Formal classification of integrity constraints in spatiotemporal database applications

Mehrdad Salehi a, b, *, Yvan Bédard a, b, Mir Abolfazl Mostafavi b, Jean Brodeur b, c

a Canada NSERC Industrial Research Chair in Geospatial Databases for Decision-Support, Canada
b Laval University, Department of Geomatics Sciences, Quebec City, Quebec, Canada
c Center for Topographic Information, Natural Resources Canada, Sherbrooke, Quebec, Canada

ARTICLE INFO

Article history:
Received 5 May 2009
Received in revised form
1 August 2010
Accepted 9 January 2011
Available online 15 January 2011

Keywords:
Conceptual model
Integrity constraint
Spatiotemporal modeling
Data quality
Spatiotemporal database

ABSTRACT

Imposing integrity constraints is an efficient way to improve data quality in databases. Effective imposition of integrity constraints requires their precise distinction and specification. Despite a few efforts for enhancing the distinction and specification of the integrity constraints in spatial and spatiotemporal databases by their classifications, these classifications fail to precisely distinguish between inherently dissimilar integrity constraints. Furthermore, the existing classifications provide imprecise definitions for the classes of integrity constraints. Such shortcomings explain why still diverse terms are used to refer to a same spatial integrity constraint. In this paper, we propose a formal and more exhaustive classification of the integrity constraints in spatiotemporal databases relying on their nature with respect to space, time, and themes. Moreover, a terminology for the integrity constraints of spatiotemporal databases is presented. Finally, we discuss the advantages of the proposed classification in the specification of integrity constraints.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Diverse emerging geographical information system (GIS) applications, for example in traffic analysis, environmental modeling, and disaster management, rely upon data describing the evolving location and shape of features, i.e., spatiotemporal data. Organizations invest millions of dollars for collecting and analyzing spatiotemporal data for their applications; however, poor data quality significantly degrades their investments. Spatiotemporal data quality is a complex topic involving several issues that refer either to the internal quality (intrinsic, respect of specifications) or to the external quality (fitness for use). ISO/TC211 [1] suggests the following elements for the internal quality: completeness, positional accuracy, temporal accuracy, thematic accuracy, and logical consistency. Integrity constraints (ICs) are defined to improve the logical consistency of databases [2].

ICs are assertions aimed at preventing the insertion of incorrect data into a database. In the spatiotemporal database domain, an IC defines mandatory, allowed, and unacceptable spatial and/or temporal relationships and values, sometimes with respect to other attribute values, other relationships, and temporal durations or geometric shapes. Examples of these ICs include allowed spatiotemporal shapes, such as “a moving-point or a moving-polygon for a moving-object, depending on its size”, spatiotemporal property such as “maximum speed of a vehicle”, and spatial and temporal relationships between features, such

---

*Correspondence to: 188 Doris Ave., Suite 1517, Toronto, ON, Canada M2N 6Z5.
E-mail addresses: mehrdad.salehi.1@ulaval.ca (M. Salehi), yvan.bedard@scg.ulaval.ca (Y. Bédard), mir-abolfazl.mostafavi@scg.ulaval.ca (M.A. Mostafavi), brodeur@RNCan.gc.ca (J. Brodeur).

1045-926X/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jvlc.2011.01.001
as “two vehicles cannot be inside a same parking spot at the same time”.

The above examples illustrate that ICs convey important semantics of the application domain. Consequently, it is necessary to specify and validate the definition of ICs at the design stage of a spatiotemporal database application. The specification of ICs is typically supported by a technology-independent model of the data called the conceptual model. A conceptual model gets rid of all technical details and focuses on semantics and business rules. This is similar to computation-independent model (CIM) proposed by the model-driven architecture [3].

Although ICs are originally introduced for improving data quality in databases, they also find applications in other fields such as database integration and interoperability, query optimization, knowledge representation, and reasoning [4]. As a result, any attempt to clarify various types of ICs is significantly useful for these disciplines.

In this respect, a few classifications of ICs in spatial and spatiotemporal databases have been proposed (e.g., [5,6]). Although these classifications provide some guidelines to distinguish a number of ICs, they suffer from two shortcomings: (1) they are not precise, i.e., the definitions of IC classes are given using imprecise natural languages such as English; and (2) they are not exhaustive enough to provide a guideline for the specification of ICs in spatiotemporal databases. Imprecise classifications may lead to ambiguity as the definition of a given IC class may be unclearly described and interpreted in different ways. For example, as we will see in Section 2, an IC like “a parcel has a polygonal geometry” falls into all the three classes of spatial ICs proposed by [5]. Moreover, the existing classifications of spatial and spatiotemporal ICs are not comprehensive enough and discriminate only among the simplest ICs. In these classifications, ICs whose specification should be done: in the data dictionary rather than the conceptual schema, by different persons with dissimilar expertise, with different categories of languages, etc., are not distinguished. Such shortcomings of the existing classifications of integrity constraints in spatial and spatiotemporal databases could explain why diverse terms are used in order to refer to even a simple IC. For example, the IC “a road does not intersect a lake” is referred to as a semantic IC [5], spatial IC [7], topo-semantic IC [8], and topological IC [9].

Therefore, the existing classifications of ICs in spatial and spatiotemporal databases are ambiguously defined using natural languages (e.g., English) and hence do not establish a satisfactory foundation for the distinction and specification of ICs. Consequently, a precise classification for these ICs is required. Two major contributions of the present paper are as follows:

1. Based on the very nature of concepts that may appear in the assertion of an IC with respect to space, time, and themes, taken alone or in combination (Section 3), we propose a formal, precise, and more comprehensive classification of ICs than the existing classifications for spatiotemporal databases at the conceptual level and a terminology for the IC classes (Section 4). This formal classification, which does not have the ambiguity of classifications defined using natural languages, provides precise definitions for the different classes of ICs.
2. By providing an example for the application of the proposed ICs classification (Section 5), we list some natural questions that spatiotemporal database designers may ask for the specification of ICs. Following each question, we discuss how further synthesis of the specification of different IC classes provides an insightful guideline for answering to these questions (Section 6).

2. State-of-the-art on classifications of ICs in spatial and spatiotemporal databases

The majority of related works on spatial ICs (e.g., [10–13]) refer to the conceptual level taxonomy of spatial ICs proposed by Cockcroft [5]. Cockcroft classifies spatial ICs into three categories: topological, semantic, and user-defined. This classification covers a number of spatial ICs. However, some limitations appear when one tries to apply it to real spatial database projects. Based on the examples given by the author, it seems that topological ICs refer to the ICs dealing with geometric properties and relationships. However, geometric properties and relationships are not limited to only “topological” ones, but include “metric” and “ordering” relationships [14]. Therefore, the label “topological IC” for the category of ICs including all geometric properties and relationships does not describe this category properly. Semantic ICs differ from topological ICs since they are concerned with the meaning of geographical features, such as “a road cannot run through a body of water”. Finally, user-defined ICs are defined as the business rules in spatial databases, such as “the distance between a school and a gas station shall be more than 300 meters”. The author calls these user-defined ICs “esoteric” and does not define them clearly.

The imprecise definitions of the two latter categories of ICs result in ambiguities and lead to an unclear distinction between a semantic IC and a user-defined IC, unless one assumes that there is a neutral, totally objective reality with meanings naturally built into the objects (a philosophical issue that has been debated without a definite answer for centuries). In fact, our point of view rather is that semantics is specified by the user (a philosophical position where semantics cannot exist without a human interpreting reality). Thus, a semantic IC is a user-defined IC since it depends on the ontology of the application. Another example of confusion in Cockcroft’s classification is found when [11] associate the IC “a parcel has polygonal geometry” to the category of topological ICs. Although this IC specifies the geometry, it also refers to “parcel”; hence, it can be interpreted as a semantic IC. Moreover, it depends on the user’s need to use a polygon or a point to represent a parcel. Consequently, this IC can be categorized as a user-defined IC as well. Due to such ambiguities, it appears that this classification does not provide satisfactory foundations to clearly distinguish ICs for spatial databases. Thus, from our point of view and philosophical perspective, Cockcroft’s classification needs to be replaced by a more rigorous solution.
We have identified two research works that directly focus on the classification of ICs at the conceptual level of spatiotemporal databases. Cockcroft [5] integrates static and transition ICs with her categories of spatial ICs. Transition ICs restrict possible transitions from one database state to another. Therefore, three classes of topological, semantic, and user-defined spatial ICs are further categorized into static or transition ICs. However, as discussed earlier, since there is no clear distinction between Cockcroft’s semantic IC and user-defined IC, the combination of these ICs with state and transition ICs gives confusing results. The second classification of ICs in spatiotemporal databases is proposed by Brisaboa et al. [6]. They argue that ICs can be defined for objects and relationships. ICs that are defined for relationships are further categorized into ICs on explicitly-stored and non-stored relationships. The missing point in this approach, however, is that no further consideration is given to the major distinctive constituents of spatiotemporal databases, i.e., space, and time, for the classification of ICs.

In addition, the above two classifications do not distinguish between spatiotemporal ICs that need different specification strategies. For example, consider IC1: “for any moving-point, there must exist at least an instant such that the speed of the moving-point is greater than zero” and IC2: “the distance between a school and a gas station must be greater than 300 meters in Canada after 1970”. While both these ICs impose restrictions in space and time, they have a number of remarkable differences. IC1 does not involve thematic information and is typically defined for the software built-in spatiotemporal data structure (i.e., property “speed” of a “moving-point”) in a spatiotemporal database engine. Hence, it has a direct impact on instances of all moving-object classes, such as vehicle, train, and airplane classes. In addition, as IC1 is concerned with the data structure, it should be defined by and with a language that is understandable and useful (easily translatable to code) for the person in charge of designing that data structure, i.e., the spatiotemporal database designer. On the other hand, IC2 involves thematic concepts (i.e., school and gas station), and is concerned with a specific application built on the top of a spatiotemporal database. Hence, it should be specified with a language understandable for this application’s domain specialist who, unlike a database designer, does not necessarily have a software engineering background. Such a language should be close to the way a non-specialist expresses ICs, such as a subset of a natural language like English. IC2 constraints instances of only two, but not all, classes in a spatiotemporal database (i.e., school class and gas station class). The specification of IC2 can be done by languages including spatial and temporal constructs (e.g., distance, after). Such languages are not required to have spatiotemporal constructs as is the case for defining IC1. The comparison of these two ICs shows that from a general perspective both of them similarly convey spatial and temporal restrictions. Further analyzing these ICs reveals that their impact on spatiotemporal data quality, the category of people involved in their specification and validation, and the type of languages for their specification are dissimilar. Therefore, these ICs are inherently of two different types and should be associated to two different IC classes. The existing classifications of ICs in spatiotemporal databases, however, do not make a distinction between these two ICs.

In summary, there is no precise and exhaustive enough classification for the ICs in spatiotemporal databases that builds a foundation for the distinction and specification of these ICs. In the remainder of this paper, we propose a formal nature-based classification of ICs in spatial and spatiotemporal databases. This classification relies on the type of concepts that may appear in the assertion of ICs regarding to three well-known concepts in GIS, i.e., space, time, and themes, and all their possible combinations.

3. The nature of concepts in the assertion of ICs

The three distinctive constituents of spatiotemporal databases are space, time, and themes [15]. Typically, in spatiotemporal databases, objects are classified based on their spatial, temporal, and thematic (i.e., non-spatial and non-temporal) characteristics, and on the combinations of these characteristics, resulting in spatial objects, temporal objects, spatiotemporal object, thematic objects, and so on [16].

Inspired by the above classification, we consider space, time, and themes for the classification of ICs in spatiotemporal databases. The principal idea is that we consider an IC as an assertion carrying a number of concepts, which we refer to as “IC concepts”. IC concepts are related to space (“shape and where”), time (“duration and when”), an application semantic themes (the “what”), and the way they are combined. We propose to classify ICs based on the nature of the IC concepts that appear in an IC’s assertion, and consequently, based on the restrictions that the IC enforces on a phenomenon and its frames of reference in space and time. We should note that, while an “IC” is an assertion per se, an “IC concept” refers to the concepts that appear in the assertion of an IC. To prevent terminological confusion, in the remainder of this paper we typically use the term “concept” to refer to IC concept. For instance, when we use the term “spatial concept”, we mean spatial concepts that appear in the definition of an IC.

In order to capture all IC concepts and their possible combinations, we consider the 3-Venn diagram [17] in Fig. 1. The three circles include “concepts referring to space” or spatial concepts (e.g., polygon, area), “concepts referring to time” or temporal concepts (e.g., after, instant), and “concepts referring to themes” or thematic concepts (e.g., house, use). The possible intersections of these 1-sets of concepts result in four mixed sets of concepts. Among these four sets, there are three 2-sets (i.e., concepts referring to space and themes, concepts referring to time and themes, and concepts referring to space and time) and one 3-set (i.e., concepts referring to space, time, and themes).

An example of a concept referring to space and themes is “a polygonal house”, and an example of a concept referring to time and themes is “a building’s construction date”. It is obvious that these two concepts can be easily expressed by spatial concepts such as “polygon” or temporal concepts such as “date”, in addition to thematic concepts such as “house”, “building”, and “construction”.


discussion earlier, since there is no clear distin
spatiotemporal concepts. A 3-Venn diagram of the concepts relating to space, time, themes, and their intersections which may be included in an IC’s assertion.

However, the concepts referring to both space and time are of two types, since they can be expressed in two different ways. Some, like “distance after 1970”, can easily be described by spatial concepts (e.g., distance) plus temporal concepts (e.g., after, 1970). On the other hand, there are concepts that refer to both space and time but are inherently cumbersome or counter-intuitive to describe by spatial concepts and temporal concepts separately, so it is more efficient to describe them by spatiotemporal concepts such as “moving-point” and “speed”. The values given to spatiotemporal properties are dependent upon both frames of reference, i.e., space and time, and they are most efficiently expressed using a single spatiotemporal concept like “speed”. Concepts that refer to all three types, space, time, and themes such as “the speed of a vehicle” or “the area of a lake during the summer”, can easily be expressed with a combination of spatiotemporal concepts (e.g., speed and thematic concepts (e.g., vehicle), or of spatial concepts (e.g., area), temporal concepts (e.g., during, summer), and thematic concepts (e.g., lake).

Consequently, the four categories of spatial concepts, temporal concepts, thematic concepts, and spatiotemporal concepts, can intuitively generate all concepts in Fig. 1. We will describe these four categories of concepts in Sections 3.1–3.4.

3.1. Spatial concepts within the assertion of ICs

In order to locate objects with respect to a framework, space is described by a datum that includes a 2D or 3D reference surface and a coordinate system located on the reference surface using an origin, directions for the axes, and a scale factor. For spatial conceptual modeling within this framework, based on a discrete object-based view of space, objects are typically represented by a number of geometrical primitives like point, line, and polygon. These primitives are the constructs for expressing more complex geometries, such as a set of lines representing a poly-line. Based on this representation of space, we can consider a number of concepts that can appear in the assertions of ICs.

We identify and categorize the concepts relevant to space that may appear in the definition of ICs as the following: “spatial primitive”, “spatial property”, “spatial relationship”, “spatial operation”, and “spatial reference system”. We now describe each category of these spatial concepts, as follows:

A spatial primitive was described earlier in this section. A spatial property is an inherent characteristic of a spatial primitive when taken as a whole, and it can be further classified as a metric property (e.g., length of a line) or a topological property (e.g., closedness of a polygon). A spatial relationship deals with the characteristics of the relations between spatial primitives. Egenhofer [14] distinguishes three principal types of spatial relationships: topological relationships, metric relationships, and ordering relationships. Topological properties and topological relationships are preserved under continuous and elastic transformations of space, such as translation, rotation, and scaling [18,19]. Metric properties and metric relationships rely on measurements and employ measuring units of a spatial reference system. Accordingly, we put the relationships known as directional relationships, such as “South” [20], in the category of metric relationships since they are tied to measurements of direction in a spatial reference system. For instance, the relationship “point B is South of point A” indicates that the direction of the axis from A to B must be measured between bearings 90° and 270° of a spatial reference system. An ordering relationship considers the order of geometries, and has an inverse, for instance “in front of” is the inverse of “behind”.

We consider a spatial operation as a function on geometries that returns values, such as geometries, numeric, or Boolean values, except those that deal with spatial properties and relationships. Spatial properties and relationships are studied explicitly as separate categories of spatial concepts, since they play a key role in our classification. Examples of spatial operations are geometric aggregation functions such as geometric union and transformation operations such as rotate. A spatial reference system or a datum, such as the North American Datum 1983, was defined earlier. A spatial reference system has units (for instance, meter, kilometer, cm² or degree) that are used to express measures along it. The highest level of precision at which the spatial data is measured is called spatial granularity.

The above categories of spatial concepts may occur in the assertion of an IC of spatiotemporal databases. In Section 4.1, we will show that these categories of spatial concepts lead to different types of spatial ICs.

3.2. Temporal concepts within the assertions of ICs

Existing spatiotemporal conceptual models typically describe time as a one-dimensional linear reference system. In addition to a linear model, a cyclic view of time is required to model cyclic phenomena (e.g., tides). For temporal conceptual modeling based on a discrete view of time, which typically underlines most database models, basic temporal primitives are defined. Common temporal primitives are the instant, i.e., a point on the
temporal domain, and interval, i.e., a temporal duration delimited by two instants on the temporal domain [21]. The basic temporal primitives generate more complex ones, such as a set of intervals (to manage presence/absence). Relying on this temporal representation, a number of temporal concepts can appear in the definitions of ICs.

Considering both a linear and a cyclic model of time, in a manner similar to the spatial domain different categories of temporal concepts that may be involved in the definition of ICs include: “temporal primitive”, “temporal property”, “temporal relationship”, “temporal operation”, and “calendar” (or temporal reference system).

A temporal property is an intrinsic characteristic of a temporal primitive, considering it as a whole, and is of one of two types: a topological property such as the closedness of a temporal interval or a metric property. A metric property of a temporal primitive is described in a metric form relative to the temporal reference system. Example for metric properties is the duration of a temporal interval in a linear or cyclic model of time. A temporal relationship refers to the relative position of temporal primitives, and falls into one of two categories. A temporal topological relationship denotes a non-metric temporal relationship [21,22]. Examples of topological relationships between temporal intervals of a linear model of time are found in [23] (e.g., before, starts, during). Topological relationships for a linear time model describe the order between temporal intervals. However, order relations, such as before and after, do not generally hold in a cyclic time model: in a series of one-day cycles, night is after morning, but it is also before next morning. In this respect, sixteen temporal topological relationships (e.g., disjoint, contained by, finishes, overlaps twice) for cyclic intervals are recognized [24]. A temporal metric relationship exploits temporal distances between instants in the temporal dimension. A temporal operation is a function on temporal primitives that returns values, such as temporal, numeric, or Boolean values. We exclude those operations that deal with spatiotemporal properties and relationships, which are considered separately, as in the case of spatial and temporal concepts. Examples for spatiotemporal operations are geometric rotation at a specific time and geometric shape at a given time. Finally, a spatiotemporal reference system is a reference system for simultaneous measurement in space and in time. The unit of a spatiotemporal reference system is, for example, meter per second (m/s).

A number of examples for spatiotemporal ICs that include these spatiotemporal concepts as well as a classification of these ICs, will be presented in Section 4.3.

3.3. Spatiotemporal concepts within the assertions of ICs

Similar to spatial and temporal concepts, we categorize spatiotemporal concepts that may appear in ICs of spatiotemporal databases. These concepts include “spatiotemporal primitive”, “spatiotemporal property”, “spatiotemporal relationship”, “spatiotemporal operation”, and “spatiotemporal reference system”.

A spatiotemporal primitive is the geometrical representation of an object that evolves in time, such as moving-point, moving-polygon [29] or expanding-polygon. A spatiotemporal property is an inherent characteristic of a spatiotemporal primitive, such as the speed or the heading of a moving-point [30]. A spatiotemporal relationship describes a relationship between spatiotemporal primitives, and deals with space and time simultaneously, for instance, the relative speed between two moving-points at a specific time, or the topological relationships between a moving-point and an expanding-polygon during a time interval.

A spatiotemporal operation is a function on spatiotemporal primitives that returns values, such as geometric, numeric, or Boolean values. We exclude those operations that deal with spatiotemporal properties and relationships, which are considered separately, as in the case of spatial and temporal concepts. Examples for spatiotemporal operations are geometric rotation at a specific time and geometric shape at a given time. Finally, a spatiotemporal reference system is a reference system for simultaneous measurement in space and in time. The unit of a spatiotemporal reference system is, for example, meter per second (m/s).

3.4. Thematic concepts within ICs’ assertions

Thematic concepts are created by humans to describe real-life phenomena and are given names in a natural language, such as “house”, “marriage”, and “to own”. Given that there is no generally accepted and all-embracing classification for these concepts [31], we group them in the single category of “thematic concept”.

4. A formal classification of ICs in spatiotemporal databases

We classify ICs according to the types of concepts that appear in their assertions. Fig. 1 shows seven separate sets of concepts suggesting seven categories of ICs. However, in Sections 4.1–4.4 we will see that the ICs whose assertions include concepts referring to space, to time, and to space and time together have, respectively, the same sub-types as those that refer to space and themes, to time and themes, and to space, time and themes. In order to prevent redundancy, we assume one super-class for each of these couples. Accordingly, a spatial IC involves
concepts referring to space (spatial concepts) or to space and themes together (spatial concepts+thematic concepts) (Fig. 2(a)); a temporal IC includes concepts that refer to time (temporal concepts) or to time and themes together (temporal concepts+thematic concepts) (Fig. 2(b)); and a spatiotemporal IC includes concepts that refer to space and time (spatial concepts+temporal concepts, or spatiotemporal concepts) or to space, time, and themes altogether (spatial concepts+temporal concepts+thematic concepts, or spatiotemporal concepts+thematic concepts) (Fig. 2(c)). The last IC class is the thematic IC, which deals only with thematic concepts (Fig. 2(d)).

The formal definition and detailed discussions of the four super-classes of ICs (Fig. 3) based on the four categories of concepts, i.e., spatial concepts, temporal concepts, spatiotemporal concepts, and thematic concepts as well as their different types and sub-classes, are presented in Sections 4.1–4.4.

4.1. Formal definition of the spatial IC, its sub-classes and types

In this section, we will formally define a spatial IC. For this purpose, we suppose that $IC_i$ is an IC of spatiotemporal databases and $C_{IC_i}$ is the set of concepts within the assertion of $IC_i$ (these concepts can be spatial concepts $C_{spatial}$, temporal concepts $C_{temporal}$, spatiotemporal concepts $C_{spatiotemporal}$, and thematic concepts $C_{thematic}$).
The complement of an arbitrary set $S$ is denoted by $S'$. The symbol $\cup$ denotes the union of sets and the symbol $\subset$ for subset. Logical connectives are denoted by $\lor$ (logical inclusive or) and $\land$ (logical and), and $\exists$ is existential quantification. Lowercase characters $c$ denotes a concept (a member of a set of concepts).

Based on the above descriptions, the set of spatial ICs is formally defined as follows:

$$IC_{\text{spatial}} = \{IC_i : (C_{ICi} \subset (C_{\text{spatial}} \cup C_{\text{thematic}})) \land \exists c(c_i \in C_{ICi} \land c_i \in C_{\text{spatial}})\}$$

(1)

Definition 1 expresses that the set $IC_{\text{spatial}}$ consists of those $IC_i$ where the set of concepts within their assertion (i.e., $C_{ICi}$) is a subset of the union of spatial concepts $C_{\text{spatial}}$ and thematic concepts $C_{\text{thematic}}$, and whose definitions include at least one spatial concept.

Fig. 4 demonstrates a UML model for the spatial IC class. In UML, each box represents a class where the name of the class is shown on top and class attributes, if any, shown underneath. The arrows represent the generalization relationship and the unfilled diamond represents the aggregation relationship. The cardinality of each relationship is illustrated beside the relationship. The dashed line linking relationships acts as a constraint for those relationships and the constraint can be expressed by English. For more information about UML, the reader is referred to [32].

The spatial IC class can be of one of five types (according to the category of spatial concepts it involves e.g., topological or metric relationships) and can be specialized into two sub-classes (taking into account whether the spatial IC includes or excludes thematic concepts). In the rest of this section, we will formally define five types of spatial ICs.

In the following formal definitions, $C_{\text{spatial}}$, $C_{\text{metric}}$, and $C_{\text{topological}}$ are the set of spatial concepts in the spatial IC ($IC_{\text{spatial}}$). $C_{\text{topological}}$ is the set of spatial topological properties and relationships, $C_{\text{metric}}$ is the set of spatial metric properties and relationships, and $C_{\text{ordering}}$ is the set of spatial ordering relationships.

**Spatial primary IC** is defined as

$$IC_{\text{spatial primary}} = \{IC_{\text{spatial}} : (C_{\text{spatial}} \subset (C_{\text{topological}} \cup C_{\text{metric}})) \land (C_{\text{topological}} \cup C_{\text{ordering}})\}$$

(2)

The above definition states that a spatial primary IC is a spatial IC that does not include topological, metric, and ordering properties and relationships. The ICs labeled 1 in Tables 1 and 2 are examples for spatial primary ICs.

**Spatial topological ICs** are defined as follows:

$$IC_{\text{spatial topological}} = \{IC_{\text{spatial}} : (C_{\text{spatial}}(IC_{\text{spatial}}) \subset (C_{\text{topological}} \cup C_{\text{ordering}})) \land (\exists c(c_j \in C_{\text{spatial}}(IC_{\text{spatial}}) \land c_j \in C_{\text{topological}}))\}$$

(3)

expressing that a spatial topological IC consists of topological properties or relationships. Examples of spatial topological ICs include ICs labeled 2 in Tables 1 and 2.

A **spatial metric IC** includes metric properties or relationships, and enforces limitations on these properties and relationships. A spatial metric IC a member of the following set:

$$IC_{\text{spatial metric}} = \{IC_{\text{spatial}} : (C_{\text{spatial}}(IC_{\text{spatial}}) \subset (C_{\text{topological}} \cup C_{\text{ordering}})) \land (\exists c(c_j \in C_{\text{spatial}}(IC_{\text{spatial}}) \land c_j \in C_{\text{metric}}))\}$$

(4)

The ICs labeled 3 in Tables 1 and 2 are spatial metric ICs.

A **spatial ordering IC** is defined as

$$IC_{\text{spatial ordering}} = \{IC_{\text{spatial}} : (C_{\text{spatial}}(IC_{\text{spatial}}) \subset (C_{\text{topological}} \cup C_{\text{metric}})) \land (\exists c(c_j \in C_{\text{spatial}}(IC_{\text{spatial}}) \land c_j \in C_{\text{ordering}}))\}$$

(5)

expressing that a spatial ordering IC does not include topological and metric properties and relationships but it involves ordering relationships. The ICs labeled 4 in Tables 1 and 2 are spatial ordering ICs.

Finally, **spatial mixed ICs** integrate and imply limitations based on at least two different types of properties and relationships of spatial primitives or objects together, such as metric and topological.

$$IC_{\text{spatial mixed}} = \{IC_{\text{spatial}} : (C_{\text{spatial}}(IC_{\text{spatial}}) \subset (C_{\text{topological}} \cup C_{\text{metric}})) \land (\exists c(c_j \in C_{\text{spatial}}(IC_{\text{spatial}}) \land c_j \in C_{\text{topological}} \lor c_j \in C_{\text{metric}}))\}$$

(6)

The ICs labeled 5 in Tables 1 and 2 are spatial mixed ICs.

In addition to five types, as it is illustrated in Fig. 4, the spatial IC has two sub-classes. The distinction between these two sub-types is important because, as we will discuss
in Section 6, they require different specification approaches. We define these two sub-classes in Sections 4.1.1 and 4.1.2.

### 4.1.1. Spatial elementary IC

Formally, the set of spatial elementary ICs is defined as follows:

\[
\text{IC}_{\text{spatial\_elementary}} = \{ \text{IC}_{\text{spatial}} : \text{C}_{\text{IC}_{\text{spatial\_elementary}}} \subset \text{C}_{\text{spatial}} \} \tag{7}
\]

In Definition 7, \( \text{C}_{\text{IC}_{\text{spatial\_elementary}}} \) is the set of concepts of the spatial IC \( \text{IC}_{\text{spatial}} \), and this definition states that a spatial elementary IC is a spatial IC whose assertion only conveys spatial concepts, and hence defines a restriction only in space. As its name suggests, restrictions defined by a spatial elementary IC are elementary constraints on spatial primitives. Table 1 illustrates a number of examples of spatial elementary ICs.

We can precisely describe the model of spatial elementary IC. The expression of a spatial elementary IC is composed of at least two spatial concepts (Fig. 5). Given that a spatial elementary IC defines restrictions on a specific type of geometry, at least one concept among its constructs is a spatial primitive. Since a spatial elementary IC is a sub-class of the spatial IC, it inherits five types: primary, topological, metric, ordering, and mixed. For instance, IC1, IC2, IC3, IC4, and IC5 in Table 1 are, respectively, spatial elementary primary, topological, metric, ordering, and mixed ICs.

### 4.1.2. Spatio-thematic IC

Formally, a spatio-thematic IC is a member of the following set:

\[
\text{IC}_{\text{spatio\_thematic}} = \{ \text{IC}_{\text{spatial}} : \exists \text{C}_i (\text{C}_i \in \text{C}_{\text{IC}_{\text{spatial}}}, \text{C}_i \in \text{C}_{\text{thematic}}) \} \tag{8}
\]

Spatio-thematic ICs constitute the majority of the ICs in the conceptual models of spatial database applications. Table 2 provides different examples of spatio-thematic ICs.
The model of a spatio-thematic IC in Fig. 6 shows that this IC is made up of at least one spatial concept and at least one thematic concept together.

Like a spatial elementary IC, a spatio-thematic IC inherits one of the five types of the spatial IC. For example IC1, IC2, IC3, IC4, and IC5 in Table 2 are spatio-thematic primary, topological, metric, ordering, and mixed ICs, respectively.

4.2. Formal definition of the temporal IC, its sub-classes and types

A temporal IC, which enforces a restriction in time, maybe in themes, but not in space, is a member of the following set:

$$IC_{\text{temporal}} = \{ IC_i : (C_{\text{temporal}} \cup C_{\text{thematic}}), \exists (C_i \in C_{\text{IC}_i}) \}$$ (9)

A temporal IC is an IC whose assertion includes temporal concepts and possibly thematic concepts but does not involve spatial concepts. A temporal IC class is of one of three types, primary, topological, and metric. In addition, depending on if a temporal IC includes or excludes thematic concepts, it can be specialized into one of two sub-classes, i.e., temporal elementary IC and temporal–thematic IC (Fig. 7). The three types of the temporal IC class are discussed in this section.

In the following definitions of the three types of the temporal IC class, we denote by $C_{\text{topological}}$ the set of temporal topological properties and relationships, $C_{\text{metric}}$ the set of temporal metric properties and relationships, and $C_{\text{temporal}} (IC_{\text{temporal}})$ the set of temporal concepts in the assertion of a temporal IC, $IC_{\text{temporal}}$.

A temporal primary IC is defined as

$$IC_{\text{temporal primary}} = \{ IC_{\text{temporal}} : (C_{\text{temporal}} \cap C_{\text{thematic}}), \exists (C_i \in C_{\text{IC}_i}) \}$$ (10)

stating that a temporal primary IC involves temporal concepts, excluding temporal properties and relationships. The ICs labeled 1 in Tables 3 and 4 are examples of temporal primary ICs.

A temporal topological IC is a member of the following set:

$$IC_{\text{temporal topological}} = \{ IC_{\text{temporal}} : (C_{\text{temporal}} \cap C_{\text{thematic}}) \subseteq C_{\text{topological}} \}$$ (11)

expressing that a temporal topological IC is defined by means of temporal topological properties and relationships. Examples of temporal topological ICs are the ICs labeled 2 in Tables 3 and 4.

Finally, a temporal metric IC is a member of the following set:

$$IC_{\text{temporal metric}} = \{ IC_{\text{temporal}} : (C_{\text{temporal}} \cap C_{\text{thematic}}) \subseteq C_{\text{metric}} \}$$ (12)

expressing that a temporal metric IC is defined by means of temporal metric properties and relationships. A temporal metric IC restricts, for example, the duration of intervals and the temporal distance between temporal instants, such as IC3 in Table 3.

As time is anisotropic and one directional, any combination of topological and metric relationships within an IC can separately be expressed by a topological IC and a metric IC. For instance, similar to a mixed spatial IC, let us consider IC4 in Table 3. This IC can be separated into a temporal elementary topological IC and a temporal elementary metric IC as follows: “two intervals overlap” and “the temporal distance between the start times of two intervals should be less than 90% of the intervals’ temporal union”. Consequently, there is no need to define a
sub-class of temporal mixed IC analogous to the class of spatial mixed ICs.

Similar to spatial ICs, there are two sub-classes of temporal ICs: temporal elementary ICs and temporal–thematic ICs (Fig. 7). We will define these sub-classes in the rest of this section.

4.2.1. Temporal elementary IC

We formally define the set of temporal elementary ICs as follows:

$$IC_{\text{temporal \_ elementary}} = \{IC_{\text{temporal}} : C_{IC_{\text{temporal}}} \subseteq C_{\text{temporal}}\}$$  \hspace{1cm} (13)

As the examples in Table 3 show, a temporal elementary IC defines basic restrictions on the temporal primitives.

The model of the temporal IC in Fig. 8 expresses that the assertion of a temporal elementary IC is made up of at least two temporal concepts. Since the temporal elementary IC defines limitations on a specific temporal primitive, at least one of the concepts involved in this IC must be a temporal primitive.

A temporal elementary IC, like its super-class, is of three types: primary (e.g., IC1 in Table 3), topological (e.g., IC2 in Table 3), and metric IC (e.g., IC3 in Table 3).

4.2.2. Temporal–thematic IC

The set of temporal–thematic ICs is formally defined as follows:

$$IC_{\text{temporal \_ thematic}} = \{IC_{\text{temporal}} : \exists C_i \in C_{IC_{\text{temporal}}} \land C_j \in C_{\text{thematic}}\}$$  \hspace{1cm} (14)

Definition 14 states that a temporal–thematic IC is a temporal IC whose assertion involves thematic concepts. Such ICs, which enforce a limit both in time and in themes, are well-recognized in the temporal database applications. Table 4 presents some examples of temporal–thematic ICs.
Fig. 9 illustrates the model of temporal–thematic ICs and expresses that these ICs comprise at least one temporal concept and at least one thematic concept.

Different types of temporal–thematic ICs are shown in Table 4. In this table, IC1 is a temporal–thematic primary IC, IC2 is a temporal–thematic topological IC, and IC3 is a temporal–thematic metric IC.

### 4.3. Formal definition of the spatiotemporal IC, its sub-classes and types

A spatiotemporal IC, which defines a restriction both in space and in time and may enforce restrictions in themes, is a member of the following set:

$$\text{IC}_{\text{spatiotemporal}} = \{ \text{IC}_i : \exists \text{CI}_j, \text{CI}_k \in \text{IC}_{\text{spatiotemporal}} \land \text{CI}_j \in \text{C}_{\text{spatial}} \land \text{CI}_k \in \text{C}_{\text{temporal}} \}$$

(15)

The above definition expresses that the assertion of a spatiotemporal IC conveys spatiotemporal concepts, or spatial concepts and temporal concepts simultaneously. Relying on the fact that a spatiotemporal IC may include or exclude thematic concepts, we consider two sub-classes, spatiotemporal elementary and spatiotemporal–thematic ICs, which will be discussed in Sections 4.3.1 and 4.3.2. Additionally, taking into account that a spatiotemporal IC conveys spatial concepts and temporal concepts, spatiotemporal concepts, or combinations thereof, we distinguish three types of spatiotemporal ICs: inherent, composite, and mixed (Fig. 10). These three types will now be formally defined.

For the formal definition of these types, let $\text{C}_{\text{IC}_{\text{spatiotemporal}}}$ be the set of concepts within a spatiotemporal IC. For simplicity, let $\text{C}_{\text{IC}_{\text{spatiotemporal}}}$ be the complete set of spatial, temporal, and spatiotemporal concepts within the assertion of $\text{IC}_{\text{spatiotemporal}}$. The three types of spatiotemporal ICs are defined in Definitions 16–18.

A spatiotemporal inherent IC is a member of the following set:

$$\text{IC}_{\text{inherent}} = \{ \text{IC}_{\text{spatiotemporal}} : \text{C}_{\text{spatiotemporal}}(\text{IC}_{\text{spatiotemporal}}) \subseteq \text{C}_{\text{spatial}} \}$$

(16)

As shown in the third column of Tables 5 and 6, the assertions of spatiotemporal ICs labeled 1 do not involve spatial and temporal concepts, and hence are both spatiotemporal inherent ICs.

A spatiotemporal composite IC is a member of the following set:

$$\text{IC}_{\text{composite}} = \{ \text{IC}_{\text{spatiotemporal}} : \text{C}_{\text{spatial}}(\text{IC}_{\text{spatiotemporal}}) \subseteq \text{C}_{\text{spatial}} \cup \text{C}_{\text{temporal}} \}$$

(17)

The ICs labeled 2 in Tables 5 and 6 are examples for spatiotemporal composite ICs.

Finally, a spatiotemporal mixed IC is a member of the following set:

$$\text{IC}_{\text{mixed}} = \{ \text{IC}_{\text{spatiotemporal}} : \exists \text{CI}_j, \text{CI}_k \in \text{C}_{\text{spatiotemporal}} \land \text{CI}_j \in \text{C}_{\text{spatial}} \land \text{CI}_k \in \text{C}_{\text{temporal}} \}$$

(18)

Referring to Table 5, IC3 includes spatiotemporal concepts, spatial concept, and temporal concepts. In Table 6,
IC3 involves thematic concepts, spatiotemporal concepts, spatial concept and temporal concepts. Therefore, both these ICs are spatiotemporal mixed ICs.

Like spatial ICs and temporal ICs, spatiotemporal ICs include two sub-classes (Fig. 10). We will explain these two sub-classes in Sections 4.3.1 and 4.3.2.

4.3.1. Spatiotemporal elementary IC

A spatiotemporal elementary is a member of the following set:

\[ IC_{\text{spatiotemporal elementary}} = (IC_{\text{spatiotemporal}} : C_{IC_{\text{spatiotemporal}}} \subset (C_{\text{spatial}} \cup C_{\text{temporal}} \cup C_{\text{spatiotemporal}})) \]  

(19)

stating that the assertion of a spatiotemporal elementary IC does not involve thematic concepts and defines a restriction in space and in time but not in themes. As its name implies, a spatiotemporal elementary IC defines elementary restrictions for either spatiotemporal primitives or spatial primitives in time. Some examples of spatiotemporal elementary ICs are given in Table 5.

We define the model of the spatial elementary IC in Fig. 11. The assertion of a spatiotemporal elementary IC includes at least one spatiotemporal primitive or a spatial primitive in addition to at least a spatiotemporal concept, a spatial concept, or a temporal concept.

Like its super-class, the class of spatiotemporal elementary IC is composed of three types. Examples of these three types are IC1, IC2, and IC3 in Table 5 which are, respectively, a spatiotemporal elementary inherent, composite, and mixed IC.

4.3.2. Spatiotemporal–thematic IC

We formally define the set of spatiotemporal–thematic ICs as

\[ IC_{\text{spatiotemporal-thematic}} = \{ IC_{\text{spatiotemporal}} : \exists C_i (C_i \in C_{IC_{\text{spatiotemporal}}} \land C_i \in C_{\text{thematic}}) \} \]  

(20)

A spatiotemporal–thematic IC is a spatiotemporal IC that conveys thematic concepts. Therefore, a spatiotemporal–thematic IC defines a restriction in space, time, and themes altogether. Table 6 shows a number of examples of spatiotemporal–thematic ICs.

As shown in Fig. 12, the assertion of the spatiotemporal–thematic IC comprised of at least one thematic concept in addition to at least either a spatiotemporal concept, or a spatial concept and a temporal concept simultaneously.
Spatiotemporal–thematic ICs are a sub-class of the spatiotemporal IC. Therefore, they inherit three types inherent, composite, and mixed. IC1, IC2, and IC3 in Table 6 are spatiotemporal–thematic inherent, composite, and mixed ICs, respectively.

4.4. Formal definition of the thematic IC

A thematic IC, which does not enforce restrictions in space and in time, is formally a member of the following set:

$$IC_{\text{thematic}} = \{ IC_i : C_{IC_i} \subset C_{\text{thematic}} \}$$

Examples of thematic ICs are a domain IC like the “type” of a “building” must be one of the followings: “residential”, “commercial”, and “industrial”, or the IC “salary must be greater than zero”. A model of the thematic IC is shown in Fig. 13.

Fig. 12. Model of the spatiotemporal–thematic IC.

Fig. 13. Model of the thematic IC.

5. Example: An agricultural database application

To illustrate how the proposed IC classification applies to real database applications, this section classifies the ICs of a spatiotemporal conceptual model for an agricultural application. The conceptual model includes six object classes, i.e., province, field, irrigation zone, operation, crop, and wind. Each object class has its own attributes represented as the arguments inside parentheses. ICs for this application are specified after the definition of object classes. We should add that the aim of this section is not to define all possible ICs for this example; rather it intends to include a selected number of illustrative ICs for the classification purpose.

Province (name, location)
Field (location, existence)
Irrigation zone (location, existence)
Operation (type)
Crop (type)
Wind (zone, speed, existence)

IC1: The location of a province is represented by a polygon.
IC2: A polygon is closed.
IC3: The existence of an irrigation zone is modeled by a temporal interval.
IC4: The start date of a temporal interval must be before its end date.
IC5: The speed of a wind is greater or equal to zero and less than 200 km/h.
IC6: The type of an operation includes: “seeding”, “fertilizing”, “soil preparation”, and “protection”.
IC7: Periods that two crops with two different types are associated to the same field must not overlap.
IC8: A field whose location is inside the location of a province whose name is “Quebec” cannot have the operation of type “seeding” after November and before April in a given year.
IC9: For a given field F and a given wind W, if the speed of W is greater than 40 km/h and the location of F is inside the zone of W and the existences of F and W
overlaps, then the operation of $F$ cannot be of type "seeding".

IC10: The location of an irrigation zone overlaps the location of at least one field.

IC11: The zones of two winds whose existences overlap must not overlap.

IC12: The existence of an irrigation zone lasts at most 3 h.

IC13: Each field should be associated with exactly one operation.

ICs 1, 2, and 10 convey spatial concepts (e.g., polygon, overlap) or spatial concepts and thematic concepts together; therefore, according to Definition 1 they are spatial ICs. IC2 bears only spatial concepts such as polygon and restrict only in space. Hence, referring to Definition 7, IC2 is a spatial elementary IC. The remaining spatial ICs, considering Definition 8, are spatio-thematic ICs since they include thematic concepts (e.g., province, field).

The assertions of ICs 3, 4, 7, and 12 bear temporal concepts (e.g., temporal interval, hour) or temporal concepts and thematic concepts. Referring to Definition 9, these ICs are temporal ICs. IC4 does not involve thematic concepts but include temporal topological relationship (i.e., before) and according to Definitions 11 and 13 is a temporal elementary topological IC. ICs labeled 3, 7, and 12 convey thematic concepts (i.e., irrigation zone) and are temporal–thematic ICs (Definition 14). IC3 conveys only temporal primitive “interval”; consequently, it is a temporal–thematic primary IC (Definition 10). IC7 involves a temporal topological relationship (i.e., overlap) and hence is a temporal–thematic topological IC (Definition 11). Finally, IC12 includes temporal distance “3 hours” and is a temporal–thematic metric IC (Definition 12).

ICs 5, 8, 9, and 11 are, based on Definition 15, spatiotemporal IC as their assertions convey spatiotemporal concepts (e.g., speed) or spatial concepts (e.g., inside) and temporal concepts (e.g., after) together. All these ICs are spatiotemporal–thematic ICs. For instance, as Definition 16 states, IC5 is a spatiotemporal–thematic IC of the type inherent since it bears thematic concepts (e.g., wind) and spatiotemporal concepts (e.g., speed) but not spatial or temporal concepts. Similarly and by referring to the Definition 17, the type of IC8, IC9, and IC11 are composite.

Considering the Definition 21, IC6 and IC13 are thematic ICs. Table 7 summarizes the results of the classification of the above ICs.

Some of the above examples (e.g., IC1) fall into all three classes of spatial ICs proposed by Cockcroft (i.e., topologic, semantic, or user-defined). In addition, using her classification, one cannot distinguish between the above temporal ICs and spatiotemporal ICs. Brisaboa et al. [6] classified ICs by taking into account if they are defined for objects or for relationships. Based on their approach, all of the above ICs are classified into two categories, and their approach does not differentiate between the ICs of this example, regarding the two important constituents of spatiotemporal databases, i.e., space and time. In the next section, we will see the advantage of such a more exhaustive classification in providing a foundation for future specification of ICs.

6. Discussion on the applications of the proposed classification

In Section 2, we stated that for referring to a simple spatial IC, diverse terms have been used in the literature.
This terminological heterogeneity should be resolved to enable meaningful information exchange or interoperability among agents that are involved in the specification and implementation of ICs. The formal classification we proposed in this paper, with the suggested nature-based vocabulary for labeling different IC classes, establishes a terminological frame of reference for ICs of spatial and spatiotemporal database applications. Such a terminological reference is expected to improve the general interoperability between different agents dealing with ICs.

Furthermore, the results of Section 5 illustrate that the number of ICs of even a simple conceptual model may be surprisingly large and diverse. Undertaking a complete, correct, and readable specification of these inherently diverse ICs at the conceptual level is not a straightforward task. To have a consistent specification, a number of key questions should be answered. This section presents these questions and discusses how the analysis of the proposed classification contributes to answering them. When we give examples, the IC number refers to the ICs of the example in Section 5.

(a) In which document of the conceptual model (i.e., conceptual schema or data dictionary) should we specify ICs?

A conceptual model has two components, a conceptual schema and a data dictionary, and ICs should eventually be specified in one of these components. Spatio-thematic primary ICs, temporal–thematic primary ICs, and spatiotemporal–thematic inherent ICs, which express the type of spatial, temporal, and spatiotemporal shapes of features, respectively, are often specified in the conceptual schemas using spatiotemporal schema modeling languages, such as spatiotemporal extensions to UML. Therefore, ICs labeled 1 and 3 are defined in the conceptual schema. Fig. 15 shows the specification of IC1 using Perceptrory’s [33] spatial pictogram. This pictogram expresses that the location of a province must be represented by a polygon.

The number of ICs involving spatial and temporal relationships, however, may end up being over several times the number of classes in a conceptual schema and may be complex to express (see for example IC9). Consequently, Bédard et al. [16] suggest defining these ICs in the data dictionary, in a complementary manner to those defined in the schema, in order to prevent the schemas from becoming too complex and unreadable. Accordingly, from a readability point of view, spatio-thematic topological, spatio-thematic metric, spatio-thematic ordering and spatio-thematic mixed ICs, temporal–thematic metric and temporal–thematic topological ICs, as well as spatiotemporal–thematic composite and mixed ICs are best defined in the data dictionary using specific languages for defining IC called integrity constraint specification languages (ICSL), ideally with the help of dedicated software such as G6 [7] or Radius Studio [34].

(b) Who should specify and validate ICs?

Typically, the ICs that involve thematic concepts (i.e., spatio-thematic ICs, temporal–thematic ICs, spatiotemporal–thematic ICs, and thematic ICs) express the semantics of a specific application. These ICs should be built on the top of the spatiotemporal DBMS technology. Application domain specialists, who know what ICs should be defined for their application and should validate the specified ICs before they are implemented, must be involved in defining these four sub-classes of ICs. Referring to the examples in Section 5, ICs labeled 1, 3, 5, 6, 7, 8, 9, 10, 11, 12, and 13 should be primarily defined and validated by application domain specialists, typically with the help of system analysts.

On the other hand, spatial elementary ICs, temporal elementary ICs, and spatiotemporal elementary ICs convey restrictions on the propriety spatial, temporal, and spatiotemporal data structure for a spatiotemporal DBMS technology. Therefore, they are typically specified by software providers when designing their software engine, by standards developers, such as ISO/TC211 or OGC, or by in-house developers of spatiotemporal applications, i.e., typically by software designers. As a result, ICs labeled 2 and 4 are expected to be specified and validated by software designers.

(c) Which type of ICSL to use?

Selecting a suitable ICSL to specify the ICs of an application is not straightforward as several factors such as understandability, readability, usability, and expressivity of the ICSL should be taken into account. The investigation of the proposed classification of ICs in this paper provides general guidelines for selecting an appropriate ICSL. These guidelines can be investigated further to provide a more concrete approach in selecting an appropriate ICSL for an application.

Different types of ICSLs exist [36]. Some formal ICSLs like OCL generally use natural languages such as English, but they still suffer from a low-level of readability and usability except for those having software engineering knowledge. OCL allows expressing ICs and queries on UML class diagrams. The specification of IC10 using the spatial extension to OCL, Spatial OCL [9], is as follows:

context irrigation_zone inv:
  self.field- > exist(f | f.location - > overlaps self.location))

A number of ICSLs use controlled natural languages, i.e., a subset of a natural language, and express the ICs similarly to the way non-specialists express ICs in natural languages. For instance, Normand [7] proposes a controlled natural language combined with spatial pictograms for expressing spatial ICs. A non-specialist of OCL can easily define IC10 using this language:

Object 1
irrigation zone

Object 2

ics should be, among others, easy to learn, to understand, and to use for the involved agents (application
domain specialist, system analyst, software developer, etc.). No language meets these criteria for all categories of agents. In this context, deciding on which category of language to use for the specification of ICs is in fact a matter of taking into account who specifies the ICs, who validates the ICs, who implements the ICs, which categories of ICs are specified in which document by whom, which system implements the ICs, etc. For example, software designers prefer formal software engineering approaches that can be easily translated into programming code. On the other hand, application domain specialists prefer natural languages as they typically do not familiar with formal languages. Consequently, a preferred ICSL for specifying spatial elementary IC, temporal elementary ICs, and spatiotemporal elementary ICs, which are defined and validated by software designers, could be OCL, for instance. On the other hand, the definition of spatio-thematic ICs, temporal–thematic ICs, spatiotemporal–thematic ICs, and thematic ICs could be done with and validated by application domain specialists, for example. Application domain specialists typically do not have the pertinent mathematical or software engineering background. Therefore, in such cases, controlled natural languages are better ICSL candidates for the specification of the above four sub-classes of ICs.

(d) What are the constructs that should be supported by an ICSL to define the different categories of ICs of spatiotemporal databases?

The particularity of the proposed classification of ICs, compared to the previous ones, is that it is based on the type of concepts that appear in the assertion of ICs. Such a classification, and the discussions in Section 3, set out a framework to identify the type of constructs that should be integrated into an ICSL to define different IC classes in spatiotemporal database applications. To express spatial ICs, spatiotemporal composite ICs, and spatiotemporal mixed ICs, the ICSL should include constructs for spatial primitives, spatial properties and relationships, spatial operations, and spatial reference systems. Likewise, in order to define temporal ICs, spatiotemporal composite ICs, and spatiotemporal mixed ICs, the language must support the constructs conveying temporal primitives, temporal properties and relationships, temporal operations, and calendars. In the same manner, for defining spatiotemporal mixed ICs, in addition to spatial and temporal constructs, the ICSL should support spatiotemporal primitives, spatiotemporal properties and relationships, spatiotemporal operators, and spatiotemporal reference systems.

It is worth noting that the majority of existing spatial ICSLs (e.g. Spatial OCL) only support spatial primitives and topological relationships, and consequently, can express a rather restricted number of spatial ICs (i.e., spatial topological ICs). In addition, existing spatial ICSLs and temporal ICSLs, e.g., Temporal OCL [37] do not include spatiotemporal constructs. Hence, it is extremely cumbersome or sometimes impossible to express a number of spatiotemporal inherent ICs and spatiotemporal mixed ICs using existing ICSLs. The study of the various concepts that may appear in the definition of ICs (Section 3) provides the necessary constructs to build a more expressive spatial ICSL as well as a spatiotemporal ICSL.

Based on the proposed classification in this paper, we have developed a formal ICSL for expressing spatial, temporal, spatiotemporal, and thematic ICs in spatial databases. The description of syntax and semantics of this ICSL as well as several examples of ICs, defined using this ICSL, are presented in [38].

7. Conclusion

Existing classifications of ICs in spatial and spatiotemporal databases suffer from two shortcomings: (1) they are not defined precisely, and (2) they only distinguish among very simple ICs. In this paper, we introduced a formal classification for the ICs in spatiotemporal databases. An IC is assumed as an assertion that conveys a number of concepts with respect to space, time, and themes, and their combinations. Relying on the nature of these concepts and the type of restrictions that ICs convey, the ICs are classified. The classification resulted in four super-classes: spatial ICs, temporal ICs, spatiotemporal ICs, and thematic ICs. Then, spatial ICs can be subdivided into five types (i.e., primary, topological, metric, ordering, mixed), the temporal IC into three types (i.e., primary, topological, metric), and the spatiotemporal IC as well (e.g., inherent, composite, mixed). In addition, by considering if an IC includes or excludes thematic concepts, each of the IC super-classes (spatial IC, temporal IC, and spatiotemporal IC) are specialized into two sub-classes: elementary and thematic. The proposed classification and vocabulary suggest a common terminological framework for the ICs in spatiotemporal database applications. Using a same term to refer to ICs whose specification strategy is similar makes the semantic interoperability between agents dealing with ICs easier.

Furthermore, we linked the proposed classification with an important phase in designing a database application, i.e., the specification of ICs. The proposed classification of ICs in this paper provides a considerably more exhaustive classification and a foundation for the specification of ICs in spatial and spatiotemporal databases. The specification of ICs is not a straightforward practice and several questions may arise during this process. Further synthesis of the specification of different IC classes provides an insightful guideline for answering to typical questions facing practitioners in the spatiotemporal database realm, such as, where, with whom, with which ICSL, and how the different classes of ICs are specified.

As future research work, there is a need to develop a system to verify data against the ICs in spatiotemporal databases. This system should prevent the insertion of data not respecting the ICs into the database. Additionally, the system should also be able to relax some ICs. Some thematic-based ICs are not always respected in the real world. For instance for sewerage, one can define a spatial-thematic IC expressing the slope must be regular from upstream to downstream. But in some well-known cases, some tubes are not correctly positioned. Hence, the system must be able to distinguish those cases from errors. Another example is the telephone system. To facilitate connecting people to the telephone system, some cables are placed in advance and they will be connected when
necessary. Here again, some relaxing system must be installed when developing the thematic-based ICs.

References

[33] Perceptory, Available at <http://sirs.scg.ulaval.ca/Perceptory/>.
[38] M. Salehi, A formal language for the specification of integrity constraints in spatial databases, in: developing a model and a language to identify and specify the integrity constraints in spatial databases, PhD. Thesis, Laval University, Canada, 2009, pp. 147–171.