Отзыв на выпускную квалификационную работу (ВКР) бакалавра Любаева Даниила Андреевича «Кодирование Java Bytecode в CIRCUIT-SAT для решения задач проверки эквивалентности программ»

Работа Любаева Д.А. посвящена актуальной теме применения комбинаторных алгоритмов к формальному анализу программ. Основная идея данного подхода состоит в следующем. Программа рассматривается как способ задания некоторой функции, свойства которой требуется исследовать. Эту функцию можно задать каким-либо другим способом, делающим применение комбинаторных алгоритмов более естественным. В работе Любаева Д.А. для этой цели используются т.н. And-Inverter графы, являющиеся по своей сути схемами из функциональных элементов над базисом, состоящим из конъюнкции и инвертора. Представление функции, задаваемой исходной программой, в виде схемы делает естественной, например, постановку задачи проверки эквивалентности двух программ. В рамках данной проблемы требуется по двум программам (возможно, реализующим различные алгоритмы или написанным на разных языках) проверить, задают ли эти программы одну и ту же функцию. Хороший пример – два разных алгоритма сортировки натуральных чисел (например, «пузырёк» и сортировка выбором) будут задавать одну и ту же функцию, а схемы, построенные по описаниям данных алгоритмов, будут эквивалентными.

Преобразование программы в схему предполагает детальный разбор работы этой программы с данными в памяти и, соответственно, высокоуровневые описания программ плохо подходят для этой цели. Гораздо естественнее в данном контексте выглядит работа с программами для виртуальных машин. В работе Любаева Д.А. для этого использовалась Java Virtual Machine (JVM). Основным результатом ВКР является программный инструмент Transbyte, который позволяет по JVM программе, задающей некоторую функцию, строить And-Inverter граф, который задает эту же функцию. Далее для функций, работающих с данными одинаковой длины, можно рассматривать задачи, относящиеся к проверке эквивалентности – конкретно, в работе была рассмотрена задача проверки по паре программ, являются ли эти программы семантическими клонами.

Для решения собственно задачи проверки логической эквивалентности двух схем можно использовать различные комбинаторные алгоритмы – чаще всего в подобных случаях используются алгоритмы решения проблемы булевой выполнимости (SAT-решатели). В настоящей работе был приведен ряд примеров решения проблемы выявления семантических клонов для некоторых JVM-программ с использованием современных SAT решателей, при том, что соответствующие задачи сводились к SAT при помощи разработанного инструмента Transbyte. Для уменьшения размера получаемых формул в конъюнктивной нормальной форме (КНФ) применялся известный инструмент для работы с булевыми схемами ABC. Экспериментальные результаты показывают, что Transbyte вполне может использоваться для решения проблемы семантических клонов для относительно небольших по объему кода JVM-программ.

Основное замечание к работе состоит в том, что Transbyte, рассматриваемый как транслятор программ в схемы, не покрывает все множество JVM-инструкций, хотя даже в своем настоящем виде Transbyte вполне можно использовать для проверки корректности обширного множества программ, содержащих арифметические операции с байтовыми данными.

Хочется дополнительно сказать несколько слов о перспективности данной работы. Представляется, что результаты, полученные в ней, демонстрируют принципиальную возможность создания программной системы, строящей схемные представления программ, написанных на языках «общего назначения» (general-purposed). В перспективе разработка такого рода системы, дающей более полное покрытие множества инструкций виртуальной машины (не обязательно JVM) видится крайне актуальной задачей в контексте формальных методов анализа программ, хотя создание такой системы, по-видимому, потребует усилий серьезной команды разработчиков.

Резюмируя, отмечу, что работа «Кодирование Java Bytecode в CIRCUIT-SAT для решения задач проверки эквивалентности программ» выглядит как законченное исследование с хорошими перспективами, выполненное на высоком научном уровне и удовлетворяющее требованиям, предъявляемым к ВКР бакалавра. Соответственно, данная работа заслуживает оценки «отлично» а ее автор, Даниил Андреевич Любаев — присвоения квалификации бакалавра по направлению «Математика и компьютерные науки».

Доцент факультета МКН СПбГУ,
Авдюшенко А.Ю.
29 мая 2023 г.