Рецензия

на выпускную квалификационную работу студентки кафедры системного программирования СПбГУ Сухановой Анжелы Кирилловны, обучающейся по направлению 09.03.04 (Программная инженерия)
Тема выпускной квалификационной работы:
«Фаззинг решателя Spacer»

В своей выпускной работе Анжела Кирилловна предлагает подход к автоматическому тестированию Хорн-решателей на основе фаззинга. В сфере верификации и автоматического тестирования ПО устойчиво используются логические решатели, например, SAT и SMT-решатели. Сейчас широкое распространение за пределами академии начинают получать т.н. Хорн-решатели, например, их уже применяют для автоматической верификации протоколов. Таким образом, проверка корректности логических решателей является критически важной задачей. К тестированию SAT и SMT-решателей уже применяют такие подходы к автоматическому тестированию, как фаззинг. Хорн-решатели же до сих пор тестировались только вручную.

В данной работе предложен способ автоматического тестирования Хорн-решателей на основе фаззинга. Спроектирован и реализован на языке Python фаззер с возможностью взвешенного выбора мутаций и возможностью кросс-валидации. Также реализован инструмент сокращения результатов фаззинга для порождения читаемых тестов. С реализацией проведены эксперименты на Хорн-решателе Spacer, в котором обнаружено 15 ошибок, которые приняты разработчиками и из которых 11 уже устранено.

Достоинства работы.

1. Исследование проведено полностью. Изучена предметная область, рассмотрены недостатки существующих решений и предложено рабочее решение, применимое на практике, с ним проведены эксперименты.
2. Текст написан в современном научном стиле, хорошо структурирован, вклад автора занимает большую часть работы.
3. В обзоре рассмотрены фаззеры SMT-решателей, которые за отсутствием фаззеров Хорн-решателей являются самым близким решением поставленной задачи. Описаны их недостатки и причины их неприменимости к задаче фаззинга Хорн-решателей, а также проведены эксперименты с ними, показывающие их действительную практическую несостоятельность для данной задачи. У меня как у эксперта в области это вызывает **отдельный восторг**: обзор, представленный в данной работе, я считаю также научным вкладом.
4. В основных разделах представлено решение поставленных задач на основе современной clock парадигмы фаззинга с практическими улучшениями на основе существующих научных работ. Тем самым, изучена и адаптирована современная парадигма фаззинга; она успешна перенесена в предметную область Хорн-решателей. В тексте присутствуют поясняющие графики и рисунки, примеры работы инструмента упрощения примеров.
5. Предложенный подход реализован в виде публично доступного инструмента на языке Python, его архитектура ясно описана. В реализации присутствует краткая документация, тесты, Docker описание. На мой взгляд, это показывает, что проделана качественная инженерная работа, которая может быть легко воспроизведена и улучшена сообществом.
6. С реализованным фаззером и инструментом упрощения примеров поставлены эксперименты путём тестирования Хорн-решателя Spacer. Хотя в работе заявлены 15

найденных ошибок, на гитхабе я нашёл 17, которые были приняты и здраво оценены разработчиками. Упрощение примеров и анализ их автором работы помогли сделать ошибки ясными для разработчиков, так что многие из них они уже смогли исправить.

7. Список литературы состоит из множества научных работ, опубликованных в престижных местах цитируемыми авторами. Большинство цитируемых работ опубликованы после 2017 года. На мой взгляд, это свидетельствует о том, что Анжела Кирилловна смогла хорошо сориентироваться в современном состоянии научной области и внести в неё важный вклад.

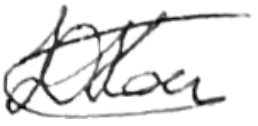
Недостатки работы.

1. Главным недостатком является тот факт, что большинство используемых в предложенном алгоритме фаззинга мутаций являются неспецифичными для решения систем дизъюнктов Хорна. В контексте Хорновских дизъюнктов было бы интереснее посмотреть на специфичные преобразования, например, на синхронизацию дизъюнктов, подстановку тел дизъюнктов вместо предикатов взятием резольвент и другие трансформации. Такое расширение дало бы сообществу не только возможность исправить ошибки в существующих решателях, но и оценить, насколько трансформации дизъюнктов могут помогать их решению.
2. Также не рассмотрены специфичные мутации для различных логических теорий. Я подозреваю, что эксперименты проводились только с дизъюнктами над теорией линейной арифметики, т.к. об этом в явном виде в тексте нигде не сказано, а также те мутации, которые специфичны для теорий, связаны именно с линейной арифметикой.
3. В тексте отсутствуют описание логики первого порядка, того, что такое теории и SMT, что важно для Хорн-решателей, которые в основном строятся поверх SMT-решателей.
4. В описании тестового стенда не сказано, с каким ограничением по времени запускался фаззинг, с каким ограничением запускались Хорн-решатели из фаззера.
5. Было бы полезно указать в разделе экспериментов, сколько времени потребовалось автору в среднем для осознания найденных инструментом ошибок и для создания соответствующих issue для разработчиков. В контексте автоматического тестирования это является очень важным.

Последние три недостатка легко поправимы в тексте. Первые два я бы отнёс скорее к перспективам дальнейшей работы, которая была бы крайне полезна для области, и которую я бы предложил выполнить в рамках магистерской работы.

Анжела Кирилловна смогла разобраться в специфических областях Хорн-решателей и фаззинга, качественно провести научное исследование и реализовать рабочий инструмент, который уже принёс практическую пользу в индустрии. Предоставленный исходный код и текст работы показывают, что автор разбирается в предметной области и обладает навыками программного инженера.

Поэтому учитывая вышеизложенные достоинства и недостатки считаю, что работа заслуживает оценки «**отлично**».



Рецензент: Костюков Ю.О.