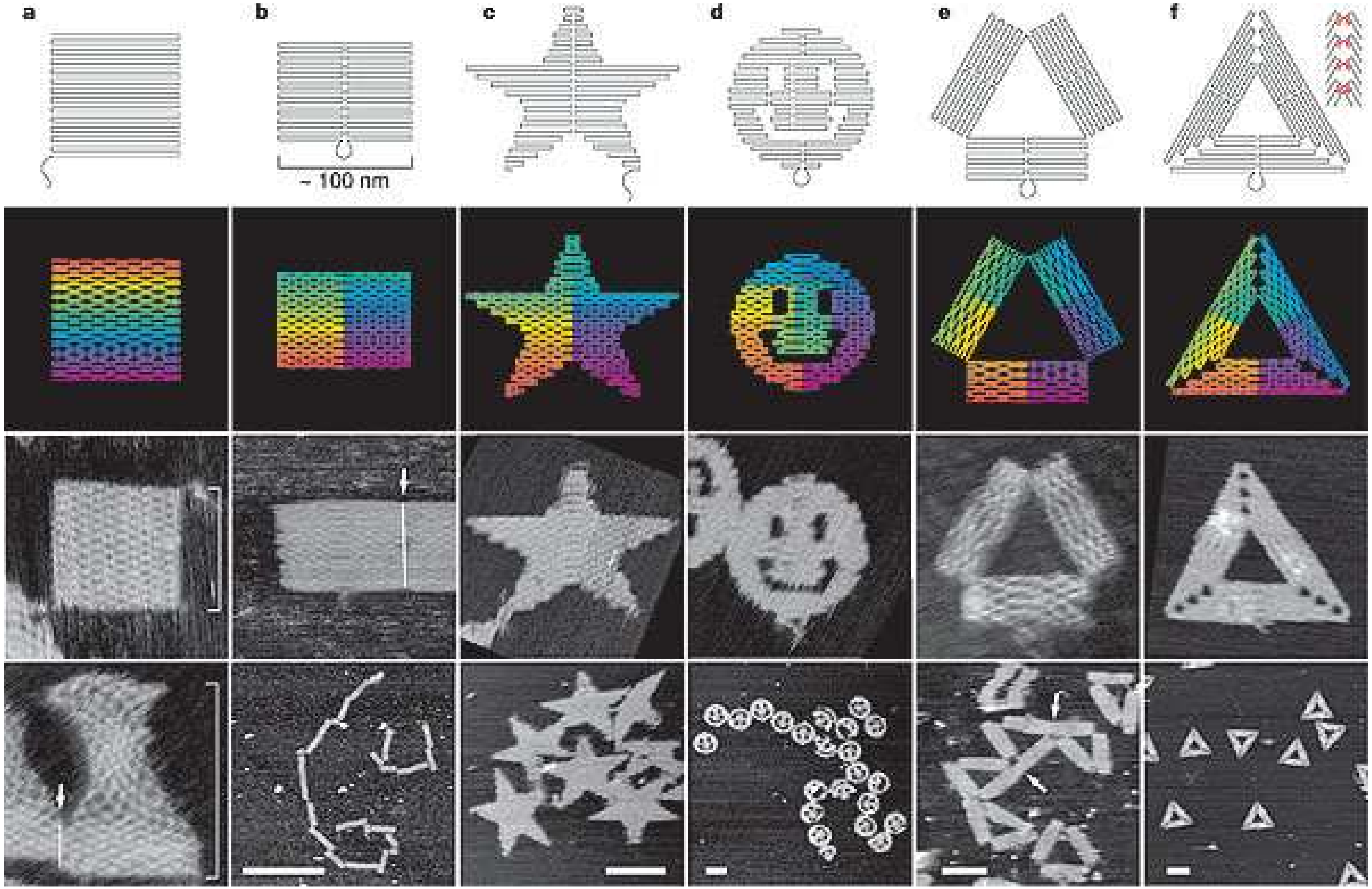


http://dx.doi.org/10.1371/journal.pbio.0020424
Algorithmic Self-Assembly of DNA Sierpinski Triangles
Rothemund, Papadakis, Winfree; PLoS Biology (2004)

self assembly



self assembly: DNA origami



Folding DNA to create nanoscale shapes and patterns
Paul W. K. Rothemund, Nature 440, 297-302 (16 March 2006)

the end …

perhaps not a computer …
but still some cool science!

PS …

DNA for cocktail parties

National Centre for Biotechnology Education
http://www.ncbe.reading.ac.uk/DNA50/cocktail.html

DNA for cocktail parties

Moisten the rim of a large test tube with lime juice then dip the rim into icing sugar.

Add about 10ml of blue curaçao to the tube.

Tilt the tube then with great care, pour about 20ml of ice-cold gin down the side of the tube to form a layer above the blue curaçao.

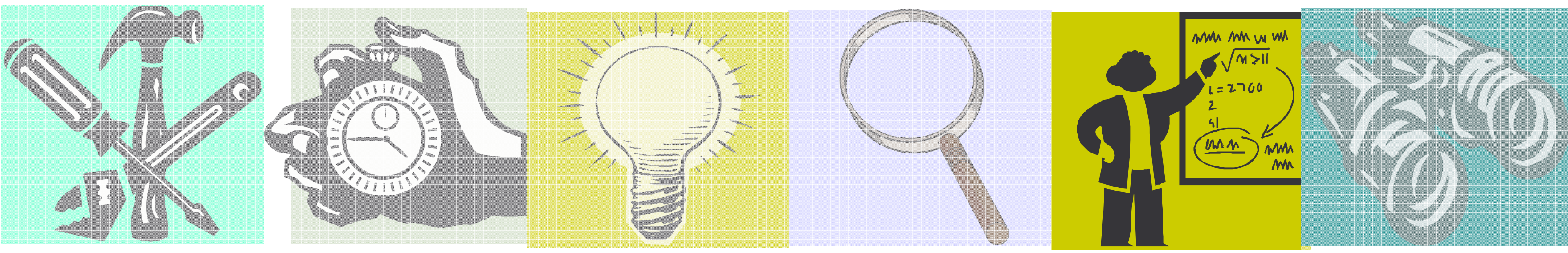
Blend the strawberries and pineapple juice for 10 seconds, then drop the purée on top of the gin. Wisps of strawberry D.N.A. will precipitate into the gin (see figure).

National Centre for Biotechnology Education

http://www.ncbe.reading.ac.uk/DNA50/cocktail.html

contents

- DNA … the tool chest
- problem complexity … P & NP
  Hamilton Path Problem
- Adleman’s algorithm
- comments
- theory … Turing machine
- recent work + future

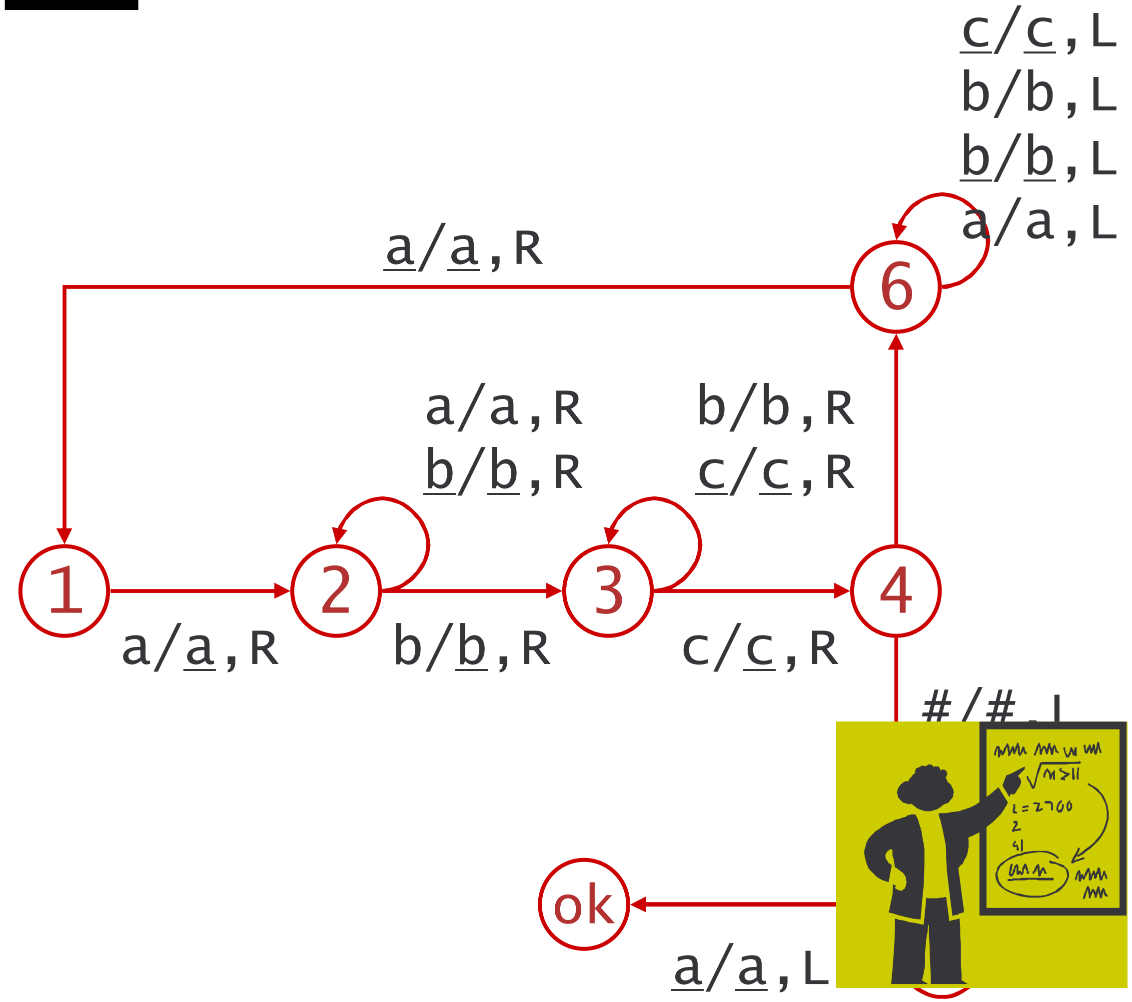


Turing machine

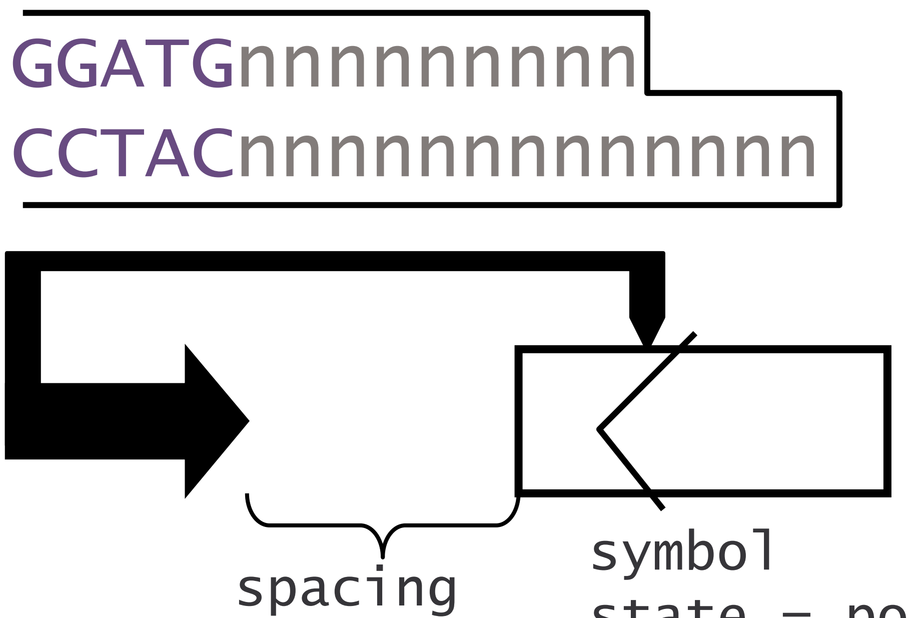
#	a	a	a	a	a	b	b	b	b	b	c	c	c	c	c	#

tape

1. mark a
2. move to b’s
   mark b
3. move to c’s
   mark c
4. if another c
5. then back to a’s
   goto 1.
   else back to a’s
6. check marks
   stop



‘universal’ Turing machine



state = position of cut

Rothemund
*Fok*I
circular DNA

- cut states with restriction enzyme
- mix ‘instructions’ with ‘tape’
- ‘activate’ instructions (cut protected end)
- ligate to form circles
- cut old symbol
- recircularize