Case based reasoning in intelligent geographic information systems for the management of logistics projects

S L Belyakov[1], A V Bozhenyuk[1], M L Belyakova[1] and S A Zubkov[1]

[1] Southern Federal University, 105/42 Bolshaya Sadovaya Str., Rostov-on-Don, Russia

Abstract. Logistics project is a special schedule that considers transportation of certain consignment from one point to another. These logistical operations should be executed within the cargo transportation process. When developing and implementing logistics projects it is necessary to make decisions that require knowledge of previously completed projects. Intelligent Geographic Information Systems (IGIS) enables developers and dispatchers to use spatial data, as well as to use solutions generated through Case Based Reasoning. This allows to consider the problem of incompleteness, uncertainty and ambiguity of data in the ever-changing real world. In this paper we investigate the mechanism for assessing the proximity of past logistics projects with specified project. The paper proposes a special conceptual model for presenting precedents that contains, in addition to the project description, a set of its permissible transformations. They do not change the meaning of the project. The function of transforming precedents for transferring experience to a given area of space is introduced. The conditions of existence of the transformed precedent are analyzed. A method for estimating the proximity of a precedent pair based on classification is proposed. The procedure of transformation past logistic projects is analyzed. An illustrate examples of the proposed approach is given. The conclusion considers the process of increasing the reliability of decisions made by accumulating the results of aposteriori analysis of precedents by experts.

1. Introduction

Any human activity, production and consumption of products require the movement of goods in space. For this reason, many transportation tasks are formulated and solved by logistics methods [1]. The task of logistics is to deliver the shipment to the recipient with a given quality. Quality is determined by the cost and time of delivery. One of the main problems of logistics is the risk of a normal completion of the planned process of cargo delivery. The reason for the risk is that transport systems, storage and handling facilities are subject to difficult not predictable influences by environment. A variety of space-time data and knowledge are required for reliable planning and reliable implementation of planned transportation [2]. This is ensured by the use of geoinformation systems (GIS) [3], which provide not only the necessary data, but also an intelligent service for the planning and execution of the logistics project. A logistics project is understood as a special schedule that determines transportation of a certain consignment from one point to another. A logistics project can be represented as a model

$$M_L = < S_L, P_S, T > \qquad (1)$$

where S_L is a set of spatially localized logistics centers; P_S is the set of transport routes between logistic centers; T is the delivery schedule. Here, the logistic center is the place where the logistics operations for the goods are performed (packaging, unpacking, loading, unloading, sorting, recast, temporary storage, etc.). The cartographic description M_L in the GIS has a link to the terrain in which the project is intended, or has already been implemented. Spatial-temporal binding, on the one hand, is a rich source of information on the topology of the environment for planned projects, on the other hand it is a source of experience in solving problems that arose earlier in the execution of projects. This allows to solve the typical logistics problem of incompleteness, uncertainty and ambiguity of data in the ever-changing real world [2].

The main means of using the previously obtained experience is Case Based Reasoning (CBR) [4], which is used to support decision making both at the stage of development of the logistics project and at the stage of its execution. Storage of precedents, choice, and training allow achieving high reliability of decision-making. The reliability is understood as the experimentally confirmed conformity of the solution to the state of the real world. As the analysis has shown, the application of CBR for planning and implementing logistics projects in the GIS environment raises a number of specific problems. So placing on the map objects from S_L to P_S reflects a fairly superficial knowledge of the essence of the project. As a result, it may turn out that an analysis of the project's similarity in metric can not give a reliable solution in principle, since it does not reflect a deep knowledge of the project's relationships with the environment. For example, in figure 1 shows the precedent of car parking with a cargo (painted rectangle). In the planned project, for some reason, it is expected to park the car at a small distance from the previously used parking place (shaded rectangle). It seems that such a decision is completely justified, since the situation and the decision taken are very close. However, the attempt to implement this solution failed: it turned out that at the given position of the car in this place it is impossible to automatically load containers. The project should be reviewed due to a critical delay caused by waiting for the desired parking space. Thus, the lack of sufficient semantic content in the description of the precedent led to an unreliable decision.

In this paper, we propose a conceptual model of knowledge representation about precedents that allows us to display in-depth knowledge about situations and solutions, as well as the organization of the CBR process that corresponds to this model. The efficiency of using the proposed approach by the criterion of reliability of the solutions being developed is investigated.

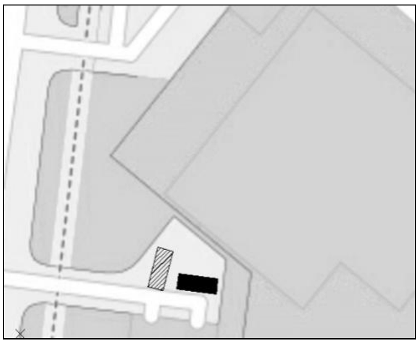


Figure 1. Example of placing a car in a logistics project.

2. Related works

The problems of using CBR for solving problems using GIS have been studied for a long time and cover all the stages of CBR: the presentation of knowledge about precedents, their storage in memory, the search for close situations, the adaptation of the known solution, and the preservation of new knowledge obtained in the implementation of the solution [4]. CBR reproduces a typical process of cartographic analysis, since GIS is a system that accumulates information in the form of precedents for observing objects, phenomena and events of the real world [3].

The main purpose of many studies related to the automation of cartographic analysis is to highlight and formalize the knowledge necessary to obtain reliable solutions. An example of this is the paper [5], in which the authors investigated the problem of selecting previously performed relevant projects for the urban planning problem. The resulting solution consisted in an automated search and issuing recommendations to the developer, on the basis of which he took the design decisions. Expert-developer, using the CBR software tools, is closely involved in the analysis cycle, retains the intellectual functions of comparing precedents and developing solutions. The drawback of this direction is the impossibility of transferring experience, known as the "mapping problem" [4].

A significant number of publications are devoted to conceptual models of knowledge representation about the past case for CBR [6, 7]. For example, in [6] the knowledge representation, including the relationships of cartographic objects, is examined. The authors included the additional component "Geographic Environment" in the knowledge base of intellectual GIS and introduced a special term "Spatial CBR", which emphasizes the importance of taking into account the topology of the space in which precedents are presented. Common to the work of this direction is the implementation of the proposed conceptual model by non-cartographic means. The product rules or decision tables play the role of a black box [5] for GIS analysts who are accustomed to using cartographic visualization. This creates a cognitive barrier to the extraction and reuse of experience. Thus, the task of constructing visualizable conceptual models that correspond to the applied field has not been studied enough.

Decision-making based on geodata always uses an important comparison of alternatives for CBR. The problem of comparison lies in the adequacy of the applied field, in the objectivity of the conclusion about the level of similarity of situations. This has been studied in a number of publications. An example is the paper [8], in which the authors set the task of estimating the weight coefficients for comparing the precedents of solutions. A known way to reduce subjectivity is the use of statistical data. However, as stated in [8], their content does not allow to reliably predict future situations. Applying a specific method of random generation of solutions and comparing their weight coefficients with a data set of a close semantic content, the authors reveal a deep knowledge of the connection between the decisions made and the topology of the terrain. Analyzing similar works, it is possible to draw a conclusion about the need for further studies of the problem of adequate comparison of situations and solutions.

It should be noted that the problem of finding and adjusting the parameters of the proximity estimation metric is in the center of attention of researchers in the field of neurocomputing. In [9-11], approaches to the implementation of CBR with the use of neural networks for the selection of similar precedents are analyzed. The decision structures in the problems under consideration are quite simple, often they are single numbers. As a consequence, the training of an intelligent system concentrates on recognizing situations that correlate with numerical solutions. The construction of solutions of a complex structure is not investigated.

In work [12] the approach to the organization of intellectual GIS, which proceeds from the goals of building the infrastructure of the geographical knowledge environment, is proposed. Geographic relations, ontologies, gazetteers and rules are the basis of "geographic reasoning" and together form a geographic knowledge base with a very wide range of applications. Systematizing modern knowledge about knowledge in intelligent GIS, the paper pays attention to the role of cartographic visualization in the extraction of knowledge and their use for solving applied problems. With reference to CBR, this indicates a poorly investigated task of cartographic description and visualization of the logic of reasoning.

The combination of cartographic visualization and intelligent methods for analyzing geodata is explored within geovisual analytics [13]. Here one of the main problems is the support of the meaningfulness of the analyzed cartographic image, the context of the analysis. For this, methods of expressive visual representation of semantic categories are necessary for the applied task. With respect to CBR, this can be interpreted as a search for a way of displaying the semantic content of visual objects used for decision making.

3. Features of the use of experience for the creation and implementation of logistics projects

The variability of the real world, the difficulty of predictability of its behavior, the uncertainty of assessing the state of the environment is the reason for the risk of implementing any logistics project. The problem of minimizing the potential loss can be solved both at the project development stage and at the stage of its implementation. It is necessary to assess the risk of the project as a whole at the design stage. To obtain it, developers need data on similar projects in terms of meaning. The involvement of experience allows us to evaluate the feasibility of the project integrally, based on the obvious assumption that similar projects are characterized by the same risk. The application of CBR in this case is aimed at finding the closest in the semantic sense of the past case, which allows either to justify the well-known, or to develop a new more reasonable solution.

The experience of eliminating emergencies plays a significant role at the stage of execution of the logistics project. In contrast to the previous case, we search for similar emergency situations in similar projects. In this case, decisions about reconfiguration or early completion of the project are generated and selected using CBR.

Analysis of the conditions in which CBR is applied made it possible to distinguish three of its features. The first is related to the use of the category "meaning" in relation to projects and emergency situations. Its intuitive interpretation by analysts, developers and dispatchers should find a formal representation using GIS. Following the ideas of conceptual semantics [14], it is advisable to associate the category of "meaning" with the figurative thinking of a person. The image accumulates in itself both actual data about the subject, and analytical information about the interpretation of these data in certain situations, about the possibilities of their modifications.

The second feature is that the state space of precedents is not uniform. This means that two situations with similar parameter values in different areas of space are essentially different. For example, parking P1 coincides in capacity and degree of congestion with parking P2, but in the area of location P1 in winter often there is a strong ice, which complicates entry and exit. From this it follows that the experience of placing a car in parking P1 can not be applied equivalently on P2. Experience must be transformed in some way when moving from one area to another. If this is not done, two situations can be compared, one of them can never be realized in a given locality. A correct comparison (from a formal point of view) of situations loses its meaningful meaning. Therefore, comparison of precedents in the CBR process should use transformation.

The third feature is related to the training of intellectual GIS. The source of knowledge for this is an expert who is able to convey in cartographic form the result of a posteriori analysis of logistics projects. Therefore, an important role in the organization of the knowledge base is acquired by the mechanism of reducing the redundancy of knowledge. The absence of such a mechanism reduces the efficiency of the expert's work and worsens the results of the CBR. There is a potential opportunity to reduce the redundancy of knowledge in GIS, which is based on the space-time data of the map. The mapping of the topology of space can be used to assess the boundaries of the application of knowledge through its transformation.

4. Representation of the meaningful content of logistics projects

The concept of "meaning" of logistics projects we propose to implement using two methods:

- by abstracting the structure of the project in the form of a model (1). In this case, each logistics project in the GIS is depicted on a separate layer in the form of a diagram. The elements of the schema are linked with references to attributive (descriptive) data, as well as with time diagrams;
- by description of the permissible transformations of each element (1), which preserve the essence of the logistics project intuitively understood by the expert. Transformations are also represented by cartographic objects on a separate layer of a common GIS map.

Let's consider an example. In figure 2a with solid lines shows the scheme of project A, described by the model:

$$M_A = << S_{A1}, S_{A2}, S_{A3} >, < P_{A1}, P_{A2} >, T = \varnothing > .$$

By dotted lines show the scheme of project B, described by the model:

$$M_B = << S_{B1}, S_{B2}, S_{B3} >, < P_{B1}, P_{B2} >, T = \varnothing > .$$

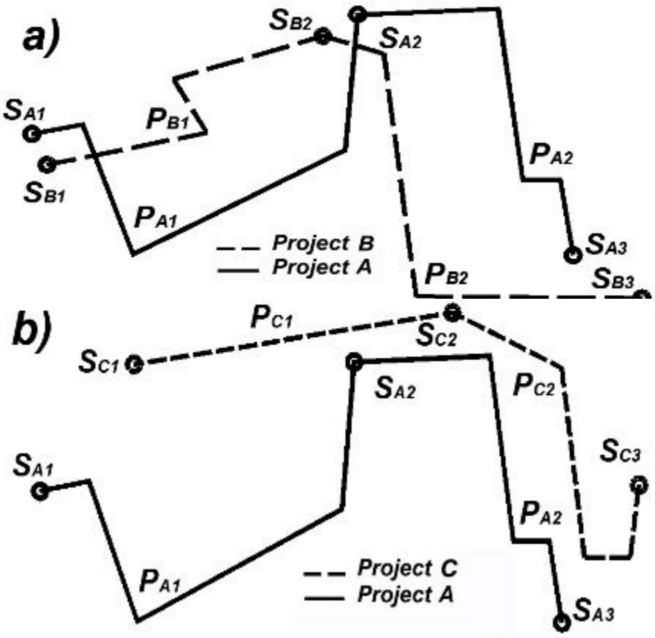


Figure 1. Schemes of logistic projects A, B, C with a link to the terrain.

Here, the transport schedule timelines are not used to simplify the presentation. The visually observed proximity of logistics centers and trajectories can not be grounds for considering projects A and B to be more similar than the similarity of projects A and C (see figure 2b). Project C is described by the model:

$$M_B = << S_{C1}, S_{C2}, S_{C3} >, < P_{C1}, P_{C2} >, T = \varnothing > .$$

An adequate measure of proximity must be used for objective comparison. The difficulty in constructing a distance metric that reliably reflects not only the geometric properties of precedent projects, but also the topology of the terrain, is obvious. Therefore, the idea arises of using a different conceptual model for describing use cases that reflects the essence of precedents. Figure 3 shows the scheme of project A with permissible transformations of the location of logistics centers and transportation trajectories. The modified model of the logistics project takes the following form:

$$M'_A = M_A \cup < H(S_{A1}), H(S_{A2}), H(S_{A3}) >, < H(P_{A1}), H(P_{A2}) >> .$$

Here, $H(x)$ denotes the set of objects that can be obtained by transforming the object x, but at the same time retain the intuitive meaning associated with it. For logistic projects, the permissible transformations can be various territorial arrangements of logistics centers, different transportation trajectories, different types of vehicles and systems, time-of-delivery restrictions, etc. For example, $H(S_{A1})$ is a polygonal area on the map, within which the location of A1 can change, while preserving all the essential characteristics of the realized project A.

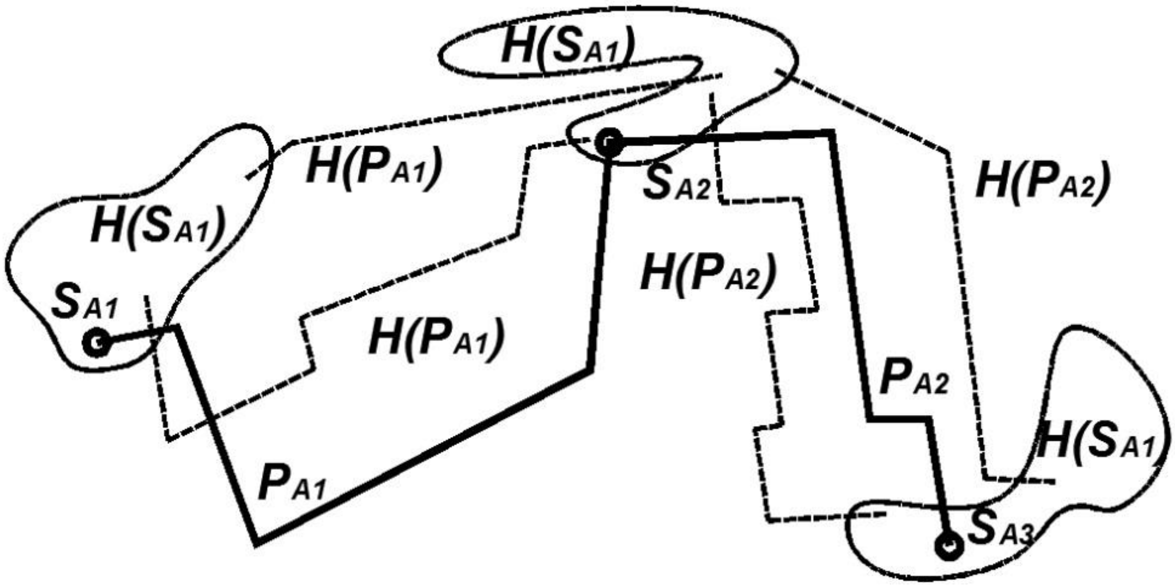


Figure 2. Scheme of the logistics project A with a set of permissible transformations.

Similarly, $H(P_{A1})$ specifies in this case two feasible transport paths, which also do not change the essence of the project. Trajectories are represented by dashed lines. If projects A and B are compared taking into account the permissible transformations of project A (see figure 4), it should be noted that neither A scheme nor its permissible modifications coincide with the project B. This leads to the idea that there is some intuitively perceived fundamental difference in the essence of projects. As for project C, figure 5 shows its position relative to project A. It can be seen that the position of the logistics centers of project C is in the areas of permissible transformations of project A, and P_{C1} and P_{C2} coincide with the permissible transformations of the project A trajectories. Thus, projects A and C are equivalent in meaning.

From a formal point of view, the precedent is represented in a way that includes two components:

$$J = < c, H(c) >, \tag{2}$$

- here c is the center of the use case image, which displays the situation actually observed. The component $H(c)$ is a set of permissible transformations that preserve the intuitive sense of the precedent Both components are represented by cartographic objects with references to non-cartographic data sources. We note some properties of the image (2):
- the image does not exist without a center ($c = \varnothing \Rightarrow J = \varnothing$). This means that the basis of figurative knowledge is an object, event or relation that can be described by a map or a scheme;
- the center of the image is a permissible transformation ($c \in H(c)$).Otherwise, the continuity of the image is violated;
- an image without a description of permissible transformations ($H(c) = \{c\}$) from the content point of view is meaningless. This means that the performing by the intellectual system of any formally correct operations over the center of the image will not lead to the appearance of meaning. The meaning arises only as a result of an external interpretation of the precedent center.

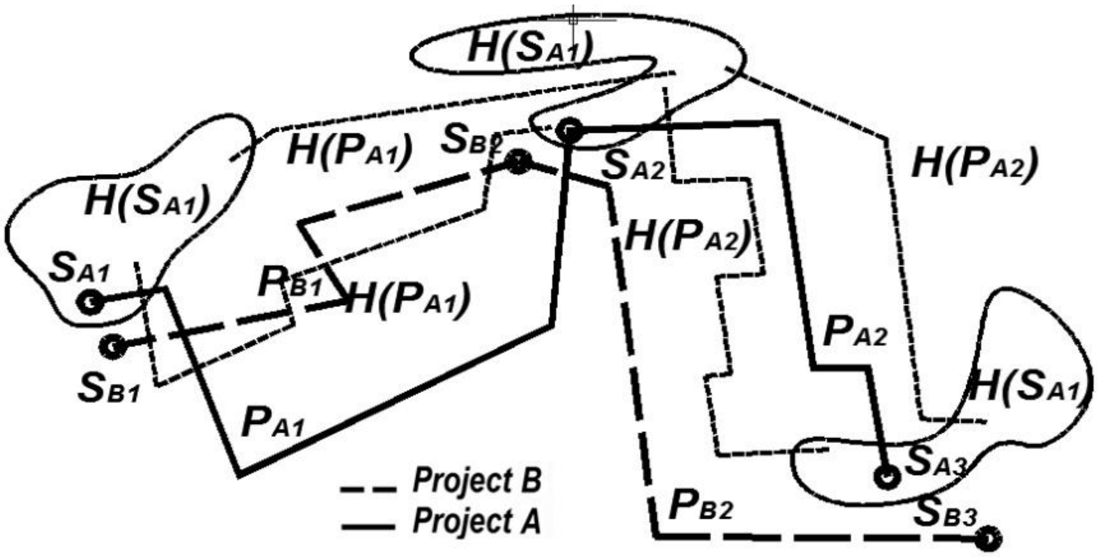


Figure 3. Project schemes A and B for comparison with allowable transformations.

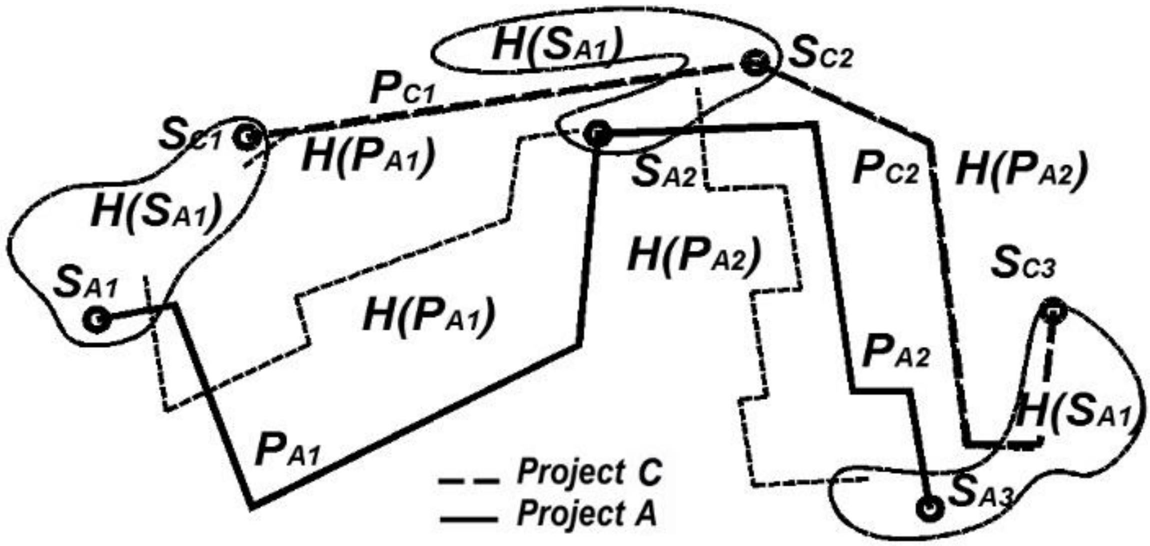


Figure 4. Scheme of projects A and C for comparison with allowable transformations

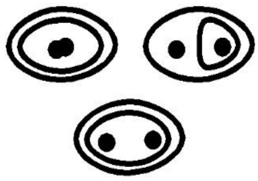
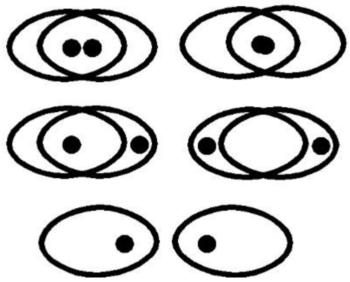
5. Comparison of the images of logistics projects

The comparison operation plays a crucial role in the implementation of the CBR. Traditional is the use of the metric, which is expressed through a weighted difference of the essential parameters of precedents [4-11]. When using the figurative representation (2), it is expedient to use a classification based on an analysis of the relative position of the centers and the regions of permissible transformations. The variants of the relative positioning are represented by the diagrams shown in figure 6. The image of the situation is represented as a polygonal area of permissible changes with the point of the center of the image inside. This makes it possible to visually display the relations "coincide", "have something in common", necessary for the logic of classification. Visibility in this case allows the expert to use the usual GIS tools to describe the rules for assessing the proximity of situations.

Evaluation of the proximity of precedents is realized by the classification of a given pair of $< J_1, J_2 >$ images. Each class is assigned a linguistic value of the proximity estimate. The distribution of the variants of the mutual arrangement of image components between classes reflects the intuitive representation of the GIS analyst about the similarity of precedents. In figure 6 shows a binary classification corresponding to the estimates of "similar images" and "not similar images."

We can see that in this case, a class of similar images refers to situations where $J_1 \supset J_2$, that is, the conclusion about similarity is deductively deduced. Such a variant of reasoning is reliable and plausible [15]. Not similar (the second column in Table 1) are images that are close by analogy, which is plausible, but not reliable. For the example above, the center of the image of the logistics project B does not fall within the permissible changes area of the project A image, i.e. is displayed by one of the diagrams of the second column of table 1. The project image center C is included in the permissible changes area of the project A image and is displayed by one of the diagrams of the first column.

Table 1. Example of classifying the relative positioning of image components

Class « similar images »	Class « not similar images»
	

6. Implementation of the transformation procedure

The ability to reproduce events and phenomena on the earth's surface with the help of GIS was one of the strong incentives for the development of systems of this class [3]. However, by now this function has not been implemented in any GIS for the following reasons: cartographic databases do not have the completeness and consistency, generalization is used for the development of maps, and the formulation of the task of simulating events and phenomena is ambiguous and ambiguous in itself. In the management of logistics projects, this task can be solved satisfactorily thanks to the abstract representation (1) and the special form of representation of experience.

In general, the transformation procedure is described by the following expression:

$$<\bar{c}, H(\bar{c})>= F_{TR}(<c, H(c)>, W_{dest}, L_{inv}, G) , \tag{3}$$

where $<\bar{c}, H(\bar{c})>$ is a transformed image and its transformations; W_{dest} is a given target map area; L_{inv} is a list of restrictions that must be met when transforming; G is the level of reliability of the result. If the precedent can not be transformed with the given input data, the procedure returns an empty set.

Consider the transformation algorithm, considering that the GIS software library includes the basic functions of designing applied map objects. For example, a path in a road network can be constructed by two points; a platform of a given type can be constructed from a set of points; the service area can be built according to the specified type of service and the location of the service centers, etc. When constructing these objects, basic knowledge of form, arrangement and topological relations is used, which exclude the construction of incorrect objects. The algorithm is described as follows:

1. Organize a loop on objects of the set $<c, H(c)>$. Perform steps 2 – 6 for each $z \in <c, H(c)>$.
2. Construct an object z in the W_{dest} region using the basic GIS function. If $z = \varnothing$ then go to step 8.

3. Set *CurrentRestrictionNumber* := 1 for the L_{inv} list.
4. Check the restriction with the number of the *CurrentRestrictionNumber*. If it not satisfied, go to step 8.
5. *CurrentRestrictionNumber:= CurrentRestrictionNumber +1*. If the L_{inv} list contains unchecked restrictions, then go to step 4.
6. Add z to $<\bar{c}, H(\bar{c})>$. If the cycle is not complete, select the next element z and go to step 2.
7. Evaluate the reliability of the result. If the reliability is not less than G then set $<\bar{c}, H(\bar{c})>$ is the result. Go to step 9.
8. Set $<\bar{c}, H(\bar{c})>$ is empty.
9. End.

The result obtained with the above algorithm has a double meaning for CBR. On the one hand, the nonempty $<\bar{c}, H(\bar{c})>$ set indicates a reasonable comparison of situations. This action is typical for the process of human thinking, when similar, but "unreal" situations are not taken into account. Transformation simulates the human imagination in this case. On the other hand, set $<\bar{c}, H(\bar{c})>$ reflects an experimentally proven solution that should participate in comparing alternatives.

7. Conclusion

The implementation of the intellectual system studied in this work is based on the expert's analytical knowledge that arises from the a posteriori analysis of the completed logistics projects. Here precedents are not only the facts actually observed, but also the results of their comprehension by the expert. The type of knowledge under consideration requires a special presentation and a mechanism for assessing proximity without affecting the CBR concept. The set of permissible transformations is simply displayed by means of GIS. Many of them are cartographic objects and are visualized. Visualization facilitates the transfer of knowledge in a figurative form.

Further development of this study is to improve the mechanisms of experience transformation, as well as the learning process of intellectual GIS.

8. References

[1] Grazia Speranza M 2018 Trends in transportation and logistics *European J of Operational Research* **264** p 830

[2] Katsumi M and Fox M 2018 Ontologies for transportation research: A survey *Transportation Research Part C* **89** p 53

[3] Shashi S and Hui X 2017 *Encyclopedia of GIS* (New York: Springer) p 506

[4] Aamodt A and Plaza E 1994 Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches *AI Communications* **7**(1) p 39

[5] Yeh A and Shi X 2001 Case-based reasoning (CBR) in development control *JAG* **3** (3) p 238

[6] Du Y, Liang F and Sun Y 2012 Integrating spatial relations into case-based reasoning to solve geographic problems *Knowledge-Based Systems* **33** p 111

[7] Wei H, Xu Q and Tang X 2011 A knowledge-based problem solving method in GIS application *Knowledge-Based Systems* **24** p 542

[8] Veronesi F, Schito J, Grassi S and Raubal M 2017 Automatic selection of weights for GIS-based multicriteria decision analysis: site selection of transmission towers as a case study *Applied Geography* **83** p 78

[9] Yan, A, Yu H and Wang D 2017 Case-based reasoning classifier based on learning pseudo metric retrieval *Expert Systems With Applications* **89** p 91

[10] De Renzis A, Garriga M, Flores A and Cechich A 2016 Case-based Reasoning for Web Service Discovery and Selection *Electronic Notes in Theoretical Computer Science* **321** p 89

[11] Relich V and Pawlewski P 2018 A case-based reasoning approach to cost estimation of new product development *Neurocomputing* **272** p 40

[12] Laurini M 2017 *Geographic Knowledge Infrastructure. Applications to Territorial Intelligence and Smart Cities* (ISTE Press Ltd)
[13] Schumann H and Tominsk C 2011 Analytical, visual and interactive concepts for geo-visual analytics *Journal of Visual Languages and Computing* **22** p257
[14] Kuznetsov O 2014 On conceptual semantics *Artificial Intelligence and Decision Making* **3** p 32
[15] Evans J S B T and Over D E 2013 Reasoning to and from belief: Deduction and induction are still distinct *Thinking and Reasoning* **3**

Acknowledgments
This work has been supported by the Russian Foundation for Basic Research, project № 17-01-00060.