CS848 Paper Submission Site

Go To CS848 Site

Submit or Update A Review For Paper #24

Go To CS848 Submissions

It is currently Monday 18th of October 2004 08:58:30 PM EDT

Your review has been finalized!

Paper # 24 (Download paper of type application/pdf, 164472 bytes)	
Title:	LEO: The DB2 Learning Optimizer
Abstract:	

You have already finalized your review for this paper.
You can no longer modify it, but you may view it.

If you made a mistake in your review and you want it "unfinalized",
you may send mail to the program chair asking them to unfinalize it

Send yourself this review by email

Attribute	Value
Are you finished with this review?	Finalize, I am done editing
Provide a short summary of the paper	This paper provides an overview of LEO, IBM's learning optimizer. The paper begins with a brief summary of some of the assumptions traditional query optimizers make about data stored in tables, and how this affects the estimates of selectivity, etc. The authors then present some related work and how LEO builds on techniques of other papers. Next, an overview of LEO is given. LEO consists of four components. An analyzer, a QEP storage system, a monitoring tool and a feedback exploitation tool. The analyzer makes an analysis offline about data received from the monitoring tool. The monitoring tool keeps track of which statistics were calculated correctly, and which were off. The feedback exploitation tool provides information to the optimizer. The QEP storage system keeps track of plans that were previously executed, and the statistics calculated for them. The four components typically work in sequence to collect data, analyze it and provide feedback directly to the engine. The authors provide an analysis of the differences between their system and those described in literature. The authors next present detailed descriptions of their implementations. Algorithms and exact methods are discussed, along with some comments on statistics generated and how they can be used. The authors present the method by which they adjust statistics during successive runs of the same or different queries. Once discussion of algorithms is complete, the paper presents some critical performance analysis of LEO. LEO consumes less than 5% additional overhead when run on a query. In addition, that 5% overhead can help make a performance gain of approximately 14 times that of the unadjusted query.

	As the paper concludes, the authors discuss augmenting LEO to perform in-flight re-optimization of a query plan using newly calculated statistics, and the learning of additional information such as bandwidth, disk speeds, etc. The authors close with a brief description of the contents of their paper.
What is the strength of the paper? (1-3 sentences)	This paper provides a novel way to allow optimizers to learn from mistakes in calculation. As well, a method of detecting correlations between data columns is presented. This paper is written at a technical level that allows even novice readers to understand it.
What is the weakness of the paper? (1-3 sentences)	The performance analysis of LEO is quite short and not very well described. More tests should have been performed and reported.
Your qualifications to review this paper	I know the material, but am not an expert
Writing Quality	Excellent
Relevance to query processing?	The paper is foundational to query processing
Experimental Methodology	**Poor**
Novelty of paper	This is very novel
Overall paper merit	**The paper is a novel or new contribution with average/weak methodology, or an incremental contribution that has good methodology. Someone in the area should read it**
In your opinion, will this paper be important over time?	Excellent
Provide additional detailed comments to the author	This paper is an excellent example of cross-discipline work. Your intellegent tuning software will surely be modeled for many years to come. Overall, this paper was very well written. Examples were clear and appropriate, and the discussions were at a technical level that could easily be understood by a novice reader. The only concern that I have with your paper is in your performance analysis. I believe either (1) you failed to run enough tests to accurately determine the overhead of your system, or (2) you failed to fully describe the tests that you ran. I believe you only ran a handful of tests, when in fact your numbers may only be reflective of that particular instance of the query. What happens when data is constantly changing, or statistics are consistently slightly out of date but not enough to cause LEO to react? In closing, excellent work. Next time, please provide a deeper performance analysis.
Additional comments to PC (not seen by author)	**This paper is quite novel. A self-tuning optimizer is a fantastic combination of artificial intellegence and database systems research. Although I do not have confidence in the performance analysis, I believe the underlying concept is quite novel. Also, the paper is very well written.**

Goto Main Index Close Window