Journal of Mathematics and Statistics Studies
ISSN: 2709-4200
DOI: 10.32996/jmss
Journal Homepage: www.al-kindipublisher.com/index.php/jmss

JMSS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

An Extended Set-Theoretic Relational Model for Multi-Valued Attributes

Sidi Mahmoud KABER[1] ✉ and Mohamedou DEBAGH[2]
[12] *Higher Institute of Accounting and Business Administration, Mauritania*
Corresponding Author: Sidi Mahmoud KABER, **E-mail**: sm.kaber@supnum.mr

| ABSTRACT
The classical relational model, as proposed by Codd, is based on a rigid set structure where each attribute takes an atomic value. However, current data modeling needs, especially in semi-structured systems, require the representation of multi-valued attributes. In this paper, we propose a rigorous mathematical model extending the relational model to allow for multivalued attributes, while maintaining set consistency. We define the formal structure of this model, illustrate its operation through examples, and discuss its advantages and possible integration with modern databases.

| KEYWORDS
Extended relational model, Multi-valued attributes, Set theory, Relational schema, Normalization, Natural join, Database systems.

| ARTICLE INFORMATION
ACCEPTED: 12 April 2025 **PUBLISHED:** 16 May 2025 **DOI:** 10.32996/jmss.2025.6.2.4

1. Introduction
Since its proposal by E. F. Codd in 1970 [1], The relational model formed the theoretical basis of relational database management systems (RDBMS). In this model, a relation is a set of tuples, and each attribute in a tuple is assumed to take an atomic value.

However, in many practical cases, some entities require attributes that can take multiple values. For example, an individual may speak multiple languages, a scientific article may have multiple authors, or a product may belong to multiple categories. The classic relational model handles these situations through separate relationships (normalization), which sometimes complicates queries and semantics.

Here we offer a relational model extended to multi-valued attributes, inspired by set theory, which explicitly allows finite sets as attribute values, while maintaining the foundations of the relational model.

2. Proposed mathematical model
2.1 Basic definitions
Soit D = {D1,D2,...,Dn} all atomic domains.

Simple attribute: an attribute Ais said to be simple if its domain D ∈ D.

Multi-valued attribute : an attribute A is multi-valued if its domain is a power of an atomic domain, i.e. D = P(V), with V ∈ D.

2

Relationship diagram: A relationship schema is a set of attributes noted :

R = {A1 : T1,A2 : T2,...,Ak : Tk}

where each Ti is a single or multi-valued domain.

Generalized tuple: a tuple t is a function t : R → Si Ti such as t(Ai) ∈ Ti for everything Ai.

Generalized relationship: a relation is a finite set of tuples :

$$r \subseteq \prod_{i=1}^{k} T_i$$

2.2 Example : people and languages spoken
Consider the following diagram :

Person(ID : N,Name : Chain,Langues : P(L))

avec

L = {fr,en,ar,es,zh}

A possible instance of the Person relationship :

ID	Name	Langues
1	Alice	{fr, en}
2	Bob	{ar, fr, es}
3	Chao	{zh}

2.3 Extended relational operations
Projection : πNom,Langues(r) = {(t(Nom),t(Langues)) | t ∈ r}

Selection by inclusion :

σfr∈Langues(r) = {t ∈ r | fr ∈ t(Langues)}

Exact equality of set :

σLangues={fr,en}(r)

Joint : Join operations are possible with single or multi-valued attributes, provided that the match condition is redefined.

2.4 Normalized decomposition
A multi-valued attribute can be represented by a separate relationship :

Parle(ID : N,Langue : L)

Relational model extended to multi-valued attributes 3

ID	Langue
1	fr
1	en
2	ar
2	fr
2	es
3	zh

This allows the use of classic relational operators, while facilitating integration with existing systems.

3. Benefits and Discussion

The proposed model :

— Formally extends the relational model without breaking its coherence.

— More accurately reflects certain natural data structures (preferences, languages, keywords).

— Simplifies representations in document-oriented or semi-structured databases.

his model can be implemented directly in NoSQL databases like MongoDB, or integrated into RDBMS via JSON structures or explicit joins.

4. Conclusion

We have presented an extended relational model to integrate multi-valued attributes in a mathematically consistent manner. This model, based on set theory, remains compatible with the foundations of the relational model and meets the needs of modern data representation. Future perspectives include the study of query optimization within this framework, as well as its integration into query languages such as SQL or Datalog.

References

[1] Codd, E. F. (1970) A Relational Model of Data for Large Shared Data Banks, Communications of the ACM, 13, 6, 377–387.
[2] Elmasri R. and Navathe S. B., (2010) Fundamentals of Database Systems, 6th edition, AddisonWesley.
[3] Kimball R. and Ross M., (2008) The Data Warehouse Toolkit, John Wiley & Sons.