

# **MOLECULAR DOCKING USING MULTI-AGENT TECHNOLOGY**

Mahabaduge Harindra Rasanjali Fernando

(139105C)

Degree of Master of Science in Artificial Intelligence

Department of Computational Mathematics

University of Moratuwa

Sri Lanka

December 2015

# **MOLECULAR DOCKING USING MULTI-AGENT TECHNOLOGY**

Mahabaduge Harindra Rasanjali Fernando

(139105C)

Thesis submitted in  
partial fulfillment of the requirements for the degree of MSc in Artificial Intelligence

Department of Computational Mathematics

University of Moratuwa

Sri Lanka

December 2015

## **Declaration**

I declare that this dissertation does not incorporate, without acknowledgment, any material previously submitted for a Degree or a Diploma in any University and to the best of my knowledge and belief, it does not contain any material previously published or written by another person or myself except where due reference is made in the text. I also hereby give consent for my dissertation, if accepted, to be made available for photocopying and for interlibrary loans, and for the title and summary to be made available to outside organization.

Name of Student

Mrs. M.Harindra.R.Fernando

Signature of Student

Date:

Supervised by

Prof.Asoka Karunananda

Signature of Supervisor

Date:

Supervised by

Dr. Roshan P. Perera

Signature of Supervisor

Date:

## **Dedication**

*I dedicate this thesis to my parents*

*Mr.Shelton Fernando and Mrs.Hemamalani Fernando.*

*I hope this achievement will complete the dream that both of you had for me all those many years ago when you choose to give the best education you could.*

## **Acknowledgements**

Foremost, I would like to express my sincere gratitude to my advisor Prof. Asoka Somabandu Karunanada for the continuous support of my M.Sc. research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor for my study. He helps me a lot to penetrate into the scientific research world with his excellent and valuable advice.

Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Nalin Wickramarachchi and Prof. N G J Dias for their encouragement, insightful comments, and hard questions.

My sincere thanks also go to Dr.Roshan Perera for offering me chemistry related knowledge, leading me working on a diverse exciting project as the co-supervisor.

It is an honor for me to mention the name of Prof. Ranil Dassanayake, Department of Chemistry, University of Colombo, who gave me massive guidance for the evaluation of the software tool. Without him the value of this project will be a dream.

I would like to thank my family: my parents, for giving birth to me, in the first place and supporting me spiritually throughout my life. Last but not the least, my special thanks should go to my loving husband, to assist me, by giving his fullest cooperation to fulfill my research work.

## Abstract

Traditional computer-based simulators for manual molecular docking for rational drug discovery have been very time consuming and a tedious task. It is evident from the literature that such computer-based solutions have been implemented merely with conventional software technologies. A large body of research publication has shown the power of Multi Agent technology for development of smart fully automated simulators.

In this research, a multi agent-based solution, named as NanoAgent, has been developed to automate the drug discovery process with little human intervention. In this solution, *ligands* and *proteins* are implemented as agents, who pose the knowledge of permitted connections with other agents to form new molecules. The system also includes several other agents for surface determination, cavity finding and energy calculation. These agents autonomously activate and communicate with each other to come up with a most probable structure over the ligands and proteins, which are participating in deliberation. Domain ontology is maintained to store the common knowledge of molecular bindings, whereas specific rules pertaining to the behavior of ligands and proteins are stored in their personal ontologies. Among other operational rules, agents are built with rules pertaining to theories of Poison Boltzmann, Vander Walls, and Monte Carlo, regarding ligands and proteins to calculate the optimal binding energy. Existing, Protein Data Bank (PDB) has also been used to calculate the space required by ligand to bond with the receptor. The drug discovery process of NanoAgent has exemplified exciting features of multi agent technology, including, communication, coordination, negotiation, butterfly effect, self-organizing and emergent behavior. Since agents consume fewer computing resources, NanoAgent has recorded optimal performance during the drug discovery process.

NanoAgent has been tested for the discovery of the known drugs for the known protein targets. It has 80% accuracy by considering the prediction of the correct actual existence of the docked molecules using energy calculations. By comparing the time taken for the manual docking process with the time taken for the molecular docking by NanoAgent, there has been 95% efficiency. The results suggest that the Multi-Agent Systems technology can be successfully applied to automate the manual molecular docking process, which is an inherently complex problem. Further work on this project can be identified as the development of automated solutions for protein-protein docking, which is a hot topic in biochemistry and allied disciplines.

# Contents

	<b>Page</b>
<b>Chapter 1 Introduction</b>	<b>1</b>
1.1. Prolegomena	1
1.2. Aim and Objectives	2
1.3. Background and Motivation	2
1.4. Problem in Brief	3
1.5. Novel Approach to Molecular Docking	4
1.6. Outline of the Thesis	5
1.7. Summary	5
<b>Chapter 2 State of the Art of Molecular Docking</b>	<b>6</b>
2.1. Introduction	6
2.2. Molecular Docking Process	6
2.3. Free Chemistry Databases	8
2.4. Computational Molecular Simulation Tools	8
2.4.1. Systematic Search	9
2.4.2. Stochastic Algorithms	9
2.5. Applications of AI Techniques for Molecular Docking	11
2.6. ParDOCK's Active Site Prediction Mechanism	13
2.7. Summary	14
<b>Chapter 3 Multi-Agent Technology and Ontology in AI</b>	<b>15</b>
3.1. Introduction	15
3.2. Overview	15
3.3. What is Agent Technology?	16
3.4. Characteristics of Software Agent	17
3.5. Emergent Behaviour of Multi-Agent Systems	17
3.6. Need of Ontology in Agents	17
3.7. Distributed Decision Making	18
3.8. Use of Ontologies in AI	19
3.9. Knowledge Representation using Ontology in Chemistry	20
3.10. Summary	21



<b>Chapter 4 Molecular Docking</b>	<b>22</b>
4.1. Introduction	22
4.2. Steps of Molecular Docking	22
4.5. Lock and Key Theory	22
4.5 Different Types of Interactions	23
4.5 Rule to Choose Appropriate Ligands	23
4.5.1 Lipinski's rule of five:	23
4.6. Energy Minimization	23
4.7. Bond Energy	24
4.8. Stretching	25
4.9. Bending	25
4.10. Rotating	26
4.11. Force Field Function	26
4.12. Summary	27
<b>Chapter 5 MAS Approach to Molecular Docking</b>	<b>28</b>
5.1. Introduction	28
5.2. Hypothesis	28
5.3. Building Blocks of Approach	28
5.3 Inputs	28
5.4 Protein Data Bank (PDB) File	29
5.5 Chemical Ontologies	30
5.6 Outputs	30
5.7 Process for the Molecular Docking	30
5.8 JADE – Java Agent Development Framework	32
5.9 Jena – A Semantic Web Framework for Java	32
5.10 SciPy - Scientific Computing Tools for Python	32
5.11 OWL – Web Ontology Language	32
5.12 Jmol - An Open-Source Java Viewer for Chemical Structures	32
5.13 MySQL - Database Management Tool	33
5.14 Features – Non Functional Requirements	33
5.15 Users of the Molecular Docking	33
5.16 Summary	33

<b>Chapter 6 MAS for Molecular Docking</b>	<b>34</b>
6.1. Introduction	34
6.2. System Integration of the NanoAgents	34
6.2.1. System Architecture	35
6.3. Searching Mechanism	37
6.4. Calculate pocket and cavities	37
6.5. Active site predictor	37
6.6. Molecular Visualizer	38
6.7. Energy Calculation	38
6.8. Summary	39
<b>Chapter 7 Implementation of MAS for Molecular Docking</b>	<b>40</b>
7.1. Introduction	40
7.5. How to find matching ligand to the protein	40
7.3. Methods to find Active Sites of the Protein	41
7.3.1. Alpha Spheres	41
7.4. Finding cavities in the protein	41
7.5. Data for Ligands and Proteins	42
7.5. Energy Calculation	44
7.6. Summary	45
<b>Chapter 8 Evaluation of MAS based Solution</b>	<b>46</b>
8.1. Introduction	46
8.2. Evaluation Methods	46
8.3. Known Receptor and Ligand Pairs	46
8.3. Test Results	47
8.4. Precision	48
8.5. Precision for Criteria 01:	48
8.6. Precision for Criteria 02:	49
8.4. Summary	49
<b>Chapter 9 Conclusion and Further Work</b>	<b>50</b>
9.1. Introduction	50
8.7. Conclusion	50

8.8.	Achievement of the Objectives	51
9.2.	Further Work to Improve Molecular Docking Tool	52
9.3.	Summary	52
<b>Appendix A MAS Based Molecular Docking Tool</b>		<b>60</b>
A.1	Screenshots	60
<b>Appendix B Free Databases and Ontologies</b>		<b>67</b>
B.1	Ligand Database	67
B.2	Protein Database	67
B.3	Atom Ontology	68

# List of Figures

	<b>Page</b>
Figure 2.1: Classification of the methods for protein-ligand docking [24]	7
Figure 2.2: Haptic device and interface	11
Figure 2.3: ParDOCK docking flow chart [50]	13
Figure 3.3: PEAS of agent	16
Figure 3.7: General Structure of Agent [64]	18
Figure 3.8: Structure of Ontology	19
Figure 4.3: Lock and Key model	23
Figure 4.7.1: Bond interactions in Methane Molecule	27
Figure 5.4.1: Structure of a PDB file [85]	29
Figure 5.5.1: Atom ontology in Protégé [61]	30
Figure 5.7.1: Flow diagram for the drug design	31
Figure 6.3: Three degrees of freedom	37
Figure 7.5.1: DOCK program algorithm	41
Figure 7.3.1.1: An Alpha sphere in 2D	43
Figure 7.4.1: Cavities formed by Gray Atoms	43
Figure 7.4.2: Class diagram for VoronoiDiagrams	44
Figure 7.5.1. Protein table structure in phpMyAdmin [107] interface	44
Figure 7.5.2. Data for protein 1RYJ	44
Figure 7.5.3. Ligand table structure in phpMyAdmin [107] interface	45
Figure 7.5.4. Data for ligand Dalfopristin [110]	45
Figure 7.5.5: Class diagram - Energy Calculation Agent	46

## List of Tables

	<b>Page</b>
Table 2.1: Protein Flexibility: Four methods	8
Table 2.2: Chemical Structure Databases available on the Internet	8
Table 5.3.1: Inputs for the Molecular Docking System	32
Table 8.3.1: Known Receptor and Ligand Pairs	60
Table 8.3.1: Known Protein-Ligand pairs and the results	51

## List of Equations

	<b>Page</b>
Equation 4.6.1: Electrostatic Potential Energy	24
Equation 4.6.2: Van der Waals Equation	25
Equation 4.7.1 Formula to calculate bond energy	25
Equation 4.7.2 Formula to find bonded energy	25
Equation 4.8.1: Bond energy for stretching	26
Equation 4.9.1: Bond Energy for Bending	26
Equation 4.10.1: Bond energy for bending	27
Equation 4.11.1: Bond energy for bending	27
Equation 8.4.1 Precision Equation [117]	49