

УДК 519.713:004.423.24

ПРИМЕНЕНИЕ СИМВОЛЬНЫХ ВЫЧИСЛЕНИЙ ДЛЯ ПРЕОБРАЗОВАНИЯ ПРЕДСТАВЛЕНИЯ КОНЕЧНЫХ АВТОМАТОВ

Доморацкий Э. А. (Университет ИТМО)

Научный руководитель – к.т.н. Лаздин А. В.

(Университет ИТМО)

Конечные автоматы применяются во многих областях информатики. Одним из видов представления КА можно считать регулярные выражения. Предлагается алгоритм, способный автоматически преобразовывать любые КА в регулярные выражения.

Введение. Регулярные выражения широко используются в промышленном программировании для решения задач обработки текста. В информатике хорошо известны алгоритмы построения конечного автомата для разбора регулярных языков, преобразования недетерминированного конечного автомата в детерминированный конечный автомат и минимизации последнего. Однако, не очень много внимания уделяется задаче обратной – получения регулярного выражения, соответствующего КА. Несмотря на малую известность задачи, для её решения существуют алгоритмы, такие как метод удаления состояний, алгебраический метод Бжозовского и метод транзитивных замыканий.

Применение конечных автоматов обусловлено хорошей математической проработкой их применения в самых разных областях проектирования программных систем. В связи с этим значительный интерес представляют вопросы проектирования КА, и/или их последующей верификации. Однако в силу того, что существует несколько способов определения КА, возникают вопросы, связанные со взаимным преобразованием таких представлений. Следует отметить, что интерес в применении формальных методов обработки представляют так же недетерминированные конечные автоматы. Некоторые сетевые протоколы используют графоаналитическую спецификацию, в то же время рассматриваются вопросы верификации сетевых протоколов в терминах сетей Петри.

Недостатком таких решений является множественное, в рамках даже одного проекта, преобразование форм представления КА, начиная от его проектирования и заканчивая его моделированием или верификацией. Авторы предлагают единый подход, основанный на представлении КА (детерминированного или недетерминированного) в виде регулярного выражения. В основе этого подхода лежит концепция символьных вычислений, в рамках которой предполагается осуществлять все действия над КА. В данной работе предлагается реализация преобразования табличного представления КА в регулярное выражение средствами символьных вычислений.

Основная часть.

Входные данные. Входными данными алгоритма является конечный автомат (детерминированный или недетерминированный) и его начальное состояние. Для удобства структура КА записывается в виде системы уравнений в регулярных выражениях.

Поиск регулярного выражения. При наличии системы уравнений в регулярных выражениях и начального состояния КА поиск регулярного выражения, соответствующего данному КА, сводится к решению системы уравнений относительно неизвестной, соответствующей начальному состоянию.

Решение системы уравнений. Решение системы уравнений в регулярных выражениях заключается в применении леммы Ардена для избавления от неизвестных с обеих сторон уравнения и подстановке переменных. Таким образом, достаточно применить лемму Ардена к уравнениям, после чего подставить правую часть вместо неизвестной в других уравнениях.

Выбор порядка решения уравнений. Для успешной работы алгоритма достаточно, чтобы уравнение, в левой части которого находится искомая переменная (соответствующая первому состоянию КА), было последним. Однако, от порядка решения других уравнений зависит длина полученного в результате вычислений регулярного выражения. Для целей данной работы был выбран алгоритм сортировки на основе обхода в глубину графа зависимостей системы уравнений, который упорядочивает вершины (уравнения) в порядке, обратном глубине спуска.

Решение уравнений. Как было упомянуто выше, решение уравнений сводится к применению леммы Ардена и подстановки левой части в остальные уравнения. При наличии абстрактного синтаксического дерева правой части уравнения применение леммы заключается в выделении части, суффиксом которой является переменная из левой части. АСД состоит из пяти типов вершин: объединение, конкатенация, итерация (звезда Клини), строка и переменная (имя). Последние два типа могут быть только листьями дерева. Если выражение удалось поделить (существуют части с суффиксом и без), достаточно заменить правую часть уравнение на новую, в соответствии с леммой.

Выводы. Поставленная задача представления КА в виде регулярных выражений решена, то есть сделан шаг на пути к единому представлению и обработке КА для последующего их анализа, т.к. к полученным регулярным выражениям в будущем могут применяться те же методы символьных вычислений.

Доморацкий Э. А. (автор)

Подпись

Лаздин А. В. (научный руководитель)

Подпись