



Extended Set Theory: A General Model For Very Large, Distributed, Backend Information Systems

D. L. Childs: Extended Set Theory, (invited paper) Third International Conference on Very Large Data Bases, October 6-8, 1977, Tokyo, Japan.

Abstract

Three distinct components comprise an Information System: INFORMATION MANAGEMENT, DATA MANAGEMENT, and STORAGE MANAGEMENT. Until recently, all three have been subsumed under data management. As applications become more demanding, as support criteria become more complex, and as storage capacity becomes very large, the need for functional independence of these three management areas has become more apparent. Recognition of this situation has been popularized through the phrase, "data independence", or more precisely, "data independence from information" and "data independence from storage".

The difficulty in achieving data independence arises through the incompatibility of a complex information space being supported by a simple storage space. The popular, but limiting approach, has been to force the information space into a restrictive record space. This achieves a deceptive compatibility allowing only the appearance of data independence at the user level. This record oriented approach has become pervasive for small databases even though it constrains user applications, requires substantial storage overhead, and imposes inherent processing inefficiencies.

As databases become very large and as distributed systems become desirable the need for inherent (not superficial) data independence becomes crucial. This paper is intended as a tutorial and will describe conditions for data independence and summarise the concepts of Extended Set Theory as a general model for expressing information systems embodying data independence. This generality will be demonstrated by considering some major problems pertinent to the design and support of very large, distributed, backend information systems.

It should be emphasized that Extended Set Theory is a formalism for expressing solutions and is not a specific solution in itself. Though "redundant membership condition", "distributed membership condition", and "set-theoretic interface" may be new concepts, Extended Set Theory does not preclude any current DBMS concepts, data structures, or existing implementations. Rather, Extended Set Theory embraces them all under a unifying model.

Copyright © 1977 by The Institute of Electrical and Electronic Engineers, Inc. (IEEE). Abstract used with permission.