

DNA Computing

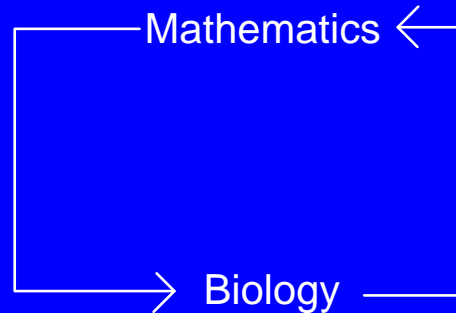
Prof.kamala Krithivasan
Dept of CSE
IIT MADRAS
CHENNAI

Road Map

1. Introduction
2. Motivation
3. About DNA
4. Adleman's experiment
5. Lipton's solution for satisfiability
6. Universality
7. Main considerations

Biological Mathematics

Mathematical Biology



The idea behind biological mathematics is the simple observation that the following two processes one biological and one mathematical are analogous

1. The very complex structure of a living being is the result of applying simple observations (copying, splicing, etc) to initial information encoded in a DNA sequence.

2. The result $f(w)$ of applying a computable function to an argument w can be obtained by applying a combination of simple functions to w .

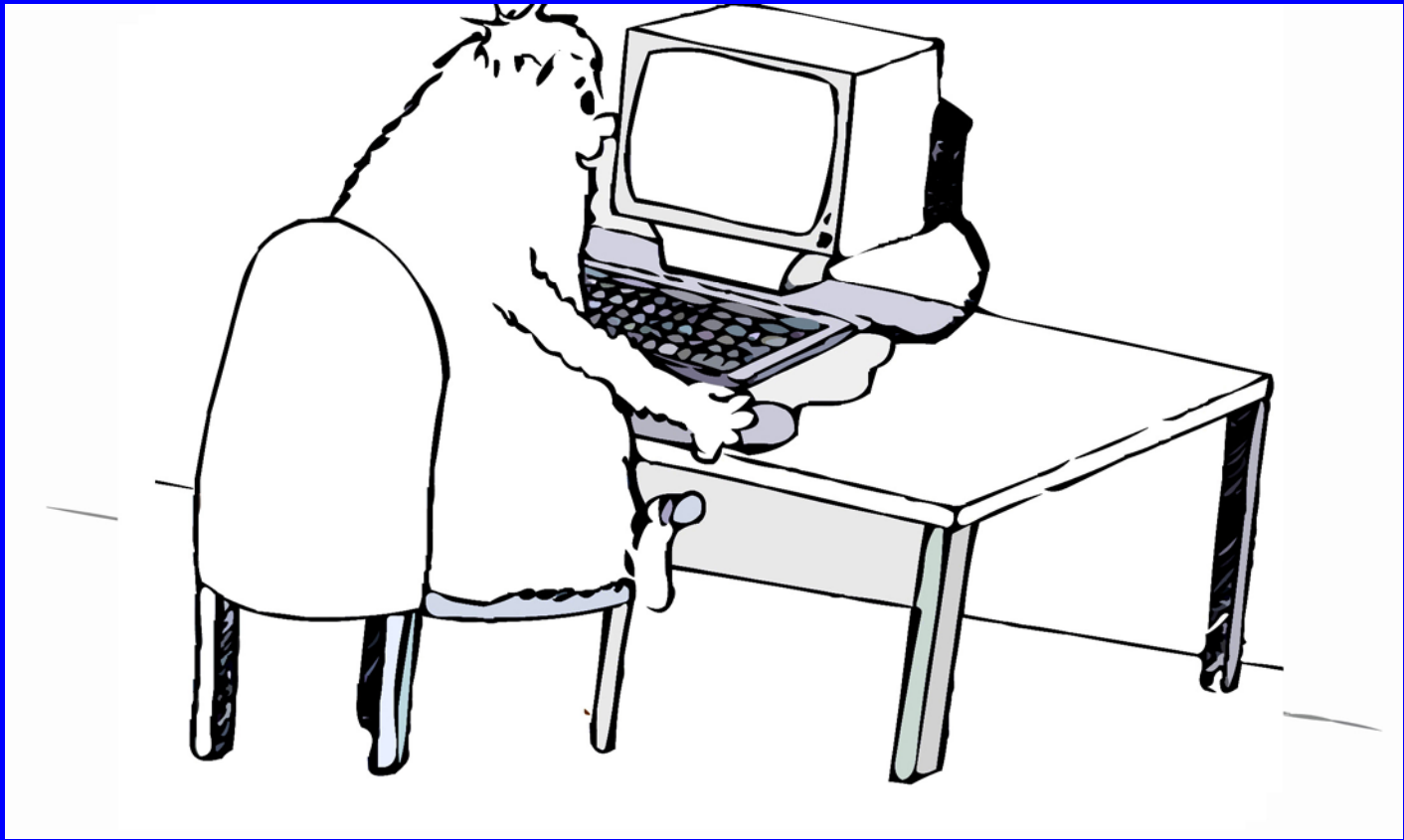
AND, OR, NOT

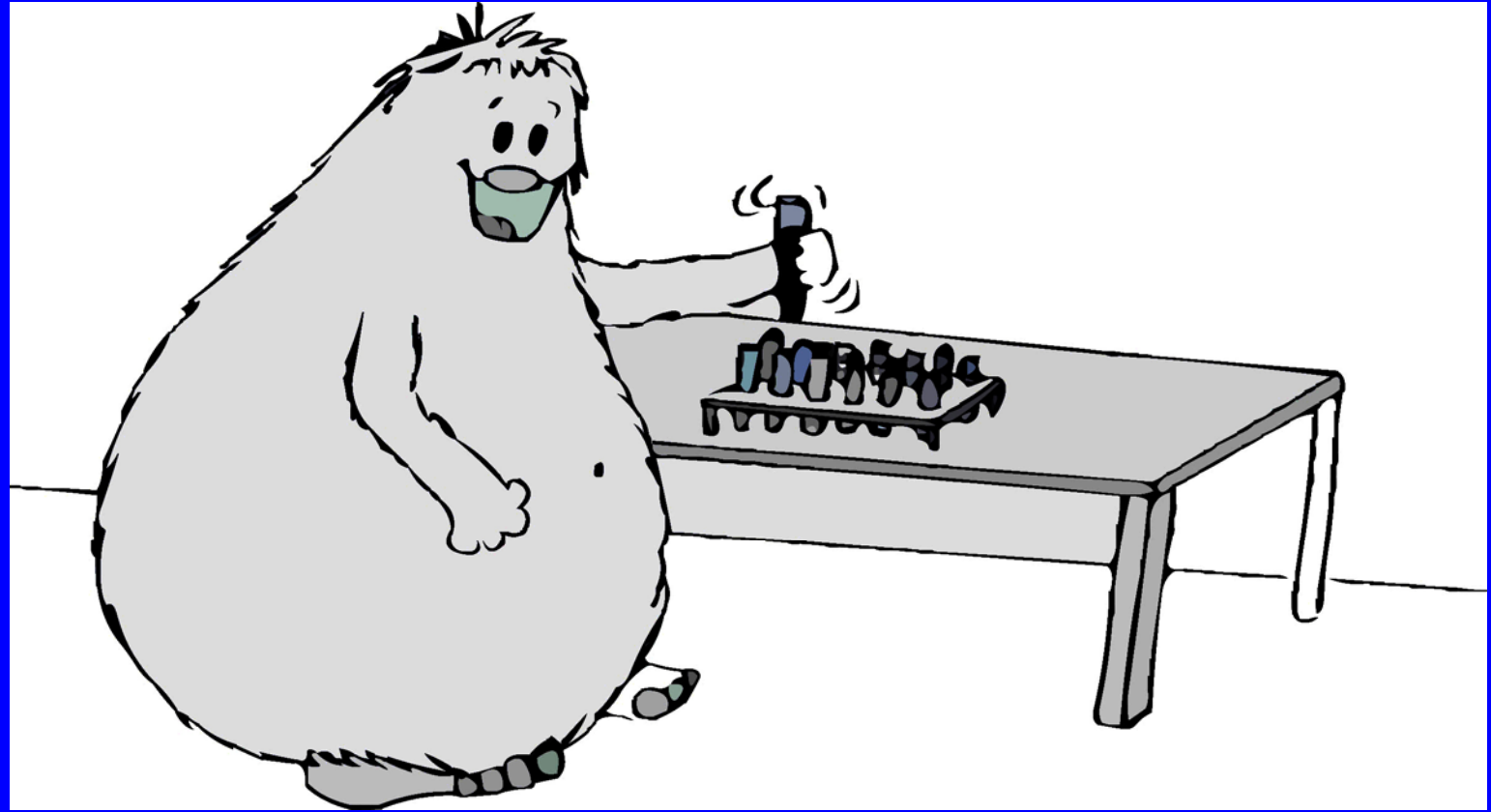
SPLICING

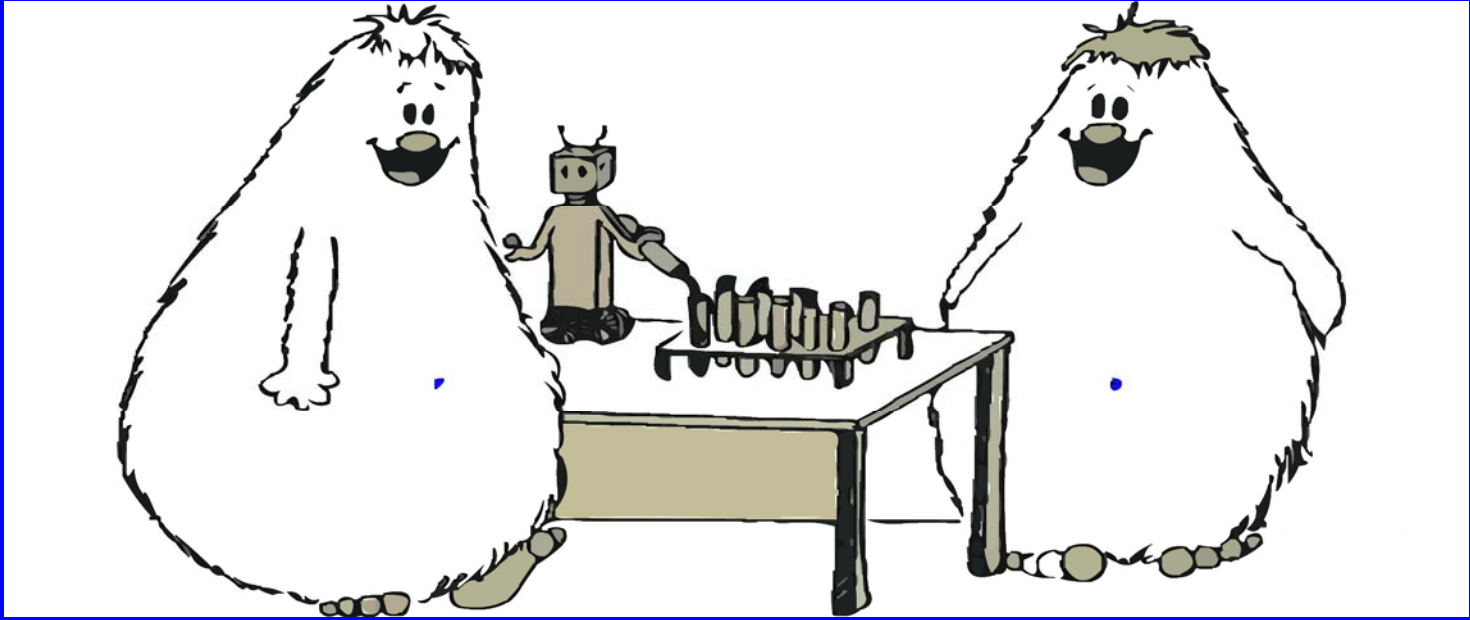
CUTTING AND PASTING

base 10

base 2







DNA Deoxyribonucleohides
Deoxyribo nuclic acid

A sugar phospahte group and consists of nitrogenous base

Purines	adenine	A
	guanine	G

Pyyrimidines	thymine	T
	cytosine	C

RNA – ribonucleic acid

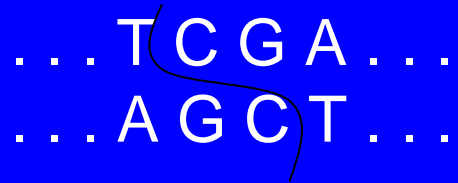
SPLICING SYSTEMS AND DNA COMPUTING

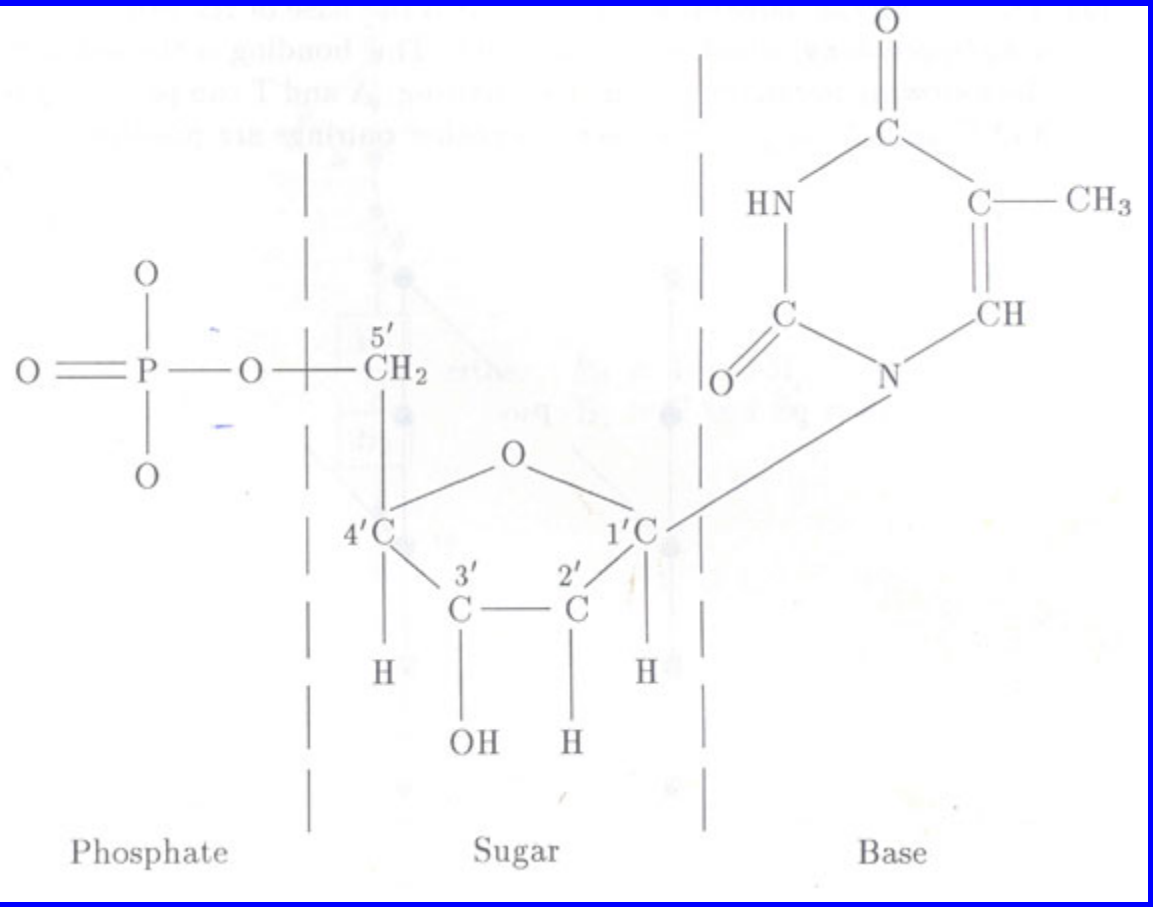
Tom Head 1987

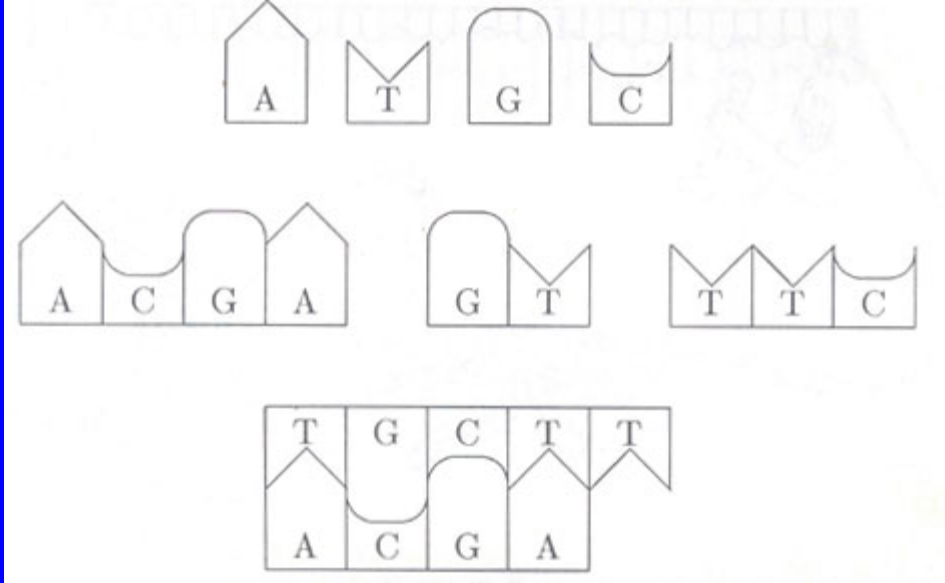
T a q I

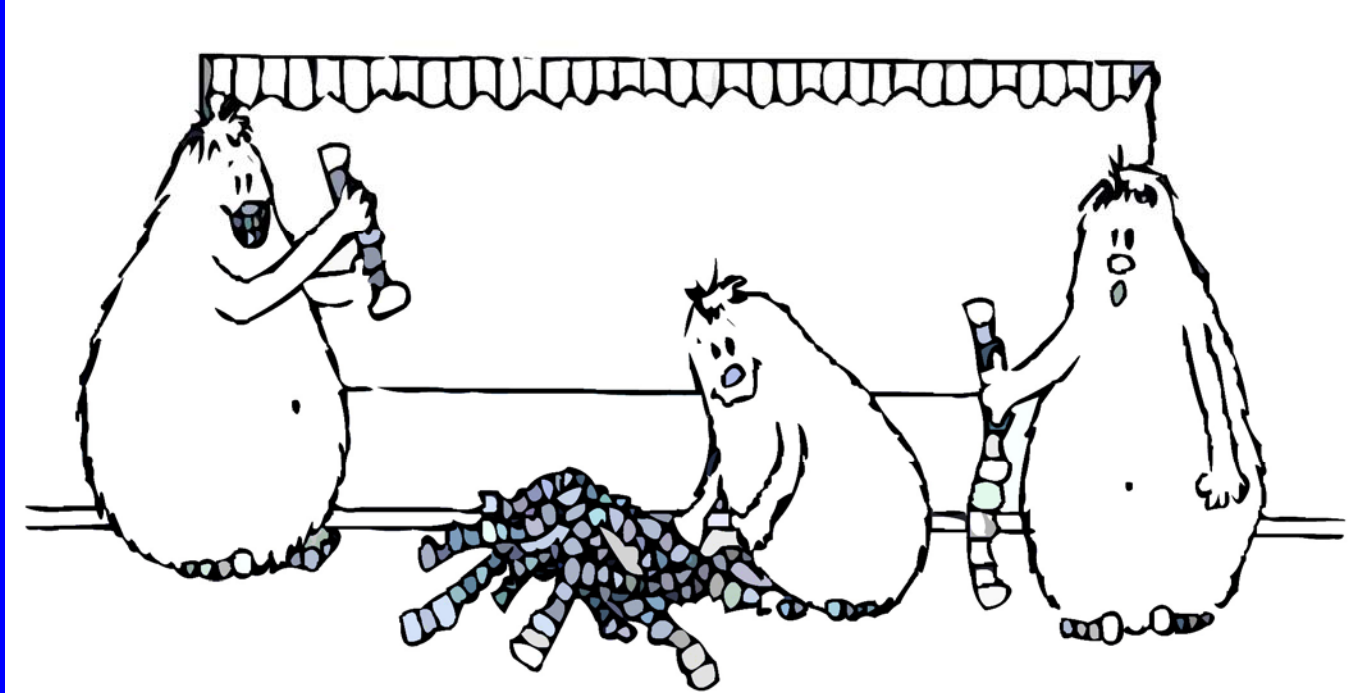
A G C T

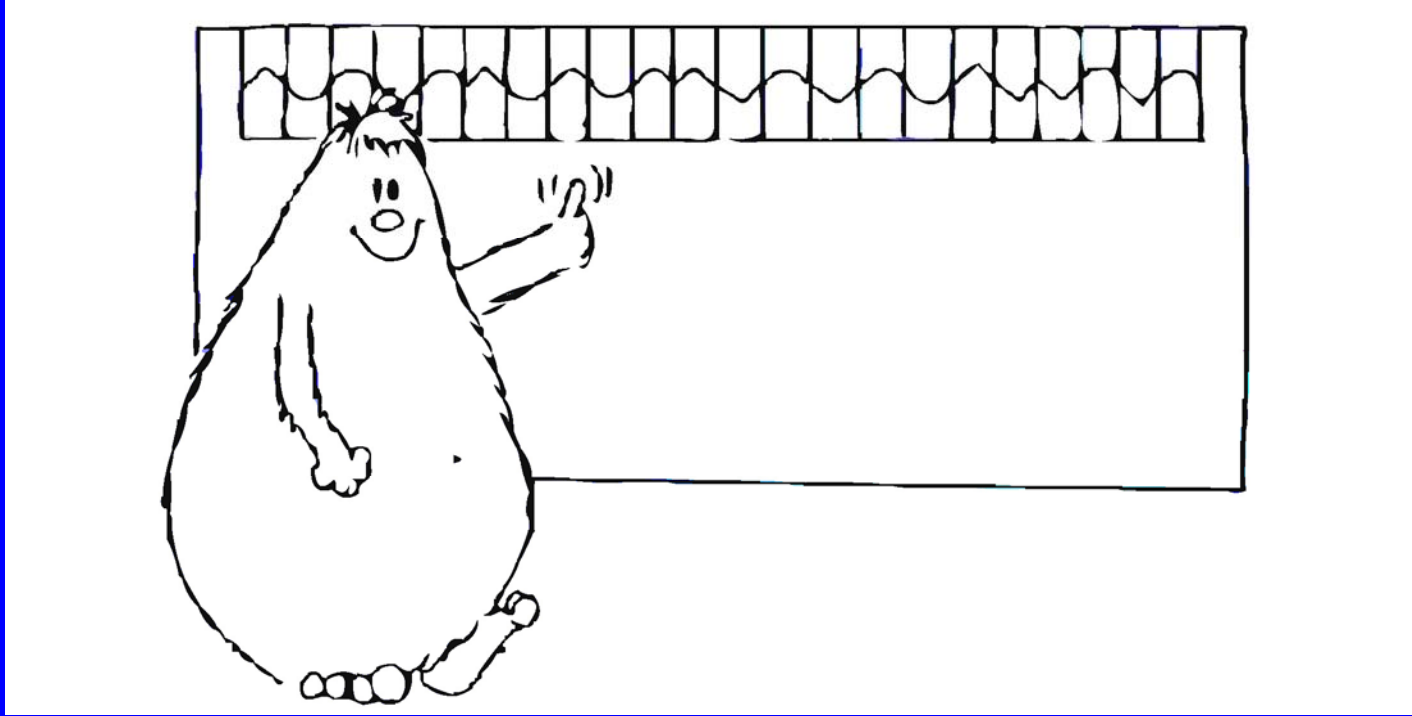
... T C G A ...
... A G C T ...

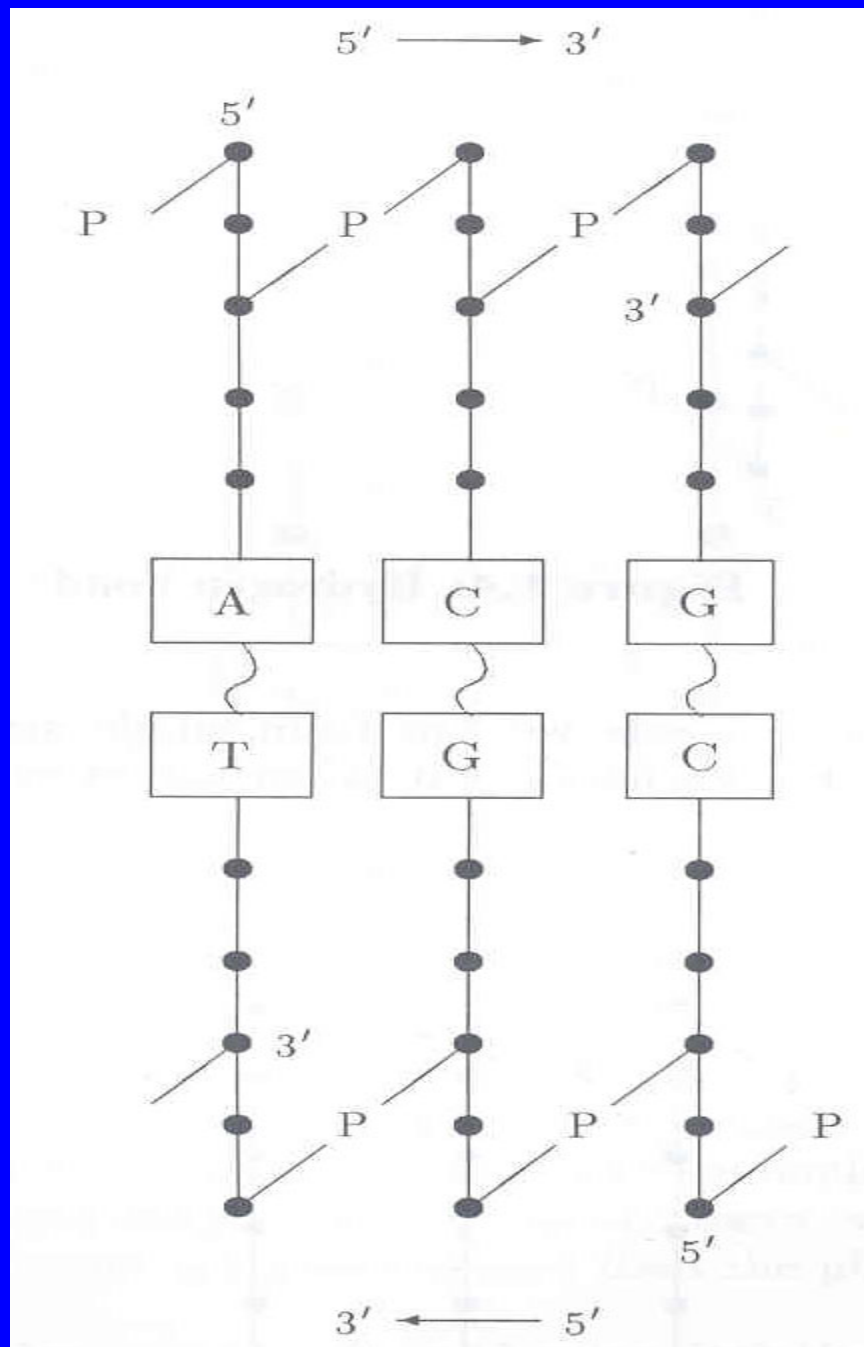








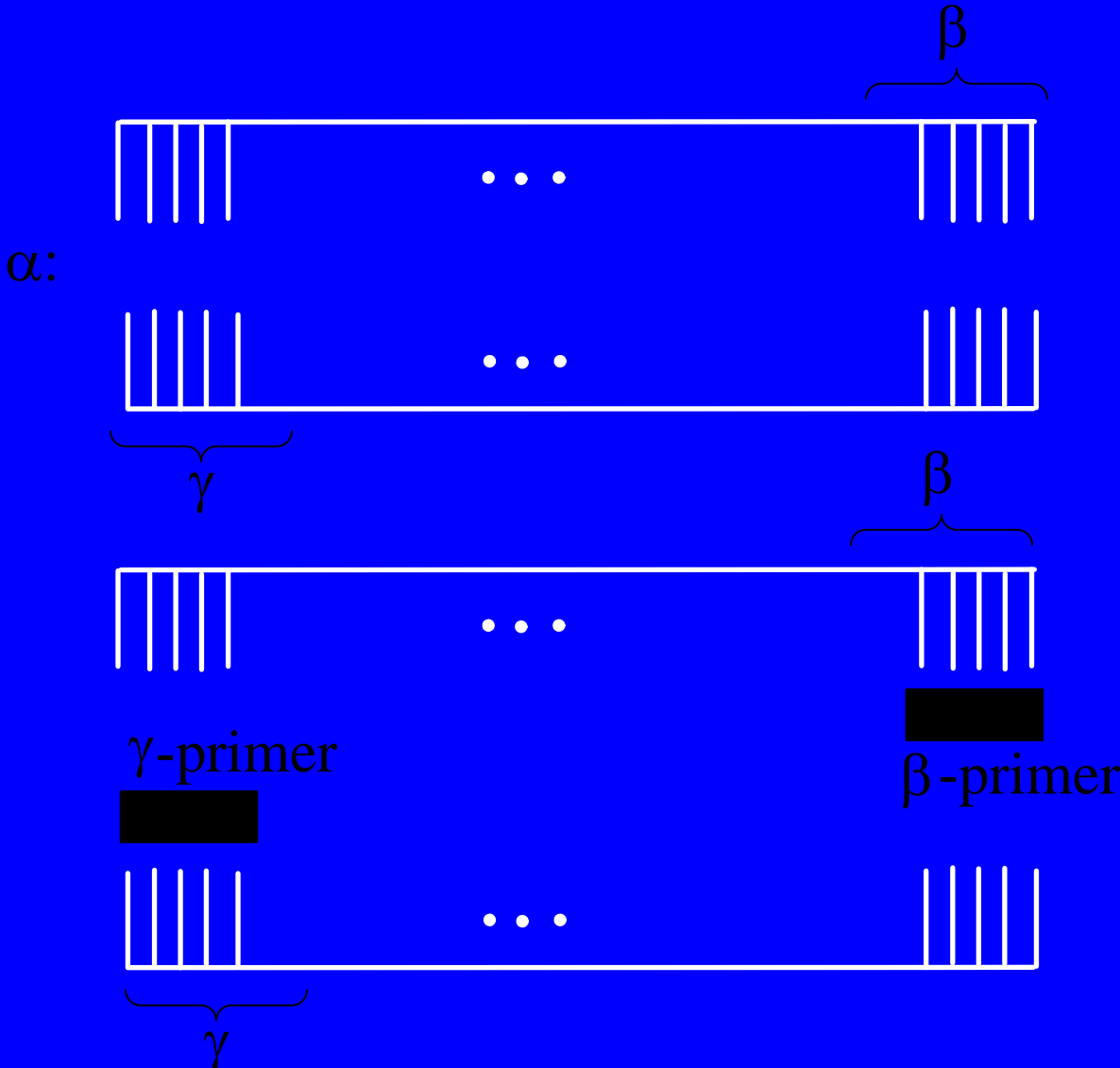




Operations on DNA Molecules

- 1) Separating and fusing DNA strands
- 2) Lengthening DNA
- 3) Shortening DNA
- 4) Cutting DNA
- 5) Linking (pasting) DNA
- 6) Modifying nucleotides of DNA
- 7) Multiplying DNA

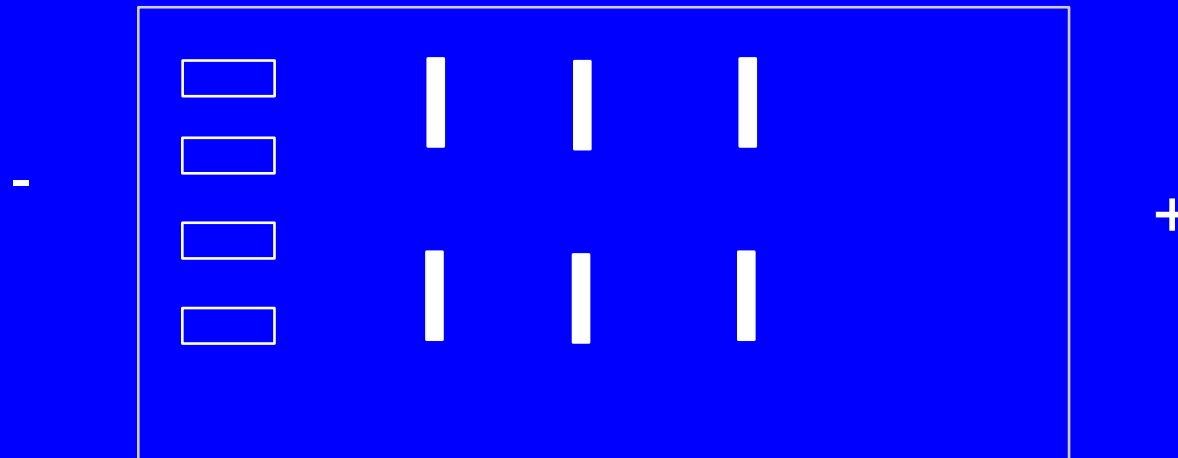
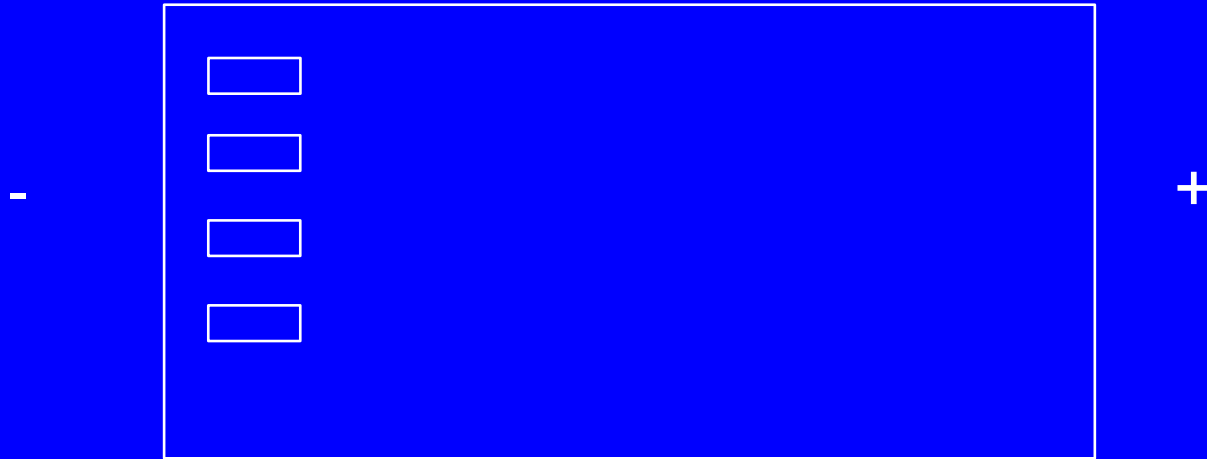
Polymerase Chain Reaction (PCR)



8) Reading out the sequence

Measuring the length of DNA molecules

Gel Electrophoresis



large

small fragments

agarose gel - large > 500 bps

Polyacrylamide gel - smaller

- 1) Staining with ethidium bromide
- 2) Attaching radioactive markers to ends of DNA

Fishing for known molecules

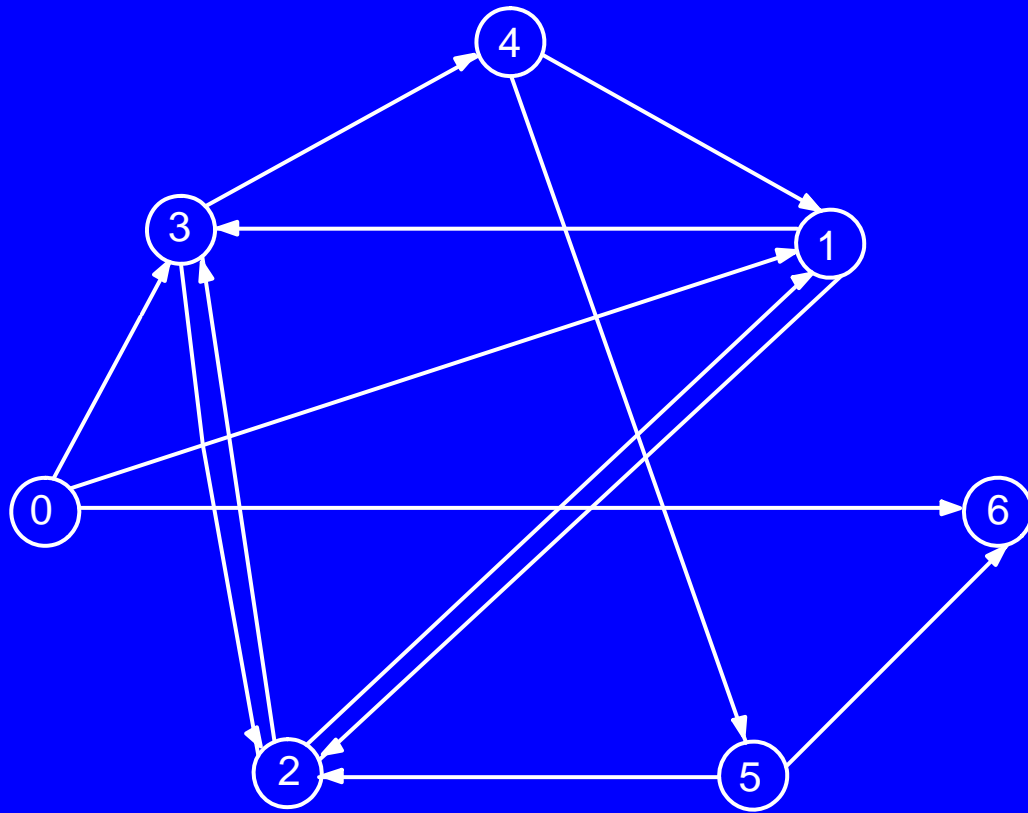
adenine A DNA (deoxyribonucleic acid)
guanine G
cytosine C
thymine T

3' : AAGCTCAG 5

5 : TTCGAGTC 3

$\Sigma = \{ A, G, C, T \}$

The practical possibilities of encoding information in a DNA sequence and of performing simple bio-operations were used by adleman to solve a 7 node instance of the Directed Hamiltonian path problems. A directed graph G with designated vertices v_{in} and v_{out} is said to have a Hamiltonian path if and only if there exists a sequence of Compatible “one way” edges e_1, e_2, \dots, e_n (that is, a path) that begins at v_{in} ends at v_{out} and enters every other vertex exactly once.



The graph in Adleman's experiment

The following algorithm solves the problem

Step 1. Generate random paths through the graph

Step 2. Keep only those paths that begin with v_{in} and end with v_{end} .

Step 3. If the graph has n vertices, then keep only those path that enter exactly n vertices.

Step 4. Keep only those paths that enter all of the vertices of the graph at least once.

Step 5. If any paths remain say “YES”; otherwise say “NO”.

To implement Step 1, each vertex of the graph was encoded into a random 20-nucleotide strand (20-letter sequence) of DNA. Then, for each (oriented) edge of the graph, a DNA sequence was created consisting of the second half of the sequence encoding the source vertex and the first half of the sequence encoding the target vertex. By using complements of the vertices as splints, DNA sequences corresponding to compatible edges were ligated, that is, linked together. Hence, the ligation reaction resulted in the formation of DNA molecules encoding random paths through the graph.

L.M Adleman: Molecular Computation of solutions to combinatorial problems

$s_2 = \text{TATCGGATCGGTATATCCGA},$

$s_3 = \text{GCTATTCGAGCTTAAAGCTA},$

$s_4 = \text{GGCTAGGTACCAGCATGCTT}.$

$e_{2 \rightarrow 3} = \text{CATATAGGCTCGATAAGCTC},$

$e_{3 \rightarrow 2} = \text{GAATTTTCGATATAGCCTAGC},$

$e_{3 \rightarrow 4} = \text{GAATTTTCGATCCGATCCATG}.$

To implement step2, the product of step1 was amplified by Polymerase chain reaction (PCR). Thus, only those molecules encoding paths that begin with vin and end with vend were amplified.

For implementing Step 3 a technique called gel-Electrophoresis was used, that makes possible the separation of DNA strands by length. (The molecules are placed at the top of a wet gel, to which an electric field is applied, drawing them to the bottom. Larger molecules travel more slowly through the gel. After a period, the molecules spread out into distinct bands according to size.)

Step 4 was accomplished by iteratively using a process called affinity purification. This process permits single strands containing a given sequence v (encoding a vertex of the graph) to be filtered out from a heterogeneous pool of other strands. (After synthesizing strands complementary to v and attaching them to magnetic beads, the heterogeneous solution is passed over the beads. Those strands containing v anneal to complementary sequence and are retained. Strands not containing v pass through without being retained.)

To implement Step 5, the presence of a molecule encoding a Hamiltonian Path was checked. This was done by amplifying the result of Step 4 by Polymerase Chain Reaction and then determining the DNA Sequence of the amplified molecules.

UNIVERSALITY

Turing Machine simulation

Beaver's model

Rothemann's model

Theoretical Developments

1. Splicing systems
2. Sticker systems
3. Watson-crick automata
4. Watson-crick L system

Main Considerations

1. Time and cost
2. Errors

Where it will be useful

1. Associative memory
2. DNA2 DNA computation
3. Cryptanalysis

THANKYOU