清华大学学位论文 LaTeX 模板
使用示例文档 v6.0.2

(申请清华大学工学硕士学位论文)

培 养 单 位 ： 软件学院

学　　科 ： 软件工程

研 究 生 ： 李　　兀

指 导 教 师 ： 顾明教授

二〇二〇年三月

An Introduction to LaTeX Thesis Template of Tsinghua University v6.0.2

Thesis Submitted to

Tsinghua University

in partial fulfillment of the requirement

for the degree of

Master of Science

in

Software Engineering

by

Li Wu

Thesis Supervisor: Professor Gu Ming

March, 2020

关于学位论文使用授权的说明

本人完全了解清华大学有关保留、使用学位论文的规定，即：

清华大学拥有在著作权法规定范围内学位论文的使用权，其中包括：（1）已获学位的研究生必须按学校规定提交学位论文，学校可以采用影印、缩印或其他复制手段保存研究生上交的学位论文；（2）为教学和科研目的，学校可以将公开的学位论文作为资料在图书馆、资料室等场所供校内师生阅读，或在校园网上供校内师生浏览部分内容。

本人保证遵守上述规定。

（保密的论文在解密后应遵守此规定）

作者签名：_____　　导师签名：_____

日　　期：_____　　日　　期：_____

摘 要

文章主要围绕 Range 那一套，写传统的静态分析方法

目的：整型缺陷分析：overflow, underflow, divide by zero

量化：Juliet、找一些工程

关键词：

Abstract

Key Words:

目 录

主要符号对照表

HPC	高性能计算 (High Performance Computing)

第 1 章　引言

1.1　研究背景与意义

突出整型缺陷会造成什么重大问题，解决这个问题十分重要。

【这里需要重新找例子，更切题一点的】

1.2　国内外研究现状

主要围绕整型缺陷这边来延伸与展开。

1.2.1　整型缺陷检测技术与工具

介绍目前识别整型缺陷的常用技术与工具，阐述它们的优缺点。

【这里可能需要额外做一些调研，然后列举在这上面】

1.2.2　抽象解释技术

论证程序正确性的方法最朴素的便是穷举程序所有可能的输入，并通过执行得到结果来判断其是否符合预期，如果运行结果符合预期，那么程序自然是正确的。然而，这种方法只是一种理论上可能的方法，在实际中，我们面对的程序输入的取值范围往往非常大，甚至无法穷举。以 C 语言的函数举例，若该函数有一个 int 型参数，由于 int 类型的表示范围是 $[-2147483648, +2147483647]$，那么单单一个 int 型参数就有 2^{32} 种取值，当输入的参数是字符串类型时，更是有无穷多中可能。因此，使用朴素的穷举来进行程序分析，其时间与空间代价在实际中是不可接受的。

当前常用的测试技术便是采用了上述思想，只不过测试技术所选择的输入是总输入空间的子集，通过边界条件分析等方法得到相对较少的输入空间。该方法的优点是大幅减少测试输入，但也带来了程序运行路径覆盖率低、需要人工参与测试输入样例的设计等问题。

相较于测试技术，抽象解释技术采用了不同的思路。抽象解释是一种对程序语义进行可靠抽象（近似）的通用理论[1]。与此同时，该理论为程序分析的设计与构建提供了一个通用的框架[2]。具体地，它是将程序语义进行不同程度的抽象，并将这种抽象及在其上的操作称为抽象域。通过将具体域中的值与抽象域中的值进行映射，从而将具体域中数量庞大甚至无穷大的取值域转化为抽象域中的有穷

的取值域。并将具体域上的操作对应到抽象域上的操作，通过在抽象域上计算程序的抽象不动点来表达程序的抽象语义。

单纯通过构建在抽象域上的与操作如迁移函数来进行建模有时并不能保证在程序的迭代分析中抽象域能快速到达不动点以获得抽象语义。因此在抽象分析中提供了加宽算子（widening），通过上近似理论来减少程序分析中的迭代次数，从而加速程序分析。由于上近似理论的可靠性，所有基于上近似抽象得到的性质，在源程序中必定成立。

抽象解释的核心问题是抽象域的设计，而如上所述，抽象解释是对程序语义的不同程度的抽象，这也就意味着抽象域并不唯一确定，针对特定问题可以设计使用特定抽象域以达到程序分析的效果。目前为止，已经出现了数十种面向不同性质的抽象域，其中，具有代表性的抽象域包括区间抽象域、八边形抽象域、多面体抽象域等数值抽象域[3]。另一方面，在开源领域出现了众多抽象域库，如 APRON[4]、ELINA[5]、PPL[6] 等。

抽象解释并不是一个已经研究成熟的课题，当下抽象解释仍然面临着很多挑战，主要包括两方面的内容：提高分析精度与拓展性。在提高分析精度方面，主要要解决的问题是基于加宽算子（widening）的不动点迭代运算的精度损失问题以及所设计的抽象域本身的表达能力具有局限性的问题。而在提高可拓展性方面，主要面临的问题是如何有效降低分析过程中抽象状态表示与计算的时空开销。

【阐述与本文的关系】

1.2.3　常用抽象域与区间抽象域

具体展开抽象域的研究，首先介绍抽象域在静态分析方法中的角色与作用，随后剖析各个抽象域的优缺点，重点介绍区间抽象域，它能解决什么问题，为什么它比较好。

【将上面的部分分一点儿到下面来】

1.2.4　总结

接上，阐述我们为什么要做区间抽象域，期望能达到什么样的一个目标，解决了传统区间抽象域的哪些痛点。

1.3　研究难点与挑战

难点大致在抽象域的设计方面：变迁规则、抽象方法（近似手段、合并操作等）

1.4　研究内容

这里搞一张图，到时候用这个说

1. 理论研究【暂未明确】
   (a) 基于程序解释的符号敏感的区间抽象域分析方法/基于线性空间的整型缺陷检测方法
   (b)【待讨论】基于二进制串的符号敏感的区间抽象域分析方法/基于环状区间的整型缺陷检测方法
   (c) 基于区间分析的数值导向型缺陷分析组合方法
2. 工具研发

1.5　研究方案

这里同样补一张图，对应于上面的研究内容。不用像开题报告那样分小章节说，直接一段话即可。

1.6　论文贡献

最后补上。

1.7　论文组织结构

最后补上。

第 2 章　基于线性空间的整型缺陷检测

交代背景；总领下面几个章节。

2.1　预备知识

写个引子，引领几个小标题

2.1.1　软件构建序列化抓取与预处理方法

因为不是自己主要做的，简单说一下原理与方法。

2.1.2　控制流自动机

主要是为了下文引用，属于知识铺垫。

2.2　基于线性空间的抽象域设计

2.2.1　Integer、Range、MultiRange、SignRange 抽象域的设计与实现

对几个抽象层逐个讲解。

2.2.2　相应变迁函数

对应展开。

2.3　基于线性空间的整型缺陷检测方法

这里介绍 checker 的原理与检测规则、生成 report 的方法。

2.4　模块实现

介绍具体实现。

2.5　实验过程与结果

这里介绍在 Juliet 与选取的几个项目上的测试结果。

2.6　本章小结

水。

第 3 章　基于环状区间的整型缺陷检测

为什么要使用环状区间、它能带来什么好处?
要解决的问题:

3.1　基于环状区间的抽象域设计与实现

【疑问】我觉得这里如果要写的话应该会被疑似抄袭那篇论文吧

3.2　基于环状区间的整型缺陷检测方法

类似于上一章，这里可能会重复阐述。

3.3　模块实现

介绍具体如何实现。

3.4　实验过程与结果

同样，在 Juliet 与项目上跑一跑。

3.5　本章小结

水。

第 4 章　基于值流图的整型变量关系分析与缺陷检测

描述一下整型变量关系分析是如何辅助帮助检测整型缺陷的。

4.1　值流图的原理与构造

仿光总

4.2　抽象域与变迁函数的设计

从论文里搞

4.3　模块实现

【待沟通】这里要不要实现？因为可能还要要求现场展示

4.4　实验过程与结果

【待沟通】

4.5　本章小结

水。

第 5 章　总结与展望

5.1　工作总结

5.2　研究展望

参考文献

[1] Cousot P, Cousot R. Abstract interpretation: a unified lattice model for static analysis of programs by construction or approximation of fixpoints // Proceedings of the 4th ACM SIGACT-SIGPLAN symposium on Principles of programming languages. 1977: 238-252.

[2] Cousot P, Cousot R. Systematic design of program analysis frameworks // Proceedings of the 6th ACM SIGACT-SIGPLAN symposium on Principles of programming languages. 1979: 269-282.

[3] 张健, 张超, 玄跻峰, 等. 程序分析研究进展. 软件学报, 2019, 30(01):80-109.

[4] Jeannet B, Miné A. Apron: A library of numerical abstract domains for static analysis // International Conference on Computer Aided Verification. Springer, 2009: 661-667.

[5] Singh G, Püschel M, Vechev M. A practical construction for decomposing numerical abstract domains. Proceedings of the ACM on Programming Languages, 2017, 2(POPL):1-28.

[6] Bagnara R, Hill P M, Zaffanella E. The parma polyhedra library: Toward a complete set of numerical abstractions for the analysis and verification of hardware and software systems. arXiv preprint cs/0612085, 2006.

[7] 薛瑞尼. ThuThesis: 清华大学学位论文模板 [EB/OL]. 2017[2019-04-27]. https://github.com/xueruini/thuthesis.

致 谢

感谢 LaTeX 和 ThuThesis[7]，帮我节省了不少时间。

声　明

本人郑重声明：所呈交的学位论文，是本人在导师指导下，独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本学位论文的研究成果不包含任何他人享有著作权的内容。对本论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明。

签　名：_____ 日　期：_____

个人简历、在学期间发表的学术论文与研究成果

个人简历

1995 年 3 月 6 日出生于内蒙古莫力达瓦自治旗。

2013 年 9 月考入大连理工大学软件学院软件工程专业，2017 年 7 月本科毕业并获得软件工程学士学位。

2017 年 9 月考研进入清华大学软件学院攻读软件工程硕士学位至今。

发表的学术论文

[1] Yang Y, Ren T L, Zhu Y P, et al. PMUTs for handwriting recognition. In press. (已被 Integrated Ferroelectrics 录用. SCI 源刊.)

研究成果

[1] 任天令, 杨轶, 朱一平, 等. 硅基铁电微声学传感器畴极化区域控制和电极连接的方法: 中国, CN1602118A. (中国专利公开号)