17. Metodolohiya vprovadzhennia tekhnolohiyi elektronnoho menedzhmentu heolohichnykh proektiv / Khlevna Yu. L., Yehoorchenkova N. Yu., Kubiavka L. B., Ivanova O. M. // Heoinformatyka: teoretychni ta prykladni aspekty. 2017.
18. Teslia I. N., Khlevna I. L., Oberemok N. V. Creation management of customized project management methodology // Project Management, System Analysis and Logistics. 2016. Issue 18.
19. Development of principles and method of electronic project management / Teslia I., Yehorchenkova N., Yehorchenkov O., Kataieva Y., Zaspa H., Khlevna I. // Eastern-European Journal of Enterprise Technologies. 2017. Vol. 5, Issue 3 (89). P. 23–29. doi: https://doi.org/10.15587/1729-4061.2017.109534

UDC 004.891.3
DOI: 10.15587/1729-4061.2018.142664

DEVELOPMENT OF A METHOD FOR THE PROBABILISTIC INFERENCE OF SEQUENCES OF A BUSINESS PROCESS ACTIVITIES TO SUPPORT THE BUSINESS PROCESS MANAGEMENT

V. Levykin
Doctor of Technical Sciences, Professor*
E-mail: levykinvictor@gmail.com
O. Chala
PhD, Associate Professor*
E-mail: oksana.chala@nure.ua
*Department of Information Control Systems
Kharkiv National University of Radio Electronics
Nauky ave., 14, Kharkiv, Ukraine, 61166

Запропоновано моделі темпоральних правил виконання дій бізнес-процесу, що застосовуються у випадку відсутності в процесній моделі повної інформації про причини виконання цих дій внаслідок втручання виконавців робіт. Правила формуються на основі аналізу порядку подій в журналі бізнес-процесу, що дозволяє визначити темпоральні умови й обмеження на виконання відповідних дій. Моделі правил можуть бути застосовані як елемент представлення знань для системи процесного управління, оскільки вони відображають записаний у журналі досвід реалізації бізнес-процесу. Використання правил дозволяє обмежити кількість можливих варіантів виконання бізнес-процесу з урахуванням його поточного стану. В результаті скорочується час прийняття рішень із процесного управління для випадку невідповідності поточного екземпляру бізнес-процесу та його моделі.

Запропоновано новий метод ймовірнісного виведення, який використовує представлені правила для того, щоб сформувати нові, допустимі послідовності дій у нетиповій ситуації, що виникла внаслідок коригування бізнес-процесу його виконавцями. Метод застосовує представлення знань на основі марківської логічної мережі, що дозволяє упорядкувати нові послідовності дій за ймовірністю їх реалізації з використанням зважених темпоральних правил. Використання комбінації правил для пар послідовних та пар віддалених в часі дій дозволяє підвищити точність обчислення ймовірності реалізації нових варіантів бізнес-процесу. Метод враховує інформацію із журналу подій при доповненні правил, що дозволяє забезпечити безперервне поповнення правил при виконанні бізнес-процесу. Зазначене обумовлює можливість практичного застосування методу при автоматизованій побудові та розширенні бази знань системи процесного управління у реальному часі

Ключові слова: бізнес-процес, марківська логічна мережа, ймовірнісне виведення, обмеження на виконання дій, журнал подій

1. Introduction

Enterprise process management is realized by means of managing business processes. The business process contains a sequence of actions in manufacture of products or creation of services that are of value to consumers. The business process management is performed using models of these processes [1].

The business process model contains all possible sequences of actions with indication of constraints, conditions and admitted results of these actions. As usual, terms and constraints specify resources (materials, executors, equipment) required for execution of actions. As a result of each action, state of the business process changes [2].

Effectiveness of process management is determined by completeness of knowledge about possible sequences of the business process actions in the current environment conditions. Adequacy of the process model depends on completeness of knowledge about the business process. Such knowledge is commonly used in modeling business processes. After completion of the business process, knowledge is refined and the model is corrected at the stage of analysis of the business process life cycle.

However, for the knowledge-intensive processes [3], solution of the problem of knowledge incompleteness is associated with the difficulties caused by knowledge flexibility and multivariances. Taking into account the current state of the subject field, qualified executors of these processes can create new sequences of the actions not included in the existing model. Executors typically change the process by using their implicit knowledge. This knowledge takes personal experience into consideration, it looks like contextual templates of actions and cannot be obtained by conventional methods of knowledge engineering. Personal knowledge is difficult for examining, therefore, after intrusion of the executor, it is necessary to evaluate the current state of the business process as well as the probability of transition from the current to the target state. Relevance of such an assessment consists in that its results can be used to support decision making when realizing the business process management.

2. Literature review and problem statement

It was shown in [1] that to solve the problems of process control, all knowledge on possible versions of execution of the business process should be included in its model. However, existing approaches make it possible just partially elucidate knowledge of executors. Study in this direction involves structuring of implicit knowledge of executors into a number of basic elements, in particular: professional understanding of tasks, corporate values, context-based experience [3]. The highlighted items are included in the knowledge objects in the business process model, as shown in [4]. From the point of view of process management, the key disadvantages of these approaches are as follows. First, the mentioned knowledge structuring does not allow one to represent the reasons for changing the sequence of the process actions in a particular situation as formal dependences, constraints and conditions of action execution. Secondly, it is impossible to prognosticate further admissible sequences of actions. Thirdly, singling out of knowledge elements requires qualified experts familiar with the processes in the given subject field.

An alternative approach to using knowledge of executors in the business process models is used in creation of virtual enterprises. As shown in [5], when this approach is used, ontology of subcontractors is formed. To complement this ontology, multi-agent technology is used. The use of ontologies enables adapting of the business process with taking into account features of the current sequence and capabilities of subcontractors. However, the subsequent selection of subcontractors and correction of the sequence of actions also depend on the formal knowledge of the business process and combined with the personal knowledge of executors. This limits application of the given approach to support the process management.

The process management support based on comparison of the known "as it should be" business process model and the "as is" model is realized with application of the process mining methods. The latter model reflects impact of implicit knowledge of executors on sequence of the business process actions. When constructing the "as is" model, event logs of the business processes are used. As shown in [6], this model can be formed as a graph with the use of the BPMN notation. Analysis of this graph allows one to represent knowledge of the business process in the form of conditions and constraints of execution of the sequence of works. In the case of more complex logs containing mixed sequences of events from several business processes, the required dependences can be obtained after filtering the log by the method proposed in [7]. In order to form such implicit dependences, application of relational nets has been proposed in [8] based on the results of analysis of the event log. Consideration of the studied approaches makes it possible to identify conditions and constraints for executing the business process actions. However, they do not ensure forecasting of the sequence of the business process actions in a new, non-standard situation, e.g. when an executor intervenes because of a conflict when accessing the resources of several business processes.

When resorting to forecasting in view of the lack of information on future external influences, it is necessary to consider not only logical but also probabilistic nature of the forecasted sequences of actions. Markov logic networks [9] are used to represent logical and probabilistic dependences. Adaptation of Markov logic networks for their application in process management was discussed in [10].

However, the problem of forecasting allowable sequences of actions with taking in account probabilities of execution of these actions relative to the current state of the business process for decision support of the process management still requires its solution.

3. The aim and objectives of the study

This study objective was to develop a method of probabilistic logical inference of admissible sequences of actions to achieve final state of the business process taking into account its current state.

To achieve this objective, the following tasks were solved:

– to develop models of temporal rules of execution of sequences of actions of a business process on the basis of analysis of its event log;

– to substantiate the possibility of the process management support based on the probabilistic inference in the Markov logic network with the use of rules of execution of the business process actions;

– to verify experimentally the proposed method of probabilistic inference.

4. Development of models of temporal rules of executing actions of the business process on the basis of analysis of its event log

Event logs have the following properties essential for probabilistic inference:

1) information on the business process execution is represented as a set of routes consisting of events;

2) each route contains a sequence of events $<e_1, e_2,..., e_j, e_{j+1},...>$ corresponding to the sequence of actions in a single execution of the business process;

3) each event records execution or change of state of the corresponding process action; events of the route are ordered according to time markers;

4) each event is characterized by a set of the same attributes with different values;

5) attributes of an event are at the same time attributes of artifacts, i.e. the objects used by the business process;

6) the set of values of the artifact attributes at the current time moment describes the conditions for starting actions of the business process at this moment.

Properties (2) and (3) determine advantage of one event over the other for each pair of consecutive events (e_j, e_{j+1}) from one route in executing the business process. This advantage is usually represented in the form of the rule "if e_j then e_{j+1}" which connects causes and consequences of relevant actions of the business process. However, in the case of the event log, this advantage is only defined in the temporal aspect: when the business process is executed, the e_j event occurs first and then the e_j+1 event occurs. Considering this aspect, advantage of the e_j event over the e_j+1 event is expedient to represent using the X(*neXt*) operator of the temporal logic [11]. This operator specifies truth of the event following the current one, so the dependence has the following form: e_jXe_{j+1}.

Definition 1. The atomic sequence for a pair of events from the log route is determined for a pair of events in the case when these events are recorded in two consecutive discrete moments of time τ_j and τ_{j+1}, that is, there is no other event between them:

$$\tau_{j+1} > \tau_j \Rightarrow e_j X e_{j+1} \Big| \forall \tau^* \neq \tau_j \ \tau^* < \tau_j \vee \tau^* \geq \tau_{j+1}. \quad (1)$$

Define the π route of the business process event log as a sequence of events in time with a given sequence for each pair of successive events:

$$\pi \models \left\{ e_j X e_{j+1} \,|\, \forall j \, \tau_{j+1} > \tau_j \right\}, \quad j = \overline{1, J-1}, \quad (2)$$

where e_j, e_{j+1} are two successive log events that arose at τ_j and τ_{j+1} time moments.

Based on the sequence for a pair of events, the entire log route is arranged making it possible to determine the possibility of achieving the final e_J event as follows:

$$\pi \models e_j F e_J \,|\, \exists e_j X e_{j+1} X \ldots X e_J, \quad j = \overline{1, J-1}, \quad (3)$$

where F is the operator of the temporal modal logic which specifies occurrence of an event in the future.

In accordance with properties (1)–(3), define constraints on the sequence of events of the business process route. The L event log consists of a set of π_i routes that were recorded during multiple execution of the business process: $L=\{\pi_i\}$. The atomic sequence for the event pairs (e_j, e_{j+1}) is strict if it is executed on all known routes of the log. Denote the strict sequence for the pairs of events with the help of index ∞ for the modal operator: X^∞. Next, define the strict sequence for a pair of events ($e_{i,j}$, $e_{i,j+1}$) from the route π_i as follows:

$$\forall \pi_i \in L \ \exists e_{i,j} X e_{i,j+1} \Rightarrow e_{i,j} X^\infty e_{i,j+1}. \quad (4)$$

The sequence $e_{i,j}X^\infty e_{i,j+1}$ specifies the set of known constraints on the admissible sequences of actions for execution of the business process. The strict sequence $e_{i,j}F^\infty e_{i,n}$ for a pair of events between which there are other events, is specified in the same way.

Definition 2. The C^∞ constraint on a sequence of events in a business process route is a conjunction of pairs of events for which a strict sequence is assigned.

The C^∞ constraints a set of potentially possible business process execution routes. That is, the condition $L=\{\pi_i|C^\infty\}$ must be met for all business process execution routes realized and known from the a priori model.

The set of rules C which define potentially admissible sequences of events consists of constraints C^∞ and conditions C^f. Unlike C^∞, C^f conditions may be violated with a non-zero probability.

Arrangement of events in accordance with (1) and (3) determines formation of two types of temporal rules. The sequence (1) specifies the rule of successive execution of two business process actions, that is, the rule of the *neXt*: $c_{neXt}=e_j, Xe_{j+1}$ type. The sequence of events (3) sets the rule of the *Future* type for execution of a sequence of actions which leads to the business process target action. This rule is set by the initial and target actions that are represented by the initial, e_j, and target, e_n, events on the log route. This rule has the form: $c_{Future}=e_j,Fe_n$. The generalized rule has the form: $c=(e_j,Xe_{j+1})\vee(e_j,Fe_n)$, $j<n\leq J$. Combining of the rules of both types to describe constraints and conditions allows one to represent in more detail admissible sequences of the business process actions.

In accordance with properties (1) and (2), the set of the business process routes recorded in the event log L contains only the sequences of actions that have been realized in practice. That is, the log contains only a subset of possible routes Π of the business process: $L\in\Pi$. The set of possible routes of the log contains a set of known routes as well as a set of ordered potentially admissible routes π_i^{new}:

$$\Pi = \left\{ L, \left\langle \pi_1^{new}, \ldots, \pi_i^{new}, \ldots, \pi_I^{new} \right\rangle \right\}. \quad (5)$$

The purpose of ordering of potentially admissible routes is to identify the routes most suitable for realization in the current environment conditions. Obviously, the routes in which the rules are fulfilled more often are typical, i.e. suitable for execution for a large number of the environment conditions. Such routes will have a smaller index in the tuple of expression (5).

Example 1. Let a double-executed business process is recorded in log L and it consists of the following completed event sequences:

$$L=\{\langle e_{1,1}, e_{1,2}, e_{1,4}, e_{1,5}\rangle, \langle e_{2,1}, e_{2,2}, e_{2,3}, e_{2,4}, e_{2,5}\rangle\}.$$

The log events are recorded according to the time of their occurrence, i. e. the second event $e_{1,2}$ appears later than the first event $e_{1,1}$: $\tau_{1,2}>\tau_{1,1}$, event $e_{1.4}$ later than $e_{1,2}$: $\tau_{1,4}>\tau_{1,2}$, etc. Each event contains a record of the business process execution. Action is defined by the second index of event, i.e. events $e_{1,1}$ and $e_{2,1}$ contain records of execution for the first action, $e_{1,2}$ and $e_{2,2}$ for the second action, etc. According to the expressions (2) and (3), the following temporal dependences between the pairs of events are fulfilled on the first route π_1:

$$\pi_1 \models \{e_{1,1}Xe_{1,2}, e_{1,2}Xe_{1,4}, e_{1,4}Xe_{1,5}, e_{1,1}Fe_{1,4}, e_{1,1}Fe_{1,5}, e_{1,2}Fe_{1,5}\}.$$

On the second route, the following dependences are true:

$$\pi_2 \models \{e_{2,1}Xe_{2,2}, e_{2,2}Xe_{2,3}, e_{2,3}Xe_{2,4}, e_{2,4}Xe_{2,5}, e_{2,1}Fe_{2,3}, e_{2,1}Fe_{2,4}, e_{2,1}Fe_{2,5}, e_{2,2}Fe_{2,3}, e_{2,2}Fe_{2,4}, e_{2,2}Fe_{2,5}, e_{2,3}Fe_{2,5}\}.$$

These dependences include conditions and constraints on the admissible sequences of the business process actions.

In accordance with expression (4), the constraint for a given business process combines all dependences that are fulfilled on both routes which makes it possible not to take into account the first index for events:

$C^{\infty}=(e_{1,1}Xe_{1,2}\wedge e_{1,1}Fe_{1,5}\wedge e_{1,2}Fe_{1,5})\wedge$
$\wedge(e_{2,1}\, e_{2,2}\wedge e_{2,1}Fe_{2,5}\wedge e_{2,}F\, e_{2,5})=e_1Xe_2\wedge e_1Fe_5\wedge e_2Fe_5.$

A set of conditions includes all other dependences that have been fulfilled on at least one of the routes of the business process. Therefore, it is expedient to delete the route index from description of these dependences:

$C^{\succ}=\{e_{1,2}Xe_{1,4},\, e_{1,4}Xe_{1,5},\, e_{1,1}Fe_{1,4},\, e_{2,2}Xe_{2,3},\, e_{2,3}Xe_{2,4},$
$e_{2,4}Xe_{2,5},\, e_{2,1}Fe_{2,3},\, e_{2,1}Fe_{2,4},\, e_{2,2}Fe_{2,4},\, e_{2,3}Fe_{2,5}\}=$
$=\{e_2Xe_4,\, e_4Xe_5,\, e_2Xe_3,\, e_3Xe_4,\, e_1Fe_3,\, e_1Fe_4,\, e_2Fe_4,\, e_3Fe_5\}.$

The conditions and constraints derived from L are in fact knowledge of the admissible behavior of new routes π_i^{new} which reflect advantage of one actions of the business process over others. Constraints C^{∞} should be fulfilled on each of these routes. The conditions for each business process route are determined by the conjunction of a subset of elements with $C^{\succ}$. In particular, the plurality of above-mentioned constraints and conditions $e_2Xe_3\wedge e_1Fe_3$ represent knowledge of such admissible new sequence of actions: 1, 2, 3, 5.

5. The method of probabilistic inference of admissible sequences of actions in relation to the current state of the business process

The proposed method uses knowledge representation based on the Markov logic network. The main feature of this representation is combination of logical and probabilistic components in describing logical facts and cause-and-effect relations to specify possible states and action sequences of knowledge-intensive business processes.

General representation of knowledge that takes into account sequence of log events is as follows:

$$KB=\left(\left\{e_{i,j}\right\},\left\{c_m^{\infty},w_m=\infty\right\},\left\{c_j^{\succ},w_j\right\},\, P(A=\{\alpha_{i,J}^{k}\}\big|e_{i,j})\right), \quad (6)$$

where c_m^{∞} is the constraint; w_m is the constraint weight; $c_j^{\succ}$ is the condition; w_j is the condition weight; $\alpha_{i,J}{}^{k}$ is the value k of the attribute of the current event $e_{\mathrm{i,J}}$; $P(A=\{\alpha_{i,J}{}^{k}\})$ is the probability of achievement of the target event $e_{i,J}$ specified by the set of values of attributes $\{\alpha_{i,J}{}^{k}\}$) in relation to the current state of the business process characterized by the event $e_{i,j}$.

In accordance with properties (4)–(6) of the event log, when calculating constraints c_m^{∞} and conditions $c_j^{\succ}$, each event is specified as the logical facts based on the values of attributes of the business process artifacts, i. e. $e_{i,j}=$ $=f_i(\{\alpha_{i,J}{}^{k}\})$. Later on, in formation of constraints and conditions, events (or the route) will be considered as a whole, without attribute detailing. The weight w_j for conditions $c_j^{\succ}$ is calculated taking into account probability of their occurrence in accordance with the expression traditional for the Markov networks:

$$P\left(A=\left\{\alpha_{i,J}^{k}\right\}\middle|e_{i,j}\right)=\frac{1}{Z}\exp\left(\sum_j w_j\, c_j^{\succ}\right), \quad (7)$$

where Z is the distribution function used for valuation.

Since the values of attributes $\{\alpha_{i,J}{}^{k}\}$ determine the final event of the route, the corresponding route π_i^{new} can be specified instead of them as an argument. The basic algorithm of calculating the weights of conditions is given in [9]. The weight of the rule determines probability of fulfillment of the corresponding dependence: the greater the weight, the oftener this dependence is fulfilled. The value of weight ∞ for constraints determines their fulfillment for all possible realizations of the business process.

The problem of probabilistic logical inference in representation of knowledge (6) has the following subproblems: verification of the possibility of successful completion of the business process in relation to its current state; identification of the best sequence of actions (events) for the successful completion of the business process.

The results of these subproblems enable support of decision making in the event that an executor changed sequence of the process and the new sequence was not executed by other versions of the business process, i. e. $\pi_i^{new}\notin L$.

The probabilistic inference method uses the following initial data: the event log L containing routes of the previous business process realizations as well as the current route; limit time for recording the event sequence, stl; conditions $C^{\succ}$for execution of actions. The method includes the following stages.

Stage 1. Filtration of the subset l of the input routes of the event log.

The business process log contains records of execution of a significant number of versions of the business process. Due to improvement of the business process, over time, the outdated routes may not correspond to its current behavior. Each event contains a time marker. Therefore, filtration of routes of the event log is performed using these markers according to the condition of the critical time of route recording, $\tau^l:\forall\pi_i\in l\tau_{i,j}\geq\tau^l$ where $\tau_{i,j}$ is the marker of time of the event j on the route, π_i.

Stage 2. Selection of conditions $C^{\succ}$ for the subset of the log routes, l.

At this stage, only those conditions $c_j^{\succ}$ are considered which connect events from a subset of routes with l, that is, the following condition is fulfilled: $\exists\pi_i$: $c_j^{\succ}$*true*.

Step 3. Selection of constraints C^{∞} and clarification of conditions for the subset of the log routes, l.

At this stage, the following dependences $c_j^{\succ}$ are selected from the set $C^{\succ}$ for which “$\pi_i\in l\; c_j^{\succ}=$*true* is fulfilled. These dependences are excluded from the $C^{\succ}$ and introduced into the set of constraints. From $C^{\succ}$, the conditions are additionally removed that include events from the set of constraints C^{∞}. For those conditions that are left, weights and values of the distribution function Z are calculated. The algorithms given in [9, 11] are used in calculation.

Stage 4. Selection of a subset of events E^l from the routes $\pi_i\in l$ that do not fulfill the constraints C^{∞} obtained in Stage 2. The condition for selection of the event e_j is as follows: “$e_j\in E^l\nexists c_m^{\infty}=$*true*. The route index is not used in this condition since events from all routes are selected. That is, one event from the E^l set can be recorded on several log routes. Additionally, the last event of the current route is deleted since it cannot be used to form new routes.

The purpose of this stage is to reduce the number of potentially possible routes that will be formed at Stage 5. The set of possible routes is defined as a sum of all possible placements for the selected set of events. The number and sequence of the events that satisfy the constraints is unchanged on all routes of the subset l which reduces the number of placements.

Stage 5. Formation of a set of potentially admissible routes, $\{\pi_i^{new}\}$.

5. 1. Formation of all possible placements from E^l by 1, by 2,..., by $|E^l|$.

5. 2. Supplement of the obtained placements with the sequences of events that satisfy the constraint of c_m^∞. Each pair of events is arranged on the route π_i^{new} according to its constraint c_m^∞.

Stage 6. Potential calculation for each route π_i^{new}. In accordance with expression (7), potential is calculated as a sum of weights of all conditions satisfied on the given route. Constraints are not counted in calculation because they are the same for all routes.

Stage 7. Calculation of probabilities of realization of each of the routes $Pot(\pi_i^{new})$ by formula (7).

Stage 8. Ordering of the routes π_i^{new} according to the values of probabilities, according to (5).

The result of this method is the set of potentially possible routes π_i^{new} ordered according to the values of probabilities.

Example 2. Let such initial data are used in the method of probabilistic inference.

1. The log L consisting of the following sequences of events:

$$L=\{\langle e_{1,1}, e_{1,2}, e_{1,3}, e_{1,5}\rangle, \langle e_{2,1}, e_{2,2}, e_{2,4}, e_{2,5}\rangle, \langle e_{3,1}, e_{3,2}, e_{3,3}, e_{3,4}, e_{3,5}\rangle, \langle e_{4,1}, e_{4,2}, e_{4,6}\rangle\}.$$

The event $e_{4.6}$ is the last event recorded in the log of the business process that is executed (that is, $\tau_{4,6}$ corresponds to the current time). The sequences of events are ordered according to the execution time of four business process versions, that is: $\tau_{4,6}>\tau_{4,1}>\tau_{3,5}>\tau_{3,1}>\tau_{2,5}>\tau_{2,1}>\tau_{1,5}>\tau_{1,1}$.

2. Limit time of recording the route which is $\tau^l>\tau_{1,5}$.

3. The set of conditions for all four sequences of events:

$$C^{\succ}=\{e_1Xe_2, e_1Fe_5, e_2Fe_5, e_2Xe_4, e_4Xe_5, e_2Xe_3, e_3Xe_4, e_1Fe_3, e_1Fe_4, e_2Fe_4, e_3Fe_5, e_3Xe_5, e_1Fe_6, e_2Xe_6\}.$$

The procedure for formation of these conditions is given in subsection 4.

The result of Stage 1. Since, $\tau^l>\tau_{1,5}$, the first sequence of events is deleted from further consideration and the subset l has the following form:

$$l=\{\langle e_{2,1}, e_{2,2}, e_{2,4}, e_{2,5}\rangle, \langle e_{3,1}, e_{3,2}, e_{3,3}, e_{3,4}, e_{3,5}\rangle, \langle e_{4,1}, e_{4,2}, e_{4,6}\rangle\}.$$

The result of Stage 2. After deletion of the first sequence of events from the set $C^{\succ}$, the dependence e_3Xe_5 associated with it is deleted and the set of conditions takes the following form:

$$C^{\succ}=\{e_1Xe_2, e_1Fe_5, e_2Fe_5, e_2Xe_4, e_4Xe_5, e_2Xe_3, e_3Xe_4, e_1Fe_3, e_1Fe_4, e_2Fe_4, e_3Fe_5, e_1Fe_6, e_2Xe_6\}.$$

The results of Stage 3. The dependences that are true on all routes with l are separated from the set of conditions. These dependences form constraints on execution of the business process actions. The conditions and constraints obtained are as follows:

$$C^\infty=e_1Xe_2\wedge e_1Fe_5\wedge e_2Fe_5,$$

$$C^{\succ}=\{e_2Xe_4(1{,}37), e_4Xe_5(0{,}47), e_2Xe_3(0{,}65), e_3Xe_4(0{,}37), e_2Fe_4(0{,}15), e_3Fe_5(0{,}2), e_2Xe_6(1{,}84)\}.$$

The conditions e_1Fe_3, e_1Fe_4, e_1Fe_6 are deleted from the set $C^{\succ}$ since they are decomposed into a sequence of conditions and constraints. In particular, the condition e_1Fe_3 is detailed through the constraint e_1Xe_2 and the condition e_2Xe_3. The weight of conditions and the value of the distribution function Z are calculated in accordance with [11]. The weight of conditions is given in the set $C^{\succ}$ in brackets after the dependence. The greater the weight the lower probability of rule violation. All constraints have weight ∞: they cannot be violated during the process. This determines exclusion from further consideration of the events related to constraints. The value of Z is 19.18.

The result of Stage 4. When forming a subset of events E^l in accordance with the constraint e_1Xe_2, the first event is not taken into consideration. Events e_2 and e_5 are also not taken into account in formation of placements because of action of the e_2Fe_5 constraint. The e_6 event represents current state of the business process and is also ignored. The resulting set has the form $E^l=\{e_3, e_4\}$.

The results of Stage 5. At Stage 5.1, the following sequences of events are formed from the events of the subset E^l: $\langle e_3, e_4\rangle$, $\langle e_4, e_3\rangle$. At Stage 5.2, the following new routes are formed:

$$\pi_5^{new}=\langle e_1, e_2, e_6, e_3, e_4, e_5\rangle,$$

$$\pi_6^{new}=\langle e_1, e_2, e_6, e_4, e_3, e_5\rangle\ \langle e_3, e_4\rangle, \langle e_4, e_3\rangle.$$

The results of Stage 6. The obtained route π_5^{new} is characterized by the following conjuncture of conditions: $e_2Fe_4\wedge e_3Xe_4\wedge e_3Fe_5\wedge e_4Xe_5$. The route π_6^{new} is characterized by the following conditions: e_2Fe_4. Potential of each of these routes forms the sum of weights of their conditions. For the given input data: $Pot(\pi_5^{new})$: 0.15+0.37+0.2+0.47=1.19; $Pot(\pi_6^{new})$=0.15.

The results of Stage 7. Probabilities of new routes have the following values: $P(\pi_5^{new}|e_6)=0.171$; $P(\pi_6^{new}|e_6)=0.087$.

The results of Stage 8. The resulting new routes are ordered by the probability value: $\langle\pi_5^{new}, \pi_6^{new}\rangle$.

The set of possible routes which, according to (5), combines the already realized and admissible sequences of events has the following form: $\Pi=\{\langle e_{2,1}, e_{2,2}, e_{2,4}, e_{2,5}\rangle, \langle e_{3,1}, e_{3,2}, e_{3,3}, e_{3,4}, e_{3,5}\rangle, \langle\pi_5^{new}, \pi_6^{new}\rangle\}$.

6. Experimental verification of the method of probabilistic inference on the example of after-sales service process

A simplified model of the business process of after-sales service of electronic devices was considered. The process description in the log contained 5 typical routes with two sequences of events. At the moment of simulation, the business process was in a state of transition to a warranty repair. This status was represented in the log by the current event $e_{i,10}$. The state was non-typical since the event $e_{i,10}$ did not belong to known completed routes.

The objective of the experimental verification was, firstly, to determine a set of potentially admissible routes π_i^{new} from the current event $e_{i,10}$ to the known final event $e_{i,8}$. Secondly, it was necessary to make sure that the new routes were in line with the existing approaches to service. In order to achieve the stated objective of the experiment, potentially admissible routes were formed, their probabilities were determined and the events of these routes were analyzed taking into account description of the relevant business process actions.

When realizing the probabilistic inference, the log of events of the business process as well as the conditions constraints on execution of actions were used.

6.1. Routes of the event log

Incoming routes of the event log are listed in Tables 1, 2.

Table 1 contains the log of the business process consisting of 5 known routes: $\pi_1, \pi_2, \pi_3, \pi_4, \pi_5$ and one current route π_6^{new}. Each event has two indices: the route index and the business process index. For example, the $e_{5,9}$ event is the third event of the route π_5 which performs the ninth action of the business process.

The indexed list of the process actions relevant to the log's events is given in Table 2.

The graph of the process is shown in Fig. 2. For the events from the potentially admissible routs, π_i^{new}, $i=\overline{6,I}$, the first index is not shown for simplicity.

Table 1

Routes of the business process event log

Routs	Events sequence
π_1	$\langle e_{1,1}, e_{1,2}, e_{1,3}, e_{1,4}, e_{1,5}, e_{1,6}, e_{1,7}, e_{1,8}\rangle$
π_2	$\langle e_{2,1}, e_{2,2}, e_{2,3}, e_{2,4}, e_{2,5}, e_{2,6}, e_{2,7}, e_{2,8}\rangle$
π_3	$\langle e_{3,1}, e_{3,2}, e_{3,3}, e_{3,4}, e_{3,5}, e_{3,6}, e_{3,7}, e_{3,8}\rangle$
π_4	$\langle e_{4,1}, e_{4,2}, e_{4,9}, e_{4,7}, e_{4,8}\rangle$
π_5	$\langle e_{5,1}, e_{5,2}, e_{5,9}, e_{5,7}, e_{5,8}\rangle$
π_6^{new}	$\langle e_{6,1}, e_{6,2}, e_{6,10}, \ldots\rangle$

Table 2

Business process actions

Event index	Description
1	Application for the device repair
2	Disassembly of the device
3	Malfunction diagnosis
4	Reconciliation with the customer on the repair
5	Purchase of an assembly or a component
6	Replacement of an assembly or a component
7	Assembly of the device
8	Payment and transfer of the device to the customer
9	Cleaning of the device
10	Transition of the faulty assembly to the warranty repair

The set of possible routes is the set of all business process event locations, excluding the current event, $e_{i,10}$. That is, the length of the possible route can vary from 1 to 10 events. However, taking into account the constraints and absence of cycles, it is necessary to exclude events $e_{i,1}$, $e_{i,2}$, $e_{i,10}$ from the potential route. Also, the last event $e_{i,8}$ is excluded since the achievement of this event indicates completion of the business process. Thus, the set of possible routes is equal to the sum of the number of placements from 6 by 1, from 6 by 2, ..., from 6 by 6 and amounts $|\{\pi_i^{new}\}|=1965$.

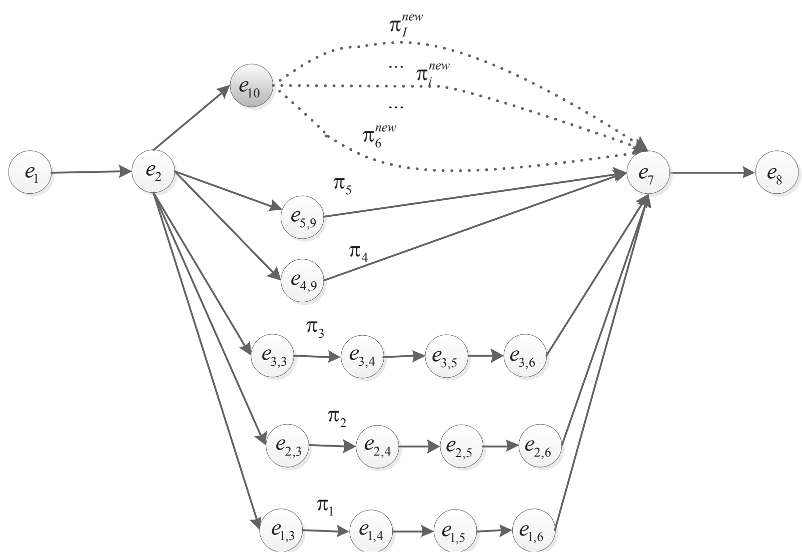

Fig. 2. Known $\pi_1, \pi_2, \pi_3, \pi_4, \pi_5$ and possible routes $\{\pi_i^{new}\}$ of the simplified business process of after-sales service

6. 2. Conditions and constraints for the sequence of actions for potentially admissible routes of the business process

This section presents the results of Stage 3 of the probabilistic inference method. Constraints are given in Table 3. All constraints were assigned weight ∞ to ensure their unconditional execution within the framework of the model of knowledge representation based on the Markov logic network [9].

Table 3

Constraints of the business process

Constraints	Content
e_1Xe_2	Application for the device repair > Disassembly of the device
e_7Xe_8	Assembly of the device > Payment and transfer of the device to the customer
e_1Fe_7	Application for the device repair >> Assembly of the device
e_1Fe_8	Application for the device repair >> Payment and transfer of the device to the customer
e_2Fe_7	Disassembly of the device >> Assembly of the device
e_2Fe_8	Disassembly of the device >> Payment and transfer of the device to the customer

The following constraints do not use the route index for events because these events belong to all log routes.

Later on, when determining new routes, the first index will also not be used since all events with the same action index reflect execution of the same action for different realizations of the business process.

As a result of applying the above constraints and taking into account the current state of the business process e_{10}, the maximum number of elements of the new route was 4: $\{e_4, e_5, e_6, e_9\}$. The number of possible routes $|\{\pi_i^{new}\}|=64$, that is, we have managed to reduce complexity of the forecasting task for this business process by about 30 times.

The conditions for the business process are presented by the rules of the *NeXt* and *Future* types given in Table 4.

This list is a subset of all rules in the set of the log routes, $\pi_1, \pi_2, \pi_3, \pi_4, \pi_5$. Each potentially admissible route π_i^{new} is characterized by a conjunction from a subset of the given conditions. The sum of weights of such conditions is the route potential and, according to (7), sets the probability of its realization.

Table 4

Conditions for execution of the business process

Condition	Weight	Routes
e_2Fe_4	0.085714	π_1, π_2, π_3
e_2Fe_5	0.042857	π_1, π_2, π_3
e_2Fe_6	0.014286	π_1, π_2, π_3
e_3Fe_5	0.085714	π_1, π_2, π_3
e_3Fe_6	0.042857	π_1, π_2, π_3
e_3Fe_7	0.014286	π_1, π_2, π_3
e_4Fe_6	0.085714	π_1, π_2, π_3
e_4Fe_7	0.042857	π_1, π_2, π_3
e_5Fe_7	0.085714	π_1, π_2, π_3
e_2Xe_3	0.900000	π_1, π_2, π_3
e_2Xe_9	2.702733	π_4, π_5
e_3Xe_4	0.900000	π_1, π_2, π_3
e_4Xe_5	0.900000	π_1, π_2, π_3
e_5Xe_6	0.900000	π_1, π_2, π_3
e_6Xe_7	0.900000	π_1, π_2, π_3
e_9Xe_7	2.702733	π_4, π_5

6. 3. The results of the probabilistic inference

The results of the method functioning are shown in Table 5. This table contains 12 routes with the highest probability of the 64 routes that resulted from the inference.

Index *i* of the π_i^{new} route is given in the first column. The column "Potential" contains the sum of weights of all rules in the "Conditions" column.

Analysis of the four most probable routes, π_{12}^{new}, π_{11}^{new}, π_{10}^{new}, π_9^{new}, has shown the following.

The route π_{12}^{new} had probability of 0.1298 and contained the following sequence of works: application for repair of the device (1); disassembly of the device (2); fault diagnostics (3); transition of the faulty assembly to warranty repair (10); replacement of the assembly or the component (6); reconciliation of repair with the customer (4); purchase of the assembly or the component (5); cleaning of the device (9); assembly of the device (7); payment and transfer of the device to the customer (8).

This route largely reflected the logic of service in the event that the component needed for repair was in stock: at first, replacement was done and then the component was purchased. The route also included an action to reconciliate repair with the customer. In practice, such a reconciliation may take place in the case of replacement of the faulty component, not by the same but by one with similar characteristics.

The route π_{11}^{new} had probability of 0.1243 and included the following key actions after transition to the warranty repair: purchase of the assembly or the component (5); Replacement of the assembly or the component (6); cleaning of the device (9). This route fully reflected logic of the warranty repair.

The route π_{10}^{new} had probability of 0.0573 and combined the action of cleaning the device (9) as well as the sequence of works from the route π_1. This route fully reflected logic of the warranty repair.

The route π_9^{new} had probability of 0.0573 and contained a sequence of works from the route π_1. This route fully reflected logic of the warranty repair.

7. Discussion of the results obtained in the study of the method of probabilistic inference

The method of probabilistic logical inference obtained in the study uses the knowledge representation based on

Table 5

Results of the probabilistic in*Fe*rence

i	Event index										Potential	Route probability	Conditions
1	1	2	3	10	4	9			7	8	2.7599	0.0426	$e_3Fe_6, e_3Fe_7, e_9Xe_7$
2	1	2	3	10	6	9			7	8	2.7599	0.0426	$e_3Fe_6, e_3Fe_7, e_9Xe_7$
3	1	2	3	10	4	6	9		7	8	2.8027	0.0444	$e_4Fe_7, e_3Fe_7, e_9Xe_7$
4	1	2	3	10	6	4	9		7	8	2.8027	0.0444	$e_3Fe_6, e_4Fe_7, e_3Fe_7, e_9Xe_7$
5	1	2	3	10	5	4	9		7	8	2.8456	0.0464	$e_3Fe_5, e_4Fe_7, e_3Fe_7, e_9Xe_7$
6	1	2	3	10	5	9			7	8	2.8884	0.0484	$e_3Fe_5, e_5Fe_7, e_3Fe_7, e_9Xe_7$
7	1	2	3	10	5	4	6	9	7	8	2.9742	0.0528	$e_3Fe_5, e_3Fe_6, e_3Fe_7, e_5Fe_7, e_3Fe_5, e_4Fe_7, e_9Xe_7$
8	1	2	3	10	6	5	4	9	7	8	2.9742	0.0528	$e_3Fe_5, e_3Fe_6, e_3Fe_7, e_5Fe_7, e_3Fe_5, e_4Fe_7, e_9Xe_7$
9	1	2	3	10	4	5	6		7	8	3.0571	0.0573	$e_3Fe_5, e_3Fe_6, e_3Fe_7, e_4Xe_5, e_4Fe_6, e_4Fe_7, e_5Xe_6, e_5Fe_7, e_6Xe_7$
10	1	2	3	10	9	4	5	6	7	8	3.0571	0.0573	$e_3Fe_5, e_3Fe_6, e_3Fe_7, e_4Xe_5, e_4Fe_6, e_4Fe_7, e_5Xe_6, e_5Fe_7, e_6Xe_7$
11	1	2	3	10	5	6	9		7	8	3.8313	0.1243	$e_3Fe_5, e_3Fe_6, e_3Fe_7, e_5Xe_6, e_5Fe_7, e_9Xe_7$
12	1	2	3	10	6	4	5	9	7	8	3.8742	0.1298	$e_3Fe_5, e_3Fe_6, e_3Fe_7, e_4Xe_5, e_4Fe_7, e_5Fe_7, e_9Xe_7$

Markov logic network. The key elements of this representation are the rules of modal logic which connect events of the business process routes and specify the conditions and constraints on execution of actions. These conditions and constraints are determined depending on the context of realization of the business process.

Experimental verification of the proposed method confirmed possibility of forecasting admissible sequences of the business process actions using rules derived from analysis of the event log.

The sets of conditions and constraints on execution of actions for a simplified business process of repair of electronic devices were determined. As a result of probabilistic inference, 64 routes were formed for transition from the non-typical state to the final state of the business process. Analysis of four routes with the highest probability value has shown that three of the four routes fully corresponded to the conventional sequence of repair works. One route determined non-standard sequence of work for the case when the components for repair are available in the stock. The indicated difference of this route related to the fact that only five routes were used for training with the subsets $\{\pi_1, \pi_2, \pi_3\}$ and $\{\pi_4, \pi_5\}$ having the same sequences of actions each.

It should also be noted that the use of conditions of *Future* type enables more accurate determination of sequences of execution of the business process actions. The estimated probability of successful completion of ten most likely sequences increased by an average 21.2 % taking into account these rules.

The proposed method does not take into consideration the cyclical execution of the business process as well as the attributes of events. This issue requires further studies. Taking into account the event attributes allows one to generalize or detail rules. The generalized rules simplify the search for potentially possible paths of execution of the business process and the detailed rules make it possible to more accurately distinguish these trajectories.

The advantage of the proposed method consists in that it provides the possibility of continuous adaptation of recommendations in the course of execution of the business process since the conditions and constraints on execution of actions can be adjusted on the basis of new events included in the log.

Forecast of promising lines of creating software ecosystems is the rather interesting field of application of the proposed method. Existing approaches are time-consuming and use parametric identification [13] and ontologies [14]. Use of the event logs of existing information systems makes it possible to identify constraints and conditions that are essential for creating a friendly behavior in the information environment. New paths of behavior of information systems can be formed taking into account these constraints and conditions.

8. Conclusions

1. Models of business process execution rules were developed. They combine constraints and conditions for execution of each possible action sequence. The set of rules defines priorities for execution of actions for a particular state of the process in the conditions of incomplete information about the reasons for execution of these actions. This makes it possible to compare versions of the business process execution in a case of unforeseen external influences on the course of work. Also, the proposed rule models ensure automating of construction of knowledge bases for the process management systems by defining the sequence in pairs of events by the routes of the business process log and further distributing the obtained dependences on the constraints and conditions for action execution. Representation of rules in the form of ordered pairs of events enables supplement of the knowledge base in real time, as new events are recorded to the business process log.

2. A method of probabilistic inference of sequences of the business process actions was developed. The essence of the method is construction and ordering, according to probabilistic assessment, of a plurality of new versions of the business process execution in relation to the current non-standard state using knowledge in the form of rules of action execution. The method makes it possible to cut the time of decision making for the business process management based on the choice of a new sequence of actions from an ordered set of versions admissible for execution. Also, the method provides the opportunity to supplement the business process model with such versions of actions that take into account changes in the state of the subject field or new external influences on activities of the enterprise at which the business process is executed.

3. Experimental verification has shown that the new sequences of actions with the highest probability values obtained by the proposed method reflect the existing logic of execution of the business process. The proposed combination of conditions and constraints on execution of actions makes it possible to reduce, at least by an order of magnitude, the number of admissible sequences of the business process works and, accordingly, the needed computing resources when using the method of probabilistic inference. Addition of conditions for arbitrary pairs of actions to conditions for pairs of sequential actions enables more precise (more than 20 %) determination of probability of the obtained routes.

References

1. Polyvyanyy A., Smirnov S., Weske M. Business Process Model Abstraction // Handbook on Business Process Management 1. 2014. P. 147–165. doi: https://doi.org/10.1007/978-3-642-45100-3_7
2. Müller D., Reichert M., Herbst J. Data-Driven Modeling and Coordination of Large Process Structures // On the Move to Meaningful Internet Systems 2007: CoopIS, DOA, ODBASE, GADA, and IS. 2007. P. 131–139. doi: https://doi.org/10.1007/978-3-540-76848-7_10
3. Gronau N., Heinze P., Weber E. Conversion and Transformation-Oriented View on Knowledge Flows // Modeling and Analyzing knowledge intensive business processes with KMDL. 2012. P. 249–268.

4. A Proposal to Model Knowledge in Knowledge-Intensive Business Processes / Vladova G., Thim C., Weber E., Ullrich A., Gronau N. // Proceedings of the Sixth International Symposium on Business Modeling and Software Design. 2016. doi: https://doi.org/10.5220/0006222600980103
5. Information support for business processes at virtual enterprises with multi-agent technologies / Pavlenko V., Shostak I., Morozova O., Danova M. // 2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies (DESSERT). 2018. doi: https://doi.org/10.1109/dessert.2018.8409189
6. Process mining using BPMN: relating event logs and process models / Kalenkova A. A., van der Aalst W. M. P., Lomazova I. A., Rubin V. A. // Software & Systems Modeling. 2015. Vol. 16, Issue 4. P. 1019–1048. doi: https://doi.org/10.1007/s10270-015-0502-0
7. Causality-based model checking in business process management tasks / Sergii C., Ihor L., Aleksandr P., Ievgen B. // 2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies (DESSERT). 2018. doi: https://doi.org/10.1109/dessert.2018.8409176
8. Implementation of search mechanism for implicit dependences in process mining / Kalynychenko O., Chalyi S., Bodyanskiy Y., Golian V., Golian N. // 2013 IEEE 7th International Conference on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS). 2013. doi: https://doi.org/10.1109/idaacs.2013.6662657
9. Richardson M., Domingos P. Markov logic networks // Machine Learning. 2006. Vol. 62, Issue 1-2. P. 107–136. doi: https://doi.org/10.1007/s10994-006-5833-1
10. Levykin V., Chala O. Method of automated construction and expansion of the knowledge base of the business process management system // EUREKA: Physics and Engineering. 2018. Issue 4. P. 29–35. doi: http://dx.doi.org/10.21303/2461-4262.2018.00676
11. Konur S. A survey on temporal logics for specifying and verifying real-time systems // Frontiers of Computer Science. 2013. Vol. 7, Issue 3. P. 370–403.
12. Lowd D., Domingos P. Efficient Weight Learning for Markov Logic Networks // European Conference on Principles of Data Mining and Knowledge Discovery. PKDD 2007: Knowledge Discovery in Databases. 2007. P. 200–211. doi: https://doi.org/10.1007/978-3-540-74976-9_21
13. A retrospective analysis technology of the Green Software Ecosystems development on the parametric identification of the Brown's model / Shostak I., Danova M., Romanenkov Yu., Kuznetsova Yu. // 2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies (DESSERT). 2018. doi: https://doi.org/10.1109/dessert.2018.8409197
14. Shostak I., Volobuyeva L., Danova M. Ontology based approach for green software ecosystem formalization // 2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies. 2018. doi: https://doi.org/10.1109/dessert.2018.8409196