DNA Computing

Molecular Computation of Solutions to Combinatorial Problems

by Leonard M. Adleman in 1994

Overview

- Motivation (DHP Problem)
- The Approach by Dr. Aldeman using DNA Computing
- Pros and Cons of the Approach

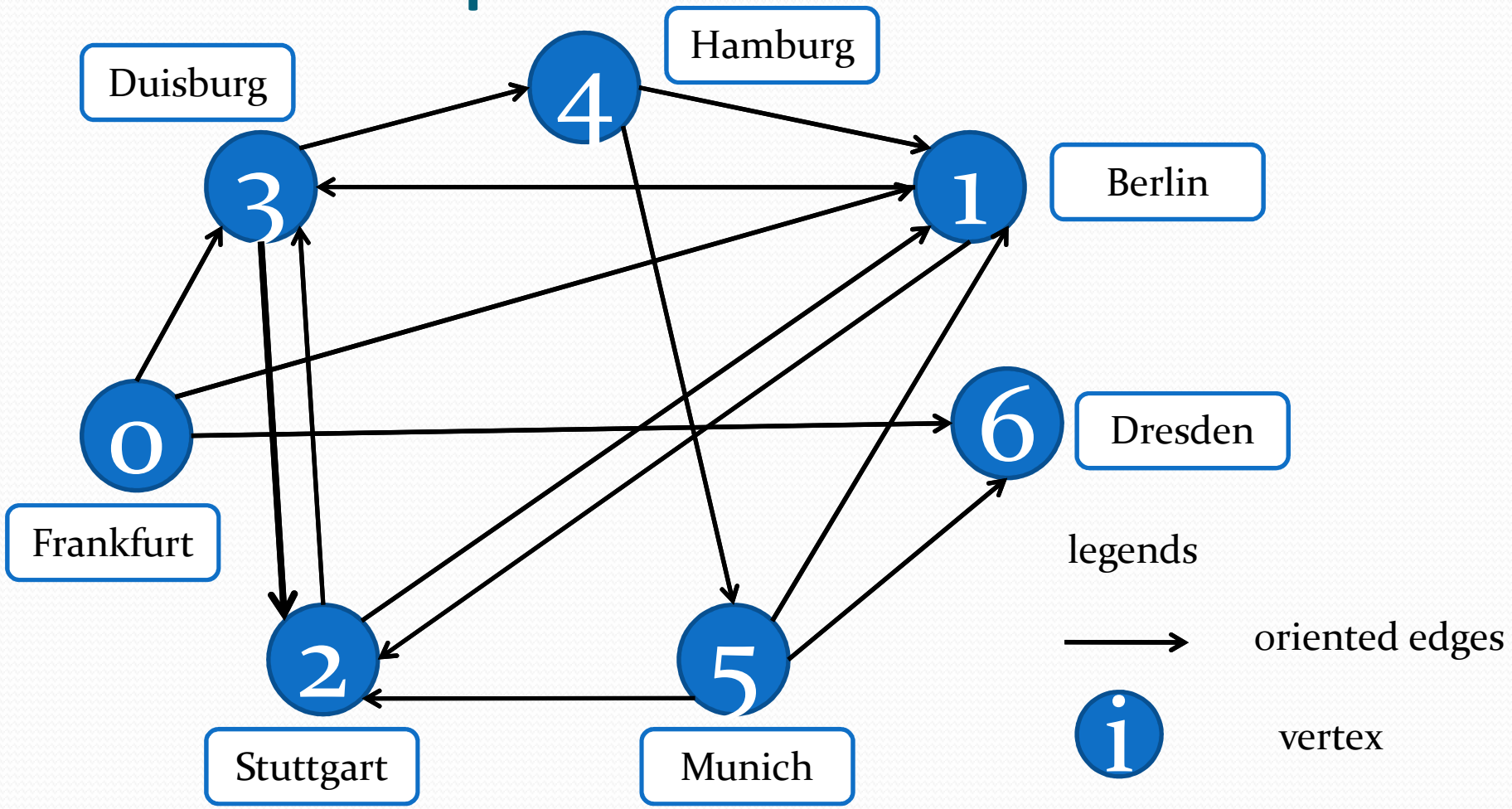
Motivation

- Directed Hamiltonian Path (DHP) Problem

  Given: an oriented graph, which consists of points (vertices) and of arrows (oriented edges).

  Problem: find a path through the graph that starts and ends at the given vertices (V_{in} and V_{out})and includes every other vertex exactly once.

Example of DHP Problem



Nondeterministic Algorithm by Aldeman

1. Generate Random paths through the graph
2. From all paths created in step 1, keep only those that start at V_{in} and end at V_{out}
3. From all remaining paths, keep only those that visit exactly n vertices.
4. From all remaining paths, keep only those that visit each vertex at least once.
5. if any path remains, return “yes”;otherwise, return “no”.

Step 1: Generate Random paths

- each vertex i in the graph is associated with a random 20-mer sequence of DNA denoted Oi.

  O_2 = ACTACGATTCCAGTACGACT

  O_3 = GGTACAGTCCATGAGCGTAT

  O_4 = CTGTGACAAGTCACGACTAT

- The reverse complementary strand is $\underline{O_i}$

  $\underline{O_2}$ = AGTCGTACTGGAATCGTAGT

  $\underline{O_3}$ = ATACGCTCATGGACTGTACC

  $\underline{O_4}$ = ATAGTCGTGACTTGTCACAG

Step 1: Generate Random paths

- Each edge i->j is presented by an oligonucleotide $O_{i->j}$ that is the 3' 10-mer of Oi followed by the 5' 10-mer of Oj

O2 = ACTACGATTCCAGTACGACT
O3 = GGTACAGTCCATGAGCGTAT
O4 = CTGTGACAAGTCACGACTAT

$O_{2->3}$ = CAGTACGACTGGTACAGTCC
$O_{3->4}$ = ATGAGCGTATCTGTGACAA

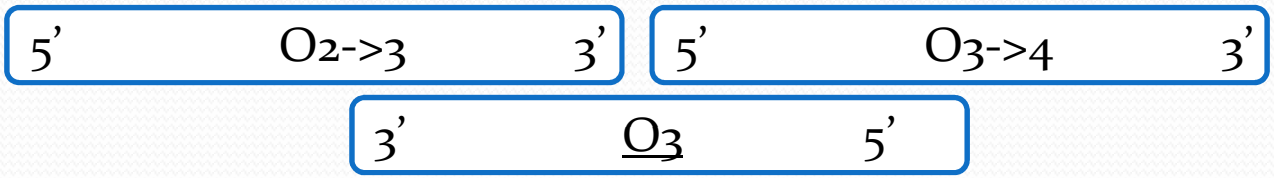
- For the edges $O_{i->j}$ invovled the begin and end (begin i = 0; end j = 6)

$O_{0->j}$ consists of 3' 20-mer of O0 and 5' 10-mer of Oj
$O_{i->6}$ consists of 3' 10-mer of Oi and 5' 10-mer of O6

- Because DNA has a 3' end and a 5' end
  $O_{2->3} \neq O_{3->2}$ (Preserves edge orientation)

Step 1: Generate Random paths

- 50 pmol of $\underline{O_i}$ and 50 pmol of $O_{i->j}$ are mixed together in a single ligation reaction



$\underline{O_3}$ = ATACGCTCATGGACTGTACC

$O_{2->3}$ = CAGTACGACTGGTACAGTCC

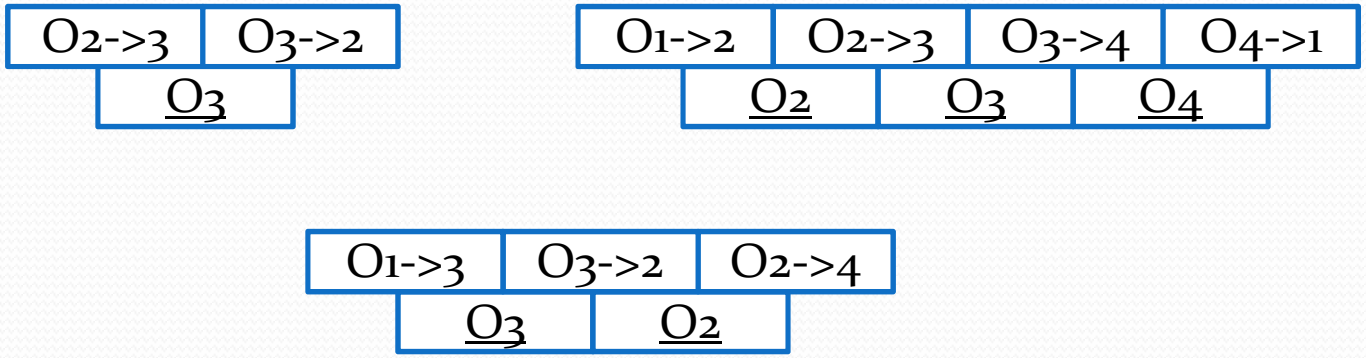
$O_{3->4}$ = ATGAGCGTATCTGTGACAA

- $\underline{O_i}$ serves as splints to bring $O_{i->j}$

Through ligation many random paths will be created

Step 1: Generate Random paths

- Examples of random paths



Step 2: keep paths that start at Vin and end at Vout

- The product of Step 1 is amplified by PCR using primers O0 and O6

-> only those molecules encoding paths that begin with vertex 0 and end with vertex 6 are amplified

Problem:

O0->3	O3->4	O4->5	O5->6
O0	O3	O4	O5	O6

Step 3: keep only those that visit exactly 7 vertices

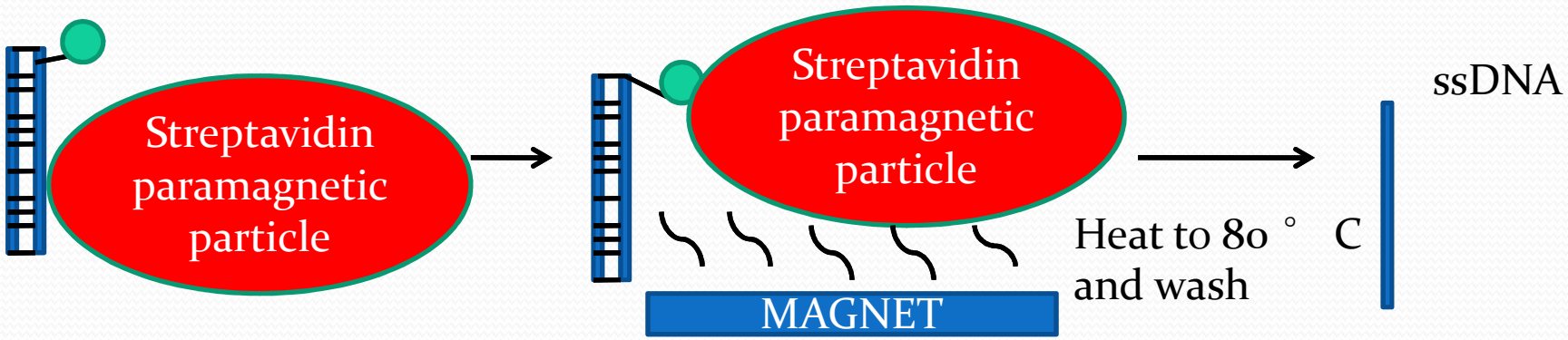
- The product of Step 2 is run on an agarose gel and the 140 bp band (for 7 vertices) was excised and soaked in ddH2O
- The product was PCR-amplified and gel-purified several times to enhance its purity

O0->3	O3->4	O4->5	O5->4	O4->5	O5->6
O0	O3	O4	O5	O4	O5	O6

Step 4: keep only those that visit each vertex at least once.

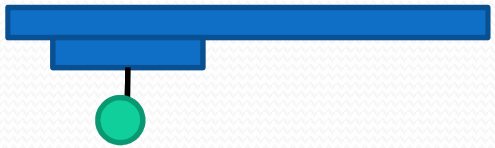
- From the product from Step 3 single stranded DNA are generated.

  PCR using primers Oo and biotinylated O6



Step 4: keep only those that visit each vertex at least once.

- ssDNA is annealed with biotinylated O1



- the complex is immobilized with MAGNET and then the sample is washed several times
  -> only ssDNA with O1 remain
- Process is repeated successively with biotinlylated O2, O3, O4, O5 -> the remaining ssDNA has the right answer

Step 5: Get The Answer

- Dr. Aldeman used “Graduated PCR”

  PCR with primers of O0 and Oi in the path the now the position of vertex i read the length of the PCR product

  example: for vertex 3

  primers: O0 and O3

O0->1	O1->2	O2->3	O3->4	O4->5	O5->6
O0	O1	O2	O3	O4	O5	O6

A band of 80 bp -> 80/20 = 4 th position in the sequence

Step 5: Get The Answer

- Today:

sequencing

Scale up

- It took 7 days to done the whole experiment.
  This Proof-of-concept experiment -> can be scaled up for 100 vertices

1. Ligation of the oligos for the edges and the vertecices
2. PCR with O0 and O99
3. Cut the right band (2000bp)
4. Using biotin-avidin magnetic beads system
5. Sequencing

It would last half a year or even more, still it is faster than normal approach.

Advantages of DNA Computing

- Theoretical higher performance (operation per second)

  one molecule is one operation

  10^{20} molecules operated in a week ≈ $1.65 * 10^{14}$ operation per sec (10^{12})
- Remarkable energy efficiency

  theoretical maximum of $34 * 10^{19}$ operations per joule (10^{9})

Disadvantages of DNA Computings

- Amount of DNA grow exponentially with the complexity of problems

  ->for 200 cities we would require an amount of DNA that weighed more than the Earth.
- Unexpected errors because of unspecific hybridization
- requires human assistance
- DNA degradation

References

- Molecular computation of solutions to combinatorial problems, L. M. Adelman
- http://foldoc.org/index.cgi?query=dna+computing&action=Search
- http://www.promega.com/tbs/tb246/tb246.pdf
- http://www.ams.org/featurecolumn/archive/dna-abc1.html
- www.cse.lehigh.edu/~lopresti/Courses/2003-04/CSE397-497/lecture_22.ppt

Thank you for your attention