Acta Geologica Polonica, Vol. 53 (2003), No. 3, pp. 247-255

Features and coverages in hydrogeological information

JANUSZ MICHALAK & PAWEŁ M. LEŚNIAK

Warsaw University, Dept. of Geology, Żwirki i Wigury 93, PL-02-089, Warszawa, Poland.
E-mail: jwm@geo.uw.edu.pl
Institute of Geological Science, Polish Academy of Science, Twarda 51/55, PL-00-818, Warszawa, Poland.
E-mail: lesniak@twarda.pan.pl

ABSTRACT:

MICHALAK, J. & LEŚNIAK, P.M. 2003. Features and coverages in hydrogeological information. *Acta Geologica Polonica*, **53** (3), 247-255. Warszawa.

In recording hydrogeological information in computer systems (databases, GIS), it became necessary to manipulate them in a consistent and efficient way. The correct functioning of these systems depends principally on the applied data models, which should be based on commonly accepted international standards of geospatial information. This is particularly important in the case of interoperability of different systems, for example a system simulating hydrogeological processes and the hydrogeological geospatial database system. The research tools of geomatics allow the construction of data models of such systems conformable to these standards. However, in the definition of such a model several ontological and semantic inconsistencies arise. Two examples are taken for consideration: hydrogeological *feature* and *coverage*. These terms are frequently applied improperly to cases of geospatial information. Definition of these terms in a hydrogeological context will prevent ambiguities in formulating models of hydrogeological data.

Key words: Geomatics, Geospatial information, Hydrogeological feature, Hydrogeological coverage, Hydrogeological object, Data model, Conceptual schema.

INTRODUCTION

The way in which geospatial information relating to hydrogeological processes is written in computer systems depends on the applied model of the data. In turn, the data model is dependent on the type of information system and the conceptual schema formulated for the designated application. From these reasons, correctness of the information encoding and effectiveness of functioning of such a system depends on both factors. The type of information system can be chosen from the large number of different possibilities and, in this case, rational criteria, based on a correctly defined conceptual schema, lead to the final results.

Because "first of all *conceptual schema* relates to human thinking, perception and human communication and then to computer logic" (SUBIETA 1999), the definition used in the particular schema requires analysis based on ontology and semantics relating to the particular field – here hydrogeology.

One of several problems is, for example, the definition of the groundwater table; this term is often mistakenly extended to an area beyond the borehole or well in which it was determined. The hydrogeological definition that it is „a surface separating the saturated zone from unsaturated one" (KLECZKOWSKI & RÓŻKOWSKI 1997), is not strictly correct from the physical point of view, and the more correct definition defines this as a surface on which pressure exerted on porewaters or fissure waters is equal to the atmospheric pressure. This is much closer to the definition of the *surface water table* or *water table in the well*. Taken as such, the notion of *groundwater table* is the only surface in the determined vertical profile, though in reality it can take complicated shapes as in a multiple aquifer or in an aquifer close to a steep cliff.

Although frequently applied, in practice the term *measurement of water table* is imprecise, because it relates to quite different phenomena, such as *hydraulic head* or *hydrodynamic potential* expressed by vertical water level. „For non-mineralized underground water, the hydraulic head is equivalent to the vertical coordinate value, if the reference level is sea level“ (KLECZKOWSKI & RÓŻKOWSKI 1997). This is true only when the vertical gradient of hydraulic potential is equal to zero, which is an exceptional situation. As a result, instead of water table measurements, we have in fact the measurement of hydraulic head, from which a map of the groundwater table is eventually constructed. In such maps, information concerning the spatial variability of the water table is shown most frequently as *isolines* of heads which, in geomatic sense, must be treated as *objects* (in so called "object-oriented maps").

Introduction to the conceptual schemas of hydrogeological information of the strictly defined geomatic terms as *feature* and *coverage* allows systematization of the notions existent in these schemas and avoidance of such ambiguities as those presented above.

BASIC TERMS AND DEFINITIONS

Determination of the relationship between the two contrasting terms *feature* and *coverage* requires application of the commonly accepted definitions of these terms. The definition presented below (with the exception of the object definition) are derived from terminological notions developed after long discussion by researchers actively participating in Open GIS Consortium and Committee ISO/TC211.

• *Geospatial information* – information, in the sense defined by computer science, but, unlike any other kind of information, it is related to a specific place (fragment of space). As a result, data specifying (in direct or indirect way) the position of this place with respect to the Earth is its essential component.

• *Geospatial feature* – the basic element (atom) of *geospatial information*. It has geospatial attributes (geometric and topological), such as shape, extent, position, relation to other features. The term *feature* is often confused with the term *object*; a feature can be an object but it does not have to be (MARK & *al*. 2001). Since in geomatics all features are geospatial, the adjective *geospatial* is usually omitted, and a shorter term – *feature* - is used.

• *Object* – A real or an abstract entity (instance), distinguishable in the modeled reality, which has a name, unambiguous identification, clearly defined boundaries, attributes and other properties, such as the type of internal structure or the structure of related data. These components (members) of the object characterize: its state (through values of attributes and associations) and its behavior (through methods, that is, operators and functions) (Subieta, 1998). In geomatics, it is assumed that an object is an instance of class, which is based upon the paradigm of object-orientation, derived from the UML (Unified Modeling Language) metamodel, which is adapted here for description of *conceptual models* (OMG 2001). In computer cartography, the definition of an object is different, close to the popular understanding of this word, for instance: military, sports, tourist object etc. (MICHALAK 2000). In separating an object from reality from an object occurring in an object computer system, the first one is called a *conceptual object* (abbreviation *object*) and the second a *programming object* (abbreviation *p-object*).

• *Conceptual schema* – formal description of a *conceptual model* in the form of a diagram of class, hierarchy and links (ISO 19103 2001, SUBIETA 1999).

• *Conceptual model* – model of processes or data structure relating to human perception and imagination, having as a purpose understanding of the problem, documentation of analyses or project in readable and abstract language form and facilitating information in human communities (SUBIETA 1999). Model that defines concepts of a universe of discourse (ISO 19103 2001).

• *Geospatial coverage* – two- (and sometimes higher-) dimensional metaphor for phenomena found on or near a portion of the Earth's surface. Fundamentally, coverages (and images) provide humans with an n-dimensional (where n is usually 2, and occasionally 3 or higher) "view" of some (usually more complex) space of geographic features. The power of coverages is in their ability to model and make visible spatial relationships between, and the spatial distribution of, Earth phenomena. A coverage is a special case of (or a subtype of) a feature. In the application schema, coverage is a function that maps a spatial domain into an attribute domain. Examples of coverage include a raster image, polygon overlay, or digital elevation matrix. Coverages are often implemented as data sets of attribute values associated to positions within a bounded space (The Open GIS 1999).

Several different definitions of these terms can be met in the literature as a result of terminology noise and these are not discussed here.

Development of the definition of *coverage* is contained in the conceptual models presented on Text-figs 2, 3 and 4. Text-fig. 5 presents a conceptual model of importance in the hydrogeology subtype coverage *grid*. These models were constructed and verified in UML language with use of Rational Rose Enterprise 2002 software. Graphical notation of UML language used in models is explained in Text-fig. 1.

RELATIONS AMONG OBJECT, FEATURE AND COVERAGE

Objects, coverages and features are the elements of conceptual models and application schemas in geoinformation systems. However, the differences between them possess a deeper ontological and semantic sense. Citation from (HERRING & KOTTMAN 1997): The answer is „Features and Coverages“ and the question was (...) „What are the two fundamentally different ways in which people think about and describe geographic information?“ Not everything that can be separated from our reality can be treated as an *object*, because not all criteria for such definition are met. However, definition of *feature* is so broad in this case that all this can be treated as a *feature*, but on condition

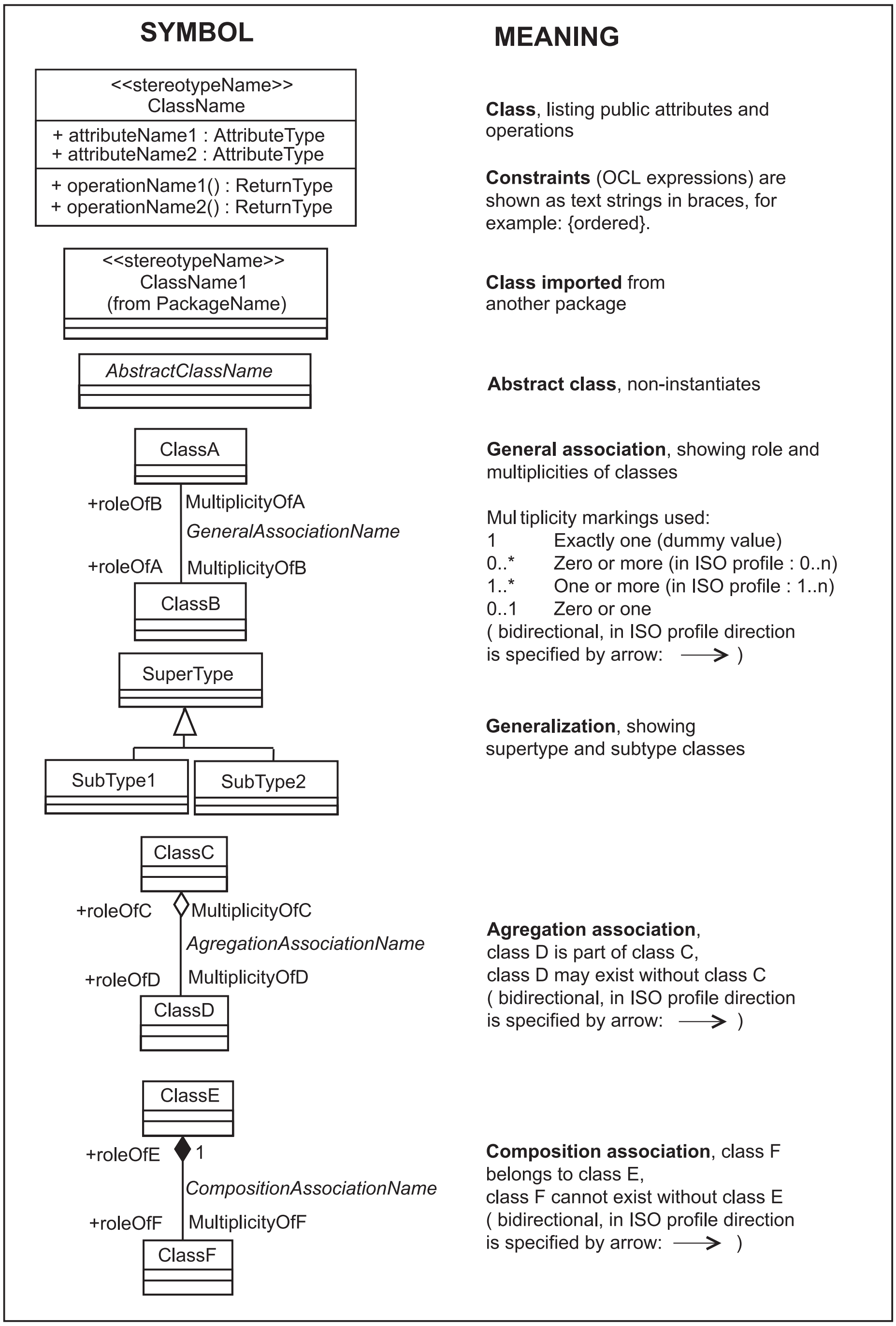

Fig. 1. Notation elements of UML class diagrams in OGC profile for geospatial data models (WHITESIDE 1999); Extended to ISO profile

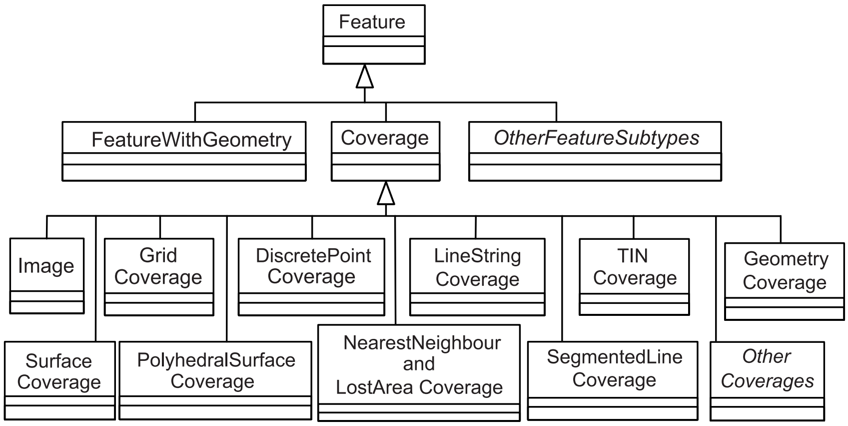

Fig. 2. Feature and Coverage subtypes hierarchy in OGC specification – a Coverage class is subtype (subclass) of Feature class. UML class diagram – logical view; Source: The OpenGIS 1999

that it has a specified extent in space and time (also, if it is an *object*). The typical examples of hydrogeological features are:

- Genuine feature such as: artesian basin, aquifer, seepage spring area.
- Fiat feature such as: area of balancing of underground water, main groundwater basin, aquifer in use.
- Fuzzy feature (genuine or fiat) such as: cone of depression, hydrogeological unit and drainage zone.
- Typical examples of hydrogeological objects are: well, hydrogeological borehole, piezometer and spring.

From the formal point of view, *coverage* is a *feature*, however, there is no example where *coverage* could be an *object* (or subtype of *object*). Text-fig. 2 presents the relationship between *coverage* and *feature* as well as their subtypes. In the geoinformation system, *coverage* can be represented by a *p-object*, however there is a very important difference between *is* and *is represented by*. Coverages, being a particular case of feature, allow conceptual modeling of real phenomena in a different way from the typical features, particularly those which are objects. The most specific property of coverage is that it

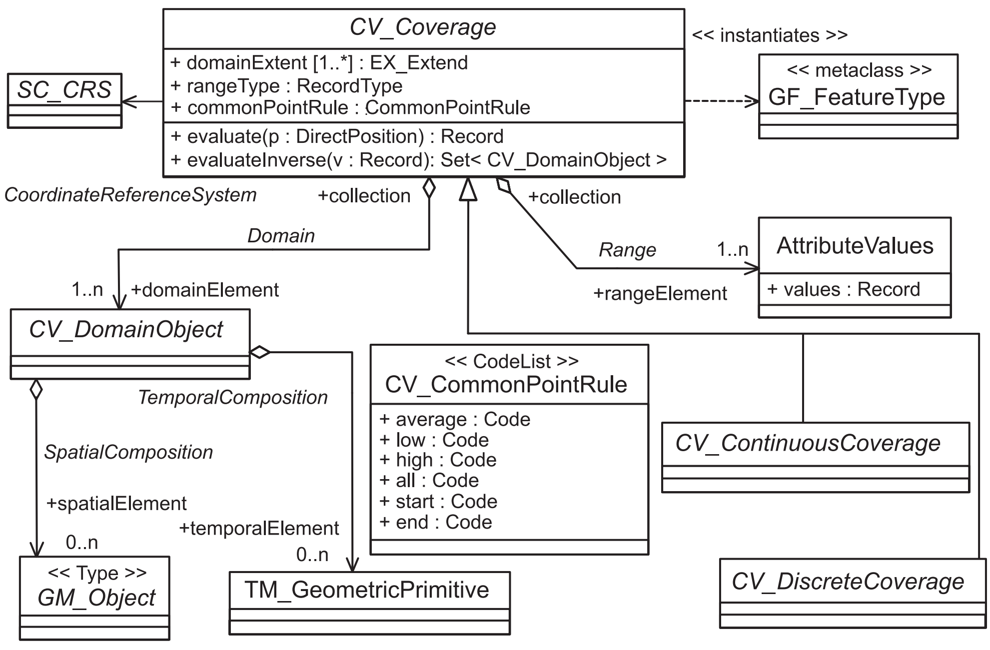

Fig. 3. Conceptual model of coverage. UML class diagram – logical view; Source: ISO 19123 2002

behaves as a mathematical function which for each point of the geospatial domain (surface, space and also discrete one) returns a value from the domain of attribute. Because of this, coverage is not only an aggregate of data, but it also has method (or methods, i.e. functions and/or operators), as is shown on Text-figs 3, 4 and 6.

The type of returned values is determined by the type of the attribute though a large spectrum of types can be applied. Examples of types are:

- Simple data types, e.g.: logical, integer, decimal (float and double), texts and character strings.
- Complex data types, e.g. qualification of colour, date, time or geological age.
- Enumerators, i.e. items containing names (or entries) and belonging to a closed ordered list.
- Items of code lists (often called dictionaries) containing connected pairs *name – text*.
- Data aggregates, e.g. data structures and *p-objects* (instances of classes) containing data.
- References, simple (pointers) and complex (URL – Uniform Resource Locator).
- Identifiers, simple (Oid – p-object identifier) and complex (URI – Uniform Resource Identifier).
- Connections, e.g.: inheritance (generalization and specialization), associations, aggregations and compositions.

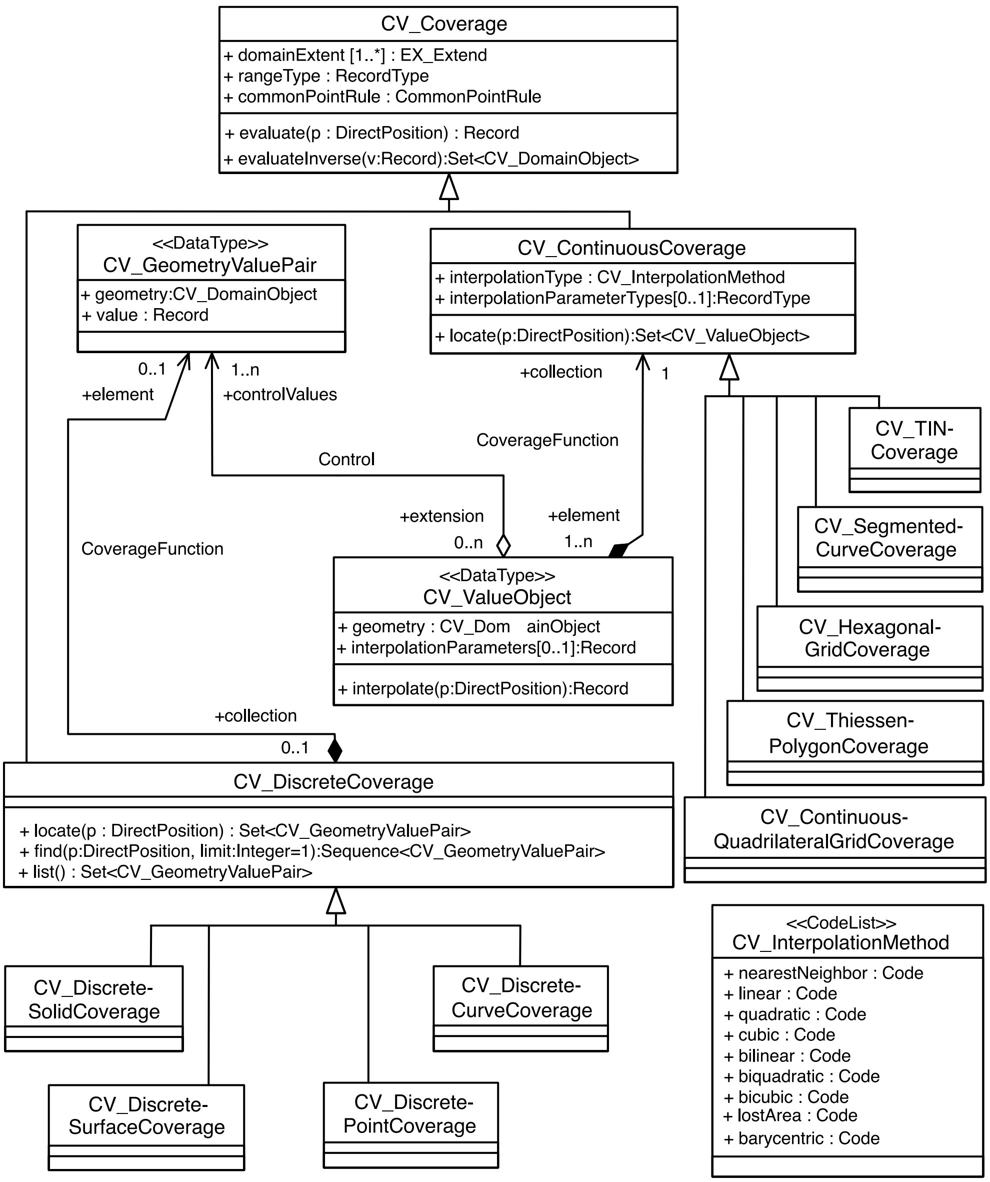

Fig. 4. Conceptual model of subtypes of Coverage. UML class diagram – logical view; Source: ISO 19123 2002

- But also methods: functions and operators (relational, logical and arithmetic) in advanced object models.

A feature that is not coverage cannot be treated as a geospatial function. Typical examples of coverage in the Earth sciences are:

- Physical and chemical fields: temperature, hydrodynamic potential, chemical potential.
- Surfaces: terrain surface, bottom of natural water pool, top and bottom of geological section.
- Spatially distributed parameters of continuous media: such as density, coefficients of electrical and heat conductivity and filtration coefficient.

COMPUTER REPRESENTATION OF FEATURES AND COVERAGES

In the geoinformation systems, feature (as a p-object) is most frequently expressed through encoding in vectorial representation as simple geometrical elements: point, line, surface or sets of these elements. Localizations of these elements are determined by a coordinate system corresponding to the Spatial Reference System (SRS). The most frequently used forms of encoding coverage are raster layers as well as images, cellular data and matrices. The common characteristic of the raster-represented types of data is tessalation of space (GAŹDZICKI 2001), when as a result a raster being a canvas (spatial structure) is generated for satisfied data associated with the particular coverage.

The fact that for both types of geospatial information (feature and coverage), vectorial and raster representations respectively are used most frequently, is not a limitation, and in several cases it is useful to apply raster representation for feature and vectorial representation for coverage. Several geoinformation systems use one or another form of representation uniquely. Examples of the application of various forms for coverage are given in Text-fig. 6. OpenGIS abstract specification for coverages determines several rules for transformation from vectorial form to raster form and reverse (The OpenGIS 1999).

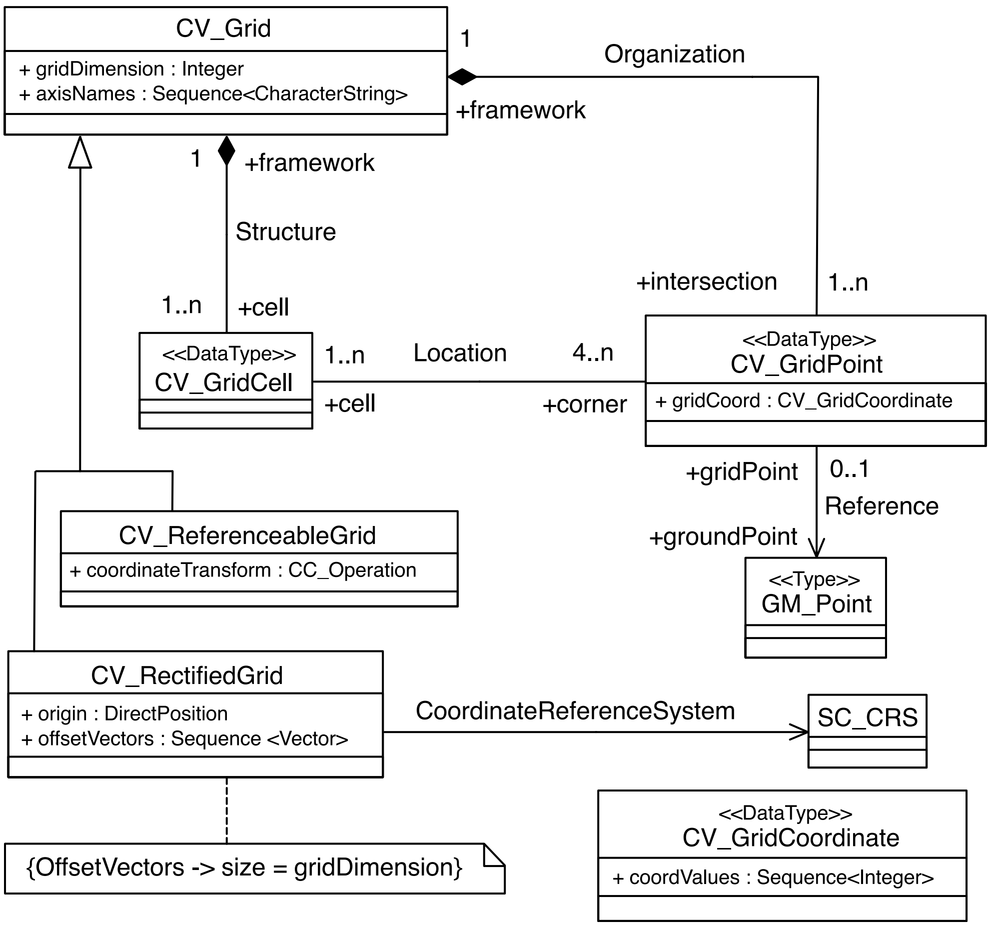

Fig. 5. Conceptual model of Grid Coverage. Inheritance of CV_GridPoint from *CV_DomainObject* is removed due to inconsistence with *CV_Coverage* model. UML class diagram – logical view; Source: ISO 19123 2002

DISTINCTIONS BETWEEN HYDROGEOLOGICAL FEATURE AND OBJECT

In hydrogeology, as in other geological sciences where the object of research is the Earth's crust, we are dealing exclusively with coverage as the geospatial information. The Earth as a planet is a unique real feature which meets the definition of an object. The Earth's crust constitutes a continuum, although differentiated and variable in time and space; however any separation of particular objects within this continuum is disputable.

Therefore, all that we can separate in the Earth's crust conforms to the type *feature* but many of them meet the definition of a *coverage*. Semantic analyses of words from the Hydrogeological Vocabulary (KLECZKOWSKI & RÓŻKOWSKI 1997) contained in the work of application of geomatics in hydrogeology (MICHALAK 2003) stated that among 95 category of entries with respect to the type of spatial feature (not being synonyms nor subtypes), 24 categories refer to types of object features. Most of them consist of technical objects, e.g. hydrogeological station, well, dam or piezometer. For genuine hydrogeological objects only sink-hole, spring, vaucluse spring and, to a certain extent, hydrogeological system meet the requirements of the definition of genuine object.

The majority of types of hydrogeological features are non-object features, because they do not meet the definition of an object. 71 categories of entries (of 95 analyzed) belong to such entry categories; representative examples of these categories are: ground water table (5 entries), aquifer (3 entries), bottom and top of the bed, aquifer (5 entries), area (15 entries) and zone

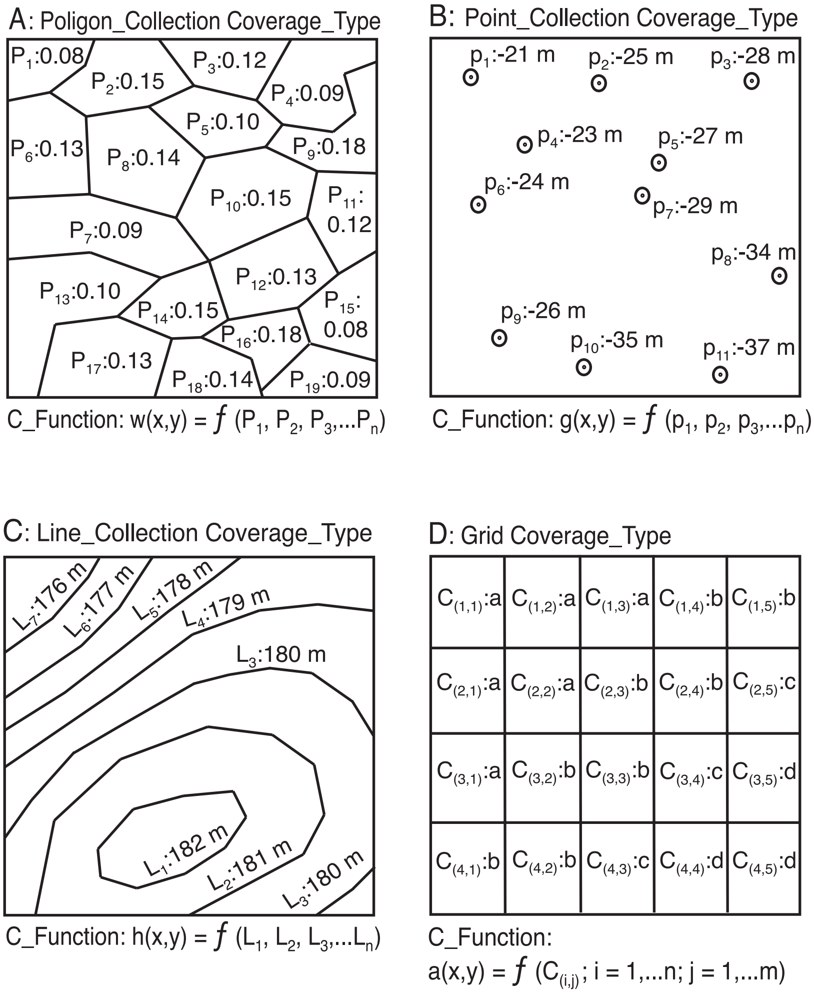

Fig. 6. Examples of coverage types for hydrogeological data; A, B, and C – coverages based on collections of features in vector representation; D – coverage based on grid (matrix) representation

(3 entries). It results that non-object feature consists of the basic ontological notion with the help of which the conceptual schemas concerning hydrogeological information can be built.

Among analyzed entries of the Hydrogeological Vocabulary, 14 of them belong to the category *isoline*, and this category need to be carefully considered. This is a non-real feature i.e. observed in reality. This is the only element of a graphical image (map or cross-section) which serves to show spatial variation of some physical field or surface. From the point of view of geomatics, isoline can be the only element of portrayal of geoinformation or component of coverage (feature treated as spatial function). In the ISO/TC211 conceptual model this type of coverage belongs to the subtype *DiscreteCurveCoverage* or *SegmentedCurveCoverage* (Text-fig. 4).

PARTICULAR KIND OF HYDROGEOLOGICAL INFORMATION – OBSERVATION

Hydrogeological information (hydrogeological data) is most frequently associated with a particular geospatial feature. It is realized in the data model using attributes (components of p-objects) or by associations between p-objects representing features and p-objects representing data. In this case, attribute constitutes a pair *name = value*, where the name determines the semantics and type of the attribute, while the value is an appearance of the element belonging to the defined domain in the type of this attribute. Associations allow much more unconstrained attachment of data to the features. Rules of these associations are defined by UML language (Unified Modeling Language) (OMG 2001).

Very often, hydrogeological data are not associated with the particular geospatial feature. In such a case one can serve himself by notion of *observation* and *measurement* defined in geomatics as particular types of point feature (COX 2002). In this work geometrical definition of measurement is as follow:

Measurement (feature) – an instance of a procedure to estimate the value of a natural phenomenon, typically involving an instrument or sensor. This is implemented as a dynamic feature type, which has a property containing the result of the measurement. The measurement feature also has a location, time, and reference to the method used to determine the value. A measurement feature effectively binds a value to location and to method or instrument.

On this basis one can build hydrogeological conceptual schemas for observation and measurements that are not associated with the real features.

WHAT KIND OF HYDROGEOLOGICAL INFORMATION IS A COVERAGE TYPE?

Coverage according to its definition and conceptual models (Text-figs 2-5), contains information which determines the spatial variability of a chosen attribute of a spatial feature or linkage to something else. There are many cases in hydrogeology in which spatial variability is described. Nearly all data composing a flow model or the transport of dissolved matter fulfill the requirements of coverage definition, and this is the most suitable form of encoding these data (MICHALAK 1997). Coverage type is also suitable for describing the majority of information contained in hydrogeological maps. Of 95 categories of vocabulary entries (KLECZKOWSKI & RÓŻKOWSKI 1997) subjected to analysis, 63 correspond to the type of feature that can be treated as coverage. Thus, when elaborating the conceptual schema for these features, one must consider them first as coverage type and then establish if this is only a feature or also an object.

Typical examples of features in hydrogeological information are: parameters of aquifer or aquitard, water table, top and bottom of aquifer, hydrochemical and hydrodynamic fields. The most frequently used coverage subtypes in hydrogeological models for these types of information are: *GridCoverage* (fig. 5) and *ContinuousQuadrilateralGridCoverage* (fig. 4). Information contained in hydrogeological maps can be presented as other coverage subtypes, for example from the group *Continuous Coverage*: *TINCoverage* and *ThiessenPolygonCoverage*. However, from the point of view of interoperability among simulation systems of hydrogeological processes and systems containing data from maps, application of other coverage subtypes should be dependent on the possibility of transformation of current subtypes into the most frequently used subtypes.

SUMMARY

The way in which geospatial information relating to hydrogeological processes is written in computer systems depends on the data model. The correctness of the conceptual schema applied in the data model is a critical factor which determines the functionality of the geoinformation system. In hydrogeology, as in other geological sciences where the object of research consists of the Earth's crust, only in few cases we are dealing with genuine objects. Instead, this is mainly geospatial information corresponding to the definition of the type coverage. The most specific property of coverage is that it behaves as a mathematical function which, for each point of the geospatial domain (surface, space) returns, a value from

the attribute domain. The majority of the data which comprise the underground flow model or transport model of dissolved solutes meet the requirements of the coverage definition and this is the most suitable form of their encoding. The *coverage* type is also suitable for encoding most information contained in hydrogeological maps. Very often hydrogeological data are not associated with particular spatial feature. In this case, one can serve himself with observation and measurement defined in geomatics as a particular type of spatial feature.

REFERENCES

Cox, S. (Ed.) 2002. Observations and Measurements. OGC Interoperability Program Report-Engineering Specification, *OpenGIS Project Document*, pp. 1-64. OGC; Wayland.

Gaździcki, J. 2001. Lexicon of geomatics, pp. 1-135. Polish Association for Spatial Information; Warszawa. [*In Polish*]

Herring, J. & Kottman, C. 1997. Features and Coverages for the Layman. *OpenGIS Newsletter*, 2 (**4**), 6-7. Wayland.

ISO 19103 (PDTR) – Geographic information – Conceptual schema language. 2001. *ISO/TC211 Document*, pp. 1-69. NTS, Oslo.

ISO 19123 (CD) – Geographic information – Schema for coverage geometry and function. 2002. *ISO/TC211 Document*, pp. 1-61. NTS, Oslo.

Kleczkowski, S. & Różkowski, A. (Eds.) 1997. Hydrogeological Vocabulary. Collective work, pp. 1-328. Wyd. TRIO, Warszawa. [*In Polish*]

Mark, D.M.; Skupin, A. & Smith, B. 2001. Features, Objects, and other Things: Ontological Distinctions in the Geographic Domain. *In*: D. Montello (*Ed.*), Spatial Information Theory, *Lecture Notes in Computer Science 2205*, pp. 488-502. Springer; Berlin/New York (URL http://ontology.buffalo.edu/smith/articles/COSIT01MSS.pdf).

Michalak, J. 1997. Object-oriented models in hydrogeology - system ASPAR [Obiektowe modele w hydrogeologii – system ASPAR], pp. 1-102. Wyd. Uniwersytetu Warszawskiego; Warszawa. [*In Polish*]

— 2003. Geomatics in hydrogeology. *Geol. Quart.*, **47** (1), 69-76.

Subieta, K. 1998. Object-orientation in design and data bases [Obiektowość w projektowaniu i bazach danych], pp. 1-283. Akademicka Oficyna Wydawnicza PLJ; Warszawa (In Polish).

— 1999. Glossary of terms in scope of object-orientation [Słownik terminów z zakresu obiektowości], pp. 1-290. Akademicka Oficyna Wydawnicza PLJ; Warszawa. [*In Polish*]

The OpenGIS Abstract Specification – Topic 0: Abstract Specification Overview – Version 4. 1999. *Open GIS Consortium Document*, pp. 1-36. OGC; Wayland (URL: http://www.-opengis.org/techno/abstract/99-AS-RFP009.pdf).

OMG Unified Modeling Language Specification – version 1.4. 2001. *Object Management Group Publication*, pp. 1-566. OMG; Needham (URL: http://cgi.omg.org/docs/formal/01-09-67.pdf).

Whiteside, A. 1999. UML Profile and Guidelines for OGC Abstract Models. *OpenGIS Project Document 99-031*, pp. 1-10. OGC, Wayland.

ISO 19123 (CD) – Geographic information – Schema for coverage geometry and function. 2002. *ISO/TC211 Document*, pp. 1-61. NTS, Oslo.

Manuscript submitted: 10th September 2002
Revised version accepted: 15th June 2003