Generation of Sequence Diagram Automatically from Use Case Model Using Genetic Algorithm

Submitted in Partial Fulfilment of the Requirements for the Degree of
Master of Science in Software Engineering

By

Heba A.Nassar

Supervisors

Dr. Aysh Alhroob Dr. Ayad T. Al-Zobaydi

Faculty of Information Technology
Master Program of Software Engineering
Al-Isra University

2016 / 2017

The undersigned have examined the thesis entitled '***Generation of Sequence Diagram Automatically from Use Case Model***' presented by ***Heba A. Nassar***, a candidate for the degree of Master of Science in Software Engineering and hereby certify that it is worthy of acceptance.

Date	Dr. Aysh Alhroob
Date	Dr. Ayad T. Al-Zobaydi
Date	Prof. Dr. Mohammad Al-Fayoumi
Date	Dr. Thamer Alrawshdeh
Date	Dean of the Faculty of IT

DEDICATION

To My great parents who never stop giving of themselves in countless ways

My dearest husband, who leads me through the valley of darkness with light of hope and support

My beloved kids: Salma and Salem, whom I can't force myself to stop loving.

To all my family, the symbol of love and giving

My friends who encourage and support me

All the people in my life who touch my heart, I dedicate this research

Heba 2017

ACKNOWLEDGMENTS

Exquisite thanks and appreciation goes to my supervisors Dr. Aysh Alhroob and Dr. Ayad T. Al-Zobaydi for their faith and support along this study and for his guidance and comments. I would like to thank Al Isra University; my second magnificent home, all members in my Faculty of Information Technology. Finally, I would like to express special words of thanks to my family for their moral and spiritual support.

TABLE OF CONTENTS

LIST OF TABLES **1**

LIST OF FIGURES **2**

LIST OF ABBREVIATIONS **3**

ABSTRACT **4**

CHAPTER 1: INTRODUCTION **5**

1.1 INTRODUCTION 5
1.2 RESEARCH QUESTION 8
1.3 RESEARCH AIM AND OBJECTIVES 8
1.4 RESEARCH METHOD 8
1.5 MOTIVATIONS 9
1.6 CONTRIBUTION(S) 10
1.7 ORGANIZATION OF THE THESIS 10

CHAPTER 2: BACKGROUND AND RELATED WORKS **11**

2.1 BACKGROUND 11
2.1.1 Natural language processing (NLP) *11*
2.1.2 Thematic Role *13*
2.1.3 Semantic Role Labelling (SRL) *15*
2.1.6 Genetic Algorithms (GA) *16*
2.2 RELATED WORKS 19
2.3 CONCLUSION OF CHAPTER 2 22

CHAPTER 3: THE DESIGN OF ISDG I-CASE TOOL **24**

3.1 THE PLAN FOR ACHIEVING GOAL OF THE RESEARCH 24
3.2 DEFINING AN ALGORITHMIC APPROACH FOR DEVELOPING SEQUENCE DIAGRAM 25
3.2.1 Withdraw Use Case Tracing Example *27*
3.3 DEFINING THE SEMI-AUTOMATED AASSD 28
3.4 CONVERTING THE SEMI-AUTOMATED AASSD TO I-CASE TOOL 30
3.4.1 The Functional Requirements of ISDG *30*
3.4.2 The Architecture of ISDG *35*
3.4.3 The Methods of the Classes *36*
3.5 CONCLUSION OF CHAPTER 3 38

CHAPTER 4: EVALUATION OF ISDG I-CASE TOOL **40**

4.1 INTRODUCTION 40
4.2 ACCURACY MEASURING APPROACH 40

4.3 DATA SET 43
4.4 ACHIEVEMENTS OF ISDG 43
4.5 CONCLUSION OF CHAPTER 4 47
CHAPTER 5: CONCLUSIONS AND RECOMMENDATIONS **48**
5.1 CONCLUSION 48
5.2 RECOMMENDATIONS 49
REFERENCES **50**
APPENDIX A: DATA SET OF ISDG **53**
A.1 TRAINING DATA SET 53
APPENDIX B: SCREEN SHOOTS OF ISDG **63**
APPENDIX C: ACCEPTANCE LETTER **66**

LIST OF TABLES

Table Page

Table 2-1: The Related Works 22
Table 3-1: Specification of Train Use Case..... 31
Table 3-2: Relationships Between Train Classes 32
Table 3-3: Specification of Test Use Case 33
Table 3-4: Relationships Between Test Classes 34
Table 3-5: Code Table Used by Codify Procedure 36

LIST OF FIGURES

Figure Page

Figure 1-1: UML Diagrams 5
Figure 1-2: Research Methodology 9
Figure 2-1: Levels of NLP 11
Figure 2-2: Results of SRL 15
Figure 2-3: Flow Diagram of Genetic Algorithm 17
Figure 2-4: Crossover and Mutation in GA 18
Figure 3-1: Plan for Achieving the Goal of Thesis 24
Figure 3-2: Flow Diagram of Developing Sequence Diagram 25
Figure 3-3: Withdraw Money from an ATM Use Case Model (Sun, 2015) 26
Figure 3-4: Sequence Diagram for Withdraw Money from an ATM 27
Figure 3-5: Enroll in Course Use Case Model (Nazeer, 2005) 28
Figure 3-6: Sequence Diagram for Enroll in Course 28
Figure 3-7: The Semi-Automated AASSD Approach 29
Figure 3-8: Methodology Used by ISDG Tool 30
Figure 3-9: Use Case Diagram for ISDG Tool 31
Figure 3-10: Sequence Diagram for Train 32
Figure 3-11: Class Diagram for Train Use Case 33
Figure 3-12: Sequence Diagram for Test 34
Figure 3-13: Class Diagram for Test 35
Figure 3-14: Class Diagram for ISDG Tool 35
Figure 4-1: Confusion Matrix for the Classification of Car 41
Figure 4-2 : Confusion Matrix Using Two Described Values 42

LIST OF ABBREVIATIONS

I-CASE	Intelligent Computer Aided Software Engineering
GUI	Graphical User Interface
NLP	Natural Language Processing
AI	Artificial intelligence
GA	Genetic Algorithm
SE	Software Engineering
UML	Unified Modelling Language
OOADM	Object Oriented Analysis and Design Method
AASSD	Algorithmic Approach for Sketching Sequence Diagram
SRS	Software Requirements Specification
SRL	Semantic Role Labelling
NomBank	Noun Bank
TP	True positive
FP	False positive
TN	True negative
FN	False negative
POS	Part of Speech
ISDG	Intelligent Sequence Diagram Generator

ABSTRACT

Sequence diagram is a method used for presenting the details of interactions between users and system's components. Sequence diagram helps in transition to a more formal level of refinement of the requirement description. Typically, system analysts are responsible for performing this process, and they usually perform the developing of sequence diagram manually.

The aim of this thesis is to develop a software tool that generates sequence diagram(s) from flow of events founded in use case model, which is called Intelligent Sequence Diagram Generator (ISDG). This tool belongs to Intelligent Computer Aided Software Engineering (I-CASE) tools that have some sort of intelligence to perform those tasks that need human intellectual skills. Thematic role principle is used to distinguish the components of sequence diagram from the statement of flow of events. Semantic Role Labelling software type of Natural Language Processing (NLP) tools is used for automatically discovering the thematic role of each word in the input statement. The proposed solution delivered by this research passes through defining a new algorithmic approach for developing sequence diagram with two versions; manual and semi automated. The final step is to convert the semi automated version to a fully one by using of Genetic Algorithm (GA) approach for selecting the classification rules of extracting elements of sequence diagram from the natural language form statements of flow of events.

This tool had been implemented using C# programming language of visual studio that support embedding other software components, and has graphical facilities for drawing the sequence diagram and developing GUI for the tool.

The evaluation of the results has been handled using confusion matrix, in which the accuracy of the ISDG reach > 77%.

Keywords: Software Requirements, UML, Sequence Diagram, Natural Language Processing, Artificial Intelligence.