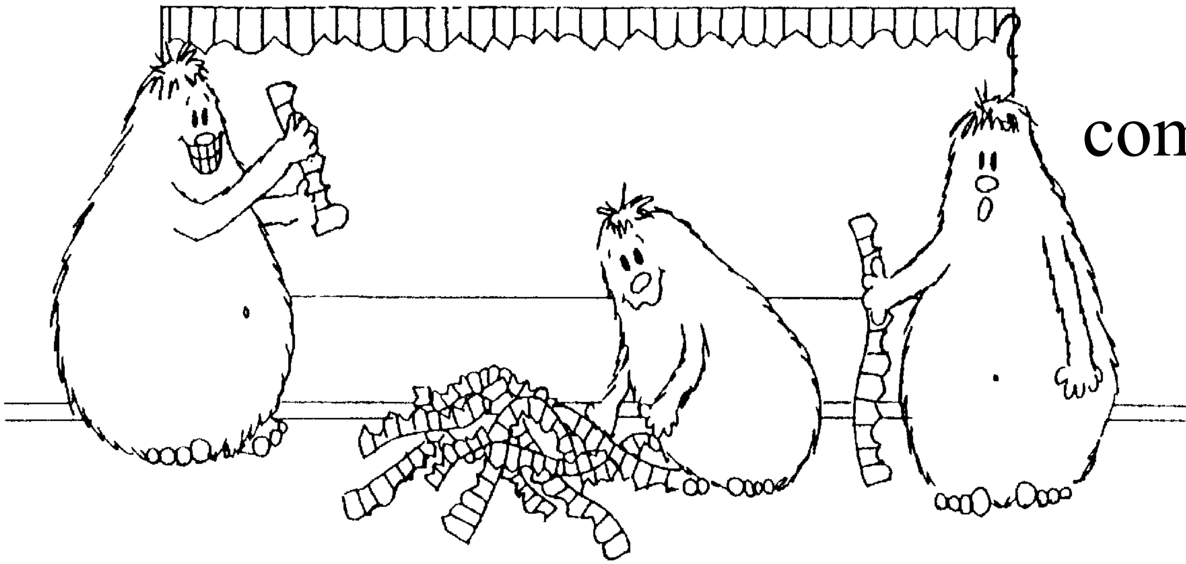
And now for something



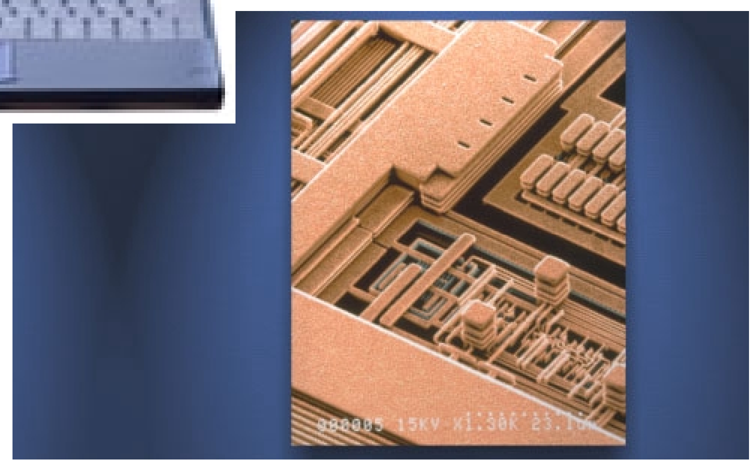
completely different…

Computing with DNA…!?

Computing

- computers….
- screen, keyboard
- RAM, ROM
- Megahertz, Gigabytes

- electronic components
- silicon substrate

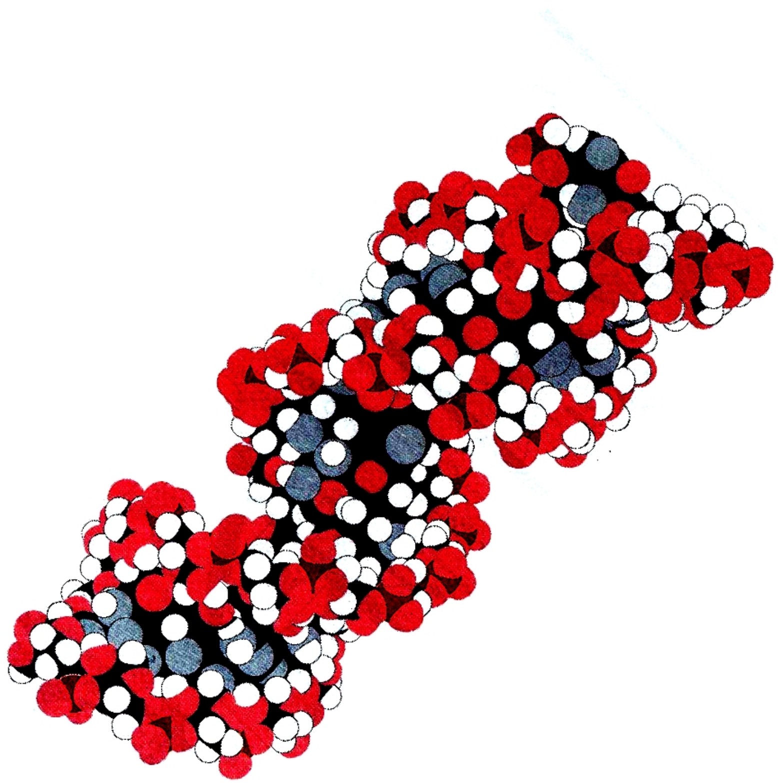


alternative forms…?

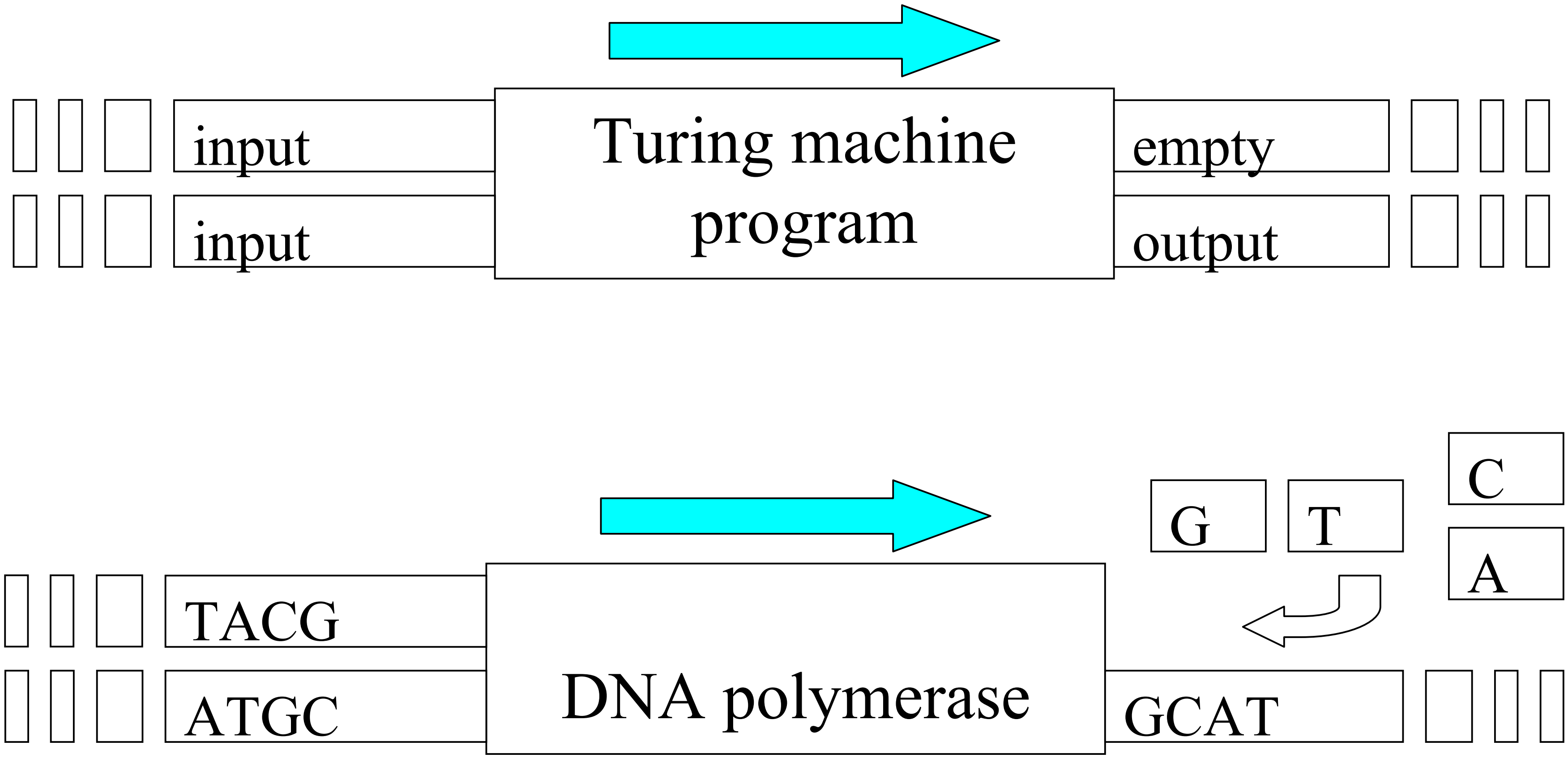
computing with DNA

- Leonard Adleman (1994) mathematician
  - AIDS modeling studies
  - molecular biology

- DNA polymerase….
- ...works like Turing´s toy computer!



Turing machine



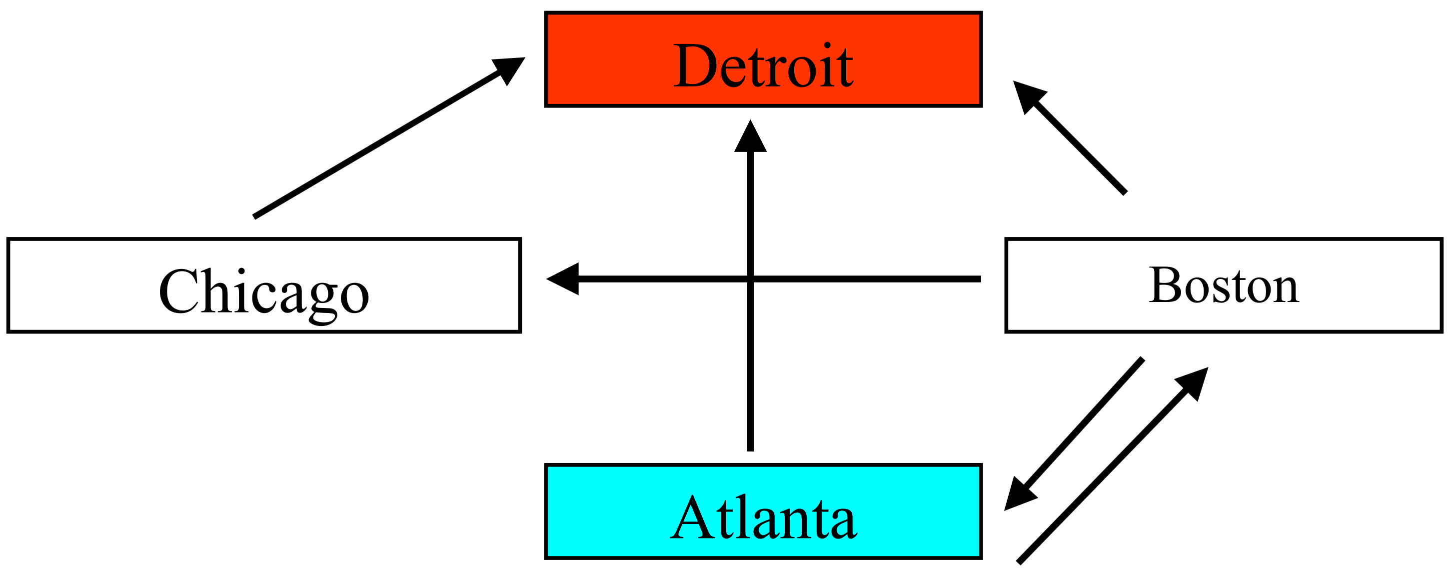
what does it take to build a computer?

- method to store information…
  - DNA
- simple operations to act on information
  - ‘hybridization logic’
  - certain enzymes

choose problem to solve…

Hamiltonian path problem

travelling salesman problem



Does a path exist that:

- commences in start city (Atlanta)

- finishes in end city (Detroit)

- passes through each remaining city only once…?

the Hamiltonian path problem, HPP

- map................**graph**
- city................**vertex**
- nonstop flight... **path**

- Hamiltonian path for a given graph is a path that:
  - starts at start vertex
  - ends at end vertex
  - passes through each remaining vertex only once

- HPP is to decide whether a Hamiltonian path exists or not….

solutions to HPP

- no efficient (= fast) algorithm found yet
- computing time needed to solve it, is increasing exponentially with the number of vertices…
  - t ~ exp (c*N)
  - best computers, best algorithms…..
  - < 100 vertices…more than 100 years
- 1970: **HPP is NP-complete**…no efficient algorithm possible

an algorithm to solve the HPP

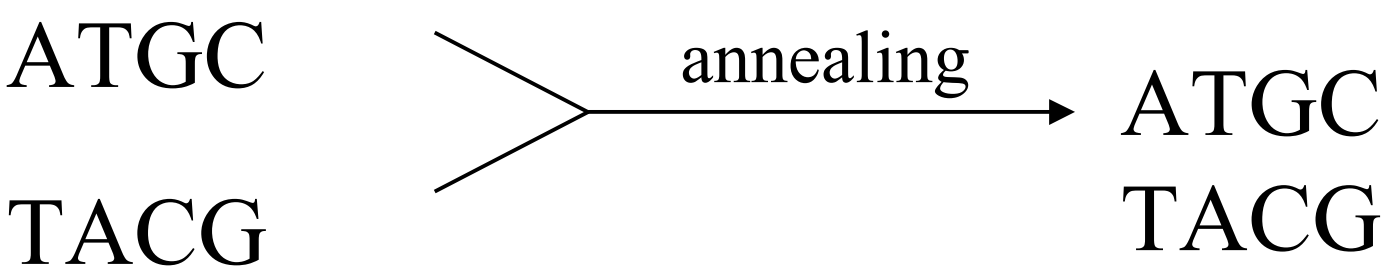
- given a graph with N vertices
  - **(1)** generate a set of random paths through the graph
- for each path in the set:
  - **(2a)** check whether path starts at start vertex and ends with the end vertex.
  - **(2b)** check if that path enters exactly N vertices.
  - **(2c)** for each vertex, check if that path enters that vertex.
- if set is
  - **(3)** not empty: Hamiltonian path found.
  - empty: no Hamiltonian path exists.

implementation in DNA…?

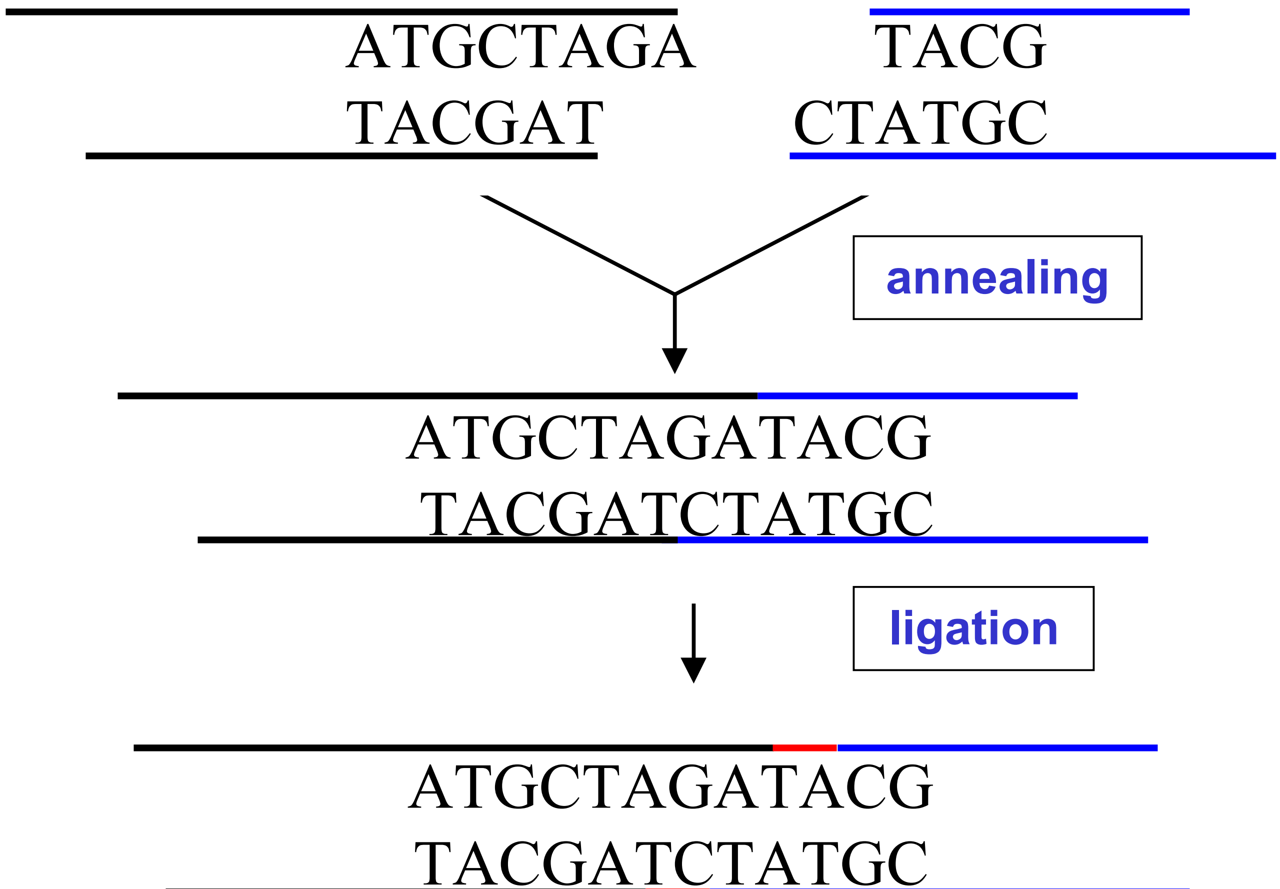
tools: Watson-Crick pairing

Rule: complementarity

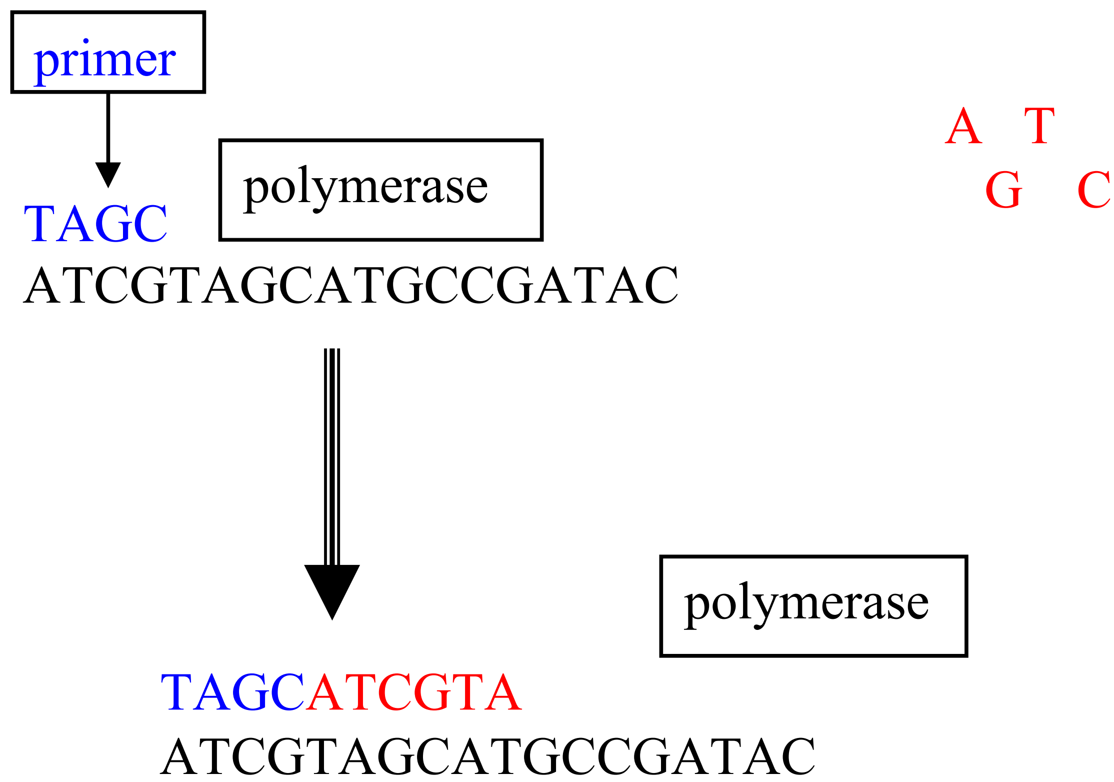
A=T, C=G



tools: DNA ligase



tools: DNA polymerases



tools: PCR

CGTCAGCCTGACCCGATACAGGCT
GCAGTCGGACTGGGCTATGTCCGA

$95^{o}C$

CGTCAGCCTGACCCGATACAGGCT

GCAGTCGGACTGGGCTATGTCCGA

$30^{o}C$

CGTCAGCCTGACCCGATACAGGCT
GCAG
GGCT
GCAGTCGGACTGGGCTATGTCCGA

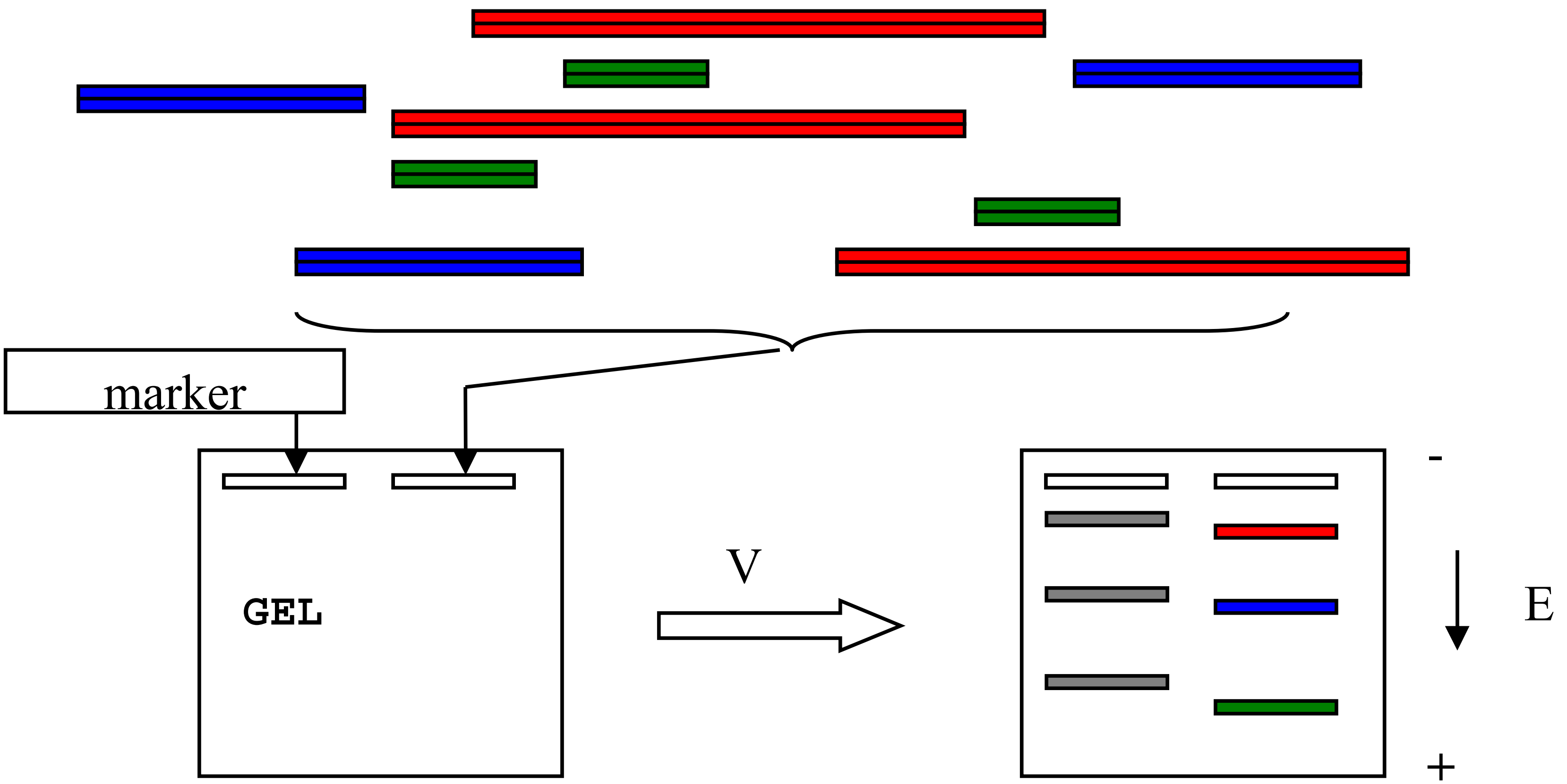
$74^{o}C$

CGTCAGCCTGACCCGATACAGGCT
5’**GCAGTCGGACTGGGCTATGTCCGA**

CGTCAGCCTGACCCGATACAGGCT 5’
GCAGTCGGACTGGGCTATGTCCGA

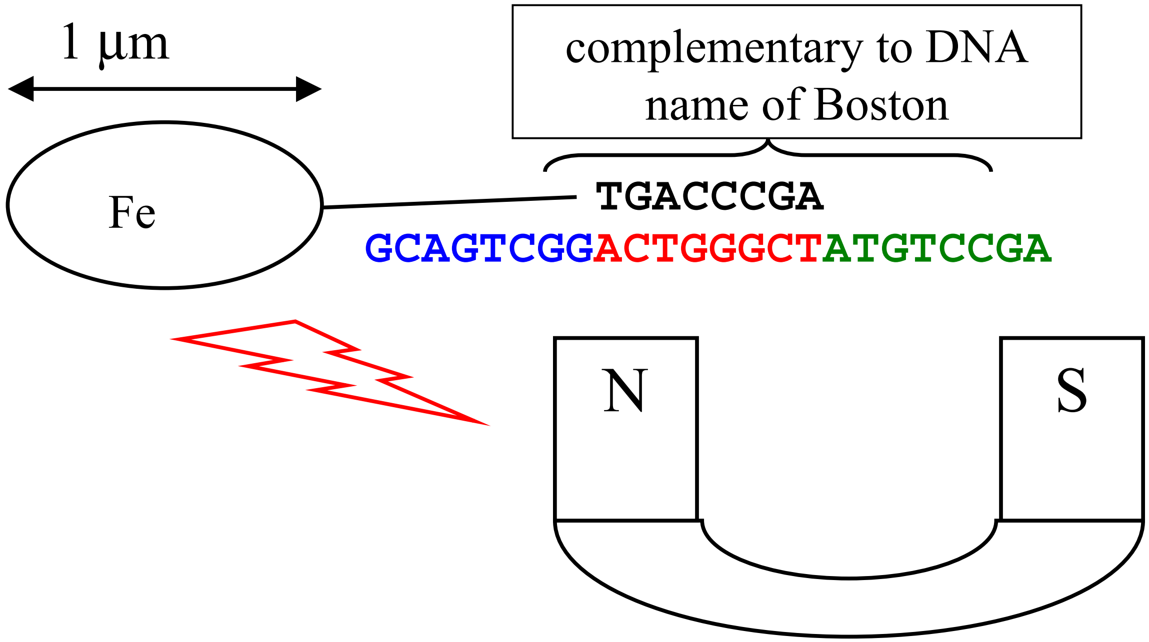
tools: gel electrophoresis

Heterogeneous mixture of DNA fragments



tools: purification

check if Boston is present…



-molecules with Boston sequence will bind to probe on beads
-keep beads in tube with magnet, throw rest of solution
-remove molecules from beads and test for presence of Chicago

HPP encoding in DNA

- assign random DNA sequence (8 nucleotides) to each city.

Atlanta – ACTTGCAG

firstname

lastname

Boston - TCGGACTG

- assign each nonstop flight a DNA flight number.

Atlanta -to- Boston

ACTTGCAG TCGGACTG

Flight #: GCAGTCGG

solving the HPP by DNA (1)

- synthesize:
  - complementary DNA city names
    - Atlanta – ACTTGCAG
      Compl. – TGAACGTC
  - DNA flight numbers
- calculation procedure
  - take 1014 molecules of each sequence
  - add water, salt and ligase
  - mix, et voila…
- the solution to the HPP is in your hand…

how does this happen…?

solving the HPP by DNA

Atlanta-to-Boston flight #:

GCAGTCGG

meets with…

complementary name of Atlanta

TGAACGTC

and Boston…

AGCCTGAC

they anneal and are ligated…

TGAACGTCAGCCTGAC
GCAGTCGG

solving the HPP by DNA

- all paths created at once by
  - simultaneous interactions of ~ 1015 molecules.
    - …… parallel processing!
  - small number cities, high number of DNA molecules…
  - almost certain that molecule encoding Hamiltonian Path is present.

Atlanta-Boston-Chicago-Detroit

TGAACGTCAGCCTGACCCGATACAGGCTCGTT

GCAGTCGGACTGGGCTATGTCCGA

how to get <u>this molecule</u> out of the mixture…?

solving the HPP by DNA (2a)

- select molecules that begin with start city and terminate with end city.
- polymerase chain reaction (PCR)
  - Needed: many copies of 2 primers
  - last name of start city: GCAG
  - complement of first name end city: GGCT
  - mix primers with Taq polymerase, nucleotides and primers and do PCR
- molecules with both right start and end cities are reproduced at exponential rate…
  - start: 1 molecule…
  - n cycles: 2^{n-1}
- molecules with only right start or end city are reproduced at much slower linear rate…
- molecules with neither right start or end city are not duplicated at all…
- at end of PCR:
  - many molecules with both right start and end city….
  - few that didn't meet criterion…

solving the HPP by DNA (2b)

- check if that path enters exactly n vertices.
  - all four cities, three fights are needed.
  - so, select only the molecules with the correct length...24 nucleotides
- technique used: gel electrophoresis.
  - Seperate DNA on length...
  - cut out and re-isolate only DNA of 24 nucleotides...

solving the HPP using DNA (2c)

- for each city, check if that path passes through that city. If not, remove that path from the set.

- technique: affinity separation
  - multiple copies of DNA ‘probe’ molecule
  - probe molecule encodes the complementary sequence of a particular city.
  - probes are attached to microscopic iron balls
  - use magnet to attract iron balls

solving the HPP using DNA (3)

- after affinity separations step is algorithm complete
- molecules left in tube are encoding Hamiltonian paths.
- determine if any DNA was present using PCR followed by gel-electrophoresis.

after sequencing, the final solution is:

GCAGTCGGACTGGGCTATGTCCGA

GCAGTCGG: Atlanta - Boston

ACTGGGCT: Boston - Chicago

ATGTCCGA: Chigago - Detroit

DNA computing… summary

- based on hybridisation logic
- massive paralellism
- potentially suited for solving computationally hard problems
- generative algorithms:
  - build solution sequence from components (shown in talk)
  - hybridisation-ligation-amplification
- substractive algorithms:
  - remove incorrect sequence solutions
  - hybridisation-cleavage-amplification
- nice proof-of-principle experiments
- but real-world problems pose difficulties:
  - HPP with 23 cities: 1 kg of DNA
  - 70 cities: 1025 kg of DNA!
- accuracy and error-rate of biochemical procedures…
  - DNA 1 in 6000
  - Intel 1 in 10^9
- technical improvement + error-correcting procedures….