

DNA Computing: A Complete Overview

Vartika Sharma

NIMS University, Institute of Management and Computer Science, Shobha Nagar, Jaipur Delhi Highway, NH 11C, Jaipur, Rajasthan, India

Abstract: DNA computing is a nascent technology that seeks to capitalize on the enormous informational capacity of DNA, biological molecules that can store huge amounts of information and are able to perform operations similar to computers through the deployment of enzymes, biological catalysts that act like software to execute desired operations i.e. Computers made of genes' building blocks. Because of their speed, miniaturization and Data Storage potential DNA computers are being considered as a replacement for silicon-based computers. Current DNA computer research has already proven that DNA computers are capable of solving complex mathematical equations and storing enormous amounts of data.

Keywords: DNA Computing, DNA Computing techniques, Operations on DNA for Computing, Applications of DNA Computing, DNA Computing as future technology.

1. Introduction

When an important property of DNA is that it can replicate, or make copies of itself. Each strand of DNA in the double helix can serve as a pattern for duplicating the sequence of bases. Therefore every DNA sequence has a natural complement. For example if sequence s is $attacgtcg$, its complement, s' , is $taatgcagc$. Both s and s' will come together (or hybridize) to form double stranded DNA. This complementarity makes DNA a unique data structure for computation and can be exploited in many ways. Error correction is one example. Errors in DNA happen due to many factors. Occasionally, DNA enzymes simply make mistakes, cutting where they shouldn't, or inserting a T for A. DNA can also be damaged by thermal energy and UV energy from the sun. If the error occurs in one of the strands of double stranded DNA, repair enzymes can restore the proper DNA sequence by using the complement strand as a reference. In this sense, double stranded DNA is similar to a RAID 1 array, where data is mirrored on two drives, allowing data to be recovered from the second drive if errors occur on the first. In biological science this process is critical when cells divide because each new cell needs to have an exact copy of the DNA present in the old cell.

To the naked eye, the DNA computer looks like clear water solution in a test tube. There is no mechanical device. A trillion bio-molecular devices could fit into a single drop of water. Instead of showing up on a computer screen, results are analyzed using a technique that allows scientists to see the length of the DNA output molecule. If you want to present the output to the naked eye, human manipulation is needed. For this reason DNA computers are spoken to by the phrase "computer in a drop of water". DNA strands produce billions of potential answers simultaneously. This makes the DNA computer suitable for solving "fuzzy logic" problems that have many possible solutions rather than the either/or logic of binary computers.

2. Classification of Problem

2.1 Finite State Problem

To compete with silicon, it is important to develop the capability of biomolecular computation to quickly execute

basic operations, such as arithmetic and boolean operations, that are executed in single steps by conventional machines. In addition, these basic operations should be executable in massively parallel fashion. Guarnieri and Bancroft developed a DNA-based addition algorithm employing successive primer extension reactions to implement the carries and the boolean logic required in binary addition (similar methods can be used for subtraction). Guarnieri, Fliss, and Bancroft prototyped the first biomolecular computation addition operations (on single bits) in recombinant DNA. They presented the development of a DNA-based algorithm for addition. The DNA representation of two non-negative binary numbers was presented in a form permitting a chain of primer extension reactions to carry out the addition operation. They demonstrated the feasibility of this algorithm through executing biochemically a simple example. However, it suffered from some limitations: (1) only two numbers were added, so it did not take advantage of the massive parallel processing capabilities of biomolecular computation; and (2) the outputs were encoded distinctly from the inputs, hence it did not allow for repetitive operations. Subsequent proposed method for basic operations such as arithmetic (addition and subtraction) allow chaining of the output of these operations into the inputs to supplementary operations, and to allow operations to be executed in massive parallel fashion. Rubin presented an experimental demonstration of a biomolecular computation method for chained integer arithmetic.[1]

2.2 Combinational Problems

The DNA computing methods were employed in complex computational problems such as the Hamilton path problem (HPP), satisfiability problem (SAT) and the like. The advantage of these approaches is the huge parallelism inherent in DNA-based computing, which has the potential to yield vast speedups over conventional electronic based computers for such search problems. For example, to solve SAT problem a DNA computer take 91 steps where as a sequential computer take around 1.6 million step.[1]

The computational problem considered by Adleman was a simple instant of the directed traveling salesmen problem (TSP) also called Hamilton path problem (HPP). Adleman's experiment represents a landmark demonstration of data

processing and communication on the level of biological molecules. It was the first DNA computer set up to solve the TSP.

Lipton argued that all NP (non-deterministic polynomial time) problems could be efficiently reduced to the HPP. He also demonstrated how DNA computing solves a two variable SAT problem. Lipton proposed a solution to the SAT. The proposed DNA algorithm can solve an n -variable m -clause SAT problem in m steps. Unlike conventional DNA computing algorithms needing an initial data pool to cover correct and incorrect answers and further executing a series of separation procedures to destroy the unwanted ones. Instead of generating the full-solution DNA library, we start with an empty test tube and then generate solutions that partially satisfy the SAT formula. These partial solutions are then extended step by step satisfying one clause in one step there by solving the entire problem and finding complete solution in steps using ligation process. If we score and sort the clauses, we can use this algorithm to markedly reduce the number of DNA strands required throughout the computing process. This algorithm is highly space-efficient and error-tolerant compared to conventional brute-force searching, and thus can be scaled-up to solve large and hard SAT problems. The algorithm that he proposed for two variables can be explained using a simple example. If x and y are binary variables which can take the values 0 or 1, a simple SAT problem would be to find all values for x and y that make the logical statement $(x=1 \text{ or } y=0)$ and $(x=0 \text{ or } y=1)$ true. In this case, $x=1, y=1$ would be a solution. The heart of Lipton's method is a process he calls extraction. When a problem has only two variables, x and y , the possible solutions are represented by the numbers 00, 01, 10, and 11—where, for instance, 01 represents $x=0, y=1$. An extraction step could remove from the test tube all DNA corresponding to strings with, say, a 1 in the second position. Such an extraction would pull out the strands encoding 01 and 11, leaving behind strands for 00 and 10.

To solve the first clause of the above SAT problem, one would first extract all $x=1$ strands (10 and 11), then extract all $y=0$ strands (10 and 00), and mix them in a new test tube. From this batch of DNA one would solve the second clause by extracting all $x=0$ strands (00), and $y=1$ strands (11). These represent the answers to the problem. [2]

3. Classes of DNA Computing Techniques of Problem

3.1 Intramolecular

To The intramolecular DNA computing operates by means of intramolecular conformational transitions. For example the Japanese project lead by Hagiya (2005) focuses on constructing programmable state machines in single DNA molecules, which using the concept of intramolecular DNA computing. It is described by a single-stranded DNA molecule of the form stopper $state^1 state^1$ stopper $state^2 state^2 \dots stopper state^n state^n$, where, in each pair (state i -state i) of states, state i denotes the state before a transition, and state i the state after the transition.[1] Each state is represented by an appropriate number of bases, called a state sequence. This process of state transitions can be

repeated in a single tube by a simple thermal program consisting of thermal cycles for denaturation, annealing, and polymerization. The state machine DNA is assumed to form a hairpin, and transitions occur in an intramolecular manner.[2]

3.2 Intermolecular

To The intermolecular DNA computing focusing on the hybridization between different DNA molecules as a basic step of computations. An example of intermolecular DNA computing is the Adleman's experiment in which he presented the solution of the HPP problem. The method represented each of the cities using different bases where by these different molecules are hybridized to trace out the possible paths from starting city to ending city. The rejection strategy is then used to find out the potential paths and the final result.[2][3]

3.2 Supermolecular

The supramolecular DNA computing, as pioneered by Winfree (2003), harnesses the process of self assembly of rigid DNA molecules with different sequences to perform computations. The self-assembly of small molecular building blocks programmed to form larger, nanometer-sized elements is an important goal of this approach. This approach is motivated by the magnificent examples occurring in nature: for instance, the supramolecular complex of the *E. Coli* ribosome consisting of 52 protein and three RNA molecules.[2] Innovation and application of supramolecular assemblies have reached impressive new heights. For example, organizations involving nucleic acids have been used for drugs or DNA delivery, and can also be efficient as sensors for detection purposes. The interactions of various low-molecular weight substances with DNA are also used in medicinal treatment.[3]

4. DNA Operations for Computing

4.1 Watson-Crick Based Pairing

According to the prior discussion, a single-stranded fragment has a phosphosugar backbone and four kinds of bases denoted by the symbols A, T, G, and C for the bases adenine, thymine, guanine, and cytosine respectively. These four nucleic acids, which can occur in any order in a single stranded DNA, paired in Watson-crick complementarity pairs to form a double strand helix of DNA. Due to the Watson-crick complementarity, A is paired with T by 2 hydrogen bonds, whereas C is paired with G by 3 hydrogen bonds, as shown in figure 1. Many pairs can form a hydrogen-bond, which is a weak bond and the point of attachment to the backbones are equally spaced, as shown in figure 6, to allow a regular helical structure. The helical structure is shown in figure 2.[5][6]

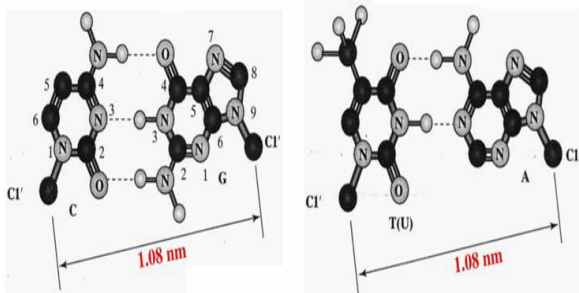


Figure 1: T-C Base Pair and A-T Base Pair

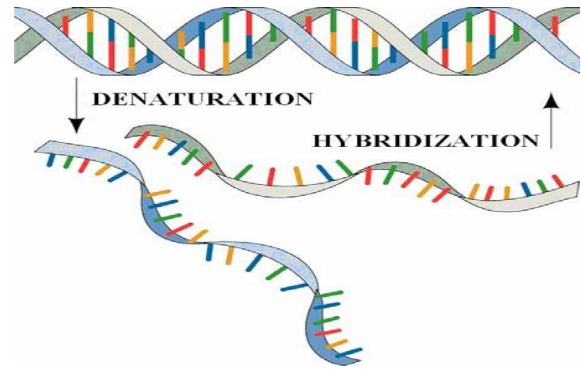


Figure 3: Bi-molecular Hybridization and Denaturation

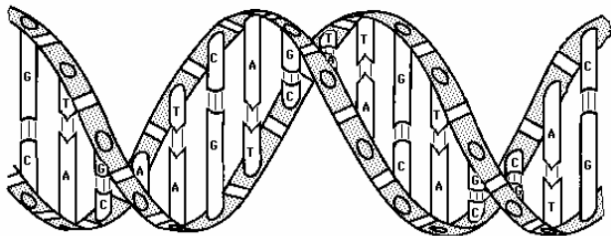


Figure 2: Helical Structure pairing of dsDNA

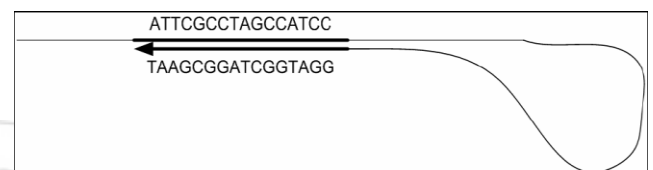


Figure 4: Hairpin Formation of DNA

4.2 Hybridization and Denaturation

Hybridization is defined as a sequence-specific annealing of two or more single stranded DNAs, forming a dsDNA product. This sequence-recognition property is very useful for DNA computing because hybridization means computation, from DNA computing sense. This operation is normally caused by cooling down the test tube reaction solution. There are basically three cases, on how the hybridization could occur: bimolecular hybridization, multi-molecular hybridization, and uni-molecular hybridization. For the first case, bi-molecular hybridization involves two kinds of ssDNAs to form a native double helix structure of DNA as shown in figure 3. For the multi-molecular hybridization, three strands are involved during the annealing. Multi-molecular hybridization is the essence of Adleman DNA computing for solving an instance of Hamiltonian path problem.[3] Thirdly, uni-molecular hybridization or self-hybridization, hairpin formation of ssDNAs could be formed if a complementary portion exists in the same ssDNAs, as depicted in figure 4.

By heating up the solution to about 85-95°C, dsDNAs will come apart because the hydrogen bonds between complementary nucleotides are much weaker than the covalent bonds between nucleotides adjacent in the two strands. The separation is called melting or denaturation. Thus, two strands can be separated without breaking the single strands, as depicted in figure 3. The same effect can be achieved by washing the double stranded DNAs in doubly distilled water.[3][5]

4.3 Ligation

Ligation is often invoked after the single DNA strands are annealed according to the Watson-crick complementarity. Many single-strand fragments will be connected in series and ligase is used as 'glue' to seal the covalent bonds between the adjacent fragments. Figure 5 shows the principle of ligation. In this figure, three kinds of ssDNAs, namely strand A, strand B, and strand C, taking part during the ligation. Strand A and strand B must be located adjacently with each other without gap, and hybridized partially with strand C. The product of ligation is a 'new' strand AB. Ligation is generally implemented in laboratory via a DNA ligase, such as T4 DNA ligase. Moreover, strand A must have a 5' PO₄ and energy required during the ligation is ATP or NAD⁺. [5]

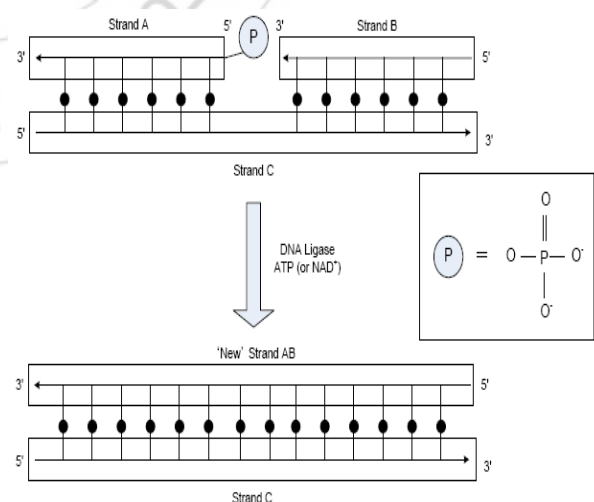


Figure 5: Ligation

4.4 Polymerization

The substrates required for polymerization are a template strand to be copied, a primer strand to be 3'-extended, incoming dNTP monomers, which act as both base and

energy sources, and DNA polymerase. DNA polymerase implements a 5' to 3' copying operation as depicted in figure 6. During the copying operation, 3' end of a primer strand is extended. Note that there is no 3' to 5' copying operation ever observed. This operation also depends on Watson-crick complementarity. In other words, A is copied to T and T is copied to A, and so on.[4]

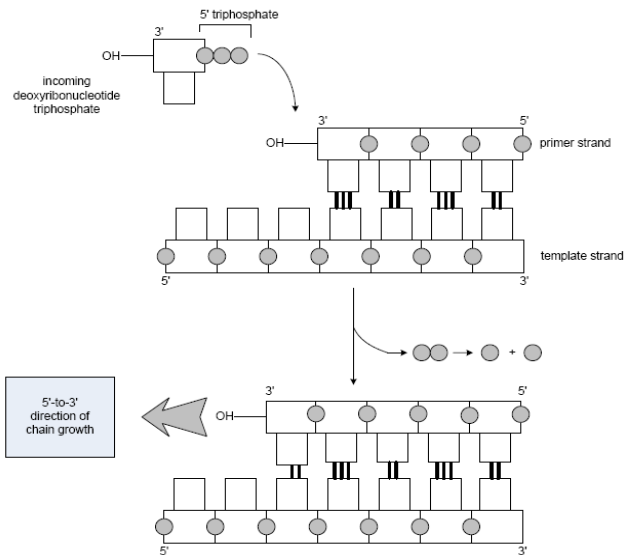


Figure 6: Polymerization in Action

4.5 Polymerase Chain Reaction (PCR)

PCR is an incredible sensitive copying machine for DNA. It also can be used for DNA detection. Given a site-specific single molecule DNA, a million or even billion of similar molecules can be created by PCR process. In n steps, it can produce 2^n copies of the same molecules. PCR needs a number of sub-sequence strands called 'primers', which are usually about 20 base long to signal a specific start and end site at a template for replication. PCR normally runs for 20-30 cycles of 3 phases: separating base pair strands of DNA at about 95°C, annealing at 55°C, and extension at 74°C. It takes about two to three hours normally in order to complete the cycles. Figure 7 shows the operations of polymerase chain reaction up to third cycles.[3][4]

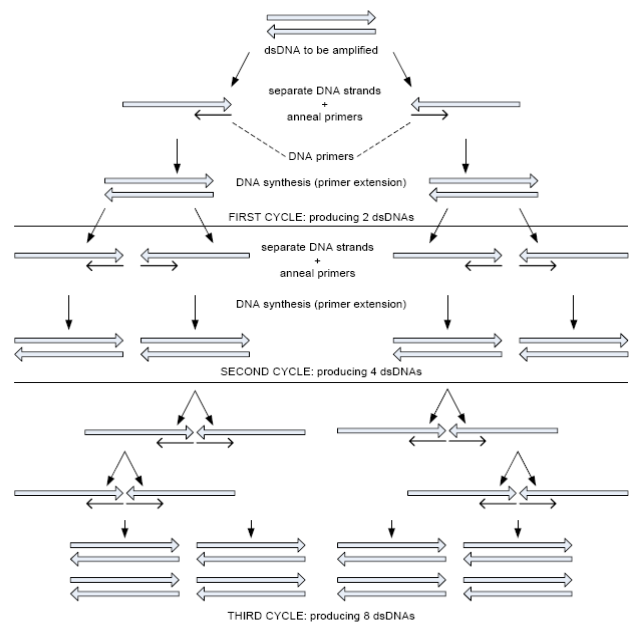


Figure 7: Polymerization in Action

4.6 Result Visualization: Gel Electrophoresis

DNA strands in a solution can be separated in terms of its length by means of gel electrophoresis. In fact, the molecules are separated according to their weight, which is almost proportional to their length. This technique relies on the fact that DNA molecules are negatively charged. Hence, by putting them in an electric field, they will move towards the positive electrode at different speed. If electrical field is applied through the gel, longer molecules will remain behind the shorter ones, as shown in figure 8. The speed of DNA mixture in a gel depends heavily on the gel porosity and the magnitude of the electrical field. Polyacrylamide gel is used for separation of shorter DNAs which is from 10bps until 500 bps. On the other hand agarose is frequently used for longer DNAs, which is more than 500bps. An example of the output of gel electrophoresis is well depicted in figure 9. In DNA computing, this technique is used to visualize the results of computation. Normally, at the end of this process, the gel is photographed for convenience.[3][5]

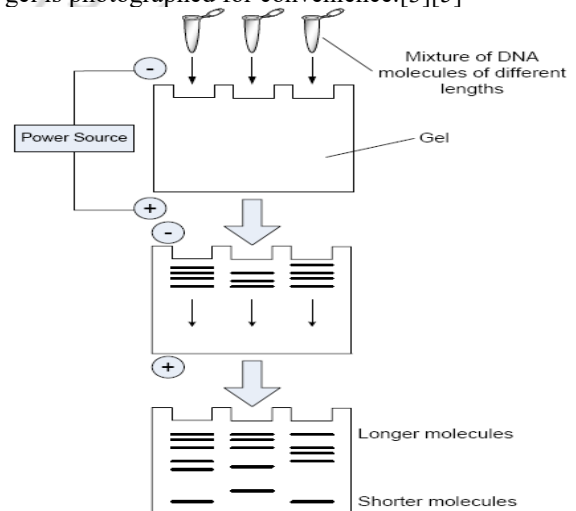


Figure 8: Molecule Behavior in electric field

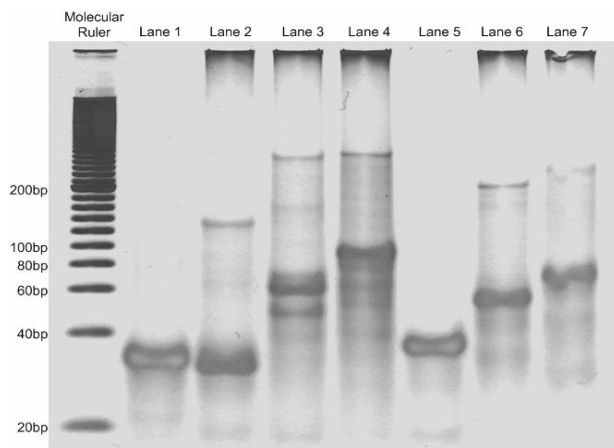


Figure 9: Example of Gel Image

4.7 DNA Extraction

Extraction is a process to filter out desired strands from a solution. This process basically helps in finding out the final result from possible solutions obtained. There are a number of different techniques used for the purpose. Some emphasized methods are described here

- **Using probes made of complementary DNA:** To perform an extraction in practice, Lipton proposed using probes made of complementary DNA. If, for example, a strand of the sequence gggtta is to be extracted out, Lipton would make its complementary sequence cccat and attach copies of it to magnetic particles. When these probes were poured into the test tube, they would latch onto the DNA strands which are represented by the sequence gggtta. Then, with a magnet, Lipton would extract the magnetic particles together with the probes and long DNA strands. In another test tube he would peel off the probes to leave the required DNA strands.[2][3]
- **Using complementary strands within solution:** Adleman and his team have found a more powerful technique for making DNA extraction where partial matching is required and the extracted strands are to be used in further computations. This method is highly used in performing digital computations efficiently. In it, Adleman starts with a large number of identical strands of DNA, each of which is thousands of bases long. By itself, one of these "memory strands" represents a long string of 0s. To put a 1 in particular positions, Adleman uses "stickers"—short strands of DNA that are complementary to the memory strands in those positions and so naturally stick there.[2]

Adleman's sticker method can match Lipton's technique, and do much more. In particular, Adleman can modify the computer's memory during the course of a calculation. Where Lipton's strands are unalterable, Adleman can change his memory strands in the middle of a series of calculations by adding stickers to every strand in the test tube at the same place on each strand.[3] This gives him a capability similar to that of silicon computers, allowing him to update his calculations. The sticker computer also avoids the need for enzymes since

none of the calculation steps involves making copies of the strands, cutting the strands, or linking two strands end to end.

- **Using complementary strands on Surface:** At the University of Wisconsin, for instance, anne condon and colleagues propose attaching strands of DNA to a surface and manipulating them there instead of letting them float free in a liquid. The idea is to avoid losing DNA molecules during the extraction steps—a likely problem for Lipton's and Adleman's designs. On the other hand, without being able to extract strands of DNA, the Wisconsin team has to rely on a more limited set of computing operations.[3][4]
- **Arrangement of strands in 2-D bricks:** At Caltech, Erik Winfree has abandoned DNA strands altogether in favour of two-dimensional "bricks" of DNA, in which several strands of DNA are combined to create structures with four corners. Winfree, a student of Adleman's, designed the bricks so they would fit together much as the DNA strands did in Adleman's original work, except that the self-assembly would be in two dimensions instead of one. The bricks assemble themselves in an inverted triangle—one in the first row, two in the second, three in the third, and so on. The answer would be read off from the sequence of the last row of bricks after the triangle is complete. Winfree is also working in three dimensions to try and assemble pyramids.[3]

4.8 Operations in Parallel

In the cell, DNA is modified biochemically by a variety of enzymes, which are tiny protein machines that read and process DNA according to nature's design. There is a wide variety and number of these "operational" proteins, which manipulate DNA on the molecular level. For example, there are enzymes that cut DNA and enzymes that paste it back together. Other enzymes function as copiers and others as repair units. Molecular biology, biochemistry, and biotechnology have developed techniques that allow us to perform many of these cellular functions in the test tube. It's this cellular machinery, along with some synthetic chemistry, that makes up the palette of operations available for computation. Just like a CPU has a basic suite of operations like addition, bit-shifting, logical operators (and, or, not nor), etc. That allow it to perform even the most complex calculations, DNA has cutting, copying, pasting, repairing, and many others. And note that in the test tube; enzymes do not function sequentially, working on one DNA at a time. Rather, many copies of the enzyme can work on many DNA molecules simultaneously. This is the power of DNA computing, that it can work in a massively parallel fashion.[2][4]

5. Challenges to DNA Computing

- All the data has to be represented in terms of DNA sequences.
- Practical protocols for input and output of data into the memory are unavailable.
- An understanding of the information capacity of the hybridization interactions in large collections of many DNA sequences is required to check information flow.
- Controlling DNA operations so that they are performed

in a required manner to produce desired results is complicated.

- Appropriate physical models to guide designs and experimentations are yet to be developed.

6. Applications of DNA Computing

6.1 Solving NP-complete and hard computational problems

To After Adleman's and Lipton's initial results, much of the work on DNA computing has continued to focus on solving NP-complete and other hard computational problems. Problems in NP-complete are such that there is no polynomial time solution known to exist using conventional computer algorithms. That is, as the complexity of these problems increase, the time required to solve them increases at an exponential rate. These problems are also said to be *intractable*, but, if within the domain of NP-complete a *tractable* solution can be found to one of these problems then it can also be used to solve for all other problems in the set. The HPP and sat problems are examples which have already been discussed.[7]

Apart from these technical problems some real life problems are also solved including most of the business planning and management science problems. Many of the cost optimization problems faced by managers are in NP-complete are currently solved using heuristic methods and other approximations. These problems include scheduling, routing, and optimal use of raw materials and correspond to problems already solved, either theoretically or experimentally, using DNA computation. Current and near-term advances in laboratory technology preclude the use of DNA computation as a method of solving problems in real time. This restricts the possible application of classical DNA computing to areas where the calculation of optimal solutions could be performed over a number of days, weeks, or months.[7] This would be able to be applied to long term production planning where initial cost commitments are very high such as chip design and manufacturing. DNA computing could also be applied extensively to optimizing airline and bus routes for planning purposes.

6.2 Storage and associative memory

DNA might also be used to mirror, and even improve upon, the associative capabilities of the human brain. Baum proposed a method for making a large content addressable memory using DNA. A truly content addressable memory occurs when a data entry can be directly retrieved from storage by entering an input that most closely resembles it over other entries in memory. This input may be very incomplete, with a number of wildcards, and in an associative memory might even contain bits that do not actually occur within the closest match. In Baum's model storing a word could be done by assigning a specific DNA subsequence to each component value pair and building a fixed length word from these subsequences. To then retrieve the word closest to the input, one would introduce marked complementary subsequences into the storage medium and chose the molecule that has the most matches to this input.[9] This technique could be further refined to more

closely approximate the brain by appending words to only store attributes that an object has, rather than wasting space using '0's to represent attributes that an object does not have.

Baum has further speculated that a memory could be constructed where only portions of the data are content-addressable and associative, with other information on an object compactly stored in addresses relative to the associative portion of the entry.

For example considering the brain's limit of about 10^{15} synapses and Feynman's low-end estimate that the brain can distinguish about 10^6 concepts, such a DNA based associative memory could have certain advantages over the brain. Without accounting for redundant molecules, Baum estimates that a large bath tub of DNA, about 50g in 1000l, could hold over 10^{20} words.[9]

6.3 DNA2DNA Applications

Another area of DNA computation exists where conventional computers clearly have no current capacity to compete. DNA2DNA computations involve the use of DNA computers to perform operations on unknown pieces of DNA without having to sequence them first. This is achieved by re-coding and amplifying unknown strands into a redundant form so that they can be operated on according to techniques similar to those used in the sticker model of DNA computation. The potential applications of DNA2DNA computation are:

- DNA Sequencing: The term DNA sequencing encompasses biochemical methods for determining the order of the nucleotide bases, adenine, guanine, cytosine, and thymine, in a DNA oligonucleotide. The sequence of DNA constitutes the heritable information that forms the basis for the developmental programs of all living organisms. Determining the DNA sequence is therefore useful in basic research studying fundamental biological processes, as well as in applied fields such as diagnostic or forensic research. Methods of DNA computing can improve the methods of DNA sequencing by motivating the use of magnetic beads, optical scanners and the like that may allow DNA to be read directly into electronic interface.[10][11][12]
- DNA Fingerprinting: DNA sequencing however provides a method for distinguishing between living organisms, but because there are millions of base pairs present, the task is very time consuming. Instead, scientists have a shorter method called the DNA fingerprinting which utilizes the repeating patterns in DNA. These patterns do not give individual fingerprints, but they are able to determine whether two DNA samples are from the same person, related people, or non-related people. Scientists use a small number of sequences of DNA that are known to vary among individuals, and analyze those to get a certain probability of a match. This approach is also widely used in determining paternity and maternity, criminal identification and forensics and for personal identification.[13]
- DNA Mutation and Population Screening: Environmental factor like ultraviolet light (sunshine), nuclear radiations and certain chemicals and the process of separation and

coping of DNA strands can cause permanent changes in DNA sequence in a gene which is called DNA mutation. Mutations in a gene's DNA sequence can alter the amino acid sequence of the protein encoded by the gene. The DNA sequence is interpreted in groups of three nucleotide bases, called codons. Each codon specifies a single amino acid in a protein. This alteration causes many genetic diseases. These diseases must be detected during infancy or childhood in order to stop them from being hazardous. For this purpose researches are conducted where by a particular segment of population is screened and preliminary tests which examine the DNA samples from dry blood spots are carried out over people. For those who found infected, diagnostic tests are conducted. DNA computing can lend a hand for such researches on larger scale.[13][14]

6.4 Use in nanocomputing

DNA combines with nanotechnology to offer us an enormous range of benefits. If you think about the structure of DNA, it consists of a double-stranded molecule. This molecule can unzip to form somewhat of a branched molecule but these branches can also be woven into structures that can self-assemble. Through the use of a double-stranded molecule of DNA, scientists can unzip the molecule partway and the branches will join with complementary sequences on the other DNA that similarly branches out.[15] With the right guidance, the DNA molecules can be joined in a way that they provide valuable structures that support many different applications. A variety of DNA nano mechanical devices have been previously constructed that demonstrate motions such as open/close, extension/contraction, and motors/rotation mediated by external environmental changes such as the addition and removal of DNA fuel strands or the change of ionic composition of the solution. For example, the DNA walker could ultimately be used to carry out computations and to precisely transport nano particles of material.[8]

6.5 Intelligent systems based on DNA Computing

In 2004 Langdon and Buxton used genetic programming in combination with a principled objective function to evolve many non linear functions of gene expression values. The approach was to whittle down the thousands of data attributes (gene expression measurements) into a few predictive ones. Intelligent DNA memory has been designed by Chen as an attempt to capture global information about a population of an organism. Furthermore, the DNA memory incorporates intelligent processing and reasoning capabilities into the test tube. After the data gathering and analysis stage is complete, the high storage capacity and parallelism of DNA are used to draw inferences on the entire *in vitro* knowledge base.

Sakakibara and Suyama proposed DNA chips with logical operations called intelligent DNA chips. They combined the DNA computing method for representing and evaluating boolean functions with the DNA coded number (DCN) method, and implemented DNA chips with logical operations executable. The developed DNA chips are considered intelligent because the DNA chips not only detected gene

expression but also found logical formulae of gene expressions. This intelligent DNA would be able to provide logical inference for such diagnoses based on detected gene expression patterns.[7]

7. Promotes to DNA Computing

- **Size:** The information density could go up to $1\text{bit}/\text{nm}^3$. i.e. DNA can hold more information in a cubic centimeter than a trillion CDs.
- **High Parallelism:** Every molecule in DNA solution could act as a small processor on nano-scale and the number of such processors per volume would be potentially enormous each of which operates at the same time on the inputs to produce simultaneous outputs from which complete solution is drawn. In an *in vitro* assay we could handle easily with about 10^{18} processors working in parallel.
- **Computational Power:** In future a droplet sized DNA computer will have more computational power than today's most powerful super computer.
- **Speed:** Although the elementary operations would be slow compared to electronic computers, their parallelism would strongly prevail, so that in certain models the number of operations per second could be of order 10^{18} operations per second, which is at least 100,000 times faster than the fastest supercomputers existing today.
- **Energy Efficiency:** 10^{19} operations per Joule. This is about a billion times more energy efficient than today's electronic devices.
- **Generate complete solution set:** for problems which are supposed to have a large number of possible solutions, a DNA computer is capable of producing the complete set of the solutions which is unattainable with conventional computers.

8. Issues with DNA Computing

- **Uncertainty in Results:** The importance of practical implementation and the potential for errors in a DNA computation have been realized from the very beginning. It has been pointed out that as the size of the problem grows, particular attention must be paid to errors and the formation of "pseudo paths", i.e., DNA strands which appear to be valid solutions, but actually are false positives. Errors can be introduced at any number of steps in a DNA computation. Care must be taken that the starting material (oligonucleotides, enzymes, etc...) Is of good quality and free of contaminants. All of these effects are produced by the fundamental chemical nature of a DNA computer, and the sensitive dependence of the computational results on the chemical conditions such as concentration of different species, and the environment, especially temperature, are critical.
- **Scalability:** to solve the problems of significant size a large volume of DNA is required. For example, it has been estimated that Adleman's Hamiltonian path problem, if enhanced to 50 or 100 cities, would require tons of DNA, it increases exponentially. Large volume of DNA requires large codes to be created for nucleotides and also make the encoding error prone.
- **Human Interaction:** every step of the computation

requires human intervention. Although, processes like hybridization, ligation etc are involuntary DNA operations, formation of complements, combination of bases, polymerization, extraction are all required to be controlled and monitored.

- **Solution Generation and Result Extraction:** for every element in every problem (e.g. Variables and clauses in satisfiability problem) to be solved however small it may be, DNA solutions are required to be generated having nucleotides along with their complements. This may not be efficient for small problems. Also DNA computer take biological inputs and produce output in the same form. Conversions are required to input the problem parameters to get its result and there by to interpret the result.

9. DNA Computers V/S Conventional Computers

Table 1: Margin specifications[16]

Basics	DNA Computers	Conventional Computers
Size	Smallest of any technology till date	Excessively outsized as compared to DNA computers.
Storage Media	Nucleic acids	Semiconductors
Memory Capacity	Ultra High	High
Operators	Biochemical operations	Logical Operations (AND, OR, NOT)
Operations	Simultaneous (Parallel)	Bitwise (Sequential)
Speed of each operations	Slow	Fast
Process	Stochastic	Deterministic
Availability	Abundant Supply as that of cellular	Silicon is less available compared to DNA
Fabrication	Biochip can be made more	Toxic material is used

10. DNA Computers V/S Quantum Computers

A quantum computer is a device for computation that makes direct use of distinctively quantum mechanical phenomena a quantum computer is a device for computation that makes direct use of distinctively quantum mechanical phenomena to perform operations on data. The principle of quantum mechanics is that the quantum properties can used to represent and structure data and that quantum mechanics, dealing with mechanical systems at atomic level, can be devised and built to perform operations on these data.

On one side where quantum computing uses physical technology for the isolation of mixed quantum states necessary to implement efficient computations solving complex problems, DNA computing operates in natural noisy environments, such as a glass of water. Using DNA

computing, embedded computing is possible without electrical power in microscopic environment. Because DNA computing is linked to molecular construction, the computations may eventually also be employed to build self-organizing partially electronic quantum computers.[16][17]

11. DNA Computers V/S Parallel Computers and Nano Computers

DNA computing provides large scale parallelism of computation. It differs greatly from parallel computing on the basis of the limits of scalability. Parallel computers cannot be scaled more than few 1000's where as DNA computers provide trillions of operations to be performed simultaneously. It also liberated from the problems such as increase in traffic, maintaining synchronization etc. faced in various architectures of parallel computing.

Nano computers are acknowledged for their reduced size. More technically it is a computer whose fundamental parts are not bigger than a few nanometers. This is undoubtedly a remarkable reduction in the size of computers, but using DNA even more compact computers can be constructed. It is a revolutionary change in the computers when size is considered. DNA computing provides enormous parallelism in one fiftieth of a teaspoon of solution more specifically; more than 10 trillion DNA molecules can fit a space no larger than a cubic centimeter.

12. Wrapping it all

While the discovery of DNA has been a significant one in the twentieth century, it will continue to revolutionize medicine, agriculture, forensics, paternity and many other important fields in society today. DNA research encompasses an evolving area of progress and continued funding and interest in its relevance will likely fuel new discoveries in the future.

Lloyd Smith and his colleagues at the University of Wisconsin-Madison caution that DNA computing is still in its infancy and a long way from challenging chip technology. Only about a dozen centers around the world are working on DNA computing and much more needs to be done. "The interest in the community of computer people is more as exploration or new models for computing than in actual practical consequences in the near term," Smith said.

However, researches are going on for incorporating DNA computing in various fields such as in agriculture, weather forecast, designing transport solutions, breaking complicated and difficult security codes, developing DNA optical storage medium and the like.[18]

DNA computing continues to be an exciting concept, delicately placed at the brink of science and biology. Not only has it become a promising technology for analyzing data but it also shows ability and power to transmit information in nanotechnology and other interesting applications. With continued research and development, DNA computing will hopefully overcome its current

challenges, paving the way for efficient and successful computing applications in a wide variety of fields.

References

- [1] G paun, —Recent Computability Models Inspired from Biology: DNA and Membrane Computing”, THEORIA, An International Journal for Theory, 2008.
- [2] Max Garzon, Vinhthuy Phan, Sujoy Roy, Andrew Neel, —In search of optimal codes for DNA computing”, Springer-Verlag Berlin Heidelberg 2006.
- [3] Ren-Hou Li, Wen Yu, —An Exploration of the principles of DNA Computation”, Chinese J. Computers 24(9), 972-978 (2001) Other: Genetic Programming 1 (Standford, 1997), Genetic Programming 2 (Wisconsin-Madison, 1998), GECCO-1999, IEEE International Conference on Evolutionary Computation (Indianapolis, 1997)
- [4] M. H. Garzon, R. J. Deaton, —Biomolecular Computing and Programming, — IEEE Transactions on Evolutionary Computation, vol. 3, pp. 236-250, 1999
- [5] G. Paun (ed.), —Computing with Biomolecules: Theory and Experiment”, Springer-Verlag, Singapore 1998
- [6] Misako Aida, —Characteristics of the Watson-Crick hydrogen bonded DNA base pair: An ab initio molecular orbital study.” Journal of Computational Chemistry, Volume 9, issue4, June 1988, Pages 362-368
- [7] Swarnendu Mukherjee, Debashis Ganguly, Swarnendu Bhattacharya, Parth Mukherjee, —A cognitive Study on DNA Based Computation”, International Journal of Recent Trends in Engineering, Vol. 1, No. 2, May 2009
- [8] Ezziiane Z.: DNA computing: applications and challenges”, Nanotechnology 17, R27-R39 (2006)
- [9] Rohani binti abu Bakar, Junzo Watada, —DNA Computing and its Applications: Survey”, ICICI International 2008 ISSN 1881-803X Springer volume 2 number 1, March 2008
- [10] Wanq JT, Rozen S, Shapiro BA, Shasha D, Wanq Z, Yin M, —New techniques for DNA Sequence classification”, J Comput Biol, 1999 Summer;6(2):209-18
- [11] Linian T.C. Franca, Emanuel Carnlho, Tarso B.L. Kist, —A review of DNA Sequencing techniques”, Quarterly Reviews of Biophysics 35, 2 (2002), pp. 169-200. —002 Cambridge University Press DOI : 10.1017/S0033583502003797 Printed in the United Kingdom
- [12] Jay Shendure, Erez Lieberman Aiden, —The expanding scope of DNA sequencing”, Volume 30 Number 11 November 2012 Nature Biotechnology
- [13] G Cui, C Li, H Li, —DNA Computing and its application to information security field”, Natural Computation, 2009 . ICNC'09., 2009
- [14] Harry Vrieling, Albert A. Van Zeeland, —Molecular Analysis of Mutations in Endogenous Genes”, Methods to access DNA Damage and Repair Interspecies Comparisons, © 1994 Published by John Wiley & sons Ltd. (Structure and Sequence)
- [15] John H Reif, Thomas H LaBean, —DNA Nanotechnology and its Biological Applications”, Chapter 13 of Block: Bio-inspired and Nano-scale

Integrated Computing(Edited by Mary Eshaghian), Wiley Publication, USA (2007)

- [16] Shazia Hassan, Humaira, Mamoona Asghar, —Imitations of silicon based computers and future prospects”, 2010 second international conference on communication software and network, 978-0-7695-3961-4/10 \$26.00 © 2010 IEEE DOI 10.1109/ICCSN.2010.81
- [17] Pedram Khalili Amiri, —Quantum Computers” redefined logic, 0278-6648/02/\$17.00© 2002 IEEE, IEEE Potentials, December 2002/January 2003
- [18] Reza M. Zadegan, Michael L. Norton, —Structural DNA Nanotechnology : From Design to Application”, International Journal of Molecular Sciences, ISSN 1422-0067, *Int. J. Mol. Sci.* **2012**, *13*, 7149-7162; doi:10.3390/jims13067149

Author Profile

Vartika Sharma received the B.Sc. (CS), MCA and M.Tech. (SE) degrees from JNVU, RTU and Pacific University in 2006, 2009 and 2015 respectively. Internship for the completion of master of computer application is done at ISRO, Jodhpur, Rajasthan. Worked as Assistant Lecture at JIET, Jodhpur, Rajasthan for one and a half year, as Post graduate teacher in Kendriya Vidyalaya Sangathan for three years and three months. Presently working as Assistant Professor at NIMS University Jaipur Rajasthan, India. Area of interests other than DNA computing are Medical Imaging and Software Engineering.