On the Complexity of Finding Set Repairs for Data-Graphs (Abstract Reprint)

Sergio Abriola[1] , **Maria Vanina Martinez**[1] , **Nina Parda**[1] , **Santiago Cifuentes**[2] and **Edwin Pin Baque**[2]

[1]ICC CONICET, Universidad de Buenos Aires
[2]Departamento de Computación, Universidad de Buenos Aires
{sabriola, mvmartinez, npardal, scifuentes, epin}@dc.uba.ar

Abstract Reprint. This is an abstract reprint of a journal by [Abriola *et al.*, 2023].

Abstract

In the deeply interconnected world we live in, pieces of information link domains all around us. As graph databases embrace effectively relationships among data and allow processing and querying these connections efficiently, they are rapidly becoming a popular platform for storage that supports a wide range of domains and applications. As in the relational case, it is expected that data preserves a set of integrity constraints that define the semantic structure of the world it represents. When a database does not satisfy its integrity constraints, a possible approach is to search for a 'similar' database that does satisfy the constraints, also known as a repair. In this work, we study the problem of computing subset and superset repairs for graph databases with data values using a notion of consistency based on having a set of Reg-GXPath expressions as integrity constraints. We show that for positive fragments of Reg-GXPath these problems admit a polynomialtime algorithm, while the full expressive power of the language renders them intractable.

References

[Abriola *et al.*, 2023] Sergio Abriola, María Vanina Martínez, Nina Pardal, Santiago Cifuentes, and Edwin Pin Baque. On the complexity of finding set repairs for data-graphs. *J. Artif. Intell. Res.*, 76:721–759, 2023.