Changing Bulletin Editors

The Bulletin practice is to appoint editors for two years. Each editor is responsible for producing two issues, one per year, during that time. Among the current editors, Surajit Chaudhuri and Donald Kossmann have now completed their two issues and are due to “retire”. Our profession depends on hardworking volunteers for the vitality of publications like the Bulletin. So I would like to thank both Surajit and Donald for their hard work and very successful results. Producing high quality issues of the Bulletin is not an accident. It is the result of the work of very capable issue editors.

I am now delighted to announce the appointment of two new editors, Sunita Sarawagi and Alon Levy. Sunita and Alon both have outstanding research reputations.

Sunita Sarawagi is currently on the faculty in the School of Information Technology at IIT Bombay. She has done research in data warehousing, OLAP, data mining and tertiary storage. Her career began as a UC Berkeley student, following which she went to the IBM Almaden Research Center. Sunita was co-author of a paper that received a 1998 SIGMOD best paper award.

Alon Levy is a Seattle neighbor of mine as he is on the faculty of the Computer Science and Engineering Department at the University Washington. Alon’s expertise includes data integration, web-site management, semi-structured data, query optimization, and the interaction of databases with AI. He is a graduate of Stanford and was on the technical staff at ATT Bell Labs. In June he co-founded a company building tools for XML data.

This Issue

Databases are increasingly used for data analysis. When done ”on-line”, this is referred to as OLAP, a field that has become increasingly important over the past five or more years. The data cube is an important part of OLAP work, but by no means the only part. The current issue looks at how summary data can be maintained and retrieved, which is surely relevant in a very broad and important way to data analysis, including the data cube.

In a certain sense, most of the work that is reported in the current issue is related to performance. Users want to summarize data so that they can browse it, explore it, understand it. The work reported here is mostly about how this summarizing can be done efficiently. Were efficiency not an issue, the functionality would not require much more than a relational database. But performance is, of course, critically important. Indeed, we need real insights in this area, where approximation and sampling play a frequent role in making summarization feasible.

Amr El Abbadi,the issue editor, has gathered for this issue a diverse collection of techniques that address the retrieval and maintenance of summary information. These include histograms, sampling, incremental maintence of cubes and more. This material will become increasingly important in the future as data analysis becomes even more pervasive. I want to thank Amr for assembling this issue, which well represents some of the most promising techniques in this important area.