ADAPTIVE DATA RESTRUCTURING FUNCTIONS

A High Performance Alternative to Indexed Data Access Structures

An early dream of computer science researchers was to develop formal techniques for organizing and adaptively restructuring data in secondary storage environments. The objective was to develop an algebra of data access and restructuring functions that would allow unrestricted access to information embedded in stored data and at the same time allow that data to be restructured for optimal access and processing by any given application.

Death Of A Dream: As the researchers quickly discovered, the mathematical machinery available at that time, to faithfully model operations and transformations on records and collections of records, did not exist.

This dream was soon dismissed in favor of programming techniques using very sophisticated secondary data structures for accessing the records of primary data structures. This family of access structures is generally referred to as 'index structures'.

Since the early 1980's, indexed data access structures have been the dominate choice for supporting low-level I/O access to secondary storage. It is assumed by most in the industry (generally those unfamiliar with early dreams) that indexed data access is the ideal technology for high performance data access and data restructuring. In principle it is not, though in fact it is - by default.

The Dream Revisited: Recent events have excited the need for Rapid Information Access (RIA). It has become painfully apparent (though not to vendors) that today's database management systems are not capable of meeting the growing need for RIA. Though few in the industry recognize it, indexed data access structures are just not capable of supporting the requirements for RIA.

Fortunately, popular consensus is not a criteria for the technical development of RIA systems. However, what is a criteria for the technical development of RIA systems is the formal mathematical foundation for modeling records and operations on records, in order to formally define adaptive data restructuring functions.

XSP Technology: Classical set theory (CST) is known to be inadequate as a formal foundation, since it does not support an adequate membership definition for modeling records as 'n-tuples'. Extended set theory (XST) provides all the capabilities of CST along with a well-defined definition for 'n-tuples'. Thus XST provides the missing mathematical foundation for realizing the early dream of adaptive data restructuring.

Based on XST, an Extended Set Processing (XSP) technology has been developed that provides all the definitions needed as a calculus for defining an algebra of operations that can be implemented to support adaptive data restructuring.

Further reading:

Childs, D L: Extended Set Theory: A General Model for Very Large, Distributed, Backend Information Systems, Third International Conference On Very Large Databases, Tokyo, Japan, 1977

Childs, D L: A Mathematical Foundation For Systems Development, NATO ASI Series, Vol F24, Database Machines, Edited by A. K. Sood and A. H. Qureshi, Springer-Verlag, 1986

[ADRF: 09/08/02]