Approaches to Representation of Knowledge of Operations on Spatial Data in Monitoring Applications

Ksenia V. Raevich(1), Yury A. Maglinets (1), Ruslan V. Brezhnev (1)

(1) The Institute of Space and Information Technology SFU, Krasnoyarsk

Abstract. Article is devoted to a research of a problem of representation of diverse knowledge and data on spatial objects and ways of operating by them at the solution of monitoring problems under control of the multi-purpose system of remote monitoring of ISIT (MpSRM). The review is presented classification of knowledge which systematizes knowledge of semantics, syntax, behavior of a monitoring object and the rule of their interpretation. The approach is considered to representation of knowledge allowing to formalize monitoring problem definition.
Keywords. Remote sensing of Earth, problem definition, ontologic approach, knowledge base, object of monitoring, model of knowledge.

1 Introduction

The current trends of the continuing increase of the scope, quality, availability of data incoming the processing centers from satellite ERS systems, development of distributed storage and access facilities, as well as principles and algorithms of spatial data processing shall potentially allow expanding the solvable applied tasks of the ground facilities' remote sensing and monitoring and increasing the qualitative values for their solution. The above factors create the "Information Blow-up" in RS data processing, such blowup being known in the information science since 70s: data are available; data processing tools are available, but since the satellite monitoring tasks are knowledge-based in many respects, the availability of experts in the data processing chain becomes a barrier to their solution. This results in a number of operations between generation of the end user/task definer query and solution receival, such operations being both automatic or requiring engagement of experts in spatial data processing algorithms which leads to both time and process expenses.

Recently some scientific teams have come to significant achievements as regards generation of the process cycles to solve the ground monitoring tasks with the known solution algorithm. Thus, some necessary technology elements were stated in Article [1]: sources and archives of ERS data, systems of primary and issue-related data processing to create the information products, special-purpose system interfaces, data presentation tools. Development of approaches detailed in [1, 2] indicates the change of interface buildup concepts from "end user - expert" to "end user - decision system". However, the current automated solutions targeted at the end user do not, as a rule, have the appropriate flexibility as they are focused on tasks' solution in the narrow statement and changes of the task characteristics make it necessary to engage experts. Thus, the knowledge engineering techniques are of immediate interest as they allow isolating the expert knowledge with the expert being taken-out from the data processing chain. The article states the main knowledge categories to be studied as part of the object-related approach to ground RS and monitoring systems, such approach being developed by the authors.

2 Main Characteristics of the End User Query to the Satellite Monitoring System

Interactive cooperation between the "Task Definer" actor (further the Definer) and the automated monitoring system (further the System) is reviewed. The key target of the Definer is to establish the remote surveillance of the ground object the Definer is interested in (further the monitoring object, MO). Thereat the Definer shall state its surveillance target and, in a while, wait for the result or cumulative results.

The Definer target in the simplest case is to see the one-time slice representative of the SO status. This is a task to solve the information query (elementary measurements). This task is methodologically important as any surveillance task as complex as need be may be solved by the set of elementary measurements and further processing of their results.

More generally the Definer target is to set the Object surveillance and routinely trace the SO changes in time. The system's role in this process is to inform the Definer of the changes. Thereat the degree of the change importance must be determined in order not to miss the important change, on the one hand, and, on the other hand, not to initiate the communication at routine events. Thus, the monitoring task is addressed.

When the Definer has not only the surveillance tools, but also the status-correction means, the Definer may generate the task of the automatic or automated SO management.

The System, in its turn, is targeted at determination of the required measurement rate and information sources; definition of deciphering features of the aspects for the Object to be found on the image with reference to each source; identification of the correspondence between the numeric values as measured on the image and the respective properties of the MO which shall be identified under the Definer query.

Thus, the key task setting parameters to be clarified interactively between the Definer and the System are:

- Name of the Object (type of objects) and its identities (properties) as known to the Definer for the Object identification by the System among the variety of the surrounding objects;
- Structural properties of the Object as known to the Definer (when significant for identification);
- Relations between SO and environment as known to the Definer (when significant for identification);
- Properties of the Object to be defined;
- Structure of the Object to be defined;
- Relations between the SO and environment to be defined.

3 Conceptual Model of the Satellite Monitoring System Knowledge

Two knowledge categories have been conceptually identified: MO knowledge; system knowledge.

3.1 MO Knowledge

MO Knowledge may generally be expressed as below:

$$O = < S_O, R, P, \Phi >, \quad (1)$$

where O is a monitoring object, S_O is a set of the Object specialties; R is a set of binary relations with other objects formalized in [5, 6]; P is a set of heterogeneous properties of the Object whereat S_{Oi} is characterized by the subset $\{p_{1i}, p_{2i}, p_{3i}, \dots, p \quad \} \in P$; Φ is a set of interpretation rules associated with the principles and algorithms of the Object features' computation in one or another specialty.

3.1.1Knowledge of the SO semantics

A descriptor being a future advance in approach (as detailed in [7]) was suggested to align the Definer mental model with the internal System knowledge representation. Statement of tasks is characterized in the below categories:

- Taxonomy of domain concepts. One or more taxonomies correspond to the specific domain, each taxonomy setting the SO hierarchical classification from a certain position. Specific element (node of a taxonomic tree) may have various specialties.
- Properties of taxonomy elements. Specific element is described by the set of properties indicative of the respective aspect of the Object. Non-representational (class of properties) and certain properties are identified, each likely matched with the range of values, type of a measurement scale, principles of measurement and possible interpretations.
- Relations between the taxonomy elements. Specific pair of elements may be described by the set of binary relations over it.
- Domain Aggregations. Describe the allowable structural combinations of the Object in question as collections of components (elements, parts).

3.1.2Knowledge of the SO behavior

Comparison of SO specialties presented by the set of the properties to be measured, each (specialty or property) being in correlation with the relevant concept taxonomy node and model of the estimated path of the property value being changed in time, peculiar for this SO specialty.

3.1.3Knowledge of the SO syntax

Refer to the collection of the key peculiarities (features) of the Monitoring Object as seen on the image:

– structural peculiarities,
– shape features,
– brightness and textural properties as seen in one or another range or a set of EM-spectrum ranges,
– spatial attitude,
– time-to-time variability (in view of each of the above peculiarities).

3.1.4Knowledge of the interpretation rules

Knowledge of the interpretation rules being indicative of the semantic values based on the syntactic marker measurements and characteristics of the initial image.

3.2 Managing (System) Knowledge.

Regulates the ERS data conversion control. The variety of procedures and algorithms support the great many various stages of data processing. It is important to identify the classes of one-function procedures for the classifier being configured in this set, such procedures, in their turn, may be subdivided to atomic actions, if required. System knowledge may be subdivided into the following sub-classes:

3.2.1Knowledge of the Data Conversion Algorithms

Knowledge of the Data Conversion Algorithms may be set in a declarative manner by lists or relational structures [3, 4], where interrelations between algorithms, tasks and rules of decision are taken into account thus allowing to formalize the task context, ways of its solution, description of inputs, principles of the interim and final results' representation and so on in the form of individual taxonomy.

3.2.2Knowledge of the Tools for Interaction with End User

Knowledge of the Tools for Interaction with End User describes the principles of the task assignment by the end use, principles of internal task interpretation and principles of the solution results representation. These principles are aimed at description of the structure of "end user - decision system" interaction. Information provided by the user when setting the task shall be of the essence when selecting the specific algorithm or set of algorithms at a particular time.

3.2.3Knowledge of Principles and Algorithms of Data Conversion

Knowledge of Principles and Algorithms of Data Conversion. In the context of the user information query resolution it is worth noting the set of algorithms of different difficulty and scope which are gathered in the respective structures united by the logics of the input conversion and input requirements. Algorithmic chaining shall be in the hands of the expert responsible for the data processing and analysis. Such expert shall also associate (assign the R set for each algorithm) the realized chains with the typical tasks of end users.

3.2.4Knowledge of the Data to be Converted

Knowledge of the Data to be Converted. The System shall be capable of searching the relevant and actual data for successive resolution of the user information query. Such data include satellite images with significant scattering of spatial resolution as well as radiometric, spectral and time resolution. Therefore, restrictions for the possible ERS data sources must be identified for various classes of tasks.

4 Implementation of the Suggested Approach

The initial design of the knowledge base shall be based on two major activities: 1) preparation and structuring of the Object knowledge O_i in the specialty of choice S_{O_i}, which includes attribution of the elements of R, P and Φ sets; 2) export of the ready structure to the data base. Software environment which supports execution of the above activities and includes such key components as Protégé OWL-ontology editor which builds up the RDF data structure describing the graph of objects and their relations in XML format; local web-interface which allows decomposing the XML data structures and export them to MySQL database, the structure of which is given on Fig 1.

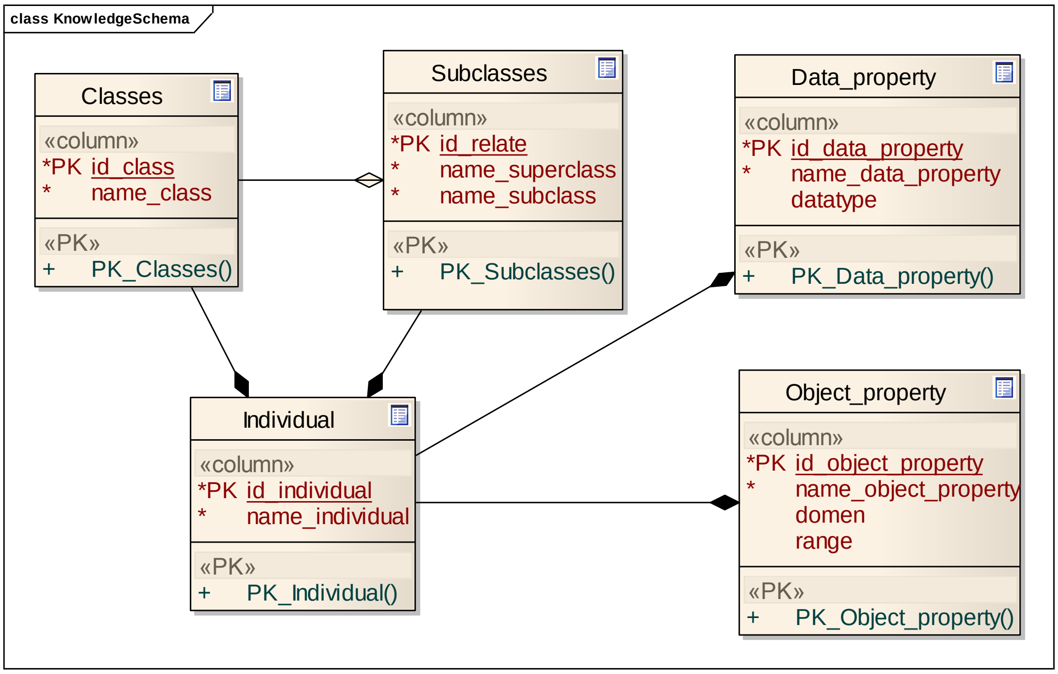

Fig 1. Database Scheme

Logical data structure includes such basic entities as Object Classes' Guide (Classes), Instance Guide for all classes (Individuals), Guide of Possible Intra-Instances' Relations (Object_property), Guide of Possible Object Properties (Data_properties), description of Class-Subclass relations (Subclasses).

The reviewed concept of the knowledge structure representation found its practical use in the automated agricultural monitoring system created by the authors [8, 9]. Data and knowledge gained in this sphere make it possible to bind the spectral and soil parameters of the Object, spatial and topological, process, economic and other groups of its properties. Such links allow responding to a number of current and future user queries (Fig. 2).

Queries can be divided into occasional and cyclic. Occasional queries are, as a rule, commands to be immediately performed. Metric calculations, for instance, area, perimeter and etc., set the example of such queries in MpSRM.

Cyclic queries imply the regular queries to change the set parameter for the given Object thus complying with the monitoring tasks. The typical examples may be monitoring of the vegetation structure, surface temperature, moisture and etc.

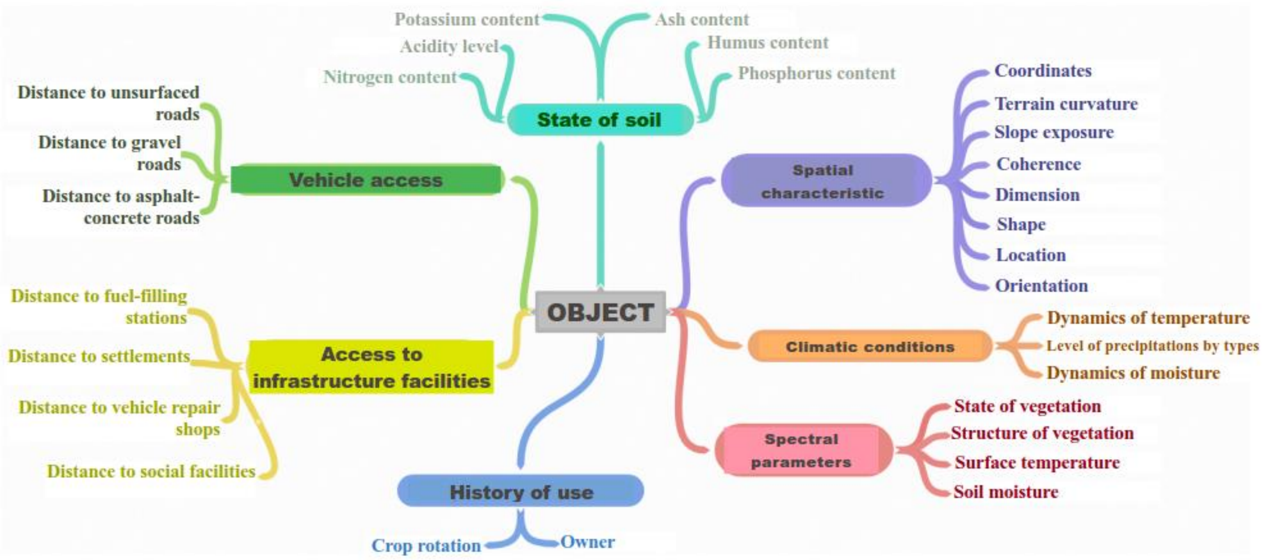

Fig 2. Typical queries of users to the decision system on the example of the agricultural land appraisal (AL)

As an example, let's consider the monitoring task for the vegetation heterogeneity within the agricultural contour, which model may be given as an expression (2):

$$\mathrm{OM} = P_{\Pi}, NS, NP, T, N, Tm, K, E, \varphi, F > \quad (2)$$

In semantic space the monitoring Object is defined by the crop $k^i \in K$ typified by the model of its development in time Tm, such model being determined by the sequence of phenophases $\varphi = \{\varphi^i\}$ and Plan of agrotechnical activities as given in (2) as an aggregate of events $E = \{e^i\}$.

A variety of Tm time ranges for the change of the Object F states is descriptive of the Object behavior and target path of its lifecycle.

Syntactic description includes a metric set describing the form (a set of coordinates P_{Π}, thickness T) and dimensions (area NS, perimeter NP) of the Object as well as the spectral NDVI index.

Rules of averaged NDVI values' interpretation – $\check{N}$ allow to compare the status $f^i \in F$ of each heterogeneous region w_j and the Object in whole [8, 9].

Knowledge base area representative of the bunch of taxonomy elements and their correlation when solving the given task is shown on Fig.3 as a direct graph.

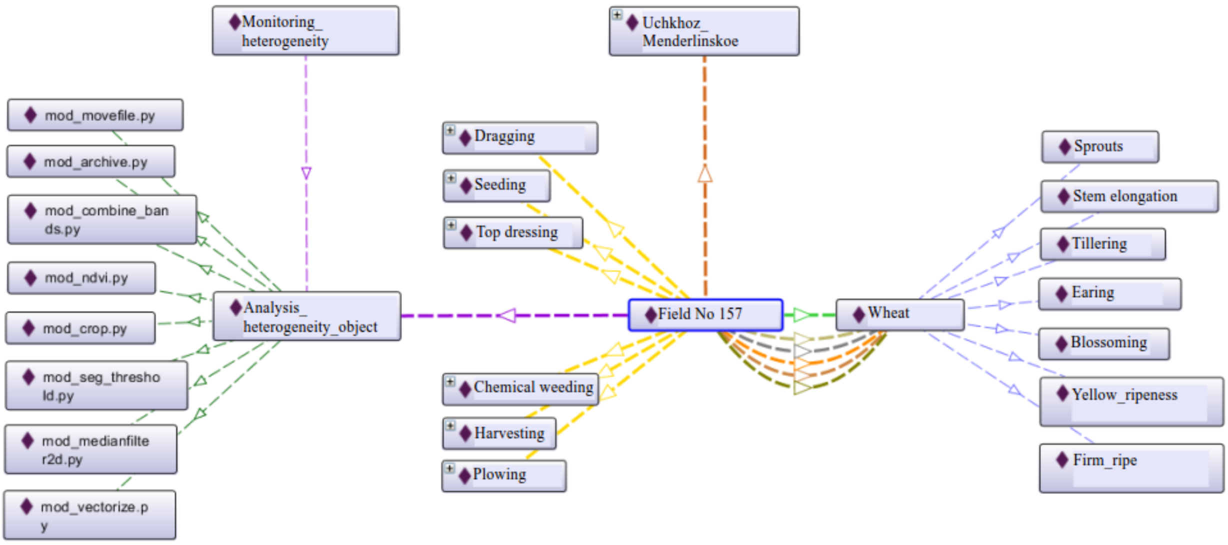

Fig 3. Correlation of the objects of MpSRM ISIT knowledge base when solving the tasks of the vegetation heterogeneity monitoring

The central graph node is "Field No 157" which belongs to the agricultural enterprise, "Uchkhoz Minderlinskoe" LLC. Wheat was grown in that field from 2013 till 2018 which is shown by a set of arcs connecting the "Field No 157" object with the "Wheat" object. Agrotechnological activities, such as dragging, seeding, top-dressing and so on, are executed at the "Field" object. Instances of the "Agrotechnological Activities" class are defined by the recommended time execution interval.

Wheat is characterized by the set of phenological development stages (sprouts, tillering, stem elongation and so on), which is shown on the graph by the "Wheat" object ties with the instances of the "Phenological Stage" class. Phenological stages, in their turn, are characterized by the reference time interval, duration and range of NDVI values [8].

Conclusion

The approach to the management of the satellite monitoring knowledge reviewed herein implies use of the suggested knowledge classification by setting the sets of concepts and relations which simulate the domain the user is interested in, as well as the all-system representations. The created knowledge base was tested with resolution of some typical SPARQL and SQL queries which simulated the information queries of users. Software module which is responsible for the "end user - decision system" chat and which allows defining the tasks of occasional or cyclic measurement of the set parameters in order to assess the state of the given object in various specialties is being prototyped.

The research was financially supported by the Russian Fund of Fundamental Research (Project No 18-47-242002 r_mk), Government of the Krasnoyarsk Territory and the Krasnoyarsk Territory Science Fund as part of the Scientific Project named "Development of the technology to build up the intellectual information systems for the object-oriented monitoring of the areas as per the RS data".

LITERATURE

[1] E.A. Lupyan, I.V. Balashov, M.A. Burtsev, V.Yu. Efremov, A.V. Kashnitski and others. Development of technologies to build up the remote monitoring information systems//Modern issues of the Earth Remote Sensing from space. 2015. V. 12. No5. P.53–75.

[2] Traynor C., Williams M.G. End users and GIS: a demonstration is worth a thousand words // Interactive Technologies. 2001. P.115-133

[3] Yu.O. Trusova Representation of “Image Analysis” domain in the form of the special thesaurus: Thesis research for Doctor of Science. Moscow, Institution of the Russian Academy of Sciences, Computation Center named after A.A. Dorodnitsyn, Russian Academy of Sciences, 2009. p.111.

[4] RF Patent Nos2003133170/09, 10.07.2005. Knowledge base for processing, analysis and recognition of images//Russia’s Patent No2256224. 2003. Bulletin No19./I.B. Gurevich.

[5] Chen J. et al. A survey of qualitative spatial representations // The Knowledge Engineering Review. 2015. Vol. 30, № 1. P.106-136

[6] Papadias D., Sellis T. Qualitative representation of spatial knowledge in two-dimensional space // The VLDB Journal. 1994 Vol. 3. NO 4. P.479-516

[7] Maglinets Yu. A., Raevich K. V., Tsibulskii G. M. Knowledge-based geoinformation technology for evaluation of agricultural lands // Procedia Engineering. 2017. Vol. 201. P.331-340

[8] Brezhnev R.V., Maglinets Yu.A., Raevich K.V., Tsibulski G.M. Modeling of Agricultural Spatial Objects with Heterogeneous Dynamically Changing Spatial Structure // CEUR Workshop Proceedings. 2018. Vol. 2210 P.316 - 322.

[9] Brezhnev R.V., Maglinets Yu.A. The dynamic model of agricultural land structure on the space images in the precision agriculture tasks // E3S Web of Conferences. 2019 Vol. 75. Article Number 01001, RPERS 2018.