OCL Specification of Inter-Register Integrity Constraints in Land Administration Systems

Miroslav Stefanović, Đorđe Pržulj, Darko Stefanović, Miloš Vukmanović, Sonja Ristić
University of Novi Sad
Faculty of Technical Sciences
Trg Dositeja Obradovića 6, Novi Sad, Serbia
{mstef, przulj, darkoste, vukmanovic.milos, sdristic}@uns.ac.rs

***Abstract**. A great number of public services rely on data stored in land administration systems (LASs) and large investments are made to collect, record and analyze these data. Therefore, the data quality in LASs is of utmost importance. Between several data quality elements we focus on data consistency and correctness. Data consistency requirements are described by means of integrity constraints (IC). Selecting a suitable IC specification language (ICSL) to specify ICs is not straightforward. In the paper Object Constraint Language (OCL) is used as ICSL. LAS data informally can be divided in two subsets: legal data and spatial data. For each of them data consistency and correctness issues would be addressed. Besides, the so-called inter-register inconsistencies may arise between the legal and spatial data subsets. In the paper, we identify them and specify corresponding ICs as an integral part of a domain model of LAS. That way, the inter-register inconsistencies may be detected. They could be consequences of incorrectnesses in legal data or in spatial data. Therefore, presented approach enables detection of possible data correctness issues in LAS data subsets.*

Keywords. land administration systems, data consistency, data correctness, domain model, object constraint language

1 Introduction

Land represents a limited resource, and good stewardship of the land is essential. Land administration system (LAS) is a formal system to identify and locate real property and to register its ownership, value and use. LASs are frequently directed at protecting the interests of landowners, but they are also instruments of national land policy and mechanisms to support stability in society and economic development. They describe real property not only with thematic attributes (legal status, value, tax data) but also with physical, spatial or topographic ones (location, dimensions, area). LAS data informally can be divided in two subsets: data about the ownership and other rights on real properties (legal data) and data about position and shape of real properties (spatial data). To support recording of such a diversity of data a LAS contains two registers: land register and cadastre. These registers complement each other.

Land register is an official record of rights on land or of deeds concerning changes in the legal situation of defined units of land. It gives an answer to the questions who owns certain property and what legal document is that ownership based on (Henssen & Williamson, 1990).

Cadastre represents a public inventory of data regarding properties within a certain country or district. Data in cadastre is based on a survey of property's boundaries. It represents register of spatial data used for describing properties and answers questions about location of certain property (Henssen & Williamson, 1990). Data that may appear in a cadastre include: geometric data (coordinates, maps), property addresses, land use, real property information, the nature and duration of the tenure, details about the construction of buildings and apartments, population, and land taxation values.

There are differences between the LASs in different countries. Land registration system may be based on deeds or titles. LAS may be realized as dual system with separate land book (title holders on real property) and land cadastre (real property) or as unified system with one register (in some countries it is called Real estate cadastre). In both cases, problem of integration of land register and cadastral data remains an important issue due to the fact that these two subsystems are poorly coupled. Also, both subsystems have their own business rules and transactions.

Different countries interpret the term "cadastre" in different ways and this can lead to some confusion. In this paper, we use previously given definitions of land register and cadastre.

The information obtained from a LAS rest upon data that represent parts of the real world. The reliability of the information and the reliability of decisions people make based on that information depends on the quality of these data. There are several elements of data quality and in the paper, we focus on logical data consistency and data correctness.

Integrity constraints (ICs) are formal statements, definitions or qualifications for describing data consistency requirements. Some of ICs can be inherently or implicitly expressed in a data model. Other ICs are usually hardcoded in the application software for the assurance of logical consistency. This separation from the data model raises problems of ICs management, adaptability and reuse in different applications. A formal specification of such ICs would overcome these problems. The approaches for modeling and enforcing integrity constraints are diverse as well as integrity constraint specification languages (ICSL). In the paper Object Constraint Language (OCL) is used as ICSL. ISO Technical Specification (ISO/TS) 19103:2015 (ISO 19103: Geographic information — Conceptual schema language, 2015) defines Unified Modeling Language (UML) in combination with OCL as conceptual schema language for specification of geographic information. Over the years UML established a position as standard for model-centric software design and development (Gogolla, Martin, & Richters,, 2007). OCL provides a framework to define constraints on UML class diagrams. As a specification language OCL is independent from its actual implementation and is universal concerning the time when constraints are enforced. Nonetheless, the tools exist that generate code in different languages (SQL, C#, Java) from OCL specifications. That enables the generation of integrity checking mechanisms automatically and eases the process of ICs implementation in a great extent.

Data are inconsistent if they violate implemented integrity constraints. Data correctness is the question of data correspondence with actual situation (legal e.g.) in reality. Detailed description of data consistency and data correctness in the context of LASs is given in Section 3. In a LAS data are recorded in two registers. Even if the registers are unified there are two parts of the register aimed to record legal and spatial data, respectively. For each of them data consistency and data correctness issues would be addressed. In addition, inconsistencies may arise between the legal data and spatial data. Such kind of inconsistencies and corresponding ICs in the paper are called inter-register inconsistencies and inter-register ICs. If integrity constraints are implemented within a data model, recorded data must not violate them. In reality, it is possible to record incomplete or incorrect data. The enforcement time of OCL constraints could be adjusted to the requirements of real world situation. Therefore, they can be enforced immediately, deferred or conditionally disabled or deferred. In the paper, we identify possible inter-register inconsistencies between legal and spatial data. OCL is used to specify inter-register ICs as an integral part of a domain model of LAS. Presented OCL ICs are specified in the context of the domain model that is result of previous research project partially reported in (Pržulj et al., 2017). The approach to modeling and enforcement of inter-register ICs presented in this paper enables detection of possible data correctness issues in two LAS data subsets. Their automated detection is very important since they may be consequences of incorrectnesses in legal data, in spatial data or in both of them. Elimination of those incorrectnesses is the first step in data quality improvement in the context of land administration.

Apart from Introduction and Conclusion, this paper is organized as follows. In Section 2 a short review of related work on subject of land administration is given. Issues of data consistency and data correctness in LASs and most common related problems are discussed in Section 3. Methodology used in this research is presented in Section 4. The extension of land administration domain model that includes OCL specification of ICs is given and discussed in Section 5.

2 Related Work

In 1998, FIG (Fédération Internationale des Géomètres) has published Cadastre 2014 after extensive research. It is based on Object–Right–Subject approach proposed in (Henssen & Williamson, 1990). Results of research are systematized in statements about cadastre for year 2014. Some of them recommend that LAS should be able to present data for complete legal situation, including public rights and restrictions on real property. There is a suggestion to abolish separation between maps and registers (Kaufmann & Steudler, 1998).

For purpose of developing information systems for land administration, International Organization for Standardization (ISO) published ISO 19152:2012 Geographic information – Land Administration Domain Model (LADM). As stated in (ISO 19152: Geographic information — Land Administration Domain Model (LADM), 2012) this model is not a complete solution for any given country, instead it should be used as foundation, permitting extensions to facilitate country's special needs.

Over the years papers proposing LADM country profiles have been published for different countries, such as (Bydlosz, 2011) for Poland, and (Govedarica et al., 2011) for Republic of Srpska. In (Vučić, Markovinović, & Mičević, 2013) data are presented showing that in Croatia only 5% of data are harmonized between two land administration registers thus clearly stating a need for creation of LADM for Croatia.

In 2001 the European Commission initiated the Infrastructure for Spatial Information in Europe (INSPIRE) that became effective in 2007 (D2.8.I.6 INSPIRE data specification on cadastral parcels – Guidelines, 2010). One of the main INSPIRE principles is the provision of access to relevant, harmonized and quality geographic information. Open

Geospatial Consortium (OGC) has paid attention to data quality by establishing a new Data Quality Working Group (DQWG) in 2007. Use of model-driven architecture in creating LAS in accordance with (ISO 19152: Geographic information — Land Administration Domain Model (LADM), 2012) and (D2.8.I.6 INSPIRE data specification on cadastral parcels – Guidelines, 2010) for Greece is discussed in (Psomadaki, Dimopoulou, & van Oosterom, 2016).

The ISO defines six data quality elements (ISO 19157: Geographic information — Data Quality, 2013). In this paper we focus on logical consistency and integrity constraint specification.

Approaches to categorizing ICs for LASs combine general database constraint types and specific constraint types for spatial and temporal databases.

(Elmasri & Navathe, 2011) divide constraints on databases into three main categories: i) constraints that are inherent in the data model – implicit constraints; ii) constraints that can be expressed in database schemas of the data model by means of data definition language (DDL) – explicit constraints; and iii) constraints that must be expressed end enforced procedurally by the application program – semantic constraints.

Approaches to classify spatial ICs can be found in (Cockcroft, A taxonomy of spatial data integrity constraints, 1997), (Cockcroft, The design and implementation of a repository for the management of spatial data integrity constraints, 2004), (Servigne et al., 2000) and (Borges, Laender, & Clodoveu Jr., 1999). Formal classification of integrity constraints in spatiotemporal database applications is given in (Salehi et al., 2011) and (Rodríguez, 2005).

A formal description of ICs is very important, especially for ICs which cannot be inherently or implicitly expressed in the data model. It enables the transfer of the integrity rules between different systems and system components. Approaches aiming at ICs' formalization can be based on different specification languages. An extensive overview of existing approaches to the formalization of ICs is given by Werder in (Werder, 2009). A general comparison of specification languages, which also includes natural languages, can be found in (Salehi et al., 2007).

Here we mention just some of them. An integration of ICs in a data modelling notation called OMT-G is presented in (Borges, Laender, & Clodoveu Jr., 1999). The paper (Bravo & Rodríguez, 2009) introduces a formalization of a set of spatial semantic integrity constraints on an extended-relational database model. The formalization extends traditional notions of functional and inclusion dependencies by adding interaction with spatial attributes. UML and OCL with corresponding spatial extensions as ICSL is used in several cases. In (Werder, 2009) GeoOCL is analyzed, and in (Bejaoui et al., 2010) SpatialOCL is proposed as ICSL.

Although consistency is a desirable property of LAS data, enforcing integrity constraints might not be always feasible. In the paper (Brisaboa et al., 2015) an importance of data-cleaning tools to detect and remove, if possible, inconsistencies in large datasets is emphasized and some solutions are presented.

Problem of maintaining consistency between spatial and non-spatial data, is pointed out in (Cockcroft, A taxonomy of spatial data integrity constraints, 1997). The same problem in the context of land register and cadastre data is addressed in (Pržulj et al., 2017) were terms geoaggregate and polygon are introduced, representing parcel and part of parcel on cadastral maps respectively. These terms are discussed in more detail in Section 5.

Even with existing standards major problem for different LAS is how to sustain correctness of data stored in those systems. In (Bittner & Frank, 2002) correctness is recognized as major issue of a land register. In the same paper, issue of differences between incorrectness and inconsistency are discussed.

The aforementioned papers in this section deal with 2D constraints. Formalization of 3D spatial ICs opens a number of new issues as it is pointed out in (Xu, van Oosterom, & Zlatanova, 2017). Specification of inter-register 3D constraints in LAS and inter-register inconsistencies in the presence of 3D constraints will require further research.

3 Consistency and correctness of LAS data

In this paper, we address two elements of data quality in LAS: data correctness and data consistency.

Data are said to be incorrect if they are not in accordance with reality. Let's suppose that in a database that holds cadastral data -100 is stored as the area of a land parcel. It is for sure incorrect because there is no land parcel with negative area. On the other side, one could misspell the area and insert 102 instead of 120 as the area of a land parcel. It is possible for a land parcel to have such an area, so area data 102 conforms to any rule in relation with allowed values for land parcel's area. The value, however, is not correct because this value is not in accordance with reality. Besides, the incorrect value may not be misspelled, but it could be incorrectly measured due to the human error or due to the measurement tool error.

In order to improve the data quality, integrity constraints have to be specified. They represent the rules to which data should adhere. Data are said to be inconsistent if they do not satisfy specified integrity constraints. At the same time, inconsistent data are by definition incorrect. If an IC specifies that the attribute area has to be positive rational number then area value –100 is inconsistent and incorrect, too. Although it is important that the ICs are specified, they won't solve the entire problem of incorrect data. It is impossible to specify the IC that detects misspelled or incorrectly measured area value that is represented with positive rational number. It is obvious that the set of inconsistent data is the proper subset of the incorrect data set. In eq. 1 let *IDBS* (incorrect data in a broader sense) denotes the set of all incorrect data of a LAS, and *ICD* (inconsistent data) denotes the set of all inconsistent data of the LAS. Let

$$IDNS = IDBS \setminus ICD \quad (1)$$

where *IDNS* stands for incorrect data in the narrow sense. In this paper we will use the term "incorrect data" to address only the data from *IDNS* set. They are not in accordance with reality and it is not possible to detect that discordance by means of ICs that are specified within only one register (land register or cadastre). Data consistency according to the ICs specified within one register in the paper is called internal consistency.

Bittner and Frank in (Bittner & Frank, 2002) define inconsistency as internal property that could be consequence of human errors or violation of property rules, while on the other hand, incorrectness is defined by situation in which information stored in cadastre does not match legal situation in reality.

Example of inconsistency of land register could be a case in which parties are registered as owners of property and sum of their shares is not equal to 1. Incorrectness of land register could be represented with an example when certain property is sold to a new owner and that transfer of ownership is still not recorded in land register.

Cadastral data can also be checked in terms of inconsistency and incorrectness as it was already mentioned for land register. In case of spatial data, correctness of geometry and topology should be checked (Vranić & Matijević, 2015). Inconsistency could be represented by an example of invalid topological relations. For example, if building is positioned outside the boundaries of parcel it belongs to. This is shown in Fig. 1 where annotation 275 represents parcel label and 1 is sequence number of building on the same parcel. On the other hand, incorrectness could be a case when, for example, shape of certain parcel, does not match the real shape measured "in field".

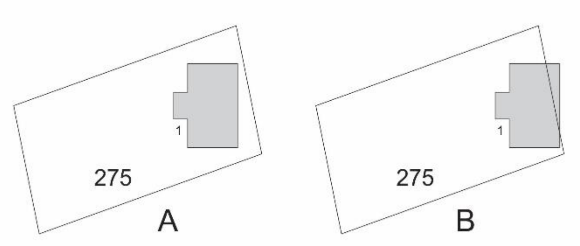

Figure 1. Cadastral data: A consistent - B inconsistent

Correctness of both registers is a problem whose origins could be traced to the fact that they are based on old "pen and paper" systems. Although these two parts of LAS should have been consistent with each other, with "pen and paper" system it was an impossible task. Also, process of digitalization was often done based on analog cadastral maps that were not of adequate quality or were of unknown provenance (Congalton & Macleod, 1994).

LAS can also be checked in terms of inconsistency and incorrectness. Inconsistency of LAS represents a case when data about a property in land register and data about the same property in cadastre do not match. For example, in land register there are data showing that on certain parcel there should be two buildings, while in cadastre there are spatial data representing only one building.

As for incorrectness, it could be said that LAS is in incorrect state when it is not representing true legal situation. This could happen only in case when data in at least one of two subsystems, land register or cadastre, are incorrect. It is not possible for LAS to be incorrect, while both land register and cadastre data are in correct state.

In general, inter-register inconsistency would always signal that there is an incorrectness in a broader sense in a system. In the case when both of LAS subsystems are internally consistent, the existence of an inter-register inconsistency in LAS is a signal that there is an incorrectness (in the narrow sense) in at least one of two subsystems.

For example, a parcel is recorded in land register as a parcel covering area of 1000m2. That data is consistent with all other data stored in land register. At the same time, in cadastre, the same parcel is registered with area of 900m2 and this data is consistent with all other data that are stored in cadastre. Unfortunately, a consistency problem exists between these two registers. Namely, the parcel covering area recorded in land register differs from the covering area of the same parcel that is recorded in cadastre. This is an example of inter-register inconsistency and, as mentioned before, it signals that there is an incorrectness in at least one of two subsystems.

There are three possible cases of incorrectness:

- land register data incorrectness – data about area are not correct,
- cadastral data incorrectness – data about area calculated from polygon shape that depicts parcel in cadastre are not correct,
- neither of data are correct.

Possible inter-register inconsistencies that could be detected in LAS and signal incorrectness in at least one subsystem are:

- for land register data about a specific parcel there are no matching data in cadastre,
- for cadastral data about parcel there are no matching land register data,
- land register data about parcel's area do not match data about matching parcel's area in cadastre,
- for land register data about part of parcel there are no matching data in cadastre,
- for cadastral data about part of parcel there are no matching land register data,

- land register data about part of parcel's area do not match data about matching part of a parcel area in cadastre,
- land register data about part of parcel's land use do not match data about matching part of a parcel's land use in cadastre.

4 Specification of integrity constraints

There is a large number of different approaches for modeling and enforcing ICs as well as for selecting most suitable ICSL. Main factors in process of selecting ICSL for specifying ICs are: understandability, readability, usability, and expressivity of the ICSL. In case of LAS, ISO Technical Specification (ISO/TS) 19103:2015 defines Unified Modeling Language (UML) in combination with Object Constraint Language (OCL) as conceptual schema language for specification of geographic information (ISO 19103: Geographic information — Conceptual schema language, 2015).

UML is a graphical language and in latest version (in time of writing this paper it was version 2.5 published from June 2015), 15 different types of diagrams were defined, categorized in two major groups: structural and behaviour diagrams (OMG Unified Modeling Language, version 2.5, 2015).

Over the years UML established a position as standard for model-centric software production (Gogolla, Martin, & Richters,, 2007). Yet, it is not possible to model all of the necessary restrictions simply by using class diagram, that is most widely used UML diagram. Therefore, an ICSL is needed to complement UML (Richters & Gogolla, 1998) (Dobing & Parsons, 2006).

Based on the ideas of Syntropy method, International Business Machines Corporation (IBM) developed Object Constraint Language (OCL) in 1995. OCL development was influenced by previously issued request for proposal (RFP), by Object Management Group (OMG), for standard object-oriented analysis and design language for UML. IBM led submission to OMG, that included OCL, was adopted in 1997, making OCL a part of UML 1.1 (Cabot & Gogolla, 2012).

OCL established a position as de facto standard modelling language for specifying system constraints that otherwise would not be possible to specify. Nevertheless because of difficult syntax, general usability of OCL is still low and this results in extended time and expense spent on design phase of software development (Salemi, Selamat, & Penhaker, 2016). It is common for model domain to be represented with class diagram, while constraints are defined using natural language instead of OCL. Needless to say that this is a big drawback to whole idea of dedicating less time for understanding subject model.

OCL is a platform-independent and generic method for modeling constraints and its main advantages are its declarative expression of constraints, connection to UML and possibility of interpretation by code engines, so generation of integrity checking mechanisms can be done automatically (Bejaoui at al., 2010).

Although OCL was initially used just as an ICSL, it soon became an important part of model-driven engineering (MDE) and nowadays it is used for different kind of model queries, manipulations, specification requirements, model transformations, well-formed rules or code generation templates (Cabot & Gogolla, 2012).

The OCL complements class diagrams since it is possible to write constraints that will "navigate" along associations to describe conditions on object states in class invariants and pre- and post-conditions of operations (Baumeister et al., 2001). Class invariants represent constraints connected to a class defining logical properties that features of all objects should satisfy. Operation preconditions represents constraints connected to operations defining what conditions should be true before execution of operation so operation in question
would execute correctly. Operation postconditions represents constraints connected to operations defining what conditions should be true after operation is executed correctly (Lano, 2009).

The OCL can track its roots to set theory, predicate logic and operational semantics. OCL supports set theoretic concepts cardinality (size), comprehension (select) and projection (collect) as well as algebra operation (union, intersection, etc.). Computational aspects are expressed using iterate construct thus providing possibilities that go beyond set theory and predicate logic, for example a possibility to sum up all elements in a specific set of integers (Baar, 2000).

5 Domain Model with inter-register constraints

In domain model proposed in (Pržulj et al., 2017), for representing spatial data, classes Geoaggregate and Polygon are used. Geoaggregate corresponds to GM_Aggregate class from OGC spatial scheme adopted in ISO 19107 (ISO 19107: Geographic information – Spatial schema, 2003). Geoaggregate represents a collection of primitive geometric shapes: points, line strings and polygons. Geoaggregate instance represents one parcel (with all its details) from land register on cadastral map. Polygon is planar surface that is topologically closed and has one external and zero or more internal boundaries (ISO 19125: OpenGIS Implementation Specification for Geographic information, 2005). In proposed domain model polygon representing part of parcel does not have explicit inner boundaries. Instead, geoaggregate referencing its polygons cares about their nesting. This way inner boundaries (if any) are detected in every polygon. Polygon instance represents spatial data about part of parcel on cadastral map. Class Theme is used for representing land uses in cadastre that in practice differ from land uses in land register.

With introduction of these new terms inter-register inconsistencies that could be detected in LAS, listed in Section 3, can be redefined as follows:

- for parcel, there is no matching geoaggregate,
- for geoaggregate, there is no matching parcel,
- parcel area does not match corresponding geoaggregate area,
- for part of parcel, there is no matching polygon,
- for polygon, there is no matching part of parcel,
- part of parcel area does not match corresponding polygon area,
- part of parcel land use does not match polygon theme.

Proposed class diagram with constraints declared using OCL is show in Fig. 2. For drawing class diagram Eclipse Java EE IDE for Web Developers, Neon Release (4.6.0) with Papyrus UML Modeller 2.0.1 was used.

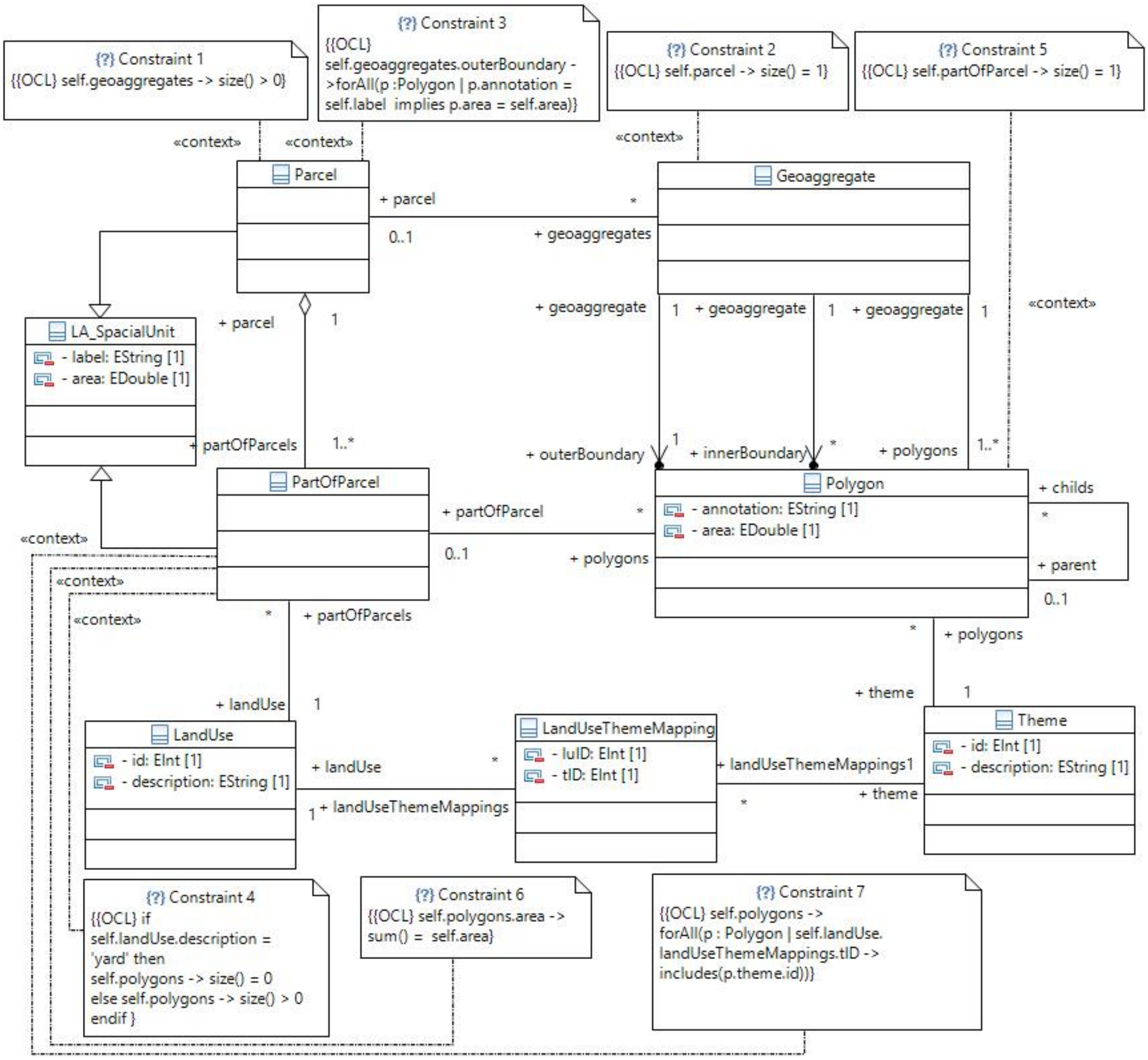

Figure 2. Class diagram for preserving data consistency in LAS

In proposed domain model classes Parcel and PartOfParcel represent abstractions of parcel and part of parcel from land register and they are both specialization of one of basic classes of LADM, class LA_SpatialUnit. Between classes Parcel and PartOfParcel aggregation is defined, modeling whole-part relation.

Class Parcel inherits attribute area from class in LA_SpatialUnit and it is used to store data about area and the same can be said for attribute area in PartOfParcel class.For every part of parcel, it is necessary to declare its land use based on class LandUse and that is why cardinality between PartOfParcel class and LandUse class is set to one. Between class Geoaggregate and class Polygon three associations are defined that represent outer and inner boundaries of parcel as well as other polygons that can appear within one geoaggregate. Following the same analogy as between classes Parcel and PartOfParcel, every geoaggregate has one polygon that covers its entire area and that represents its outer boundaries. In cadastral systems, it is possible that one parcel is nested within another, so geoaggregate besides mandatory polygon that represents outter boundary can have polygons that represent inner boundaries if they exist. This is modeled with association innerBoundary between Geoaggregate and Polygon classes.

Attribute area in class Polygon is used to store data about area. In case when polygon represents outer boundary of a parcel, value recorded in area attribute represents true area of parcel, where areas of all inner boundary polygons, if any, were deducted. It's important to notice that certain tolerance between data representing area in land register and cadastre is acceptable. This acceptable level of tolerance is defined by countries' legislation and it should depend on technologies and tools used for surveying purposes.

Analog to association between classes PartOfParcel and LandUse, association between classes Polygon and Theme is created.

Inter-register dependencies between land register and cadastre are created through associations between classes Parcel and Geoaggregate, classes PartOfParcel and Polygon and classes LandUse and Theme through class LandUseThemeMapping. Instance of LandUseThemeMapping class is pair that maps land use from land register with corresponding theme from cadastre map.

Attribute label, that class Parcel inherits from class LA_SpatialUnit, is used for matching parcel corresponding outer boundary polygon through its attribute annotation.

Cardinality between classes Parcel and Geoaggregate allows existence of parcel that has no representation in spatial data and also a possibility for one parcel to be represented by several geoaggregates. It is necessary for model to allow a possibility of having more than one geoaggregate associated with parcel when there are data collected from different sources of spatial data about the same property. One geoaggregate can be associated with at most one parcel, while it is possible that there is a geoaggregate that is not associated with any of parcels. Minimal cardinality set to zero in both cases is necessary to allow storing of data from various sources that are often in inconsistent state.

Cardinality between classes PartOfParcel and Polygon allows possibility that part of parcel will not have a representation in spatial data and this is intended to support existence of part of parcel added to land register as a taxation deductible ('yard'). If there are multiple parts of parcel with the same land use they are represented by one part of parcel record in land register, but by multiple polygons on cadastral map.

It is evident that some constraints could not be implemented using just UML class diagram, so those constraints are defined using OCL as follows:

- Constraint 1 – It is necessary that every parcel has matching geoaggregates.
  ```
  context Parcel:
  self.geoaggregates > size()>0
  ```
- Constraint 2 – It is necessary that every geoaggregate has a matching parcel.
  ```
  context GeoAggregate:
  self.parcel > size()=1
  ```
- Constraint 3 – It is necessary that every parcel and matching geoaggregate have the same area.
  ```
  context Parcel:
  self.geoaggregates.outerBoundary
  forAll(p:Polygon|p.annotation =
  self.label implies p.area =
  self.area)
  ```
- Constraint 4 – It is necessary for every part of parcel, unless its land uses description is 'yard', to have matching polygon(s).
  ```
  context PartOfParcel:
  if self.landUse.description =
  'yard'
  then self.polygons > size()=0
  else self.polygons > size() > 0
  endif
  ```
- Constraint 5 – It is necessary that every polygon has matching part of parcel.
  ```
  context Polygon:
  self.partOfParcel > size()=1
  ```
- Constraint 6 – It is necessary that every part of parcel has the same area as a sum of areas of matching polygons.
  ```
  context PartOfParcel:
  self.polygons.area >
  sum() = self.area
  ```
- Constraint 7 – It is necessary that every part of parcel and matching polygon(s) have appropriate land use/theme mapping.
  ```
  context PartOfParcel:
  self.polygons >
  forAll(p:Polygon|self.landUse.
  landUseThemeMappings.tID >
  includes(p.theme.id))
  ```

6 Conclusion

Process of establishing LAS has to support storing of all available data that have previously been held in "pen and paper" systems. Without possibility of implementing sustainable IC validation in "pen and paper" systems, it is to be expected that newly established LAS will contain data that are inconsistent and therefore incorrect.

Presented solution proposes ICs that should be implemented and used to detect inter-register inconsistencies in LAS. Detected inter-register inconsistencies indicate incorrectnesses in at least one of LAS subsystems that should be addressed and corrected by responsible authorities. Following step towards LAS data consistency should be enforcement of inter-register ICs. LAS data that are brought to consistent state by detecting and correcting identified violations of inter-register ICs can be preserved in consistent state by enforcing those ICs.

Inter-registered ICs presented in this paper could be a starting point for future modeling of land register and cadastre internal consistency. Future research would be directed at identification of processes in LAS that could cause violation of inter-register constraints proposed in this paper.

Platform independent model of inter-register ICs expressed by means of OCL could be transformed into different platform specific models aimed at ICs' implementation. These transformations could be the subject of future research, as well as implementation and testing of proposed domain model on real world data.

References

Baar, T. (2000). Experiences with the UML/OCL-approach to precise software modeling: A report from practice. *Proc. Net. Object-Days.* Erfurt.

Baumeister, H., Hennicker, R., Knapp, A., & Wirsing, M. (2001). OCL component invariants. *Proc. Wsh. Montery-Engineering Automation for Software Intensive System Integration*, (pp. 208-215). Monterey.

Bejaoui, L., Pinet, F., Schneider, M., & Bedard, Y. (2010). OCL for formal modelling of topological constraints involving regions with broad boundaries. *GeoInformatica*, 353-378.

Bittner, S., & Frank, A. U. (2002). A formal model of correctness in a cadastre. *Computers, environment and urban systems*, 465-482.

Borges, K. V., Laender, A. F., & Clodoveu Jr., A. D. (1999). Spatial data integrity constraints in object oriented geographic data modeling.". *Proceedings of the 7th ACM international symposium on Advances in geographic information systems*, (pp. 1-6).

Bravo, L., & Rodríguez, M. A. (2009). Semantic integrity constraints for spatial databases.

Brisaboa, N. R., Rodríguez, M. A., Seco, D., & Troncoso, R. A. (2015). Rank-based strategies for cleaning inconsistent spatial databases. *International Journal of Geographical Information Science*, 280-304.

Bydlosz, J. (2011). Towards LADM country cadastral profile–case Poland. *Proceedings of the 5th Land Administration Domain Model Workshop*, (pp. 24-25). Kuala Lumpur.

Cabot, J., & Gogolla, M. (2012). Object constraint language (OCL): A definitive guide. In *Formal methods for model-driven engineering* (pp. 58-90). Berlin.

Cockcroft, S. (1997). A taxonomy of spatial data integrity constraints. *GeoInformatica*, 327-343.

Cockcroft, S. (2004). The design and implementation of a repository for the management of spatial data integrity constraints. *Geoinformatica*,, 49-64.

Congalton, R. G., & Macleod, R. (1994). Change detection accuracy assessment on the NOAA Chesapeake Bay pilot study. *Proceedings of the International Symposium on Spatial Accuracy of Natural Resource Data Bases: unlocking the puzzle*, (pp. 78-87). Williamsburg.

D2.8.I.6 INSPIRE data specification on cadastral parcels – Guidelines. (2010).

Dobing, B., & Parsons, J. (2006). How UML is used. *Communications of the ACM*, (pp. 109-113).

Elmasri, R., & Navathe, B. S. (2011). *Database systems: Models, languages, design and application programming.* Pearson Global Edition.

Gogolla, M., Martin, F., & Richters,, M. (2007). USE: A UMLbased. *Science of Computer Programming*, 27-34.

Govedarica, M., Ristić, A., Sladić, D., & Pržulj, Đ. (2011). LADM profile for Republic of Srpska. *Congress on cadastre in Bosnia and Herzegovina*, (pp. 1-11). Sarajevo.

Henssen, J. G., & Williamson, I. P. (1990). Land registration, cadastre and its interaction - a world perspective. *FIG XIX Congress*, (pp. 14-43). Helsinki.

ISO 19103: Geographic information — Conceptual schema language. (2015).

ISO 19107: Geographic information – Spatial schema. (2003).

ISO 19125: OpenGIS Implementation Specification for Geographic information. (2005).

ISO 19152: Geographic information — Land Administration Domain Model (LADM). (2012).

ISO 19157: Geographic information — Data Quality. (2013).

Kaufmann, J., & Steudler, D. (1998). Cadastre 2014: A vision for a future cadastral system. Bern: FIG.

Lano, K. (2009). *UML 2 semantics and applications.* John Wiley & Sons.

OMG Unified Modeling Language, version 2.5. (2015).

Pržulj, Đ., Radaković, N., Sladić, D., & Radulović, A. (2017). *Domain model for cadastral map with land use component.* Unpublished manuscript: University of Novi Sad, Novi Sad, Serbia.

Psomadaki, S., Dimopoulou, E., & van Oosterom, P. (2016). Model driven architecture engineered land administration in conformance with international standards–illustrated with the Hellenic Cadastre. *Open Geospatial Data, Software and Standards*, 3-14.

Richters, M., & Gogolla, M. (1998). On formalizing the UML object constraint language OCL. *International Conference on Conceptual Modeling*, (pp. 449-464). Berlin.

Rodríguez, A. (2005). Inconsistency issues in spatial databases. *Inconsistency tolerance*, 237-269.

Salehi, M., Bédard, Y., Mostafavi, M. A., & Brodeur, J. (2007). On languages for the specification of integrity constraints in spatial conceptual models. *Advances in conceptual modeling – Foundations and applications*, 388-397.

Salehi, M., Bédard, Y., Mostafavi, M. A., & Brodeur, J. (2011). Formal classification of integrity constraints in spatiotemporal database applications. *Journal of Visual Languages & Computing*, 323-339.

Salemi, S., Selamat, A., & Penhaker, M. (2016). A model transformation framework to increase OCL usability. *Journal of King Saud University-Computer and Information Sciences*, 13-26.

Servigne, S., Ubeda, T., Puricelli, A., & Laurini, R. (2000). A methodology for spatial consistency improvement of geographic databases. *Geoinformatica*, 7-34.

Vranić, S., & Matijević, H. (2015). Workflows for ensuring consistency of cadastral data. *FIG Working Week 2015.*

Vučić, M., Markovinović, D., & Mičević, B. (2013). LADM in the Republic of Croatia–making and testing country profile. *Proceedings*, (pp. 329-344). Kuala Lumpur,.

Werder, S. (2009). Formalization of spatial constraints. *In 12th AGILE International Conference on Geographic Information Science*, (pp. 13-26). Hannover.

Xu, D., van Oosterom, P., & Zlatanova, S. (2017). A methodology for modelling of 3D spatial constraints. *Advances in 3D geoinformation, Lecture notes in geoinformation and cartography*, (pp. 96-117).