

基于影响域分析的软件补丁兼容性检测

(申请清华大学工程硕士学位论文)

培养单位: 软件学院
学 科: 软件工程
研 究 生: 潘 晓 梦
指 导 教 师: 贺 飞 副 教 授

二〇一五年五月

Thesis Submitted to
Tsinghua University
in partial fulfillment of the requirement
for the professional degree of
Master of Software Engineering

by
Pan Xiaomeng
(Software Engineering)

Thesis Supervisor : Professor He Fei

May, 2015

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

我还不知道写什么，就先这样吧。

关键词：

Abstract

The software system will continue evolving in its whole life time, and patch is a common way to accomplish the task to update or fix bug and so on. As the software evolving is rapid, how to identify whether a patch for a specific version of software is applicable for a newer version of the same software. We here present a so called patch compatibility analysis to try to fix this problem. It consists of three kind of analysis—program differencing which is used to output the structural difference of the two versions, change impact analysis which is used to analyze the impact of every change including dependency and impacted expressions/statements and so on. And at last, the conflict analysis used to determine whether the two patches' impact are conflictual.

Key words: T_EX; L_AT_EX; CJK; template; thesis

目 录

第 1 章 绪论	1
1.1 研究背景和意义	1
1.2 本文所要解决的问题与主要工作	2
1.2.1 应用场景	2
1.2.2 问题解答	2
1.2.3 主要工作	3
1.3 本文组织结构	4
第 2 章 相关工作	5
2.1 程序间差异分析	5
2.2 程序变更影响分析	5
2.3 相关工具	8
2.3.1 git	9
2.3.2 Beyond Compare	9
2.3.3 jpf-regression	9
2.3.4 ASTro	11
第 3 章 分析方法	12
3.1 问题定义	12
3.1.1 补丁版本迁移	13
3.1.2 语义影响范围分析	16
3.1.3 兼容性分析	17
3.2 解决方案	18
3.2.1 整体架构	18
3.2.2 补丁版本迁移	20
3.2.3 语义影响范围分析	21
3.2.3.1 程序差异性分析	21
3.2.3.2 变更影响分析	23
3.2.4 冲突分析	24
3.3 本章小结	26

目 录

第 4 章 工具设计与实现	27
4.1 整体架构	27
4.2 补丁版本迁移	27
4.3 语义影响范围分析	29
4.3.1 程序间差异分析	29
4.3.2 变更影响分析	32
4.3.2.1 分析流程	33
4.3.2.2 影响追踪	35
4.3.2.3 错误记录	36
4.3.2.4 大规模分析	37
4.3.2.5 Bug 修复	39
4.4 冲突分析	40
4.5 本章小结	40
第 5 章 实验结果与分析	41
5.1 实验设计	41
5.1.1 实验平台	41
5.1.2 实验数据	41
5.2 实验结果与分析	42
5.2.1 补丁版本迁移	42
5.2.2 语义影响范围分析	45
5.2.2.1 程序间差异性分析	45
5.2.2.2 变更影响分析	45
5.2.3 冲突分析	46
5.3 本章小结	47
第 6 章 结论	48
6.1 工作总结	48
6.2 未来工作	48
参考文献	49
致 谢	50
声 明	51

第1章 绪论

1.1 研究背景和意义

软件维护（Software Maintenance）是软件开发周期中耗时最长、开销最大的过程。随着外部环境和用户需求的不断变化，软件系统需要随之进行适应和调整，同时也需要修复在实际运行中暴露出来的问题。

在软件演进的过程中，软件系统可能会由于各种各样的原因而发生更新行为，这是软件维护周期中的常见活动。为了进行软件

补丁（Patch）就是这样一类可以用于完成修补程序漏洞、增强软件功能、改善程序性能等任务的程序。补丁在工业界中得到了广泛的实际应用，是软件维护过程的重要组成部分。

补丁一般是通过 Diff 工具而产生的文本数据，用于表示以行为单位的程序间差异性。一般而言，Diff 工具属于程序差别分析工具中的一种，单纯进行文本的差异性比较，并给出行级别的差异性信息。然而这样产生出来的差异性信息（Patch），只包含了文本差异，而失去了语法、语义等差异信息。

虽然得到了广泛应用，但补丁程序仍然具有一定的局限性，它一般只针对某个专门软件版本而开发，对于软件演进过程中获得的新版本而言，我们无法确定补丁程序是否也能适用于新版本的程序。然而补丁程序的应用一般都具有特定的目的，例如功能升级、漏洞修补等，新版本的程序中很可能仍然需要补丁程序的应用来完善自身。因而在实际的软件维护过程中工程师往往不得不重新去开发针对该新版本的补丁程序，对人力成本和时间资源造成了许多浪费。

究其缘由，主要在于补丁程序通常会引入软件变更（Software Change），使得软件源代码在应用这些变更后可以实现功能修复、版本升级等目的。使用 Diff 工具所产生的补丁文件一般会以行为单位引入变更。

然而，软件变更不止会局限于对被修改的某行产生作用，事实上，由于软件代码的耦合性，单行变更就足以将变更带来的影响广泛地传播到软件系统的其他部分中去。因此，软件变更不仅会使软件系统发生结构化的变更，还会产生语义上的变化。例如，对赋值语句的修改可能会影响到后续引用到该变量的条件语句。

因此，如何确定补丁程序对于软件其他版本的适用性就成为了一个较复杂的问题，这主要是由于补丁程序所引入的软件变更和版本升级所引入的软件变更可能是互斥的，或者说是部分互斥的。更准确的来讲，即变更所影响到的程序语法和语义层面之间的互斥性。

可见，要完全确定补丁程序对其他版本代码的兼容性是一件事实上比较困难的事情，目前就我们所知尚无这方面的相关工作出现。而且这方面的工作比较具有实用价值，能够解决工业界的实际问题，因而本文尝试去解决这样一个软件补丁兼容性检测的问题。

1.2 本文所要解决的问题与主要工作

软件系统总是处于不断演进的过程中的。补丁(patch)是一种常见的软件系统版本演进的方法，常用于修补漏洞和缺陷、改进性能、升级等。在实际应用中，补丁往往针对某一个特定版本的代码而设计，然而现代软件系统往往更新换代较快，因而面临着将补丁应用于其他版本的代码时可能失效的问题。为了解决这个问题，我们为此提出了进行补丁兼容性分析的想法。

为此，我们考虑这样一个问题，针对某软件系统 s ，假设已有版本 v_1 和 v_2 ，其中补丁 p_1 是针对版本 v_1 而设计，应用后能使其演进到版本 v_3 ，问 p_1 是否能够应用于版本 v_2 ，并保证不会发生冲突。

从该问题中可以发现， $p_1 = \text{diff}(v_1, v_3)$ ，设 $p_2 = \text{diff}(v_1, v_2)$ ，那么该问题就可以规约为，给定两个补丁 p_1 和 p_2 ，问二者是否会发生冲突。

其中 $\text{diff}(v_i, v_j)$ 函数的定义可以参考定义3.17。

1.2.1 应用场景

本文中的主要问题在于如何判定一个针对给定程序版本的补丁，它是否能够应用于其他版本，并且与其兼容？为了使读者对这一问题更加清晰，我们可以考虑这样一个应用场景，并且提出相应的前提假设：

- 某个团队正在对一软件进行开发工作，该软件已推出一个正式版 v_1 ，现在正在进行 v_2 版本的开发。
- 该团队使用版本控制系统进行版本控制，使用 Github 作为代码托管和团队协作工具。
- 该软件存在一个第三方开发的针对 v_1 版本的的补丁 p ，该补丁可以增强该软件的功能。

我们希望知道，该补丁 p 是否能够应用于正式版 v_2 ，并且补丁 p 中的变更不仅能够被正确地应用，还不影响软件版本 v_2 的功能正确性？

1.2.2 问题解答

事实上，对该问题而言，其答案可以分为多个层面来回答。

从语法角度出发，则该问题主要关注的是在应用过程中是否会造成语法结构上的错误，如果能够将补丁 p_1 成功应用于程序版本 v_2 ，则认为该补丁是可以兼容于新版本 v_2 的。语法结构上的错误可能是多种多样的，例如补丁中的要修改的代码不在原位置或已被删除、补丁中要添加的代码已经在该文件中存在等等，不一而足。

从语义角度出发，则单纯的语法兼容并不能够完全解答这个问题。某行修改可能会影响多处源代码，从而导致程序的行为发生变化。因而从语义层面而言，我们需要保证补丁 p_1 对程序版本 v_2 造成的语义影响不会波及到从版本 v_1 演进到 v_2 时所造成的语义变化。也就是说我们需要保证补丁 p_1 和 p_2 之间不会存在冲突——即互斥子集。

语法层面的回答很容易就能给出，现有的版本控制系统等都能在一定程度上给出相应的答案。而对于语义层面的答案来说，就目前所知尚未有这方面的工作。

因而本文将着重从语义层面去尝试解决这个问题。更具体的问题描述可以同样参考章节3.1。

1.2.3 主要工作

目前而言，本文的主要工作包括以下几个部分：

1. 补丁版本迁移。
2. 语义影响范围分析。
3. 补丁冲突分析。

所谓的补丁版本迁移工作是指，由于补丁 p_1 本来是适用于程序版本 v_1 的，如果想要适用于程序 v_2 ，可能需要进行一定的版本迁移工作。该部分工作可以从语法层面给出兼容性的答案，并为了实现补丁兼容性而进行一定的语法结构修正，主要使用版本控制系统和文本合并工具完成。

所谓的语义影响分析工作是指主要采用程序间差异分析、变更影响分析等手段来对界定补丁 p_i 对程序 v_i 所造成的语义影响范围。在有了补丁 p_1 和 p_2 的语义影响范围后，我们对其进行重叠判断，即可获得互斥的子集。该语义影响范围分析（Semantic Impacted Area Analysis）过程可以参考章节3.1.2中的详细叙述。

而在有了语义影响范围之后，就可以进行冲突分析的工作。通过影响范围重叠的判断，我们可以获得两个补丁间的互斥子集，通过人工分析的辅助，我们可以判定该互斥子集是否存在误报（false positive），并筛选出真正的互斥子集。

我们可以简单的认定如果两个补丁间如果存在真互斥子集，那么他们之间就存在冲突。而实际上，真互斥子集中的一些互斥变更，根据我们的定义他们是互斥的，但在实际应用中他们其实仍然可能是互相兼容的。

如果对真互斥子集进行人工的冲突分析，可以将这些情况进一步剔除。

可见，我们给出的冲突定义是比实际冲突更加严格的定义，可能会造成一定的过高估计（over-estimate）。在第5章的实验中我们会给出具体的介绍。

1.3 本文组织结构

本文主要包括六个章节，第一章是绪论，介绍本文的研究背景和主要工作；第二章主要介绍与本文所述内容相关的国内外的工作；第三章主要介绍了补丁兼容性分析的具体方法；第四章主要介绍了如何将各阶段的不同分析方法进行整合，形成具体的流程；第五章介绍了实验过程和结果；第六章主要对本文的工作进行了总结，并提出了进一步的工作方向。

第 2 章 相关工作

2.1 程序间差异分析

对于软件演进分析而言，如何确定一个程序的不同版本之间的变更是一个关键性的问题。程序间差异分析能够通过分析同一程序的不同版本间的差异，来确定版本间的变更集合。

按分析的深度而言，程序间差异分析可以分为三类：

- 文本差异：单纯的对比文本间的不同，这是最简单也最广泛应用的分析方法，如 Unix Diff 工具。
- 语法差异：对比并获得源代码间语法结构上的不同。
- 语义差异：对比并获得源代码间语义层面上的不同。

现有的帮助工程师进行软件维护和演进活动的工具往往都受限于低质量的变更信息。例如现在源代码的变更信息，往往都存储于版本管理系统中，如 CVS 等。他们会追踪对某个特定文件的文本行的增加/删除等操作，并没有考虑代码中的结构化变更。

考虑到源代码可以抽象语法树（Abstract Syntax Tree）的形式进行表达，可以考虑采用树间差异分析的方法来抽取出这些变更信息。Change Distilling 就是这样一类进行树间差异分析的算法。该算法能够从两棵 AST 之间寻找匹配节点，并找到一个能够令一棵树转化为另一棵树的最小变更集合，该变更集合即为所求的程序间差异。由于是从 AST 的实体和语句中抽取信息，该算法可以获取结构化的变更信息。

Change Distilling 中采用二元字符串相似性来匹配源代码语句，并使用子树相似性来匹配源代码结构（如语句、循环等）。在寻找变更集合时，它采用基本的树变更操作来描述源代码的变更，包括更新、删除、增加等。在实际的使用过程中，该算法可能会受限于如何找到合适数量的移动操作。

2.2 程序变更影响分析

软件维护是软件开发周期中最复杂、高成本和劳动密集的活动。软件产品天然的需要跟随系统需求的变更而进行适应和变化，以满足用户的需求。软件变更 (software change) 是软件维护 (software maintainace) 中的基础组件。变更可能从新的需求、缺陷修复、变更请求等而来，当变更应用于软件时，他们不可避免的会

带来一些副作用，也可能会与原软件的其他部分产生不协调。

而 CIA 正是这样一类用于确定变更对于软件其他部分的影响的技术集合，它在软件开发、维护和回归测试等过程中都起到着重要的作用。一般而言，CIA 可以用于程序理解、变更影响预测、影响追踪、变更传播、测试用例的选取等。

CIA 方法的分类：

- 基于追踪性的 CIA，追踪两个不同抽象级别上的元素间的依赖性，目的在于链接不同类型的软件工件（需求、设计等）
- 基于依赖的 CIA，致力于衡量变更的潜在影响，通常尝试去分析程序的语法关系，他们表示了程序实体间的一些语义依赖关系，或者说他们在同一个抽象级别上对软件工件实施了影响分析。现在最主要的这类 CIA 主要在源代码级别上进行研究。

软件系统上的变更可能导致不可意料的副作用，CIA 的目标就在于辨认这些副作用 (side effect 或者 ripple effect)，并防止之。CIA 从分析变更请求和源代码开始，以便获得变更集合 (change set)。最后实际获得的影响集合叫做 EIS (estimated impact set)，而真实的集合叫做 AIS (actual impact set)。

整个软件变更影响的分析过程可以大致分为以下流程，更详细的流程可以参考图 2.3。

1. 确定变更集合，需要针对变更请求进行分析，这步一般叫做特征定位 (feature location)，用于确认源代码中实现了对应功能的起始位置。
2. 衡量由变更集合引入的影响。这一步是目前大部分 CIA 技术的着重点。若按分析技术的类型划分，则主要的两类包括静态分析和动态分析。
 - 静态分析：历史分析、文本分析、结构分析。静态分析通过分析程序的语法、语义或者历史依赖实现，通常会产生很多错报 (false positive)。
 - 结构分析着重于分析程序的结构依赖性，构建依赖关系图
 - 文本分析根据注释和标识符提取出概念依赖性
 - 历史分析从多个软件演进版本中挖掘出信息

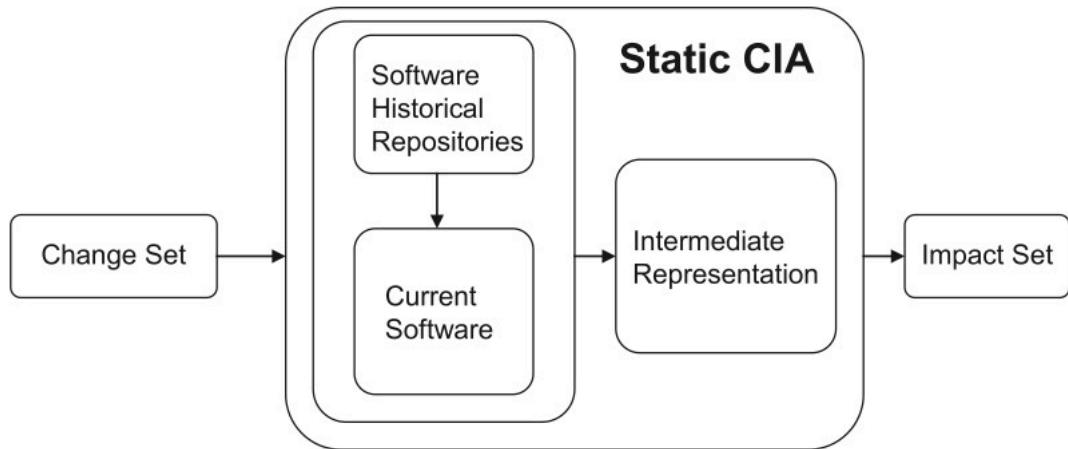


图 2.1 静态分析过程

- 动态分析：在线分析和离线分析。考虑特定的输入，并且依赖于程序运行时收集到的信息进行分析（如运行时的追踪信息、覆盖信息等）。其得到的影响集合往往比静态分析京都更高，但其开销也相应更大，而且通常包括误报（false negative）。
 - 在线 CIA 使用程序运行时收集的信息实施所有的分析

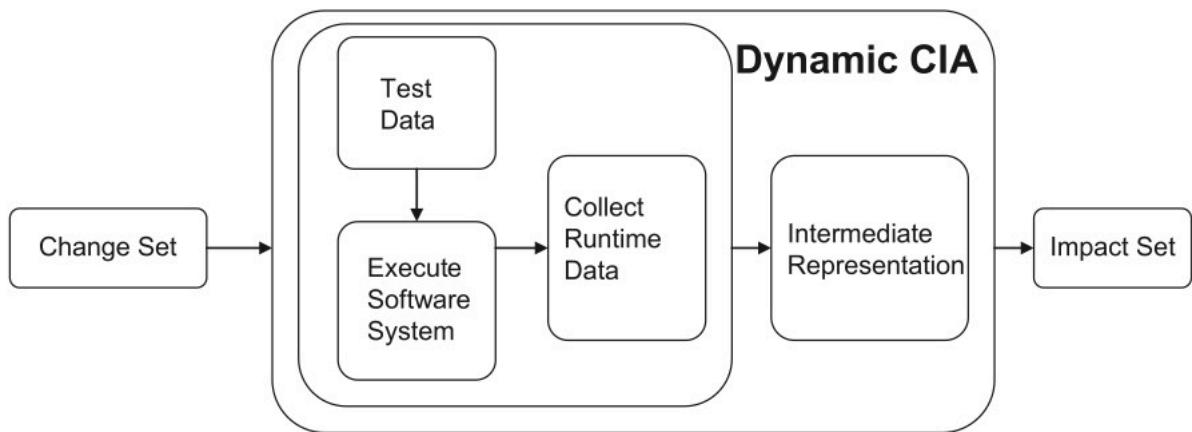


图 2.2 动态分析过程

近年来，有一些以 CIA 技术作为支撑而实现的工具，他们一般在实现 CIA 技术之后，利用得到的中间结果进行后续分析，帮助进行软件维护和演进。下面将对他们进行简要的介绍：

1. Chianti：应用于 Java 程序，Eclipse 插件

应用场景：

- (a) 使用回归测试来看变更以前的功能是否能正常使用
- (b) 若失败了，利用 Chianti 进行变更影响分析，将 change 拆分成若干原子

change，并分析他们之间的依赖关系，结合原程序生成中间表示形式，看可能是哪些部分影响到了 test 的运作

网址：<http://aria.cs.vt.edu/projects/chianti/> Ren X, Shah F, Tip F, et al. Chianti: a tool for change impact analysis of java programs[C]//ACM Sigplan Notices. ACM, 2004, 39(10): 432-448.

2. JRipples: 应用于 Java 程序, Eclipse 插件

应用场景：利用 dependency graph 等，自动标注可能受到被变更 class 影响的其他 class，并展示出对应的变更传播路径。可利用人工标注来修正结果网址：<http://jripples.sourceforge.net/>

3. ROSE: 应用于 Java CVS 工程, Eclipse 插件

应用场景：挖掘软件代码仓库，当用户对代码变更时提示用户可能相关的变更（类似于“变更了这个函数的人通常还变更了另一个函数”）网址：<https://www.st.cs.uni-saarland.de/softevo/erose/>

4. jpf-regression: 应用于 Java 程序, Eclipse 插件/命令行工具

应用场景：利用程序切片技术进行 change impact analysis，利用得到的 impact sets 来驱动 symbolic execution，获取被改变部分代码的行为（path condition）

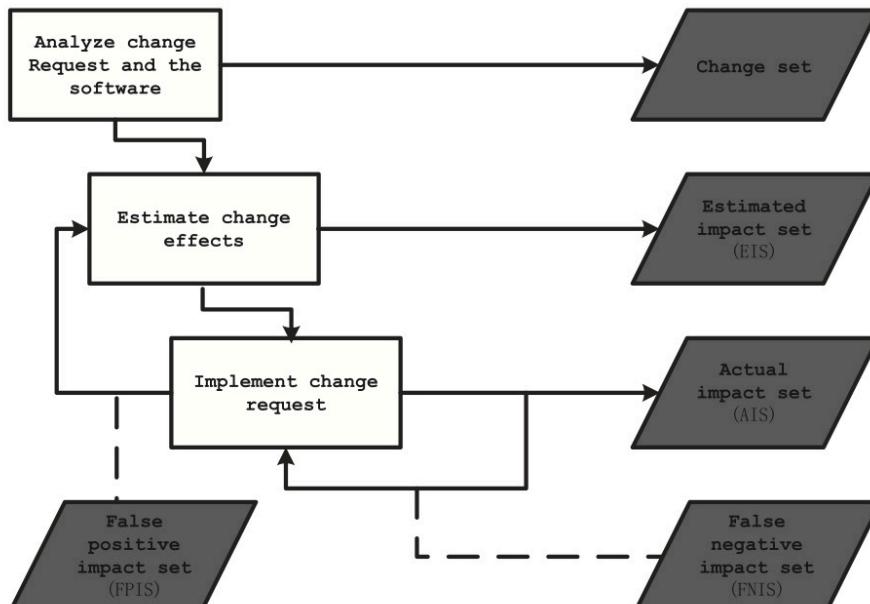


图 2.3 CIA 过程

2.3 相关工具

本节将主要介绍本文中采用到的相关工具。

2.3.1 git

git 是一个分布式的版本控制系统，最初由 Linus Torvalds 在 2005 年为 Linux 内核而开发，现在已经成为最流行的版本控制系统。

与 CVS 和 SVN 等集中式的 C/S 版本控制系统不同，git 是分布式的版本库，每个本地的 git 工作目录都包含了完整的历史数据和版本追踪能力，无需网络连接或服务器端。

本文中主要考虑以 git 作为版本管理系统的应用场景，类似于 GitHub，假定为项目开发了 new version 和 patch version 两个不同的分支，并使用 git 的分支合并策略实现补丁的版本迁移过程。

2.3.2 Beyond Compare

Beyond Compare 是一款内容比较工具，可以用于文件、目录、压缩包的比较，横跨 Windows、Mac OS X、Linux 三大操作系统，可用作版本控制系统的文本比较和合并工具，例如 git。

本文中主要采用其作为 git 的文本比较和合并工具，用于解决补丁版本迁移时的冲突（conflict）问题。

2.3.3 jpf-regression

变更影响分析常用于衡量软件变更的潜在影响。CIA 的结果通常可用做其他程序分析技术的输入，例如回归测试可利用 CIA 来确定程序的哪些部分需要进行再分析。而由于软件开发过程的演进特性，并且软件系统的大小和复杂性越来越高，刺激了人们对于有效的变更影响分析的需求。

一般而言，“单行变更”就足以引发广泛而未知的影响，因而变更影响分析在软件演进和维护过程中扮演着重要的作用。

目前，大部分的自动分析工具都将变更的影响以程序句法结构的形式表现出来，例如函数和程序语句。基于依赖的分析方法则通过分析程序组件间的内部关系来衡量变更的影响。这类技术能够以程序位置信息来描述变更的影响，然而却缺失了与受影响位置相关的程序运行路径的信息，而这类信息往往对于程序行为的验证、调试等起到很大的作用，而且能够将关注的范围缩小，只关注受到影响的程序行为集合。

（Directed Incremental Symbolic Execution）DiSE 能够结合静态分析的效率和符号化执行的精度，分析方法内部的变更影响，并描述程序变更对其运行行为的影响。

DiSE 的静态分析部分采用程序切片技术来衡量变更对代码中其他位置的影响，其生成的影响集合可以用于引导符号化执行去分析受到影响的程序行为，并生成对应的路径条件（path condition），这些路径条件用于描述受到影响的程序行为。这些路径条件之后可以利用 SMT 等技术进行求解，用于后续的验证、调试等过程。

在 DiSE 方法中，程序变更影响分析是其进行后续分析过程的基础，它的变更影响分析过程主要特点在于：

1. 粒度：语句级别。即具体的影响分析过程中，以程序语句为基本单位进行影响集合的计算。
2. 影响范围：方法内部。即以方法的作用域作为单次影响集合并计算的限制范围，最后得到变更对其所处的方法内部其他语句所造成的影响。也就是说不计算变更对于其他方法的影响，即不考虑对方法内部函数调用的影响。
3. 影响来源：主要从控制流和数据流两个层面考虑变更所造成的影响，实际上采用了语句间的控制依赖和数据依赖关系进行计算。

DiSE 中，受影响的集合主要分为两类：

- ACN: 受影响的条件语句节点（affected conditional nodes）。
- AWN: 受影响的赋值语句节点（affected write nodes）。

为了计算得到这两类影响集合，DiSE 中一共使用了四条规则来迭代计算，从最初的变更集合出发，不断应用规则向外扩展，最后计算得到闭包，即为所求的影响集合：

1. 如果 ACN 中有一个节点 ni ，且存在一个条件节点 nj ，其中 nj 控制依赖于 ni ，那么将 nj 加入到 ACN 中。
2. 如果 nj 是一条赋值语句，且控制依赖于 ACN 中的节点 ni ，那么 nj 加入到 AWN 中。

前两条公式表明，如果条件语句和赋值语句控制依赖于变更后的 CFG 中的节点，那么他们就应当被加进来。

3. 如果 AWN 中的一条赋值语句节点 ni 对于变量的赋值在条件语句 nj 中被使用了，且 CFG 中存在一条从 ni 到 nj 的路径，那么将 nj 加入到 ACN 中。
4. 如果一个写语句节点 ni 对于变量的赋值在 AWN 或 ACN 中的某个节点 nj 被使用了，且 CFG 中存在一条从 ni 到 nj 的路径，那么将 ni 加入到 AWN 中。

jpf-regression 是 DiSE 的工具实现，它是基于 Java Path Finder 框架而实现的工具，可作为 Eclipse 的插件使用，支持 Java 语言。

2.3.4 ASTro

本文中所采用的 AST 比较工具是由内布拉斯加大学林肯分校的 Josh Reed, Suzette Person 和 Sebastian Elbaum 等人所开发的 ASTro, 它也是 jpf-regression 中使用的前置工具, 用于对比两个不同版本的源代码, 并获取其抽象语法树上的差异, 以 XML 文件的格式进行输出。该工具支持 Java 语言。

第3章 分析方法

3.1 问题定义

第一章中对本文所要解决的问题进行了简要介绍。下面对该问题进行正式的定义。

问题 3.1：给定针对某同一版本代码 v_k 的补丁 p_i 和 p_j ，问这两个补丁对于该版本代码的影响是否会产生冲突？

为了回答这个问题，我们需要明确的知道到底什么是补丁对代码的影响，以及什么是冲突？

本文首先讨论什么是影响。

定义 3.1：影响 (Impact)——假如某个代码结构对其他代码结构存在依赖关系，那么就说这两者互相影响。其中代码结构指程序的语法结构，包括类、方法、语句等不同的类型。

事实上，如果给出不同的依赖关系的定义，那么就会得到不同类型的影响。可见，这里所谓的影响即是程序代码间的耦合关系。常见的依赖关系包括控制依赖和数据依赖等。

因此，这里可以给出影响的形式化定义。

定义 3.2：Code——指源代码。

假设存在某个版本的源代码 v_i ，即 $\exists v_i \subset Code$

定义 3.3： $Structure \rightarrow Class | Method | Statement | \dots$

定义 3.4： $Struct : Code \mapsto \{Structure\}$ 。Struct 函数定义为一个从源代码到该源代码的所有代码结构集合的映射。

定义 3.5： $im : Structure \mapsto Structure$ 。影响关系定义为一个从代码结构到代码结构的映射关系。

定义 3.6： $depend : Structure \mapsto Structure$ 。依赖关系定义为一个从代码结构到代码结构的映射关系。

定义 3.7: $\forall structure_i, structure_j \subset Struct(v_i), im(structure_i) = structure_j \iff depend(structure_i) = structure_j.$

考虑到补丁间不兼容的实质是因为在应用时和应用后出现了错误，可以具体将错误分为两类：

定义 3.8: 兼容性错误——将补丁 p_i 和补丁 p_j 先后应用于同一版本代码 v_k 时和之后所出现的程序错误。主要分为两类：

- 语法错误：即在应用一补丁 p_i 之后，再次应用另一补丁 p_j 时出现语法结构上的错误。易由编译器发现。
- 语义错误：将补丁 p_i 中的变更和补丁 p_j 中的变更应用于源代码之后，由于语义上的变更互相冲突而导致出现了功能上的错误，较难发现。

因而，可见为了判定补丁间是否相容，我们主要需要判定两个补丁所造成的变更之间是否存在语义上的冲突。

定义 3.9: 语义冲突 (Semantic Conflict)——指两个补丁所对应的不同变更集合之间存在某两条互斥的变更。其中两条变更互斥指二者对同一行代码进行了修改，或者他们所修改了不同行代码，但其影响传播到了某行相同的代码。

上述定义是对冲突的非形式化的描述，为了更清楚的理解，可以给出如下的形式化定义。

定义 3.10: 形式化定义

$$\exists c_i \subset p_1 \exists c_j \subset p_j$$

在实际的工作中，我们发现该问题可以拆分为若干个子问题。因而我们可以采用分治的办法去逐步解决这些子问题，最后将其组合成一套解决方案。下面将对这些子问题及其解决方法分别进行阐述。

3.1.1 补丁版本迁移

补丁程序是专门为某一版本的软件系统所设计，因而往往不能直接应用到其他版本的代码上，而且强行应用往往会出现很多错误。常见的问题可能包括：

1. 补丁程序中所提及的行不在原位置。
2. 补丁程序中所要删除的行其实已被删除，或者所要添加的行其实已被添加。
3. 补丁程序中所提及的文件未找到。

由于有这些可能的问题存在，导致我们无法直接将补丁程序应用到其他版本的补丁上。在这里我们先不考虑应用之后可能出现的语义错误，而尝试去解决可能出现的语法错误。因而这个问题可以如下定义之。

问题 3.2： 由于补丁 p 针对版本 v_1 设计，当尝试将其应用到版本 v_2 上时，可能会出现程序语法结构上的错误。如何实现补丁的版本迁移，将 p 成功应用到版本 v_2 上，并且不会出现程序语法错误？

为了解决这个问题，我们实际上可以采用版本控制系统的版本合并功能。

版本控制系统中经常需要将其他分支中的版本合并到当前分支中，为了解决不同版本之间可能存在的语法冲突，通常会采用三路归并算法（3-way merge）来对两个不同版本的代码进行合并，合并时以二者共同的祖先版本作为依据进行。

而我们所面临的问题也是类似的，由于 $patch(p_1, v_1) = v_3$ ，为了将 p_1 应用到版本 v_2 上获得版本 v_4 ，我们可以采用类似的想法，将版本 v_3 和版本 v_2 进行合并即可。合并后我们就等于同时拥有了 $p_1 = diff(v_1, v_3)$ 和 $p_2 = diff(v_1, v_2)$ 这两个补丁中的变更。其中 $patch(p_i, v_j)$ 函数的定义参见定义3.21。

这实际上也是现在软件开发过程中的常见手段。例如使用 git 作为版本控制系统时，为了修复某个功能性的 bug，我们可以新建一个分支 FixBug，然后切换到该分支进行漏洞修复，等到修补完毕后，再将该分支合并到主分支 Master 中即可。

该过程如图3.1所示。

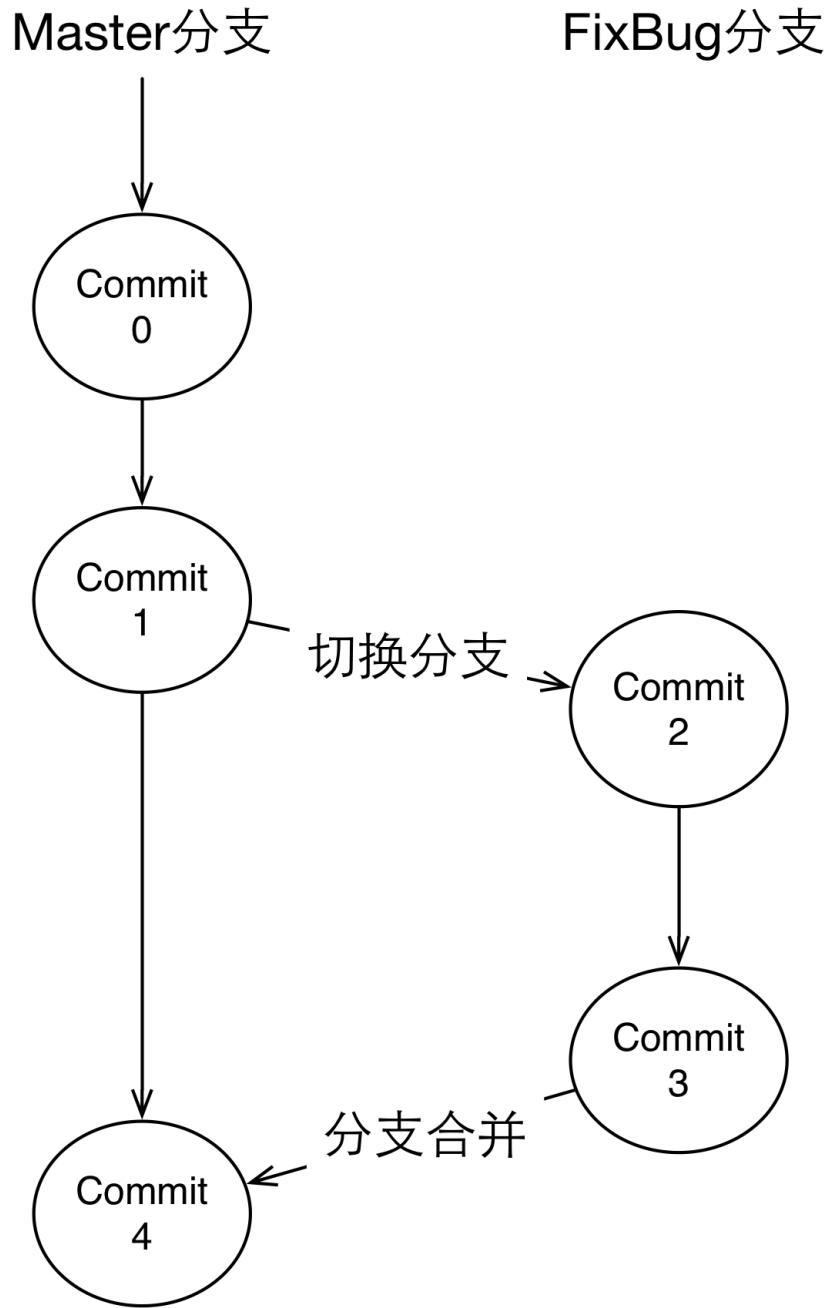


图 3.1 git 分支切换与合并

因而，我们可以将补丁版本迁移过程定义如下。

定义 3.11： 补丁版本迁移——将原本为软件版本 v_1 所设计的补丁 p 修改为可以在另一软件版本 v_2 上使用的补丁，使得应用补丁之后不会出现程序语法结构错误。

3.1.2 语义影响范围分析

为了确定两个补丁之间是否存在互斥子集，我们需要从语义的角度入手。根据本文对互斥的定义，我们可以发现，我们主要需要从代码中挖掘中两类信息：

1. 代码的变更集合，即两个版本之间代码的差异性。
2. 变更的影响集合，即变更集合对代码中其他哪些部分会造成影响。

在有了这两类集合之后，我们就可以界定出软件变更在语义层面上对代码的影响范围，我们可以将这整个过程称为语义影响范围分析（Semantic Impacted Area Analysis）。之后，我们可以通过比较两个补丁对代码的影响范围是否存在重叠来判定补丁间的互斥性。

因而，本节中的子问题主要在于如何界定软件变更对于程序的影响范围。

定义 3.12： 软件变更（Software Change）——对软件代码做出的修改。其中变更可以划分为三类不同的表示方法：

- 文本：只考虑程序代码在文本行上的变化。
- 语法：考虑程序代码在语法结构上的变化。
- 语义：考虑程序代码在语义层面上发生的变化。

定义 3.13： 软件变更集合（Software Change Set）——补丁 p 中所涉及到的软件变更的集合。

定义 3.14： 程序结构——指程序的基础语法结构，包括类、方法、基本块、语句等不同级别。

定义 3.15： 语义影响范围（Sematic Impacted Area）——在程序的某个限定范围内，直接或间接受到变更的语义影响的程序结构集合。

问题 3.3： 给定一个程序和他的变更集合，如何界定该变更集合对于其他程序结构的语义影响范围？

事实上，可以将程序中受变更影响的部分划分为不同的粒度，从而获得不同程度的影响，我们考虑对面向对象的编程方法中的影响元素级别进行划分：

1. 类：探讨变更对于其他类和对象的影响。对于面向对象的程序设计方法而言，这是最高级别的粒度。
2. 方法：探讨变更对于其他方法的影响。
3. 基本块：探讨变更对于其他基本块的影响。
4. 语句：探讨变更对于其他语句的影响。

而所谓的影响范围也需要界定其粒度，不同层级的粒度显然会对语义影响范围分析的精度产生影响。

1. 类间：考虑变更的影响可能延伸到其他类、对象。
2. 方法间：考虑变更的影响可能延伸到其他方法内部。
3. 方法内部：考虑变更的影响只在本方法的内部延伸。

在实际情况中，不同级别的影响范围均可分别采用不同的影响元素级别。

通过语义影响范围分析，我们就可以从代码中挖掘出所需要的变更影响信息，从而可以进行后续的兼容性分析工作。

在实际应用中，语义影响范围分析主要由两类分析过程组成：

1. 程序间差异性分析：获取两个软件版本间的差异性，获得结构化的软件变更信息。该分析过程即为 $diff(v_i, v_j)$ 函数的具体实现。
2. 变更影响分析：分析软件变更对软件其他部分是否存在影响，并找到受影响的集合。该分析过程即为 $impact(p_i, v_i)$ 函数的具体实现。

我们可以将整个分析过程定义为如下的函数：

定义 3.16： $s = ia(v_i, v_j) = impact(diff(v_i, v_j), v_i), i \in \mathbb{N}, j \in \mathbb{N}$

定义 3.17： $p = diff(v_i, v_j)$ ——求两个不同版本的程序 v_i 和 v_j 之间的差异性，其结果即为补丁 p。

定义 3.18： $s = impact(p_i, v_j)$ ——求补丁 p_i 对代码版本 v_j 的影响，其结果即为影响集合 s。

3.1.3 兼容性分析

我们最后还是要落实到补丁兼容性的判定问题上来。在有了前面两个子问题的铺垫之后，我们就可以着手进行核心的补丁兼容性分析过程。因而该子问题可以定义如下。

问题 3.4： 给定两个补丁和他们对代码的语义影响范围，问这两个补丁是否兼容，即不存在语义上的冲突？冲突的定义参见定义 3.9。

在有了软件变更的语义影响范围之后，我们可以通过简单地比较两个补丁对代码的影响范围是否存在重叠来判定其互斥性。

定义 3.19： 语义影响范围的重叠——即两个语义影响范围间是否存在交集，如果存在交集，则重叠，反之亦然。

显然，如果两个补丁对代码的影响范围不存在重叠，那么他们两者之间自然是兼容的，因为他们不仅本身的变更互不干涉，并且他们所影响到的程序结构也互不干涉。在这样的情况下，补丁是可以成功应用到其他版本上的代码，并且可以完成补丁本身的目标的。

如果影响范围发生了重叠，那么我们就可以认为补丁之间是不相容的，因为补丁所作的变更会对相同的程序结构产生影响。实际上这样的影响是否能够兼容需要人工的判定，因为在不同的情况下，他们可能是兼容的，也可能是不兼容的。而我们无法从代码中获取到足够的信息来进行这样的判定，需要人类对于变更的期望信息作为辅助。

我们可以将这个过程定义为如下类型的函数：

定义 3.20: $isCompatible(\text{diff}(v_2, v_1), \text{diff}(v_2, v_4)) :: Patch \rightarrow Patch \rightarrow Boolean$

该函数接受两个补丁作为参数，并返回一个布尔类型的判定结果。

3.2 解决方案

考虑上两节中所提到的应用场景和子问题划分，我们可以给出一个通用的解决方案来解决整个兼容性问题的判定。

3.2.1 整体架构

如前所述，为了解决该问题，本文中提出了一套通用的补丁兼容性分析解决方案。该解决方案的整体架构如图3.2所示。

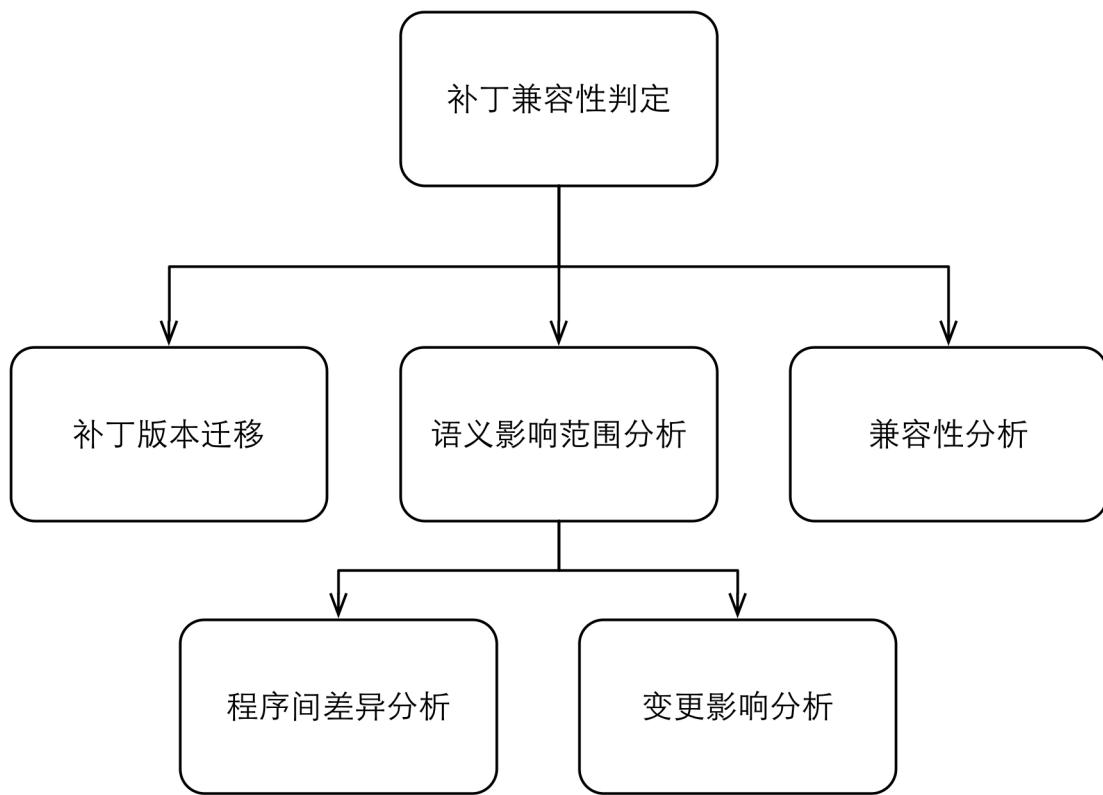


图 3.2 整体架构

可见，问题3.1的解决主要需要采用对应的解决方法分别解决三个子问题，最后将再将其组合起来即可，是一个应用了分治法的解决方案。

整个解决方案的流程可以参考图3.3。

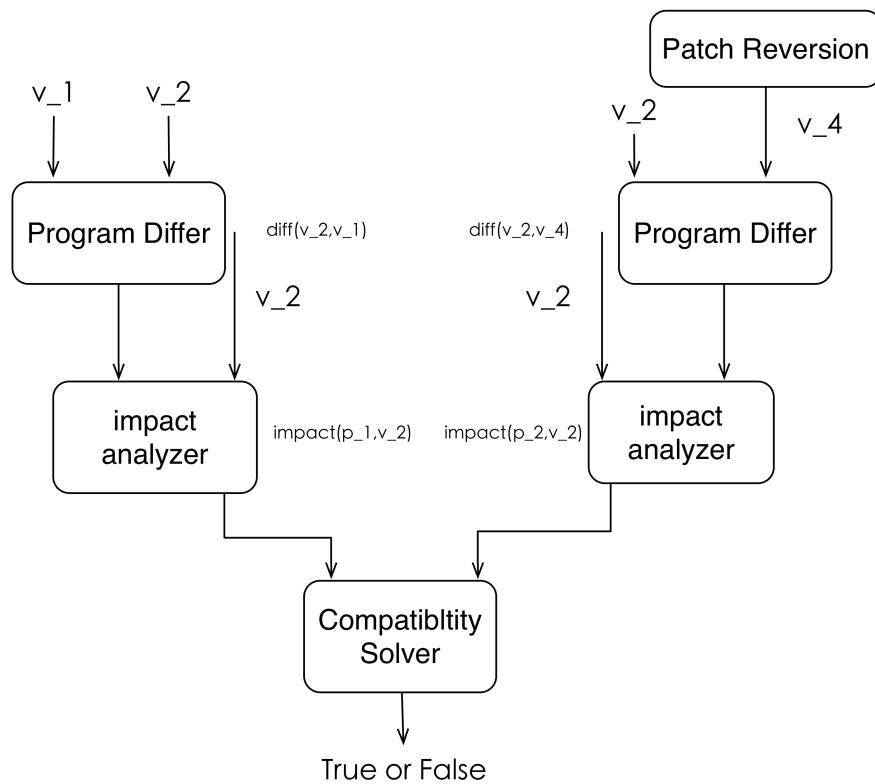


图 3.3 解决方案

其整体流程可以简述如下：

1. 采用版本控制系统进行代码版本管理，并进行补丁版本迁移，得到应用于新版本的补丁后代码版本 v_4 。
2. 根据得到的三个版本代码 v_1 、 v_2 、 v_4 ，分别分析其程序间差异性，生成对应的程序变更集合。
3. 根据得到的程序变更集合，进行不同版本间的变更影响分析，生成程序影响集合。
4. 根据得到的语义影响范围，进行冲突分析。

可见，本文中给出了一个通用的补丁兼容性分析解决方案，可以通过更改相应的分析模块来提高解决方案的普遍适用性。解决方案中所涉及到的各个分析模块都可以更改其核心算法，采取不同的分析算法来完成相应的任务。

3.2.2 补丁版本迁移

如问题3.2中所述，我们可以采用常见的版本控制工具提供的三路归并算法进行补丁的版本迁移工作。具体来讲，其流程可以简述如下：

1. 采用 git 进行代码的版本管理，在主分支 master 中提交代码版本 v_1 。

2. 切换到分支 new，并提交新版本代码 v_2 。
3. 将补丁 p_1 应用到版本 v_1 ，获得旧版本应用补丁后的代码，即其版本为 $v_3 = patch(p_1, v_1)$ 。
4. 切换到分支 patch，并提交代码版本 v_3 。
5. 切换回分支 new，将分支 patch 进行入分支 new，获得新版本应用补丁后的代码，其版本为 v_4 。
6. 解决分支合并中可能出现的冲突问题。

定义 3.21： $patch(p_i, v_j)$ ——将补丁 p_i 应用于版本 v_j 。

其中，通过三路归并算法进行的合并工作中可能会出现冲突，这说明版本 v_2 和 v_3 之间存在语法冲突，即这两个版本在语法层面上不兼容。我们可以通过人工修改的方式进行修复，实现语法层面上的兼容性。

在解决冲突的过程中可以采用第三方的合并工具，利用现成工具的高效算法快速解决冲突。

3.2.3 语义影响范围分析

如问题3.3所述，我们需要进行语义影响范围分析来获取两个补丁的影响范围，用于后续分析使用。

实际上，语义影响范围分析主要分为两个分析过程，即程序间差异分析和变更影响分析，通过这两个分析的协作来完成整个语义影响范围的分析。

3.2.3.1 程序差异性分析

程序差异性分析主要用于分析两个不同版本的程序间的差异性，其结果即我们所需要的程序变更集合。近年来程序差异性分析方面有不少工作，实现了一些较为成熟的比较工具，我们可以使用这些工具来实现该分析过程。

在本文的组合架构中，程序差异性分析的主要任务是接受两个不同版本的源代码，并返回代码间的结构化差异信息。结构化的差异信息可以视为补丁的一种，只不过它具有比常见的采用 Unix diff 工具生成的 .patch 类型的补丁文件更丰富的信息，能够描述以程序语法结构的形式对软件变更进行描述。

该分析过程应该满足如下需要：

- 输入为两个不同版本的代码。
- 输出为源代码间的软件变更集合。
- 每条变更描述语句或基本块级别的变更。
- 每条变更描述新旧程序结构的相关信息和其关联关系。

- 每条变更描述了其所属的作用域。

选择这样的分析结果类型是为了后续分析过程的方便，因为变更影响分析需要我们提供软件变更集合作为输入，而 *.patch* 类型的补丁文件只描述文本行的变更，不包含语法信息，我们无法从中提取出所需的语法层面的变更信息。

一种比较好的选择是采用 AST 差异性分析，因为抽象语法树中包括了足够多的语法结构信息。

该过程可以描述如图3.4。在实际过程中，Program Differ 组件可以根据需要任意选择，只要其输入输出信息满足上述需求即可。图中所描述的软件变更集合可以定义如下：

定义 3.22: $change_set = \{(old_i, new_i) \mid old_i \in Structure, new_i \in Structure, i \in \mathbb{N}\}$

定义 3.23: $Structure \rightarrow Statement \mid BasicBlock \mid Method \mid Class$

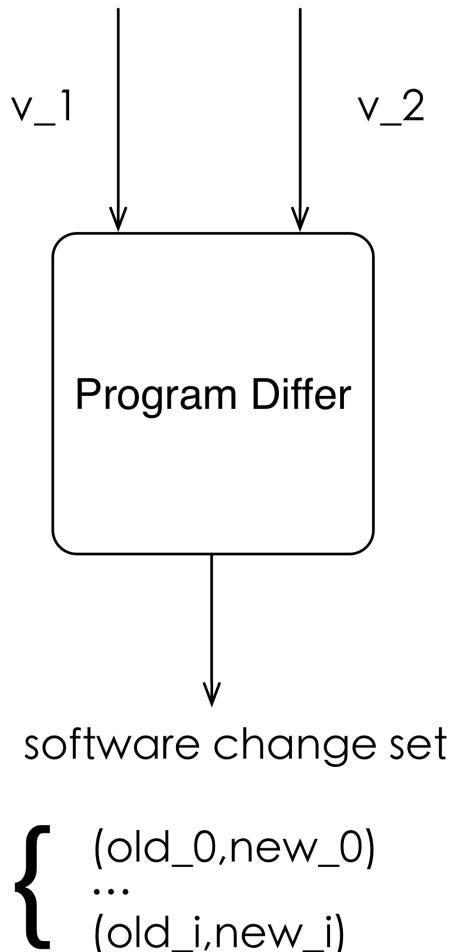


图 3.4 程序间差异分析

3.2.3.2 变更影响分析

程序变更影响分析主要用于获取变更集合对其他程序结构的影响集合，该集合所包含的程序结构直接或间接地受到变更集合中的元素影响。近年来这方面比较成熟的工作也有不少，因而可以直接选择合适的变更影响分析算法作为该过分析过程的具体实现。

在本文的组合架构中，该分析过程应当接受两个不同版本代码间的变更集合作为输入，并输出变更集合所对应的影响集合，这也就是我们所需要的变更集合的语义影响范围。变更影响分析的过程可以通过控制流、数据流等信息分析出变更集合中每条变更对其他程序结构是否存在影响，并进行闭包计算即可。

本文对于该分析过程的要求如下：

- 接受两个不同版本代码间的变更集合作为输入。
- 计算得到该变更集合所对应的影响集合。
- 计算过程中可以指定受影响的范围和元素类型。
- 影响集合中的元素按照其不同类型进行分类。
- 具有影响追踪系统，将计算影响集合的过程进行记录。方便后续的兼容性分析过程进行回溯。

其中影响追踪系统可以定义成如下类型的函数。该函数接受一个变更影响分析函数 $ia(v_i, v_j)$ ，并返回对应的影响依赖关系集合。

定义 3.24: $impact_track(ia(v_i, v_j)) :: (Code \rightarrow Code \rightarrow ImpactSet) \rightarrow ImpactDependency$

该分析过程如图3.5所述，其中的 impact analyzer 组件可以任意选择某一能够满足上述要求的变更影响分析算法图中所描述的变更影响集合可以定义如下：

定义 3.25: $impact_set = \{(structure_i) \mid structure_i \subset Structure, i \in \mathbb{N}\}$

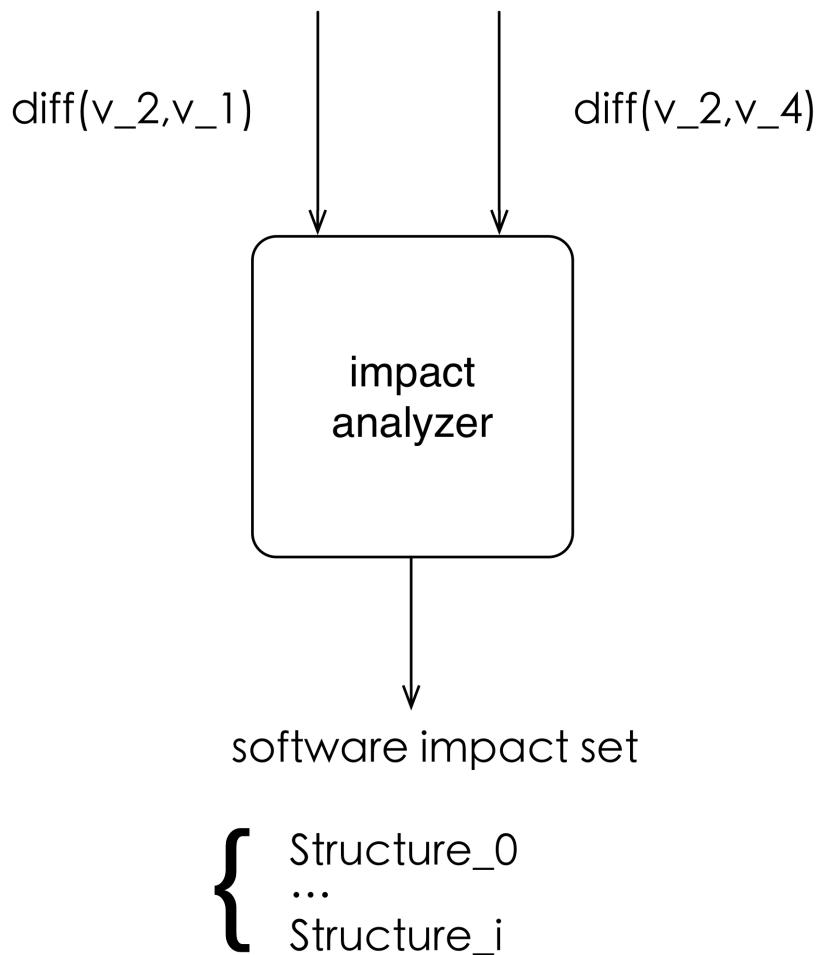


图 3.5 程序间差异分析

在完成了上述两项子分析过程后，我们就完成了整个语义影响范围分析过程，并获得了两个不同版本代码的变更集合和对应的其语义影响范围。接下来就可以进行具体的兼容性分析工作。

3.2.4 冲突分析

在冲突分析中，我们将对比两个版本代码的影响范围，确定他们是否重叠，以此为依据判断其兼容性。

理论上来讲，这种简单比对即可发现两个版本间的兼容性问题，因为一旦发生重叠，那么重叠的代码部分显然是会发生兼容性问题的。然而在实际情况中，受限于工具的精度，我们往往不能达到理论上的准确度，而可能会发生误报（false positive）等情况。

显然，如果重叠不存在，则兼容性是可以得到满足的。而对于如何界定重叠

部分的兼容性，则需要人工分析的介入，因为这部分代码的兼容性与补丁的功能目标密切相关，而我们无法从代码中获取到这种信息，因而只有依靠外界来提供，并以此为依据进行详细分析过程，界定这部分代码是否真的存在冲突。

在进行人工分析的过程中，我们不仅需要知道哪部分代码出现了重叠现象，而且还需要知道是哪些变更影响到了这部分代码，因而我们需要引入影响追踪系统来记录变更影响分析过程的轨迹。影响追踪系统可以记录下变更影响分析过程中的每一步，从而获取到程序结构间的影响关系链，事后通过回溯即可追踪到具体的软件变更可见，本过程中主要需要影响追踪系统的回溯模块的支持。

可以看得出来，由于使用了更严格的冲突定义，我们的方法会造成一定的过高估计的结果。

该分析过程可以参考算法1。

Algorithm 1 冲突分析

Require: $s_1 = ia(v_2, v_1), s_2 = ia(v_2, v_4)$

$t_1 = impact_track(ia(v_2, v_1)), t_2 = impact_track(ia(v_2, v_4))$

Ensure: $isCompatible(diff(v_2, v_1), diff(v_2, v_4))$

```
1: if  $s_1 = \emptyset \vee s_2 = \emptyset$  then
2:   result  $\leftarrow True$ 
3: else
4:    $s \leftarrow s_1 \cap s_2$ 
5:   if  $s = \emptyset$  then
6:     result  $\leftarrow True$ 
7:   else
8:     result  $\leftarrow Manual\_analysis(s_1, s_2, t_1, t_2)$ 
9:   end if
10: end if
11: return result
```

整个分析过程的架构如图3.6所示。

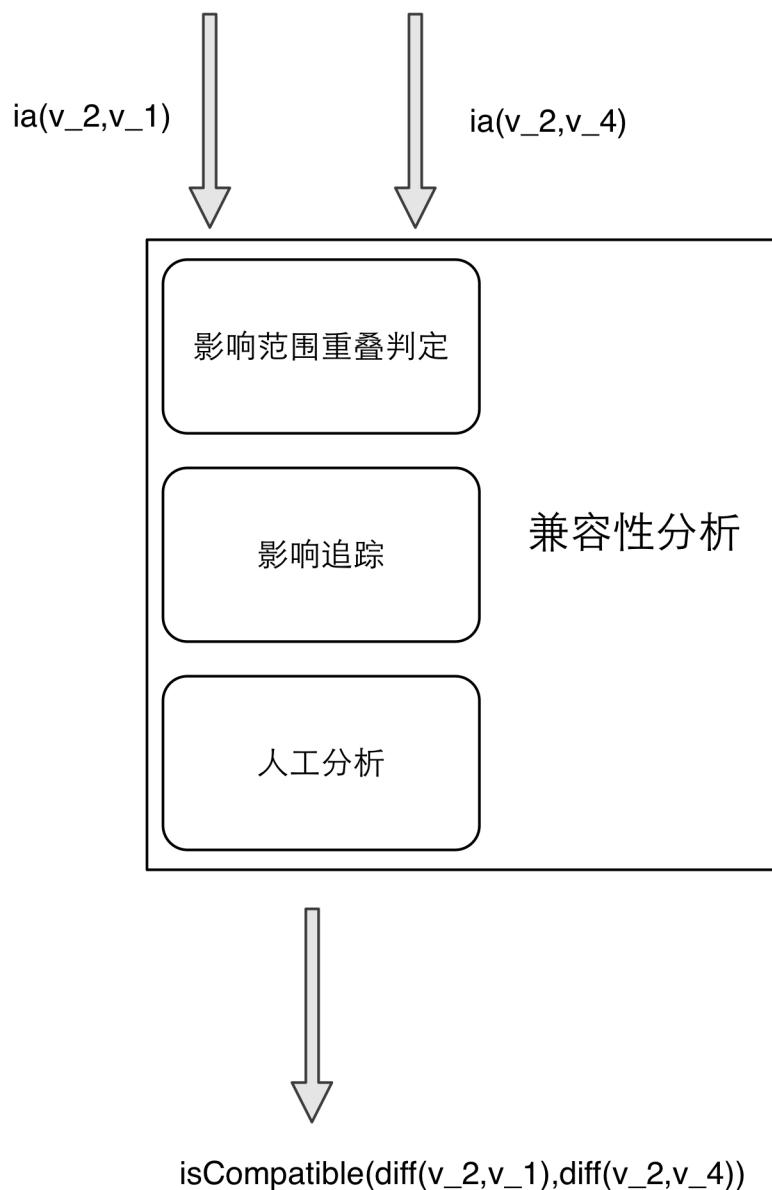


图 3.6 兼容性分析

3.3 本章小结

第4章 工具设计与实现

本章中将主要介绍如何将补丁版本迁移、语义影响范围分析、冲突分析等过程具体实现之，并组合成整体流程。

4.1 整体架构

4.2 补丁版本迁移

如前所述，补丁版本迁移主要用于解决问题3.2。我们在实现该过程时，主要采用了 git 工具和 Beyond Compare 工具来实现版本合并和冲突解决的过程。

我们下面将利用 git 的分支功能进行版本合并。首先介绍一些前提假设：

1. 假设 git 的工作目录如图4.1所示。其中 git_working_dir 为 git 目录，下属目录包括源代码仓库 src 和 git 目录的隐藏文件夹.git。另外与 git 目录平级的目录 patch 下面则存储了具体的补丁文件。
2. 假设补丁 $p_1 = old.patch$ ，并且采用 git diff 或其他 Unix diff 命令生成。

```
+/git_working_dir
|+/.git
|+/src
+/patch
|+old.patch
```

图 4.1 git 工作目录

下面将具体介绍整个版本合并的过程。

首先在工作目录下将 git 切换到主分支 master。然后在主分支 master 中提交代码版本 v_1 。

```
1 git checkout master
2 git add .
3 git commit -a -m "old.version.committed"
```

然后新建并切换到分支 new 中，并提交代码版本 v_2 。

```
1 git checkout -b new
2 git add .
3 git commit -a -m "new.version.committed"
```

再次从主分支新建并切换到分支 patch，然后将补丁 p_1 应用到版本 v_1 ，获得旧版本应用补丁后的代码，其版本为 v_3 ，并提交代码版本 v_3 。此时我们所应用的补丁 p_1 是专为版本 v_1 设计，所以应用时不会出现问题。

```
1 git checkout master
2 git checkout -b patch
3 git apply ..patch/old.patch
4 git add .
5 git commit -a -m "patched version from old version committed"
```

然后再切换回分支 new，将分支 patch 合并入分支 new，获得新版本应用补丁后的代码，其版本为 v_4 ，并使用 Beyond Compare 工具解决可能出现的冲突问题。将冲突解决完毕后，再提交版本 v_4 。

```
1 git checkout new
2 git merge patch
3 git mergetool
4 git commit -a -m "patched version from new version committed"
```

如果确实有冲突，那么 git mergetool 命令会调用第三方的可视化合并工具并引导你去解决冲突。这里我们采用的合并工具即为 Beyond Compare，它会展开一个可视化的界面，并给出冲突位置的提示，方便进行人工选择、合并。

整个过程可以参考图4.2。

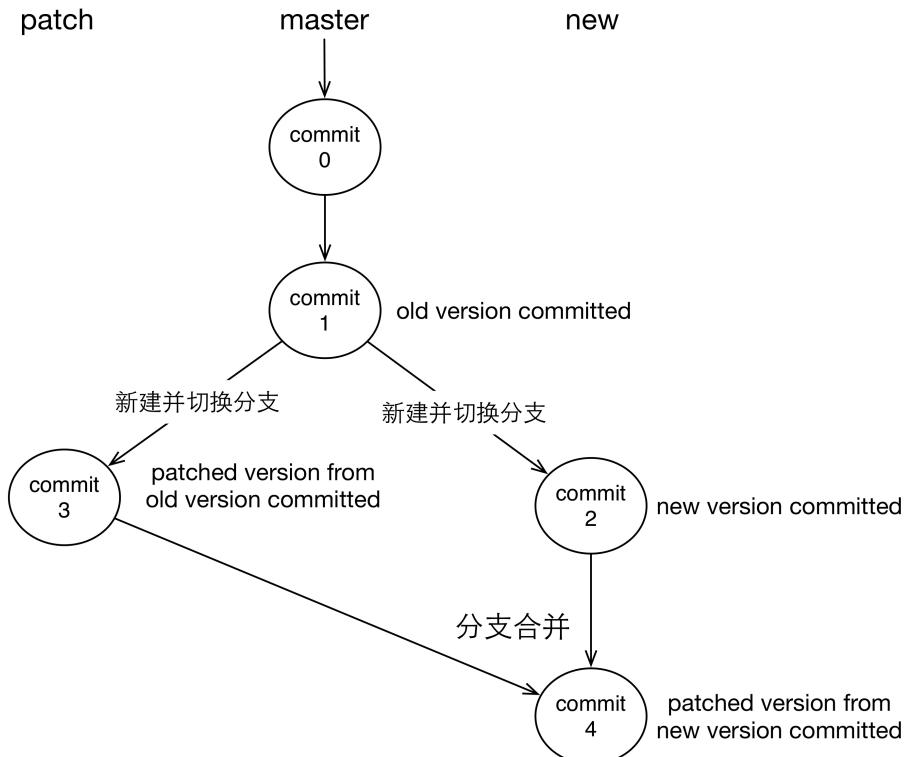


图 4.2 git 版本合并

4.3 语义影响范围分析

如前所述，语义影响范围分析主要用于解决问题3.3。语义影响范围分析需要首先定义好影响范围和影响元素的级别。在实际情况中，我们主要采用了较为简单的分析过程，即将影响范围限制在方法内部，但是将受影响元素设置为程序语句级别，以保证精度。

由于我们采用了现有工具来完成具体的程序间差异分析和变更影响分析过程，因而这两个子过程的分析算法已经无需我们自己实现。我们可以将主要精力放在如何整合两个子过程的工作上，并将其工具转化为适用于本问题的具体情况。

在具体的实现过程中，我们发现一些主要需要解决的问题包括：

1. 修正 ASTro 的分析结果，以提高精确度。
2. 改进 jpf-regression 工具，包括：
 - (a) 修复工具中自带的 Bug。
 - (b) 修改工具的分析过程。
 - (c) 增加影响追踪系统。
 - (d) 增加错误记录系统。
3. 实现实验过程的批量化和自动化。

4.3.1 程序间差异分析

本文中主要采用 jpf-regression 工具提供的前置工具 ASTro 进行程序差异性分析工作。

ASTro 支持对 Java 代码的比对。它会比对两个文件的抽象语法树，并从中抽取出对应的不同之处，形成语法结构上的差异性，并输出为 XML 格式的文件以供后续分析过程使用。

该工具输出的 XML 文件将源代码按照抽象语法树的格式进行输出，其节点级别为基本块，并提供了丰富的差异性信息，如两个版本代码的对应节点是否发生了变更等。

利用这些输出信息，我们从中提取出程序的变更集合，从而进行后续的变更影响分析，

在实际使用中，为了满足本文的需要，在将工具整合进来时进行了一些改进。

首先，本文中采用了 shell 脚本完成了分析过程的批量化和自动化，能够自动对整个软件系统的所有代码进行批量化处理，并循环调用 ASTro 进行分析。若需要对不同的代码进行修改，只需要修改对应的实验数据存放位置和其代码所依赖的 JAR 包即可。在这部分工作中，脚本代码主要完成了以下内容：

表 4.1 ASTro 参数对照表

参数名	描述	启用
-file	分析目标的名字	是
-dir	输出路径	是
-ocp	旧版本代码的 Classpath	是
-mcp	旧版本代码的 Classpath	是
-original	旧版本代码的位置	是
-modified	新版本代码的位置	是
-xml	以 XML 格式输出结果	是
-cs	以变更脚本 (Change Script) 格式输出结果	否
-heu	以启发式的方式进行匹配	是

- 实验数据定位，包括 Java 源代码和编译后的 Class 文件等。
- 根据代码的存放路径，计算其对应 Class 文件的位置。
- 获取代码文件名，以确定本次分析的对象。
- 实验数据的依赖 JAR 包定位。
- 创建输出文件目录。
- 定义 ASTro 的输入参数，包括输入文件位置、输出文件位置、查找路径等。
- 调用 ASTro 进行单次分析。

其中 ASTro 工具的使用格式可参考如下，其具体各参数的定义参考表4.1。

¹ ASTDiffer 3/27/2013

² USAGE: java ASTDiffer –original <file>.java –modified <file>.java

³ –dir <output folder>

⁴ OPTIONAL: –file <fileName> –ocp <classpath> –mcp <classpath>

⁵ –oco <outputDir> –mco <outputDir> –cs –xml

其次，我们采用了 shell 脚本完成了对后续分析过程的支持，能够自动批量化创建变更影响分析所需的配置文件。配置文件为自定义的 JPF 格式，通过类似键值对的方式定义了各项属性的值。JPF 文件的格式等来自于 Java Path Finder 框架的设计，是该框架运行所必须的配置文件。该配置文件的具体属性为自定义设置，可以参考表4.2所述。

在使用 shell 脚本调用 ASTro 工具进行分析和输出变更影响分析的配置文件时，考虑到实际使用中，我们需要将新版本 v_2 作为对比的基准，以获取一致的行号。因而在进行变更影响分析时，我们需要进行相应配置，使得：

- $s_1 = \text{impact}(\text{diff}(v_2, v_1), v_2)$ ，求得补丁 $p_1 = \text{diff}(v_2, v_1)$ 对版本 v_2 的影响范围 s_1 。
- $s_2 = \text{impact}(\text{diff}(v_2, v_4), v_2)$ ，求得补丁 $p_2 = \text{diff}(v_2, v_4)$ 对版本 v_2 的影响范围

表 4.2 JPF 属性对照表

属性名	描述
target	分析的目标
sourcepath	源代码路径
rse.ASTResults	ASTro 工具的输出文件位置
rse.newClass	新版本代码的 Class 文件位置
rse.oldClass	旧版本代码的 Class 文件位置
rse.dotFile	jpf-regression 工具的 Dot 格式输出文件位置

s_2 。

实际上，也就是说：

- 补丁 $p_1 = \text{diff}(v_2, v_1)$ ，即将新版本 v_2 视为“旧版本”，将旧版本 v_1 视为“新版本”。
- 补丁 $p_2 = \text{diff}(v_2, v_4)$ ，即将新版本 v_2 视为“旧版本”，将新版本应用补丁后的版本 v_4 视为“新版本”。

在实际操作中，我们只需做这样的版本交换即可。

受限于 ASTro 工具的具体实现，其输出结果的存在一定的问题，主要包括：

1. 对某些代码文件无法完成差异性分析。
2. 对某些代码文件输出结果不准确，存在过高估计（over-estimate）的问题。

对于第一个问题，由于无法知道该工具的源代码，我们无法解决，不过这只是极少数现象。

对于第二个问题，我们分析其结果可以发现，其结果中存在误报的情况，即某些代码行并未发生变更，然而工具却报告其发生了诸如移动、先删后增之类的伪变更。同样由于无法知道该工具的源代码，我们无法从算法的角度进行修改，不过对于这样的情况，我们可以对其输出结果进行一定的预处理，将这些误报的情况进行过滤，保留一个真变更子集合即可。

预处理算法可以用伪代码2进行描述。

我们可以归纳证明这种预处理操作的正确性。由于变更对于代码的影响是链式的，对于某次变更影响分析的结果集合 $s_{i,j} = ia(v_i, v_j)$ 而言，假设对于其中任意一个受影响的元素 e_k ，其中 $k \in \mathbb{N}$ ，其影响来源可能包括如下几种可能：

1. 其影响仅来源于变更 c_1 。
 - 如果 c_1 为真变更，那么删除所有伪变更对于 e_k 没有影响。
 - 如果 c_2 为伪变更，那么删除所有伪变更会导致 e_k 从集合 $s_{i,j}$ 中被删除，但此时 e_k 本身即为伪影响，集合 $s_{i,j}$ 的正确性会得到提高。
2. 其影响来源于多条变更 $c_1, c_2 \dots, c_m$ ，其中 $m \in \mathbb{N}$ 。

Algorithm 2 XML 结果过滤算法

Require: $c_1 = \text{diff}(v_2, v_1), c_2 = \text{diff}(v_2, v_4)$

Ensure: $\text{filter}(c_1, c_2)$

```

1:  $del_1 \leftarrow \emptyset$ 
2:  $del_2 \leftarrow \emptyset$ 
3: for  $i = 0$  to  $\text{sizeof}(c_1)$  do
4:    $tc_1 \leftarrow c_1[i]$ 
5:   for  $j = 0$  to  $\text{sizeof}(c_2)$  do
6:      $tc_2 \leftarrow c_2[j]$ 
7:     if  $tc_1 == tc_2$  then
8:        $del_1.add(tc_1)$ 
9:        $del_2.add(tc_2)$ 
10:    end if
11:   end for
12: end for
13:  $c_1 \leftarrow c_1.delete(del_1)$ 
14:  $c_2 \leftarrow c_2.delete(del_2)$ 

```

- 假若所有变更均为真变更，那么删除所有伪变更对于 e_k 没有影响。
- 假若来源变更集合中包括某几条伪变更，那么删除所有伪变更之后，仍然存在其他真变更，这些真变更仍然会在变更影响分析中导致 e_k 被添加到集合 $s_{i,j}$ 中，因而也不会使集合 $s_{i,j}$ 的正确性下降。
- 假若所有变更均为伪变更，那么删除所有伪变更会导致 e_k 从集合 $s_{i,j}$ 中被删除，但此时 e_k 本身即为伪影响，集合 $s_{i,j}$ 的正确性会得到提高。

可见，我们的预处理操作是正确的，它不会导致结果集合 $s_{i,j} = ia(v_i, v_j)$ 的正确性降低。

整个程序间差异性分析过程可以参考图4.3。

图 4.3 程序间差异性分析流程

4.3.2 变更影响分析

本文中主要采用 jpf-regression 工具进行变更影响分析的工作。jpf-regression 是 DiSE 方法在 Java Path Finder 软件框架下的具体实现，提供了方法内和方法间的程序语句级别的变更影响分析。

然而 jpf-regression 工具中变更影响分析只是其中的一个子模块，主要用于为其后续的 DiSE 分析过程服务。因而在实践过程中，我们采用的解决办法是重用 jpf-regression 的代码，并对其加以改造，主要的变化包括：

1. 修改分析流程。
2. 增加影响追踪系统。
3. 增加错误记录系统。
4. 使其适应大规模批量化分析的需要。
5. 已知 Bug 修复。

下面分别进行介绍。

4.3.2.1 分析流程

jpf-regression 作为 Java Path Finder 框架的一个插件，事实上在使用时需要遵守该框架的约束，有明确的执行流程规定。然而在实际使用中，该流程约定与我们的实际情况并不合用，因而我们对此进行了一定的修正。

事实上，原执行流程约定，每次分析以源代码文件中的 Main 函数作为入口，探索并分析 Main 函数所调用的其他函数。该流程对于大部分情况而言是具有实际意义的，并且由于只考虑 Main 函数及其调用的函数，工具可以节约分析的开销，更快的得出结论，而忽略掉其他事实上并未在执行过程中被涉及到的函数。

然而对于我们的分析需要而言，该流程只能覆盖到部分情况，对于其他类型的软件系统而言可能并不适用，例如以 Eclipse JDT Core 项目而言，该项目主要用于为 Eclipse 软件系统的其他组件提供服务，因而在实际中以 JAR 包的形式作为库函数而存在。对于这类以库函数形式对外提供服务的源代码而言，他们并不存在入口函数，也无法预知到底会有哪些函数会被外界所调用。因而对于这类情况而言，我们需要在分析过程中覆盖其所有函数，以保证结果的完整性和正确性。

我们对于流程的修改可以参考图4.5进行对比。首先我们去掉了图4.4中所示的 JPF 框架启动流程，直接调用 jpf-regression 的核心功能。其次，如图4.5所示，我们将待计算的方法集合从 Main 函数的调用函数集合修改为文件中所有方法的集合，并在相应的影响集合计算过程中增加了影响追踪系统。

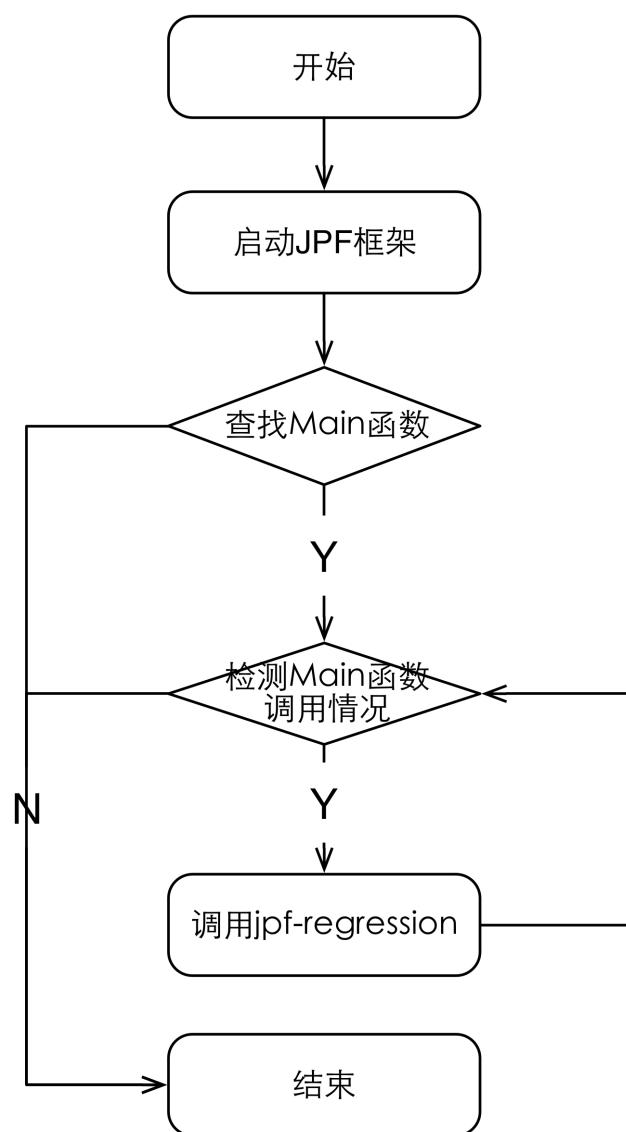


图 4.4 jpf 框架启动流程

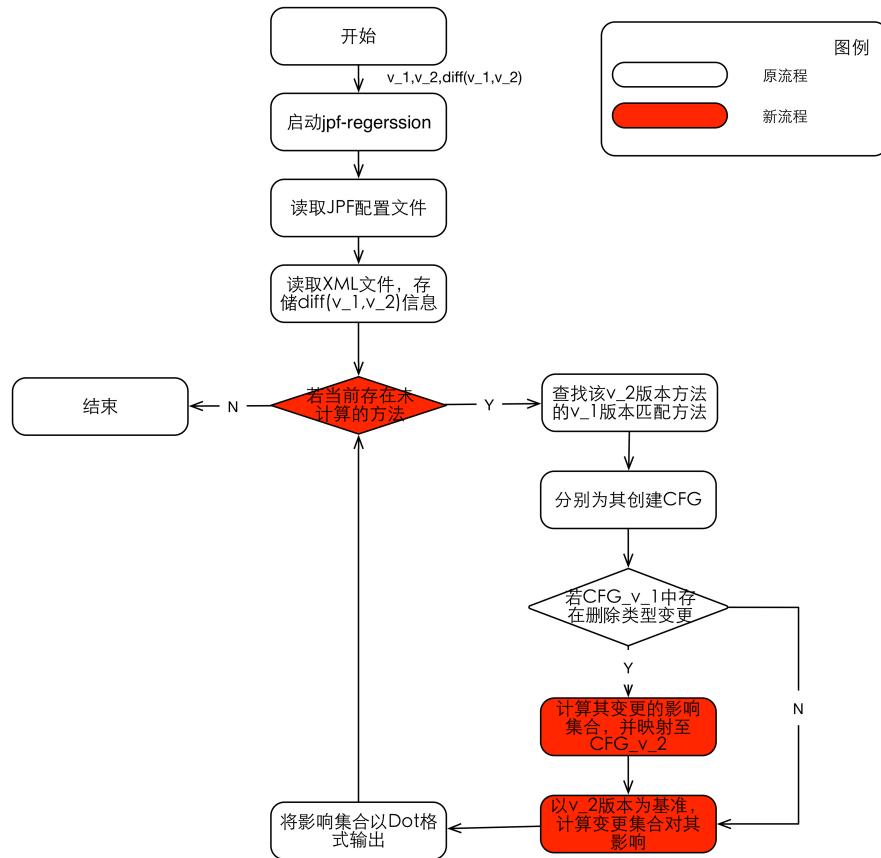


图 4.5 jpf-regression 原流程及变化

4.3.2.2 影响追踪

在后续的兼容性分析过程中，对于得到的冲突结果，我们需要对其进行追根溯源，挖掘其起始的影响来源，找到对应的影响来源变更集合，以进行人工分析对比，判定该情况是否确实冲突。

因此，我们需要在变更影响分析的阶段也加入影响追踪系统的记录模块，以便记录下变更影响的轨迹，根据这些信息为后续的分析过程提供便利。

为了实现影响追踪系统，我们需要存储程序结构间的影响关系。如前所述，影响的来源主要有两类，即：

- 控制依赖
- 数据依赖

因而，为了描述该影响关系，本文设计了 Dependency 类族，参见图4.6。该类中使用了一个二元组的数据结构 depend 来存储影响来源和受影响对象，并使用了多态机制来区分影响关系的类型，即是控制依赖还是数据依赖。Dependency 类族中重写了 hashCode() 方法和 equals() 方法，以便能够放入集合中进行存储。

表 4.3 影响关系数据结构

数据类型	数据结构	用途
Dependency	dependency	单个影响关系
Control	dependency	单个控制依赖影响关系
Data	dependency	单个数据依赖影响关系
Map<Integer, Set<Dependency>>	depend	存储计算过程中的全部影响关系

影响追踪系统中具体使用到的数据结构可以参考表4.3。

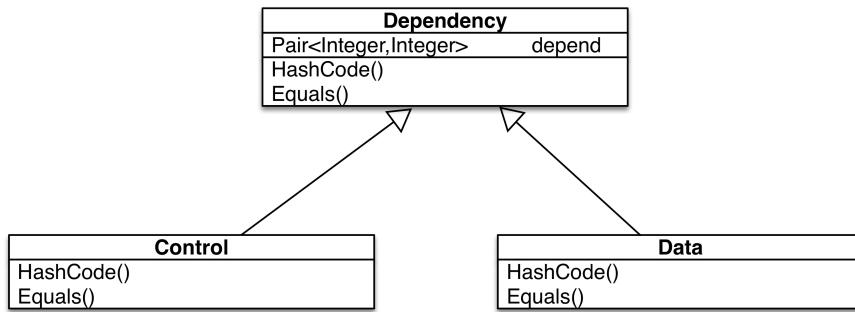


图 4.6 Dependency 类族

影响关系的创建需要在进行变更影响分析过程的同时进行，以便记录下整个影响关系链。最后将整个影响关系输出到 Dot 文件去即可。

4.3.2.3 错误记录

原有的 jpf-regression 工具由于是单次分析过程，因而一旦在运行过程中遇到问题，就会采用抛出异常终止运行的方式结束分析。然而我们在实际情况中需要进行大规模的分析作业，如果仅仅在其中单个文件的分析过程中出错就终止整个分析作业，会造成极大的时间和计算资源的浪费。

因而我们对该工具的异常处理方式进行了修改，使其在单次分析过程中如果遇到问题，则会及时抛出异常，但并不终止整个程序的运行，而采取了继续往下执行并分析其他文件的策略。然而分析错误是确实存在的，为了不丢失这类错误信息，我们为工具添加了错误记录系统，不断记录单次分析过程中遇到的问题。

我们对程序运行过程中可能出现的报错情况进行了分类，并设计了专门的错误统计类以专门别类的对错误情况进行记录和统计。该类的设计可以参考图4.7。其中使用到的数据结构可以参考表4.4的说明。

表 4.4 错误记录数据结构

数据类型	数据结构	用途
String	xml_not_found	错误: XML 文件未找到
String	class_not_found	错误: Class 文件未找到
String	interface_class	错误: 接口类
String	null_class	错误: 类中无具体实现(如抽象类)
Set<String>	precise_analysis	记录有多少方法在影响计算过程中出错
Map<String, String>	class_error	记录有哪些类出现了哪些错误
void	print()	输出错误记录

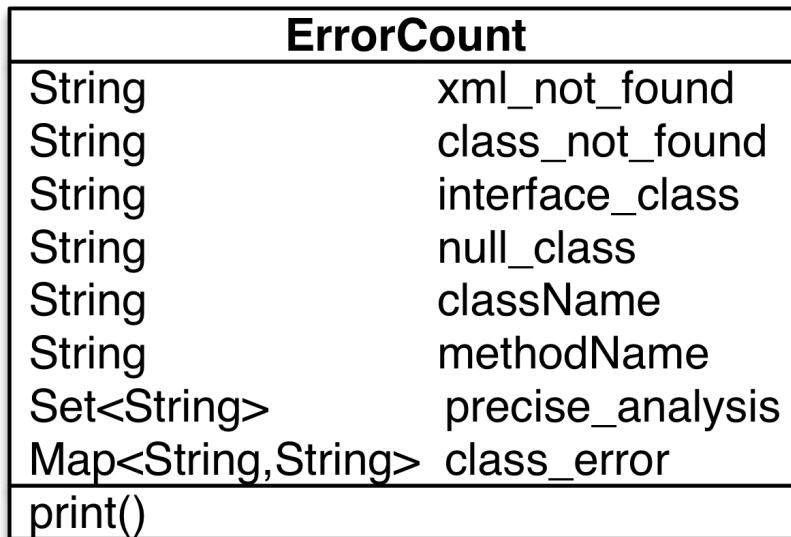


图 4.7 ErrorCount 类

4.3.2.4 大规模分析

原有的 `jpf-regression` 工具只能支持单次分析过程，在实际情况中我们需要工具具备大规模批量化分析的能力，以应对大规模软件系统的实际需求。为此我们可以保留原有的单次分析过程，然后在上层进行封装，循环多次调用单次分析过程，以达到批量化自动分析的效果。这个过程由于进行了封装，对于用户而言是透明的。

同时，在进行大规模分析的时候，输出文件的命名格式也需要修改。在原单次分析过程中，输出文件直接采用被分析的方法名进行命名。对于分析小型文件而言，这种设计就足够了，然而在大规模分析的时候，我们需要进行一定的优化。

由于大规模分析时，可能存在一些现象，例如：

1. 函数重载
2. 不同版本间的代码其方法可能无法一一匹配。例如有的方法仅在单个版本中出现。

在这种情况下，我们采用的命名格式为：

MethodName + HashCode(MethodName) + ExtensionName

其中，*MethodName* 即为方法名，无法保证方法名的唯一性。再利用 Java 中的 *HashCode* 方法，对 *MethodName* 计算其 *HashCode* 作为其后缀，以保证方法名的唯一性。最后 *ExtensionName* 即为文件扩展名，在 *jpf-regression* 中 *ExtensionName = .dot*。

同时我们也保留了原有的单次分析能力，以满足实际情况中的其他需要，例如进行小规模的案例分析。

其次，在进行大规模分析的情况下，由于实验数据量的庞大，我们无法按照单次分析过程中那样手动查看并分析实验结果。因此，为了适应这种需要，我们还增加了数据统计模块，使得程序具备一定的自动化分析实验结果的能力。

相关的类设计可以参考图4.8和图4.9。其中涉及的数据结构及其说明参考表4.5和表4.6。

RunAll	
<code>Set<String></code>	<code>filenames</code>
<code>Set<ErrorCount></code>	<code>errors</code>
<code>void readFileName(String path)</code>	
<code>static void main(String[] args)</code>	
<code>void bakDot(String p)</code>	
<code>void deleteDot(String p)</code>	

图 4.8 RunAll 类

表 4.5 RunAll 数据结构

数据类型	数据结构	用途
Set<String>	filenames	存储待分析的文件名
Set<ErrorCount>	errors	存储每次分析中出现的错误
void	readFileName(String path)	从文件中读取待分析的文件名
static void	main(String[] args)	实现大规模分析过程
void	bakDot(String p)	备份分析结果
void	deleteDot(String p)	删除分析结果

表 4.6 Runjpf 数据结构

数据类型	数据结构	用途
String	file	待分析文件名
String	old_to_new	$impact(diff(v_2, v_1), v_2)$ 过程的配置文件位置
String	new_to_patch	$impact(diff(v_2, v_4), v_2)$ 过程的配置文件位置
void	change(String file)	改变当前需要读取的配置文件
ErrorCount	run_otn(String filename)	运行 $impact(diff(v_2, v_1), v_2)$ 过程
ErrorCount	run_ntp(String filename)	运行 $impact(diff(v_2, v_4), v_2)$
static void	main(String[] args)	实现单次分析过程

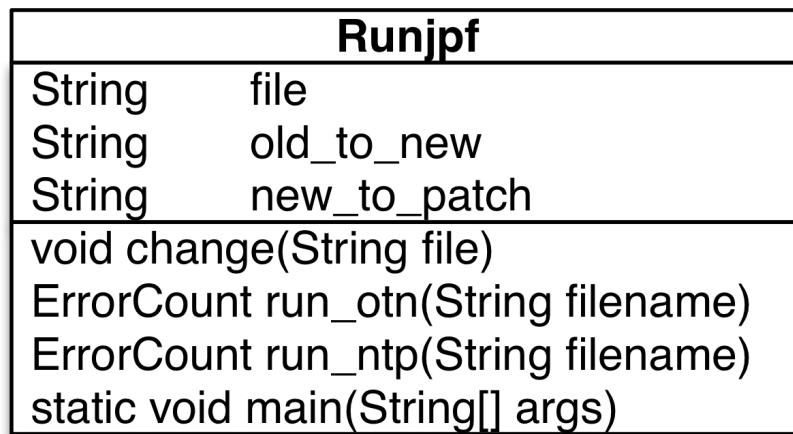


图 4.9 Runjpf 类

4.3.2.5 Bug 修复

在实际使用 jpf-regression 进行实验的过程中，我们发现该工具存在一些 Bug，这些 Bug 或多或少的导致了分析结果的正确性和精度降低。我们对其中力所能及

表 4.7 Bug 报告

Bug	危害	修复
内部类无法进行方法匹配	小	否
只有单个版本存在时无法进行方法匹配	小	是
将 CFG_{v_1} 的影响集合映射到 CFG_{v_2} 时判断条件出错	大	是
依赖 JAR 包 jpf_guided_test 出错	小	否
依赖 JAR 包 jpf_symboc 出错	小	否

的 Bug 进行了修复，并对这些 Bug 进行了总结。

目前已知的 Bug 及其修复情况可以参见表4.7。

4.4 冲突分析

如前所述，本文中给出了一个较为简单的自动化冲突分析方法，就目前而言，更精确的分析结果需要人工分析过程的辅助。

在实现冲突分析的过程中，主要涉及到的流程包括：

1. 读取语义影响范围分析的结果
2. 计算是否发生影响范围重叠
3. 对于发生了重叠现象的影响范围，判定为冲突
4. 回溯不兼容部分代码的影响来源并输出其影响关系链
5. 根据得到的影响关系链，人工分析重叠部分，判定是否确实冲突

可见，冲突分析过程中的主要工作包括：

- 计算影响范围重叠
- 进行影响追踪，输出影响关系链

4.5 本章小结

第 5 章 实验结果与分析

本章主要介绍如何对本文中所提出的补丁兼容性分析方法进行实验，并且给出了实验结果和分析。

5.1 实验设计

5.1.1 实验平台

下面就本文中所采用的实验平台，对其配置说明如下：

- 操作系统：Mac OS X 10.9
- CPU：2.4 GHz Intel Core i5
- 内存：8 GB 1600 MHz DDR3
- 硬盘：251 GB APPLE SSD SM0256F Media

5.1.2 实验数据

本文中主要采用 Eclipse JDT Core 项目作为实验对象。JDT Core 是 Eclipse 工具的 Java 基础组件，它主要提供了如下功能：

- Java 编译器。
- Java 文档模型。
- Java 模型。
- 编程帮助。
- 搜索。
- 源代码排版。

该项目的相关信息可以参见表 5.1。

Groovy-Eclipse 是一个 Eclipse 的插件集合，用于为 Groovy 项目提供 Eclipse 的工具支持。Groovy 是一门类似于 Java 的面向对象编程语言，Groovy 代码可以

表 5.1 Eclipse JDT Core

信息	描述
语言	Java
文件数	约 1100 个
代码量	约 3W 行

表 5.2 实验数据

代码	发行版本	逻辑版本
Eclipse JDT Core	4.3.2	v_1
Groovy-Eclipse JDT Core	4.3.2	v_3
Eclipse JDT Core	4.4	v_2
Eclipse JDT Core	4.4.2	v_2
Eclipse JDT Core	4.3	v_2
Eclipse JDT Core	4.3.1	v_2
Eclipse JDT Core	4.2	v_2
Eclipse JDT Core	4.2.1	v_2
Eclipse JDT Core	4.2.2	v_2

被编译器转化为 Java 字节码，从而在 JVM 上运行。由于这一特性，使得 Groovy 可以调用其他 Java 语言编写的库，从而大大丰富了其可用性。

Groovy-Eclipse 中提供了一个补丁后的 JDT Core 版本，该补丁主要用于增强 JDT Core 的功能，使其能够无缝编译 Groovy 代码。因而我们选择了该补丁作为实验数据中逻辑版本 v_3 的来源。

JDT Core 项目截止目前已推出发行版 4.4.2，我们从中选择了若干个发行版作为实验数据中逻辑版本 v_1 和 v_2 的来源。逻辑版本的具体说明可以参考表5.3，具体的实验数据选择可以参考表5.2，其中由于逻辑版本 v_4 由版本合并而来，因此在表5.2中不予以列出。

从表5.2中可见，我们一共选择了 7 个发行版本作为逻辑版本 v_2 的来源。也就是说，我们将以固定的逻辑版本 v_1 和 v_3 为基准，并选择不断变化的逻辑版本 v_2 来进行实验。

这样做的好处是我们可以更加清晰的了解本文中所提出的补丁兼容性分析方法的实用性，包括：

- 可针对工业界实际项目进行分析
- 切实贴合工业界的实际需求
- 对同一软件系统的多个不同版本具有普遍适用性

5.2 实验结果与分析

5.2.1 补丁版本迁移

根据我们所选择的实验数据，对于所有版本的提交和合并过程可以参考图5.1。可见该过程中我们以 Eclipse JDT Core 发行版 4.3.2 作为版本 v_1 ，以 Groovy-Eclipse

表 5.3 逻辑版本对照表

逻辑版本	描述
v_1	旧版本
v_2	新版本
v_3	应用补丁 p_1 后旧版本
v_4	版本 v_2 和版本 v_3 合并后版本，相当于应用补丁 p_1 后新版本

表 5.4 Patch Reversion 结果

代码	发行版本	冲突数量
Eclipse JDT Core	4.4	约 310
Eclipse JDT Core	4.4.2	约 350
Eclipse JDT Core	4.3	约 40
Eclipse JDT Core	4.3.1	约 40
Eclipse JDT Core	4.2	约 45
Eclipse JDT Core	4.2.1	约 50
Eclipse JDT Core	4.2.2	约 50

JDT Core 发行版 4.3.2 作为版本 v_2 ，以其他 Eclipse JDT 发行版作为版本 v_3 ，并为每个版本 v_3 创建了自己的分支，用于进行版本合并。

在版本合并的过程中，我们检测到的待解决冲突数量可以参见表5.4。根据实验结果，这些冲突 95% 以上都可以采用 Beyond Compare 工具自动给出的方案进行解决，只有极少部分代码需要进行手工操作。

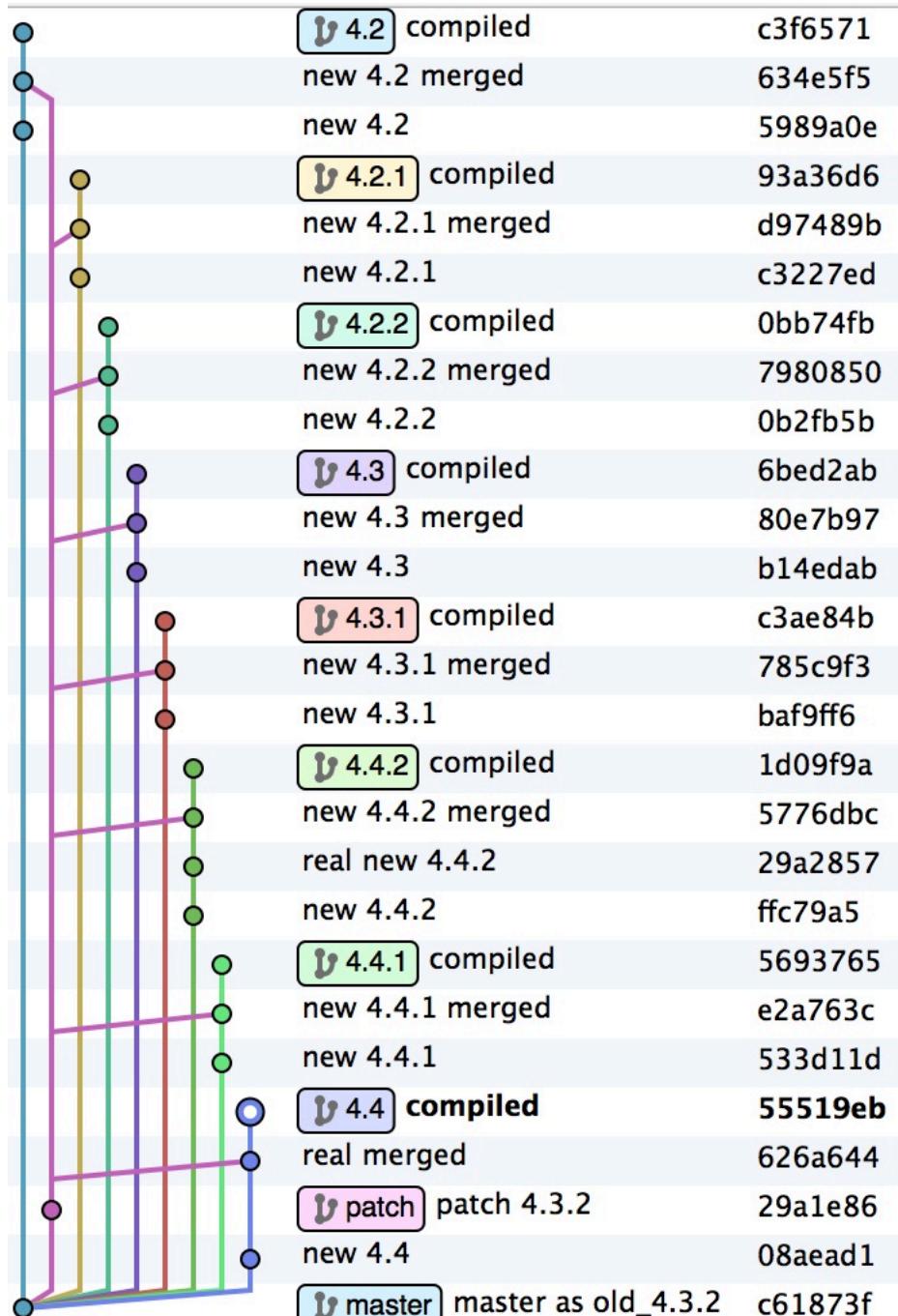


图 5.1 git 版本合并

在实际操作过程中，由于后续的语义影响范围分析需要提供 Java 代码编译后产生的 Class 文件，我们对于合并后的版本 v_4 还需要进行编译。实验过程中，99% 的文件都能够正常编译通过，只有极少部分文件由于合并出错等原因而无法编译通过。

可见，我们的补丁版本迁移方法是正确且有效的。

5.2.2 语义影响范围分析

由于整个语义影响范围分析可以分为程序间差异性分析和变更影响分析两个子过程，下面将分别叙述之。

5.2.2.1 程序间差异性分析

在程序间差异性分析过程中，我们主要关注成功进行分析的文件数量。

对于 $diff(v_2, v_1)$ 过程而言，使用预处理器后的结果如表5.5所示。

表 5.5 Program Differ 结果

代码	发行版本	预处理数	输出文件	实际文件
Eclipse JDT Core	4.4	1077	1127	
Eclipse JDT Core	4.4.2	1087	1127	
Eclipse JDT Core	4.3	1126	1127	
Eclipse JDT Core	4.3.1	1126	1127	
Eclipse JDT Core	4.2	1109	1127	
Eclipse JDT Core	4.2.1	1115	1127	
Eclipse JDT Core	4.2.2	1109	1127	

对 $diff(v_2, v_4)$ 过程而言，使用预处理器后的结果如表5.6所示。

表 5.6 Program Differ 结果

代码	发行版本	预处理数	输出文件	实际文件
Eclipse JDT Core	4.4	1077	1127	
Eclipse JDT Core	4.4.2	1087	1127	
Eclipse JDT Core	4.3	1124	1127	
Eclipse JDT Core	4.3.1	1123	1127	
Eclipse JDT Core	4.2	1109	1127	
Eclipse JDT Core	4.2.1	1116	1127	
Eclipse JDT Core	4.2.2	1109	1127	

可见，我们的预处理器算法对原有的 Program Differ 工具输出结果进行了有效的过滤。

5.2.2.2 变更影响分析

变更影响分析过程中，我们主要关注到底有多少文件成功完成了分析过程。

对 $impact(p_1, v_2)$ 过程而言，应用 impact analyzer 模块后，分析结果如表5.7所述。

表 5.7 impact analyzer 结果

代码	发行版本	分析结果数	分析文件数
Eclipse JDT Core	4.4	881	1127
Eclipse JDT Core	4.4.2	892	1127
Eclipse JDT Core	4.3	930	1127
Eclipse JDT Core	4.3.1	930	1127
Eclipse JDT Core	4.2	918	1127
Eclipse JDT Core	4.2.1	923	1127
Eclipse JDT Core	4.2.2		1127

对 $impact(p_2, v_2)$ 过程而言，应用 impact analyzer 模块后，分析结果如表5.8所述。

表 5.8 impact analyzer 结果

代码	发行版本	分析结果数	分析文件数
Eclipse JDT Core	4.4	881	1127
Eclipse JDT Core	4.4.2	892	1127
Eclipse JDT Core	4.3	930	1127
Eclipse JDT Core	4.3.1	925	1127
Eclipse JDT Core	4.2	916	1127
Eclipse JDT Core	4.2.1	924	1127
Eclipse JDT Core	4.2.2		1127

5.2.3 冲突分析

在冲突分析过程中，应用本文提出的冲突分析算法后，再通过影响追踪系统进行辅助人工分析，可以得到如表5.9的结果。

表 5.9 结果

代码	发行版本	冲突文件数	影响范围重叠数	分析文件数
Eclipse JDT Core	4.4	2	2	881
Eclipse JDT Core	4.4.2	3	3	892
Eclipse JDT Core	4.3	3	3	930
Eclipse JDT Core	4.3.1			
Eclipse JDT Core	4.2	4	4	918
Eclipse JDT Core	4.2.1	3	3	923
Eclipse JDT Core	4.2.2	4	4	932

然而，在不使用预处理器的情况下，实验结果如图5.10所示。

表 5.10 结果

代码	发行版本	冲突文件数	影响范围重叠数	分析文件数
Eclipse JDT Core	4.4			
Eclipse JDT Core	4.4.2			
Eclipse JDT Core	4.3		69	930
Eclipse JDT Core	4.3.1		66	925
Eclipse JDT Core	4.2	4		918
Eclipse JDT Core	4.2.1	3	63	923
Eclipse JDT Core	4.2.2	4		932

5.3 本章小结

第 6 章 结论

6.1 工作总结

如前所述，本文主要的工作是在于解决问题3.10。为此，本文提出了一套具有普遍适用性的补丁兼容性分析解决方案，该解决方案主要由三个组件构成，即：

- 补丁版本迁移
- 语义影响分析
 - 程序间差异性分析
 - 变更影响分析
- 冲突分析

其中语义影响分析组件的两个子组件可以自由替换成其他相应的算法，以提高解决方案的实用性。

目前冲突分析组件提出了一种较简单的自动冲突分析方法，更精确的分析结果目前需要人工分析的辅助，为此解决方案中需要变更影响分析组件实现影响追踪系统，以便追溯语义影响的来源。

可见，本文的主要成果包括：

- 提出了一套补丁兼容性分析的通用解决方案。
- 利用相应的已有工具实现了语义影响分析组件，将解决方案实例化，并对中型项目 Eclipse JDT Core 的八个不同版本实施，论证了该解决方案对工业界实际项目的可用性。

6.2 未来工作

对于本文中所提到的补丁兼容性解决方案，其可能的未来工作方向包括：

- 根据解决方案的要求，进一步整合已有工具，实现解决方案的工具化。
- 更换语义影响分析组件中所使用的工具并继续实验，论证该解决方案的普遍适用性，并论证兼容性分析的精度与所采用的工具之间的关系。
- 对更大型的项目进行实验。

参考文献

致 谢

感谢贺飞老师对我的悉心指导，他的言传身教令我受益匪浅，并在我完成毕设工作时给出了许多有用的意见。

最后要感谢 ThuThesis，帮助我完成了本文的写作。

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： _____