

THE INTENSIONAL RELATION

JOHN PLAICE

*Département d'informatique, Université Laval
Ste-Foy (Québec), Canada G1K 7P4
E-mail: John.Plaice@ift.ulaval.ca*

and

JOEY PAQUET

E-mail: jpaquet@ift.ulaval.ca

ABSTRACT

This paper includes an intensional overview of the universal relation in database theory. It is argued that all of the concepts therein can be intensionalized, and in so doing, can address several attacks made against the universal relation.

1. Extended abstract

The universal relation is a concept that is now commonly used to express the semantics of dependencies in relational databases ^{1,2,3,7,6}. According to this idea, an interface is provided by a database to give the impression that all queries and all updates are made on a *single* relation. This *universal* relation is in fact the natural join of the relations that make up the database.

Before the introduction of the universal relation, assumptions about the real world had to be expressed using functional dependencies and multivalued dependencies. The former are fairly easy to discern and enumerate; however, the same does not hold for multivalued dependencies. The universal relation, when appropriate, allows the multivalued dependencies to be automatically inferred from a single join dependency.

Databases built using the universal relation assumption make several basic assumptions, presented below.

1. Each attribute A is expected to play a single role. In particular, puns are not allowed. For example, attribute ADDRESS cannot refer to both SENDING ADDRESS and to RECEIVING ADDRESS.
2. Each combination $[X]$ of attributes is expected to have a unique meaning. For a given set X of attributes, there may well be several relationships, but there should be one that is most basic.
3. All access paths to compute the connection on X represent the same “flavor” of relationship among the attributes in X . This means that if there are several possible paths to compute X , they all have the same basic meaning.

Much of the theoretical work in databases in the early eighties dealt with the problems of building databases with universal relation interfaces. In most of these systems, an actual universal relation is never really built. Rather, there are a number of base relations that participate in a join dependency, which defines the universal relation interface. As a result, a user never need worry about how exactly to navigate through a database.

Notwithstanding certain critiques (see ^{4,5,8} for a fuller discussion), the universal relation is very useful for expressing the semantics of dependency theory. Nevertheless, it should be clear that the three basic assumptions described above cannot be applied in all situations.

Some problems can be resolved in particular situations. Names can be lengthened to ensure that all attributes play a single role. But if a meta-database combining information from many different databases were created, the information therein might be of such a different nature that it would be difficult to create a universal relation.

What is important to note is that each of the three notions can be intensionalized. Suppose that there were a natural notion of possible world in databases. Then the universal relation assumption would be

transformed into the intensional relation assumption as follows: “In each possible world, each of the three assumptions must be satisfied.”

It turns out that the natural concept of possible world already exists: it is the concept of “maximal objects”. These correspond to the largest objects in which one is willing to navigate automatically. These can be constructed automatically from the functional dependencies (essentially they ensure that there are no functional dependency cycles in a maximal object).

In fact, the only navigation that should be explicitly given is navigation between maximal objects. In fact, this was actually suggested in a 1983 paper ⁶. The banking example in that paper uses two maximal objects: in one CUSTOMER means DEPOSITOR and in the other CUSTOMER means BORROWER. It is suggested that aliases be allowed to indicate in which maximal object the object is considered to lie.

The intensional relation is therefore the universal relation in which one can navigate from one maximal object to another. It is not just a technical trick: it should simplify many problems of interoperability between heterogeneous databases. The full paper will present all of these notions formally.

2. References

1. V. Brosda and G. Vossen. Update and retrieval in a relational database through a universal schema interface. *ACM Transactions on Database Systems*, 13(4):449–485, 1988.
2. R. Fagin, A. O. Mendelzon, and J. D. Ullman. A simplified universal relation assumption and its properties. *ACM Transactions on Database Systems*, 7(3):343–360, 1982.
3. P. C. Kanellakis. Elements of relational database theory. In *Handbook of Theoretical Computer Science: Formal Models and Semantics*, pages 1073–1156. Elsevier–North Holland, Amsterdam, 1990.
4. W. Kent. Consequences of assuming a universal relation. *ACM Transactions on Database Systems*, 6(4):539–556, 1981.
5. W. Kent. The universal relation revisited. *ACM Transactions on Database Systems*, 8(4):644–648, 1983.
6. D. Maier and J. D. Ullman. Maximal objects and the semantics of universal relation databases. *ACM Transactions on Database Systems*, 8(1):1–14, 1983.
7. D. Maier, J. D. Ullman, and M. Y Vardi. On the foundations of the universal relation model. *ACM Transactions on Database Systems*, 9(2):283–308, 1984.
8. J. D. Ullman. On kent’s “consequences of assuming a universal relation”. *ACM Transactions on Database Systems*, 8(4):637–643, 1983.