Editor's Notes

Welcome to the September 2019 issue of the ACM SIGMOD Record!

This issue starts with the Database Principles column featuring an article by Wijsen that discusses the problem of query answering on inconsistent databases. A database is called inconsistent if it violates some integrity constraints; the key issue is what information can and cannot be inferred from inconsistent data in the process of query answering. Intuitively, an inconsistent database can be viewed as a collection of consistent databases, each of which results from some minimal corrections, or repairs, to the original data. Then, an answer to a query on such a database is called consistent if it can be obtained by processing the query on each of the repaired databases. The article summarizes core concepts and theoretical developments that have arisen in foundations of consistent query answering in the past twenty years. The report on the rich body of theoretical results covers studies for common classes of integrity constraints and emphasizes results in computational and descriptive complexity. In general, consistent query answering is intractable. However, as the intractability usually arises in worst-case results, the article calls for research in refinement of the complexity results for practically important cases. The article also discusses open problems, and provides references on related areas.

The Research Articles column features two articles. The first article, by McCullough, Mokfi, and Almaeenjad, studies the problem of accuracy of SQL software for statistical purposes. The study focuses on testing elementary statistical calculations with a collection of benchmark tests known as Wilkinson's tests. These tests have a long track record of being applied to uncover flaws in statistical packages; this article describes their application to six well-known SQL packages. The article discusses the flaws that were identified in the analysis of statistical functions in the SQL packages, provides pointers to relevant algorithms, and gives the best choice of data types for statistical purposes in each package. The authors call for developers to fix the discovered errors, and propose that accuracy of algorithms be incorporated into the SQL standards.

The second article in the Research Articles column, by Cavero, Vela, and Cáceres, discusses the problem of simulating SQL assertions via materialized views. While assertions have been in the SQL standard since 1992 as a powerful means of specifying cross-row and cross-table constraints, they are usually not supported in commercial database-management systems, and are typically implemented by triggers or application programs. The article shows how assertions can be simulated using materialized views that count the number of violations of the assertion's conditions. With the help of a series of tests, the authors demonstrate that materialized views are easier to program and less error prone than triggers or application programs, as well as more efficient than triggers in some situations. The article provides recommendations on application scenarios that can benefit from the proposed assertion-codification approach, and specifies relevant DMBS requirements.

The Distinguished Profiles column features Michael Franklin, inaugural holder of the Liew Family Chair of Computer Science and senior advisor to the provost on computation and data science at the University of Chicago. Before that, for many years, Mike was a professor at Berkeley, where he also served as a chair of the Computer Science division. Mike was a co-founder and director of the Algorithms, Machines, and People Lab at Berkeley, better known as the AMPLab, a leading academic big data analytics research center that received a National Science Foundation CISE "Expeditions in Computing" award. At the AMPLab, Mike was one of the creators of the Spark (now Apache Spark) data analytics and machine learning platform. Mike is a Fellow of the Association for Computing Machinery and a two-time recipient of the ACM SIGMOD "Test of Time" award. Mike's Ph.D.

is from the University of Wisconsin Madison. In this interview, Mike talks about his experience with building computer science and data science at the University of Chicago in a way that integrates these disciplines into the fabric of the university, with opportunities for people with widely varying interests to work together. He shares his views on the prospects for real-time streaming analytics, and discusses how people fit into overall systems architectures. Mike also outlines aspects of computational and data literacy that an educated person in the 21st century needs to know, gives advice on research, and shares information about his own advising style and research.

The Reports column features two articles. The first article, by Palpanas and Beckmann, reports on the First and Second Interdisciplinary Time Series Analysis Workshops (ITISA), which took place in Paris in June and December 2016. Time-series data arise naturally in many applications, and their analysis can push the computational power and other resources to their limits. Over 80 participants in the two ITISA workshops participated in discussions of the challenges and requirements on the new technologies and algorithms. The workshops included 14 keynote talks, hands-on sessions, and panel discussions. The article summarizes the ideas presented and discussed in the workshops, highlighting the relevant state-of-the-art techniques and advancements in time-series management and analysis. The ITISA programs and slides are available from the authors of the article.

The second article in the Reports column, by Kondylakis, Stefanidis, and Rao, reports on the outcomes of the First International Workshop on Semantic Web Technologies for Health Data Management (SWH). The workshop took place in 2018 in Monterey, CA USA, in conjunction with the International Semantic Web Conference. The SWH workshop aimed to bring together an interdisciplinary audience, to discuss challenges in healthcare data management and to propose novel and practical solutions for next-generation data-driven healthcare systems. The article summarizes the outcomes of the workshop, and outlines key observations and emerging research directions.

On behalf of the SIGMOD Record Editorial board, I hope that you enjoy reading the September 2019 issue of the SIGMOD Record!

Your submissions to the SIGMOD Record are welcome via the submission site:
https://mc.manuscriptcentral.com/sigmodrecord

Prior to submission, please read the Editorial Policy on the SIGMOD Record's website:
https://sigmodrecord.org/sigmod-record-editorial-policy/

Rada Chirkova

September 2019

Past SIGMOD Record Editors: