ОТЗЫВ

официального оппонента, доктора физико-математических наук Галатенко Владимира Антоновича на диссертационную работу Мандрыкина Михаила Усамовича «Моделирование памяти Си-программ для инструментов статической верификации на основе SMT-решателей», представленную к защите на соискание ученой степени кандидата физико-математических наук по специальности 05.13.11 – «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей»

Актуальность темы диссертационной работы. Методы моделирования памяти в инструментах дедуктивной верификации, использующих SMT-решатели, отличаются как от методов моделирования памяти в инструментах статического анализа, так и от методов моделирования памяти в инструментах дедуктивной верификации, использующих интерактивные средства доказательства теорем. Для инструментов на основе SMT-решателей лучше разработаны методы моделирования памяти в языках программирования, имеющих высокоуровневую семантику, таких как Ada, Eiffel, ML и др. Для языков с низкоуровневой семантикой, существенно зависящей от расположения данных в памяти программы и их интерпретации в различных контекстах как объектов различного типа, моделирование памяти представляется более сложным. Такие языки, в частности Си, широко применяются для реализации низкоуровневых систем с повышенными требованиями к надежности, таких как ОС и гипервизоры. При этом методы моделирования памяти для языка Си часто порождают сложные для анализа формулы из-за необходимости точно моделировать низкоуровневую семантику языка. Поэтому задача разработки новых методов моделирования памяти Си-программ, снижающих сложность генерируемых формул, для инструментов дедуктивной верификации, использующих SMT-решатели, является **актуальной**.

В работе М.У. Мандрыкина предложен метод моделирования памяти Си-программ для инструмента дедуктивной верификации, использующий для снижения сложности генерируемых формул вспомогательные аннотации и результаты простого статического анализа указателей. Метод адаптирован для верификации модулей ОС Linux. Для метода приведено доказательство точности моделирования в виде теорем о корректности и полноте. Предложенный метод был реализован в инструменте дедуктивной верификации Jessie.

В работе также предложен метод моделирования памяти Си-программ для инструмента автоматической статической верификации, использующего предикатную абстракцию. Этот метод реализован в инструменте CPAchecker.

Диссертация М.У. Мандрыкина состоит из 4-х глав, введения, заключения, списка литературы и свидетельств о регистрации программ, имеется одно приложение. Общий объем диссертации составляет 207 страниц.

Введение содержит обоснование актуальности работы, формулировку ее цели и задач, научной новизны, значимости, а также ссылки на публикации по теме работы и список конференций и семинаров, на которых работа представлялась с целью апробации.

В **первой** главе приведен обзор методов моделирования памяти Си-программ в инструментах статической верификации. Обзор опирается на обширный список литературы и содержит анализ методов моделирования памяти как в инструментах автоматической статической верификации, так и в инструментах дедуктивной верификации.

Во **второй главе** рассмотрены проблемы существующих моделей памяти, используемых в инструментах статической верификации, указана область применения этих инструментов в рамках проектов по формальной верификации LDV и Astraver, сформулированы цель и задачи работы.

В **третьей главе** дается описание модели памяти для инструмента автоматической статической верификации CPAchecker. Основное отличие предложенной модели памяти – улучшенная поддержка адресной арифметики, часто используемой в модулях ядра ОС Linux. Реализация предложенной модели памяти в инструменте верификации CPAchecker позволила уменьшить число ложных срабатываний инструмента верификации на задачах, полученных в рамках проекта LDV.

В **четвертой** главе предложена модель памяти для инструмента дедуктивной верификации Си-программ. Преимущества предложенной модели памяти заключаются в том, что она:

- не требует дополнительных аннотаций для поддержки вложенных структур и массивов;
- упрощает порождение формул за счет использования результатов известного метода анализа указателей, имеющего линейную сложность;
- моделирует защищенную память, операции выделения и освобождения памяти;
- обладает полнотой (при использовании вспомогательных аннотаций) для подмножества языка Си, включающего объединения и произвольные приведения типов указателей.

В главе приведены доказательства теорем о корректности и полноте предложенной модели памяти для подмножества языка Си, включающего вложенные структуры, массивы, объединения и произвольные приведения типов указателей. Предложенная модель памяти для инструмента Jessie, а именно реализация поддержки вложенных структур и массивов и поддержка приведения типов указателей на целочисленные типы данных, позволила применить инструмент для дедуктивной верификации модуля безопасности ядра ОС Linux.

Научная новизна. В работе представлены следующие новые научные результаты:

- модель памяти на основе теории неинтерпретируемых функций для автоматической статической верификации Си-программ с использованием предикатных абстракций, уточняемых с помощью интерполяции Крейга;
- полная модель памяти с поддержкой переинтерпретации типов указателей и автоматизированного разделения на непересекающиеся области (регионы) для дедуктивной верификации Си-программ;
- формализация низкоуровневой семантики практически значимого подмножества языка Си;
- доказательство корректности и полноты модели памяти для дедуктивной верификации Си-программ с разделением на непересекающиеся регионы;
- метод автоматизированного разделения на непересекающиеся регионы для соответствующей модели памяти.

Практическая значимость результатов исследования. Полученные научные результаты могут использоваться в исследованиях по методам верификации программного обеспечения и в учебных курсах. Разработанные программные средства опубликованы под открытыми лицензиями и могут применяться в других программных проектах.

Достоверность и обоснованность научных положений и выводов работы. Достоверность полученных результатов подтверждается апробацией основных результатов работы на 5 научных конференциях и семинарах. Результаты представлены в 10 научных трудах соискателя, которые опубликованы в журналах, рекомендованных ВАК Минобрнауки России (из них 9 работ в изданиях, индексированных в базе данных Scopus).

Замечания. В изложении материала диссертации можно отметить следующие недостатки:

1. Излишне пространная обзорная часть, занимающая более половины основного текста диссертации.
2. Приложение А («Теории, поддерживаемые современными решателями») не относится к результатам работы автора, поэтому могло быть заменено ссылками на соответствующие источники в обзорной части.
3. Форма представления результатов сравнения в таблицах 3.3 и 3.5 не позволяет достаточно содержательно проинтерпретировать изменения за счет использования предложенной модели памяти.
4. Имеются опечатки (см., например, с. 15, 21, 108 диссертации, с. 3, 14, 20 автореферата, в диссертации ссылка [35] относится к элементу библиографии 36).

Перечисленные замечания не носят принципиального характера и не влияют на общую положительную оценку работы.

Заключение

Диссертационная работа соответствует всем требованиям ВАК РФ, предъявляемым к диссертациям на соискание ученой степени кандидата физико-математических наук, а Мандрыкин Михаил Усамович заслуживает присуждения ему ученой степени по специальности 05.13.11 – «математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей». Автореферат соответствует основному содержанию диссертации.

Заведующий сектором ФГУ ФНЦ НИИСИ РАН,
доктор физико-математических наук В.А. Галатенко

«28» ноября 2016 г.