The University of Western Ontario

Computer Science 881a, Fall 1996

TOPICS IN COMPUTING AND BIOLOGY: DNA COMPUTING

General Information

Description

In 1994 Adleman astonished the scientific community when he solved an instance of the Directed Hamiltonian Path Problem in a test tube, just by handling DNA strands. Lipton followed by showing how to solve the Satisfiability Problem and other NP-complete problems by DNA computing.

Two fundamental questions, formulated by Adleman and Lipton, were left open:

(1) Can **every** algorithm/program be simulated by a DNA algorithm?

(2) Is it possible, to design **programmable DNA computers**, i.e., "test tube computers" able to run any arbitrarily given program?

In this course we investigate the above problems by studying several computational models for DNA computing. More precisely, we explore ways of encoding information in DNA sequences, molecular procedures used for computation, classes of problems that can be solved by DNA computing, and the feasibility and advantages of a DNA computer.

The construction of a DNA computer would revolutionize the way we think about computing. In the particular case solved by Adleman, the estimated speed exceeded that of current supercomputers by more than a thousandfold, the energy efficiency was 10^{10} times greater, and storing information in DNA molecules was 10^{12} times more economic.

Prerequisites

The course has no *formal* prerequisites. However, some knowledge of computability theory will be assumed. Students with other backgrounds should contact the instructor.

Instructor

Lila Kari, 361 Middlesex College, Office Hours 4.00-5.00pm Wednesday, Tel. 679 2111 ext. 6894, E-Mail lila@csd.uwo.ca, WWW page http://www.csd.uwo.ca/~ lila

Class meetings

Tuesdays, 1:00-3:00pm, MC 354 (tentative).

Topics

- Introduction to DNA computing
- Encoding information in DNA
- Bio-operations
- DNA models of computation
- DNA algorithms
- Error rates in DNA computing

Topic Schedule and Readings[1]

1. Introduction – Prof. Lila Kari.
2. Adleman's first experiment in DNA computing, [1]*, [2], [34]*, [31] – Prof. Lila Kari.
3. Basic concepts used in DNA computing (proteins, amino-acids, nucleotides, ribosomes, DNA), [70], [4]*, [16] – Prof. Greg Gloor.
4. Potential bio-operations: molecular procedures (PCR, biotin-avidin purification, hybridization, etc.), [58] – Prof. Greg Gloor.
5. Models of computation: DNA computing based on contextual insertions and deletion or splicing, [35]*, [28], [13], [22]*, [46], [61], [66] – Prof. Lila Kari.
6. The genetic code, [38]*, [62]* – Dr. Mark Gijzen.
7. TBA.
8. Encoding information using DNA, [6]*, [7], [44]*, [18].
9. Algorithms I: Using DNA to solve NP-complete problems, [23], [39]*, [40]*, [42], [41], [33].
10. Implementation techniques: [53]*, [3], [43]*.
11. Algorithms II: using DNA to solve algebraic problems (addition, matrix multiplication, determinants), [27]*, [45]*, [56].
12. The computational power of DNA computing, [8]*, [9], [54], [12], [55], [67], [68]*, [69].
13. Error rates in DNA computing, [5]*, [10], [37], [60]*, [36].

Evaluation

Seminar Presentation (35 % of final grade)
Each student will lead a seminar on the issue of the day, presenting appropriate information, and initiating and moderating discussion. Presenters will also prepare a 7-page handout for classmates, consisting of a 300 word abstract summarizing their conclusions on the issue, with an annotated bibliography based on 3 to 5 relevant resources. The grade will be based on peer- and self-evaluation.

Essay Exam (50 % percent of the final grade)
The take-home final exam will consist of three sections of two questions each. Students will choose one question from each section, and answer each in 1200 words or less (not including references). The questions will require examinees to evaluate and integrate research on DNA computing and related topics, and identify their implications.

Participation (15 % percent of the final grade)
Participation includes preparation for class discussions, and consistency and quality of contributions. The grade will be determined by the instructor and the student.

Requirements
The handouts are to be written in HTML, in order to be integrated in the WWW course home-page. Hard copies of the handouts are to be distributed to the classmates one week prior to the presentation.

[1] The readings marked with * are required. Copies will be available via 2 hour reserve in the Taylor library.

Remark Depending on the number of participants, the way of conducting the course may have to be modified.

References

[1] L.Adleman. Molecular computation of solutions to combinatorial problems. *Science* v.266, Nov.1994, 1021– 1024.

[2] L.Adleman. On constructing a molecular computer, *ftp: /ftp/pub/csinfo/papers/adleman/molecular_computer.ps.*

[3] L.Adleman, P.Rothemund, S.Roweis, E.Winfree. On applying molecular computation to the data encryption standard[2].

[4] B.Alberts et al. Introductory chapter. *Molecular biology of the cell*, 3rd ed., New York : Garland Pub., c1994.

[5] M.Amos, A.Gibbons, D.Hodgson. Error-resistant implementation of DNA computations. See footnote 2.

[6] E.Baum. Building an associative memory vastly larger than the brain. *Science*, vol.268, April 1995, 583–585.

[7] E.Baum. DNA sequences useful for computation. See footnote 2.

[8] D.Beaver. The complexity of molecular computation. *http://www.transarc.com/~beaver/research/alternative/molecute/molec.html.*

[9] D. Beaver. Computing with DNA. *Journal of Computational Biology*, (2:1), Spring 1995.

[10] D.Boneh, R.Lipton. Making DNA computers error resistant. See footnote 2.

[11] D.Boneh, R.Lipton, C.Dunworth. Breaking DES using a molecular computer. *http://www.cs.princeton.edu/~dabo.*

[12] D.Boneh, R.Lipton, C.Dunworth, J.Sgall. On the computational power of DNA. *http://www.cs.princeton.edu/~dabo.*

[13] E.Csuhaj–Varju, R.Freund, L.Kari, G.Paun. DNA computing based on splicing: universality results. *First Annual Pacific Symposium on Biocomputing*, Hawaii, 1996, also *http://www.csd.uwo.ca/~lila.*

[14] E.Csuhaj–Varju, L.Kari, G.Paun. Test tube distributed systems based on splicing, *http://www.csd.uwo.ca/~lila.*

[15] K.Culik, T.Harju. Splicing semigroups of dominoes and DNA. *Discrete Applied Mathematics*, 31(1991) 261–277.

[16] J.E.Darnell, H.Lodish, D.Baltimore. Introductory chapter. *Molecular cell biology*, New York : Scientific American Books, c1986.

[17] J.Dassow, V.Mitrana. Splicing grammar systems. *Computers and AI.* To appear.

[2] Proceedings of the *2nd Annual Meeting of DNA based computers*, Princeton, 1996

[18] R.Deaton, R.Murphy, M.Garzon, D.Franceschetti, S.Stevens Jr.. Good encodings for DNA-based solutions to combinatorial problems. See footnote 2.

[19] K.L.Deninnghoff, R.W.Gatterdam. On the undecidability of splicing systems. *International Journal of Computer Mathematics*, 27(1989) 133–145.

[20] C.Ferretti, S.Kobayashi, T.Yokomori. DNA splicing systems and Post Systems. *First Annual Pacific Symposium on Biocomputing*, Hawaii, 1996.

[21] R.Freund. Splicing systems on graphs. *Proc.Intelligence in Neural and Biological Systems*, IEEE Press, May 1995, 189–194.

[22] R.Freund, L.Kari, G.Paun. DNA computing based on splicing: the existence of universal computers. T.Report 185–2/FR–2/95, TU Wien, Institute for Computer Languages, 1995, also *http://www.csd.uwo.ca/~lila.*

[23] M.Garey, D.Johnson. *Computers and Intractability. A Guide to the Theory of NP–completeness.* W.H.Freeman and Company, San Francisco, 1979.

[24] R.W.Gatterdam. Splicing systems and regularity. *International Journal of Computer Mathematics*, 31(1989) 63–67.

[25] R.W.Gatterdam. DNA and twist free splicing systems, in *Words, Languages and Combinatorics II*, Eds.:M.Ito and H.Jürgensen, World Scientific Publishers, Singapore, 1994, 170–178.

[26] D.K.Gifford. On the path to computation with DNA. *Science* 266(Nov.1994), 993–994.

[27] F.Guranieri, C.Bancroft. Use of a horizontal chain reaction for DNA based addition. Proceedings of the *2nd Annual Meeting of DNA based computers*, Princeton, 1996.

[28] T.Head. Formal language theory and DNA: an analysis of the generative capacity of recombinant behaviors. *Bulletin of Mathematical Biology*, 49(1987) 737–759.

[29] T.Head. Splicing schemes and DNA, in: *Lindenmayer systems – Impacts on Theoretical Computer Science, Computer Graphics and Developmental Biology.* Springer Verlag, Berlin 1992, 371–383.

[30] D.Hofstadter. *Gödel, Escher, Bach: an eternal golden braid*, New York, Basic Books, 1979.

[31] D.Hofstadter. *Metamagical themas: questing for the essence of mind and pattern*, Basic Books, Inc., Publishers, New York, 1985.

[32] F.Hoppensteadt. Getting started in mathematical biology. *Notices of the AMS*, vol.42, no.9, Sept.1995.

[33] N.Jonoska, S.Karl. A molecular computation of the road coloring problem. Proceedings of the *2nd Annual Meeting of DNA based computers*, Princeton, 1996, 148-159.

[34] L.Kari. DNA computers: tomorrow's reality. *Bulletin of EATCS*, summer 1996.

[35] L.Kari, G.Thierrin. Contextual insertions/deletions and computability. To appear in *Information and computation.*

[36] P.Kaplan, G.Checchi, A.Libchaber. DNA based molecular computation: template-template interactions in PCR. Proceedings of the *2nd Annual Meeting of DNA based computers*, Princeton, 1996, 159-172.

[37] S.Kurtz, S.Mahaney, J.Royer, J.Simon. Active transport in biological computing. See footnote 2.

[38] A.L.Lehninger, D.L.Nelson, M.M.Cox. *Principles of Biochemistry*, Chapter 26: Protein metabolism, Worth Publ., New York, 1993, 892–922.

[39] R.Lipton. DNA solution of hard computational problems. *Science*, vol.268, April 1995, 542–545.

[40] R.Lipton. Using DNA to solve NP–complete problems. *http://www.cs.princeton.edu/~dabo.*

[41] R.Lipton. Using DNA to solve SAT. *http://www.cs.princeton.edu/~dabo.*

[42] R.Lipton. Speeding up computations via molecular biology. Manuscript.

[43] Q.Liu, Z.Guo, A.Condon, R.Corn, M.Lagally, L.Smith. A surface based approach to DNA computation. Proceedings of the *2nd Annual Meeting of DNA based computers*, Princeton, 1996, 206-217.

[44] K.Mir. A restricted genetic alphabet for DNA computing. See footnote 2..

[45] J.Oliver. Computation of DNA matrix multiplication. Proceedings of the *2nd Annual Meeting of DNA based computers*, Princeton, 1996, 236-249.

[46] G.Paun. On the splicing operation. *Discrete Applied Mathematics*, to appear.

[47] G.Paun. On the power of the splicing operation. *International Journal of Computer Mathematics*, to appear.

[48] G.Paun. Regular extended H systems are computationally universal. *Journal of Information Processing and Cybernetics, EIK*, to appear.

[49] G.Paun, G.Rozenberg, A.Salomaa. Computing by splicing. Submitted, 1995.

[50] G.Paun, A.Salomaa. DNA computing based on the splicing operation. Submitted, 1995.

[51] D.Pixton. Linear and circular splicing systems. *Proc.Intelligence in Neural and Biological Systems*, IEEE Press, May 1995, 181–188.

[52] J.Reif. Parallel Molecular Computation. To appear in SPAA'95, also at *http://www.cs.duke.edu/~reif/HomePage.html.*

[53] S.Roweis et all, A sticker based architecture for DNA computation. See footnote 2.

[54] D.Roos, K.Wagner. On the power of DNA computers. Manuscript.

[55] P.Rothemund. A DNA and restriction enzyme implementation of Turing machines. Abstract at *http://www.ugcs.caaltech.edu/~pwkr/oett.html.*

[56] T.Leete, M.Schwartz, R.Williams, D.Wood, J.Salem, H.Rubin. Massively parallel DNA computation: expansion of symbolic determinants. Proceedings of the *2nd Annual Meeting on DNA Based Computers*, Princeton, 1996, 49-67.

[57] A.Salomaa. *Formal Languages.* Academic Press, New York, 1973.

[58] J.Sambrook. E.F.Fritsch, T.Maniatis. Chapter 5 from *Molecular cloning : a laboratory manual*, second edition, 1989.

[59] D.Searls. Formal grammars for intermolecular structure. Proceedings of the *International IEEE Symposium on Intelligence in Neural and Biological Systems*, 1995, pp.30-37.

[60] N.Seeman et all. The perils of polynucleotides: the experimental gap between the design and assembly of unusual DNA structures. Proceedings of the *2nd Annual Meeting of DNA based computers*, Princeton, 1996, 191-206. 1969.

[61] W.Smith, A.Schweitzer. DNA computers in vitro and in vivo, *NEC Technical Report 3/20/95.*

[62] C.Starr, R.Taggart. *Biology: The Unity and Diversity of Life* (7th edition), Chapter 14: From DNA to Protein, 1995, 218-233.

[63] A.M.Turing. On computable numbers, with an application to the Entscheidungsproblem. *Proc.London Math Soc.*, Ser.2, 42(1936), 230–265.

[64] A.Yasuhara. *Recursive Function Theory and Logic.* Academic Press, New York, 1971.

[65] T.Yokomori, S.Kobayashi. DNA evolutionary linguistics and RNA structure modeling: a computational approach. *Proc.Intelligence in Neural and Biological Systems*, IEEE Press, May 1995, 38–45.

[66] T.Yokomori, S.Kobayashi, C.Ferretti. On the power of circular splicing systems and DNA computability. Rep. CSIM–95–01, University of Electro–Communications, Dept.of Comp.Sci. and Information Mathematics, Chofu, Tokyo 182, Japan.

[67] E.Winfree. Complexity of restricted and unrestricted models of molecular computation. *http://dope.caltech.edu/winfree/DNA.html.*

[68] E.Winfree. On the computational power of DNA annealing and ligation. *http://dope.caltech.edu/winfree/DNA.html.*

[69] E.Winfree, X.Yang, N.Seeman. Universal computation via self-assembly of DNA: some theory and experiments. See footnote 2.

[70] S.L.Wolfe. Introductory chapter, *Molecular and cellular biology*, Belmont, Calif. : Wadsworth Pub. Co., c1993.