

**DATABASE AND APPLICATION
DEVELOPMENT IN THE '90s:**

THE CONFERENCE

Chicago, October 1-4, 1991

NDN
Norm De Nardi Enterprises
Presents

**DATABASE AND APPLICATION
DEVELOPMENT
IN THE '90S:**

THE CONFERENCE

Chicago
October 1 - October 4, 1991

(c) Copyright Norm De Nardi Enterprises 1991

All rights reserved. No part of this material may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electric, mechanical, photographic or otherwise, without the written permission of the copyright owners.

It should be noted that the copyright for the material as a whole belongs to Norm De Nardi Enterprises. The copyright for each individual contribution belongs to the author.

**DBMS Technology Directions
and What They Mean
David McGoveran
Alternative Technologies**

OBJECT ORIENTATION: TWO MANIFESTOS

- **THIRD-GENERATION DATA BASE SYSTEM MANIFESTO
(4/9/90)**
 - **2 TENETS AND 13 PROPOSITIONS**

- **THE OBJECT-ORIENTED DATABASE SYSTEM MANIFESTO
(7/5/89)**
 - **13 MANDATORY CHARACTERISTICS**
 - **5 OPTIONAL CHARACTERISTICS**

 - **4 OPEN CHARACTERISTICS (DESIGNERS CHOICE)**

- **OBJECT MANAGEMENT GROUP**

OBJECT ORIENTED TERMINOLOGY

- **ENCAPSULATION**
- **INHERITANCE**
- **POLYMORPHISM**
- **METHODS**
- **MESSAGES**
- **METHODS DICTIONARY**
- **MESSENGER**

OBJECT ORIENTED PRINCIPLES AND FEATURES

- **RELATIONAL ELIMINATED LOOPS**
- **OOP ELIMINATES DECISION TREES?**
- **ORDER**
 - ✚ **Contrary to common opinion, oop takes a neutral approach toward ordering - there is no top to the call structure, no hierarchy, no application-wide functional decomposition.**
- **BUT... this assumes a complete orthogonal basis**

OBJECT ORIENTED PRINCIPLES AND FEATURES

■ ORTHOGONAL BASIS

A minimal set of classes (called the basis) exist such that:

- ☞ every class to be implemented consists solely of properties and services held by the union of this set of classes**
- ☞ each class of the basis shares no properties or services with any other class of the basis**
- ☞ no superclass of any of the basis classes will be created - i.e. the classes to be implemented require only multiple inheritance from classes in the basis.**
- ☞ all objects well defined, stable, & "universal"**
- ☞ every thing an "object"**
- ☞ stable requirements**
- ☞ there are no incompatible views of the world**

OBJECT-ORIENTED DATABASE PRODUCTS

- ONTOS
- VERSANT
- SERVIO
- POSTGRES
- TRELLIS
- etc.

KNOWLEDGE BASES: TERMINOLOGY

- **WHAT IS A KNOWLEDGE BASE?**
- **EXPERT SYSTEMS: INFERENCE ENGINE + KNOWLEDGE BASE + EXPLAINER**
- **DEDUCTION AND INDUCTION**
- **PREDICATES AND HORN CLAUSES**
- **RULES**
- **PROBABILITY TREES**
- **CONFIDENCE FACTORS AND MULTIVALUED LOGIC**
- **FORWARD AND BACKWARD CHAINING**
- **BACKTRACKING**

KNOWLEDGE BASES: FEATURES

- **RULE BASE AND STORAGE METHODS**
- **WORLD OR DOMAIN INFORMATION**
- **INFERENCE ENGINE**
- **INTERFACES: DEVELOPMENT VS. USE**
- **SOME KEY MACHINE LEARNING ALGORITHMS**
 - **ID3: generates decision trees to classify data**
 - **AQ: generates rules more powerful than decision trees**
 - **INDUCE: generates rules for structured descriptions**
 - **CLUSTER: clusters data and discovers structure**
 - **RX and RADIX: statistically deal with time-oriented data**
 - **ENTAIL: generates inexact rules from inexact data**
 - **Holland's Classifier: classifies patterns based on features**

KNOWLEDGE BASES: PRODUCTS AND USES

- **AN EXAMPLE: KBMS**

- **USES**
 - **ONLINE DECISION MANAGEMENT (OLDM)**
 - **AUTOMATED REASONING**
 - **MACHINE LEARNING**
 - **PLANNERS**
 - **TRADING SYSTEMS**
 - **DISTRIBUTION SYSTEMS AND ROUTING**

KNOWLEDGE BASES: POTENTIAL AND LIMITATIONS

- **HIGHLY RESPONSIVE SYSTEMS**
- **REPLICATION OF EXPERT KNOWLEDGE**
- **INTELLIGENT DATABASES**
- **REQUIRES EXPERTISE TO DEVELOP**
- **REQUIRES DOMAIN KNOWLEDGE**
- **GENERALLY FRAGILE SYSTEMS**
- **LONG TRAINING TIME**
- **RELATIONAL COMPATIBILITY: HIGHER LEVEL OF ABSTRACTION**

INTELLIGENT DATABASES

- **WHAT IS AN INTELLIGENT DATABASE?**
- **DEAL WITH INFORMATION RATHER THAN DATA**
- **A SYNTHESIS OF...**
 - **RELATIONAL**
 - **OBJECT ORIENTED**
 - **MULTIMEDIA**
 - **EXPERT SYSTEMS**

INTELLIGENT DATABASES: AN EXAMPLE

- INTELLIGENCE/COMPILER
- AUTO/INTELLIGENCE
- EXPERT/MEASURE
- DATABASE/SUPERVISOR: ERROR DETECTION
- DVT: THE DATABASE VISUALIZATION TOOL
- HYPER/OBJECT
- DATA/INTELLIGENCE - DECISION SUPPORT/EIS
- RAPID APPLICATION ENVIRONMENT

INTELLIGENT DATABASES: AN EXAMPLE II

■ IXL: THE DISCOVERY MACHINE

- ✎ IXL analyzes data and discovers patterns, rules, and unexpected relationships. It generates rules in formats that can be used by other tools. Discovery forms include correlations, clusters, logical relationships, time-oriented relationships, and structural relationships.

■ FIVE MODULES

- User Interface
- Data Dictionary
- Discovery Module
- Induction Engine
- Database Interface

INTELLIGENT DATABASES: USES

- **AUTOMATIC RULE AND PATTERN DISCOVERY**
- **INEXACT AND FUZZY QUERY PROCESSING**
- **AUTOMATIC SQL APPLICATION GENERATION**
- **AUTOMATIC ERROR DETECTION**
- **DATA VISUALIZATION AND SIMULATION**

INTELLIGENT DATABASES: POTENTIAL AND LIMITATIONS

- **A NEW BUZZWORD?**
 - **INGRES (ASK)**
 - **BRITTON LEE IDM**
 - **SYBASE OPEN SERVER**

- **COMPATIBILITY WITH THE RELATIONAL MODEL**

MULTIMEDIA DATABASES: TERMINOLOGY AND FEATURES

- **TEXT**
- **IMAGE**
- **VOICE**
- **BLOB**
- **HYPertext/HYPERMEDIA**

MULTIMEDIA DATABASE: FEATURES

■ FREE FORM TEXT

- LARGE ARBITRARY LENGTH LIMITS
- NO DATA ENTRY SYNTAX OR SEMANTICS
- REQUIRES LOTS OF STORAGE
- HARD TO INDEX AND SEARCH

■ STRUCTURED TEXT

- PREDICTABLE STORAGE REQUIREMENTS
- PERMITS AUTOMATED MANAGEMENT
- LIMITED DATA ENTRY FLEXIBILITY
- USE AND MANAGEMENT COMPLEXITY

MULTIMEDIA DATABASE: FEATURES

■ BITMAP IMAGE

- EXACT IMAGE REPRODUCTION
- ARBITRARY COMPLEX DATA
- ENORMOUS STORAGE REQUIREMENTS
- DATA COMPLETELY UNSTRUCTURED

■ VECTOR IMAGE

- COMPLETE GEOMETRIC MANIPULATION
- RICH HIERARCHICAL DATA STRUCTURES
- UNSUITABLE FOR ARBITRARY IMAGE CAPTURE
- COMPLEX DATA MANAGEMENT

MULTIMEDIA DATABASES: SOME EXISTING PRODUCTS

■ INTEGRATED (DIRECT) SUPPORT

- INTERBASE
- SYBASE
- INGRES
- RDB/VMS

■ "OPEN" SUPPORT

- SYBASE
- INGRES
- INFORMIX

■ DEDICATED PRODUCTS

- TOPIC (VERITY)
- ARCINFO (EIS)

MULTIMEDIA DATABASES: USES

- **TEXT ONLY DOCUMENTS**
- **MEMO FIELDS**
- **EDI**
- **MAIL MERGE FORMS**
- **DIGITIZED PHOTOGRAPHS AND VIDEO**
- **FACSIMILES**
- **CAD DRAWINGS**
- **SCALABLE FONTS**

MULTIMEDIA DATABASES: POTENTIAL AND LIMITATIONS

- **INTEGRATED DATA REPOSITORY**
- **INHERENT NON-UNIFORMITY**
- **POTENTIAL PERFORMANCE PROBLEMS**
- **DIFFICULT ADMINISTRATION PROBLEMS**

MULTIMEDIA DATABASES: RELATIONAL COMPATIBILITY

- **STORAGE: FILES, BLOB, STRUCTURED**
- **ACCESS METHODS (B-, R-, & QUAD-TREES, HASHING)**
- **TRANSACTION MANAGEMENT & RECOVERY**
- **DATA SHARING: LOCKS & VERSION CONTROL**
- **SECURITY: FIELD CONTENT LEVEL OR VIEWS**
- **OPERATORS: RELATIONAL, AGGREGATE, NON-STANDARD, PRECEDENCE**
- **OPTIMIZATION: SELECTIVITY AND JOIN ALGORITHMS**
- **COLLATING SEQUENCES**
- **INTEGRITY ISSUES**

GRAPHICAL USER INTERFACES

- WHAT IS POINT AND CLICK?
- IS GUI OBJECT ORIENTED?
- BLOBs AS ICONS OR DATA
 - PICK LARGE OBJECT AND PLACE?
- EVENTS, COMMANDS, AND MESSAGES
- PROBLEMS ISSUING DDL AND DML
- PROBLEMS REPRESENTING SET PROCESSING
 - LARGE TABLES
 - TRANSACTION MANAGEMENT
 - REMOTE PROCESSING AND DISTRIBUTION

NATURAL LANGUAGE: TERMINOLOGY AND FEATURES

- SEMANTICS VERSUS SYNTAX
- REPRESENTATIONS
- PARSERS
- GENERATORS
- SEMANTIC NETWORK
- MORPHEMES AND TAGMEMES
- CONVERSION FROM WRITTEN ENGLISH TO SQL

NATURAL LANGUAGE: EXAMPLE PRODUCTS: I

EASYTALK (INTELLIGENT BUSINESS SYSTEMS)

- **REPOSITORY**

- **NATURAL LANGUAGE SYSTEM**
 - ☞ **Produces the conceptual meaning representation**

- **DATABASE EXPERT SYSTEM - SEMANTICS**

- **NAVIGATOR/QUERY GENERATOR - produces SQL.**
 - ☞ **To the DBMS system catalog are added semantic column descriptions: entity, time period, unit of measure, level of aggregation, application specific characteristics. Easy Talk uses synonyms to identify database objects.**

NATURAL LANGUAGE: EXAMPLE PRODUCTS: II

DATATALKER - NATURAL LANGUAGE INCORPORATED

- EIGHT MODULES
 - FRONT END
 - PARSER
 - SEMANTIC INTERFACE
 - GENERATOR
 - INTERPRETER
 - DATA BASE INTERFACE
 - CONVERSATION MONITOR
 - DICTIONARY

- THREE CONSTRUCTS
 - CONCEPTS
 - RULES
 - TRANSCRIPTS

NATURAL LANGUAGE USES

- **FACTORY AUTOMATION**
- **ORDER ENTRY**
- **AUTOMATED HELP SYSTEMS**
- **AUTOMATED TELEMARKETING**
- **VOICE MAIL**
- **DATABASE QUERY SYSTEMS**

NATURAL LANGUAGE: POTENTIAL AND LIMITATIONS

- **COMPATIBILITY WITH THE RELATIONAL MODEL**
 - **IMPORTANCE OF ORDERING**
 - **HIERARCHICAL STRUCTURES**
 - **RECURSION**

- **FULL INTERACTIVE DATABASE QUERY IN ENGLISH?**

- **ADVICE:**
 - ☛ **Get a linguistics expert to try to break the system in a sample application.**

NATURAL LANGUAGE: POTENTIAL AND LIMITATIONS

EVALUATING NLI

- **SPEED**
- **SENTENCE COMPLEXITY**
- **MULTIPLE USERS**
- **AMBIGUITY**
- **INSTALLATION COMPLEXITY AND COST**
- **COMPLETENESS OF SQL SELECT GENERATION**
 - ☛ **GROUP BY clauses, joins, subqueries, and aggregates**

NATURAL LANGUAGE: POTENTIAL AND LIMITATIONS

EVALUATING NLI

- UPDATE STATEMENT SUPPORT
- TRANSACTION SUPPORT
- MULTILINGUAL SUPPORT
- EASE OF USE
- HELP FACILITIES
- RESOURCE REQUIREMENTS
- PLATFORMS, NETWORKS, & DBMS SUPPORT

SHARING DATA

METHODS

- BROADCAST UPDATE
- READ OLD WRITE BOTH
- READ NEW WRITE BOTH
- 2PC
- BATCH SYNCHRONIZATION (TO OLD? TO NEW?)
- GATEWAYS (EXTERNAL, OPEN SERVERS)