

GUEST EDITORS' FOREWORD

The Eighteenth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (PODS 1999) was held in Philadelphia, Pennsylvania, on May 31–June 2, 1999. The conference call for papers attracted 116 submissions. The nine papers in this special issue were selected from the thirty-two papers presented in Philadelphia. Their authors were invited to submit more complete and revised versions for this issue for a round of careful refereeing.

The paper titled “Tracking Join and Self-Join Sizes in Limited Storage,” by Alon, Gibbons, Matias, and Szegedy, examines the problem of incrementally approximating the size of join and self-join results when storage is limited. It presents several algorithms, accompanied by both theoretical and experimental results.

“A Framework for Measuring Differences in Data Characteristics,” by Ganti, Gehrke, Ramakrishnan, and Loh, provides a framework and algorithms to measure differences between data sets with respect to the data-mining models they induce. Its approach is quite general and captures several types of models as well as several measures of difference.

The paper “Maximizing Sharing of Protected Information,” by Dawson, De Capitani Di Vimercati, Lincoln, and Samarati, deals with the issue of security in database systems, and particularly with the issue of preventing sensitive information from leaking to the outside. It contains an approach that classifies information at the desired level and enforces constraints on what a user may derive through inference or data associations.

In “Aggregate Operators in Constraint Query Languages,” Benedikt and Libkin extend constraint query languages with both standard and spatial aggregate operators, including approximate operators. They present several results regarding expressiveness and closure of such extended languages.

“Querying Incomplete Information in Semistructured Data,” by Kanza, Nutt, and Sagiv, considers queries on semistructured data (as in xml) which admit incomplete (i.e., not fully instantiated) answers. The authors tackle the difficulties and choices in defining the semantics of such queries, and they develop algorithms for them.

Semistructured data are also the subject of “Rewriting of Regular Expressions and Regular Path Queries,” by Calvanese, De Giacomo, Lenzerini, and Vardi. They address the problem of rewriting regular-expression (path) queries in the context of semi-structured data, and provide a method that modifies such a query and generates an optimal rewriting based on a collection of available views.

In the paper “Inherent Complexity of Recursive Queries,” Cosmadakis points out the limitations of Datalog optimisations by proving lower bounds on the

complexity of evaluating certain Datalog queries. He adapts to this new goal the Ehrenfeucht–Fraïssé games classically used to establish inexpressibility in first-order logic.

The paper titled “Polymorphic Type Inference for the Relational Algebra,” by Van den Bussche and Waller, revisits relational algebra from the point of view of type inference. Given a relational query, the authors show how to represent, succinctly, all possible relational schemas (a set of attributes for each relation) so that the query is well-typed.

Finally, “Hypertree Decompositions and Tractable Queries,” by Gottlob, Leone, and Scarcello, deals with an even more primitive query language: conjunctive queries. They come up with “the right” parameter of a conjunctive query (its *hypertree width*) whose boundedness can be detected efficiently, and also implies tractability of important computational problems related to the query.

We believe that the papers in this special issue reflect the diversity of problems that are being addressed currently by the database research community, and database theorists in particular. A mix of rigorous investigations of novel domains and papers revisiting time-honored themes in a new light, they represent a cross-section of an exciting and rich field as it expands into new areas of data management. We would like to thank the authors of all papers for their efforts. Our deep appreciation also goes to the program committee members of PODS 1999 and especially to those who participated in the extra reviewing phase that helped in improving the papers in this issue.

Yannis Ioannidis

Univ. of Athens, Hellas

Christos Papadimitriou

Univ. of California-Berkeley, CA, USA