Situations representation and retrieve in the case-based reasoning system for managing a complex technological object

Igor Glukhikh and Dmitry Glukhikh

University of Tyumen, 6 Volodarskogo Street, Tyumen, 625003, Russian Federation

gluhihdmitry@gmail.com

Abstract. Modern urban infrastructure systems are complex technological objects. Their stable operation is important for facilitating a comfortable and safe urban environment. These systems are supported by monitoring and quickly addressing potentially dangerous situations. Due to the high complexity and high level of responsibility of decision-making in dangerous situations, the problems of intelligent decision-making support are relevant. The article explores the use of the case based reasoning (CBR) method for solving these problems. In CBR, the knowledge base of a decision support system contains cases: situations and solutions that are applicable in such situations. When a dangerous situation arises, the system turns to the knowledge base to search for a case with a prepared solution. For realization of the CBR method, the paper proposes a situation is represented through the states of the elements of a complex object and the relationships between them. For the retrieve of cases in the knowledge base, an approach that takes into account the structural and parametric similarity of situations is proposed.

Keywords: case-based reasoning, intelligence monitoring, decision support systems, urban infrastructure.

1. Introduction

Modern urban infrastructure systems (electricity, gas, water supply, and heating) are complex technological objects (CTO). The safety and stability of their processes are important not only for supporting the life in the city, but also for the preservation of the environment, health and lives of people.

These systems are supported by monitoring their condition and prompt troubleshooting. The tasks of monitoring complex objects in order to prevent emergencies are relevant for enterprises supplying heat, water, gas, and energy to the region, as well as for security and urban management services.

Quite a lot of recent papers are dedicated to the research in the field of monitoring of technological objects. Primary data collection and processing are a priority, and various technical methods, devices and communication channels are being developed for this [1-4]. The next big task is to analyze the data and predict the dynamics of changes in the condition of the object and to identify emergency situations. For this, methods of data mining, as well as machine learning and artificial neural networks are used [5-12]. However, after identifying dangerous situations, a complex of tasks at the following levels appears. These are the tasks of decision-making in dangerous situations and preventing their consequences, as well as the tasks of analyzing, identifying and addressing the causes of such situations in order to prevent them from occurring in the future. This paper is dedicated to the tasks of the level of decision-making support in the prompt elimination of dangerous situations. The case based reasoning (CBR) method known from artificial intelligence studies is considered as a base for this.

The CBR method involves maintaining a knowledge base (KB) where cases, descriptions of complex situations known from past experience and solutions that have been recommended or used in these situations before, are stored. When a new dangerous situation arises, the KB contains a case with the same or a similar situation and a solution that is provided to users. The solutions found in this way can be used directly or adapted to the situation at hand.

Case based reasoning is widely used in different subject areas. One of the promising areas is associated with decision-making when managing complex technical and organizational objects [13-16]. At the same time, due to the complexity and diversity of the objects under consideration, each problem area still requires its own research, starting with the search for models for formalizing the representation of objects and continuing with algorithms for inference and adaptation of solutions.

The goal of this research is to develop generalized (universal) models of representation of situations for the application of the case based reasoning method when making decisions in dangerous situations on complex technological objects of urban infrastructure. The article first describes the content and stages of the CBR method, then develops a general ontological model (GOM) of a complex object. Then, based on this model, a formalized representation of situations arising at a technological object is developed. Further, an approach and methods for assessing the similarity of situations are proposed, taking into account both the parameters and the structure of situations on a complex object. After that, the results are discussed and tasks for further research are proposed.

2. Materials and methods

The case based reasoning method was widely used in different subject areas. In the 80s and 90s of the last century, a number of commercial systems were developed using CBR (CLAVIER, CHEF, CASEY, JULIA, etc.) that convincingly demonstrated the effectiveness of this method [17]. Today, the development of CBR is associated, on the one hand, with the expansion of the scope of applications and the development of practical applications; on the other hand, with the research that will aid in combining the methods of presentation of knowledge and machine learning for an integrated, neurosymbolic approach to the creation of artificial intelligence [18-20]. In a CBR system, a knowledge base is a case base, each of which is a pair <Sit, Sol>, where Sit is a situation that required its solution; Sol is the solution for this situation. As a solution, one can write down what has already been applied in practice in that particular situation and has shown positive results. Or it can be a solution that was specially developed by experts in advance for such situations.

The main stages of inference in the CBR system are:

1. Identifying the actual situation $Sit_{act;}$
2. Retrieve the Sit* situation, which is closest to the current situation Sit_{act}.

Applying a solution from a pair of < Sit* , Sol*> to resolve the Sit_{act}. In this case, if the situation Sit* is not close enough to the Sit_{act}, the solution Sol* acts as a basis for generating the adapted $Sol_{act} \Leftrightarrow Sit_{act}$.

3. Analyzing the new pair < Sol_{act}, Sit_{act} > and saving it in the KB for later use.

In order to implement these steps of the CBR method in the area under consideration, it is necessary to formalize the representation < Sol, Sit >. Meanwhile, it is necessary to consider that in general, a complex technological object includes elements of different types, such as technical devices, software and hardware communication and management systems, servicing and operating organizations (staff), resources, and other environments.

Further, we will develop a representation of situations by presenting the states of such elements and the relationship between them.

At the same time, by the situation at a complex object, we will understand the state of affairs, which is characterized by the current state of the elements of the object and the connections between them.

3. Results

3.1. The complex technological object structure

The CTO structure contains elements of different types. We will single out the following elements: equipment, personnel, software and information complex, resources, buildings, environmental objects and phenomena. Buildings, environmental objects and phenomena belong to the environment, but are considered as part of the CTO as they have a connection with the CTO and are able to influence it.

In the ontological representation, a complex object CTO is described by the quaternion <*O*, *S*, *R*, *A*>, where *O* is a set of elements. These elements include: equipment, personnel, software and information complex, resources, buildings, environmental objects and phenomena;

S is a set of states:

$$S=\{ S_{ij} \mid \forall_i \in I ;\ \forall_j \in J_i\}, \quad (1)$$

where I is a set of indices of CTO elements,

J_i is a set of the i^{th} element state indices. A typical set of states includes "Running", "Stopped", "Operational", "Not operational", "Present", "Absent", "Available", "Unavailable," etc.

R is a set of relationships between elements of a complex object:

$$R=\{ R_k \mid \forall_k \in K\}, \quad (2)$$

where K is a set of indices of relationships between the elements of a CO; it contains typical relationships Part-of, Has-a, Kind-of, etc. Additional relationships characteristic of a particular object can be added as well.

A is a set of axioms representing certain necessary combinations of connections between the elements of the object.

3.2. Representation and retrieve of situations in the case base

Representation of the situation at CTO. In terms of the developed ontological model, a situation at a complex object will be understood as a set of elements states and a connection between them at a particular moment. In other words, the Sit_z situation is a projection of an ontological model to a specific environment, which identifies the specific values of elements, connections, and states:

$$Sit_z=\langle O_z, S_z, R_z\rangle, \quad (3)$$

where: $O_z \subseteq O$, $S_z \subseteq S$, $R_z \subseteq R$, z is an index of the situation in the set of situations stored in the knowledge base.

Thus, from the overall representation of a complex object with the GOM, we transition to presenting situations that reflect the states of the CTO elements and the relationship between them. Location of elements in (3) according to current data allows to identify the current situation Sit_{Act}, after which it becomes possible to compare it with the Sit_z situations recorded in the knowledge base.

Retrieve the situational knowledge base. Two similarity metrics are used to retrieve close situations: similarity in the space of relationships SSim (structural similarity) and similarity in the space of states PSim (parametric similarity). The first will reach its maximum value in the event that in two situations there is a complete similarity between the links of the elements and the minimum value, if in two situations there are no identical connections between the elements of a complex object. The second will reach its maximum value when all CTO elements in each of the situations are in the same state and the minimum value when all the elements of the object are in different states.

First, we will consider the set of relationships on the CTO elements. We will introduce a graph G_k, which will show kth relationship of the elements of a complex object. The union of relationship graphs represents the entire set of interaction relationships in the ontological model.

Further, we will represent the relationship graph G_k with the adjacency matrix M, in which cells contain 1 if the corresponding elements of the object are in a relationship from a set R_z and 0 otherwise.

Let $M_{k,act}$ be a matrix for the k^{th} relationship in the current situation, and $M_{k,z}$ be a matrix for the k^{th} relationship for the z^{th} situation from the knowledge base.

Then we can determine the similarity matrix of the two situations in a relation R_k:

$$M_k (z, act) = M_{k,z} * M_{k,act} , \tag{4}$$

Where:

* is the operation of elemental multiplication of matrices.

The following formula is used to assess the similarity Sim_k to R_k:

$$Sim_k (Sit_Z, Sit_{Act}) = N / \max \{N_{act}; N_z\}, \tag{5}$$

Where:

N is the number of non-zero cells in the matrix M_k (z, act), this number shows the number of connections between the elements that are in the first and second situations;

N_{act}, N_z are the number of non-zero cells in the matrices $M_{k, act}$ and $M_{k,z}$, which means the number of non-zero connections between elements in situations Sit_{Act}, Sit_z, respectively. In this manner, Sim_k (.) demonstrates the share of matches between two situations in a relations R_k.

Then the overall estimate of similarity in the space of the relationship is calculated using a weighted sum:

$$SSim (Sit_Z, Sit_{Act}) = \sum \alpha_k\ Sim_k(Sit_Z, Sit_{Act}), \tag{6}$$

Where:

$\alpha_k \in [0, 1]$, $\sum\alpha_k = 1$ is the coefficient of relative importance of k^{th} relationship and the SSim value lies in the range of 0 to 1.

A similar approach is used to assess similarity in the state space.

A situation matrix representation is also used to assess the similarity of the situation in the state space. In this case, in the situation matrix, the rows denote the elements of a complex object, and the columns denote the states in which these elements can be located. Then, if the element O_i is in the state S_j, then 1 appears at the intersection of the corresponding row and column.

Line-by-line comparison of two matrices allows calculating the similarity between the elements Oi in the state space $PSim_i$ (Sit_Z, Sit_{Act}). In this case, $PSim_i$ (Sit_Z, Sit_{Act}) takes on the value 1 - if it is in the same position in both matrices, i.e. the elements are in the same state and 0 otherwise. Then the general affinity between situations in the state space is:

$$PSim(Sit_Z, Sit_{Act}) = \sum \beta_i\ PSim_i (Sit_Z, Sit_{Act}), \tag{7}$$

Where:

$\beta i \in [0, 1]$, $\sum\beta i = 1$ is the relative importance factor of the i^{th} element in determining the similarity of situations.

The final similarity of the situation assessment is formed by a pair:

$$Sim (Sit_Z, Sit_{Act}) = \langle SSim (Sit_Z, Sit_{Act}), PSim(Sit_Z, Sit_{Act})\rangle, \tag{8}$$

Sequential retrieve from the knowledge base is done in two stages:

- retrieve by the criterion: SSim (Sit_Z, Sit_{Act}) $\geq$ (1- ε) , where ε is a certain threshold on which the number of situations retrieved at this stage will depend;

- retrieve of the closest situation in the KB by the criterion: PSim (SitZ, SitAct) $\rightarrow$ max.

The experiment. For experimental verification of the described technique, the case is considered: a malfunction at a complex object of urban infrastructure. The individual heating station of the building is taken for a complex object. The technological scheme is an independent two-circuit heating system, where an external coolant through a heat exchanger transfers thermal energy to the coolant of the home heating system.
The elements of a complex object are formed into groups:

- technological (pump, heat exchanger, internal pipeline);
- providing (software, electricity, other equipment);
- personnel (electrician, plumber, emergency service);
- environment (substation, neighboring buildings, natural objects, natural phenomena).

The base of 300 situations with different states was generated using the developed algorithm in the Visual Basic code language (VBA). The generated situations were written as state matrices.

Also, the actual situation (Sit_{act}) was generated and wrote in the form of a state matrix. The Sit_{act} was designed to simulate a pump stopping against a background the providing equipment some malfunction. Wherein, the pump has a working condition.

A retrieve of similar situations was carried out from the generated base using the developed algorithm in the VBA, according to the method described above, by the formula 7. As a result of the selection, the algorithm produced two situations which having PSim(Sit_Z, Sit_{Act}) = 0.92, i.e. similarity of 92% with the current one. Both situations have a difference in the state of only one element. The state matrices of actual and retrieved situations presents in the tables 1-3.

Table 1. The state matrices of the actual situation (Sit_{Act}).

Sit_{Act}	workable	faulty	in work	stopped	not available	available	interfere	not interfere
Input	1							
Heat exchanger	1							
Pump				1				
Other equipment		1						
IT		1						
Electricity						1		
Emergency						1		
Plumber					1			
Electrician						1		
Buildings								1
Nature object								1
Nature event							1	
Location								1

Table 2. The state matrices of the actual situation No1 (Sit_Z No1).

Sit_Z No1	workable	faulty	in work	stopped	not available	available	interfere	not interfere
Input	1							
Heat exchanger	1							
Pump			1					
Other equipment		1						
IT		1						
Electricity						1		
Emergency						1		
Plumber					1			
Electrician						1		
Buildings								1
Nature object								1
Nature event							1	
Location								1

Table 3. The state matrices of the actual situation No2 (Sit_Z No2)

Sit_Z No1	workable	faulty	in work	stopped	not available	available	interfere	not interfere
Input	1							
Heat exchanger	1							
Pump		1						
Other equipment		1						
IT		1						
Electricity						1		
Emergency						1		
Plumber					1			
Electrician						1		
Buildings								1
Nature object								1
Nature event							1	
Location								1

However, during the analysis, there was a semantic difference between these situations. In the first selected case, the pump remained working against a background the providing equipment some malfunction. In the second case, the pump also had a broken condition like providing equipment.

In practice, the solution for each situation will be different, since a working pump and a broken pump obviously require different actions. Consequently, a collision arises when selecting a similar situation from the base.

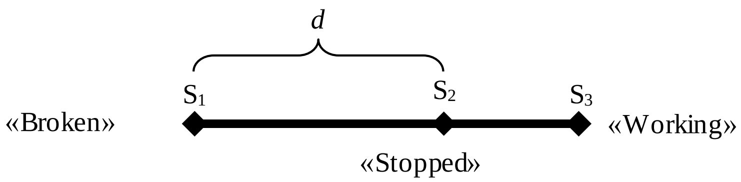
There is introducing the distance between states to exclude such collisions. The distance between the states of an element is determined by the following formula:

$$d_i = \left\| S_{i,act} - S_{i,z} \right\|, \tag{9}$$

where:

$S_{i,\,act}$, $S_{i,\,z}$ - the state of the i^th^ element in the actual situation and the z^th^ situation from the base, respectively.

For the case under consideration, the possible states of the pump are ordered on the interval [0:1] in such a way that S_1 is at point 0, S_3 is at point 1, and the other states take values between them:



Thus, the degree of closeness of situations in the state space is determined by the formula:

$$\mathrm{PSim}(\mathrm{Sit}_Z,\ \mathrm{Sit}_{Act}) = \sum \beta_i * (1 - d_i)\ , \tag{10}$$

where:

β_i –relative importance factor introduce above, which determines by the formula:

$$\beta_i = \frac{r_i}{\sum r_i}, \tag{11}$$

where

r_i – the significance coefficient of an element in a complex object;

d_i – the distance between the states in which the i^th^ element of a complex object is in the compared situations.

The r_i coefficients are formed expertly using the significance scale, in which the following values are used: 0 - not significant, 1 - weak significance, 3 - medium significance, 5 - strong significance.

Thus, after the modernization of the formulas due to the distance d, cases were excluded when two critically different situations with different solutions are issued in response to the appearance of the same Sit_{Act}.

4. Discussion

In this study, the authors proposed universal models of formalized representation of situations and solutions for managing a complex object based on the CBR method.

Initially, a general ontological model of a complex urban infrastructure object was developed, which reflects the elements of the object and the relationships between them. The situation is considered as a projection of the ontological model onto the specific states of the CTO elements and the relationships between the elements at a given time. The decision made in the current situation is represented as a discrete process of transition from the initial state to the target state. The pairs <Sit, Sol> are recorded in the knowledge base, where each situation is associated with a solution recommended for it. The presence of such a knowledge base during situational management of a complex object allows to quickly find solutions in critical situations applying cases from the knowledge base.

The proposed way of formalizing CTO allows for a broader view of emerging situations. The inclusion of elements of CTO environment into its formal representation allows us to take into account not only the technical aspects of the technological object, but also the influence of many external factors (the state of the surrounding objects, organizational systems, climatic conditions, etc.) in the presentation of situations and decision-making. Such comprehensive assessment of urban infrastructure in the context of its surroundings allows us to consider the object from the point of view of environmental safety as well.

Each situation in the knowledge base is associated with a solution that can be recommended to end users: service personnel, emergency crews, operational dispatch services of the city and the operating organization. In general, this solution can be presented as the following set of components:

- R1, instructions on how to act in a situation (technological map);
- R2, list of contacts of responsible persons and necessary organizations;
- R3, required reporting documents (templates, forms, acts, etc.);
- R4, additional background information (references to similar situations, expert recommendations, etc.).

It is assumed that during management of complex interactions between different services, components are addressed to different executors and controllers involved in resolving situations.

Proposed models allow to organize a view of situations by presenting the final set of states of the elements and the relationship between the elements of a complex object. Accordingly, the comparison of the two situations during the retrieve in the knowledge base will go through the comparison of elementary states and relationships. Of particular interest is the level of states.

Detailing to the level of elementary states reduces the complex task of comparing situations to simpler tasks of comparing (recognizing) elementary states. The number of such states is limited and does not change during the operation of the system. However, their combination across the many elements of a complex object allows us to get an almost unlimited in practice set of situations. Various metrics and technologies, including trained neural networks, can be used to solve these state comparison (recognition) tasks. This will complement the methods of retrieve situations based on machine learning metrics, in which the correctness of choice is based on past experience and training data.

The use of neural networks to implement the CBR method may be one of the areas of further development of this research. The fact is that distance metrics, which require calculating multiple parameters [13-15], are traditionally used for comparing situations in CBR. At the same time, the results can be seriously influenced by both the choice of the metric and the quality of parameter measurement. The states of different elements of the object will be described in a variety of ways, from the exact values of numerical parameters to the quality of the characteristics and graphic images (on maps, photographs, etc.). For each of these methods, its own state comparison technology can be used. This opens up possibilities for creating hybrid choice models, where both metrics and recognizing neural networks will be used to retrieve situations in the knowledge base.

5. Conclusion

The study produced the following key results: structure of complex object, models of formalized representation of situations, as well as an approach and metrics for the retrieve of situations that take into account the structural and parametric similarity of them were proposed.

These results are important for further development of the methods and technologies of applied CBR systems. For example, on the basis of these representation models, inference problems can be solved under conditions of uncertainty of the states of a complex object and the relations between them. Representation of a complex situation through a combination of elementary states from a limited set creates the basis for using neural networks to retrieve cases in the knowledge base of a CBR system through recognition of its elementary states. Thus, this study is important for the development of the neurosymbolic artificial intelligence approach as applied to the tasks of managing complex organizational and technical objects.

In order to further develop the results, we plan to address such tasks as developing methods for analyzing and comparing states described in different parametric spaces, as well as generalizing the results for uncertainties of complex object states and the relations between them.

6. Acknowledgments

The research was funded by RFBR and Tyumen Region, project number 20-47-720004.

References

1. Kiziroglou, M. E., Boyle, D. E., Wright, S. W., Yeatman, E. M. Acoustic power delivery to pipeline monitoring wireless sensors. Ultrasonics, 77, 54-60 (2017).
2. Zarifi, M. H., Deif, S., Daneshmand, M.: Wireless passive RFID sensor for pipeline integrity monitoring. Sensors and Actuators A: Physical, 261, 24-29 (2017).

3. Junie, P, Dinu O., Eremia C., Stefanoiu D., Savulescu I.: A WSN Based Monitoring System for Oil and Gas Transportation through Pipelines. IFAC Proceedings Volumes 6(45), 1796-1801 (2012).
4. Zrelli, A., Ezzedine, T.: Design of optical and wireless sensors for underground mining monitoring system. Optik, 170, 376-383 (2018).
5. Zengqiang Xing at al.: Design and Application of Transmission Line Intelligent Monitoring System. E3S Web of Conferences 185, 01063 (2020).
6. Jia, Z.: Pipeline abnormal classification based on support vector machine using FBG hoop strain sensor. Opti. 170, 328-338 (2018).
7. He, P., Wang, J.: Statistical process monitoring as a big data analytics tool for smart manufacturing. Journal of Process Control 67, 35-43 (2018).
8. Ariyaluran Habeeb, R. A., Nasaruddin, F., Gani, A., Hashem, I., Ahmed, E., Imran, M.: Real-time big data processing for anomaly detection: A Survey. International Journal of Information Management, 45, (2018).
9. Rukmani, P., Teja, G.K., Vinay S., Reddy B.P.K.: Industrial Monitoring Using Image Processing, IoT and Analyzing the Sensor Values Using Big Data. Procedia Computer Science, 133, 991-997 (2018).
10. Z. Sh. Juraev, D. T. Muhamediyeva, D. M. Sotvoldiev: Construction of hybrid intellectual monitoring and decision-making systems. Journal of Physics: Conference Series, 1546, 012083 (2020).
11. Andrzej Bieleckia, Mateusz Wójcik: Hybrid AI system based on ART neural network and Mixture of Gaussians modules with application to intelligent monitoring of the wind turbine. Applied Soft Computing Journal, 108, 107400 (2021).
12. Andrey Yu. Puchkov, Maksim I. Dli, Ekaterina I. Lobaneva: Application of deep neural networks in models of complex technological objects. Bulletin of the Saint Petersburg State Institute of Technology (Technical University), 52, 104-110 (2020).
13. Eremeev, A., Varshavskiy, P., Alekhin, R.: Case-Based Reasoning Module for Intelligent Decision Support Systems. Proceedings of the First International Scientific Conference "Intelligent Information Technologies for Industry" (IITI'16), Vol. 1. Part III, pp. 207-216, Springer International Publishing (2016).
14. Galina S. Maltugueva, Alexander Yu. Yurin: Improving case-based reasoning with the aid of multi-criteria and group decision-making methods. Proceedings of the International Convention MIPRO, IEEE, 1031 – 1036 (2019).
15. Xiaoyan Jiang at al.: A Decision Method for Construction Safety Risk Management Based on Ontology and Improved CBR: Example of a Subway Project. International Journal of Environmental Research and Public Health. 17, 3928 (2020).
16. Hoda Nikpour, Agnar Aamodt: Inference and reasoning in a Bayesian knowledge-intensive CBR system. Progress in Artificial Intelligence. Progress in Artificial Intelligence volume 10, pages 49–63 (2021).
17. Kolodner, J.: Case-Based Reasoning. Morgan Kaufmann Publishers, Inc., San Mateo (1993).
18. Eisenstadt V., Langenhan C., Althof K.: Generation of Floor Plan Variations with Convolutional Neural Networks and Case-based Reasoning An approach for transformative adaptation of room configurations within a framework for support of early conceptual design phases, Proceedings of 37 eCAADe and XXIII SIGraDi Joint Conference, p. 79-84 (2019).
19. Mathisen, B.M., Bach, K. & Aamodt, A.: Using extended siamese networks to provide decision support in aquaculture operations. Applied Intelligence (2021).

20. M.F. Abdelwahed et al. : Solving the motion planning problem using learning experience through case-based reasoning and machine learning algorithms. Ain Shams Engineering Journal, 11, 133–142 (2020).