

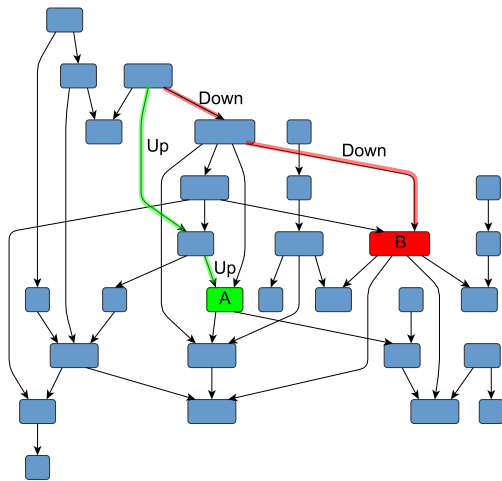
# Multiple-Source Context-Free Path Querying in Terms of Linear Algebra

Arseniy Terekhov, Vlada Pogozhelskaya, Vadim Abzalov, Timur Zinnatulin, **Semyon Grigorev**

JetBrains Research, Programming Languages and Tools Lab  
Saint Petersburg University

March 24, 2021

# Formal Language Constrained Path Querying



Navigation through a graph

- Are nodes A and B on the same level of hierarchy?
- Is there a path of form **Up<sup>n</sup> Down<sup>n</sup>**?
- Find all paths of form **Up<sup>n</sup> Down<sup>n</sup>** which start from the node A

# Context-Free Path Querying (CFPQ)

- Applications
  - ▶ Static code analysis
  - ▶ Graph segmentation
- Theory
  - ▶ !!!
  - ▶ !!!

# Context-Free Path Querying (CFPQ)

- Applications
  - ▶ Static code analysis
  - ▶ Graph segmentation
- Theory
  - ▶ !!!
  - ▶ !!!
- Integration with real-world systems
  - ▶ !!!!
  - ▶ Kuijpers for Neo4j: too slow to be practical

# Context-Free Path Querying (CFPQ)

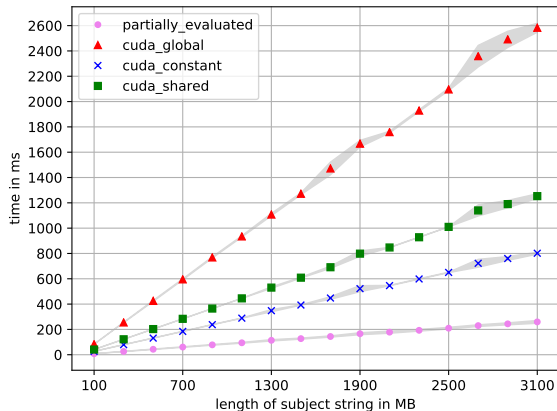
- Applications
  - ▶ Static code analysis
  - ▶ Graph segmentation
- Theory
  - ▶ !!!
  - ▶ !!!
- Integration with real-world systems
  - ▶ !!!!
  - ▶ Kuijpers for Neo4j: too slow to be practical
  - ▶ RedisGraph

# Linear Algebra Based CFPQ Algorithm

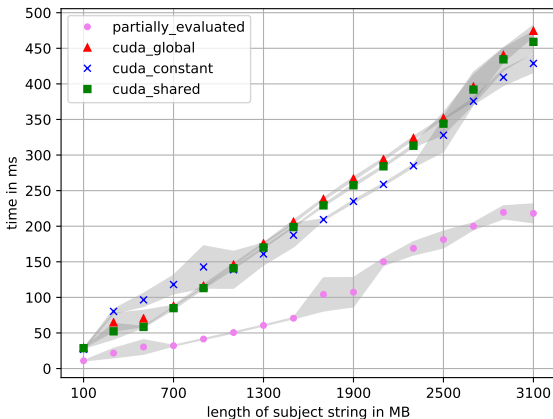
- Definition
  - ▶ Special DSL which can be specialized and compiled
  - ▶ Ahead-of-time specialization
- Impractical memory consumption
  - ▶ Naïve multiple substring matching
  - ▶ 2D convolution
- Context-Free grammars are too hard to be used by end-users
  - ▶ **GTX-1070**: Pascal architecture, 8GB GDDR5, 1920 CUDA cores
  - ▶ **Tesla T4**: Turing architecture, 16GB GDDR6, 2560 CUDA cores

# Proposed Solution

- RedisGraph
- Cypher<sup>1</sup>
- Multiple-Source CFPQ



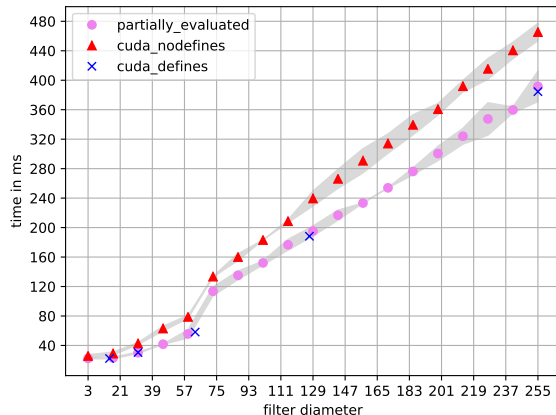
Results for GTX-1070



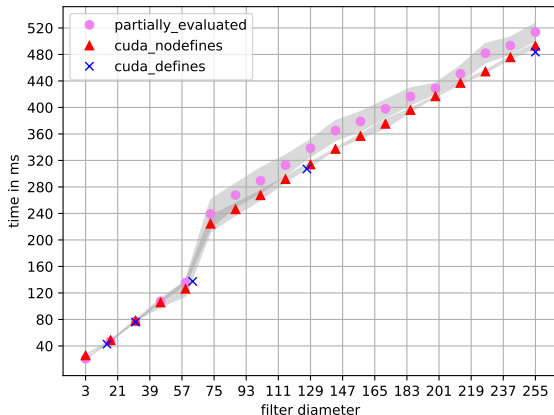
Results for Tesla T4

# Multiple-Source CFPQ

- Application: image processing
- Subject image: random image of size 1GB
- Filters: random square filters with diameter 3 to 255



Results for GTX-1070

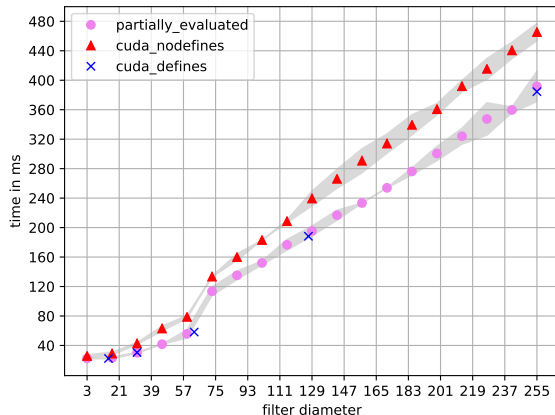


Results for Tesla T4

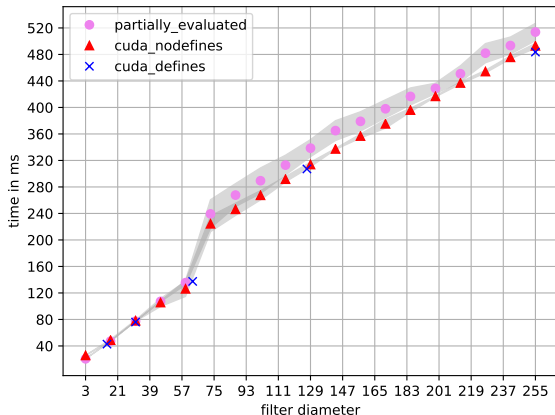


# Implementation Details

- Application: image processing
- Subject image: random image of size 1GB
- Filters: random square filters with diameter 3 to 255



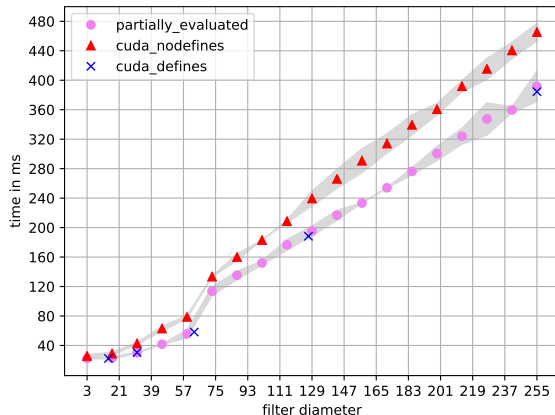
Results for GTX-1070



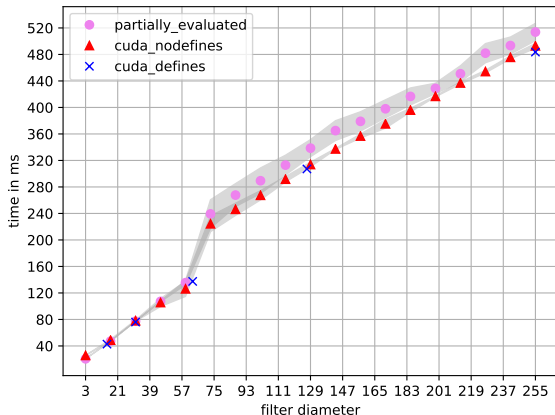
Results for Tesla T4

# Cypher Extension

- Application: image processing
- Subject image: random image of size 1GB
- Filters: random square filters with diameter 3 to 255



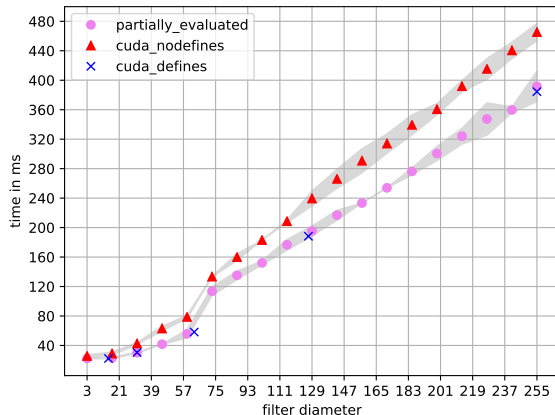
Results for GTX-1070



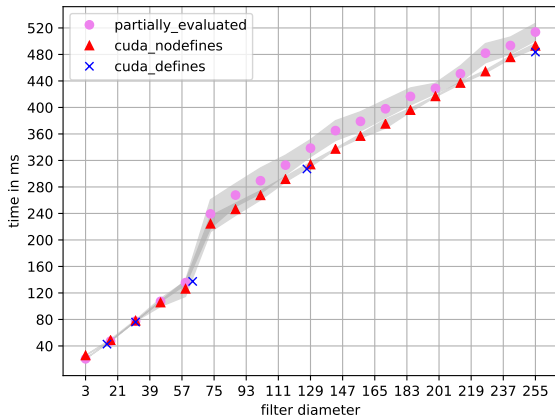
Results for Tesla T4

# Queries Examples

- Application: image processing
- Subject image: random image of size 1GB
- Filters: random square filters with diameter 3 to 255



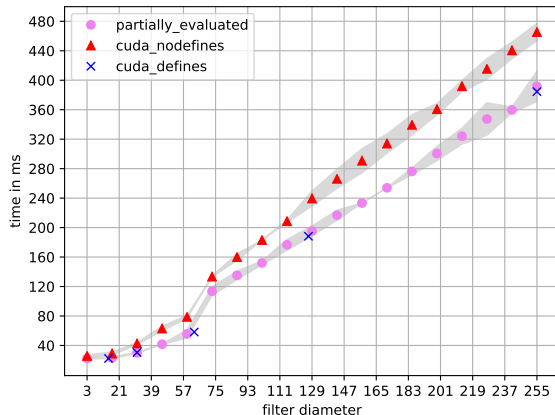
Results for GTX-1070



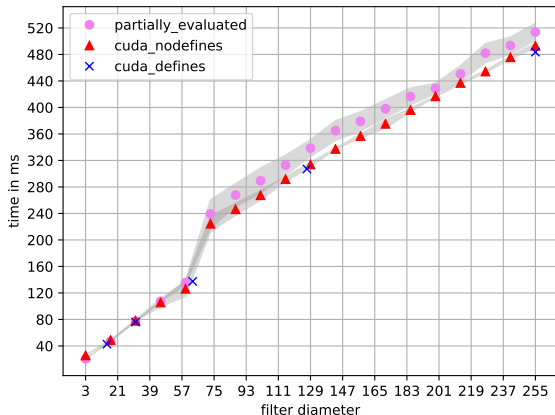
Results for Tesla T4

# Evaluation Setup

- Application: image processing
- Subject image: random image of size 1GB
- Filters: random square filters with diameter 3 to 255



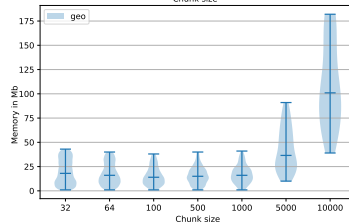
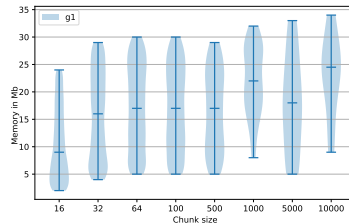
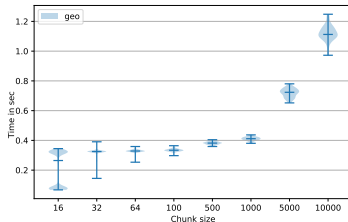
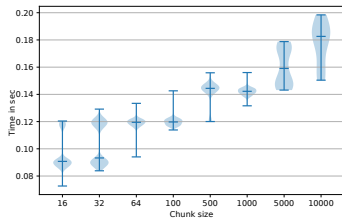
Results for GTX-1070



Results for Tesla T4

# Evaluation Results

`eclass_514en`



- Full-stack support for CFPQ in real-world graph query languages on the top of real-world graph database
- Reasonable performance of context-free path queries
- Filters: random square filters with diameter 3 to 255

# Future Research

- Migration to CUDA C partial evaluator
  - ▶ LLVM.mix: partial evaluator for LLVM IR
- Reduction of specialization overhead
  - ▶ To be applicable in run-time
- Integration with shared memory register spilling
  - ▶ “RegDem: Increasing GPU Performance via Shared Memory Register Spilling” (Putt Sakdhnagool et.al. 2019)
- Evaluation on real-world examples
  - ▶ Homology search in bioinformatics
  - ▶ Regular expression matching for traffic analysis, log processing
  - ▶ Graph database querying
  - ▶ Ray tracing, path tracing

# Contact Information

- Semyon Grigorev:
  - ▶ s.v.grigoriev@spbu.ru
  - ▶ Semen.Grigorev@jetbrains.com
- Arseniy Terekhov: simpletondl@yandex.ru
- Vlada Pogozhelskaya: pogozhelskaya@gmail.com
- Vadim Abzalov: vadim.i.abzalov@gmail.com
- Timur Zinnatulin: teemychteemych@gmail.com
  
- Try it out (Docker image with all included): !!!
- Sources of RedisGraph extended with CFPQ: !!!
- Sources of Cypher parser extended with path patterns: !!!

Thanks!