Semantic Technology of Intellectual Geoinformation Systems Development

Sergei Samodumkin
Belarusian State University of Informatics and Radioelectronics
Minsk, Republic of Belarus
Samodumkin@bsuir.by

***Abstract*—This paper is devoted to the creation of a special technology for designing intelligent geo-information systems that use knowledge of terrain objects to solve applied tasks in problem areas.**

***Keywords*—semantic technology, geographic information system, topologic-geographical relations**

I. Introduction

Knowledge and data about terrain objects can interest us not only as spatial data and knowledge, moreover they are an integration basis for various subject domains. Formalization of such knowledge and its presentation in knowledge requires both determination of subject relations for description of properties and consistent patterns inherited in observing subject domain using terrain objects, but also determination of geometric characteristics, that are able to bind objects on area. In addition, taking in account life of information and terrain object themselves, there is possibility of retrospective analysis, which makes possible to observe terrain objects in subject domains not only from their spatial position and and semantic attributes, but also to take into account temporal aspect of existing of terrain objects. As intellectual systems are designed to meet information needs of users, these prerequisites promote an expansion of subject domains and an addition of new functionality in the design of intelligent geographic information systems [1-6, 11].

However, until now, knowledge about terrain objects was considered as cartographic data and was a result of mapping of search queries with applying result to terrain maps. At the same time, due attention was practically not given to the data on the terrain as an integrating element of different subject domains, the dynamics of their change, taking into account their actualization in time counts. Well-known studies aimed at ensuring consistent information exchange between spatial and subject knowledge to ensure, that semantic interoperability were conducted for systems based on RDF, RDFS, OWL Semantic Web technology stack and as shown in [7] the OWL web language presents advanced capabilities to describe the subject domains of interacting systems and gives machine-interpretable definitions of fundamental concepts of subject domain and the relationship between those in the ontology. However, formalization of the subject domains and ontological engineering is only one of stages in the design technology of intelligent systems and are not sufficient in themselves to implement conclusions based on knowledge, because ontological engineering allows describing declarative knowledge of subject domain, while procedural knowledge allows designing task solvers and implementing knowledge based conclusions.

The availability of developed design technology and tools is an important point that reduces, on the one hand, period of development of intelligent systems, and on the other hand, increases functionality of intelligent systems that use knowledge of objects of terrain as an integration. At the same time, the design technology should ensure reusing of information and functional components of the system in order to shorten the design and development time of the applied systems. The above requirements are possessed by the semantic technology for designing intelligent systems — OSTIS, the advantage of which is its extensibility both in terms of expanding the subject domains, types of knowledge used and their formalization, and in terms of functionality.

Despite the advantages of the OSTIS [8] technology, knowledge of the objects of the terrain as integrating elements, the semantic interoperability of such knowledge is not paid attention to, the procedures for interaction of spatial knowledge with knowledge of subject domains, and the possibility of correlating objects of terrain in time have not been established. Thats why actual task is, first, to design spatial ontologies taking into account the life cycle of terrain objects and based on their solution to the problem of semantic interoperability of knowledge of subject domains, as well as solving the problem of managing metadata and improving search, access and exchange in the context of growing volumes of spatial information and services provided by numerous sources of geo-information, secondly, the implementation of conclusions based on knowledge using spatial and mathematical information as components of knowledge of terrain objects; thirdly, the integration of the cartographic interface as a natural way of presenting information about terrain objects to humans.

This work is devoted to the creation of a private technology of designing intelligent systems that use knowledge about objects of the terrain to solve applied problems in subject domains. In this connection, the paper provides

1) a formal basis for representing knowledge about terrain objects taking into account the spatial component (cartographic knowledge description language) and formalizing the thematic component of terrain objects based on ontological engineering of subject domains using knowledge about terrain objects;
2) the organization of conclusions based on knowledge using knowledge about the objects of the terrain;
3) integration of the cartographical interface.

Private technology of designing intellectual geo-information systems considered in the work is based on the semantic component technology of designing intellectual systems. Graph-dynamic models of a special kind - semantic models of knowledge representation and processing based on semantic networks - are used as a formal basis of the projected intellectual systems and the basis of abstract logical-semantic models of intellectual systems.

II. Ontology of terrain objects as a formal basis for integrating knowledge of subject domains in the design of intelligent geographic information systems

In geographic information systems (GIS), terrain data, which is called geo-information data (geodata), is considered as the basis for solving specific applied problems. To solve such problems, it is proposed to use ontology as a conceptual model that allows to represent objects of a terrain at the semantic level and at the knowledge level. The advantage of using ontologies is that the described spatial data in the form of a semantic code, formally interpreted (understood) by a computer, can also ensure the integration of geodata obtained from various sources and in various forms of representations. In addition, different experts in different subject domains (for the implementation of interdisciplinary tasks) completely differently represent the same data sets in GIS. First of all, it is connected with different cartographic coordinate systems, which can be reference (that is, adapted to a specific part of the Earth's surface) and common, use ellipsoids with different parameters and, accordingly, have different coordinates of the same physical object of the terrain. Thus, when presenting spatial objects, there is semantic inconsistency, for eliminating which it is necessary to perform ontological engineering and identify key objects of subject domains that use spatial data to represent terrain objects taking into account the life cycle of terrain objects as well as typical relationships over terrain objects.

The selection for each of the classes (types) of terrain objects of the main, inherent only to him, semantic characteristics is the basis of ontological engineering of terrain objects. At the same time, metric characteristics of terrain objects do not possess this property. To indicate the semantic properties of the terrain feature classes, it is proposed to use the topographic information classifier displayed on topographic maps and city plans NCRB 012-2007 [9].

According to the that classifier, each terrain object has unique unequivocal designation. The hierarchical classification has eight grading stages: class code, subclass code, group code, squad code, suborder code, species code, subspecies code. Thus, thanks to the coding of already defined generic connections, reflecting the interrelations of various classes of terrain objects, the characteristics of classes of terrain objects are also established. Due to the fact that these properties and relations are not related to specific physical objects, but to their classes, this means that specific meta-information objects are a set of these meta-information of an intellectual geo-information system.

The ontology of terrain objects includes the description of the following classes of terrain objects:

– water objects and hydrotechnical structures;
– settlements;
– industrial, agricultural and socio-cultural objects;
– road network and road structures;
– vegetation cover and priming.

The ontology of terrain objects is a classification tree in accordance with the hierarchy shown in Figure 1. For each class of terrain objects, generic links are established. For instance, Figure 2 shows the hierarchy of water objects.

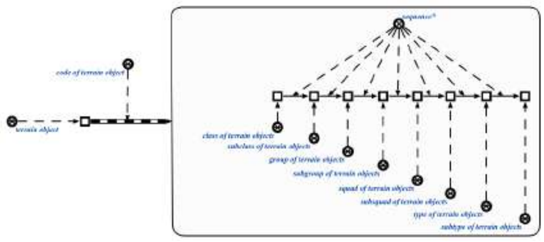


Figure 1. Levels of hierarchy for classes of terrain objects

For each class of terrain objects, semantic features that characterize terrain objects are set. At the same time, for each class of terrain objects, its own characteristic set of features is highlighted (for example, in Figure 3, for all objects of the type "river", the relations "eigenvalue*", "width on a scale*", "sign of shipping*", "water quality features*").

Thus, the considered ontology of terrain objects and the method of its formal setting allow us to describe all the main classes of terrain objects and establish for these

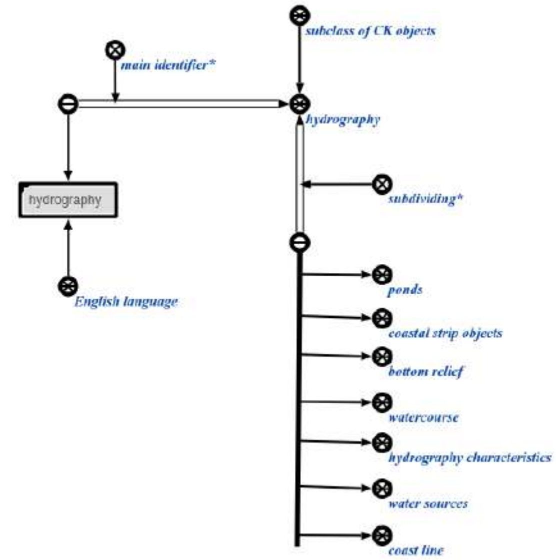


Figure 2. Hierarchy for terrain water objects

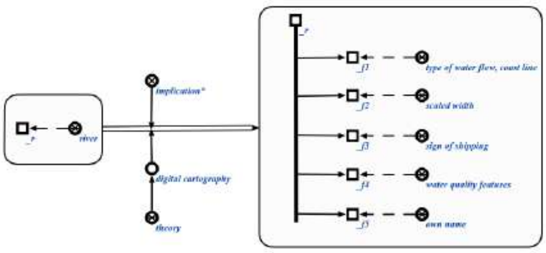
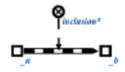
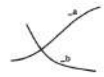
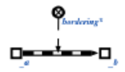
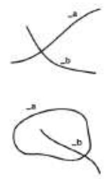
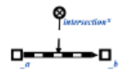
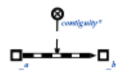


Figure 3. Assignment of semantic features for terrain object «river»

classes a set of features characteristic of the considered class of terrain objects.

At the same time, the designed ontology of terrain objects is not enough to provide conclusions based on knowledge, since it is required to set relationships over terrain objects. The representation of terrain objects in a graph-dynamic model required the identification of types of relationships that can be defined over terrain objects. In addition to subject connections for processing knowledge in intelligent geographic information systems, it is necessary to set topological and geographical relations. To this end, we select the types of terrain objects: areal objects, linear, polylines and points. Table I shows all established relationships, the types of objects for which they are established, as well as a schematic representation of the relations and structures for their storage in the language of semantic networks used in OSTIS technology.

Table I
TOPOLOGICAL RELATIONS, SET ON TERRAIN OBJECTS

Relation type	Object type	Scheme	SC notation
Inclusion	Areal and linear (multilinear) objects Areal objects and points Areal objects		
Bordering	Areal objects		
Intersection	Linear (multilinear) objects Areal and linear (multilinear) objects		
Contiguity	Linear (multilinear) objects		

III. CONCLUSIONS BASED ON KNOWLEDGE USING SPATIAL AND THEMATIC INFORMATION

The implementation of the findings, based on knowledge, is implemented by domain solvers using agents [10]. At the same time, in connection with the specifics of geographic information systems, operations for working with topological-geographical relations, operations for semantic comparison of objects of terrain, and integration of subject solvers are proposed.

A. Operations with topologic-geographical relations

The search for the inclusion relationship is possible between areal, areal and point, areal and linear (multilinear) objects. As particular cases - the search for relations of inclusion between administrative objects (regions, districts, settlements). To solve this problem, algorithms of computational geometry are used. In general cases, the definition of an inclusion relationship between objects of a locality is reduced to an algorithm for determining whether a polygon contains a point. For areal and linear (multilinear) and for areal objects, all points of one object are checked for containing all points of other object.

At the first stage, when checking the inclusion relationship between areal and linear (multilinear) objects, a check is performed for the nesting of rectangles in which the object data is entered. If the rectangles are not inscribed, then no further verification is performed.

This check allows you to significantly save computing resources and execution time. Then, to verify the inclusion of point objects and areal, the algorithm for determining whether a point belongs to a polygon is executed, and for other objects that have successfully passed the first check, this algorithm is performed for each point of the object. As an algorithm for determining whether a point belongs to a polygon, the ray tracing method is chosen.

After defining inclusion relationship between terrain objects, they are formed in graph structures designed to store the relationship (Fig. 4).

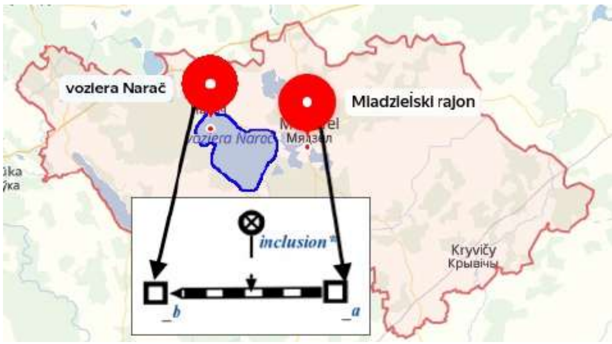


Figure 4. Formation in memory of the established relationship of inclusion between the terrain objects ("Myadel district includes Lake Naroch")

The intersection relation is distinguished between linear (multilinear) and polygon objects. The fact of intersection of at least one part of one object and at least one part of another object is checked. On the first stage rectangles are determined. If the rectangles have no intersections or inclusion, there is no further checking. Else procedure is being repeated for segment objects. If the fact of intersection is established, the objects intersect, and then they are entered into formed graph structures designed to store this type of relationship (Figure 5).

The "bordering" relationship stands out between areal objects. The fact of coincidence of at least one part of one object and at least one part of another object of the terrain is checked. If there is a coincidence and the objects do not overlap, then a “border” relationship is established for them, which is written into the graph structure (Fig. 6).

The relation "contiguity" is allocated between linear (multilinear) objects. The fact of contiguity of the ends of an object with any part of another object is checked. At the first stage, the rectangles in which the scanned objects are inscribed are determined. If the data rectangles do not intersect and do not fit into each other, further verification is not performed. Otherwise, the procedure is repeated for the first object of each segment of the second. If the rectangles that are inscribed inscribed object and a segment intersect or fit into one another, then the algorithm checks the belonging from the ends of the first

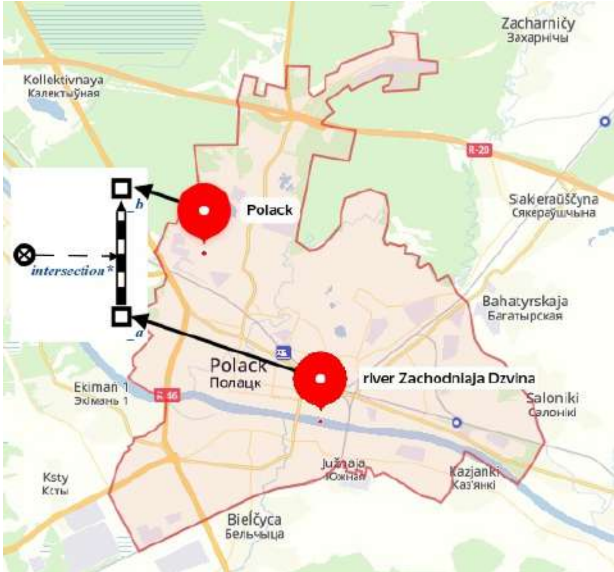


Figure 5. Formation in memory of the established relationship of intersection between terrain objects ("The Zachodniaja Dzvina River crosses the city of Polack")

Figure 6. Formation in memory of the established relationship "bordering" between the objects of the area ("Belarus is bordered by the Lithuanian")

object to this segment. In the event that the ownership is established, the objects are adjacent to each other.

If a membership established, the objects have a common border to each other and the corresponding graph structure is formed in the memory (Fig. 7).

When recording the junction relation for the rivers, the fact that the rivers consist of segments having different attributes or their meanings is taken into account. On the basis of this, an additional relationship is established for the rivers, decomplexing. Each river is decomposed into segments, and those in turn adjoin each other.

It should be noted that the formed topological-

Figure 7. Formation in memory of the established relationship of junction between objects of the terrain ("The Svislach River contiguity the Biarezina River")

geographical relations between the objects of the terrain may have their own interpretations in the subject domain. For example, the formed relation "The Svislach River contiguity the Biarezina River" in the knowledge base of the intellectual system will be interpreted as the "Svislach River adjoins into the Biarezina River".

B. Operations for semantic mapping of terrain objects

In addition to establishing topological relationships, for the correct and unambiguous storage of terrain objects in the knowledge base of an intelligent system, it is necessary to carry out a semantic comparison of geographic objects. The semantic comparison of geographical objects of the map occurs according to the following principle:

- the object class is determined;
- subclass, type, subspecies, etc. are determined. object in accordance with the classification of terrain objects, i.e. types of terrain objects in ontology;
- attributes and characteristics that are inherent in this class of terrain objects are determined;
- defines the values of the characteristics for this class of object;
- the homonymy of identification is eliminated;
- corresponding links are established between the map object, knowledge base object and object attributes;
- establishes topological relationships between map objects related to specific classes and types.

As a result of these actions, the first version of the knowledge base is formed, in which there are map objects with established topological relationships. As a result, new knowledge is formed.

The second version of the knowledge base is the result of the integration of the knowledge base, obtained at the first stage and the external subject-oriented knowledge bases. Such integration allows you to fill a knowledge base with new types of knowledge, as well as to eliminate the homonymy of geographic objects. For example, using the knowledge base of settlements with codes according to the SOATDS (system of designations of objects of administrative-territorial division and settlements) allows identifying settlements with the same name but belonging to different administrative-territorial units in a one-to-one manner.

C. Integration of subject solvers

An advantage of OSTIS technology used is a possibility of using problem solvers. Thus, such problem as implementing a search for routes between specific terrain objects, calculating geometric characteristics (for example, an area of a territory) are performed by agents that can be implemented within the OSTIS technology, and also be third-party applications initiated to solve applied domain problems.

On Figure 8 a specification of task for the subject domain is shown. The result of processing with intellectual agents is shown on Figure 9.

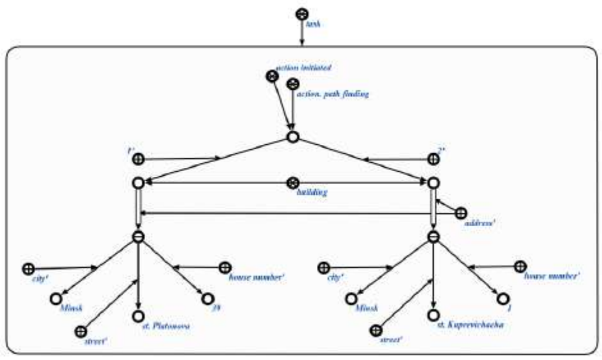


Figure 8. Initial specification of an agent processing

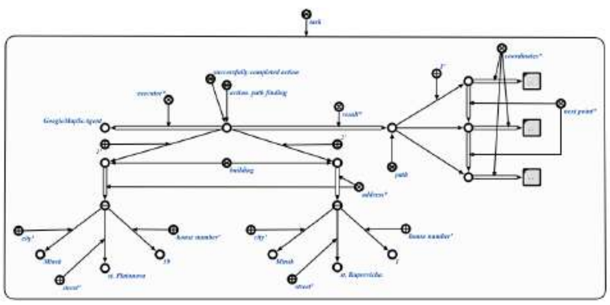


Figure 9. Result of an agent processing

IV. Integration of the cartographic interface as a natural way for people to represent information about terrain objects

The proposed principle of coding geo-information allows the comparison of terrain objects described within the framework of the subject domain and currently known services of cartographic information in the Internet environment. Thus, visualization of terrain objects stored in knowledge bases using Google, Yandex and Open Map Street services is possible, as well as semantic comparison of geographical objects with a knowledge base with stored knowledge about these terrain objects based on the described principles in this article.

On Figure 10 a scheme of cartographic data visualization is shown.

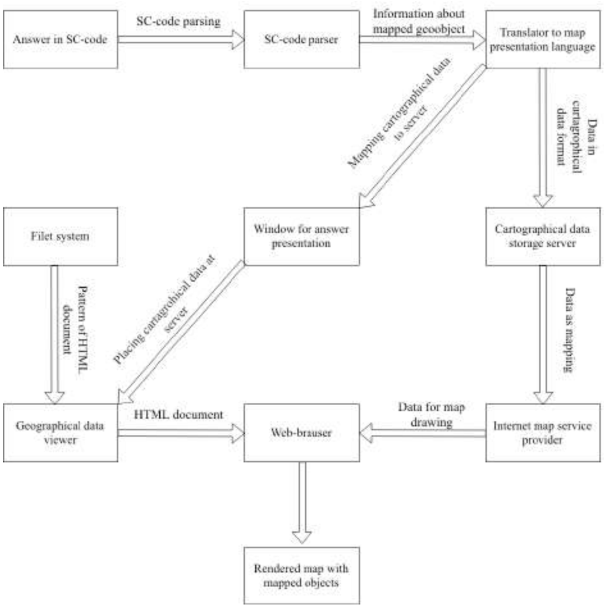


Figure 10. Scheme of cartographic data visualization

V. Example of usage

One of the "advantages" of intelligent systems developed using the OSTIS technology is the solution of objective problems when there is no clear specification and algorithm to solve it. This is achieved by forming products that are recorded and stored also in the knowledge base Design technology of knowledge processing machines and models for solving problems.

As an example, consider the solution to the following problem: "Determine whether there is a water route between the Min cities If there is such a waterway, display the result on the map". The following statements of the knowledge base are the basic data:

1) The Svislach River flows through the city of Minsk.
2) The Dnieper River flows through the city of Rechitsa.
3) The Svislach River is a tributary of the Biarezina River.
4) The Biarezina River is a tributary of the Dnieper River.

During the first iteration, a "contiguity" will be established at the first iteration, analyzing the topological relations:

1) The Svislach River flows into the Biarezina River (i.e. the Svislach River near the city of G.);
2) The Biarezina River flows into the Dnieper River (i.e. the Biarezina River adjoins);

And it will be concluded that the Svislach River can be used.

At the second iteration the city of Minsk and the Svislach River, as well as the city of Rechitsa and the Dnieper River. As a result, the path that is displayed on the map of the area will be found.

References

[1] Xaun Shi "The Semantics of web Services: An Examination in GIScience Application," ISCPRS Inetrnational Journal of Geo-Information, pp. 888-907, February 2015.

[2] Burcu Alinci and Hassan Karini "CAD and GIS Interoperability through Semantic web services," Electronic Journal of Information Technology in Construction.

[3] A. Diostenu and L. Cotfas "Agent Based Knowledge Management Solution using Ontology, Semantic Web Services and GIS," Informatics Economics, vol. 13, pp. 90–98, April 2004.

[4] Gregor J.Nalepa and T. Furmanska "Review of Semantic Web Technologies for GIS," Automatyka, vol.13, pp. 485–492.

[5] F. Fonseca, M. Egenhofer, P. Agouris and C. Camara "Using Ontologies for Integrated Geographic Information Systems," Transactions in GIS, vol.6, March 2002.

[6] Mahadi Farnaghi and Ali Mansourian "Multi-Agent Planning for Automatic Geospatial Web Service Composition in Geoportals," ISCPRS Inetrnational Journal of Geo-Information, July 2018.

[7] Din Le Dat and V. Serebryakov "Designing and Development of Formal Ontologies of Geographical Spatial Data and Services," Robototechnics Journal, vol. 2, pp. 85–89, 2008.

[8] V. Golenkov "Ontology-Based Design of Intellegent Systems," Materials of Scientific and Technical International Conference OSTIS ,vol. 1, pp. 37–56, February 2017.

[9] "Digital maps of the area. Topographic information displayed on topographic maps and city plans," NCRB 012-2007.

[10] D. Shunkevich "Ontology-Based Design of Knowledge Processing Machines," Materials of Scientific and Technical International Conference OSTIS ,vol. 1, pp. 73–94, February 2017.

[11] V.B. Taranchuk "New computer technologies, analysis and interpretation of geodata," MATEC Web of Conferences IPICSE-2018. V. 251, 04059. VI International Scientific Conference "Integration, Partnership and Innovation in Construction Science and Education" (IPICSE-2018). – P. 1-8.

СЕМАНТИЧЕСКАЯ ТЕХНОЛОГИЯ ПРОЕКТИРОВАНИЯ ИНТЕЛЛЕКТУАЛЬНЫХ ГЕОИНФОРМАЦИОННЫХ СИСТЕМ

Самодумкин С.А.

Настоящая работа посвящена вопросам создания частной технологии проектирования интеллектуальных геоинформационых систем, использующих знания об объектах местности для решения прикладных задач в проблемных областях.

Received 10.01.19