Graph analysis

Semyon Grigorev

Saint Petersburg State University

April 28, 2022

Research areas

- High-performance graph analysis
  - Formal languages constrained path querying
  - High-level languages for high-performance computing
- Path problems with constraints
  - Formal languages constrained path querying
  - High-level languages for high-performance computing
- Graph databases
  - Formal languages constrained path querying
  - High-level languages for high-performance computing

Team

- Semyon Grigorev (Lead)
  - PhD (2016)
  - Associate professor (2016, SPbSU)
  - s.v.grigoriev@spbu.ru
  - High-performance graph analysis
  - Graph databases
  - dblp: https://dblp.org/pid/181/9903.html
- Ekaterina Shemetova
  - PhD student
  - Path problems with constraints
  - Fine-grained complexity
  - Dynamic graph problems
- Rustam Azimov
  - PhD student
  - Linear algebra based graph analysis
  - GraphBLAS API
  - Algebraic path problem

Team: Master Students

- Alexandra Istomina
  - Master student
  - Fine-grained complexity
  - Path problems with constraints
  - Algebraic path problem
- Vladimir Kutuev
  - Master student
  - Linear algebra based graph analysis
  - GraphBLAS API
  - Parallel programming
- Julia Susanina
  - Master student
  - Linear algebra based graph analysis
  - Probabilistic graph analysis
  - GPGPU programming

High-Performance Graph Analysis

- Linear algebra based algorithms for graph analysis
  - Parallel algorithms on CPU and GPGPU
  - Sparse linear algebra for graph analysis
  - GraphBLAS API[1]

[1] https://graphblas.org/

High-Performance Graph Analysis

- Linear algebra based algorithms for graph analysis
  - Parallel algorithms on CPU and GPGPU
  - Sparse linear algebra for graph analysis
  - GraphBLAS API[1]
- Research directions
  - GraphBLAS-based algorithms design, implementation and evaluation
  - Portable multi-GPGPU implementation of GraphBALS-like API
  - GraphBLAS API analysis

[1] https://graphblas.org/

Path Problems With Constraints: Algebraic Path Problems

- Semiring-like structures to specify constraints on paths
  - Reachability — boolean semiring
  - Shortest paths — tropical semiring
  - ...

Path Problems With Constraints: Algebraic Path Problems

- Semiring-like structures to specify constraints on paths
  - Reachability — boolean semiring
  - Shortest paths — tropical semiring
  - ...
- Linear algebra friendly algorithms
  - Transitive closure using matrix-matrix multiplication
  - APSP using matrix-matrix multiplication
  - BFS-like traversals using matrix-vector multiplication
  - ...

Path Problems With Constraints: Algebraic Path Problems

- Semiring-like structures to specify constraints on paths
  - Reachability — boolean semiring
  - Shortest paths — tropical semiring
  - ...
- Linear algebra friendly algorithms
  - Transitive closure using matrix-matrix multiplication
  - APSP using matrix-matrix multiplication
  - BFS-like traversals using matrix-vector multiplication
  - ...
- Compositionality
  - Having two semirings one can create a new one
  - Single solution for similar problems
    - ⋆ Generic solution
    - ⋆ Configurable solution

Formal Language Constraint Path Querying

- Particular case of algebraic path problem
  - Multiplication is not associative
  - Multiplication is not commutative
  - ...

Formal Language Constraint Path Querying

- Particular case of algebraic path problem
  - Multiplication is not associative
  - Multiplication is not commutative
  - ...
- Examples
  - Regular path querying (RPQ)
  - Context-free path querying (CFPQ)

Formal Language Constraint Path Querying

- Particular case of algebraic path problem
  - Multiplication is not associative
  - Multiplication is not commutative
  - ...
- Examples
  - Regular path querying (RPQ)
  - Context-free path querying (CFPQ)
- Applications
  - Graph analysis
  - Interprocedural static code analysis
  - Graph database querying

Formal Language Constraint Path Querying

- Particular case of algebraic path problem
  - Multiplication is not associative
  - Multiplication is not commutative
  - ...
- Examples
  - Regular path querying (RPQ)
  - Context-free path querying (CFPQ)
- Applications
  - Graph analysis
  - Interprocedural static code analysis
  - Graph database querying
- Research directions
  - New algorithms development
  - Complexity analysis
  - New classes of languages investigation
  - High performance algorithms implementation and evaluation

Our Results

- Tools
  - Spla: sparse linear algebra framework for multi-GPU computations based on OpenCL
  - SPbLA: library of GPGPU-powered sparse boolean linear algebra operations

Our Results

- Tools
  - Spla: sparse linear algebra framework for multi-GPU computations based on OpenCL
  - SPbLA: library of GPGPU-powered sparse boolean linear algebra operations
  - CFPQ_PyAlgo: set of GraphBLAS-based FLPQ algorithms
  - LDBC Graphalytics extension for evaluation of formal language constrained path querying

Our Results

- Tools
  - Spla: sparse linear algebra framework for multi-GPU computations based on OpenCL
  - SPbLA: library of GPGPU-powered sparse boolean linear algebra operations
  - CFPQ_PyAlgo: set of GraphBLAS-based FLPQ algorithms
  - LDBC Graphalytics extension for evaluation of formal language constrained path querying
  - GLL4Graph: CFPQ for Neo4j
  - CFPQ for RedisGraph
- Papers (> 10)
  - SPbLA: The Library of GPGPU-Powered Sparse Boolean Linear Algebra Operations (GrAPL@IPDPS)
  - Evaluation of the context-free path querying algorithm based on matrix multiplication (GRADES-NDA@SIGMOD)
  - Multiple-Source Context-Free Path Querying in Terms of Linear Algebra (EDBT, Core A)
  - Context-free path querying by matrix multiplication (GRADES-NDA@SIGMOD)

Possible Ways for Collaboration

- Algebraic Path Problem framework applicability for network analysis
  - Which constraints can be specified in terms of semirings?
    - Length minimality
    - Nodes to visit
    - ...
  - Is it flexible enough?
- High-performance network analysis
  - GraphBLAS-based solution
  - Algorithms development and analysis
  - Algorithms implementation and evaluation