HAL
open science

# Handling imperfect spatiotemporal information from the conceptual modeling to database structures

Asma Zoghlami, Cyril de Runz, Herman Akdag, Montaceur Zaghdoud, Henda Ben Ghezala

## To cite this version:

Asma Zoghlami, Cyril de Runz, Herman Akdag, Montaceur Zaghdoud, Henda Ben Ghezala. Handling imperfect spatiotemporal information from the conceptual modeling to database structures. International Symposium on Spatial Data Quality, 2011, Coimbra, Portugal, France. pp.165–170. hal-00632620

**HAL Id: hal-00632620**

**https://hal.science/hal-00632620v1**

Submitted on 14 Oct 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Handling imperfect spatiotemporal information from the conceptual modeling to database structures

*Asma Zoghlami[1,2], Cyril de Runz[1], Herman Akdag[1,3], Montaceur Zaghdoud[2] & Henda Ben Ghezala[2]*

[1] CReSTIC, Université de Reims Champagne-Ardenne
asma.zoghlami@etudiant.univ-reims.fr, cyril.de-runz@univ-reims.fr
[2] RIADI-GDL, ENSI Tunis
montaceur.zaghdoud@ensi.rnu.tn, henda.benghezala@ensi.rnu.tn
[3] LIP6, Université Paris 6
Herman.Akdag@lip6.fr

## Abstract

*Many real world systems and applications require a management tool that provides support for dealing with imperfect data. The aim of this paper is to handle the imperfection of spatiotemporal data from the conceptual modeling to the database conception. We propose to add new pictograms in PERCEPTORY in order to build imperfect spatiotemporal class diagrams such as those made using Fuzzy UML. Using those models, we organize the database as a three layer organization: data layer, metadata layer, multivalued layer. Those interlinked layers give a more accurate interaction.*

**Keywords**: Conceptual model, imperfection, pictograms, database, multi-layer, multivalued set.

## 1 Introduction

As said in (Goodchild, 2006), "the quality of spatial data, as indeed of any data, is crucial to its effective use". As data imperfection is a part of (spatial) data quality, Geographic Information Systems (GIS) has to deal with, for instance, uncertain, imprecise and/or incomplete knowledge.

This paper considers imperfect information modeling in GIS, either on its descriptive, temporal or spatial levels. Imperfect information that characterizes the knowledge is most often manifested by the vagueness but also sometimes by uncertainty or lack of data. Much work about handling imprecise information in Information Systems and in spatial field has been done (Devillers and Jeansoulin, 2006; Jeansoulin *et al.*, 2010). Therefore, a lot of conceptual data models have been extended to model fuzzy data (Ma *et al.*, 2010); fuzzy set theory is the best possible tool that represents data imprecision.

The starting point of our conceptual data modeling is an UML type model which is the class diagram. It is just through this model that different levels of fuzziness were introduced into the concept of class, object and relations between classes.

Thus, the concepts of fuzzy class, fuzzy association, fuzzy aggregation and fuzzy generalization were set in the work of (Ma, 2008).

PERCEPTORY, presented in (Bédard *et al.*, 2004), is a modeling tool that is mainly based on UML and extended with spatial and spatiotemporal stereotypes. These stereotypes have been developed through two distinct PVL (plug-in for visual language): the spatial PVL, for the representation of spatial data and the spatiotemporal PVL that is used to model spatial data and/or temporal data (Brodeur *et al.*, 2000). However, it does not provide a sufficiently explicit way to cope with information imperfection, such as uncertainty and imprecision.

Even though one considers that we can represent imperfect data using the dictionary associated to the model, the automation of the translation into GIS is not easy. This may be possible using an approach that splits data, imperfection and meta-data storage.

This paper aims to highlight the imperfect information in the PERCEPTORY model through the introduction of new visual symbols to manage imperfect spatial, temporal and descriptive data.

Then, we manage these imperfections from the conceptual data model to the database through a multi-layer approach. In this approach, a multi-valued layer is added to the more classical, *data layer* and *meta layer*.

This work is organized as follows. Section 2 is devoted to the insertion of visual variables in PERCEPTORY for the management of data and knowledge imperfection. Then, the structure of the built system is exposed (section 3) with a focus on the explanation of the imperfect management layer so called multivalued layer. The last section (section 5) presents the conclusion.

## 2 Highlighting imperfect knowledge in PERCEPTORY: new visual variables

There are a lot of ways to highlight the imperfection of knowledge in conceptual models. This section introduces the ones we have to deal in the objective of building an agronomical observatory, called Observox, in the Vesle Basin (the project challenges are described in (De Runz and Desjardin, 2009).

### 2.1 Imperfect spatial symbol

Spatial pictograms in PERCEPTORY allow us to define the geometry chosen for spatial elements of a class model. The main geometries are: point, line and polygon. If one can't define with precision the boundaries of a spatial object, there is some imprecision on the geometric shape of this object at the class level. Thus in the PERCEPTORY model, we propose to use the three following basic geometries with dashed outline as shown below in figure 1.

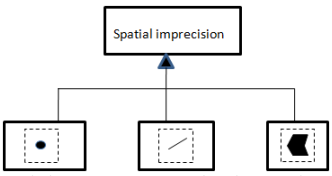


**Figure 1.** Spatial vagueness on the form of spatial objects

### 2.2 Imperfect temporal symbol

The temporal modeling in PERCEPTORY is based on the concepts of existence and evolution. The existence of an object corresponds to its period of life. Objects, having an instant existence, are represented by a pictogram indicating a date while objects that have a sustainable existence are represented by a pictogram indicating a time interval. At this level, there may be some imprecision in the definition of a date or of a time interval. The question is: when was an object present and when did it disappear? The two temporal pictograms are used with dashed outline to express this imprecision.

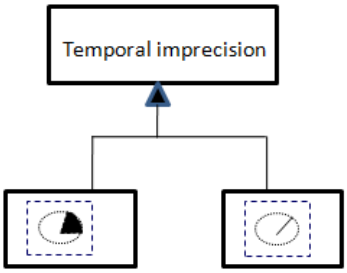


**Figure 2.** Temporal imprecision

### 2.3 Imperfect object attribute symbol

Some attributes in the class model may be defined by fuzzy sets or belief masses. To model this level of imprecision, the keyword **IMP** is introduced and placed in front of the attribute name.

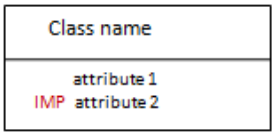


**Figure 3.** Modeling of imperfect attribute

### 2.4 On modeling of class relationship imperfection

The modeling of relationships between classes with their imperfection, we use the UML object diagram in which we associate to the link between the two classes a membership degree.

The relation between two classes (A and B) may also have different value depending of the class instances. In order to model that, we link each instance of A to each instance of B with a membership degree. The figure 4 shows an example of modeling an uncertain relationship between two classes A and B according to class instances.

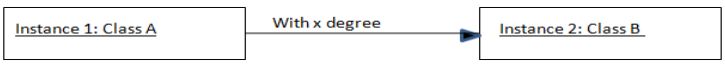


**Figure 4.** Modeling uncertain relationships between two classes through class instances

### 2.5 Membership degree of an object to a class

An object can belong to a class with a membership degree. This is shown in the object diagram through the introduction of the word "with a membership degree" after the instance name.

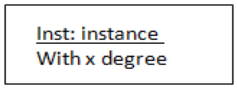


**Figure 5.** An object belonging to a class with a membership degree

## 3 Structuring the database: a multilayer approach

3.1 A multilayer approach

At the implementation level, a first classical layer is implemented. This layer, called *data layer*, contains data in a crisp modeling: the geometric data with the shape and the location of an object, the descriptive data referring to all the descriptive attributes of an object, the temporal data... In our database, the spatial data are represented according to the vector mode in which the objects are represented by points, lines and polygons instead of the raster mode because the vector approach has a lower storage cost.

The *data layer* is followed by a *meta layer* that represents the metadata. Metadata usually concerns the content, data sources, data identification, data quality, spatial representation, spatial reference and any other useful characteristic that may qualify the data. It can also store specific ontology, database schema, *etc.*

A third layer will allow us to link the *data layer* and the *meta layer* to a modeling tool that takes into consideration their imperfection. This layer allow to represent the imprecision and uncertainty through a multivalent approach *(*De Runz *et al., 2010)*.

The principle of the multivalent approach lies in the introduction of several truth values that modulate the information in order to focus on the natural language imperfection. Thus, linguistic expressions such as "very little", "a lot", etc can be used (Akdag et al., 2008). By building a link between data, metadata and imperfection modeling, we try to provide a more accurate view of the processed information by linking together these three layers through putting a link interface between them (figure 6).

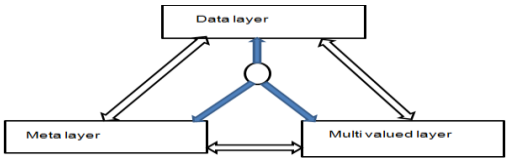


**Figure 6.** Relations between layers

3.2 Multivalued layer

The multi valued layer deals with all the spatial, temporal and descriptive imperfections that may be present on the two more classic layers (*cf.* figure 7).

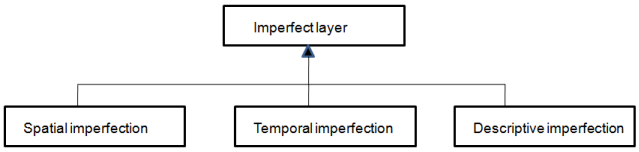


**Figure 7.** Managing imperfection in the multi valued layer

According to (Fisher, 1999; Dubois and Prade, 2009), the modeling of imperfect (spatial) data may be done using a lot of theories (probabilities, possibilities, fuzzy sets, belief functions, etc.). All of those theories use a paradigm of the attribution of weights (between [0;1]) to each element of the studied domain ($\mathbb{R}^2$ or $\mathbb{R}^3$ for space, $\mathbb{R}$ for time, $\mathbb{R}^+$ for quantitative information, *etc.*).

In order to reduce the cost and the complexity of storage and also in order to maintain the possibilities of exploitation, an approach based on the α-cute principle – the domain of values for which the weight is higher or equal to α – is developed. Then, the modeling data have been putting into a multivalued form. This view is then adaptable to the more frequent uncertain representation. It allows users to choose the mode of uncertainty representation for every data. The interoperability between theories should after be done by the systems (it is one of our future goals).

The imperfection layer has an impact on the data layer through dealing with the imperfect relations between objects, imperfect object class relations, imperfect attributes, etc. To deal with possible uncertain relations between objects belonging to different database classes, one must associate a membership degree to the object identifiers in a new database table that indicates at what degree they may have a relation between them.

### 3.3 Links between layers (example)

A geographical entity (De Runz *et al.*, 2010) is composed by a fuzzy spatial area and a set of fuzzy quantities. To handle the vagueness at the database level, the imprecise information is stored in a specific table connected to the geographical entity table through an intermediate table which references the fuzzy quantity values stored in the fuzzy quantity table (see figure 8.).

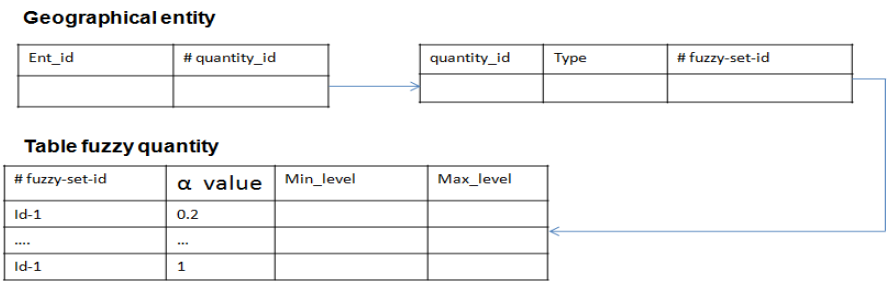


**Figure 8. Illustration of the storage of fuzzy quantities as multivalent set of values.**

## 4 Conclusion

In this paper, we started from the PERCEPTORY conceptual data model to handle spatial, temporal and descriptive data imperfections through the introduction of new visual symbols. At the database level, a multilayer structure is implemented.

Thus, a communication between a multi valued layer, a data layer and a meta layer is established.

In perspective, an application of our approach will be done in the archeological domain and more specifically archeological digs in the city of Reims (ARCHEOCHAMP project).

## Acknowledgments

We would thank the Champagne-Ardenne Region Council, France and European Union (FEDER) for their funding of the AQUAL Project.

## References

Akdag H., Truck I. (2008), "Uncertainty Operators in a Many-valued Logic", *Encyclopedia Of DataWarehousing And Mining*, 2nd edition, Vol 4, Idea Group Publishing, pp. 1997-2003

Bédard Y, Larrivée S., Proulx M.J., Nadeau M. (2004), "Modeling Geospatial Databases with Plug-Ins for Visual Languages: A Pragmatic Approach and the Impacts of 16 Years of Research and Experimentations on Perceptory", *In* S. Wang et al. (eds.) *Proceedings of COMOGIS Workshops ER2004*, LNCS 3289, Springer, pp. 17–30.

Brodeur J., Bédard Y., Proulx MJ. (2000), "Modelling Geospatial Application Database using UML-based Repositories Aligned with International Standards in Geomatics", *In ACMGIS 2000*, Washington DC, USA

De Runz, C., Desjardin, E., Akdag, H. (2010), "Study of spatial fusion of geographical entities and quantitative information in accordance with their imprecision: application to agricultural information in Observox", *In* Tate, N.J., Fisher, P.F. (eds.) *Proceedings of Accuracy 2010,* Leicester, UK, pp. 33-36

De Runz, C., Desjardin, E. (2009), "Issues about qualified spatialization of agricultural practices and data for water contamination study in the Vesle basin", *In Proceedings of Spatial Analysis and GEOmatics*, Paris, France.

Devillers, R., Jeansoulin, R. (2006), *Fundamentals of Spatial Data Quality*, ISTE, London, UK, 312p

Dubois, D., Prade, H. (2009), "Formal representations of uncertainty", *In* Bouyssou, D., Dubois, D., Pirlot, M., Prade, H. (eds.) *Decision-making - Concepts and Methods*, Wiley, 3, pp. 85-156

Fisher, P.F. (1999), "Models of uncertainty in spatial data", *In Geographical Information Systems*, John Wiley & Sons, New-York, pp. 191-205.

Goodchild, M.F. (2006), "Foreword". *In* Devillers, R., Jeansoulin, R. (eds.), *Fundamentals of Spatial Data Quality*. ISTE, London, pp. 13–16.

Jeansoulin, R., Papini, O., Prade, H., Schockaert, S. (2010), *Methods for Handling Imperfect Spatial Information*, Studies in Fuzziness and Soft Computing, Springer, 400p.

Ma, Z. M., Yan, L. (2010), "A Literature Overview of Fuzzy Conceptual Data Modeling", *J. Inf. Sci. Eng*, vol. 26, n° 2, pp. 427-441

Ma, Z. M. (2008), "Fuzzy conceptual information modeling in UML data model", *In Proceedings of the International Symposium on Computer Science and Computational Technology*, vol. 2, IEEE, pp. 331-334