

行政院國家科學委員會專題研究計畫成果報告

DNA 計算之研究 Study on DNA Computation

計畫編號：NSC 88-2213-E-009-018

執行期限：87 年 08 月 01 日至 88 年 07 月 31 日

主持人：陳榮傑 國立交通大學 資訊工程系

E-mail: rjchen@csie.nctu.edu.tw

一、中文摘要

1959 年 Richard Feynman 提出以分子等級(molecular level)來做計算的夢想，然而直到 1994 年才由 Leonard Adleman 提出一個突破性的 DNA 計算模式來實現。他首先利用 DNA 計算來解決有方向性的 Hamiltonian 路徑問題(此問題為 NP-complete)，主要的理論根據在於，只要我們適當的對 DNA 進行編碼以符合問題的需求，再透過現今生物技術中，對於 DNA 的實驗操作包括結合反應(ligation)、聚合體連鎖反應(polymerase chain reaction)、凝膠電泳法(gel-electrophoresis)、酵素親和純化作用(biotin-avidin affinity purification)等化學反應，由於 DNA 在實驗的過程具有自主性，因此相當於在同一時間內，有大量的處理器在作平行計算，大大地降低計算的時間複雜度，實驗結果也證實其時間複雜度與輸入資料呈線性關係。之後，Donald Beaver 提出分子杜倫機(molecular Turing machine)的模型，作為 DNA 電腦的計算理論基礎。

本計畫蒐集整理了 DNA 計算領域目前最新的研究發展情形，提供一個完整而詳盡的報告，作為國內在此領域的入門基礎。其次就演算法的觀點，將 DNA 電腦實際運用於解決傳統計算模式下圖論和數論兩領域的問題，並提出創新的 DNA 計算模型與 DNA 演算法。最後對於我們所設計出的 DNA 演算法，和傳統計算模式下已經發展的演算法作比較，分析其優缺點，歸納整理出其特性，得出 DNA 電腦最適合解決也最有威力的問題類型。

關鍵詞：DNA 計算、平行計算、分子杜倫機、DNA 電腦、DNA 演算法

Abstract

DNA computation has its roots in 1959, when Richard Feynman first introduced the visionary idea of computing at a molecular level. In 1994, Leonard Adleman used biological experiments with DNA strings to solve the directed Hamiltonian path problem, which is considered to be intractable because of its NP-completeness. The idea is to use strands of DNA to encode the problem, and to manipulate them using techniques commonly available in any molecular biology laboratory, such as ligation, PCR(polymerase chain reaction), gel-electrophoresis, and biotin-avidin affinity purification, etc. After Adleman's experiment, there has been a flurry of interest in the idea of using DNA to perform computations. And some computation theories about DNA computer are developed, such as molecular Turing machine. The potential benefits of using this particular molecule are enormous due to the massive inherent parallelism of performing concurrent operations on trillions of strands.

We survey the field in DNA computation and provide the foundation of this field for people who are interested in it. We then perform DNA computation in dealing with problems which were solved by using conventional algorithms, including graphic theory and number theory, and propose novel DNA

computing models and DNA algorithms. Finally, we compare the DNA algorithms with the conventional algorithms, analyze their advantages and disadvantages respectively, characterize the properties of DNA computers, and find classes of problems for which DNA-based computers will out-perform conventional, silicon-based machines.

Keywords: DNA computation, parallel computation, molecular Turing machine, DNA computer, DNA algorithm

二、緣由與目的

近幾年來，非傳統的計算模式陸續被提出：如量子電腦(Quantum Computer)和 DNA 電腦(DNA Computer)。Leonard M. Adleman[1]於1994年提出一個生物實驗：利用 DNA 來解決有方向性的 Hamiltonian 路徑問題(此問題為 NP-complete)。他透過對 DNA 進行編碼，以及現今生物技術中，對於 DNA 的實驗操作，經由試管成功的解決此問題，並證明其時間複雜度與輸入資料呈線性關係。過去傳統計算理論受限於 Von Neumann 架構下只能循序執行的限制，於是有 NP-complete 難題的產生，之後雖然有平行機器和平行演算法的發展，但是其平行處理的能力始終有一定的限制；然而在 DNA 電腦之下，我們可以很輕易地達到 10^{20} 以上的平行處理能力，因為每個 DNA 分子就像是一個處理器(processor)，而一莫耳(mole)溶液中就有 6.02×10^{23} 個分子。相較於傳統的計算模式，DNA 的計算模式在平行處理能力上，提供非常大的潛力。

生物學上的 DNA 是由四種不同核苷酸(nucleotide)所組成的雙螺旋體，包括：腺嘌呤(adenine ;A)、胸腺嘧啶(thymine ;T)、鳥糞嘌呤(guanine ;G)、胞嘧啶(cytosine ;C)。其中腺嘌呤(A)和胸腺嘧啶(T)互補，鳥

糞嘌呤(G)和胞嘧啶(C)互補；即在 DNA 的雙螺旋體結構中，A 和 T、G 和 C 必定是面對面相連。故由雙螺旋體中一股(strand)的排列順序即可推知另一股的排列順序。例如：

一股為： \uparrow AGTC \equiv 5'-AGTC-3'

另一股為： \downarrow AGTC \equiv 3'-TCAG-5'

透過對 DNA 結構及生物技術的認識，讓我們更能掌握 DNA 計算的原理。之後陸續有多篇報告針對不同的 DNA 編碼方式和使用的生物技術，將 DNA 的計算模式分類。在 Beaver[5]中更提出了分子杜倫

機(molecular Turing machine)的觀念，將 Turing machine 中的 configuration(包括磁帶內容、現在狀態、以及磁頭位置)用 DNA 分子來編碼，狀態的轉換則是經由生物步驟(bio-step)來達成，建立了 DNA 計算理論的基礎。另一方面，為了能將 DNA 計算以高階(High-level)的模式來描述，Roos[16]提出一種分子程式語言：DNA-Pascal，將 DNA 計算中出現的操作(operation)加以整理分類，以函數的方式表現，配合傳統 Pascal 的架構，如 For 迴圈、If-then 架構，將 DNA 計算以演算法的型態表示，並證明使用不同操作集合(operation set)的計算模式，其計算能力也不相同：例如 Lipton[14]模式下線性時間能解的問題屬於 P^{NP} (oracle Turing machine)。

在應用方面，Baum[3]提出 DNA 電腦中記憶體的設計—關聯性記憶體(associative memory)；[6]則利用 DNA 來解決密碼學上破解 DES(Data Encryption Standard)的問題；[4]介紹如何在 DNA 電腦上面執行動態程式設計(dynamic programming)的演算法，利用 DNA 計算來解決 knapsack 問題，不同於之前[1][14]所提出的暴力窮舉式的演算法(brute force algorithms)，[4]運用和傳統相同的演算法，並提出在 DNA 電腦中如何做加法的運算。[15]中則將 DNA 計算運用於矩陣的乘法及乘冪，提出利用分子濃度做乘法運算的方式。

針對 DNA 電腦發展的潛力，我們蒐集整理了最新的研究發展情形，提供一個完整而詳盡的報告，作為國內在此領域的入門基礎。另外為了協助 DNA 電腦成為一般目的(general-purpose)的計算機，我們提出創新的 DNA 計算模式和 DNA 演算法來解決傳統計算模式下的問題，並與傳統的演算法做比較，藉由分析比較中，歸納整理出 DNA 計算的特性，得出 DNA 電腦最適合解決也最有威力的問題類型。

三、結果與討論

我們蒐集了到目前為止有關 DNA 計算的研究報告，整理完成 DNA 計算的模型假設與理論基礎[9]，提供給國內作為在此領域的入門基礎。

其次，對於傳統計算模式下的問題，我們分別對圖學理論中的推銷員旅行問題(traveling-salesman problem)以及代數理論中的 0 與 1 整數規劃(0-1 integer programming)，分別提出不同的 DNA 演算法[9]，詳細說明其編碼方式以及操作步驟，並證明其時間複雜度為線性；相較於在傳統計算模式下，這兩個問題都是 NP-complete，DNA 計算的確提供了更有效率的計算方式。

據此，我們完成本計畫的目的並在 DNA 計算的領域上有所貢獻。

四、計畫成果自評

依上節所提之結果，我們達成了此計畫預期的目標。此計畫的研究結果除了為國內其他研究者提供 DNA 計算領域的詳盡介紹，另外透過設計 DNA 演算法的過程，參與的研究人員能實際瞭解 DNA 電腦的發展架構，並啟發非傳統計算模式的思考，分析 DNA 電腦未來的發展方向。成果極具有學術上的價值與貢獻，相當適合學術期刊上發表。

五、參考文獻

- [1] L. M. Adleman, "Molecular Computation of Solutions to Combinatorial Problems," *Science* 266:1021-1024, December 1994.
- [2] M. Amos, A. Gibbons, and D. Hodgson, "Error-resistant Implementation of DNA Computations," In *Proceedings of the Second Annual Meeting on DNA based Computers*, 1996.
- [3] E. B. Baum, "A DNA associative memory potentially larger than the brain," In *DNA based computers*. Vol. 27 in DIMACS, 1996
- [4] E. B. Baum and D. Boneh, "Running dynamic programming algorithms on a DNA computer," In

Proceedings of the Second Annual Meeting on DNA based Computers, 1996.

- [5] D. Beaver, "A universal molecular computer," In *DNA based computers*. Vol. 27 in DIMACS, 1996.
- [6] D. Boneh, C. Dunworth, and R. J. Lipton, "Breaking DES using a molecular computer," In *DNA based computers*. Vol. 27 in DIMACS, 1996
- [7] David Freifelder, *Molecular Biology*, Jones and Bartlett, 1987.
- [8] V. Gupta, S. Parthasarathy, and M. J. Zaki, "Arithmetic and Logic Operations with DNA," In *Proceedings of the Third Annual Meeting on DNA based Computers*, 1997.
- [9] J. S. Hwu and R. J. Chen, *Study on DNA Computation*, Master thesis, the National Chiao Tung University, (1998).
- [10] J. S. Hwu and R. J. Chen, "DNA Solution to the Traveling Salesman Optimization Problem", In *Proceedings of workshop on algorithms*, pp.44-50, International Computer Symposium, 1998.
- [11] L. Kari, G. Gloor, and S. Yu, "Using DNA to Solve the Bounded Post Correspondence Problem," *Technical Report*, University of Western Ontario, 1997.
- [12] L. Kari, G. Gloor, and S. Yu, "Towards a Solution for the Shortest Common Superstring Problem," *Technical Report*, University of Western Ontario, 1997.
- [13] R. M. Karp, C. Kenyon, and Orli Waarts, "Error-resilient DNA Computation," In *7th ACM-SIAM Symposium on Discrete Algorithms*, pp.458-467, SIAM, 1996.
- [14] R. J. Lipton, "DNA Solution of Hard Computational Problems," *Science* 268:542-545, 1995.
- [15] J. S. Oliver, "Computation with DNA-Matrix Multiplication," In *Proceedings of the Second Annual Meeting on DNA based Computers*, 1996.
- [16] Diana Roo β and K. W. Wagner, "On the Power of DNA Computing," In *Information and Computation* Vol. 131, pp. 95-109, 1996.