Universidade Federal do Rio Grande do Norte
Centro de Ciências Exatas e da Terra
Depto de Informática e Matemática Aplicada
Programa de Pós-graduação em Sistemas e Computação
Campus Universitário - Lagoa Nova
59072-970 Natal, RN
T (55) (84) 3215 3814 ext. 232
F (55) (84) 3215 3813
mam@dimap.ufrn.br

REPORT

**On the dissertation presented by Rustam Azimov,
entitled « Context-Free Path Queries Using Linear Algebra »
submitted for the Candidate of Science in Physics and Mathematics degree
in specialization 2.3.5 Mathematical and software support for computers, complexes and computer networks**

The thesis investigates and proposes new algorithms to implement context-free path queries in graph databases (as well as graph data structures in general). The proposals explore linear algebra and use matrices to represent data graphs. This allows the use of well known algorithms/operations over matrices to be adapted as query algorithms. Among the advantages of this approach there are the ample theoretical background on matrix manipulation, as well as the efficient implementation for these algorithms when using GPUs.

The motivation for the research work relies on the utility of context-free queries (and, in general, formal language-constrained path query problems) in a number of practical problems that arise in the programming languages and database areas.

The dissertation begins with a very good, didactic revision on the basic concepts used in the work. This includes the mathematical concepts about graphs and matrices, and formal languages and automata. The chapter also includes a revision of existing context-free path query algorithms, and graph analysis using linear algebra. The chapter ends with a comparative analysis of tools/libraries for linear algebra operations and their deployment in parallel architectures (including CPU and GPU implementations).

The main contribution of the dissertation is presented in chapters 2 to 4. The presentation begins with a general, linear algebra-based approach to context-free path queries. This approach defines the basis for the subsequent implementations in chapter 3 and 4.

Chapter 3 is dedicated to the definition, proof of correctness and implementation of a context-free path query based on matrix multiplication. The algorithm requires a grammar in weak Chomsky normal form (WCNF). Reachability aspects, as well as single-path and all-path semantics of the queries are analysed. Mechanisms for achieving these query semantics are discussed. The correctness and asymptotic behaviour of the approach are also shown for each query semantics. The chapter includes a small but very detailed, step-by-step demonstration of the algorithm, including its matrix manipulation for each query semantics. Finally, chapter 3 ends with a discussion on the implementation of the proposals.

33-06-274 от 07.03.2023

Universidade Federal do Rio Grande do Norte
Centro de Ciências Exatas e da Terra
Depto de Informática e Matemática Aplicada
Programa de Pós-graduação em Sistemas e Computação
Campus Universitário - Lagoa Nova
59072-970 Natal, RN

Chapter 4 defines a Kronecker product-based version of the path query algorithm that does not require a grammar in WCNF. As in the previous chapter, reachability, as well as single-path, and all-path semantics are considered. Correctness and asymptotic time complexity are also addressed. The chapter includes a detailed example as well as a description of the implementation.

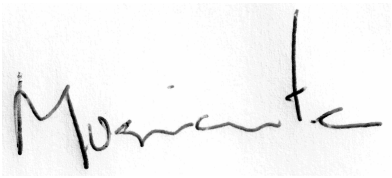
Chapter 5 is devoted to the performance evaluation of the approach (in its several versions). It includes figures for queries in a number of contexts and cases, showing that the approach outperforms previous proposals, both in time and memory consumption.

Chapter 6 is a succinct, *a posteriory* comparison of the proposed method with other approaches. It is informative and places the work in context with other scientific work in the area.

Considered as a whole, the thesis represents a step forward the state of the art in the area of non-regular path query languages. The contents is compatible with a PhD work, and the publications of the candidate are also consistent with a world-level expectations for this kind of work.

In this context, I can fairly conclude that the thesis entitled “Context-Free Path Queries Using Linear Algebra”, by Rustam Azimov, deserves to be awarded the degree of PhD in Computer Science.

Natal, Brazil, February 13th, 2023



Prof. Martin A. Musicante
Computer Science Department
Federal University of Rio Grande do Norte (UFRN)