Intelligent GIS Conceptualization

Vasily V. Popovich

Abstract The concept of an intelligent geographic information system (IGIS) is not new. However, the term represents a complex set of ideas and methods that cannot be combined into one method, theory, or concept. This chapter attempts to describe this new domain in geographic information system science and technology. Here, we present selected results of theoretical investigations and technological developments to share the vision of IGIS.

Keywords GIS · IGIS · Artificial intelligence · Inference machine · Ontology

1 Introduction

The history of the intelligent geographic information system (IGIS) is comparatively short. The first references to IGIS can be found in literature from the mid-1980s on the geographic information system (GIS), in which the intelligent subsystem was considered to be intended for solving a certain class of tasks, mainly related to data mining and cartography.

GIS applications for decision-making support systems (DMSS), which were considered to be a convenient interface and a key technology, created new requirements and challenges for conventional GIS. The requirement of intelligence is imperative for obvious reasons. One of the main challenges is the complexity of GIS, which is a key factor in computer science development, particularly in the process of software development. The entire evolution of the software development process from first-level languages to the concept of the object-oriented approach (OOA), which includes analysis, design, and programming, is based on

V. V. Popovich (✉)
St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS), 39, 14 Linia VO, St, Petersburg 199178, Russia
e-mail: popovich@oogis.ru

V. Popovich et al. (eds.), *Information Fusion and Geographic Information Systems (IF&GIS 2013)*, Lecture Notes in Geoinformation and Cartography,
DOI: 10.1007/978-3-642-31833-7_2,

dealing with complexity (Booch 1995) . OOA applications inspired the development of large and complex systems. However, some problems were exposed in the application of this methodology, such as in business logic development for a subject domain or a system of subject domains combined in one project. The business logic development process for a subject domain stretches into an infinite chain of software development cycles, according to Edward Yourdon (1989). These conditions thus threaten the implementation of heterogeneously distributed system development as a whole.

In addition, the complexity of decision making in a DMSS is a challenge, especially in real-time systems. This complexity is mainly caused by the human factor phenomenon (Brooks 1975). The intellectualization of the software development process and the decision-making process created a universal approach to solving specified groups of problems. All of these reasons were key factors in the development of GIS intellectualization.

IGIS is defined here as GIS that includes integrated tools and/or systems of artificial intelligence (AI) (Popovich 2003, Popovich et al. 2011). IGIS incorporates a number of basic components, such as:

- an inference machine (IM) and expert system (ES);
- a knowledge-based (KB) system (ontology);
- a visual environment for developing classes and objects of a subject domain;
- a visual environment for developing the models (script for acts) of objects functioning in GIS;
- a system for scenario implementation in real time or/and a user-defined arbitrary scale with a visual display as symbols or images on the background of electronic maps;
- recommendations for decision makers during the scenarios playing within the real system on research design, business games, situation analysis, and learning and training of personnel.

The proposed approach considers an ES and ontologies to be a system of AI. Thus, ES is used dually as a conventional ES, which assists the decision makers, and as a system to manage various processes working under control of some "scenario."

ES applications in GIS led to the emergence of many new scientific directions and technologies, such as:

- visual development of geospatial subject domain ontologies;
- development of ESs based on rules and objects;
- knowledge representation of the spatial environment based on ontologies.

Our experience indicates that IGIS applications increase the development speed and reliability of general and special software. IGIS also increases the efficiency of various information systems, including those for decision making on various levels, including strategic, operative, and tactical.

Many scientific methods and technologies related to intellectualization are united under the common name of AI. AI has been defined as "the scientific direction that poses and solves the problems of hardware or program modeling of human activities that traditionally are considered intellectual" or as the "property of intelligent systems to carry out functions, for instance, the creative ones, which traditionally are considered to be a human prerogative" (Sorokin et al. 2005). However, the notion of AI can have different meanings in different languages. For example, consider the remark made by Dr. T. A. Gavrilova, the chairman of St. Petersburg Chapter of the Russian Association for AI:

> for Russian-speaking readers …in English the word-combination AI does not bear that slightly fantastic anthropomorphous tinge that it has acquired in quite unsuccessful Russian translation. The word intelligence means "ability for reasoning", and not "intelligence" for which there exists an English analogue—"intellect."

This chapter discusses the systems and methods of AI rather than "intellect" itself.

Scenario is an important notion for IGIS. Here, the scenario is considered to be the selfsame algorithm with a capacity for parallel execution in some of its branches; the scenario has an ontological representation and is interpreted by the IM. The presence of such basic systems as the ES, IM, and scenario approach opens up high-grade and flexible possibilities for IGIS use in DMSS. In addition, AI could be used in the conventional AI sense as well as a real-time control system. AI applications in IGIS have the following advantages:

- the possibility of developing the business logic of the decision maker in the form of scenario ontology, thus assuring independence of software code and the possibility of visual development and correction for each scenario in the system's development and operation;
- a sharp decrease in cost for business analytics development;
- a decrease in DMSS possession cost and depreciation of DMSS possession;
- a decrease of the number of errors in business logic;
- the possibility of a fast, practically instantaneous, change of business logic, thus avoiding the traditional software upgrade cycle.

In addition to the intellectualization problem, IGIS solves some other problems, as demonstrated by the history of GIS formation.

1.1 Universal Model of Data

Historically, the majority of GIS researchers and producers are oriented toward the production of geographic chapter maps, as was dictated by the old cartographic tradition that faithfully served mankind for many centuries. Upon the appearance of computers, the situation changed dramatically; nevertheless, the newly emerged capacities were slowly introduced into practice. The old tradition preserves the whole cycle, from user interfaces to data models and exchange formats. The production (printing) tool of chapter maps penetrated into other subject areas, such as modeling systems, automated systems of decision-making support, and many others that are far from map production. The notion of layers is one of the main concepts of "chapter" cartography that found a wide application in electronic cartography. Layers became the basis for many data models and cartographic formats. As long as GIS was primarily used as the means for cartographic information visualization, the problem was insignificant; however, implementation of GIS in modeling systems and DMSS constituents obviously indicated the unpromising use of the layer notion and layer-based technologies.[1]

It is worth noting that the situation in cartographic production changed dramatically. According to the National Oceanic and Atmoshpheric Administration, the USA began using a new system of electronic cartography that incorporates the chapter maps printing system as one of the functions within the integrated cartographic system instead of as a separate technologic process.

This chapter focuses on the scientific and technical aspects of GIS development as a tool for decision-making support in real-time (or close to real-time) systems. In the considered systems, the electronic imitation of chapter maps does not work because the system should first and foremost assist the user in the decision-making process rather than simply displaying information in a user-friendly graphical form. More importantly, the system should not overload the decision maker with too much information, which may be redundant and hinder the user's perception instead of facilitating it.

To implement GIS in DMSS, detailed attention should be paid to GIS development basics, starting with data models (Popovich et al. 2004). Data models are the base for the subject area description—the main logical and mathematical scheme that is the foundation for all data, information, and knowledge storage. Eventually, the data model determines the complete GIS configuration and architecture.

The main assumption in data model development is that operation with the environmental information (subject area) cannot be accurately specified by a certain aggregate of analytical and/or other functions; that is, the unambiguous (isomorphic) transformation of one parameter into another based on the set cannot be determined. The considered approach matches the Immanuel Kant thesis "thing-in-itself" (Ding an sich), which has become increasingly more understood in a broad sense (Kant 1889). In general, certain phenomena can be observed such that any phenomenon stays partially incognizable. In this regard, information about the phenomenon should be replenished continuously from outside sources, such as a measuring apparatus and/or data source, so that analytical and statistical concepts concerning the entity and phenomenon are continuously refined and edited. For example, in digital mapping, real physical parameters such as temperature, salinity, pressure, and speed measured by sensors should be considered rather than abstract parameters such as current, isobaths, cyclones, and ice, which are "assumed" by man. Some conventional and/or other abstractions could be

[1] Useful model for the presentation of information in the GIS.

exclusively formed based on primary data. This is important because the isomorphism condition is rarely used in data transforms. Furthermore, the situation could occur when other processing methods are applied to already processed data, thus causing an impetuous increase of data inaccuracy. The formal substantiation of these ideas can be found in higher algebra model theory (Tarski 2004).

Avoiding the layer notion and receiving (when needed) direct access to measurements requires an introduction to the abstract data model or environment universal model (EUM). Such a model is based on an abstract space extraction in the form of a multidimensional vector, allowing for adjustment to a definite subject area. This model accounts for environmental changes by conventional electronic navigation maps, topographic maps or specifically generated subject area maps, diagrams, reports, objects, and three-dimensional (3D) dynamic images based on modern computer paradigms or other information technologies.

1.2 General Architecture of IGIS

Depending on the theoretic research and practical implementation, the major requirements for IGIS general architecture could be defined as follows:

- OOA should be used in IGIS as a primary approach instead of a cartographic (topographic, topological, etc.) approach for describing maps and objects on maps. It should include complex object support, class property inheritance, encapsulation, polymorphism, overlapping, reboot of functions (object models), etc. Here, object orientation does not mean that IGIS is written using object-oriented tools; the idea is that the same mechanisms should be rendered to the system user in an interactive mode to allow for organization of the space–time data.
- Different IGIS users should be given an opportunity to simultaneously attribute any objects on the electronic map to various classes and to construct multilevel “subjects” (classification systems with different bases), possibly for each user’s application or definite research.
- Maximal openness of IGIS architecture assumes that, by default, IGIS represents the mechanisms of data exchange with external systems based on open data formats and the possibility of IGIS function extension. The presence of an internal restricted language is insufficient; it is required that IGIS provides access to its functions and internal data from external programs, written by other developers based on the developers’ application programming interface. Distributed and rapidly developing systems should be able to be integrated with other developments existing in different subject areas.
- It should be possibile to establish all necessary topological relationships between the objects on the digital map. The importance of system support for the whole spectrum of topological relationships between objects is so high that, if the above mechanisms are omitted, IGIS turns into a map editor. However,

IGIS is primarily intended for analysis of information distribution in space and for modeling situations and processes as needed for prognosis. In particular, without the inter-objects and inter-layers topological relations the problems of modeling (analysis and prognosis) the development for complex network systems.

- IGIS should be easily managed in a distributed environment (including the Internet).
- The presence of a well-developed system ensuring rights of access to cartographical and semantic data is mandatory. So far, in most modern IGIS, a well-developed system that ensures rights of access to cartographical information does not exist.
- Embedded IM and ES should be included.
- An embedded database (DB) system (ontology system) should be included.

Figure 1 schematically depicts the general architecture of IGIS (Popovich et al. 2005).

The central part of IGIS is a KB that encompasses an aggregate of ontologies describing the subject area entities and their relationships. The other part of the KB is an aggregate of objects representing real entities of the subject area. An ES that includes IM and a set of logical rules is an important part of IGIS. Other architecture components are quite conventional for GIS. The structures of some constituents will be further considered in detail.

The object server is intended to support the uniform object model of the whole system. The server's main functions are to give access to the subject area objects and perform operations on them, eventually building objects based on the subject area meta-model stored in ontology, destructing objects, and modifying the properties of objects.

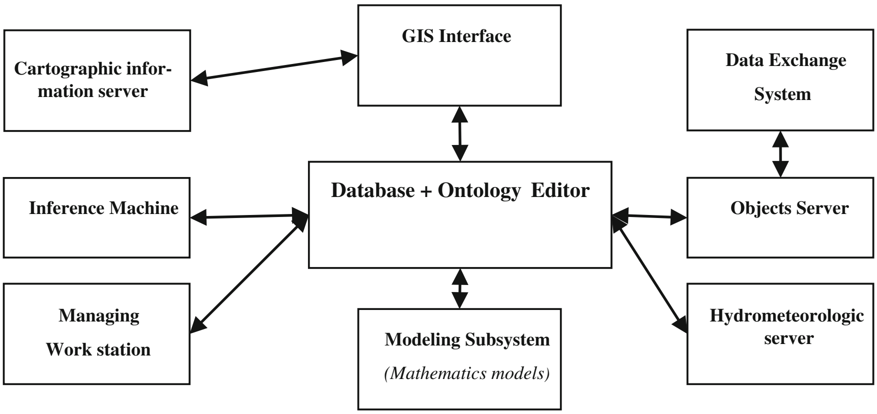

Fig. 1 General architecture of IGIS

The GIS interface is composed of minimal set of constituents, executing interactions between the system and the user and integration of software constituents in order to support an appropriate system functionality. The GIS interface renders a set of interfaces for interaction with the other software constituents at a technological level. The GIS interface represents the current situation on the electronic map and arranges for interaction between the user and application. To adjust the whole system and the GIS interface to a certain subject area, plug-ins are used to change the appearance of the interface accordingly. Localization of the user's location, workstation, keen client, worksheet, or communicator is performed via the GIS interface.

The complex architecture of IGIS cannot be realized within one programming system or software development environment. Organization of different levels of servers is required to support interactions between all constituents.

2 Service-Oriented Architecture

As previously discussed, IGIS incorporates a number of interacting concepts, ideas, approaches, and technologies. Theoretically, an abstract IGIS theory can be the subject of separate research, resulting in the construction of a system with rigorous axiomatic approaches. Moreover, practical orientation is one of the important factors in GIS science. In this regard, special attention should be paid to the problem of different approaches' interaction on a technological level and to developing ready-made applications on a theoretical level. In light of the above, let us consider the idea of interaction as proposed by the concept of service-oriented architecture (SOA), which is currently an advanced concept in information systems' interaction.

2.1 General SOA of IGIS

To realize the SOA, it is necessary to logically construct the architecture and then arrange for its realization through a certain set of technologies, products, and platforms for different purposes. SOA consists of three major logic components:

- *Users*—someone using the service proposed by a supplier;
- *Suppliers*—someone offering a definitive service or function;
- *SOA infrastructure*—a basis for arranging a normal consumer–supplier interaction.

IGIS users might be heavy desktop users, or they may use interfaces such as web browsers, mobile devices, portlets, or other services generalizing information arriving from lower-level servers. The GIS interface is the main interface for the heavy IGIS user. The GIS interface is a software component for the visual representation of space data of various geographic digital formats and the subject area's objects stored in databases. It supports user–system interactions and provides the interface necessary for solving applied problems. The GIS interface allows the user to map the cartographic information in two-dimensional and 3D formats. Plug-ins are provided to get IGIS attuned to a definite subject area. These plug-ins accordingly adjust the specifications of the functional tasks being solved by the system.

SOA infrastructure incorporates the services for supplier–consumer interactions and the service backup. The service backup contains different components supporting the SOA mechanisms' normal functioning. It is intended for developing the SOA infrastructure for further deployment of technical services, including five standard functions: directory, security, management, orchestration, and semantics. A detailed description of these functions is beyond the scope of this chapter.

The service suppliers define the service query format and publish it for detection and multiple uses. The following entities may be included by the service suppliers in IGIS:

- cartographic information server;
- hydrometeorological information server;
- object server;
- IM;
- simulation and mathematical model server;
- administration (management) server;
- interaction with external systems server.

The services in IGIS can be divided into three main groups:

1. *Application services*. Application services provide data and control parameters for organizing common work in one unified computer space. In our case, we are operating with many heterogeneous applications in different layers: one workstation, local network, or global network.
2. *DBs, KBs, and ontology services*. DBs and KBs and ontology subsystems require many kinds of services. A DB system is heterogeneous and as a rule is distributed inside a local or global network. Therefore, for every part of a DB, a particular kind of service should be selected. The DB environment forms unified data, information, and knowledge spaces. This is a key issue in developing large-scale practical applications and systems in IGIS.
3. *End user's services*. The basis for the end user's services is a web interface, except for real-time applications.

Let us consider some of these groups in detail.

2.2 Application Services

Application services realized in IGIS can be conditionally divided into subgroups based on these service suppliers.

Cartographic information servers provide the following applications:

- accessing spatial data in different formats, including Shape, S57, SXF, OpenStreetLayer, and VPF;
- geodesic functions for a given projection and Earth model;
- visualizing cartographic data;
- importing/exporting spatial data to/from different formats, etc.

Administration (management) servers provide the following applications:

- resource distribution;
- managing access for the user and other services;
- managing the IGIS tuning and operating modes;
- IGIS operation journaling, etc.

Hydrometeorological information servers provide the following applications:

- selection of and attuning the channel to receive hydrometeorological information from a source via transmission control protocol, file transfer protocol, and e-mail.
- filtering the flow of recipient information in accordance with the required types of hydrometeorological weather reports;
- sentence analysis of the recipient's hydrometeorological information reports according to the international (World Meteorological Organization) and regional code forms;
- decoding and storing the values of given meteorological parameters;
- controlling the server's given operation processes.

Modeling and mathematical models' servers provide the following applications:

- universal time;
- mathematical problems (e.g., search theory, radio location, hydroacoustic);
- multilevel data processing (Popovich and Voronin 2005);
- simulation;
- 3D representation of results;
- object recognition and classification as well as tactical situations.

Interaction with external systems servers provide the following applications:

- receiving information from external mobile objects and systems (e.g., transportation monitoring systems, permanently functioning network of oceanographic stations based on drifting gauge-buoys [ARGO project], and others);

- receiving locations of sea and river vessels in the World Ocean based on automatic identification system data;
- receiving locations of aircrafts based on the data from Automatic Dependent Surveillance-Broadcast system's transponders;
- receiving locations of spacecraft and calculation of their coverage zones based on open-source data and a number of other services, as determined by a definite subject area.

The object server sustains the unified object model of the whole system. The main server's services provide access to the subject area objects and operations with the objects, generate objects based on the subject area metamodel, store the ontology, destruct objects, and change the objects' properties.

2.3 Database, Knowledge Base, and Ontology Services

The functions of providing data access are assigned to DBs, KBs, and ontology. The KB is an important component of any intelligent system. KBs provide the data necessary for the functioning of the IGIS ES component. The ES is a main reason why GIS can be called "intelligent." It is intended to seek techniques to solve problems in a certain subject area based on DB records and on a situation described by users. The ES can serve to solve two problems. The first problem is conventional for ESs: making recommendations in situations that present difficulties in decision making. The second problem consists of managing (controlling) the complex modeling modes. The ES core—IM, using, for example, the Rete (Forgy 1982) algorithm—can describe complex concurrent processes by sets of simple rules while providing high levels of managing (control) and modeling efficiencies. The ES uses ontology services in addition to DB services. Ontology is intended for overall and detailed formalization of the IGIS subject area by using certain conceptual schema. Interaction between the KB and ontology (Gruninger, 1995) objects is a "frame" intended for representation of concepts and their links (relationships) within the application's subject area. The universal IGIS should load the various scenarios in the KB and, thus, be attuned to different subject areas. DB services and ontologies are consumers of BD services that furnish information for their functioning.

As a rule, information in IGIS is divided into the following groups:

1. Cartographic data
2. Object data
3. Data received from external sources
4. ES scenarios

Because each of the above groups is intended to realize different functionalities, the requirements each group is expected to meet are rather diverse. Cartographic data are aimed at the user-oriented visual representation of information in the form of geographic maps (Figs. 2, 3). The map is generated based on data sets. Several sets may be used while solving certain tasks, such as sets containing sea and ground maps, vector graphs of satellite shots and drawn roads, and topographic maps with superimposed height nets. The sets are divisible by types, including vector, raster, and regular value matrices; each set has its own format.

Object data contain knowledge about the dynamic information model of the subject area. The dynamic information model in the considered case is organized in accordance with a certain rules system totality of situation elements; at each time instance, it contains data matching factual parameters of the objects and medium of their functioning (Popovich et al. 2005).

2.4 End User Services

IGIS provides the end user with a great number of services, such as:

1. Access to cartographic information
2. Access to hydrometeorological information
3. Mathematical problem modeling and solving (in theory of search, hydro-acoustics, radio location, etc.)

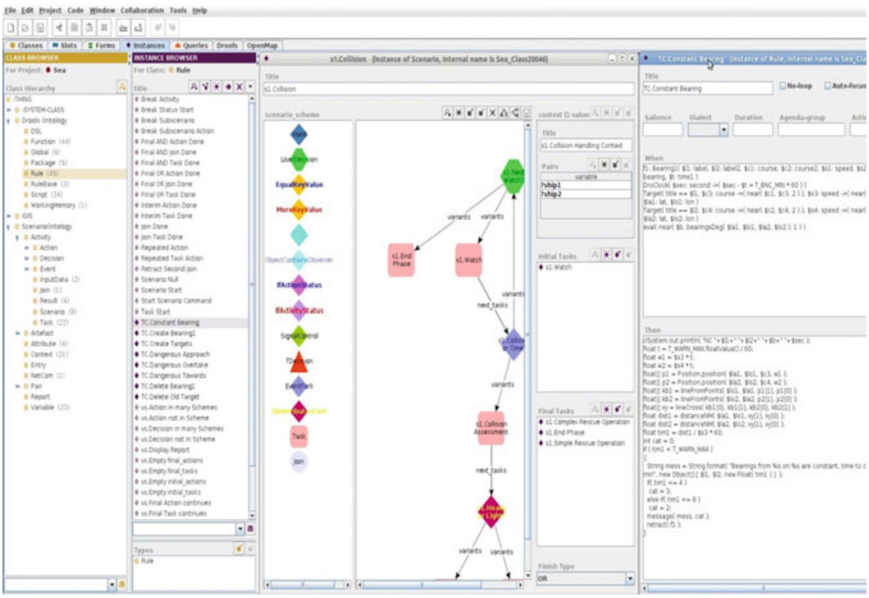

Fig. 2 Visual medium for scenario development

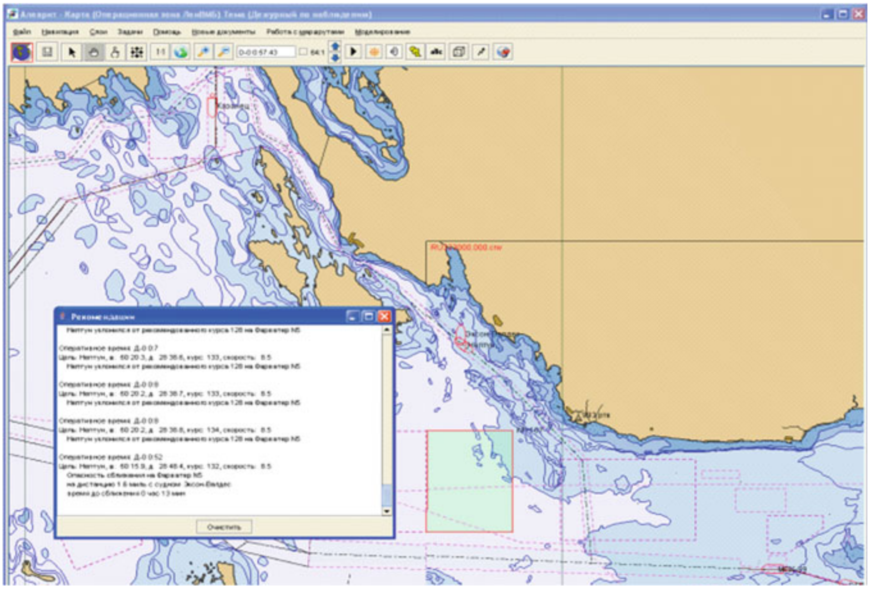

Fig. 3 An example of expert system operation in the form of interactive prompting

The dynamic information model of the subject area underlies all users' services. As mentioned, the model provides access to all objects of the subject area and allows one to solve all IGIS applied problems. Access to the model elements is performed through an aggregate of users' services of the object server. The object server is specified by the following service groups:

1. Subject area universal description;
2. Inheritance mechanism, also including multiple inheritance of objects;
3. Separation of an object's stationary and transient data;
4. Universal mechanism of relationships;
5. Access to the history of the objects' property states;
6. Information filtering for different user groups.

The subject area universal description describes objects within the uniform model of data representation from unified standpoints. The subject area universal description by default assumes access to the objects' classes with an indication of their peculiar properties along with their values. It also lists possible object states based on business process requirements, expressed through the values of an object's congruent properties and through the list of relationships, binding the objects while realizing the business function.

A variety of problems/tasks from different subject areas that are being solved by IGIS may involve the same information model's objects; this fact requires different object behaviors. Therefore, while solving tasks of this kind, the object can be classified by different attributes. The mechanism of multiple inheritance lets such objects be described in the information model.

Different behaviors of the same objects in different tasks/problems leads to one more issue. In the simultaneous modeling of different business processes, the properties of the same object may take on different values. The mechanisms separating constant and transient data avoid the conflict of simultaneous altering of the same object properties. The services accessing ontology objects and the topic objects are responsible for this task solving.

The universal mechanism of relationships describes the complex loaded links between different objects of the information model that arise while realizing business processes.

To analyze the information model's dynamics and characteristics of its state change in-depth to separate object properties and (when needed) to restore the state to the given time instance, it would be necessary to use the mechanism storing the history of object property states.

GIS information resources may be available to a great number of users of different levels and access rights. In order to solve certain problems, the different user groups will need different information about the same processes and participating objects. Services filtering information for different groups of users are intended to confine access to the necessary information.

The overall aggregate of specific problems being solved by IGIS is constructed based on the above services of the object server.

3 Ontology

Ontology is not a new phenomenon in computer science, but it is often asked why ontology is needed for a particular project. Here, we present some arguments in favor of using ontology in IGIS:

- creating a uniform system of subject area notions and their relationships;
- forming a uniform space at data, information, and knowledge levels;
- providing a uniform space for the users, applications, and data (knowledge) bases;
- assuring the possibility of knowledge reuse and multiple use within the subject area;
- explicit formalizing of suppositions and assumptions within the subject area;
- separation of knowledge and live data within the subject area;
- knowledge analysis within the subject area.

Development of uniform information fields within the IGIS constituents, as well as within other applications, is the priority task of ontology approach implementation. Existence of the universal model of environment and uniform ontology systems significantly facilitates the use of different cartographic data sources, such as shape formats, C57, VPF, CXF, and others. The service system is also ordered and does not cause any conflicts at interactions between various applications or agents. A well-thought-out ontology system also simplifies the problem of data transfer between various models, because it does not need special mechanisms or format development.

3.1 Knowledge Representation by Ontology

Determination of subject area knowledge is an important property of the ontology approach. Here, the main point is that the knowledge system is separated from realization (i.e., from the software code), which is impossible in conventional OOA. Thus, quite a flexible system is possible, which allows for altering business logic without interfering with the substantial regulations in the subject area.

It should be noted that data analysis in the subject area is possible when the terms' declarative specifications exist. The terms' formal analysis is extremely valuable for the existing ontologies' reuse, as well as their extension.[2]

As a rule in computer science, the subject area ontology is not a target in itself. Ontology development is somewhat similar to the determination of data sets and their structure when intended for use by other programs. The idea itself is somewhat analogous to OOA and earlier approaches, such as the concept of data abstract types. When realizing the task-solving methods and developing software components, the use of KBs as ontology data is rather effective. This dramatically simplifies the development of software code and architecture of the subject area classes.

Notions such as classes, as realized in the ontology software system, become information objects and are transferred from the category of software code to the category of data that could be manipulated by IGIS, as well as by the interacting environment.

The proposed approach has many advantages:

- subject area object classes and their relationships are described as data and not as software code, which significantly increases the software system's flexibility;
- alteration of the class characteristics does not imply a necessity to change the system software code; consequently, maintenance of the class libraries is simplified;
- class libraries can be stored in DBs, which means that the system's DB structure becomes simpler and does not require any alteration to class descriptions.

A number of languages have been developed for meta-class descriptions. The best-known languages are briefly described here.

---

[2] OWL Web Ontology Language Guide URL: http://www.w3.org/TR/2004/REC-owl-guide-20040210/#import.

Consortium WWW (W3C) developed resource description framework (RDF), a language for knowledge coding. This technology was used for the first time in forming web-page content. Since then, the use of RDF has been extended. In general, it can be used to perform global navigation when searching for data or knowledge needed for software components, executing an information search, or realizing definite functions.

Based on the four-level meta data model and unified simulation language, the unified modelling language developed a special language, Web Ontology Language (OWL), which was intended for metadata exchange between various modeling tools via the Internet. This language allows one to describe any ontology of the subject area.

3.2 Open Technology for Ontology Development

Currently, the following systems for ontology management are in wide use: Protégé, SUMO (Suggested Upper Merged Ontology), KAON2 (The Karlsruhe Ontology and Semantic Web tool suite), LarKC, and CYC. The functional capacities of these systems are different, and it is somewhat difficult to single out uniform criteria for comparison. Each of these systems has some advantages and disadvantages, which could be applicable in solving a definite class of problems.

The Protégé system can be considered as an example. It is a free distributed ontology editor and a tool intended for KB development.[3] The system was developed by a research group in medical informatics at Stanford University. The system has two versions: a system based on graphs of descriptive logic (DL; Protégé-OWL) and a system of frame-based representation (Protégé-Frames).

Protégé, based on the frame model of knowledge representation OKBC (Open KB Connectivity), is intended for ontology development, visualization, and manipulation using various languages for ontology descriptions. Classes, instances, slots (describing the properties of instances and classes), and facets (specifying additional information about slots) are the main constituents of such a system.

Protégé has an open architecture that enables the extension of its functionality based on modules called plug-ins. It has an extension adapted for editing models stored in different formats (standard, textual, in DB JDBC, UML, languages XML, XOL, simple HTML ontology extensions [SHOE], RDF and RDFS, DAML + OIL, OWL). So far, more than 60 modules of the kind are developed for Protégé.

Protégé can be used as a self-contained subsystem. Conjointly with GIS, the system can play various roles, such as a subsystem in the subject area of business analytics, ES, and/or IM. In some cases, Protégé can be used as an IGIS kernel.

[3] Protege. www.protege.stanford.edu.

3.3 Ontology for Subject Domains

Even relatively simple IGIS applications can hardly be squeezed into the frame of a single subject area. Therefore, when dealing with a real application, a system of ontologies should be implemented rather than just one ontology. To understand how to determine the subject area list or the domain list, consider the principles of OOA (Booch 1995). When developing large distributed systems, division into several ontologies is an extremely important issue and a series of principles should be followed:

- all class sets of different ontologies must be of a disjointed nature;
- ontology development, editing, and modification should be performed by the subject area professional rather than the programmer;
- ontology interface systems should be altered only for exceptional cases (detection of conflicts and/or errors, system upgrades, etc.).

3.4 OWL and GML

The web ontology language OWL can describe the classes and their relationships that are proper for Web documents and applications. OWL ontologies may contain descriptions of classes, properties, and their instances. OWL allows for the correct determination of the terms declared in one application (system) in another system, independently of the systems' technical and work scenario specifications. Based on DL, OWL provides tools intended for the logical description of notion semantics (sense); therefore, the information can be coordinated for use by people as well as by applications in different information systems (Websites, DBs, ESs, DMSS, etc.). Additionally, the OWL language supports the complete computer processing of Internet content better than XML, RDF, and RDF Schema (RDF-S); therefore, it provides an additional terminology dictionary along with formal semantics.

SHOE was the earliest language for Web ontology representation. DAML + OIL (DAML + OIL, 2001) was an OWL predecessor; it used the integrated projects of DALM (DARPA Agent Markup Language) and OIL (Ontology Inference Layer) to form the upper level of languages intended for Web ontology description.

3.5 Data and Workflow

Each document in the subsystem of an electronic workflow can be considered as a separate information flow that has its own model of data representation. The component task consists of transforming the information contained in the documents into a form matching the geoinformation system ontology. This task solving is complicated by the fact that documents contain verbal or weakly structured information. Therefore, it is necessary to accurately interpret all introduced notions and to impart meaningful content to all objects and relationships. Weakness of structures enlisted in model development requires a rigid semantic mapping between the objects being operated with.

Signs of a document's verbal elements are determined to be two dimensional or static; the signs' combination can only be linearly ordered (more accurately, monolinearly ordered) and discrete (i.e., the signs in combinations preserve the separability property).

Signs of the structural components are determined to be one or two dimensional. Their combination should be discrete and polylinearly ordered; otherwise, the order and relationships between characters and signs are imposed by net, hierarchical, relational, and discrete parametric schemata.

Thus, the subsystem of an electronic workflow is intended for the documents' information transformation and its further use in geoinformation systems.

As a rule, the workflow subsystem is adjusted to a definite subject area in accordance with the developed system's purpose.

4 Scenario Approach

The scenario is one of the key IGIS notions. This notion describes various processes almost seamlessly as compared with, say, an algorithm. In a way, the scenario is a generalization of an algorithm.

4.1 Scenario Definition

In the spatial environment being modeled by GIS, various processes take place, including spatial ones (which are the most complex). The process scope encompasses natural process and various human activities. These processes can be complex and include interactions between many heterogeneous objects. The process complexity may be hierarchical, when more global processes are a result of conjoint running and mutual influence between separate, specific processes, which in turn are decomposed into more elementary processes.

Different events can be external manifestations of spatial processes in various points of space, as well as some changes in space.

The following events can be separated:

- appearance and disappearance of objects in different space points;
- motion of point objects on different trajectories;
- change of elongated objects in shapes and sizes according to different laws, etc.

For visual computer modeling, the spatial processes can possibly be described by scenarios.

Let us give some definitions:

- *Scenario* can formally be determined as a sequence of stages and decisions.
- *Stage* is an aggregate of elementary actions performed sequentially or concurrently.
- *Decision* is a point where the process flow may change its direction depending on certain conditions arising at the current moment. So, formally the decision is determined as an aggregate of branches (directions of scenario continuation). Realization of the decision-making procedure (i.e., a choice of direction for further scenario development) can be different: it can be automatic (program) or manual (including the human assistance).
- *Actions* are "construction" blocks for scenarios. They represent specific elements of the scenario participants' actions, which may have different realizations. Some events or changes in space may take place as a result of actions (Fig. 2).

Unlike algorithms, the actions belonging to some stage can be executed sequentially as well as concurrently.

Using mathematical modeling and simulation, it is theoretically possible to model processes with any desired accuracy. The reason is that any action could be described by its *particular scenario* whose elementary actions could in turn have their particular execution scenarios, and so on ad infinitum. Consequently, in scenario decomposition, it is important to determine the level of elementary action detailing. One of the OOA principles—namely, the redundancy depletion at the event or entity formal description—should be kept in mind.

The detailing level first should be determined by modeling objectives, not just by computer tool capacities. The detailing level of the constituents of separate processes should be almost the same; otherwise, the advantage of detailed modeling for some actions would be nullified by rough modeling for other actions. In other words, the same scale order should be considered. In mathematics, such an idea could be illustrated by infinitesimals of different orders.

In visual modeling, the scenarios of spatial processes are played back on a computer in real or arbitrary time. In GIS, such modeling results are displayed on the background of a map as different spatial events or changes. The results of multiple playbacks of scenarios in spatial processes could be used to elaborate their statistical estimations and substantiate decisions for definite action modes. The scenarios are used not just for modeling; they can also be extremely fruitful in formalization of the subject area's business logic and for realization of real-time systems.

4.2 Expert Systems

An expert system is an aggregate of the interacting software tools using expert knowledge (i.e., experts' or specialists' knowledge) to assure highly effective problem solving within a certain subject area. The previously mentioned software systems, as a rule, represent knowledge symbolically, can explain the reasoning of (study) processes, and are intended for the subject areas where humans need to be specially trained for years to acquire the necessary skills. For an ES, the development technology itself requires a specific form of interaction between the ES developer (usually called a knowledge engineer) and one or several subject area experts. The knowledge engineer "extracts" from the experts' procedures, strategies, and empiric rules, which are later used in problem solving and are embedded into ESs.

Standard composition of an ES integrated in GIS consists of the following main constituents:

1. KB, which contains facts (data) and rules (or other data representations) that use these facts as a basis for decision making.
2. IM (mechanism), which is used in the search for the process of problem solving based on knowledge from the KB.
3. User interface, which supports external interactions with the ES. Taking into consideration the ES integration into geoinformation systems, the user interface would be first and foremost the GIS interface.

Apart from these components, the GIS-integrated ES may include some other software tools, as determined by specific characteristics of the problems to be solved.

The KB is undoubtedly the major and most valuable constituent of the ES for IGIS. The process of development differs from the development of conventional software tools, determined by a specific characteristic of the special information type of knowledge.

At least five steps can be singled out in the process of transforming data into knowledge:

- *Knowledge is true*. The system containing knowledge trusts it completely: its statements are true. Should the system reveal certain faults in some of its data fragments, the fragments would immediately lose the status of knowledge and would be deleted or properly modified;
- *Knowledge is abstract*. Disregarding its empirical origin, knowledge is separated and extracted from the subject content and separated and removed from reality;
- *Knowledge is interpretable*. Knowledge tends to express its authenticity in some definite subject realization;
- *Knowledge is active*. Knowledge possesses an intrinsic capacity to be activated and to assign on its own will the states to conjugated information structures;

- *Knowledge is structured.* Knowledge is inside a stable frame originated by diverse structural units and levels.

4.3 Inference Machine

A number of IMs currently exist, with the Rete algorithm being one of the most well-known. The Rete algorithm was developed in 1979 by Dr. Charles L. Forgy at Carnegie-Mellon University, USA (Forgy 1982).

Historically, Rete algorithm-based IMs have been implemented in ESs. These systems are called "expert" because the experts' knowledge in some subject areas is represented there via certain rules. This knowledge mainly has an empiric nature and is accumulated based on the experts' experience in certain subject areas. At the rules' actuation, certain facts can be verified—say, whether the current distance between vessels has become less dangerous or if the distances can be calculated.

In IGIS, the ES may be used to solve the following problems:

- To acquire new information quality at different processing levels in accordance with the JDL model (e.g., measurements, signals, objects, traces, tactical situations, threats);
- To solve some complex hydrometeorological problems, such as prognosis, forecast, subject maps, etc;
- To assure a vessel's navigation safety;
- To manage tactical situations within the zone of shipping control system responsibility;
- To assist the decision maker with emergency management and many other tasks.

Within the frames of conventional use, the IM deals with *empirical* rules put into the ES DB.

However, Rete algorithms, along with empirical knowledge areas, can be successfully used in well-structured knowledge areas where rigorous mathematical theories containing axioms and formal rules exist. The Rete algorithm ideally matches to get complex parallel processes modeled. Processes of this kind are being studied by geoinformation scientists.

4.4 Inference Machine as a System Supervisor

To start, let us consider a comparatively simple task of monitoring navigation safety. To facilitate the given example, consideration of only the reciprocal vessels' collision, particularly the "high-seas version," will be used. This problem could be reduced to defining the distance between all pairs of vessels being monitored and comparing distances with the safe distance. Using standard programming, this task solving is performed by a dual cycle: externally according to all vessels and internally according to vessels yet in question. At a total number of 100 vessels, this requires 5,000 verifications in one iteration.

How often should such verifications be performed? The answer depends on the current distance between the vessels in pairs and their approach speed. In close proximity, even a verification each second might be insufficient; however, verifications might be done once an hour when vessels are far apart (say, hundreds of miles). Nevertheless, an algorithm based on a dual cycle should work at a frequency to assure the navigation safety (i.e., with a maximal frequency). Thus, such a solution is as ineffective as the case of the IM linear realization.

Using an IM based on the Rete algorithm, this task could be optimally solved by one rather simple rule. The verifications' frequency would functionally depend on current distances between vessels and their speeds. The above rule in CLIPS language is as follows:

```
(Rule 1)
(Current – Time ?cur.)
(Check-Dist. (vessel1 ?v1)(vessel2 ?v2) (time ?time) )
( test ( < ?time ?cur. ) )
=>
( if ( < ( Distance ?C1 C2 )( Foul-Distance ?v1 v2 ) )
then
( assert (Risk of collision (vessel1 ?v1)(vessel 2 ?v2) )
else
( assert (Check-Dist. (vessel1 ?v1)(vessel 2 ?v2)) (time
( + ?cur. ( min - Heading in - time ?v1 ?v2 ))) )) ) )
```

The rule content will be reduced to:

1. Fact (*Current time? cur*) is given each second (or with other maximal frequency).
2. Slot value (*time*? *tm*) for a fact. (*Check-Dist*) stores the astronomical time when it is necessary to run a regular verification of a distance between the given pair of vessels. Should the current time override the astronomical time, the rule would trigger.
3. At the rule triggering, the distance between vessels is verified against the fact of the distance becoming shorter than the dangerous distance. Matching functions are evoked to calculate the above distances.
4. If the current distance between vessels became shorter than the dangerous distance, then the fact alerting the collision danger is confirmed.
5. In the opposite case, the fact is confirmed that would evoke the distance retesting in a time specified by a function: *min-time-appr*.

4.5 Expert System as the Core of the Decision-Making Support Subsystem

Implementation of the ES as a tool for GIS-based managerial support and decision making assures the appearance of qualitatively new possibilities in modeling geospatial processes and in substantiation of complex managerial decisions. These new possibilities include the following:

- Fast recognition of typical situations in the applied subject area and a well-grounded offer for an operator (user) of the matching actions;
- Assurance of the operator's (user's) self-control of actions and decisions in regard to the management of applied subject area processes;
- Monitoring the state of controlled processes in the subject area over different criteria in real time;
- Arranging the information and reference assistance for the operator (user) at the stage of adaptation to IGIS functional capacities;
- Intelligent analysis of spatial-time activity of heterogeneous moving objects;
- Visual development of functioning models (scenarios, actions) for the objects in GIS;
- Playback of objects' action scenarios in real and arbitrary times as accompanied by visual mapping in the form of conventional signs on an electronic map background;
- Recommendations for decision makers in scenarios playing at the stage of research design for systems, business games, situation analysis, personnel learning, and training.

ES is used as a standard tool for a decision maker's intelligent support and as a system intended for managing various processes, such as modeling.

4.6 Case Studies of the Scenario Approach

The results of ES work based on the scenario approach can be represented to the user as interactive promptings or dialogues within a specified research area. Figure 3 depicts an example of an ES composed of interactive promptings. Obviously, it would be hardly possible to visually distinguish the ES work from the script language. However, the advantages of the proposed approach become obvious when accounting for the fact that scenarios and DBs can be developed in a visual medium using the full power of OOA. Figure 3 depicts the tactical situation of modeling special actions within a closed area; resulting from the previous situation analysis, the ES embedded in GIS generates prompts for decision makers (Fig. 4).

These prompts could be accompanied by voice and some other effects, such as visual cues. Luckily, the scenario allows the realization of practically any functionality and logic for making the most complex decisions based on first-order predicate algebra.

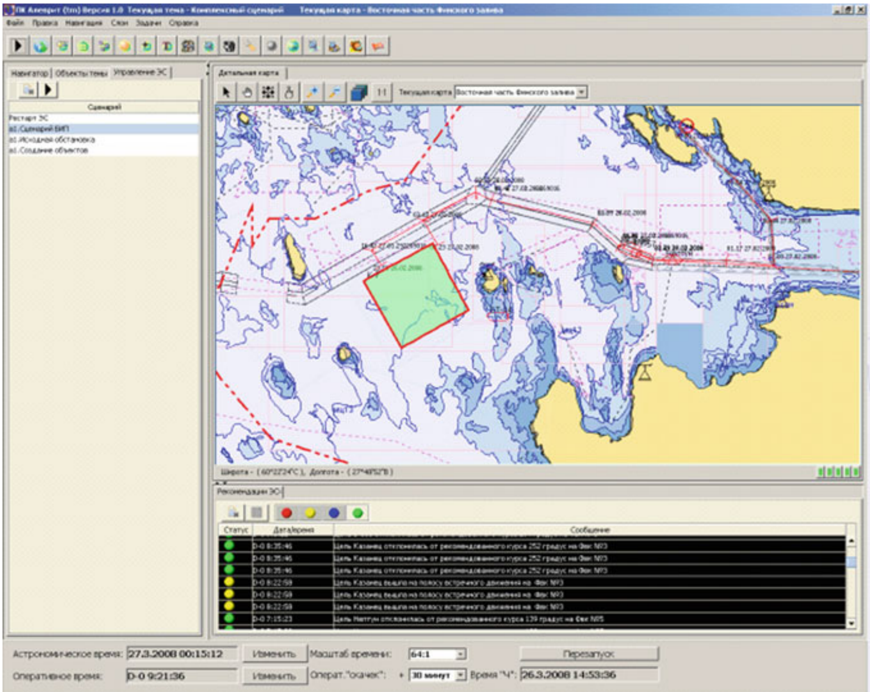

Fig. 4 The general structure of a DMOS based on a JDL model

5 Case Study

As an example of the ideas proposed in this chapter, the realization of the maritime surveillance and recognition systems (MSRS) developed by SPIIRAS and SPIIRAS-NTBVT Ltd. will be discussed. The given system architecture was developed based upon the concept of data harmonization, integration, and fusion in IGIS (Popovich and Voronin 2005, Popovich 2009).

The process of information harmonization assumes a distribution of main notions (concepts) and their relationships (ontologies) over the matching subject areas and/or responsibility areas. For instance, the partition could be executed over existing knowledge areas, such as hydroacoustics, radiolocation, theory of search, etc.

Information integration allows the fusion of information from heterogeneous sources and access to information resources for the current task being solved (modeling). The distinguishing feature of information integration is the fact that the result is always aimed at solving a definite class of problems.

In the literature (e.g., Blasch 2002), the notions of data fusion and information fusion are often separated. Data fusion is understood as the combination of organized data intended for analysis and decision making, whereas information fusion is the combination of data intended for knowledge acquisition or as a process of fusing data from heterogeneous sources. The concept of “data fusion” was introduced in the beginning of the 1990s (Blasch 2002). In IGIS, fusion is understood as the appearance of a new quality of information rather than an improved quality of information (see Fig. 5).

MSRS is used for information, model, and graphic support while typical functional tasks are being solved by managing officials at all situation coverage levels. The system is an information base for an integrated automated control system at all control levels and stages for different customers: military, emergency management, vessel traffic managers, regional governments, etc.

MSRS is capable of information and technical integration with necessary information sources, observation stations, data acquisition, and processing stations at horizontal and vertical levels. An example of a visual representation of the air, surface, and coastal situation for the Northern Baltic Sea is shown in Fig. 6.

The main functions of MSRS include the following:

- Integrated information space generation for specific tasks and analysis of the current situations. An example of a visual representation of navigation situation based on the S57 format is shown in Fig. 7.

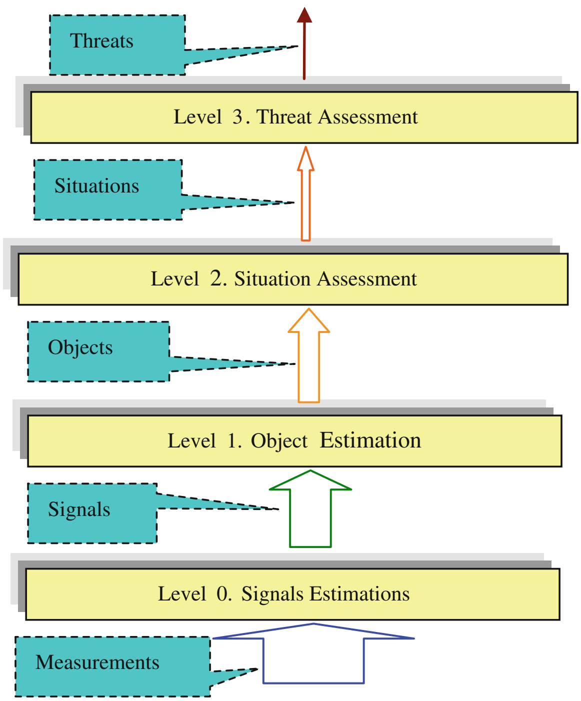

Fig. 5 Information fusion levels

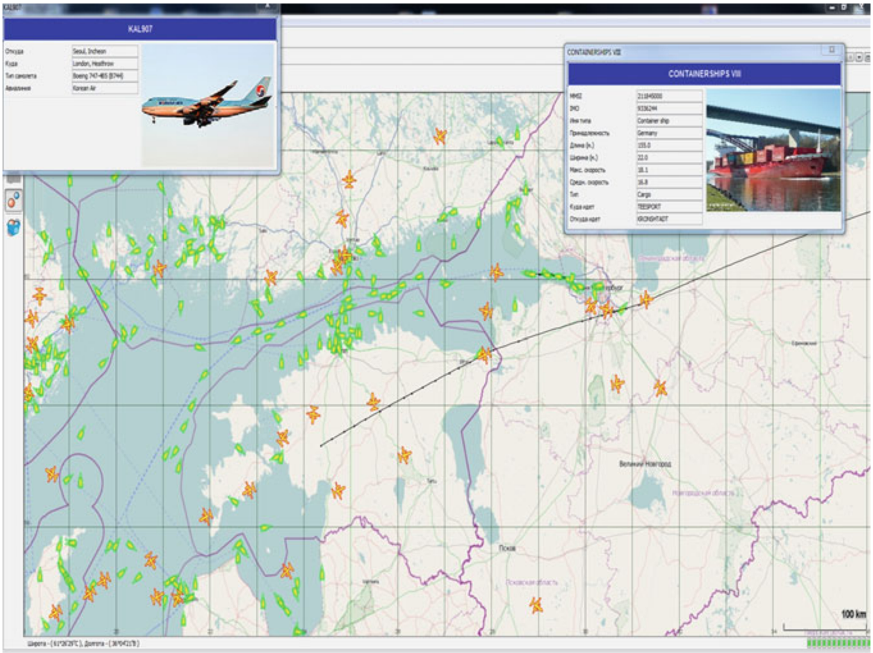

Fig. 6 An example of a visual representation of air, surface, and coastal situations

- Processing of heterogeneous information in a unified coordinate environment at the data entry pace, as well as its timely distribution and transfer to consumers at a desired pace and scope (see Figs. 6, 7).
- Continuous monitoring of moving targets, assessment of their state and environment (including weather conditions), and elaboration of summary data about the current operational situation in the zone of responsibility of the typical observation station.
- Early detection of dangerous targets, situations, and threats, as well as intelligent support of decision making using the scenario-based approach.
- Generation of recommendations on the adoption of adequate measures for strategic, operational, and tactical levels to prevent potential threats (weather, ice conditions, the threat of terrorists and pirates), based on the current situational data.
- Management of mobile surveillance equipment for operational data validation and refining.
- Simulation for assessing and predicting situations during the planning phase for optimal decision making.
- Surveillance system optimization at the construction and functioning stages, based on simulation and modeling using means and methods of artificial intellectuality.

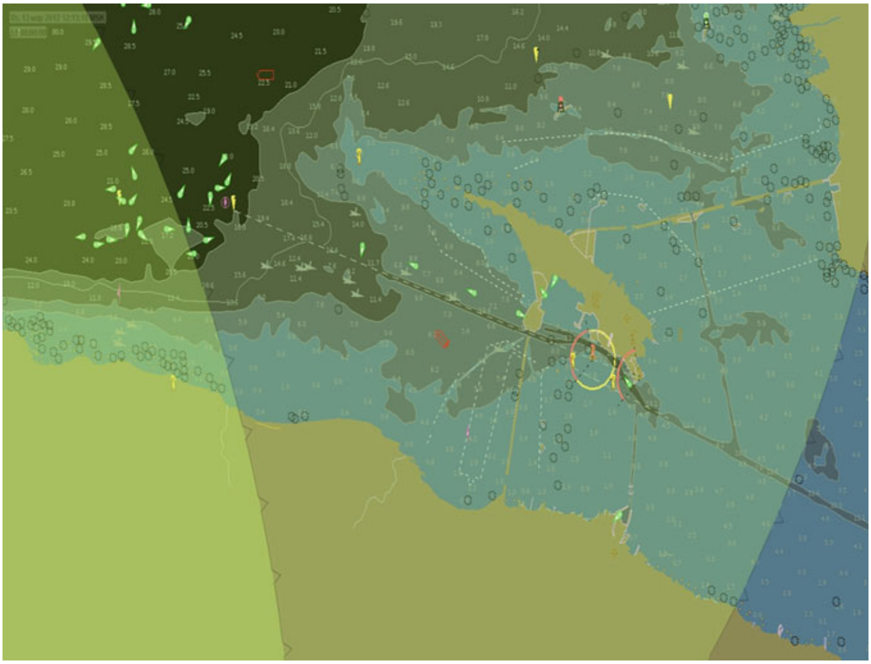

Fig. 7 An example of a visual representation of a navigation situation based on the S57 format

The system is capable of real-time (information update cycle: 1–5 min) tracking up to 100,000 moving targets. It is possible only by using IM. The scenario approach allows one to change business logic on the fly while the system is working.

Let us consider a very interesting and important subsystem of MSRS—the sonar calculation subsystem (SCS) (Popovich et al. 2009). SCS provides calculations (modeling) of acoustic fields for the estimation of sonar range, calculation and simulation of observation zones, and tactical situation conduct based on sonar data. The results of these calculations allow one to resolve a number of challenges related to the optimal placement of a positional sonar's optimization search operation, such as fishing, searching for sunken objects, searching for moving underwater and surface targets, etc.

The IGIS interface allows a 3D + t picture of transmission losses (see Fig. 8) and target detection zone distribution on a digital map. ES and IM help end users to understand current tactical situations and make the right decisions. MSRS and SCS contain different sets of ontologies, but they work together under the system of scenarios and sets of mathematical and simulation models.

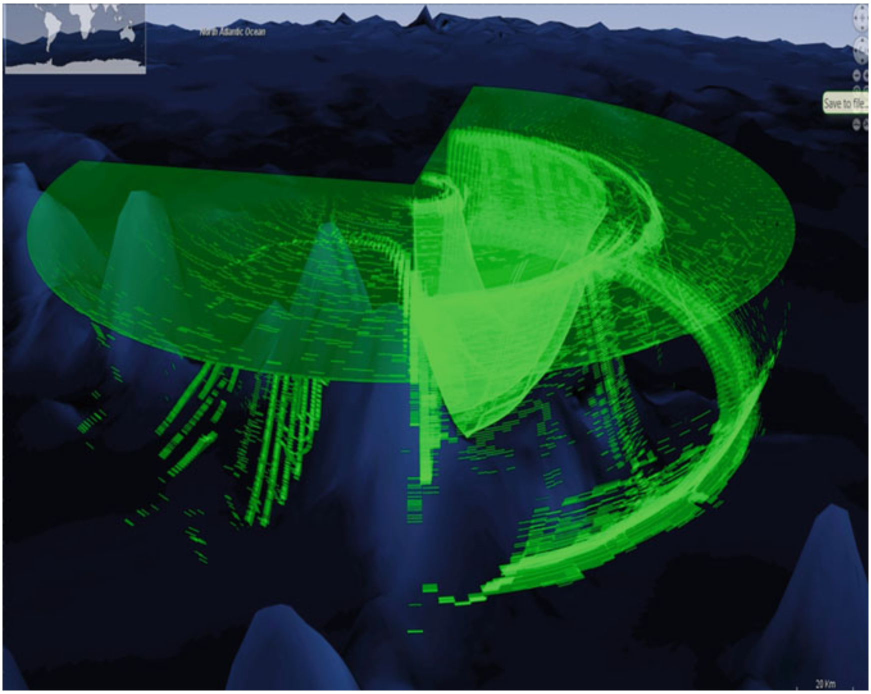

Fig. 8 3D visual representation of sonar's capabilities

6 Conclusion

The ideas presented in this chapter regarding IGIS could be joined in a specific subject domain. IGIS theory is one more attempt to fight complexity. This complexity is considered in two aspects: software development complexity and the complexity of decision-making support in some given location at given time. The proposed approach conditionally could be divided into three interrelated constituents:

1. Universal model of geospatial data: The core of the new approach in geoinformatics that allows for abandoning the old chapter maps heritage.
2. An intelligent subsystem that incorporates at least a KB using the ontology approach and an IM using algebra of first-order predicate logic.
3. A scenario approach based on the concept-generalized notion of algorithms. This approach opens up a wide variety of possibilities in formalization and automation of the subject areas' business logic.

Practical approbation of the proposed theoretical and technological ideas realized through serially developed information systems (heterogeneous and distributed) encourages the realization of the given ideas in a number of subject areas related to geoinformatics and software system development in general.

References

Blasch E (2002) Fundamentals of information fusion and applications. Tutorial, TD2, Fusion

Booch G (1995) Object solutions: managing the object-oriented project. Pearson Education, New Jersey, ISBN 0-8053-0594-7

Brooks F (1975) The mythical man-month. Addison-Wesley, Cambridge, ISBN 0201835959

DAML + OIL (2001) Reference description URL: http://www.w3.org/TR/daml+oil-reference

Edward Yourdon (1989) Modern structured analysis. Prentice Hall, New Jersey

Forgy C (1982) Rete: a fast algorithm for the many pattern/many object pattern match problem. Artif Intell 19(198):17–37

Gruninger M (1995) Methodology for the design and evaluation of ontologies. In: Proceedings of the workshop on basic ontological issues in knowledge sharing, IJCAI-95, Montreal

Kant I (1889) Kritik der reinen Vernunft. 1781

Popovich VV (2003) Concept of geoinformatic systems for information fusion. In: Proceedings of the 1st international workshop on information fusion and geographic information system, St. Petersburg, 17–20 Sept 2003

Popovich V (2009) Concept for corporate information systems based on GIS technologies development. In: Proceedings of IF&GIS-09, St. Petersburg, Springer//Lecture Notes in Geoinformation and Cartography, 17–20 May 2009, p 371

Popovich V, Voronin M (2005) Data harmonization, integration and fusion: three sources and three major components of geoinformation technologies. In: Proceedings of IF&GIS, St. Petersburg, 25–27 Sept 2005

Popovich VV, Shpakov MV, Shaida SS (2004) Ontology modeler for geographic information systems applications. In: Proceedings of 10th EC-GI&GIS workshop. Warsaw, 23–25 June 2004

Popovich VV, Potapichev SN, Sorokin RP, Pankin AV (2005) Intelligent GIS for monitoring systems development. In: Proceedings of CORP2005, University of Technology Vienna, Vienna, 22–25 Feb 2005

Popovich V, Aviloff C, Ermolaev V, Zinyakov Y, Leontev Y, Korolenko K, Popov O (2009) Sound propagation modelling on intelligent GIS base. In: Papadakis JS, Bjørnø L (eds) Proceedings of UAM 2009, June 21–26 2009, Nafplion

Popovich V, Vanurin S, Kokh S, Kuzyonny V (2011) Intellectual geographic information system for navigation safety "Systems", vol 25(8), pp 29–31, Aug 2011

Sorokin R (2005) Application of artificial intelligence methods in geographic information systems. In: Ivakin Y, Sorokin R (eds.) Proceedings of international workshop information fusion and geografic information systems (IF&GIS2005), St. Petersburg, 25–27 Sept 2005, pp 105–114

Tarski A (2004) Life and logic. Anita Burdman Feferman; Solomon Feferman, Stanford University, California. Hardback. ISBN: 9780521802406

http://www.springer.com/978-3-642-31832-0

Information Fusion and Geographic Information
Systems (IF&GIS 2013)
Environmental and Urban Challenges
Popovich, V.; Claramunt, C.; Schrenk, M.; Korolenko, K.
(Eds.)
2014, IX, 324 p. 152 illus., 33 illus. in color., Hardcover
ISBN: 978-3-642-31832-0