Published in: International Conference on Advanced GeographicInformation
Systems & Web Services, GeoWS 2009, pp. 146-151,

Categories of Geospatial and Temporal Integrity Constraints

Stephan Mäs and Wolfgang Reinhardt
AGIS, University of the Bundeswehr Munich, 85577 Neubiberg, Germany
Stephan.Maes | Wolfgang.Reinhardt @unibw.de

Abstract

Integrity constraints play a major role when the logical consistency of a data set has to be evaluated during quality assurance. The paper reviews established approaches of integrity constraint classification and proposes a profound categorisation which distinguishes semantic integrity constraints according to the semantic domains of the involved properties. The categorisation provides the preconditions for formalisation and the identification of conflicts and redundancies in sets of integrity constraints. Therewith the management and usability of integrity constraints will be improved which will also result an improvement of spatial data quality.

1. Introduction

Integrity or sometimes also called consistency is a term originally used for the property of database systems of being free of logical contradictions within a model of reality. This model also contains defined integrity constraints **(IC)** that must hold on the database to grasp the semantics intended by the model.

In the last decade much research has been done in the area of GI quality assurance with integrity constraints [1][2], IC definition, formalisation, exchange and management [3][4][5][6] and validation of internal consistency of IC sets [7]. Nevertheless most of these approaches only consider specific types of IC. The aim of this paper is to review available proposals of IC classification and to develop a profound categorisation as a basis for further work on IC formalisation, management and validation. A particular focus is laid on the particularities of spatiotemporal data.

IC are associated with logical consistency, which specifies the "degree of adherence to logical rules of data structure, attribution and relationships" [8]. The constraint categorisation presented in this paper does not include constraints concerning the accuracy and completeness of data, which describe a lack of contradictions with reality.

2. Approaches to IC categorisation

This chapter reviews existing approaches of IC categorisation starting with general constraint categories in database systems and thereafter the specific constraint types for spatiotemporal data.

2.1. General IC for database systems

2.1.1. Specification technique. Some fundamental categories of IC for data modelling have been defined by Elmasri and Navathe [9]. They classify IC according to their specification technique, to the type of conditions they specify and the number of database states they constrain.

The specification technique of constraints in a database system can be inherent, implicit or explicit. **Inherent constraints** are directly associated with the constructs of the data model itself and do not need to be specified in the data schema. **Implicit constraints** are, as well as the inherent constraints, contained in the database schema but are specified by the data definition language (DDL) during the database schema definition. They describe each entity type, attribute and relationship through the specification possibilities implied by the particular schema DDL. An example is the uniqueness constraint that is put on an attribute when it is specified as a key of an entity type. More complex constraints which are not expressed by the DDL, and therefore have to be additionally specified, are called **explicit constraints** [9]. Typical examples are the general semantic IC, mentioned below.

2.1.2. Specified conditions. A second classification method identifies the following types of IC regarding the specified conditions [9]:

- **domain constraints** restrict the allowed types of values of an attribute
- **key and relationship constraints** refer to the possibility to define key values (i.e. unique values) for entity classes, cardinalities for relationships between entity classes and participation requirements
- **general semantic integrity constraints** are explicitly specified and usually more complex. They

refer to semantics of the modelled entity classes which are not representable through the other two categories. They specify relations between the modelled concepts which are usually not explicitly represented in the data.

2.1.3. Number of constrained database states. Elmasri and Navathe [9] distinguish IC based on whether they restrict a single database state (**state constraints**) or multiple states (**transition constraints**). Thereby one database state includes all data of the database at a particular point in time.

2.1.4. Involved data. A fourth classification approach was made by [10]. They differentiated IC according to the involved data:

1. IC referring to an attribute of a single entity
2. IC referring to at least two attributes of a single entity
3. IC referring to all entities of a single entity class
4. IC referring to an entity and its associated entities of various classes
5. IC referring to operations of entities.

2.2. Spatial data integrity constraints

The categorisations listed so far do not address the particularities of spatial data. These specific properties allow for spatial analysis methods, which can also be used for integrity checking and thus enable the definition of spatial IC. These constraints particularly deal with things like location, extent, shape or topology. This subchapter summarises some approaches to classify spatial IC in order to comprehend the peculiarities of spatial data.

Servigne et al. [1] define three kinds of spatial IC that apply to structural, geometric and topo-semantic conditions. Structural errors result from an insufficient implementation of the data model based on the data structures provided by the GIS. To overcome this shortcoming structural constraints have to be defined. Structural constraints are “programming tricks” used to handle entities that can’t be appropriately represented by the available data structures. In the categorisation following in the next section we do not consider structural errors since we assume that the data structures sufficiently represent the data model.

Geometric constraints refer to the general geometric and topological assumptions of the geometry types of the data model. They define and restrict properties and relations of geometric and topological primitives independently of the semantics of specific entity classes. A very common example of a geometric IC is considering the closeness of polygons:

‘Polygons must be closed.’

The topo-semantic constraints of [1] refer to topological relations between two entities. Since the validity of the topological relation depends on the semantic of the entities topo-semantic constraints are a subtype of the general semantic IC defined by [9]. In [4] topo-semantic constraints are subdivided into semantic and user defined constraints. The former are based on the nature and the physics of the objects, e.g. ‘roads are not running through lakes’. User defined IC are of more artificial nature and describe social or business rules or laws defined by humans e.g. ‘a fuel station should not be within a certain distance of a school.’ A more extensive classification of constraint origins or contexts can be found in Frank [11]. He demonstrated how IC are part of a GIS ontology and showed that different constraints are appropriate to different tiers of the ontology. This points out the importance to treat the constraints not independently of the context for which they are valid.

3. A refined categorisation of IC

The IC categorisations mentioned in the previous section are either not practical for constraint formalisation and validation, leaving out spatial aspects or do only cover some spatial aspects like topology. In this paragraph a refined categorisation (see figure 1), which particularly incorporates the different aspects of spatial data integrity, will be introduced. It distinguishes the constraints according to the involved types of conditions and profoundly differentiates the aspects of semantic IC. This is particularly useful for the validation of the internal consistency of IC sets, as it has been demonstrated in [7] and [12].

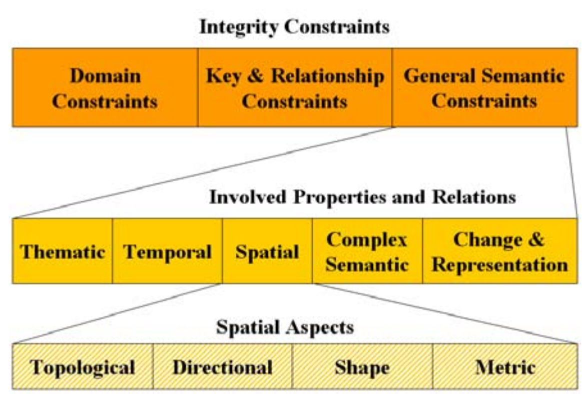

Figure 1. A categorisation of integrity constraints.

The categorisation shall provide a basis for further work on IC formalisation, management and validation. It points out the properties, which can be constrained through the integrity rules and thereby identifies aspects, where presently the potential of the constraints is by far not exploited.

The new concepts extend the well established classification of database IC of [9]. The top level of figure 1 contains these three basic types of IC. The remainder of this chapter provides definitions and examples of the defined IC classes with particular focus on the IC required by spatial data. It is structured in analogy to the presented categorisation.

3.1. Domain constraints

Domain constraints restrict the allowed types of values of an attribute. The value domain of an attribute is usually defined by a data type such as numerical data types like integer or real, character and string data types, Boolean data types, date and time data types or enumerated data types [9]. Database systems and schema description languages commonly include these data types and have the corresponding IC inherently defined. Moreover some applications require so called user-defined data types which are defined by the designers of the (database) schema.

Spatial and temporal information require data types and corresponding domain constraints which restrict all kinds of defined primitives and complexes like for example geometric / topological and temporal primitives / complexes. The primitives and their corresponding IR are defined independently of the semantics of entity classes. Currently, a variety of database systems is already capable to handle the particular requirements of spatial data and provides predefined data types. For these systems the corresponding IC don't have to be explicitly specified.

Domain constraints on geometric and topological primitives correspond to the geometric constraints defined by [1], which have been itemised in the previous chapter. International standards (e.g. [13]) specify names and geometric definitions for geometry types and geometric primitives. Since the constraints are part of the geometry type definition of the data model they are also included in these standards.

Temporal primitives also define domain constraints but since time is unidirectional temporal primitives are less complex than the potential geometrical primitives of the 3-dimensional space. For many applications interval is the only temporal primitive beside the mentioned common date and time data types. A simple constraint like:

'the beginning of an event (i.e. interval) must be before the ending'

is the only internal assumption for time intervals and it sufficiently assures their integrity. Temporal primitives are also standardised by the ISO [14].

3.2. Key and relationship constraints

For the definition of key and relationship constraints we also refer to [9]. These constraints are mostly inherently or implicitly incorporated in the schema of the database and therefore don't need to be additionally specified. Examples are cardinality restrictions of associations. Since spatial and temporal information doesn't have particular requirements on key and relationship constraints they are no further researched in this paper.

3.3. General semantic integrity constraints

Following the definition of [9] general semantic integrity constraints **(SIC)** are based on relations between the involved entities or on specific properties of a single entity. The validity of the relations is based on the semantics of the entities. SIC are defined on the level of the entity classes and have to be explicitly defined. A further subdivision of SIC can be made by grouping the restricted properties corresponding to their semantic domain. Extending the approaches of [1] and [4] we do not only consider topological relations. The categorisation (see figure 1) includes thematic, temporal, spatial and change relations as well as relations for the connection between multiple representations of an entity. Additionally we define complex SIC that combine relations of more than one of the semantic domains, mereological relations and domain specific associations. Furthermore we do not restrict constraints to concerning only two entity classes. SIC allow for the definition of restrictions affecting single or multiple entity classes in combination with their attributes and relations in a single constraint.

The following subchapters provide definitions and examples of the subtypes of SIC and refer to relations and properties appropriate for constraint definition.

3.3.1. Thematic SIC. Thematic SIC refer to the consistency of thematic attributes. They restrict the ranges of attributes of a single entity by specifying relations between the values of two or more attributes or one attribute value and a defined other value, for example for a road entity class with the attributes *road_type* and *number_of_driving_lanes*:

‘Roads of the type ‘Autobahn’ must have at least two driving lanes.’

The applied comparison operators are well established and have been used in many standards, e.g. in [15]. Beside the common order relations like (= , < , > , >= , <= , < >) it includes operators for String and NULL comparison as well as fundamental arithmetic operators for addition, subtraction etc..

3.3.2. Temporal SIC. To assure the logical consistency of temporal information temporal SIC can be defined. These constraints substantiate rules that examine all temporal characteristics of the data. A simple example is the definition of a minimal duration of a particular time interval or event. More complex temporal IC refer to the relation of particular points in time or time intervals like events, processes, states or actions. An example could be for an entity class bridge with the temporal attributes *construction_time* and *official_opening_date*:

‘The construction time of a bridge must be before the date of the official opening.’

In [16] Allen defines a set of 13 binary relations between times intervals which can be used for the definition such constraints. The intervals represent events through an ordered pair of points of time with the first point earlier than the second on a time scale (domain constraint). Every relation of two intervals can then be determined by not more than two relations between the start- and endpoints of the involved intervals.

3.3.3. Spatial SIC. For the definition of SIC, which restrict the spatial arrangement of entities and their spatial properties, relations specifically considering the spatiality have to be defined. Spatial relations between entities are usually not explicitly stored, but can be inferred from the geometries, shapes or extents of the entities and are therefore also used for spatial analysis. Such spatial relations can be subdivided according to the different aspects of spatial relations into topological, directional, shape and metric relations. This differentiation has also been made on the lower level of the classification in figure 1. The distinction of the different spatial aspects, the concepts of mereology and change is similar to the structure proposed by [17] in their overview on qualitative spatial representation and reasoning. For a more extensive survey of qualitative spatial relations we refer to this article.

Topological SIC. Topology is probably the most fundamental perception of space [17]. Topology is a purely qualitative concept, independent of any quantitative measures and concerns the spatial connectedness of entities. Topological relations stay invariant under linear and affine transformations like rotation, translation and scaling. Topological relationships between two entities have been extensively studied in the literature. For an overview we refer to [17].

Topological relationship predicates like *Intersects* or *Overlaps* are commonly used in SIC, e.g.:

‘Lakes are not allowed to intersect with contour lines’

Topological SIC can also restrict the topology of a single entity. Therefore properties which specify things like the internal connection, the number of components and presence/absence of holes of a single entity are constrained, for example for lake entities with polygon geometry:

‘The inner rings of a lake (which represent the islands) are not allowed to intersect with each other or with the outer ring.’

Such restrictions extend the geometry type definitions with regard to the semantic of the concerned entity classes.

Directional SIC. Directional SIC refer to orientation relations of entities. These relations are based on the definition of a vector space. They are invariant under translation and uniform scaling. Generally two groups of directional relations can be distinguished regarding their requirement of a fixed reference system.

Cardinal directional relations are very often used for verbal descriptions when men explain the relative position of entities in geographic space or when they reason about these entities. Anyhow, IC which specify cardinal directional relations are hard to find, because usually these relations do not restrict the occurrence or the characteristics of entities. For example there is hardly any entity type which has to be north of another one (except the North Pole). More relevant might be relations in the 3D space like above and below, which also refer to a reference system.

The second group of directional relations deals with the order of entities in space independently of a fixed reference system. The relations are typically used for entities with a well defined front or back region or a forward / backward orientation. [18] defines a set of relations based on a left/right and front/back dichotomy for oriented entities. Entities like houses, which usually have an intrinsic front side, segment the space into a front and back semi-plane. A resulting SIC could be:

‘The backyard must be in the back of a house.’

Shape SIC. Shape SIC restrict the geometry of an entity in terms of form, shape and stature. Since they ensure semantic integrity, their restrictions result from the entity’s semantics, i.e. the concepts of the entity classes which are represented in the data model.

Shape is a concept which is difficult to describe qualitatively [17] and for many entity classes there is no general valid shape property definable. Because of this, shape is rarely used for the definition of SIC but nevertheless shape constraints seem to be convenient, in particular for 3D data (e.g. roof types of houses).

Metric SIC. Metric properties involve geographic distances between geospatial entities or their constituent parts. They change with scale but stay invariant under rotation and translation. Corresponding metric SIC restrict distances and size.

Some quantitative metric properties like length, area size and radius are based on operations on the entity’s geometry. Some of them are considered in international standards as defined methods of particular geometry types. For example the ISO [13] defines for Curve and LineString geometries a *Length* method and for Surface geometries an *Area* method, which return the length of the curve respectively the area of the surface. Such methods can be used to define constraints that restrict things like the minimum length of a linear entity or the maximum size of an area. The ISO norm also provides the spatial analysis operators *Distance* or *Buffer*, which are also expedient for metric SIC. An example of a metric SIC, which results from a national law, is:

‘A petrol station must be at least 300m away from a school.’

3.3.4. Complex SIC. Many SIC combine relations of more than one of the semantic domains used in the categorisation above. Thus we introduce complex SIC, which restrict relations or properties of several semantic domains in one SIC. An example of a complex constraint is:

‘A butterfly valve must not intersect a pipe if the diameter of the pipe is greater than 40cm’ [4]

This constraint describes a situation by a combination of the topological relation *intersect* between butterfly valve and pipe and the thematic relation *greater_than* of the pipe attribute diameter. The thematic relation is used to define a subset of pipe entities for which the constraint is valid.

Beside the relations of the elucidated semantic domains complex SIC can include mereological relations and domain specific associations that are defined in the data model. Mereology is dealing with relations between parts and their respective wholes. For spatial data the concepts of mereology and topology are usually not completely independent; the interactions between the two notions have been investigated in [19]. The correspondence of spatial aggregation and the consistency of the thematic attributes of the sub-entities has been pointed out by [20]. As an example constraint they define:

‘The number of inhabitants of a country is the sum of the numbers of the inhabitants of its administrative districts.’

In this example administrative district entities are related to the country entity through a *partOf* relation. It also shows the strong connection between mereology and topology, since the districts are usually topologically contained by the country.

3.3.5. Change and Representation SIC. According to [21] changes refer to operations performed on an object or a group of objects. Thereby the changes can either preserve the object identity or result in a change or a deletion of identity. Change SIC restrict these modifications of the entities and their properties, i.e. the relations between the two versions or states of the entities.

In general there are two types of change SIC to consider. First, there are transition constraints, which are restrictions on multiple states of a database [9]. Transition constraints are the only type of IC that can only be checked during a database transaction. All other constraints can be proven independently of a transaction. Many GIS applications integrate temporal changes in their model and thus store multiple states of an entity in a single database state. That’s why we include constraints, which restrict the changes between multiple consecutive versions of a single entity or group of entities of one database state as a second type of change SIC.

Possible relations that represent the change operation in SIC can be found in [21]. They classified operators for change actions of single entities (e.g. create, deconstruct), aggregates of objects (e.g. combine, compound) as well as their attributes and relations (e.g. add, remove).

An example of a change SIC could apply on a numerical attribute, allowing the attribute value only to increase in case of change. The corresponding checking algorithm would have to compare the two versions of the entity or the attribute in the two database states.

Some geodatabases contain multiple representations of a property, for example when an entity has a separate geometry stored for each level of detail. As shown by [22] the consistency of those representations can also be evaluated by SIC. They described a framework to assess the topological consistency of multiple representations. An example is:

‘The number of represented islands of a lake (e.g. holes of the lakes geometry) is not allowed to increase when the geometry is generalised.’

Relations that represent the connection between multiple representations in IC can be *generalisation*, *detailing* or for two specific levels of detail something like *LOD1toLOD2*. In GIS the change of representation mostly results in a generalisation of the entity’s geometry, but thematic attributes can also be involved. The differences between two representations are generally comparable to differences resulting from change operations. Therefore we define a combined category for change and representation IC.

The restricted differences of change and representation constraints might refer to any relation or property of the former mentioned semantic domains. For example [22] restrict the changes of the internal topological properties of an entity and of topological relations between entities.

4. Conclusion

The proposed categorisation provides a basis for further work on IC definition, formalisation, management and validation of internal consistency of IC sets. We consider it as a starting point for further discussion and refinement. We want to point out, that the assignment of an IC to one of the sub-classes of the categorisation is not always unambiguous. This is due to influences and overlaps between the domains of the constraint classes. In such cases we suggest to assign the constraints to the more abstract domain concept.

At present the potential of the constraints is by far not exploited and they are hardly supported by the available GIS or spatial database systems [6]. But in an environment of distributed and interoperably accessible geodatabases we estimate formalised and machine interpretable descriptions of integrity constraints as one of the great demands in GI Science.

5. References

[1] Servigne, S., Ubeda, T., Puricelli, A., Laurini, R., “A Methodology for Spatial Consistency Improvement of Geographic Databases”, GeoInformatica, 4, 2000, p. 7 – 34.

[2] Mäs, S., Reinhardt, W., Kandawasvika, A., Wang, F., “Concepts for quality assurance during mobile online data acquisition”, 8th AGILE Conference on Geographic Information Science - Proceedings. pp 3-12, Estoril, 2005

[3] Wang, F., Reinhardt, W., “Extending Geographic Data Modeling by Adopting Constraint Decision Table to Specify Spatial Integrity Constraints”, AGILE 2007, Lecture Notes in Geoinformation and Cartography.

[4] Cockcroft, S., “The Design and Implementation of a Repository for the Management of Spatial Data Integrity Constraints”, GeoInformatica, 8, 2004, pp. 49 – 69.

[5] Mäs, S., Wang, F., Reinhardt, W., “Using Ontologies for Integrity Constraint Definition”, ISSDQ 2005 Proceedings, 2005, pp 304-313.

[6] Louwsma, J., Zlatanova, S., Lammeren, R.v., Oosterom, P.v., “Specifications and implementations of constraints in GIS- with Examples form a Geo-Virtual Reality System”, GeoInformatica, 10, 2006, 531-550.

[7] Mäs, S., “Reasoning on Spatial Semantic Integrity Constraints”, COSIT 2007, LNCS 4736, 2007, pp. 285–302.

[8] ISO 19113, Geographic Information – Quality Principles

[9] Elmasri, R., Navathe, S. B., “Fundamentals of database systems”, 2nd Edition (Addison-Wesley), 1994, pp. 638-643.

[10] Friis-Christensen, A., Tryfona, N., Jensen, C.S., “Requirements and Research Issues in Geographic Data Modeling”, 9th ACM int. Symp. on advances in GIS, 2001.

[11] Frank, A.U., “Tiers of ontology and consistency constraints in geographical information systems”, Int. Journal of Geographical Inf. Science, 15, 2001, pp. 667 - 678.

[12] Mäs, S., “Reasoning on Spatial Relations between Entity Classes”, GIScience2008, LNCS 5266, 2008, 234-248

[13] ISO 19107, Geographic Information - Spatial Schema.

[14] ISO 19108, Geographic Information - Temporal Schema

[15] OGC, Filter Encoding Implementation Spec., Version: 1.1.0, OGC 04-095, 3 May 2005.

[16] Allen, J. F., “Maintaining Knowledge about Temporal Intervals”, Communications of the ACM, 26, 1983.

[17] Cohn, A.G., Hazarika, S.M., “Qualitative spatial representation and reasoning: an overview.” Fundamenta Informaticae, 46, 2001, pp. 1 – 29.

[18] Freksa, C., “Using Orientation Information for Qualitative Spatial Reasoning”, LNCS 639, 1992, 162-178.

[19] Varzi, A.C., “On the Boundary between Mereology and Topology”, Philosophy and the Cognitive Sciences, Schriftreihe der Wittgenstein-Gesellschaft, 21, 1994.

[20] Plümer, L., Gröger, G., “Achieving Integrity in Geographic Information Systems - Maps and Nested Maps.” GeoInformatica, 1, 1997, pp. 345 – 367.

[21] Hornsby, K., Egenhofer, M.J., “Qualitative Representation of Change” COSIT 1997, LNCS 1329, pp. 15 - 33.

[22] Egenhofer, M., Clementini, E., Felice, P.D., “Evaluating Inconsistencies Among Multiple Representations”, Sixth Int. Symp. on Spatial Data Handling, 1994, pp. 901-920.