

Санкт-Петербургский государственный университет

Кафедра Системного программирования

Свитков Сергей Андреевич

Реализация поиска путей с
КС-ограничениями в рамках библиотеки
УС.QuickGraph

Курсовая работа

Научный руководитель:
ст. преп, к. ф-м. н. Григорьев С. В.

Санкт-Петербург
2016

Оглавление

| | |
|-----------------------------|----|
| Аннотация | 3 |
| Введение | 4 |
| 1. Обзор предметной области | 5 |
| 2. Постановка задачи | 6 |
| 3. Основная часть | 7 |
| 4. Эксперименты | 8 |
| Заключение | 9 |
| Список литературы | 10 |

Аннотация

Большинство промышленных языков для написания запросов к графовым базам данных являются регулярными, например, Cypher, используемый в Neo4J. Но регулярные языки не применимы в ряде задач, поэтому актуальным является создание контекстно-свободного (в дальнейшем — КС) языка запросов.

Предлагаемое в статье решение было реализовано как расширение библиотеки UC.QuickGraph. Полученные результаты могут быть применены в проектах, реализованных с использованием языка C# или F#. В данной работе рассматривается реализация контекстно-свободных запросов к ориентированным графам с помеченными ребрами.

Введение

Модель представления данных в виде графов, в частности, в виде орграфов с метками на ребрах, имеет широкую область применения и используется в биоинформатике, социальных исследованиях (например, при представлении социальных графов), semantic web, при реализации графовых баз данных.

При наличии представления данных в виде определенной структуры становится актуальным вопрос их обработки, а именно — получения из всего набора только тех данных, которые представляют какой-либо конкретный интерес. Для этого используются языки запросов. Из всего множества таких языков стоит выделить Cypher [1], применяемый при работе с Neo4J (другие языки запросов к графовым базам данных во многом схожи с ним, поэтому их рассмотрение можно опустить). Но данный язык является регулярным, а значит, не может применяться в некоторых задачах. Например, при решении задачи о поиске пар n поколения потомков от общего предка, нужно найти строки вида $parent^n child^n$. Такую строку нельзя задать с помощью регулярной грамматики, но можно с помощью КС-грамматики с правилами вывода $N \rightarrow parent\ child, N \rightarrow parentN\ child$.

Существуют работы, предлагающие различные подходы к реализации КС-запросов. Но данные решения в большинстве своем теоретические, а те, что реализованы на практике, предоставляют довольно бедный функционал, или же имеют слишком узкую специализацию. Поэтому было принято решение реализовать библиотеку, позволяющую использовать КС-грамматики в качестве языка запросов и поддерживающую возможность представления результата запроса в форме подграфа, множества путей или кратчайшего пути.

Реализация выполнялась как расширение библиотеки UC.QuickGraph [2] с использованием языка спецификаций грамматик YARD [3] для запросов. Полученное расширение предоставляет набор функций для исполнения запросов и представления результата запроса в удобной для пользователя форме.

1. Обзор предметной области

1

2. Постановка задачи

1

3. Основная часть

1

4. Эксперименты

1

Заключение

1

Список литературы

- [1] Neo4j. Cypher // Neo4j official page. — URL: <https://neo4j.com/developer/cypher/> (online; accessed: 22.11.2016).
- [2] YaccConstructor. YC.QuickGraph // YaccConstructor official page. — URL: <http://yaccconstructor.github.io/QuickGraph/> (online; accessed: 22.11.2016).
- [3] YaccConstructor. YARD // YaccConstructor official page. — 2015. — URL: <http://yaccconstructor.github.io/YaccConstructor/yard.html> (online; accessed: 22.11.2016).