PhD Thesis

Title: **Using UTP and PVS for Formal Verification of Composition and Coordination Models**

Name : Muhammad Saqib Nawaz Khan

Student ID : 1501110042

Affiliation : School of Mathematical Sciences

Major : Applied Mathematics

Supervisor : Prof. Meng Sun

Research Division : Cyber-Physical Systems

2019.06

Copyright

摘要

在分布式系统与并发系统的开发中，由于系统本身的复杂性，它们往往被拆分为多个组件，而组件之间通过协议（Protocol）进行交互。本文提出了一个基于定理证明系统PVS并集协议建模、分析与推理于一体的框架，主要用于对分布式系统与并发系统中的协议进行形式化验证。

我们首先使用该框架对由协调模型Reo中的连接件所描述的交互协议进行建模与验证。在Reo中使用基本信道（channels）来构建连接件（connectors）。其中基本信道又包含不同的种类，可以分别描述不含时间约束的行为，包含时间约束的行为以及带有概率约束的行为。继而我们将该框架进行扩展，通过对组合算子的形式化描述，我们可以用该框架对云服务进行刻画，并对它们的性质、精化与等价关系进行证明。此外，我们还基于程序的统一理论（UTP）对遗传算法中的交叉算子和突变算子进行形式化建模，并证明了一些相关辅助性质及其等价关系。

该框架具有很强的表达能力，但同时存在一个主要问题：在进行证明时需要用户与PVS进行大量手工交互。为解决这个问题，需要设计更好的证明引导和证明搜索算法。我们在PVS中添加了一种基于序列模式挖掘（SPM）的证明引导算法。该算法通过对已有证明的分析来对当前证明目标进行提示。这项工作可以有效提高我们在交互式证明复杂连接件性质时的效率。其次，通过引入遗传算法中一系列交叉算子与变异算子，我们将证明引导算法进一步扩展为证明搜索算法，它能够在PVS库中对证明所需的相关定理和引理进行自动搜索。

关键词：统一程序设计理论, 原型验证系统, 并发, Reo, 云服务, 遗传算法, 交叉, 变异, 序列模式挖掘, 证明导览与搜索

Using UTP and PVS for Formal Verification of Composition and Coordination Models

Muhammad Saqib Nawaz Khan (Applied Mathematics)
Directed by Prof. Meng Sun

ABSTRACT

In this thesis, we present a unified framework for the modeling, analysis and reasoning about protocols that are explicitly used to bind different components in distributed and concurrent systems so that they can interact and communicate with each other. Prototype Verification System (PVS) that is based on higher-order logic is used for the framework development. The unified framework is used first to model and reason about component connectors in Reo coordination language. Untimed, timed and probabilistic/random channels are used in Reo as basic primitives with its own composition operators to construct complex connectors. The framework is then extended to model cloud services and their composition operators. Within the framework, different properties of component connectors and cloud services as well as the refinement/equivalence relation between them can be naturally formalized and proved in PVS. Moreover, design models based on unifying theories of programming (UTP) semantic framework is developed for operators (crossover and mutation) of genetic algorithms (GAs) and some crossover operators are later encoded in PVS, followed by proving their rudimentary properties and the equivalence relation between them.

The unified frameworks has one main limitation: Heavy user interaction is required with the PVS proof assistant in the proof development process, which makes the proof guidance and proof searching the two most desired properties in theorem provers. For that, a proof guidance approach based on sequential pattern mining (SPM) is proposed that can be used to guide the interactive proof development process. Moreover, the proof guidance approach is extended to provide an evolutionary proof searching approach, where a GA with different crossover and mutation operators is used to search and optimize the proofs for theorems and lemmas in PVS theories.

KEYWORDS: UTP, PVS, Concurrency, Reo, Cloud Services, Genetic Algorithms, Proof Guidance, Sequential Pattern Mining.

Table of Contents

Chapter 1 Introduction **1**

1.1 Background and Motivation 1

1.2 Unifying Theories of Programming 6

1.3 Concurrency 10

1.4 Coordination Language Reo 12

1.5 PVS 15

1.6 Contributions 18

1.7 Overview 20

Chapter 2 Formal Modeling and Verification of Component Connectors **21**

2.1 Introduction 21

2.2 Basic Definitions in PVS 22

2.3 Reo Channels and Operators 25

2.3.1 Primitive Channels 26

2.3.2 Operators Modeling in PVS 30

2.4 Reasoning about Connectors 31

2.5 Related Work 37

2.6 Summary 39

Chapter 3 Formalization of Probabilistic Reo Connectors **40**

3.1 Introduction 40

3.2 Basic Definitions in PVS 41

3.3 Probabilistic Channels and Operators 44

3.3.1 Random and Probabilistic Channels 44

3.3.2 Operators 48

3.4 Reasoning 49

3.4.1 Refinement and Equivalence 50

3.5 Related Work 54

3.6 Summary 55

Chapter 4 Modeling and Verifying Cloud Services and their Composition 57
4.1 Introduction 57
4.2 Basic Definitions 59
4.3 Cloud Services Modeling in PVS 61
4.4 Modeling of Cloud Service Composition Operators 67
4.4.1 Sequential Composition 67
4.4.2 Parallel Compositions and Conditional Choice 67
4.4.3 Merge and Router 68
4.4.4 Locate 68
4.5 Related Work 72
4.6 Summary 73

Chapter 5 Design Models for Genetic Algorithms Operators and its Encoding in PVS 75
5.1 Introduction 75
5.2 The UTP Observational Model 77
5.3 GA Operators as Designs 78
5.3.1 Crossover and Mutation Operators 79
5.3.2 Refinement and Equivalence 87
5.4 PVS Specifications 88
5.5 Equivalence Proof in PVS 93
5.6 Related Work 94
5.7 Summary 95

Chapter 6 Proof Guidance and Searching in PVS with SPM and GA 97
6.1 Introduction 97
6.2 Proof Corpus Mining with SPM 99
6.3 Experiments 106
6.3.1 Results and Discussion 106
6.4 Proof Searching with a GA 111
6.5 Results and Discussion 117
6.6 Related Work 122
6.7 Summary 124

Chapter 7 Conclusions **126**

7.1 Conclusion 126

7.2 Future Work 127

References **130**

List of Publications **140**

Dedication **142**

Acknowledgements **143**

Abbrevations **144**

北京大学学位论文原创性声明和使用授权说明 **145**

List of Figures

1.1 Generated state space for four concurrent algorithms 4
1.2 Some primitive channels in Reo 13
1.3 Operators for channel composition 14
1.4 Untimed connector (Take-Cue Regulator) and timed connector (Lower Bounded FIFO1) 14
1.5 Sum function in PVS and the proof tree for the theorem 16
1.6 The hypothetical model 19

2.1 Operators for channel composition 30
2.2 A connector composed of a Sync and a FIFO1 channel 30
2.3 Connector for take-cue regulator 34
2.4 Sequencer connector 34
2.5 Lower Bounded FIFO1 Connector 35
2.6 Expiring *FIFO1* connector 35
2.7 $2 \times t$ Timer Connector 37

3.1 Formalization infrastructure for reasoning about Reo probabilistic / random connectors 41
3.2 A probabilistic Reo connector 49
3.3 Connectors refinement example 51
3.4 Proof tree for connectors refinement proof 52
3.5 Connectors equivalence example 53

4.1 Flight Information Query Service 62
4.2 The PVS proof tree for theorem 4.3.1 64
4.3 Holiday Reservation Service 65
4.4 Proof steps for theorem 4.3.4 66
4.5 Remote Printing Service 69
4.6 PVS proof steps for lemma 4.4.1 71
4.7 PVS proof steps of theorem 4.4.2 72

5.1 Example of operators equivalence 88
5.2 PVS graphical proof tree for theorem 5.5.1 94

6.1 SPM-based approach to learn the proof process 100
6.2 Comparison of TKS and CM-SPAM with varying *k* 109
6.3 Sequential rules discovered in corpus 109
6.4 Flowchart of the GA for proofs searching and optimization in PVS 112
6.5 Time used by the GA to find the first ten matched *PS* 121
6.6 Total time and generations for the *PReo* theorem with GA 122

List of Tables

1.1 Concurrency in Reo 12
1.2 TDS semantics for Reo primitives 15
1.3 Comparison of ITPs 17

2.1 Formalized Reo channels in PVS 39

3.1 Formalized probabilistic / random Reo channels in PVS 55

4.1 Formalized compositional operators for cloud services 74

5.1 Formalized crossover and mutation operators 95

6.1 A sample of proof corpus 101
6.2 Conversion of proof sequences to integer sequences 102
6.3 Discovered proof steps with Apriori Algorithm 106
6.4 Extracted proof steps / patterns with TKS algorithm 107
6.5 Frequent proof patterns extracted with CM-SPAM 108
6.6 Results for TPC+ prediction model 110
6.7 Accuracy of prediction models 111
6.8 Results for different crossover and mutation operators of GA 118
6.9 Average total generation count for various crossover and mutation operators . 119
6.10 SD_t, mean and total time for GA with different combination of operators . . . 120
6.11 Mining techniques and respective algorithms in SPMF 125

List of Publications

1. M. Saqib Nawaz and Meng Sun. Using PVS for Modeling and Verification of Probabilistic Connectors. *8th International Conference on Fundamentals of Software Engineering (FSEN 2019)* (**To Appear**) (Chapter 3).
2. M. Saqib Nawaz, Meng Sun and Philippe Fournier-Viger. Proof Guidance in PVS with Sequential Pattern Mining. *8th International Conference on Fundamentals of Software Engineering (FSEN 2019)* (**To Appear**) (Chapter 6).
3. M. Saqib Nawaz and Meng Sun. Reo2PVS: Formal Specification and Verification of Component Connectors. In *Proceedings of 30th International Conference on Software Engineering and Knowledge Engineering (SEKE 2018)*, pages 391-396, KSI Research Inc. and Knowledge Systems Institute, 2018 (Chapter 2).
4. M. Saqib Nawaz and Meng Sun. Using PVS for Modeling and Verifying Cloud Services and Their Composition. In *Proceedings of 6th International Conference on Advanced Cloud and Big Data (CBD 2018)*, pages 42-47, IEEE, 2018 (Chapter 4).
5. M. Saqib Nawaz and Meng Sun. A Formal Design Model for Genetic Algorithms Operators and its Encoding in PVS. In *Proceedings of 2nd International Conference on Big Data and Internet of Things (BDIOT 2018)*, pages 186-190, ACM, 2018 (Chapter 5).
6. Hong Weijiang, M. Saqib Nawaz, Zhang Xiyue, Li Yi and Meng Sun. Using Coq for Formal Modeling and Verification of Timed Connectors. In *Proceedings of 15th International Conference on Software Engineering and Formal Methods (SEFM 2017) Collocated Workshops*, Revised Selected Papers, Vol. 10729 of LNCS, pages 558-573, Springer, 2018.
7. M. Saqib Nawaz, M. Ikram Ullah Lali and Meng Sun. Formal Modeling, Analysis and Verification of Black White Bakery Algorithm. In *Proceedings of 9th International Conference on Intelligent Human Machines Systems and Cybernetics (IHMSC 2017)*, pages 407-410, IEEE, 2017.

Submitted for Publication

1. M. Saqib Nawaz, Meng Sun, Basit Shahzad, M. IkramUllah Lali, Tariq Umer and Shaohua Wan. Quality of Service in IoT Protocol as Designs and its Verification in PVS.

Transactions on Emerging Telecommunications Technologies.

2. M. Saqib Nawaz, Moin Malik, Yi Li, Meng Sun and M. IkramUllah Lali. A Survey on Theorem Provers in Formal Methods. *Journal of Computer Science and Technology*.

Dedication

To my late mother who died before I started my PhD studies

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor, Prof. Meng Sun, for his continuous encouragement, support and guidance that enabled me to successfully complete my PhD studies. I was fortunate enough to work under his supervision in the Software Theory and Formal Methods (STFM) group. STFM members have always been very kind and helpful to me and I wish them success in their research work.

I am indebted to Dr. M Ikram Ullah Lali, who put me on this path at the first place. He monitored the progress of my work and was always available for discussion, on top of his insightful feedbacks at various stages. I would also like to thank Prof. Philippe Fournier-Viger for taking time out of his busy schedule to help me complete the last part of the thesis. Many thanks to Prof. Wim Hesselink and the active members of the PVS mailing list, particularly Sam Owre, for replying and providing answers to my questions.

There are a number of people at Peking who deserve special appreciation. I would like to thank my group members especially Yi Li, Yuanyi Ji and Xiyue Zhang for their help both in research and in academic matters that were in Chinese language. They spent a lot of time to assist my works. I have spent some quality and memorable time in Beijing with my fellow Pakistani students in PKU (Asif, Yousaf, Khurram, Wassem, Usman, Imran, Zeeshan, Hamid, Salman, Maqbool, Bashir, Ali) and friends from BIT (Umair, Saeed, Shahzad). Special thanks to Satria Sambijantoro for accepting our invitation to visit Pakistan and my hometown during the 2017 winter break.

Last but not least, I am eternally grateful to my family for their love and encouragement. I am very lucky to have a family that value and put higher education in high esteem. I must acknowledge my father, Prof. Muhammad Nawaz, whose devotion for knowledge and research has always been a great source of inspiration behind my academic success. I would also like to mention my brothers (Sajid, Shoaib and Zohaib). Zohaib showed his interest in my work as he is also working in the same area in his masters thesis and has a clear idea of my work.

I finally acknowledge that my PhD could not have been completed without the financial support from the Chinese Government Scholarship (CSC). I wish that CSC will continue their unwavering support for students in the future.

Abbrevations

PVS	Prototype Verification System
UTP	Unifying Theories of Programming
ITPs	Interactive Theorem Provers
ATPs	Automated Theorem Provers
GAs	Genetic Algorithms
SPM	Sequential Pattern Mining
HOL	Higher-Order Logic
FOL	First-Order Logic
TKS	Top-k Sequential
CPT	Compact Prediction Tree
ERMiner	Equivalence Class Based Sequential Rule Miner
NB	Naive Bayes
SPMF	Sequential Pattern Mining Framework