

Отзыв научного руководителя
на выпускную квалификационную работу студентки группы 16.Б10-мм
Люлиной Елены Сергеевны
“Система анализа поведения студентов при обучении программированию”

Автоматическая генерация подсказок при решении задач по программированию может быть полезна студентам при самостоятельном обучении, а также преподавателям для эффективной организации уроков и экономии ресурсов на проверку решений задач. Для языка программирования Python существуют алгоритмы генерации подсказок, однако они плохо работают на больших и сложных задачах в силу вариативности решений задач. Однако, существуют алгоритмы генерации подсказок для блочного языка программирования, показавшие хорошие результаты. Они используют понятие пространства решений, что требует дополнительных данных о решениях задач. Перед Еленой Сергеевной стояла задача собрать необходимые данные и разработать алгоритм генерации подсказок к решениям на языке Python с использованием пространства решений, что позволит генерировать подсказки в том числе для сложных задач.

Елена Сергеевна изучила существующие подходы генерации подсказок для разных языков программирования, оценила их применимость для языка Python и выбрала подход с использованием пространства решений. Для реализации этого подхода необходимо было собрать данные о том, как студенты решают задачи по программированию, содержащие динамики решения задач. Еленой было создано расширение для сред программирования, основанных на IntelliJ Platform, которое записывало динамику решения задачи пользователем. При помощи этого инструмента был проведен сбор данных, в котором приняло участие более 120 человек. Собранные данные были обработаны и использованы для построения пространства решений, для сжатия которого был применен алгоритм каноникализации абстрактных синтаксических деревьев (AST), соответствующих фрагментам решений. Сам алгоритм генерации подсказок, разработанный Еленой, заключается в поиске подходящего фрагмента в пространстве решений для текущего фрагмента пользователя. Этот алгоритм учитывает удаленность текущего решения от финальных решений в пространстве, структуру решений, количество вершин в синтаксических деревьях фрагментов, опыт и возраст пользователя, количество пройденных тестов, популярность предлагаемого фрагмента у предыдущих пользователей. Алгоритм был протестирован на самой простой и самой сложной задаче из тех, что предлагались пользователям при сборе данных. Для этого вручную было размечено по 50 подсказок для каждой задачи. Алгоритм показал хорошие результаты на простой задаче (80.85% логичных подсказок) и достаточно хорошие на сложной (60% логичных подсказок), что может быть объяснено различиями в вариативности решений и в количестве собранных данных. Кроме того, собранные в процессе работы данные представляют особую ценность, так как до настоящего времени не существовало аналогичных наборов. Эти данные можно использовать для анализа паттернов поведения студентов при программировании, поиска типичных ошибок, непонятных концепций языка и так далее, что открывает много возможностей для будущей работы.

Елена Сергеевна провела очень большую работу и выполнила даже больше задач, чем я планировал успеть до защиты ВКР. Студентка продемонстрировала как высокую самостоятельность, так и умение работать в команде (проект выполнялся совместно с другой студенткой матмеха Анастасией Бирилло). Считаю, что работа соответствует всем требованиям к ВКР бакалавра и заслуживает оценки “отлично”.

Брыксин Т.А.,

к.т.н., доцент кафедры системного программирования СПбГУ