Dynamic integrated expert systems: automated construction features of temporal knowledge bases with using problem-oriented methodology

Rybina G.V., Sorokin I.A., Sorokin D.O.
National Research Nuclear University MEPhI (Moscow Engineering Physics Institute)
Moscow, Russian Federation
galina@ailab.mephi.ru

***Abstract*—The work is focused on the problems of integration in the dynamic integrated expert systems (IES) architecture in the context of integration paradigm of artificial intelligence with models, methods and tools from other domains. These systems are developed basing on the problem-oriented methodology and the AT-TECHNOLOGY WorkBench. In this paper models, methods and software, for implementing a combined method of acquisition temporal knowledge from various sources of knowledge (experts, natural language texts, databases) are considered.**

***Keywords*—dynamic integrated expert systems, problem-oriented methodology, AT-TECHNOLOGY WorkBench, temporal knowledge bases, integration, temporal knowledge acquisition, temporal knowledge representation, temporal processing, simulation**

I. Introduction

At educational-scientific laboratory "Intelligent systems and technologies" of the Department of Cybernetics of NRNU MEPhI are developing dynamic integrated expert systems (IES) with AT-TECHNOLOGY — software platform that implements problem-oriented methodology [1], [2]. This platform supports and automates processes of prototyping and maintaining IES throughout their lifecycle. For a number of criteria (such as knowledge representation models, reasoning tools, object-oriented design support, etc.) AT-TECHNOLOGY is comparable to G2 (Gensym corp., US) [3] — leading software platform for real-time expert systems. Considering the built-in subsystem of outer world simulation, AT-TECHNOLOGY even goes ahead of G2. While G2 and some other tools lack automated knowledge acquisition, out platform offers original hybrid knowledge acquisition tools enabling fuzzy and temporal knowledge acquisition from various sources The dynamic version of the AT-TECHNOLOGY WorkBench that supports problem-oriented methodology for building dynamic IES is being actively developing (described in [1], [2], [4], and others.).

II. Some integration features of modern IES in the dynamic IES architecture

In the context of solving the modern IES construction problems (in particular for the control of complex discrete systems), problem-oriented methodology [1], which is conceptually related set of models, methods, algorithms and standard procedures to create applied IES of different typologies and level of complexity, has the following properties [1], [2]: a powerful combination method of acquiring knowledge that supports the automated process of acquiring knowledge from the sources of knowledge of different typology (experts, databases, texts) is used to gain knowledge; generalized knowledge representation language designed for building models of problem areas in dynamic IES allows to represent temporal knowledge, based on a modified interval Allen logic [5] and time control logic [6], together with the basic knowledge, including those containing knowledge with uncertainty, imprecision and vagueness; supports the use of various output means (universal AT-SOLVER and a specialized temporal solver designed for dynamic tasks); in the context of enhanced functionality and principles of the components IES deep integration provides the possibility of implementing simulation techniques for modeling the external environment and how to interact with them: the high efficiency of the a large number of applied IES development, including dynamic areas of concern;instrumentally supported by a modern software such as WorkBench (complex AT-TECHNOLOGY).

It should be noted that the conceptual basis of this methodology is a multi-level integration into the IES processes model, modeling specific types of tasks, relevant traditional expert systems technologies, methods and ways of building software architecture IES and its components at every level of integration, etc. In particular, in the dynamic IES important place is given to the integration of methods and means of temporal information presentation and processing with the methods and means of the outside world simulation, in this case discrete complex technical systems(CTS) in real time. This leads to expansion of the intelligent control systems architecture, built on the concept of dynamic IES relevant subsystems adequately reflecting all the processes and laws of functioning simulated systems, as an integral phase of building dynamic IES.

Using [1], [2], let's take the example of a integration model components of simulation discrete CTS models in language RAO^{AT} [2] with the other components of dynamic IES

Language RAO^{AT} is used here for a description of discrete CTS simulation models. This language was developed on the basis of well-known Russian method of RAO (Resources, Actions, Operations), [7] which allow to maximize fully reflect the behavior of any discrete type CTS (change of state resources, the emergence of regular and irregular events, and so on.)

It is important that the description of the CTS resources and description of the objects in the problem areas of the developed language of knowledge representation are conceptually close, allowing the temporal solver at AT-TECHNOLOGY complex to use parameters of the resources from the simulation model of discrete CTS through the working memory in the temporal withdrawal. So are all the prerequisites for the integration of simulation technology with dynamic IES technology in modern intelligent control systems and etc.

III. Automated acquisition, representation and processing of temporal knowledge in dynamic integrated expert systems

Acquisition, representation, and processing of temporal knowledge (i.e., knowledge considering time as an entity of a problem domain) play important role in the context of the systematic approach to development of dynamic integrated expert systems. In [1], [2] we have described basic models of knowledge representation and inference tools of AT-TECHNOLOGY AT-SOLVER.

Here we consider temporal aspects of inference perfomed on knowledge bases containing some unreliable knowledge, i.e., knowledge with such negative factors as uncertainty, inaccuracy, fuzziness, and with constraints on variables. Generalized model of temporal inference with production rules for dynamic integrated expert systems involves processing of knowledge with temporal dependencies together with basic knowledge of the problem domain.

We see the purpose of temporal inference in construction of the event flow model interpretation and in generation of a list of controlling actions for the problem domain. So we applied Allen's logic [5] with some enhancements together with Osipov's logic of control over time [6]. We define global event flow model by a set of temporal objects (events and intervals). Local event flow models in rules are defined by formulas of Allen's logic. Event flow model interpretation may be given as a set of timestamps of events and intervals.

To represent temporal knowledge in dynamic integrated expert systems we enhanced knowledge representation language of AT-TECHNOLOGY — AT-KRL [8]. Now it allows representing temporal knowledge together with basic knowledge including knowledge with uncertainty, inaccuracy, and fuzziness. To do so, we introduces new basic types of objects: events and intervals; new type of object properties: condition of event occurrence.

We modified the structure of rule antecedents: we added local event flow model requirements. At last, Osipov's control over time concepts lead to adding new rule types into AT-KRL: *reactions*, aimed to provide quick response to certaion, usually urgent, events in probles domain,and *periodic rules*, aimed to track certain duty cycles.

Reaction rules generally correspond to enhanced Allen's logic. Their antecedents are formulas where each operand is a single temporal object (an event or an interval). Antecedents of periodic rules contain extra condition with firing period. These enhancements of AT-KRL allow us to describe temporal relationships of objects in a problem domain by rules. Decision making is now performed taking into account actual event flow of the problem domain.

As for inference process the major changes were made in matching procedures: we implemented evaluation of formulas containing temporal arguments in rule antecedents and construction of event flow interpretation on each inference loop. When forming event flow interpretation, events and intervals are bound to the time axis by identifying the facts of their occurrence and considering the history of events. Processing of temporal parts of antecedents uses the results of event flow interpretation construction. For active rules AT-SOLVER matches local event flow models with constructed event flow interpretation.

Thus, the synergy of AT-SOLVER and temporal tools addresses both static and dynamic domains. Note that complex discrete systems produce input data for temporal inference in dynamic integrated expert systems. Issues related to models, methods, algorithms, and software for simulation modeling are considered in a number of papers, for example [8], [9].

In expert systems development we can automate experts' work by implementing methods and tools for detecting and extracting temporal knowledge from natural language texts (NL-texts)(Text Mining and Natural Language Processing) and from databases (Data Mining and Knowledge Discovery in Databases). In world practice there is a number of approaches to temporal dependencies acquisition but most of them are focused on processing of English-language texts. Moreover, they do not consider obtaining temporal knowledge for temporal knowledge bases for dynamic intelligent systems and for dynamic integrated expert systems in particular.

The combined method of knowledge acquisition (CMKA) [1], [2] has proven its efficiency in development of static integrated expert systems with AT-TECHNOLOGY. It automates interviewing of experts using natural sublanguage (business prose style), data mining, and verification of knowledge bases. In [2], [4] we described the client-server architecture and tools for knowledge acquisition from geographically distributed sources of knowledge of various types: experts, natural language texts, databases.

In dynamic integrated expert systems methods of automated detection of temporal knowledge remain an unexplored problem. In particular, for extracting information about time from texts in Russian, only the few approaches are proposed that pertly help to automate these processes, e.g. [10], [11]. Therefore, we focused on further evolution of combined method of knowledge acquisition and especially its temporal enhancement by developing new methods and tools for automated construction of temporal knowledge bases in dynamic integrated experts systems.

Our approach to knowledge acquisition (directly from experts by automated inter-viewing) bases on original technique of using patterns for solving typical problems [1]. We have put meta knowledge about strategies of solving into heuristic solving patterns for specific cases: diagnostics, engineering, planning, control, learning, and some other. To support these solving patterns we developed a number of methods and tools for modeling dialog scenarios used in interviewing. These methods [1] address both thematic structures of dialog, i.e. problem solving pattern, and local structure of a dialog, i.e. dialog steps –specific actions and reactions between n the expert and the system.

As computer-aided interviewing of an expert goes on, the problem solving pattern fills with structured data that can be exported to knowledge base. To derive the "action-reaction" model of dialog we use several techniques, e.g., simulation of consultation. Interviewing of experts is carried out automatically by dialog scenario interpreter.

The interpreter also generates dialog screens for entering answers and data including such things as uncertainty, imprecision, fuzziness. The specialized linguistic processor and s asset of dynamically replenished dictionaries [1] support knowledge acquisition process. Computer-aided interviewing of expert, natural language processing, data mining from databases are tightly coupled in AT-TECHNOLOGY [2], [4], [8], [9].

We developed a technique of detection and interpretation of the simplest temporal pointers (i.e., independent individual words and phrases denoting time) within a single sentence. We used generic classification of temporal pointers presented in [10], [11] together with vocabulary of Russian-language lexemes indicating temporal relations. To model dialogs we used business prose linguistic model for medical diagnostics and specialized linguistic processor [2].

We adopted the Random Forest algorithm [12] to work with databases containing temporal data [13]. The ensemble of decision trees is constructed in accordance with the basic

Random Forest algorithm. We use multidimensional feature space, one of which is the timestamp. However, the calculation method of the partitioning criterion value has changed to the arithmetic mean of entropy values. Also, the construction of the tree is carried out until all the elements of the subsample are proceeds without using cut – off procedure [4], [9].

To convert the ensemble of decision trees to knowledge base format we use some helper tools: The main object containing all features of the feature space as attributes; The counter to measure time; Vote counters. Each leaf is converted to a rule of the following form: f the duration of all intervals corresponding to vertices on the path to the root is greater than zero, then increment the vote counter for the class that corresponds to the current vertex.

When we extract knowledge containing temporal data from various sources of different types (experts, text, and databases) , we get multiple fragments containing objects, types, and rules. To merge all the fragments of knowledge together we use methods, algorithms, and software tools of the combined method of knowledge acquisition from distributed sources taking into account temporal data. Knowledge verification is not considered in this paper but implemented in AT-TECHNOLOGY - as well.

We studied the distributed knowledge acquisition process where temporal databases in medical domain were used as additional sources of knowledge. We noticed that knowledge base growth ratio lies between 12-25 % with 15 % in average. [4], [9]

IV. Some aspects of the interactions between the temporal solver, AT-SOLVER and the simulation modeling subsystem

The important feature of the temporal solver is the close interaction with the all-purpose AT-SOLVER and the subsystem of the simulation modeling of the external environment (external world), which is an obligatory component of any dynamic IES. [2] The temporal solver, as well as the subsystem of simulation modeling, acts on the times and process of the interaction between the temporal solver, while the subsystem of simulation modeling is carried out by data and command exchange in the asynchronous mode.

Figure 1 [8] shows the chart of the interactions between the temporal solver, all-purpose AT-SOLVER, and the simulation modeling subsystem. These interactions are provided by joint functioning support facilities. Moreover, the components interact with the total working memory. It should be noted that the interaction is carried out in two modes: the development of the applied dynamic IESs (including the adjustment of a series of IES prototypes) and functioning of the final prototype of the dynamic IESs. The first mode that is needed for dynamic IES construction is the first that was considered in the present paper.

The interaction between the components begins after the support facilities receive a message about the start of the adjustment of the developed prototype of the dynamic IES. At the initial stage, this is the configuration of components, including the setting of the duration of the operation cycle for the simulation modeling subsystem, indicating the database for the temporal solver and the AT-SOLVER.

The joint functioning support facilities provide the synchronization of component operation by sending messages with start or stop commands. The selected objects, whose collection of attributes describes the system state, are presented in the working memory. The knowledge base contains the temporal rules that are necessary to solve the formulated problem, as well as to describe the events and intervals. As a result of temporal inference on rules, the system state changes, i.e., the attributes of the working memory objects vary according to the solved problem. A synchronous interaction would mean that the simulation modeling subsystem goes to the standby mode up to the inference completion by the temporal solver following data transmission (from the subsystem to the solver). Otherwise, the asynchronous interaction is said to be the ability to continue the operation of the subsystem without waiting for the temporal inference.

Asynchronous interaction allows higher productivity due to the use of the time of the processing of general situations by the temporal solver to execute the next modeling cycle. Note that similar asynchronous interactions are applied in real practice, when it is impossible to react to an event immediately. This is the reason that the temporal solver and subsystem are synchronous [8].

The functioning of the simulation modeling subsystem and the inference facilities is an asynchronous process that is executed in parallel. The functioning of the temporal solver and AT-SOLVER is an synchronous process that is executed sequentially. The interaction between the components of the AT-TECHNOLOGY complex is a very difficult process that requires the development of models, methods, and software facilities to support interactions. The functions of the modules and blocks are the following: The configuration block carries out the component configuration. It sets the duration of the cycle of the discrete model time and the assignment of the names for the simulation modeling subsystem objects, temporal solver, and AT-SOLVER.

The model time generation block counts the cycles of the discrete model time according to the cycle duration specified by the configuration block. The working memory scanning block observes the changes in the working memory.

The control effect calculation block implements the target function of the interaction model. As a result of the block operation, the target component and the control effect, which should be set, are defined. The control effect generation block forms the control effect as the message to the certain component. The interface module of the message exchange with the components processes the input messages and sends control effects. For joint functioning support facilities, the special adjustment tools allow one to emulate the combined work of the components of AT-TECHNOLOGY in both the stepbystep and realtime modes. The use of these tools allows the study of the operation of the main components of the dynamic version of AT-TECHNOLOGY in the fullest manner.

Conclusion

These experimental investigations showed the advantages of the developed software tools compared to similar ones according to such criteria as the KRL power, operation speed, and reduction of the lead time of dynamic IESs. Verification of the performance and efficiency of these tools was done by developing a set of basic components, which is the minimum that is needed for dynamic IES operation.

Acknowledgements

This work was partially funded by the RFBR (Russian Foundation for Basic Research), project No. 18-01-00457.

References

[1] G. V. Rybina, *Theory and technology of construction of integrated expert systems. Monography.* Moscow: Nauchtehlitizdat, 2008.

[2] G. Rybina, *Intelligent systems from A to Z. A series of monographs in three volumes. Vol. 2. Intelligent interactive systems. Dynamic Intelligent Systems.* Moscow: Nauchtehlitizdat, 2015.

[3] G2 platform. [Online]. Available: http://www.gensym.com

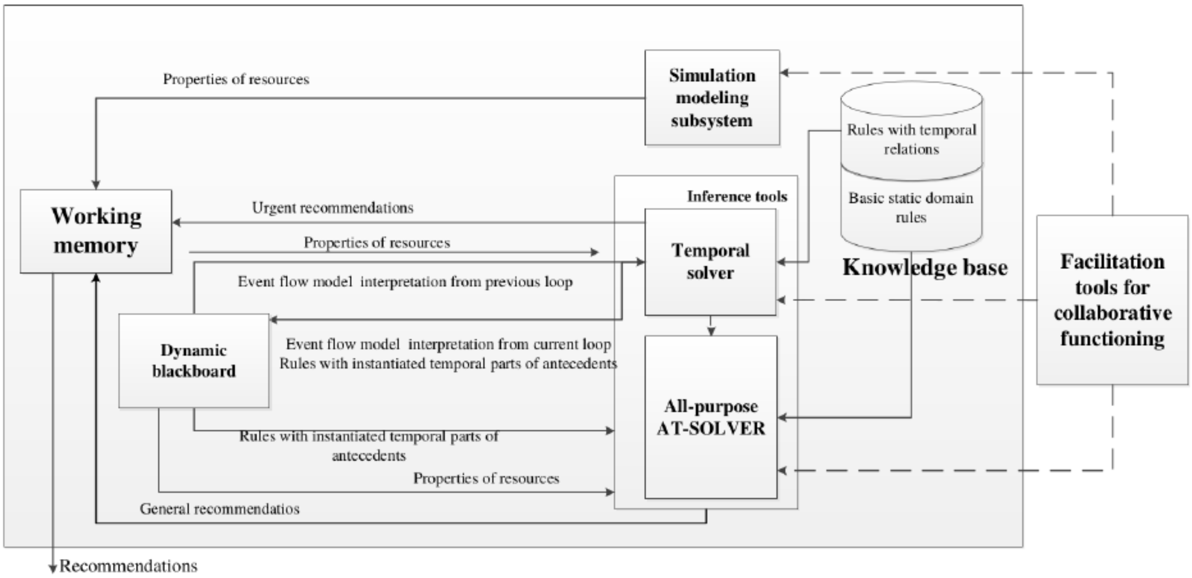

Figure 1. Interaction diagram of temporal solver, all-purpose solver, and simulation modeling subsystem

[4] G. Rybina and I. Danyakin, "Combined method of automated temporal information acquisition for development of knowledge bases of intelligent systems," *Proceedings of the 2017 2nd International Conference on Knowledge Engineering and Application. London IEEE*, pp. 117–123, 2017.

[5] J. Allen, "Maintaining knowledge about temporal intervals," *Communications of the Association for Computiong Machinery*, pp. 832–843, 1983.

[6] G. Osipov, "Dynamic intelligent systems," *Iskusstvennyj intellekt i prinyatie reshenij*, no. 1, pp. 47–54, 2008.

[7] V. Emelyanov and Y. S.I., "Introduction to intelligent simulation of complex discrete systems and processes. language rbd, anvik," 1998.

[8] G. Rybina and A. Mozgachev, "The use of temporal inferences in dynamic integrated expert systems." *Scientific and Technical Information Processing*, no. 6, pp. 390–399, 2014.

[9] G. Rybina and D. V. Demidov, "Automated acquisition, representation and processing of temporal knowledge in dynamic integrated expert systems. postproceedings of the 9th annual international conference on biologically inspired cognitive architectures, bica 2018 (ninth annual meeting of the bica society), vol.145, 2018. p. 448-452."

[10] I. Efimenko, "Semantic of time: models, methods, and identification algoritms for nlp-systems," *Vestnic Moskovskogo gosudarstvennogo oblastnogo universiteta Linguistics*, pp. 179–185, 2007.

[11] N. Arutyunova and Y. T. E., "Logical analysis of language: Language and time," *lndrik*, 1997.

[12] A. Tzacheva, A. Bagavathi, and P. Ganesan, "Mr - random forest algoritm for distributed action rules discovery," *International Journal of Data Mining & Knowledge Managment Process (IJDKP)*, no. 6(5), pp. 15–30, 2016.

[13] M. Kaufmann, A. Manjili, P. Vagenas, P. Fisher, D. Krossmann, F. Faerber, and N. May, "Timetable index: A unified data structure for processing queries on temporal data in sap hana," *Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data*, pp. 1173–1184.

ДИНАМИЧЕСКИЕ ИНТЕГРИРОВАННЫЕ ЭКСПЕРТНЫЕ СИСТЕМЫ : ОСОБЕННОСТИ АВТОМАТИЗИРОВАННОГО ПОСТРОЕНИЯ ТЕМПОРАЛЬНЫХ БАЗ ЗНАНИЙ НА ОСНОВЕ ЗАДАЧНО-ОРИЕНТИРОВАННОЙ МЕТОДОЛОГИИ

Рыбина Г.В., Сорокин И.А., Сорокин Д.О.

Развитие инженерии знаний привело к появлению новой профессии, в которой активно востребованы как профессиональные компетенции, так и индивидуальные качества личности. Проанализирован методический и технологический опыт автоматизированного построения компетентностно-ориентированных моделей специалистов в области инженерии знаний, в частности, специалистов по профессии "системный аналитик с использованием обучающих интегрированных экспертных систем.

Received 10.01.19