

1086-68-1072

Dan E Willard* (dew@cs.albany.edu). *On the Linear and Quasi-Linear CPU Time for Most Relational Calculus and Data Mining Queries Using $O(N)$ Space.*

Let Q denote a relational calculus or data mining query that seeks to search an unindexed set of size I to produce an output of size U . This talk will define a subset of relational calculus queries of this type, called RCS, that has the property that every RCS query runs in time $O(U + \text{Log}^d I)$ where the exponent d is a constant that depends on the query (and which usually equals zero or one).

A result of this type was previously announced by us in two stages in year 1996 and 2002 articles in the Journal of Computer and System Sciences. During those times, our results may have seemed mostly theoretical because their $O(U + \text{Log}^d I)$ CPU time and $O(I+U)$ space would have appeared prohibitively expensive (on account of mostly the amount of main memory space used). A theme of our current talk is that these algorithms are now very tempting in the context of the 4G bytes that most modern computers have available.

The current talk will both review our prior results and explore the main open questions that they raise. (These include examining RCS queries from a more parameterized perspective.) (Received September 19, 2012)