Abstract

In this paper a constraint-based generalised object-oriented database model is adapted to manage spatio-temporal information. This adaptation is based on the definition of a new data type, which is suited to handle both temporal and spatial information. Generalised constraints are used to describe spatio-temporal data, to enforce integrity rules on databases, to specify the formal semantics of a database scheme and to impose selection criteria for information retrieval.

Keywords: Spatio-temporal information modelling, object-oriented database model, constraints.

1 Introduction

A constraint can formally be seen as a relationship, which has to be satisfied. With respect to information and knowledge-based systems, constraints are considered to be an important and adequate means to define the semantics and the integrity of the data [1, 2, 3, 4, 5]. This is especially true for spatial data and for temporal data. An instance then belongs to the information or knowledge base in as far as it satisfies all of its defining constraints.

In practice, spatial data usually consist of line segments, and therefore linear arithmetic constraints are particularly appropriate for representing such data [5]. For example, if spatial geographical information is handled, constraints can be used to define the borders of a country, a city, a region, to define a river, a highway, etc. This is illustrated in Figure 1 and in Table 1, which respectively represent a map of France (a real map will be defined by many more constraints, but the basic ideas are the same) and some geometrical descriptions.

Figure 1: Spatial information: map of France

Constraints can also be used to impose selection criteria for information retrieval. In this case, each constraint defines a condition for the instances to belong to the result of the retrieval [6, 7]. Every instance belongs to the result in as far as it satisfies all the imposed criteria. For example, if someone wants to retrieve all the young persons who live in...
In this paper an extension of a constraint-based fuzzy object-oriented database model is presented. This extension is based on a generalisation of linear arithmetic constraints. In the following section, the main concepts of the fuzzy object-oriented database model are introduced. The modelling of fuzzy spatio-temporal information, by means of generalised linear arithmetic constraints, is discussed in Section 3. Finally, the achieved results and future developments are summarized in the concluding section.

2 Generalised object-oriented database model

The employed fuzzy object-oriented database model [3, 14] has been obtained as a generalisation of a crisp object-oriented database model that is consistent with the ODMG de facto standard [13]. The model is build upon a generalised algebraic type system and a generalised constraint system, which are both used for the definition of so-called generalised object schemes and generalised database schemes.

Table 1: Geometrical descriptions.

<table>
<thead>
<tr>
<th>Area</th>
<th>Geometrical Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annecy:</td>
<td>$(x \geq 10.9) \land (x \leq 11.1) \land (y \geq 6) \land (y \leq 6.2)$</td>
</tr>
<tr>
<td>Seine:</td>
<td>$((y \leq 11) \land (y + 0.7x = 15.2) \land (y \geq 9.6))$</td>
</tr>
<tr>
<td></td>
<td>$\lor((y \leq 9.7) \land (y - 0.2x = 8) \land (y \geq 9.6))$</td>
</tr>
<tr>
<td></td>
<td>$\lor((y \leq 8) \land (y + 1.2x = -20) \land (y \geq 9.7))$</td>
</tr>
<tr>
<td>Haute-Savoie:</td>
<td>$(y - 0.1x \leq 8) \land (y + 7x \leq 89.6)$</td>
</tr>
<tr>
<td></td>
<td>$(y - 0.4x \leq 1.5) \land (y + 0.2x \geq 8.2)$</td>
</tr>
<tr>
<td></td>
<td>$(y - 4x \geq 36.7)$</td>
</tr>
</tbody>
</table>

Annecy, two constraints can be imposed: one that selects all the young persons and another that selects all persons living in Annecy.

Spatio-temporal information can be fuzzy and/or uncertain [8, 9]. There has been considerable amount of research regarding fuzziness in spatio-temporal databases [8, 10, 11, 12]. In this paper an extension of a constraint-based fuzzy object-oriented database model [3] is presented. This extension is based on the introduction of a new data type and on the generalisation of linear arithmetic constraints.

In the following section, the main concepts of the fuzzy object-oriented database model are introduced. The modelling of fuzzy spatio-temporal information, by means of generalised linear arithmetic constraints, is discussed in Section 3. Finally, the achieved results and future developments are summarized in the concluding section.

2.1 Type system

To support the definition of types, a (generalised) type system $GTS$ has been built [3]. In order to be consistent with the ODMG data model, the type system supports the generalised definitions of literal types, object types and reference types (which enable to refer to the instances of object types and are used to formalize the binary relationships between the object types in the database scheme).

The semantic definitions of a type $\tilde{t}$ are based on domains, sets of domains and sets of operators (cf. [15]) and are determined by:

- a set of domains $D_{\tilde{t}}$
- a designated domain $dom_{\tilde{t}} \in D_{\tilde{t}}$
- a set of operators $O_{\tilde{t}}$ and
- a set of axioms $A_{\tilde{t}}$

The designated domain $dom_{\tilde{t}}$ is called the domain of the type $\tilde{t}$ and consists of the set of all the possible values for $t$. Every domain value is represented by a fuzzy set, which is defined over the domain of the corresponding ordinary type $t$. In order to deal with “undefined” values, a type specific bottom value $\bot_t$, has been added to the domain of every ordinary type $t$. The set of operators $O_{\tilde{t}}$ contains all the operators that are defined on the domain $dom_{\tilde{t}}$. The set of domains $D_{\tilde{t}}$ consists of all the domains that are involved in the definition of the operators of $O_{\tilde{t}}$, whereas the set of axioms $A_{\tilde{t}}$ contains all the axioms that are involved in the definition of the semantics of $O_{\tilde{t}}$.

The instances of a literal type, an object type and a reference type are respectively called literals, objects and reference instances. Every instance is characterised by its type and a domain value of this type (also called the state of the instance). Objects either have a transient or a persistent lifetime. Persistent objects are additionally characterised by a unique object identifier and a set of object names.
2.2 Constraint system

Constraints are used to enforce integrity rules on databases (e.g. domain rules, referential integrity rules, etc.) and to specify the formal semantics of the database scheme (e.g. the applicability of null values, the definition of keys, etc.). To support the definition of constraints, a (generalised) constraint system GCS has been built. The set of generalised constraint definitions supported by the constraint system can be partitioned into the subset of constraints which can be applied to objects independent of any existing database (e.g. domain constraints) and the subset of “database” constraints which are defined for database objects (e.g. referential integrity constraints) [3, 4].

The semantics of a constraint are defined by means of a function, which associates with every object a fuzzy set

\[
\{(\text{True}, \mu_{\text{True}}), (\text{False}, \mu_{\text{False}}), \quad (\bot_{\text{Boolean}}, \mu_{\bot_{\text{Boolean}}})\}
\]

which represents the extended possibilistic truth value [16] of the proposition

“the object satisfies the constraint”

The membership degrees \(\mu_{\text{True}}\) and \(\mu_{\text{False}}\) indicate to which degree this proposition is respectively true and false. The membership degree \(\mu_{\bot_{\text{Boolean}}}\) denotes to which degree the proposition is not applicable and is used to model those cases where the constraint \(\hat{\rho}\) is (partially) not applicable to \(\hat{o}\).

2.3 Object schemes

The full semantics of an object are described by an object scheme \(\hat{o}s\). This scheme “in fine” completely defines the object, now including the definitions of the constraints that apply to it. Each object scheme is defined by an identifier \(id\), an object type \(\hat{o}\), a “meaning” \(\hat{M}\) and a conjunctive fuzzy set of constraints \(\hat{C}\), which all have to be applied onto the objects of type \(\hat{o}\) independently of any existing database

\[
\hat{o}s = [id, \hat{o}, \hat{M}, \hat{C}]
\]

The “meaning” \(\hat{M}\) is provided to add comments and is usually described in a natural language. The membership degree of an element of \(\hat{C}\) indicates to which degree the constraint applies to the object type \(\hat{o}\).

An instance \(\hat{o}\) of the object type \(\hat{o}\) is defined to be an instance of the object scheme \(\hat{o}s\), if it satisfies (with a truth value which differs from \(\{(\text{False}, 1)\}\)) all the constraints of \(\hat{C}\) and all the constraints of the sets \(\hat{C}\) of the object schemes, which have been defined for the supertypes \(\hat{o}\) of \(\hat{o}\).

2.4 Database schemes

A database scheme \(\hat{d}s\) describes the full semantics of the objects which are stored in a generalised database and is defined by the quadruple

\[
\hat{d}s = [id, \hat{D}, \hat{M}, \hat{C}_{\hat{D}}]
\]

in which \(id\) is the identifier of the database scheme,

\[
\hat{D} = \{\hat{o}s_i | 1 \leq i \leq n, i, n \in \mathbb{N}_0\}
\]

is a finite set of object schemes, \(\hat{M}\) is provided to add comments, and \(\hat{C}_{\hat{D}}\) is a conjunctive fuzzy set of “database” constraints, which imposes extra conditions on the instances of the object schemes of \(\hat{D}\) (e.g. referential constraints between two object schemes). Again, the membership degrees denote the relevance of the constraints. Every generalized object scheme in \(\hat{D}\) has a different object type. If an object scheme \(\hat{o}s\) in \(\hat{D}\) is defined for an object type \(\hat{t}\) and \(\hat{t}'\) is a supertype of \(\hat{t}\), or \(\hat{t}'\) is an object type for which a binary relationship with \(\hat{t}\) is defined, then an object scheme \(\hat{o}s'\) in \(\hat{D}\) has to be defined for \(\hat{t}'\).

Every persistent instance \(\hat{o}\) of an object scheme \(\hat{o}s\) in \(\hat{D}\) of a database scheme \(\hat{d}s\) has to satisfy all the constraints of \(\hat{C}_{\hat{D}}\), with a truth value which differs from \(\{(\text{False}, 1)\}\).

2.5 Database Model

The generalised database model is finally obtained by extending the formalism with data definition (DDL) and data manipulation operators (DML) [14] (see Figure 2).
3 Modelling of spatio-temporal information

The generalised object-oriented database model presented in the previous section is extended in order to support the modelling of both temporal and spatial information. This is done by adding a new generic literal type SpaceTime to the type system. The domain of this new type consists of all the fuzzy sets which are defined over the points of a given geometrical space, which on its turn is defined by a finite number of axes, which all have only one point in common. Each axis either represents a time dimension or a spatial dimension. Generalised linear arithmetic constraints are defined and are used to describe the domain values of the SpaceTime type.

In the next subsection the focus is on the modelling of one-dimensional temporal information. In this special case, the literal type SpaceTime has to describe a temporal space, hereby using the set \( \mathbb{R} \) of real numbers as a representation of the continuum of physical time points [8]. However, in prospect of the generalisation discussed in Subsections 3.2 and 3.3, the type SpaceTime is introduced.

The type SpaceTime is structured and consists of a finite number of components. Each component either represents a temporal dimension or a spatial dimension. In this subsection only one (temporal) component is considered, so that the specification of SpaceTime is defined as:

\[
\text{SpaceTime \( id(id_1 : TimeDim) \)}
\]

where \( id \) is the identifier of the type and \( id_1 \) is the identifier of the component.

The domain of type

\[
\text{SpaceTime \( id(id : TimeDim) \)}
\]

(soon written as \( \text{dom}_{\text{id}} \)) is defined by

\[
\text{dom}_{\text{id}} = \hat{\gamma}(\{(x)|x \in \text{dom}_{\text{TimeDim}}\}) \cup \{\bot_{\text{SpaceTime}}\}
\]

where \( \hat{\gamma}(U) \) denotes the set of all fuzzy sets, which can be defined over the universe \( U \) and

The domain of TimeDim is defined by

\[
dom_{\text{TimeDim}} = \mathbb{R} \cup \{\bot_{\text{TimeDim}}\}
\]

where \( \mathbb{R} \) denotes the set of real numbers and \( \bot_{\text{TimeDim}} \) represents an “undefined” domain value.

The considered operators are the binary operators \( =, \neq, <, >, \leq, \geq, +, -, * \) and / and a null-ary operator \( \bot \), which always results in an “undefined” domain value. When restricted to the set \( \text{dom}_{\text{TimeDim}} \setminus \{\bot_{\text{TimeDim}}\} \), all binary operators have the same semantics as their counterparts within \( \mathbb{R}^2 \). For the bottom value \( \bot_{\text{TimeDim}} \), the semantics are:

\[
\forall x \in \text{dom}_{\text{TimeDim}} : \text{op}(x, \bot_{\text{TimeDim}}) = \text{op}(\bot_{\text{TimeDim}}, x) = \bot_{\text{TimeDim}}
\]

where “op” is a variable copula whose successive values are respectively the symbols \( =, \neq, <, >, \leq, \geq, +, -, * \) and /.
\( \downarrow_{SpaceTime} \) represents an “undefined” domain value.

With the previous definition, every regular value of \( \text{dom}_{\text{id}} \) is a fuzzy set, which is defined over the continuum of physical time points. In order to describe the values of \( \text{dom}_{\text{id}} \), linear arithmetic constraints are generalised. This is done by generalising the comparison operators \( =, \leq \) and \( \geq \).

Traditionally, these operators allow to describe crisp subsets of \( \text{dom}_{\text{id}} \), e.g. \( \forall \ t \in \text{dom}_{\text{id}}, \ x = t \) is a description of the fuzzy set \( \{(t, 1)\} \), \( x \leq t \) represents the set \( \{(x, 1) | x \leq t\} \) and \( x \geq t \) describes the set \( \{(x, 1) | x \geq t\} \).

For the generalisation, a normalised fuzzy set \( \hat{V} \) has been associated with each operator. This fuzzy set is defined over the universe of valid distances — the set \( \mathbb{R}^+ \) of positive real numbers — and the boundary condition \( \mu_{\hat{V}}(0) = 1 \) must hold for it.

If \( d(x, x') \) denotes the Euclidean distance between the defined elements \( x \) and \( x' \) of \( \text{dom}_{\text{TimeDim}} \) — i.e. \( d(x, x') = |x - x'| \) — then the membership functions of the fuzzy sets described by the generalised operators \( =_{\hat{V}}, \leq_{\hat{V}} \) and \( \geq_{\hat{V}} \) are defined as follows:

\[
\forall \ x, t \in \text{dom}_{\text{TimeDim}} \setminus \{\downarrow_{\text{TimeDim}}\}:
\]

- \( \mu_{x =_{\hat{V}} t}(x) = \mu_{\hat{V}}(d') \), with
  \[
d' = \min\{d(x, x') | x' \in \text{dom}_{\text{TimeDim}} \land x' = t\}
\]

- \( \mu_{x \leq_{\hat{V}} t}(x) = \mu_{\hat{V}}(d') \), with
  \[
d' = \min\{d(x, x') | x' \in \text{dom}_{\text{TimeDim}} \land x' \leq t\}
\]

- \( \mu_{x \geq_{\hat{V}} t}(x) = \mu_{\hat{V}}(d') \), with
  \[
d' = \min\{d(x, x') | x' \in \text{dom}_{\text{TimeDim}} \land x' \geq t\}
\]

Figure 3 illustrates the membership functions that result from the application of the generalised comparison operators \( \leq_{\hat{V}}, \geq_{\hat{V}} \) and \( =_{\hat{V}} \) to a given fuzzy set \( \hat{V} \).

![Figure 3: Application of the generalised comparison operators](image)

Linear arithmetic constraints have been generalised by replacing all regular comparison operators by (adequate) generalised comparison operators and by replacing the regular logical operators \( \land, \lor \) and \( \neg \) by their fuzzy counterparts \( \land_{\hat{V}}, \lor_{\hat{V}} \) and \( \neg_{\hat{V}} \), which semantics have been defined as follows:

- the impact of the \( \land \) operator is reflected by applying Zadeh’s (standard) intersection operator [17] onto the fuzzy sets described by the arguments of the operator, i.e. with arguments \( U \) and \( V \), the membership degree of \( (x) \), \( x \in \text{dom}_{\text{TimeDim}} \) in the resulting fuzzy set equals
  \[
  \min(\mu_{\hat{U}}(x), \mu_{\hat{V}}(x))
  \]

- the impact of the \( \lor \) operator is reflected by applying Zadeh’s (standard) union operator [17], i.e. with arguments \( U \) and \( V \), the membership degree of \( (x) \), \( x \in \text{dom}_{\text{TimeDim}} \) in the resulting fuzzy set equals
  \[
  \max(\mu_{\hat{U}}(x), \mu_{\hat{V}}(x))
  \]

- the impact of the \( \neg \) operator is reflected by applying Zadeh’s (standard) complement operator [17], i.e. with argument \( U \), the membership degree of \( (x) \), \( x \in \text{dom}_{\text{TimeDim}} \) in the resulting fuzzy set equals
  \[
  1 - \mu_{\hat{U}}(x)
  \]
For example, with appropriate fuzzy sets $\tilde{U}$, $\tilde{V}$ and $\tilde{W}$, the fuzzy temporal information “around” time point 60 and from time point about 100 until time point about 120” can be described as:

$$(x = \tilde{U} 60) \tilde{V} ((x \geq \tilde{V} 100) \tilde{W} (x \leq \tilde{W} 120))$$

### 3.2 Modelling of spatial information

The data type SpaceTime can be adapted to model spatial information. A distinction is made between one-dimensional, two-dimensional and n-dimensional data.

In order to model spatial information a new data type SpaceDim is defined. The definition of this type is similar to the definition of the type TimeDim, introduced in the previous section: the domain of SpaceDim is defined by

$$\text{dom}_{\text{SpaceDim}} = \mathbb{R} \cup \{ \perp_{\text{SpaceDim}} \}$$

where $\perp_{\text{SpaceDim}}$ represents an “undefined” domain value; furthermore, the same operators $=, \neq, <, >, \leq, \geq, +, -, \ast, /$ and $\perp$ have been defined.

#### 3.2.1 One-dimensional spatial data

The type SpaceTime is also suited for the modelling of spatial data. One-dimensional spatial data can be handled by considering one (spatial) component. The specification of SpaceTime then becomes:

$$\text{SpaceTime id(id}_1 : \text{SpaceDim})$$

where id remains the identifier of the type and $id_1$ remains the identifier of the component. In the one-dimensional case, the modelling of spatial information is then completely analogous to the temporal case discussed in the previous subsection.

#### 3.2.2 Two-dimensional spatial data

Two-dimensional spatial data can be modelled by considering two (spatial) components for the type SpaceTime, i.e. by considering the specification:

$$\text{SpaceTime id(id}_1 : \text{SpaceDim}, id_2 : \text{SpaceDim})$$

In this case, the domain $\text{dom}_{id}$ is defined by

$$
\text{dom}_{id} = \tilde{\nu}(\{(x, y)|x, y \in \text{dom}_{\text{SpaceDim}}\})
\cup \{ \perp_{\text{SpaceDim}} \}
$$

With this definition, each regular value of $\text{dom}_{id}$ is a fuzzy set, which is defined over the continuum of points in the plane defined by the two spatial axes with identifiers $id_1$ and $id_2$.

The generalisation of the comparison operators $=, \neq$ and $\geq$ is obtained analogously as in the one-dimensional case. A normalised fuzzy set $\tilde{V}$, which is defined over the universe of valid distances and for which the boundary condition $\mu_V(0) = 1$ holds, is associated with each operator.

If $d((x, y), (x', y'))$ denotes the Euclidean distance between the defined elements $(x, y)$ and $(x', y')$ of $\text{dom}_{\text{SpaceDim}} \times \text{dom}_{\text{SpaceDim}}$—i.e. $d((x, y), (x', y')) = \sqrt{(x - x')^2 + (y - y')^2}$—then the membership functions of the fuzzy sets described by the generalised operators $=\tilde{V}, \leq\tilde{V}$ and $\geq\tilde{V}$ are defined as follows:

$$
\forall (x, y) \in (\text{dom}_{\text{SpaceDim}} \setminus \{ \perp_{\text{SpaceDim}} \}), \forall m, l \in \mathbb{R}:
$$

- $\mu_{x+my=\tilde{V}l}(x, y)) = \mu_\tilde{V}(d')$, with
  
  $$
d' = \min\{d((x, y), (x', y'))|(x', y') \in (\text{dom}_{\text{SpaceDim}})^2 \land x' + my' = l\}
$$

- $\mu_{x+my\leq\tilde{V}l}(x, y)) = \mu_\tilde{V}(d')$, with
  
  $$
d' = \min\{d((x, y), (x', y'))|(x', y') \in (\text{dom}_{\text{SpaceDim}})^2 \land x' + my' \leq l\}
$$

- $\mu_{x+my\geq\tilde{V}l}(x, y)) = \mu_\tilde{V}(d')$, with
  
  $$
d' = \min\{d((x, y), (x', y'))|(x', y') \in (\text{dom}_{\text{SpaceDim}})^2 \land x' + my' \geq l\}
$$

For example, with the fuzzy set $\tilde{V}$ of Figure 3 and the fuzzy set $W = \{(0, 1)\}$, “the environment of Annecy” can be modelled by

$$(x = \tilde{V} 11) \tilde{V} (y = \tilde{V} 6.1)$$
and “the neighbourhood of the Lake of Geneva in Haute-Savoie” can be modelled by

\[(y \geq \bar{V} 6.7) \land (y-x \geq \bar{V} -3.7) \land (y \leq \bar{W} 6.7) \land (y-x \geq \bar{W} -3.7) \land (x \leq \bar{W} 11.9) \land (x \geq \bar{V} 11)\]

Both examples are illustrated in Figure 4 (drawn to scale).

Figure 4: Illustration of the modelling of two-dimensional spatial data

3.2.3 \(n\)-dimensional spatial data

In order to model \(n\)-dimensional spatial data, the type \(SpaceTime\) can be constructed with \(n\) spatial components. In this case, the domain \(dom_{id}\) is defined by

\[dom_{id} = \tilde{\gamma}((x_1, x_2, \ldots, x_n)|x_i \in dom_{SpaceDim}\}
\]

The comparison operators =, \(\leq\) and \(\geq\) can be generalised straightforwardly and analogously to previous cases by considering the Euclidean distance

\[d((x_1, x_2, \ldots, x_n), (x'_1, x'_2, \ldots, x'_n)) = \sqrt{(x_1 - x'_1)^2 + (x_2 - x'_2)^2 + \cdots + (x_n - x'_n)^2}\]

and an associated, normalised fuzzy set \(\tilde{V}\), which is defined over the universe of valid distances and for which the boundary condition \(\mu_{\tilde{V}}(0) = 1\) holds.

3.3 Literal type \(SpaceTime\)

In general the literal type \(SpaceTime\) can have both spatial and temporal components. This allows to model spatio-temporal information in its most general form. The type specification then becomes:

\[SpaceTime id(id_1 : t_1, id_2 : t_2, \ldots, id_n : t_n)\]

where \(id\) remains the identifier of the type, \(id_i, i = 1, 2, \ldots, n\) remain the identifiers of the axes represented by the components and \(t_i \in \{TimeDim, SpaceDim\}, i = 1, 2, \ldots, n\) denote the nature of the axes.

The domain of \(SpaceTime\) is defined by

\[dom_{id} = \tilde{\gamma}((x_1, x_2, \ldots, x_n)|x_i \in dom_{t_i}, i = 1, 2, \ldots, n) \cup \{\bot_{SpaceTime}\}\]

where \(\bot_{SpaceTime}\) represents an “undefined” domain value.

Because by definition all the elements of the domain are fuzzy sets, operators have been provided for the handling of fuzzy sets. Among the considered operators are: \(\cup, \cap, co, normalise, support, core, a-cut\) and \(\alpha-cut\). Each operator preserves its usual semantics. Additionally, a null-ary operator \(\bot\), which always results in an “undefined” domain value, is added.

4 Conclusion

A new approach for the handling of spatio-temporal information is presented. The rationale behind this approach is the assumption that linear arithmetic constraints are particularly appropriate for representing such information.

The approach is presented as an extension of a constraint-based fuzzy object-oriented database model, but its application is definitely not restricted to database models. Central to the approach is the introduction of a new generic type \(SpaceTime\), which is suited to handle fuzzy multi-dimensional temporal and/or spatial information. The description of domain values of \(SpaceTime\) by means of so-called generalised linear arithmetic constraints, which have been obtained by generalising the definition of the comparison operators =, \(\leq\) and \(\geq\), is typical. Future work includes the definition of appropriate data definition and data manipulation operators.
References


