Review of DNA Based Computers III
Editors: Harvey Rubin and David Harlan Wood
Publisher: American Mathematical Society
Series: DIMACS Series in Discrete Mathematics and Theoretical Computer Science Volume 48
ISBN: 0-8218-0842-7
375 pages
Hardcover, $79

Martyn Amos*
Department of Computer Science, University of Liverpool, Liverpool L69 7ZF, England

The concept of computing at a molecular level back to Feynman in the 1960s, and was developed by Conrad and others in the 1980s. Recently, there has been a surge of interest in the field as a result of Adleman's 1994 paper, which showed how molecular level computation may be achieved using standard laboratory operations on strands of DNA.

Since Adleman's paper, DNA computing has developed into a rapidly growing area of enquiry; several major conferences now dedicate special sessions to the field, and the Annual Workshop on DNA Computing held its fifth meeting this year, at MIT. This volume, DNA Based Computers III, contains papers presented at the earlier Third Annual Workshop, held in June 1997 at the University of Pennslyvania. The twenty-five contributions in this volume were collected by the editors, Harvey Rubin, a medical doctor at the University of Pennslyvania, and David Harlan Wood, a Computer Scientist at the University of Delaware. The diverse backgrounds of the editors are a testament to the truly inter-disciplinary nature of the field.

Several "threads" dominating the workshop are represented by various papers in this volume, although this fact is not at all clear on initial inspec-

*E-mail: martyn@csc.liv.ac.uk

tion of the contents page. However, the preface clarifies matters somewhat, and the different sessions are clearly delineated.

The *Second* Annual Workshop was characterised by a general feeling that the field is an exciting one to be in, but that a lot more experimental work needs to be done in order to establish DNA Computing as a viable discipline in the long-term. Although we are still some way away from this position, the volume under consideration describes encouraging movement towards it. Over one-third of the papers describe the results of (sometimes preliminary) laboratory experiments, and a casual flick through the book reveals a healthy number of gel visualisations! The "Making It Work in the Lab" sessions addressed the "real-world" problems inherent to experiments in DNA computing. Often, these problems are non-obvious, due to the unusual nature of the protocols used or the level of accuracy required. The paper by Wetmur (p. 1-24) provides a useful review of the nature of DNA hybridization, and describes various approaches to and advantages of limiting experimental conditions in order to minimize the occurrence of "undesirable" molecular complexes. The paper by Hartemink and Gifford (p. 25-37) also focuses on DNA hybridization, this time presenting a a simulation system for optimizing the design of DNA strands used as computational components. As the authors rightly state, "...it becomes clear that many of the proposed computational systems will need to move 'from the blackboard into the laboratory' in order to iron out the subtle details associated with their implementation", and this system may well prove to be an invaluable tool in achieving this. The MIT group of Khodor and Gifford then reports a series of experimental observations on the efficiency of DNA extraction operations (p. 39-46), a central task in any DNA computation. They conclude that the community should adopt a standard set of benchmarks for key primitive operations, such as separation. On a similar note, Chen and Harlan Wood propose in their paper (p. 47-56) a new separation technique with a low error-rate. Their method uses circular DNA strands, rather than the more common linear variant, and the authors claim the main benefits of their approach to be simplicity and improved performance over traditional, affinity-based methods. As the use of circular DNA strands has increased since the 1997 meeting, this method should certainly be reviewed again by practitioners in the field. The paper by Hagiya *et al.* (p. 57-72) then describes the tantalising prospect of "single-pot" evaluation of Boolean formulae. As the number of manipulations of the contents of a tube increases, so does the potential for error, so massively-parallel computations performed without human intervention may well be one characteristic of the

elusive "killer application" of DNA computing.

The "Proposed Applications" session begins with the description of a proposed "killer app", suggested by Lipton, Landweber and Rabin (p. 161-172). The idea involves operating on "unknown" sequences of DNA, a technique which may have applications in DNA sequencing and fingerprinting. Although the "DNA^{2}DNA" method may prove beneficial in these areas, its relevance to general computational issues is questionable. The paper by Ellington *et al.* (p. 173-184) then makes the extremely important point that nucleic acids can act not only as passive information storage media, but as functional computational components. Ellington *et al.* describe the exciting possibility of nucleic acids such as ribozymes acting as transistors. This "non-passive" view of nucleic acids is currently being championed by several prominent researchers in the field, and this paper is essential reading for anyone working in or considering research in DNA computing.

The highlight of the "Proposed Algorithmic Approaches" session is a paper by Fraenkel (p. 101-122), in which he discusses the theoretical relationship between the protein folding problem (an incredibly significant problem in molecular biology) and NP-complete problems. He relates the problem of finding the final conformation of a protein (given only its amino acid sequence) to that of finding the ground state of spin glasses. This work implies that it may be possible to synthesise proteins to solve specific instances of the spin glass problem.

The "Models of DNA Computers" session opens with a description of "Local Parallel Biomolecular Computation" by John Reif (p. 217-254). Containing an extensive and extremely useful bibliography, this paper advocates a move away from "distributed parallelism" in DNA computation, where operations are performed in parallel on large numbers of different molecules, towards "local parallelism", where operations are executed in parallel on *each* given molecule. Reif describes methods that draw on developments in nano-fabrication as one possible route towards large-scale assembly of molecular complexes to solve specific problems. Ogihara and Ray (p. 255-264) then describe one possible method for reducing the amount of DNA required to solve NP-complete problems, using a technique known as DNA *counting*. Although interesting, this paper appears rather dated in the light of recent compelling arguments that the main benefit of using DNA computation is unlikely to derive from the solution of NP-complete problems. The paper by Blumberg (p. 265-279) then addresses the important issue of communication in parallel computation, with particular reference to DNA computation, and presents a direct DNA simulation of the Connection Ma-

chine. Conrad and Zauner (p. 281-287) then describe a model of a DNA conformational processor, which utilizes the ability of DNA to switch from one form to another, based on the presence or absence of specific chemicals. These molecules could, potentially, be evolved to perform certain pattern recognition functions.

The final "Computability Using DNA" session, as its title suggests, relates "traditional" computability theory to the area of molecular computing. The paper by Freund, Paun, Rozenberg and Salomaa (p. 297-327) describes the notion of "Watson-Crick Complementarity", which guarantees universal computation in any model of DNA computation that provides sufficient input/output capabilities. Kari *et al.* (p. 329-346) then present several characterizations of recursively enumerable languages using insertion/deletion systems. Insertion and deletion of "words" of DNA is a very common biological phenomenon, and these operations can be used as the sole primitives in DNA computations.

As the field of DNA computing is moving so rapidly, some of these papers may have lost some of their immediate relevance. Although hardly constituting part of the DNA computing "fossil record", this book will be invaluable to researchers in the field, as it contains the various stages of evolution of current ideas. However, it may well prove to be a daunting read to the new student of DNA computing, and the fields equivalent of Goldberg's "Genetic Algorithms in Search, Optimization and Machine Learning" still remains to be written. To conclude, it is perhaps worth noting that the Sixth Annual Workshop on DNA Based Computers (previously always held in the United States) will be hosted by Leiden University, The Netherlands, in the year 2000.