
Levykin Viktor, Doctor of Technical Sciences, Professor, Head of Department of Information Control System, Kharkiv National University of Radio Electronics, Ukraine, ORCID: http://orcid.org/0000-0002-7929-515X, e-mail: viktor.levykin@nure.ua
Iuriev Ivan, Assistant, Department of Information Control System, Kharkiv National University of Radio Electronics, Ukraine, ORCID: http://orcid.org/0000-0002-5178-519X, e-mail: ivan.iuriev@nure.ua

UDC 004.891.3
DOI: 10.15587/2312-8372.2019.160205

Chala O. DEVELOPMENT OF INFORMATION TECHNOLOGY FOR THE AUTOMATED CONSTRUCTION AND EXPANSION OF THE TEMPORAL KNOWLEDGE BASE IN THE TASKS OF SUPPORTING MANAGEMENT DECISIONS

The current state of organizational management is characterized by the wide use of knowledge bases (KB) to support management decisions for solving partially structured and unstructured tasks. The process of making management decisions involves the formation of a set of alternative context-sensitive sets of management actions, choice and implementation of the chosen decision. This choice is made by the decision maker (DM). Knowledge-oriented
decision support is implemented under uncertainty conditions as a result of incomplete information about the state of the management object and therefore requires constant updating of knowledge taking into account the evolution of the behavior of the object of organizational management.

This shows the relevance of methods and technologies development for knowledge bases creation to support management decisions in order to provide an opportunity to effectively identify the problem situation in the organization. Along with the creation the set of alternative decision to get out of this situation.

2. The object of research and its technological audit

The object of research is the process of building knowledge bases which involve the development of formal knowledge representation, knowledge extraction, verification of their contradictions and inclusion in the KB.

Extracting or externalizing knowledge consists in their transformation from implicit to explicit form [1]. Implicit includes expert knowledge and built-in knowledge. The latter are contained in typical procedures, processes, documents of the organizational system and can be obtained by analyzing the sequences of its states.

Then a formalization of the received explicit knowledge is carried out in accordance with the developed knowledge representation, which provides the possibility of using the latter for decision-making support.

The audit of the object of research showed the possibility of building a KB based on the analysis of the behavior of the organizational system, presented in the form of a sequence of its states. The relationship between states reflects the knowledge about control actions that were implemented consistently in time when performing management decisions.

3. The aim and objectives of research

The aim of research is development of the information technology of automated construction and replenishment of the temporal KB to provide effective support for the rational choice of administrative decisions of decision maker in the context of domain constraints.

In order to achieve the goal, the following tasks must be implemented:

1. To improve the model of logical-probabilistic presentation of temporal knowledge taking into account the characteristics of the proposed model of generalized temporal dependence and a hierarchical description of the management decision context.

2. To improve the method of construction and support of temporal KB taking into account the attributive description of the states of the control object, as well as the hierarchical model of the context of the management decision.

3. To integrate the logical-probabilistic representation of temporal knowledge and the method of constructing and supporting temporal KB in the framework of unified information technology for the creation of temporal KB.

4. Research of existing solutions of the problem

Management decision is a component of the management process and implemented under uncertainty [2]. To improve the effectiveness of management decisions at the tactical level of the organizational hierarchy apply decision support systems based on knowledge [3]. Such systems use input from transaction processing systems [4]. While constructing KB solutions for decision support systems, communicative methods of knowledge extraction are used [5]. These methods transform the implicit knowledge of experts into an explicit form [1]. The usage of such methods, in contrast to withdraw the embedded knowledge [6], requires a great deal of time spent by experts and knowledge engineers. This does not always allow you to maintain the KB in its current state.

In order to overcome the shortcomings of communicative methods, the paradigm of automated control of KB [7] is used. Automated management of KB realizes tasks of development and support of knowledge base [8], as well as the use of KB in search engines and information systems [9]. However, existing models and methods of automated construction of KB use declarative knowledge without taking into account temporal dependencies. Such dependencies are essential for building a management process, taking into account the sequence of processing elementary objects – artifacts [10]. This does not allow constructing a management decision in the form of a sequence of control actions associated with temporal dependencies [11].

The analysis of the approaches to the KB construction gives an opportunity to conclude that there is a discrepancy between the needs of knowledge-oriented support for management decisions and the possibilities of existing methods and technologies of interactive and automated construction of the KB. This indicates the promise of developing the technology of automated construction of the temporal knowledge base.

5. Methods of research

The proposed technology integrates a number of models and methods developed by the author [11–13]. Within the framework of technology, knowledge based on the proposed temporal rules is used. Temporal rules specify the time dependencies between the states $S_i$ of the organizational system or its components. Such dependencies reflect the control actions that led to the transition between these states. The state of the organizational system combines a set of states of elementary objects – artifacts, for example, documents, parts, products used in the process of its activities. The state of each artifact is determined by the current values of its attributes. The values of these attributes are recorded and entered into the database by the transaction processing system. As one of the attributes, a timestamp is always used, which provides the ability to set the temporal dependencies. The sequence of the states has the form:

$$\Pi = \{S_0, S_1, \ldots, S_{n+1}\}.$$  

Each state $S_j$ is defined through the set of variables $a$ that correspond to the attributes of the artifacts, as well as using the timestamp:

$$S_j = \{a: a \in A_{type} \lor a \in A_{id} \lor a \in A_{temp} \lor a \in A_{time} \} \quad (1)$$

where $A_{type}$ – identification attributes; $A_{id}$ – attributes of the action that led to the state $S_j$ of the control object; $A_{temp}$ – attributes of the time.
$A_{\text{infact}}$ - attributes that determine the properties of artifacts in a state $S_j$; $A_{\text{new}}$ - attributes of the time of occurrence of the state $S_j$.

The temporal knowledge of the control object behavior is presented in the form of states occurrence facts and temporal dependencies between them. The fact $f_j$ of occurrence of a state $S_j$ is given through a predicate $Q$ on the set of its attributes:

$$f_j = Q(A): A \subseteq S_j,$$

where $A = \{a_1^j, \ldots, a_k^j, \ldots a_l^j\}$ - the attributes of the state $S_j$.

The truth region $Y_j$ of such a predicate is determined by the admissible set of attribute values of the event:

$$Y_j = \{a_1^j: a_k^j \in \Lambda_{\text{type}} \vee a_k^j \in \Lambda_{\text{back}} \vee a_k^j \in \Lambda_{\text{infact}}\},$$

where $a_k^j$ - value $k$ - variable of the state $S_j$; $\Lambda_{\text{type}}, \Lambda_{\text{back}}, \Lambda_{\text{infact}}$ - the set of valid attributes values of the event type, actions and properties of the artifacts.

Temporal rules determine the ordering of states $S_j$ in time. The same relationships between states at the same time may occur for different sequences $\Pi$, that had implemented at different intervals. Therefore, the temporal rules in the knowledge base specify the order of facts in time. Such rules are formed based on patterns of temporal dependencies between these facts. The general representation of the temporal dependence $r$ has the following form:

$$r = f_j \text{KO} f_n,$$

where $K$ - the quantifier, which defines a set of state sequences for which this dependence will be executed; $O$ - temporal operator.

The developed temporal rules are used by the quantifiers of temporal logic $E$ (Exists) and $A$ (All). The first quantifier specifies the truth of the fact $f_n$ in at least one sequence of states $\Pi_n$, and the second - in all known sequence of states.

Temporal operators determine the type of connection between states. A pair of successive states is given by the operator $X$ (NeXt), a couple of states among which there are other states, is described by the operator $F$ (Future). A pair of states with given transition conditions between them is determined by the operator $U$ (Until).

A logical-probabilistic model of the temporal knowledge representation is developed using the Markov logical network apparatus to reflect the logical and probabilistic aspects of this knowledge. The logical aspect of the model is represented by the set of weighted rules $r$ by templates (4). The truth of these rules depends on the meaning of the facts $f_j$:

$$Kn_n = \{H_i, (r_i; w_i): r_i \in R_n; r_i = f_j \text{AO} f_n \Rightarrow w_i = \omega, r_i = f_j \text{EO} f_n \Rightarrow 0 \leq w_i < \Omega\},$$

where $w_i$ - the weight of rules $r_i$; $R_n$ - a subset of rules $R$, selected on the basis of a priori set restrictions for the subject domain; $\Omega$ - the maximum possible value of the weight of the rule with a quantifier $E$ for a given set of state sequences $\Pi=\{\Pi\}$; $H_i$ - a hierarchical description of the context based on the ratio between the artifacts.

Subsets of artifacts $A^r_j$ and state sequences $\Pi$ are used as input in the formation of a comprehensive management decision. They set the values of predicate $Q$ variables at specific moments of time.

For each temporal rule $r_i$ that describes the dependencies between states $S_i$ and $S_n$ on a set of sequences $\{\Pi\} \in \Pi'$, the factor-function is set: $\phi_i: \phi_i = \exp(w_i r_i)$. This function reflects the effect of the weight of the rule on the probability of a pair of states occurring in a plurality. The probability of implementing an arbitrary sequence of states is determined by the product of its factors:

$$P(\Pi') = 1/Z \prod_{j} \phi_i: r_i = f_j \text{KO} f_n, S_j, S_n \in \Pi',$$

where $Z$ - the partition function.

The function $Z$ is used for normalization and is given through the achievement of factors for all sequences from the set $\Pi'$:

$$Z = \prod_{i} \phi_i: r_i = f_j \text{KO} f_n, S_j, S_n \in \Pi'.$$

The probability of implementation of a new sequence of states $\Pi'$ as a possible management decision has the form:

$$P(\Pi') = 1/Z \exp(\sum w_i (f_j \text{KO} f_n)),$$

The expression (8) defines the probabilistic component of the temporal knowledge representation. The dependence $f_j \text{KO} f_n$ has a true value only if there is an appropriate relation between states $S_i$ and $S_n$ on the sequence $\Pi$. Therefore, in calculating the probability, let’s only take into account the weight of the rules that describe this sequence of states.

The proposed logical-probabilistic presentation of temporal knowledge combines the logical description (5) and function (8) to calculate the probability of possible alternative solutions $\Pi'$ and has the following form:

$$Kn = \{Kn_n, P(\Pi')\}.$$

In general, the knowledge representation (9) allows describe the possible sequences of management actions in the management solution, the contextual dependence of these actions, taking into account the hierarchy of artifacts, as well as determining the likelihood of implementation of the specified sequences.

The construction of the temporal knowledge base using the representation (9) is implemented with the help of the advanced method of automated construction and support of the temporal knowledge base. As the input the method use sequences $\Pi$, the existing context hierarchy $H$, and the deviation threshold $\varepsilon$ for actual and calculated probabilities $P(\Pi')$. The method includes the following stages:

1. **Stage 1.** Choice of patterns for the representation of temporal dependencies. Such dependencies can be represented in the form (4) or in the typical artifact processing sequences [10].

2. **Stage 2.** Construction or clarification of the hierarchical model of the context of execution of control actions. This model has the form of a hierarchy of classes of artifacts according to the results presented in [12]. The context hierarchy defines the basic knowledge structure in the subject area.
Stage 3. Selection of a subset of input data by matching the set of input artifacts and the context hierarchy. The feature of this stage is the ability to detail the data for a specific level of the context hierarchy.

Stage 4. Construction of facts \( f_j \). At this stage, the values of the variables characterizing the state \( S_j \) are substituted into the predicate (2).

Stage 5. Construction of temporal rules \( r \). It performed by the method [11].

Stage 6. Finding the weights of temporal dependencies. At this stage, let’s use the method [13], which is based on the method of searching the weights in Markov’s logical network and takes into account the features of temporal rules and constraints.

Stage 7. Placing rules in the knowledge base and checking the weight of the temporal dependencies.

Step 7.1. Calculate the probabilities of the input sequence of states according to (8) using all the rules in the knowledge base defining the specified sequences.

Step 7.2. Comparison of the obtained probability values \( P_{\text{new}}(\Pi) \) with the probability \( P_{\text{old}}(\Pi) \) of the input data. In case, if \( |P_{\text{new}}(\Pi) - P_{\text{old}}(\Pi)| > \varepsilon \) stage 6 and 7 are repeated for new and existing rules regarding \( \Pi \).

Stage 8. Addition of a hierarchical context description of the management actions implementation in the KB. At this stage, the hierarchical representation of the context in the KB is supplemented by attributes and their properties corresponding to the variable states in the input data.

6. Research results

The proposed technology is intended for the construction of temporal dependencies, which determine possible sequences of control actions (or corresponding sequences of the control object states) as a part of the management decision.

As the input use the results of the transaction processing systems work, which can store the organization’s states by monitoring results. An example of the status description of one of the processes in an enterprise that contains the properties of a workgroup, product, executors, and timestamp, as well as actions leading to the current state, is shown in Fig. 1.

![Fig. 1. An example of a state description in the input data](image)

Group=«L40 3rd»
organization_country=«fr»
resource=«Sophie»
role=«D_1»
concept_name=«Completed»
product=«PRD609»
lifecycle_transition=«Resolved»
timestamp=2012-05-14T11:04:40:02:00»

The technology contains the following stages:

Stage 1. Construction or selection of temporal dependencies for the knowledge representation. At this stage, a selection of a subset of temporal rules is performed using a combination of quantifiers and temporal operators.

Stage 2. Construction of a hierarchy describing the execution context of control actions. Depending on the subject area, the description can reflect the hierarchy of tasks, the organizational hierarchy, or determine the degree of detail of the processes performed in the organization. At this stage, links are established between classes of artifacts that set hierarchical links in a subject area.

The example on the Fig. 1 shows some levels of the organizational hierarchy: country «fr»; Working Group «L40 3rd»; Position (or role) of the artist «D_1».

Stage 3. Select a subset of input data according to the context hierarchy.

At this stage, the selection of input data is performed according to a certain level of the hierarchy. For example, data for a group of employees with the code «L40 3rd» are selected.

Stage 4. Select a subset of input data by a temporal sign.

The importance of this phase connected to the fact that the sequence of tasks can be changed as the evolution of the requirements and goals of the organization evolves. Data selection for a given temporal interval is performed according to the time stamp.

Stage 5. Construction of temporal rules based on selected and stage 1 temporal dependencies.

Step 5.1. Determining facts \( f_j \) for states from all \( \Pi \).

For example, for represented in Fig. 1 state description \( S_i \) will be the fact \( f_j \) for the rule-type \( \Pi \).

Step 5.2. Construction of temporal rules. At this step, the temporal rules for the state pairs \( S_i \) and \( S_a \) the method [12] are formed.

Stage 6. The weights determination of the received temporal rules by the method [13] taking into account the probability of occurrence of incoming sequences of states.

Stage 7. The context hierarchy clarification in the knowledge base using SQL queries that compare the existing hierarchy and the hierarchy received from the input data.

Stage 8. Adding new rules to the knowledge base. This stage is implemented with the help of SQL queries in the relational database – the storage of the knowledge base.

Step 8.1. Before modifying the KB for each rule, the same rule is checked, but with another weight.

Step 8.2. In the case that one of the rules (new or existing) is a constraint and then the constraint is placed into the KB.

Step 8.3. In the case of discrepancy of the weight of the rules, stage 8 is performed without checking the difference between the probabilities of the sequence of states.

Stage 9. Clarification of the weight of the temporal rules according to stage 7 of the method of construction and support of the temporal KB.

Stage 10. Examination of the received temporal rules’ semantic by expert. It is verified the conformity the new rules with the existing one and the description of the context of the execution of control actions. According to the results of the check, the rules can be removed from the KB.

Let’s consider an illustrative example of technology implementation.

The input sequences of the states and the intervals of their execution time are given:

\[
\Pi_1 = \{S_{11}, S_{12}, S_{13}, S_{14}, S_{15}\}, \quad 01.09.18-15.09.18,
\]

\[
\Pi_2 = \{S_{21}, S_{22}, S_{23}, S_{24}, S_{25}\}, \quad 15.09.18-31.10.18,
\]

\[
\Pi_3 = \{S_{31}, S_{32}, S_{33}, S_{34}, S_{35}\}, \quad 01.10.18-15.12.18.
\]
A hierarchical description of the context is represented by the distribution of states into two divisions of the organization, respectively:

\[ \{s_{11}, s_{21}, s_{22}, s_{23}, s_{31}, s_{32}\} \quad \text{and} \quad \{s_{31}, s_{31}, s_{31}, s_{24}, s_{25}, s_{36}, s_{36}\}. \]

It is necessary to build a knowledge base with the temporal rules of the types \( X \) and \( F \) for the tasks solved or solved, and the use of communicative methods of knowledge engineering. Therefore, the specified sequence of input data has the form:

\[ \Pi_1 = \{s_{21}, s_{23}, s_{32}\}, \quad \Pi_2 = \{s_{24}, s_{25}\}, \quad \Pi_3 = \{s_{34}, s_{35}\}. \]

Taking into account temporal constraints let’s obtain:

\[ \Pi_1 = \{s_{11}, s_{31}, s_{31}\}, \quad \Pi_2 = \{s_{24}, s_{25}\}. \]

For them, the following rules are formed:

\( (f_1, AX f_1), \quad (f_1, EF f_1), \quad (f_1, EX f_1). \)

The first rule is executed on both sequences and therefore is a constraint. According to [10] its weight is infinite. The weight of the other two rules is the same considering \( P(\Pi_1) = P(\Pi_2) = 0.5 \). It is calculated by the method [13] and is 4.1. According to the input data, the received rules belong to the level of the hierarchy of the second subdivision. At the clarification stage the weight rule \( (f_1, AX f_1) \) will replace the rule \( (f_1, EX f_1) \) if it is in the knowledge base. In the given example, three rules completely define the sequences \( \Pi_1 \) and \( \Pi_2 \), since they describe all dependencies between the three states. Therefore, weight refinement will only be needed if there are rules \( (f_1, EF f_1) \) or \( (f_1, EX f_1) \) that have other weights in the knowledge base.

7. SWOT analysis of research results

Strengths. The key advantage of the proposed technology is the possibility of obtaining temporal rules without the use of communicative methods of knowledge engineering. This allows you to replenish the knowledge base promptly, as far as it arrives information about the next state of the control object. Also, the probabilistic representation of temporal knowledge makes it possible to form a set of sequences of states of the object of management from the current to the target state and, thus, enable the decision maker to choose a rational management solution. Weaknesses. Weaknesses of this technology are associated with the assumption of completeness and the absence of contradictions in the input data. It is expedient to use technology only if there are such attributes of each state in the input data, allowing to uniquely determining its level in the context hierarchy, as well as the absence of errors when recording the time stamps of these states.

Opportunities. Additional features of the proposed technology are related not only to the temporal aspect, but also to the object and spatial characteristics of the input data. This will expand the application area of the results to geographic information systems.

Threats. The lack or inaccuracy of incoming data in case of failures when they are formed by transaction processing systems affects the correctness of the received temporal dependencies. Thus, late fixation operations of delays in the writing operations to the database can lead to the formation of false temporal rules. Also, as the transaction processing systems evolve, attribute and artifact names can be changed, which may lead to errors in the records of the occurrence of the control object state.

8. Conclusion

1. The logical-probabilistic model of temporal knowledge representation is improved by taking into account a hierarchical description of the context of a management decision. A hierarchical description of the context makes it possible to select a subset of elementary objects that are necessary for management, and thus simplify the construction of the solution. The proposed model provides the opportunity to support a rational choice from the set of alternative managerial decisions based on the probability of transition to the target state of the control object.

2. The method of automated construction and support of the temporal knowledge base is improved by taking into account the attributive description of the control object states, as well as the hierarchical model of the management decision context. The method involves the construction of temporal dependencies patterns, the formation and verification of the logical-probabilistic knowledge representation in the form of weighted temporal rules, as well as the formation of a context description of managerial decisions. The method provides the operational construction of the temporal knowledge base, rely on the usage of built-in implicit knowledge in the support managerial decision tasks.

3. Integration of the temporal knowledge representation model, as well as methods of constructing knowledge bases, identifying knowledge, finding the weight of temporal rules within the framework of a unified information technology for the automated creation of temporal knowledge bases is implemented. Information technology combines the possibility of forming presentation and semantic knowledge verification patterns, performed by a specialist in the subject area. And also the advantages of automatic construction weighted temporal rules are based on the detection of dependencies in the sequences of the control object states. This makes it possible to quickly identify new temporal dependencies characteristic for the subject area and place them into the knowledge base after semantic verification by an expert.

References


Chala Oksana, PhD, Associate Professor, Department of Information Control Systems, Kharkiv National University of Radio Electronics, Ukraine, e-mail: oksana.chala@nure.ua, ORCID: http://orcid.org/0000-0001-8265-2480