



REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPOLAIRE  
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA  
RECHERCHE SIENTIFIQUE

Université Elhadj Lakhder, Batna  
Faculté des sciences de l'ingénieur  
*Département d'informatique*  
Option : Informatique industrielle

## Mémoire

Pour l'obtention du diplôme de magister

# *Systemes bio-inspirés pour le traitement de l'information : Application du DNA computing à la résolution de problèmes NP-complets*

Présenté et soutenu par

AKSA KARIMA

Le :

Directeur de Mémoire : Pr M.C.Batouche

Composition du JURY

Dr A.Zidani	Maître de conférence	Président
Dr M.Benmohammed	Maître de conférence	Examineur
Dr D.E.Saidouni	Maître de conférence	Examineur

## Résumé

La solution d'un problème combinatoire, quand elle existe, peut être déterminée par l'énumération d'un ensemble fini E de possibilités. S'il y a un algorithme simple pour tester si un élément de cet ensemble est une solution, on dispose d'un algorithme pour trouver une solution au problème en testant toutes les possibilités. Mais le cardinal de E est souvent exponentiel par rapport à la taille des données, ce qui rend l'énumération impossible dans un temps raisonnable.

Il existe de très nombreux problèmes d'un grand intérêt pratique pour lesquels on ne connaît pas d'algorithme plus efficace qu'un test portant sur tous les sous ensembles d'un ensemble, ce qui implique un temps de calcul exponentiel par rapport à la taille de l'ensemble.

De nos jours, certains problèmes nécessitent, pour être résolus une capacité de calcul phénoménale. L'ordinateur actuel étant limité dans sa puissance de calcul constitue une barrière dans la résolution de tels problèmes. C'est dans cette optique que de nouveaux concepts d'ordinateurs sont développés. Ainsi on tente de mettre sur pied des ordinateurs quantiques et des ordinateurs moléculaires. Ces derniers représentent certainement l'alternative de demain aux ordinateurs actuels.

En utilisant l'ADN, contrairement aux ordinateurs actuels, cela permet de réaliser des opérations en parallèles (multitask) avec des vitesses de calculs phénoménales. De plus, l'ADN constitue un excellent moyen de stockage de données et ne demande qu'un apport infime en énergie.

L'expérience qui a éveillé le monde des ordinateurs à ADN est celle d'Adleman. En 1994, Adleman, qui est un mathématicien de l'université de Californie du sud et qui possède également des connaissances en biologie, a résolu grâce à un ordinateur moléculaire le problème de chemin hamiltonien (HPP : Hamiltonian Path Problem).

Depuis l'expérience originale d'Adleman, des inondations d'idées ont été proposées pour la résolution des différents problèmes NP-complets. Dans ce mémoire, deux nouveaux algorithmes ont été proposés pour résoudre deux problèmes NP-complets : le Problème de Voyageur de Commerce (TSP : Traveling Salesmen Problem), et celui de la Satisfaisabilité (SAT : Satisfiability problem). Théoriquement, ces deux algorithmes proposés et d'autres types d'algorithmes peuvent être exécutés également avec succès. Mais pratiquement ils seront très difficiles pour l'application, car ici on travaille avec la biologie qui nécessite la prudence pour appliquer des opérations biologiques, tel que la dénaturation, sur des molécules vivantes.

Malgré le fait que les ordinateurs moléculaires (ordinateurs d'ADN) semblent être très performants, ils possèdent de nombreux aspects négatifs. Ils ne peuvent résoudre que des problèmes combinatoires (pas de possibilité de traitement de texte ou de jeu sur de tels ordinateurs), ils peuvent être très lents dans la résolution de problèmes simples pour des ordinateurs classiques. Les réponses qu'ils fournissent peuvent être extrêmement compliquées. D'autre part la fiabilité de ces ordinateurs peut être remise en cause du fait de la capacité de la mutation de l'ADN.

**MOTS CLES :** Calcul moléculaire, Calcul de l'ADN, Ordinateur à l'ADN, Bio-calcul, Problèmes NP-complets, Bio-informatique, Systèmes bio-inspirés, Biologie moléculaire, ADN.

# Abstract

The solution of a combinatorial problem, when it exists, can be determined by the enumeration of a finished set E of possibilities. If there is a simple algorithm to test if an element of this set is a solution, one has an algorithm to find a solution of the problem by testing all the possibilities. But the cardinal of E is often exponential with regard to the size of data, what returns impossible enumeration in a reasonable time.

There are very numerous problems of a big practical interest for which one does not know algorithm more effective than a test concerning all the sub-sets of a set, what implies a time of exponential calculation with regard to the size of the set.

Nowadays, certain problems require, to be resolved a phenomenal capacity of calculation. The current computer being limited in its power of calculation constitutes a barrier in the resolution of such problems. It is in this optics that new concepts of computers are developed. So one tries to set up quantum computers and molecular computers. These last ones represent certainly the alternative of tomorrow to the current computers.

By using the DNA, contrary to the current computers, it allows to realize operations in parallel (multitask) with phenomenal speeds of calculations. Furthermore, the DNA constitutes an excellent means of stocking of data and asks only for a tiny contribution in energy.

The experience which has to wake the world of computers to DNA is that of Adleman. In 1994, Adleman, who is a mathematician of the University of South California and who also possesses knowledge in biology, has resolved due to a molecular computer the Hamiltonian Path Problem (HPP).

Since the original experience of Adleman, floods of ideas were proposed for resolutions of various NP-complets problems. In this memoir, two new algorithms were proposed to resolve two NP-complets problems: Commercial traveller's problem Traveling Salesmen Problem (TSP), and that of Satisfiability problem (SAT). Theoretically, these two algorithms proposed and other types of algorithms can be executed successfully. But practically they will be very difficult for application, because here one works with biology which necessitates prudence to apply biologic operations, such as denaturation, on alive molecules.

In spite of the fact that the molecular computers (DNA computers) seem to be very successful, they possess numerous negative aspects. They can resolve only combinatorial problems (no possibility of text processing or game on such computers), they can be very slow in the resolution of simple problems for classic computers. The answers which they supply can be extremely complicated. On the other hand the reliability of these computers can be questioned because of the capacity of the DNA mutation.

**WORD KEY :** Molecular computing, DNA computing, DNA computer, Biocomputing, NP-complets problems, Bio-informatics, Bio-inspired systems, Molecular biology, DNA.

# Table de matières

<b>Résumé.....</b>	<b>1</b>
<b>Table des matières .....</b>	<b>3</b>
<b>Liste des figures .....</b>	<b>6</b>
<b>Liste des tableaux .....</b>	<b>9</b>
<b>Introduction générale .....</b>	<b>10</b>
<b>Chapitre1 : Les systèmes bio-inspirés</b>	
1.1 Introduction .....	12
1.2 Intelligence artificielle .....	12
1.3 Vie artificielle.....	13
1.3.1 Les critères de la vie artificielle.....	13
1.3.2 Les domaines de la vie artificielle .....	14
1.4 Les systèmes bio-inspirés.....	14
1.4.1 Pourquoi étudier les systèmes biologiques ?.....	14
1.4.2 Les types principaux des systèmes bio-inspirés .....	15
1.4.2.1 L'intelligence d'essaim 'Swarm Intelligence' .....	15
1.4.2.1.1 L'optimisation par Colonies de fourmis .....	16
1.4.2.1.2 La communication et l'auto-organisation chez les colonies d'abeilles.....	21
1.4.2.2 Les boids de Reynolds .....	23
1.4.2.3 Les essaims de particule (Particle Swarm).....	25
1.4.2.4 Le calcul à l'ADN (DNA computing).....	26
1.4.2.5 Les réseaux de neurones artificiels (Artificial Neural Network).....	26
1.4.2.6 Algorithmes évolutionnaires (EA : Evolutionary Algorithms) .....	30
1.4.2.7 Système Immunitaire Artificielle (AIS : Artificial Immune System).....	34
1.4.2.8 Les automates cellulaires .....	36
1.4.2.9 L-système .....	38
1.5 Conclusion.....	39

**Chapitre2 : L'étude biologique de l'ADN**

2.1 Introduction .....40

2.2 L'ADN, le miracle .....40

2.3 Où trouve-t-on l'ADN dans les organismes?.....41

2.4 La structure de l'ADN .....43

2.5 Les outils de la biologie moléculaire .....45

2.5.1 Les enzymes .....45

2.5.1.1 Les polymérase.....46

2.5.1.2 Ligase .....48

2.5.1.3 Enzymes de restrictions.....48

2.5.2 Electrophorèse .....50

2.5.3 PCR.....51

2.5.4 L'ADN recombinant.....53

2.5.5 Séquençage d'ADN .....54

2.5.6 Synthèse d'ADN .....55

2.6 Conclusion.....56

**Chapitre3 : L'expérience d'Adleman & les ordinateurs à l'ADN**

3.1 Introduction .....57

3.2 Le langage .....57

3.3 Les machines de Turing et les systèmes d'insertion délétion .....60

3.4 Les machines à l'ADN concrètes .....63

3.4.1 L'expérience d'Adleman.....63

3.4.1.1 Le problème du chemin hamiltonien (Hamiltonian path problem (HPP)).....63

3.4.1.2 Algorithme d'Adleman .....64

3.4.1.3 Exemple d'application de l'algorithme d'Adleman.....67

3.4.2 Les travaux qui suivent le travail d'Adleman .....73

3.4.3 Les limites actuelles des ordinateurs à l'ADN .....74

3.5 Conclusion.....75

**Chapitre4 : Résolution des problèmes NP-complets par le calcul d'ADN**

4.1 Introduction .....76

4.2 Généralités sur la complexité des problèmes .....76

4.2.1 Notion de complexité : Définition .....	76
4.2.2 Problèmes de décision et d'optimisation .....	77
4.2.3 Classement des problèmes .....	77
4.2.3.1 Problèmes « faciles », problèmes « difficiles » .....	77
4.2.3.2 Les classes P et NP .....	77
4.2.3.3 Classe NP-complet.....	78
4.2.3.4 Relation entre les classes.....	80
4.3 Les problèmes NP-complet les plus célèbres.....	80
4.3.1 Problème de satisfiabilité : SAT.....	80
4.3.2 Problème du voyageur de commerce.....	81
4.3.3 Problème de sac à dos (Knapsack problem).....	82
4.3.4 Problèmes de 8 reines .....	82
4.4 Résolution de quelques problèmes NP-complet en utilisant le calcul d'ADN .....	83
4.4.1 Le problème de la satisfaisabilité .....	83
4.4.2 Le problème de voyageur de commerce (TSP : Traveling Salesman Problem) .....	85
4.5 Deux algorithmes proposés pour la résolution de TSP et SAT par le calcul d'ADN .....	88
4.5.1 Nouvel algorithme proposé pour la résolution du TSP.....	88
4.5.1.1 Etapes de l'algorithme .....	89
4.5.1.2 Exemple d'application .....	90
4.5.1.3 Discussion .....	93
4.5.2 Nouveaux algorithmes proposés pour la résolution de K-SAT.....	94
4.5.2.1 Algorithme1.....	94
4.5.2.1.1 Les étapes de l'algorithme1 .....	94
4.5.2.1.2 Exemple d'application .....	95
4.5.2.1.3 Discussion.....	97
4.5.2.2 Algorithme2 .....	98
4.5.2.2.1 Les étapes de l'algorithme2.....	98
4.5.2.2.2 Exemple d'application .....	99
4.5.2.3 Discussion .....	101
4.6 Conclusion .....	101
<b>Conclusion générale.....</b>	<b>102</b>

**Bibliographie**