

Molecular Computing and Molecular Programming

Masami Hagiya

*JST CREST and Department of Computer Science,
Graduate School of Information Science and Technology,
University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, JAPAN.
E-mail: hagiya@is.s.u-tokyo.ac.jp*

Abstract

In this talk, I first describe the directions of current research in the field of DNA and molecular computing by summarizing a recent international conference in this field: *DNA8, the Eighth International Meeting on DNA Based Computers*. Research in this field has clearly shifted from purely mathematical computations to broader applications in nanotechnology and biotechnology, and the principles and methods for designing molecular systems with information-processing capability for such applications are considered important. We call research into designing such molecular systems *molecular programming*. In the talk, I review existing models of DNA and molecular computation and analyze the results of these models, and then briefly describe some methods for molecular programming, including sequence design. I finally touch on molecular machines made of DNA, the current focus of molecular programming.

Introduction

Several years have passed since the direction of research in the field of DNA and molecular computing shifted from purely mathematical computation to much broader applications in nanotechnology and biotechnology. This means that principles and methods developed in the field are now being applied to non-mathematical problems, such as constructing molecular machines and analyzing the human genome. A few years ago, while developing his DNA computer for solving 3-SAT problems, Suyama insisted that if a DNA computer could solve 3-SAT problems accurately, then the computer could be used to analyze the human genome and produce reliable results that could be used for medical diagnosis.

Another possible direction in research is towards applying evolutionary computation to molecular evolution. There was once interest in performing evolutionary computation with molecules, but now it is thought that applying methods established in evolutionary computation, such as the genetic algorithm or programming, to molecular evolution will be more fruitful. Since evolution can be regarded as a kind of computation, this direction is considered to be yet another application of molecular computing in a broad sense.

DNA8

These new directions in the field were also topics at a recent international conference on DNA and molecular computing, i.e., *DNA8, the Eighth International Meeting on DNA Based Computers* [1], which was held in Sapporo, Japan, on June 10-13, 2002.

- * Classical Adleman-style DNA computing seems to be disappearing, while applications to nanotechnology and biotechnology are becoming central issues.
- * The field of DNA nanotechnology treats both the construction and motion of structures.
- * Even sequence design is geared towards DNA nanotechnology, as secondary structure design is becoming an important issue.
- * Applications in genomic analysis are also becoming realistic. For example, Suyama proposed using his DNA computer for gene expression analysis and SNP analysis. Cell engineering is not a dream any more. Weiss *et al.* are trying to engineer cells to make sensors. This is a movement towards *in vivo* molecular computing.
- * In the past, DNA computing flourished as a result of theoretical contributions related to the splicing system (or H-system) and membrane system (or P-system). There were some interesting and important talks in the theory session of DNA8. Nevertheless, theoretical contributions seem stalled, and the theory of DNA computing needs breakthroughs.

Molecular Computing

Research in this field is increasingly being directed towards developing

principles and methods for designing molecules and molecular systems that not only solve purely computational problems, but also the broader problems mentioned above. To solve such problems, a computing perspective is still important, because some kind of information processing is always involved. For example, molecular machines should have information processing capability at the molecular level. Analyzing human genomes is nothing but processing information on genes and their products. Any molecular system that is capable of information processing is considered a subject of this field.

In order to construct such molecular systems, we need to have a deep understanding of what molecules can compute. This is exactly the question that researchers in the field of DNA and molecular computing have been hoping to answer, and many models of DNA and molecular computation have been proposed. In this talk, I very briefly summarize the models used for molecular computation and the results of analyses of the computational power of the models, including computability and complexity.

I first touch on the Adleman-Lipton paradigm, and two of its refinements, Suyama's DNA computer and Sakamoto's SAT Engine. I also explain the computational paradigm based on self-assembly, proposed by Seeman and Winfree, with exciting results on its computability.

Molecular Programming

The increasing power of these models is allowing the field to expand as they are applied to broader applications. This requires research on design and construction. I use the term *molecular programming* to include research into designing molecules and molecular systems with information processing capabilities. The word *programming* suggests that designing molecules and molecular systems is like programming electronic computers. A typical example of molecular programming is designing DNA nucleotide sequences. Sequence design is one of the most important aspects of DNA computing, because it greatly influences the accuracy and efficiency of DNA computation. By hybridization, DNA molecules interact with each other and form complex structures or even machines encoded in their sequences. Moreover, the sequences of molecules with enzymatic functions, such as RNA and proteins, control their own reactions. Designing the sequences of such molecules is simply programming their behavior.

Needless to say, programming molecules and molecular systems

involves many complex computational problems that are themselves good research subjects in computer science. In this talk, I briefly describe the problem of designing huge sequence sets, and that of designing DNA sequences that fold into given secondary structures (inverse folding).

Finally, I touch on molecular machines made of DNA, which are currently the main target of molecular programming: Seeman's DNA motor by B-Z transition, Hagiya's whiplash PCR, Yurke's molecular tweezer, Seeman's PX-JX2 switch, and Shapiro's DNA automata.

Concluding Remarks

In Japan, a research group on molecular programming funded as *Priority Area Research* by the Ministry of Education has recently been formed. In addition, a CREST project has been funded by the Japan Science and Technology Corporation to research *molecular addressing* for constructing molecular memory out of DNA and related molecules. It is also intended to explore new directions in molecular computing, including *optical molecular computing*. The preparation of this talk was supported by both of these projects.

The full paper of this abstract can be obtained from [2].

References

[1] Masami Hagiya and Azuma Ohuchi, eds.: *Preliminary Proceedings of the Eighth International Meeting on DNA Based Computers*, June 10--13, 2002, Hokkaido University, 2002.

<http://hagi.is.s.u-tokyo.ac.jp/dna8/>

To appear as *Lecture Notes in Computer Science*, Vol.2568, Springer.

[2] Masami Hagiya: Towards Molecular Programming, *available at*

<http://hagi.is.s.u-tokyo.ac.jp/members/hagiya.html>