Solving the Set-Splitting Problem in Sticker-Based Model and the Lipton-Adelmann Model

Weng-Long Chang¹, Minyi Guo², and Michael Ho³
¹,³Department of Information Management, Southern Taiwan University of Technology, Tainan, Taiwan 701, R.O.C.
changwl{@csie.ncku.edu.tw, @mail.stut.edu.tw}
²Department of Computer Software, The University of Aizu, Aizu-Wakamatsu City, Fukushima 965-8580, Japan
minyi@u-aizu.ac.jp

Abstract. Adleman wrote the first paper in which it was demonstrated that DNA (DeoxyriboNucleic Acid) strands could be applied for dealing with solutions to an instance of the NP-complete Hamiltonian path problem (HPP). Lipton wrote the second paper in which it was shown that the Adleman techniques could also be used to solving the NP-complete satisfiability (SAT) problem (the first NP-complete problem). Adleman and his co-authors proposed sticker for enhancing the Adleman-Lipton model. In the paper, it is proved how to apply sticker in the sticker-based model for constructing solution space of DNA for the set-splitting problem and how to apply DNA operations in the Adleman-Lipton model to solve that problem from solution space of sticker.

1 Introduction

Nowadays, it is possible to generate roughly $10^{18}$ DNA strands that fit in a test tube through advances in molecular biology [1]. Adleman [2] wrote the first paper in which DNA strands could be applied to manipulate solutions for an instance of the NP-complete Hamiltonian path problem. Lipton [3] wrote the second paper that demonstrated that the Adleman techniques could be employed to solving the NP-complete satisfiability problem (the first NP-complete problem). Adleman and his co-authors [14] proposed sticker for enhancing error rate of hybridization to the Adleman-Lipton model.

In the paper, we use sticker in the sticker-based model to constructing solution space of DNA for the set-splitting problem. Simultaneously, we also use DNA operations in the Adleman-Lipton model to develop one DNA algorithm. It is proved from the main result of the proposed DNA algorithm that the set-splitting problem is resolved with biological operations in the Adleman-Lipton model from solution space of sticker. Furthermore, this work represents obvious evidence for the ability of DNA-based computing for resolving the NP-complete problems.

The rest of this paper is organized as follows. In Section 2, the Adleman-Lipton model is in detail introduced and the comparison of the model with other models is also
given. Section 3 introduces a DNA algorithm for solving the set-splitting problem from solution space of sticker in the Adleman-Lipton model. In Section 4, the experimental result of simulated DNA computing is also given. Conclusions are drawn in Section 5.

2 DNA Model of Computation

In subsection 2.1, the summary of DNA structure and the Adleman-Lipton model is in detail described. In subsection 2.2, the comparison of the Adleman-Lipton model with other models is also in detail introduced.

2.1 The Adleman-Lipton Model

Due to [1, 16], distinct nucleotides are detected only with their bases, which come in two sorts: purines and pyrimidines. Purines include adenine and guanine, abbreviated 4 and 5. Pyrimidines contain cytosine and thymine, abbreviated C and T. Because nucleotides are only distinguished from their bases, they are simply represented as 4, 5, C, or T nucleotides, depending upon the sort of base that they have.

The DNA operations in the Adleman-Lipton model, cited from [2, 3, 11, 12], are described below. These operations will be used for figuring out solutions of the set-splitting problem.

The Adleman-Lipton model:

A (test) tube is a set of molecules of DNA (i.e. a multi-set of finite strings over the alphabet \{4, C, 5, T\}). Given a tube, one can perform the following operations:

1. Extract. Given a tube 4 and a short single strand of DNA, 5, produce two tubes +4(5, 5) and −4(5, 5), where +4(5, 5) is all of the molecules of DNA in 4 which contain the strand 5 as a sub-strand and −4(5, 5) is all of the molecules of DNA in 4 which do not contain the short strand 5.

2. Merge. Given tubes 41 and 42, yield ∪4(41, 42), where ∪4(41, 42) = 41 ∪ 42. This operation is to pour two tubes into one, with no change of the individual strands.

3. Detect. Given a tube 4, say ‘yes’ if 4 includes at least one DNA molecule, and say ‘no’ if it contains none.

4. Discard. Given a tube 4, the operation will discard the tube 4.

5. Read. Given a tube 4, the operation is used to describe a single molecule, which is contained in the tube 4. Even if 4 contains many different molecules each encoding a different set of bases, the operation can give an explicit description of exactly one of them.

2.2 The Comparison of the Adleman-Lipton Model with Other Models

Quyang et al. [4] proved that restriction enzymes could be used to solve the NP-complete clique problem (MCP). The maximum number of vertices that they can process is limited to 27 because the size of the pool with the size of the problem exponentially increases [4]. Arito et al. [5] described new molecular experimental techniques for searching a Hamiltonian path. Morimoto et al. [6] offered a solid-phase