SEMANTIC DISTRIBUTION

C. ESCULIER.

SEMANTIC DISTRIBUTION

C. ESCULIER

IRIA (Projet SIRIUS) & STERIA

FRANCE

Abstract

To get cooperation in a distributed data management system between different data bases we must deal with "semantic distribution". We call semantic distribution the variety of points of views of a perceived reality. This variety appears in the different models used to formalize this perception and in the names given to entity-types and relations.

The paper emphasizes some problems we meet in the use of the data models on a semantically distributed data base, mainly concerning :

- the duality of fundamental concepts
- the implicit context
- the little part of data semantics supported by any data model.

0. - INTRODUCTION

Distributed data bases emphasize many problems related with the D.B. management.

Besides the numerous technical problems - DDBMS, protocols, interfaces, concurrency, ... - we would like to pay attention to another kind of problem : the semantic distribution.

Until now, any centralized data base concerns a very limited population. We don't mean limited in number (which concerns essentially the technical aspects) but limited on a semantical point of view : all the population is able to cope with few defined schemas.

With a large distributed data base, the population can be much more numerous and a very large variety of points of view will have to be managed.

We call this variety the "semantic distribution". To deal with it, a conceptual schema of the common perceived reality must be established.

To do that :

- we must identify the common perceived reality,
- we must build the common schema.

The two steps involve models :

- to analyze the different data bases schemas
- to formulate the common elements in a common language.

Several models are used as conceptual models of data bases schemas.
We would like to consider the proposed models and their use when we must cope with semantic distribution.

We propose three steps in this approach, concerning :

- the fundamental concepts,
- the meta-concepts,
- the semantic supports.

1. - THE FUNDAMENTAL CONCEPTS

Every proposed model is built on few fundamental concepts - called meta-entities by C.W. BACHMAN, namedcategories by M.E. SENKO - and since several years a "model war" brings an interesting auto-inter-intra-infra-criticism.

With comparisons about generality, naturalness, mathematical aspects, schematic representation, semantic completeness, number of concepts, ... we could guess that everything has been said.

Anyway, when a distributed data base implies the contribution of several data bases elaborated with different models, we must "live" with them.

It means : first to understand them, then to translate data from one model to another one.

We shall consider the first point.

1.1. - Fuzzy concepts

The proposed models are articulated on the notions of entity and/or relation, and/or property (named like that or not).

These notions seem simple and natural.

Every proposed model illustrates it because the fundamental concepts are either defined in a very intuitive way (1) (4) (14) (17) or are not defined at all.

TSICHRITZIS in (24) shows very well the problem :

"An entity is something that has reality and distinctness of being in fact or in thought. Of course the same can be said for properties, values and relationships. What distinguihes entities from others is not something intrinsec but something subjective : interest".

The problem arises in distributed data bases because the interest can be quite different for one partof the population and another part.