

УТВЕРЖДАЮ
зам. директора по научной работе
ИПМ им. М.В. Келдыша РАН,
— м.н., профессор

М.В. Якобовский
29 ноября 2022 г.

ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ

Федерального государственного учреждения «Федеральный исследовательский центр Институт прикладной математики им. М.В. Келдыша Российской академии наук»
о диссертации Головешкина Алексея Валерьевича
на тему «Устойчивая алгоритмическая привязка к коду программы»,
представленной на соискание учёной степени кандидата технических наук
по специальности 2.3.5 — математическое и программное обеспечение
вычислительных систем, комплексов и компьютерных сетей.

Диссертация Головешкина Алексея Валерьевича посвящена вопросу устойчивой алгоритмической привязки к коду программы. Привязкой автор называет построение и сохранение описания участка кода в виде набора структур специального вида, называемых контекстами. Устойчивость привязки определяется успешностью перепривязки — поиска проекции на новую версию программы исходного участка кода по его описанию, сохранённому в момент привязки. Предполагается, что в ходе перепривязки актуальная версия искомого участка кода должна устойчиво находиться, кроме очевидного случая, когда этот код был удалён.

Актуальность работы обусловлена тем, что она позволяет в значительной мере решить часто возникающие при разработке программ проблемы, отнимающие большое количество времени у разработчика: проблемы поиска участков кода, обеспечивающих определенную функциональность, и поиска актуальной документации для этих участков.

Устойчивая алгоритмическая привязка к некоторому коду позволяет впоследствии автоматизировать поиск проекции этого кода на отредактированную программу. Модели и алгоритмы, реализующие устойчивую привязку, используются автором диссертации как основа для предлагаемого инструмента разметки кода, интегрируемого в различные среды разработки. Разметка кода позволяет программисту материализовать информацию о расположении участков кода, имеющих отношение к определенной функциональности. Также она позволяет связать с кодом пометки и комментарии — дополнительную информацию, поясняющую его назначение, связь с предметной областью и особенности реализации. Благодаря алгоритмам перепривязки однажды помеченные участки кода автоматически разыскиваются в отредактированной программе, т.е. связь между кодом и пометками не теряется в процессе развития программы.

Научная новизна работы состоит в предложенном автором методе легковесного парсинга на основе упрощённых грамматик, оказавшегося достаточным для устойчивой алгоритмической привязки участка кода к синтаксическим элементам программы. Практическая значимость полученных в работе результатов заключается в возможности построения на их основе продуктивных алгоритмов поиска участков кода с заданной функциональностью, в том числе с построением отношений на множестве участков кода, образующих прорезающие функциональности.

Результаты диссертационного исследования прошли апробацию на 8 всероссийских и международных конференциях, на 3 научных семинарах, опубликованы в 10 печатных работах, из них 1 работа проиндексирована в международной базе данных Scopus, 4 статьи опубликованы в журналах из перечня ВАК. Получены 3 свидетельства о государственной регистрации программы для ЭВМ.

Структура работы

Работа состоит из введения, пяти глав, заключения и списка литературы.

Во введении автор обосновывает актуальность диссертационного исследования, приводит сведения о степени разработанности темы, формулирует новизну и значимость работы, перечисляет методы исследования, обосновывает достоверность результатов, приводит информацию об апробации работы и публикациях, содержащих результаты работы.

В первой главе проводится обзор предметной области исследования. Анализируется широкий спектр научных публикаций, посвящённых вопросам прорезающих функциональностей, разработки программы с прорезающими функциональностями, запоминания участка кода и последующего его поиска, легковесного парсинга программ. Автор позиционирует своё диссертационное исследование относительно существующих методов, отмечая и обосновывая отличия выбранного им подхода.

Во второй главе вводится понятие упрощённой грамматики, позволившее построить алгоритмы синтаксического анализа для осуществления легковесного парсинга программы. Упрощение грамматики основывается на применении специального терминального символа *Any*, который используется для обозначения областей, анализ которых не требуется в рамках решаемой задачи. Приводятся данные эксперимента, демонстрирующего успешную разработку упрощённой грамматики языка C# в соответствии с предлагаемым подходом.

В третьей главе описываются расширения алгоритма синтаксического анализа, обслуживающие обработку символа *Any*. Предлагаются основанные на обработке *Any* алгоритмы восстановления от ошибок. По итогам экспериментальной проверки легковесных парсеров для языков C# и Java делается вывод о том, что парсеры, сгенерированные по грамматикам с символом *Any* и использующие модифицированные алгоритмы парсинга и алгоритмы восстановления от ошибок, успешно разбирают программы, разбираемые полным парсером, за исключением случаев, когда полный

парсер использует специальные проверки, выходящие за рамки контекстно-свободного анализа.

В четвёртой главе описываются модели контекстов и алгоритмы для устойчивой привязки к крупным синтаксическим элементам программы. При привязке запоминаются тип элемента и набор из пяти контекстов. Каждый контекст хранит информацию о некоторой части самого элемента либо его окружения. При перепривязке элемент ищется среди элементов того же типа, присутствующих в новой (отредактированной) версии файла. На основе ряда произведенных на реальных программах замеров утверждается, что достигнута устойчивость привязки, близкая к 100%.

В пятой главе описываются дополнительные алгоритмы, обслуживающие привязку не к крупной синтаксической единице, а к отдельной строке или произвольной группе строк.

В заключении перечисляются научные и практические результаты, полученные в работе, и направления дальнейшего исследования.

Основные результаты работы. Разработаны и реализованы средства задания участков кода, отвечающих за определенную функциональность. Главное достижение автора — предоставленные разработчикам средства, позволяющие разыскать в многократно измененном коде участки, отражающие ранее отмеченную функциональность. Поиск и предъявление разработчикам таких участков производится чрезвычайно устойчиво, с высокой вероятностью успеха.

Рекомендации по использованию результатов работы. Созданные инструментальные средства можно рекомендовать организациям-разработчикам масштабных программных проектов, реализуемых в течение длительного времени, таким как ИСП им.В.П.Иванникова РАН др. Применение этих средств может существенно облегчить навигацию по протяженному программному коду, которая нередко отнимает весьма заметную часть времени и усилий разработчика.

Замечания

1. Не всегда соглашаешься с предлагаемой лексикой: "легковесная грамматика" звучит несколько пренебрежительно, вряд ли автор рассчитывал на такое прочтение.
2. В предлагаемой постановке можно было бы поговорить о полноте охвата всевозможных полезных для заявленной задачи атрибутов, свойств рассматриваемого фрагмента программы: все ли такие атрибуты пришли в голову диссертанту? Или хотя бы учтены ли наиболее продуктивные? Хотя результаты нескольких проверок на реальном коде выглядят весьма убедительно, однако кашу маслом не испортишь, подумать об усилении постановки задачи имело бы смысл.
3. Как-то обойден молчанием вопрос о том, как быть, если за прошедшее время образовался, но не был соответствующим образом размечен еще один, новый фрагмент кода, относящийся к данной функциональности. Вот тут-то, по-видимому, и пригодились бы дополнительные механизмы поиска?
4. Хотелось бы услышать побольше соображений относительно включения предлагаемых средств в алгоритмические языки. Как такое включение могло бы выглядеть? Услышать ответ диссертанта на очевидный вопрос: почему до сих пор в языках отсутствуют подобные интересные и важные конструкции?

Заключение. Приведенные замечания не влияют на общую положительную оценку представленного диссертационного исследования. Диссертационная работа «Устойчивая алгоритмическая привязка к коду программы» является законченной научно-квалификационной работой, самостоятельно выполненной автором на высоком уровне. Работа удовлетворяет всем требованиям ВАК, предъявляемым к кандидатским диссертациям, а её автор, Головешкин Алексей Валерьевич, заслуживает присуждения учёной степени кандидата технических наук по специальности

2.3.5 — математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей.

Диссертация и составленный отзыв на нее были заслушаны и обсуждены в ИПМ им.М.В.Келдыша на заседании семинара им.М.Р.Шура-Бура, протокол № 9 от 29 ноября 2022 г.

Главный научный сотрудник
ИПМ им.М.В.Келдыша РАН,
д.ф.-м.н.

М.М.Горбунов-Посадов

Сведения об организации:

Федеральное государственное учреждение «Федеральный исследовательский центр Институт прикладной математики им. М.В. Келдыша Российской академии наук»

Адрес: 125047, город Москва, Миусская пл., д. 4

Телефон: +7 499 978-13-14

E-mail: office@keldysh.ru

Веб-сайт: <https://www.keldysh.ru>