

Rank-aware Query Processing and Optimization

Ihab F. Ilyas

School of Computer Science

University of Waterloo

Waterloo, Ontario N2L 3G1, Canada

ilyas@uwaterloo.ca

Walid G. Aref*

Department of Computer Science

Purdue University

West Lafayette IN 47907-1398, USA

aref@cs.purdue.edu

Efficient execution of ranking query is increasingly becoming a major *challenge for database technology*. Nowadays, many applications have requirements that can only be matched by a *combination* of information retrieval systems and DBMSs. DBMSs provide efficient update, indexing, concurrency and recovery. On the other hand, IR on text and multimedia requires techniques involving uncertainty and *ranking for effective retrieval*. A true integration is likely to require significant changes in the standard database techniques for indexing and query optimization and may require new query languages.

The main goal of this seminar is to give an in-depth look on supporting ranking queries as an increasingly interesting area of research. We cover the state-of-the-art techniques in research prototypes and industry-strength database engines for efficient handling of ranking and top-*k* queries. We give an inclusive background on ranking, voting and rank-aggregation algorithms. Then we give a detailed coverage of ranking query models—covering top-*k* selection and top-*k* join queries, and the various approaches recently proposed by researchers to support these queries in database systems. We focus primarily on how to integrate ranking as a new query processing and optimization dimension, with the aim of supporting ranking queries as a basic and core functionality. The seminar identifies several challenges that need to be addressed towards a true support for ranking and effective retrieval in database management systems.

This seminar is targeted at general database researchers. The seminar is also of interest to audience with industrial background as it describes and summarizes different attempts to integrate new functionalities in industrial database management systems. The seminar gives several motivating examples and chal-

lenging applications that are in real need for efficient handling of ranking queries. More importantly, the seminar highlights some interesting challenges in rank-aware query processing and optimization.

Speakers

Ihab F. Ilyas is an assistant professor at the school of computer Science, University of Waterloo. He obtained his Ph.D. degree in 2004 from Purdue university. His main research interests include advanced query processing and optimization, self-managing and adaptive computing, and non-traditional database systems (e.g., multimedia and spatial databases). In his recent Ph.D. thesis, he has introduced novel query processing and optimization techniques for top-*k* queries in relational database systems. For more information, visit <http://db.uwaterloo.ca/~ilyas>

Walid G. Aref is an associate professor of computer science at Purdue University. His research interests are in developing database technologies for emerging applications, e.g., spatial, multimedia, genomics, and sensor databases. He is also interested in indexing, data mining, and geographic information systems (GIS). His research has been supported by the NSF, Purdue Research Foundation, CERIAS, Panasonic, and Microsoft Corp. In 2001, he received the CAREER Award from the National Science Foundation. He is in the editorial board of the VLDB Journal and is a member of the ACM and the IEEE. For more information, visit <http://www.cs.purdue.edu/~aref>

*This author acknowledges the support from the NSF under grant numbers IIS-0093116 and IIS-0209120.