

# ContextFree Wars: The RedisGraph Strikes Back

Arseniy Terekhov  
simpletondl@yandex.ru  
Saint Petersburg State University  
St. Petersburg, Russia

Vadim Abzalov  
!!!@!!!  
Saint Petersburg State University  
St. Petersburg, Russia

Vlada Poguzelskaya  
!!!@!!!  
Saint Petersburg State University  
St. Petersburg, Russia

Semyon Grigorev  
s.v.grigoriev@spbu.ru  
semyon.grigorev@jetbrains.com  
Saint Petersburg State University  
St. Petersburg, Russia  
JetBrains Research  
St. Petersburg, Russia



Figure 1: Episode IV: A New Hope

## ABSTRACT

A long time ago in a galaxy far far away...

## 1 INTRODUCTION

CFPQ is a way to use grammars.

CFPQ is widely spread and gain popularity last years.

Application for real-world data analysis is a problem. First of all, the performance problems. Jochem Kuijpers [2]. There are no full-stack solutions, only separated algorithms. For example, recently graph segmentation in data provenance analysis was reduced to CFPQ [? ], but authors faced the problem during the evaluation of the proposed approach: no one graph database support CFPQ. Thus, it is necessary to provide !!!

In [1] Rustam Azimov propose a matrix-based algorithm for CFPQ. This algorithm is one of promissing way to provide appropriate solution for real-world daata analysis: it was shown that !!! But the proposed algorithm

All-pairs is a classical problem. What about single-source and multiple-source? We propose a matrix-based multiple-source CFPQ algorithm.

Also, we provide full-stack support of CFPQ. We implement a Cypher query language extension<sup>1</sup> that allows one to express context-free constraints, and extend the RedisGraph graph database to support this extension. In our knowledge, it is the first full-stack implementation of CFPQ.

The following contribution.

<sup>1</sup>!!!

- (1) Multiple-source matrix-based CFPQ algorithm. Single-source as a partial case.
- (2) Evaluatideon of two versions of this algorithm.
- (3) RedisGraph extending to provide full-stack support of CFPQ.

## 2 PRELIMINARIES

In this section we introduce common definitions in graph theory and formal language theory which will be used in this paper. Also, we provide brief description of Azimov's algorithm which will be used as a base of our solution.

### 2.1 Graphs

Labelled digraph, matrices, ...

*Definition 2.1.* Digraph

### 2.2 Languages

Grammars, normal forms, ...

*Definition 2.2.* Digraph

### 2.3 Matrix-Based Algorithm

Description

Pseudocode.

What about arbitrary (not CNF) CFG?

## 3 MATRIX-BASED MULTIPLE-SOURCE CFPQ ALGORITHM

New algo description.

### 3.1 Implementation Details

Algo implementation details: python, graphBLAS, smthng else?  
Two versions.

### 3.2 Algorithm Evaluation

And comparison. With combinators, GLL (.NET version).

Evaluation setup.

Tables.

Results.

Conclusion.

## 4 REDISGRAPH EXTENDING

In order to provide full-stack support of CFPQ we choose an appropriate graph DB.

### 4.1 Cypher

parser extending, proposal, ....

Examples of queries.

### 4.2 RedisGraph

CFPQ to matrix expressions, etc.

Limits, restrictions, etc.

### 4.3 Evaluation

Small basic evaluation on real-world graph (geo?). In order to show, that performance is reasonable.

## 5 CONCLUSION

Conclusion

Future research

Formal translation of Cypher to linear algebra. In order to formalize limits and restrictions. There are works on a subset of SPARQL to linear algebra translation. But they are very limited.

For real-world solutions is important to provide a unified algorithm for both RPQ and CFPQ. A matrix-based algorithm is not a better choice. Experiments on a unified evaluation of RPQ and CFPQ (tensors?).

Multiple-source for Neo4j (non-linear algebra based approaches evaluation).

More applications.

## REFERENCES

- [1] Rustam Azimov and Semyon Grigorev. 2018. Context-free Path Querying by Matrix Multiplication. In *Proceedings of the 1st ACM SIGMOD Joint International Workshop on Graph Data Management Experiences & Systems (GRADES) and Network Data Analytics (NDA) (GRADES-NDA '18)*. ACM, New York, NY, USA, Article 5, 10 pages. <https://doi.org/10.1145/3210259.3210264>
- [2] Jochem Kuijpers, George Fletcher, Nikolay Yakovets, and Tobias Lindaaaker. 2019. An Experimental Study of Context-Free Path Query Evaluation Methods. In *Proceedings of the 31st International Conference on Scientific and Statistical Database Management (SSDBM '19)*. ACM, New York, NY, USA, 121–132. <https://doi.org/10.1145/3335783.3335791>