УТВЕРЖДАЮ

Директор Института систем
информатики им. А.П. Ершова
Сибирского отделения
Российской академии наук,
доктор физико-математических наук

_____ Марчук А.Г.
24 ноября 2017 г.

ОТЗЫВ

ведущей организации на диссертационную работу
Михайлова Андрея Анатольевича
«Методы декомпиляции объектного кода Delphi»
представленную на соискание учёной степени кандидата технических наук
по специальности 05.13.11 — Математическое и программное обеспечение
вычислительных машин, комплексов и компьютерных сетей

Диссертационная работа А.А. Михайлова посвящена декомпиляции программ, т.е. автоматическому восстановлению исходного текста программы из объектного кода. Входным языком выбран объектно-ориентированный язык программирования Delphi. Этот язык являлся весьма популярным, и на нём был разработан большой объём программного обеспечения, сопровождение которого в настоящее время сталкивается с серьёзными проблемами, в частности, с тем, что исходный код был либо утерян, либо изначально недоступен пользователям. Декомпиляция программ может стать одним из инструментов (хотя бы частичного) решения этой проблемы.

Первая глава диссертации содержит постановку задачи декомпиляции и обширный обзор существующих декомпиляторов. Основное внимание в обзоре уделяется близким по духу языкам, таким как Java и языки платформы .NET. Выделяются основные проблем декомпиляции, такие как разделение кода и данных, анализ потоков управления, восстановление типов данных и др. Проводится сравнительный анализ декомпиляции с машинного языка, объектного кода и байт-кода.

Суть предлагаемого в диссертации метода можно охарактеризовать аналогией процессов декомпиляции и компиляции. Декомпиляция не сводится просто к обнаружению в объектном коде шаблонов кодогенерации, а может требовать нетривиального глобального анализа структур управления и потоков данных, различных фаз и форм промежуточного представления и т.д. При этом декомпилятор может

использовать, во-первых, знание о специфике конкретного компилятора, и, во-вторых, в некоторых случаях оставленную компилятором вспомогательную информацию, например, необходимую для целей отладки.

Вторая глава работы содержит описание методов и алгоритмов, используемых в декомпиляции объектного кода Delphi. Здесь приводится перечень определений — графа потока управления, базового блока, дерева доминаторов, и т.п., а также алгоритмы поиска доминаторов и интервального анализа, хорошо известные в теории программирования. При восстановлении управляющей структуры программы могут быть использованы ставшие уже классическими методы сведения графов, которые были адаптированы для характерных для Delphi шаблонов.

Третья глава описывает разработанную автором систему декомпиляции, её архитектуру, пользовательский интерфейс и примеры использования. Приводятся результаты тестирования на одной, хотя и достаточно сложной программе.

Последняя, четвёртая, глава посвящена задаче, лишь косвенно связанной с задачей декомпиляции, — визуализации управляющего графа.

Актуальность темы диссертации обуславливается наличием реальной производственной необходимости сопровождения большого объёма существующего программного обеспечения, разработанного на языке Delphi.

Предложенные в работе методы декомпиляции программ, скомпилированных с языка Delphi на платформу .NET, в своей полноте и комплексности являются **новыми**, их **обоснованность** опирается на использование теоретико-графовых моделей и алгоритмов, корректность которых формально доказана. **Практическая значимость** работы состоит в том, что предлагаемый в работе метод дает эффективное решение указанной задачи, что продемонстрировано успешным практическим использованием системы, подтверждённым соответствующими свидетельствами о регистрации и актами о внедрении.

Все основные результаты диссертации опубликованы с достаточной полнотой в российских и иностранных изданиях и прошли апробацию на конференциях и семинарах. Автореферат правильно отражает содержание диссертации.

Полученные в диссертации результаты представляют интерес для научно-исследовательских и учебных организаций (таких, как ИПС РАН, ИДСТУ СО РАН, СПбГУ, НГУ, МГУ и др.) а также для специалистов в области разработки трансляторов и другого системного программного обеспечения. Результаты рекомендуется использовать в указанных организациях, а также в других университетах РФ при подготовке специалистов в указанных областях.

Диссертация изложена на 155 страницах, состоит из введения, четырёх глав, заключения, списков использованных сокращений, условных обозначений и терминов, списка использованной литературы и 12 приложений, в которых помимо примеров результата работы декомпилятора, содержатся копии свидетельств, актов и дипломов,

полученных автором. Работа написана ясным языком, хорошо структурирована и логично оформлена.

Замечания по диссертационной работе:

1. В тексте диссертации имеются опечатки и пунктуационные ошибки. Например, стр. 8, 5 строка снизу, стр. 15, середина, стр. 38, 6 строка сверху и 11 стр. снизу, стр. 43, строка 2 снизу, стр. 52 строка 9 снизу и др.
2. Определение понятия «декомпиляция» даётся разными способами дважды: определение 1.0.1. и определение 1.2.1.
3. В первом предложении раздела 1.4.3 слово «анализа» следует заменить на «декомпиляции».
4. Несколько раз (стр. 38, стр. 39, и др.) утверждается что «задача декомпиляции неразрешима», что либо неверно, поскольку перевод всегда можно осуществить как эмуляцию, либо подразумевает более чёткое определение понятия декомпиляции.
5. Вероятно, на стр. 45, «статическое одиночное присваивание (SSA)» означает то же, что «статическое единичное присваивание (СЕП)». Сокращение SSA далее в работе не используется.
6. В определении 2.4.1. ориентированный граф определяется как G(X,U), но U нигде не используется.
7. Рисунки 2.10 и 3.3. абсолютно одинаковы.
8. Глава 4 представляется наиболее слабой в диссертации. Существует большое количество инструментальных средств, которые не только позволяют использовать пиктограммы, но и ориентированы на графы управления. Приведённые в диссертации алгоритмы достаточно наивны, о чем свидетельствует далеко не лучший результат их работы на рис. 4.5. Кроме этого,
9. Раздел 4.3. неожиданно начинается с рисунка 4.3, который потом дублируется как рис. 4.4.
10. Раздел 4.3.1 пуст.
11. Приложение В «Процедура инициализации опкодов CIL» объёмно, но не информативно и может быть удалено.

Перечисленные замечания не снижают общее положительное впечатление о диссертационной работе. В целом диссертационная работа А.А. Михайлова представляет собой законченную научно-квалификационную работу, выполненную на высоком математическом и техническом уровне.

Диссертация удовлетворяет всем требованиям, предъявляемым к диссертациям на соискание учёной степени кандидата технических паук Положением о присуждении ученых степеней, а ее автор Михайлов Андрей Анатольевич заслуживает присуждения

ему ученой степени кандидата технических наук по специальности 05.13.11 — Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей. Обсуждение диссертации проведено на заседании специализированного семинара «Системное программирование» Института систем информатики СО РАН 6 апреля 2017 года, протокол № 3.

Заведующий лабораторией смешанных вычислений
Института систем информатики СО РАН
кандидат физико-математических наук,
старший научный сотрудник

Бульонков Михаил Алексеевич

24.11.2017

Федеральное государственное бюджетное учреждение науки «Институт систем информатики им. А.П. Ершова Сибирского отделения Российской академии наук», 630090, Российская Федерация, г. Новосибирск, проспект Академика Лаврентьева, 6, (383) 3308652, www.iis.nsk.su, iis@iis.nsk.su.