

QFix: Demonstrating Error Diagnosis in Query Histories

Submitted by grigby1 on Tue, 03/07/2017 - 1:56pm

Title	QFix: Demonstrating Error Diagnosis in Query Histories
Publication Type	Conference Paper
Year of Publication	2016
Authors	Wang, Xiaolan , Meliou, Alexandra , Wu, Eugene
Conference Name	Proceedings of the 2016 International Conference on Management of Data
Publisher	ACM
Conference Location	New York, NY, USA
ISBN Number	978-1-4503-3531-7
Keywords	data cleaning , mixed-integer linear programming , pubcrawl170201 , query provenance

Abstract

An increasing number of applications in all aspects of society rely on data. Despite the long line of research in data cleaning and repairs, data correctness has been an elusive goal. Errors in the data can be extremely disruptive, and are detrimental to the effectiveness and proper function of data-driven applications. Even when data is cleaned, new errors can be introduced by applications and users who interact with the data. Subsequent valid updates can obscure these errors and propagate them through the dataset causing more discrepancies. Any discovered errors tend to be corrected superficially, on a case-by-case basis, further obscuring the true underlying cause, and making detection of the remaining errors harder. In this demo proposal, we outline the design of QFix, a query-centric framework that derives explanations and repairs for discrepancies in relational data based on potential errors in the queries that operated on the data. This is a marked departure from traditional data-centric techniques that directly fix the data. We then describe how users will use QFix in a demonstration scenario. Participants will be able to select from a number of transactional benchmarks, introduce errors into the queries that are executed, and compare the fixes to the queries proposed by QFix as well as existing alternative algorithms such as decision trees.

URL <http://doi.acm.org/10.1145/2882903.2899388>

DOI [10.1145/2882903.2899388](https://doi.org/10.1145/2882903.2899388)

Citation Key wang_qfix:_2016



[data cleaning mixed-integer linear programming pubcrawl170201 query provenance](#)
