Analysis of the structure of the medical decision support rules system based on the dependency graph

Ruslan Vafin, Rashit Nasyrov, Rustem Zulkarneev

***Abstract*— Currently, one of the promising areas in medicine is personalized medicine, which allows selecting the optimal treatment for each individual patient. Within the framework of the national program "Digital Economy of the Russian Federation", artificial intelligence technologies are the most promising for solving the problems of personalized medicine. The most important areas of application of artificial intelligence is the analysis of a large volume of medical texts, for example, electronic medical records or various professional medical literature, as well as image recognition of various formats. Currently, a large amount of medical knowledge about various diseases has been accumulated, formalized in the form of clinical guidelines (CG). For general practitioners, the need to regularly master a large volume of textual medical information leads to information overload, and, accordingly, a decrease in the effectiveness of the processes of diagnosis, treatment and rehabilitation of patients. The main direction for solving this problem is the use of clinical decision support systems (CDSS), which are based on the decision-making rules formulated on the basis of the CG. This article discusses an analysis of a structure of relationship in decision rules system for CDSS, implemented in the logic programming language Prolog.**

***Keywords*— decision support system, decision rule, decision graph, graph structure, clinical guidelines.**

I. Introduction

Decision support is a necessary component of activity in fields with a high level of uncertainty, possible inconsistencies in data interpretation and a high chance of errors. This situation is due to the large intensive information exchange between the subjects of activity, which has both formal and informal nature. The most important representative of this type of activity is medical activity.

From a practical point of view, the problem of providing timely medical care is one of the urgent problems of public health. In many small towns and various villages there is a problem in the availability of qualified medical workers. The main problem is that doctors do not have enough experience in the treatment of various diseases, quite a few doctors act in a stereotyped manner, and apply the same treatment regimens for different patients, thereby reducing the effectiveness of treatment. In Russia, it is legally established that doctors should be guided by clinical guidelines when treating patients. Doctors can deviate from recommendations if they can prove the need to make changes in treatment, and this does not lead to a worsening of the patient's condition.

According to the information on the official website of the Ministry of Health of the Russian Federation, at the moment, 187 clinical guidelines have been approved by the Scientific and Practical Council of the Ministry of Health of the Russian Federation. There are also a fairly large number of clinical guidelines that have been compiled by other organizations, such as the Russian Respiratory Society, which issues guidelines for various respiratory diseases. These documents are quite large, which can reach hundreds of pages of text in A4 format. For general practitioners, such as therapists, it is difficult to remember several dozen documents containing prescriptions for each group of patients (men, women, pregnant women, smokers, etc.) and many additional conditions, i.e. there is information overload [1]. In connection with this overload, decisions made by doctors can become overly stereotyped and boil down to the fact that the same treatment rules are applied for all types of patients. This reduces the effectiveness of treatment, since the individual requirements of the patient are not taken into account. To solve this problem in medicine, CDSS are used, which are used to help doctors more accurately diagnose and prescribe treatment for patients [2].

In medicine, CDSS are used to solve a wide range of tasks: recognition and delivery of alarms and reminders, support in the diagnostic process, search for suitable cases (precedents), planning and monitoring of therapy, pattern recognition and interpretation. An important function of the CDSS is the dissemination of "best practices", incl. international. Most often, CDSS are used precisely to help with diagnosis, prescribing treatment and, if necessary, adjusting the prescribed treatment [3].

This paper discusses CDSS that are intended for use in a medical technological process. A medical technological process is a system of interrelated necessary and sufficient scientifically grounded therapeutic and diagnostic measures, the implementation of which allows the most rational way to carry out treatment and ensure the achievement of the maximum correspondence of scientifically predicted results

Manuscript received July 29, 2021. This work was partly supported by RFBR grant № 19-07-00780.

Ruslan Vafin Author is with the The Faculty of Informatics and Robotics, Ufa State Aviation Technical University, Ufa, Russia (corresponding author to provide e-mail: vafrus74@gmail.com).

Rashit Nasyrov Author is with the The Faculty of Informatics and Robotics, Ufa State Aviation Technical University, Ufa, Russia (e-mail: nrash@yandex.ru).

Rustem Zulkarneev Author is with the Medical faculty, Bashkortostan State Medical University, Ufa, Russia (e-mail: zrustem@ufanet.ru).

to real ones while minimizing costs [4]. In this case, the CDSS helps doctors follow the necessary measures, which are determined by the current standards in the field of health care, as which clinical guidelines are considered in this work. In clinical guidelines, descriptions and conditions for performing various activities may be misinterpreted or initially contain errors, which will lead to errors in the CDSS based on those documents.

II. State of the art

An analysis of publications in Web of Science, PubMed and RSCI revealed that the existing modern CDSS have problems with the effectiveness of decision-making and organizational problems, such as the lack of sufficient flexibility for doctors, who in some rare cases can make more effective decisions, compared to CDSS.

A systematic analysis [5] of randomized controlled trials evaluating the effectiveness of the use of CDSS integrated with electronic medical records. As a result of the analysis, it was revealed that the use of CDSS did not affect mortality, but moderately affected the outcome of the disease.

According to the results of the EHR analysis in [6], only 1 out of 20 quality indicators demonstrated significantly better efficiency of visits with CDS compared to EHR visits without HSSP.

A randomized controlled trial of CDSS in cardiovascular disease in 614 patients showed that the use of a computerized clinical decision support system did not provide any additional benefit over the indications in the chart (similar to the SCORE scale) and could make it difficult to communicate evidence to individual patients [7].

In an experiment based on 20 clinics and 31 thousand patients, Teen BPCDSS had a significant positive effect on the recognition of adolescent hypertension along with a moderate increase in the effectiveness of treatment, compared with the use of traditional clinical guidelines [8].

A system using a combination of fuzzy logic and if-then rules to determine the severity of a patient showed a severity score of 99.44% compared to a manual score of 86.04% [9].

In a review of the literature on the use of CDSS used in the field of otolaryngology, it was shown [10] that the accuracy of various systems for diagnosis and differential diagnosis is quite high and is about 90%. Nevertheless, most of the above systems are designed for a narrow range of tasks, such as a system for determining the severity of acute tonsillitis in children with predictive significance of symptoms. In this review, there are no holistic systems that would solve the entire spectrum of problems associated with at least one disease, namely, diagnosis, treatment, and rehabilitation.

Based on the results of the analysis of publications on the effectiveness of the CDSS, it is difficult to make an unambiguous conclusion about a significant improvement in the quality of diagnosis and treatment. Most of the studies cited show an improvement in the quality of the diagnosis and treatment process using CDSS relative to traditional decision making.

An analysis of works on this topic also revealed various specific CDSSs that support the entire treatment and diagnostic process, or part of it. The feature lies either in the use of remote analysis of patient data, or in a specific interaction between the system and the CDSS, which is different from simply entering patient data and obtaining the result. Here are some examples of works that describe the data of the CDSS.

In [11], an adaptive CDSSis presented, which provide the necessary information about a patient, but without limiting doctors in making decisions. The authors believe that conventional CDSS restrict doctors by forcing them to follow recommendations directly, even if it is necessary to deviate from the recommendations, in the event that the HSSP gives suboptimal recommendations.

In [12], CDSS is described for solving problems in the field of telemedicine, in particular, analyzing data received from patients to search for trends in quantitative indicators and exceeding acceptable values in order to effectively manage the treatment and diagnostic process of various patients and redistribute resources for the most needy patients. The system is implemented on the basis of a client-server architecture, where clients are patients with diagnostic equipment, and a server on the side of the clinic makes decisions. At the time of publication of the work, the system is at an early stage of development, and the current implementation.

The aim of work [13] is to develop a rule-based expert system for a mobile phone-based remote heart failure monitoring system. The set of rules was developed by interviewing experts. It is assumed that the use of the system will improve the quality of life of patients with cardiovascular diseases.

In addition to the diversity of the architectures of the CDSS itself, we analyzed the works devoted to the architecture of the knowledge base, the core of CDSS, and various features related to rules, such as rules based on fuzzy logic.

A description of the architecture of the CDSS and the rules used are given in [14]. Physical questionnaires are scanned or manually filled by operators into a system that, based on clinical algorithms, makes a preliminary diagnosis of the patient. The Knowledge Base is a digitalized clinical guidebook in Arden Syntax format and provides an HL7-compatible interface.

The developed system [15] is intended for identification of drugs, their dosages, methods of administration, frequency, duration and reasons for taking in the discharge. This system is not a full-fledged CDSS, but it can be an important part of larger systems, since it allows to extract the necessary information with a sufficiently high accuracy and provide decision support based on clinical extracts. At the time of this writing, the system was presented as a prototype, and it is planned to improve the volume and accuracy of the information extracted. The results of this work can also be used to extract information from arbitrary medical texts.

The DSS architecture using rules with temporal conditions, such as the conditions of the sequence of application of rules or the application of rules after a certain time, which are based on CG, is given in [16]. This paper presents the architecture of the system in the form of a set of software modules that are easily modified, which allows flexible configuration of the system for various CG. In the appendix, various decision rules have been classified, for which rule templates are allocated, which allows to add new rules by editing the templates. In [17], a more extensive classification of decision-making rules was carried out, which will allow standardizing the representation of knowledge in CDSS, for faster and more accurate development of knowledge bases for various CDSS.

In [18], a system is presented that simultaneously uses formal rules based on CR and, if there are no suitable rules, then uses models trained on the case histories of various patients. The article presents UML models of such a system without implementation.

The work [19] describes the phased application of decision rules in the form of trees with the addition of patient information for the treatment of non-metastatic breast cancer based on CG and patient records. A knowledge base development environment was developed in the form of decision trees and CDSS for working with this knowledge base.

This article [20] discusses a rule-based approach to rule base formation using the evidence-based argumentation approach (RIMER). At RIMER, the rule base is designed with confidence levels embedded in all possible consequences of the rule. This rule base has been tested on a small sample of patient test data and needs further validation.

The article provides a method for classifying textual information based on symbolic rules [21]. This paper provides a classification of the stages of cancer of tumor nodes and metastases of the lung from the pathology reports. It uses the well-known open source natural language processing system General Architecture for Text Engineering (GATE) using the same name natural language processing architecture. The disadvantage of the GATE system is the lack of support for the Russian language out of the box.

The paper [22] describes the process of systematization and explicit transformation of documentary knowledge into integrated into the CDSS workflow. The work shows the process in the form of UML activity diagrams. This paper focuses on the process of manually annotating text and extracting guidelines from the text into a knowledge representation format in the form of a Guideline Elements Model (GEM). The model of clinical guidelines obtained using this approach at the time of the release of the work had no implementation in the form of software codes for DSS.

In paper [23], a system was presented that generates a set of sequential rules based on training on the correlation between nursing diagnoses, outcomes and medical interventions, which is based on a prefix tree.

Based on the results of a review of current scientific works in the field of CDSS, it was concluded that most of the existing systems work with a rather narrow area of knowledge, for example, accounting for drug interactions or detecting hypertension. At the moment, there are few systems that would support the entire medical technological process for specific diseases, and existing systems are mostly closed systems. The main part of the systems is intended to support the diagnosis, treatment or rehabilitation of various aspects of diseases, but no relevant systems have been found for comprehensive formal support of the medical technological process based on clinical guidelines.

Thus, the analysis of publications shows that the use of clinical guidelines as a basis for the formation of rules for making diagnostic and treatment decisions as part of CDSS is an urgent task. There are currently no satisfactory solutions in this area to support a wide range of medical technology processes for various types of diseases. Based on the analysis, a hypothesis is formed that it is possible to create a universal structure of decision-making rules. Using this structure, it is possible to develop a subsystem for the automatic inference of diagnoses and prescriptions for any type of disease. The leading contradiction is that the universal structure of the rules on the peculiarities of diagnostics and treatment of various types of diseases does not always correspond to the specific requirements of the subject area.

The main goal of this article is to highlight and describe dependencies in the rule system. To achieve this goal, it is necessary to develop a system of rules, develop a methodology for analyzing dependencies in rules, develop a prototype of the CDSS and a language for describing the knowledge base to test the performance of the proposed system of rules. To achieve this goal, it is necessary to develop a general approach to converting the text of clinical guidelines into formal constructions based on the “if-then” scheme and to analyze their mutual dependencies. Then need to justify a way to transform these constructs into predicate constructs. The final task is to test the rules using the tools of the logic programming language Prolog.

III. Description of the Object of Analysis

Clinical guidelines - documents based on proven clinical experience, describing the actions of a doctor in the diagnosis, treatment, rehabilitation and prevention of diseases, helping him to make adequate and accurate clinical decisions [24]. The document defines the types, volume and indicators of the quality of medical care for citizens with a specific disease, syndrome or clinical situation.

The structure of clinical guidelines is specified in GOST R 56034-2014. The document is a semi-structured text in which information is divided into sections and subsections that are specified in GOST. The text in the sections is unstructured text, as the standard only prescribes the structure and content of clinical guidelines, but does not contain precise wording for the entire structure of the document.

Some of the information in a document contains a rigidly defined structure. For example, information about the level of evidence and the level of persuasiveness of recommendations is presented as follows: “Strength of recommendation D (level of evidence IV)”. Evidence confidence level is the degree of confidence that a found benefit from a medical intervention is true. Strength of a recommendation is the degree of confidence that the effect of an intervention is credible and that following the recommendation will do more good than harm in a particular situation.

IV. Description of the rule system

Conversion of clinical guidelines into the knowledge base for CDSS is carried out in several stages. At the first stage, decision rules are automatically extracted using text mining, or, less efficiently, manually. Further, the resulting set of rules in a human-readable format, for example, in an algorithmic language, is verified by experts in the field of knowledge, and, if necessary, corrected, which makes it possible to correct errors when manually or automatically extracting the rules of their source text. The resulting knowledge base in an intermediate format is translated into code in a language supported by a specific CDSS.

In general, the structure of rule relationships can be represented as a graph, where nodes represent rules and arrows represent dependencies in conditions on other rules. The data graph can be both connected and disconnected. In the case of a connected graph, the result of applying a set of rules can be one node or more than one if no rule was triggered on the input dataset. As a result of the conclusion based on the graph $G_1 = (X, A)$, the solution $S_1 = (P, x_s)$ will be obtained, where X is the set of nodes (rules), A is the set of dependency arrows, P is the set of output data, x_s is the node of the graph, which is the final solution for the input.

In the case of a disconnected graph, the calculation of the result for each connected component is carried out in parallel. This can be applied in rules like drug interactions where incompatible drugs are calculated based on the input set of prescribed drugs. As a result of the conclusion based on the graph $G_2 = (X, A)$, the solution $S_2 = (\{(P_i, x_i)\}, P)$ will be obtained, where X is the set of vertices (rules), A is the set of dependency arrows, i is the number of the solution, P_i is the set of output data of solution i, P is the combined set of output data of all decisions P_i, x_i is the vertex of the graph, which is solution i. Also, several solutions may arise in cases where there are several drains in the connectivity component, for example, during diagnostics, it was revealed that the patient has several pathologies at the same time. In the general case, for a comprehensive study of a patient, it is necessary to pass the input data through several disconnected graphs, each of which produces a certain result, adds new data to the patient's input data, and passes the next rule graph.

In a clinical practice guideline, chapters and sections group several semantically related rules into separate meaningful blocks, such as a patient diagnosis block. An example of such a rules system based on two-level dependency graph is shown in Figure 1. This figure shows groups of rules that are consistently applied to patient data. For each group of rules, a conclusion is displayed, for example, the degree of severity, the necessary treatment, etc.

In the second block of rules as an example each rule (condition) depends on the result of the previous one, such a chain of rules can be the determination of the severity. If the patient's oxygen saturation is below 80%, then the severity is "severe" and there is no point in checking other conditions. The dash-and-dot arrows on the two-level structure of rules show connections between rules located in different blocks.

Each rule can contain several conditions, separated by the logical operation "or". If one part of the condition is dependent on another rule, and the second part is independent, then the entire rule containing such a condition becomes partially dependent. In Figures 1 and 2, such rules are indicated by the dashed arrows included in them.

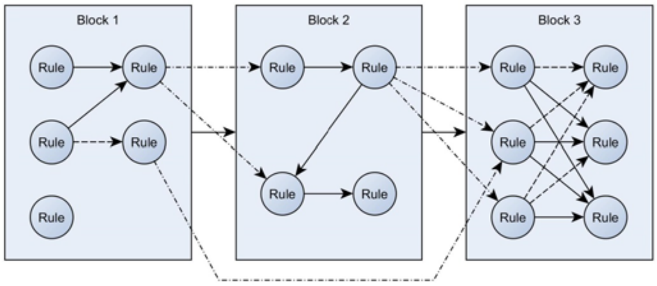

Figure 1. An example of rules system based on two-level dependency graph.

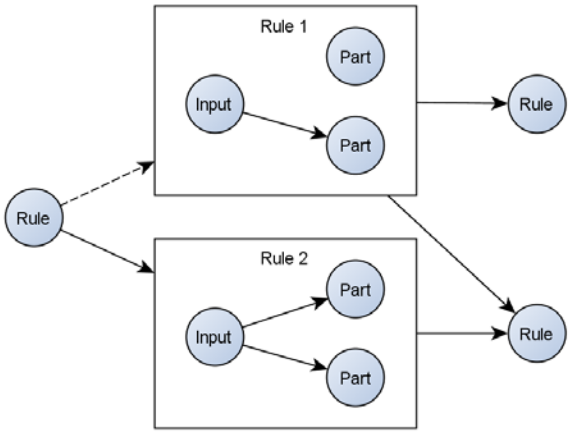

Figure 2. An example of rules with logical condition "OR"

This figure shows the internal structure of the rules, where the "input" node denotes the values on which the parts of the boolean condition "part" depend. In the first rule "rule 1" the condition is: "part1 OR part2 (input)", in the second: "part1 (input) OR part2 (input)". The first rule can be triggered both with a triggered input value, and without it; in the second rule, all parts of the logical condition depend on the input value.

Combining rules into blocks occurs at the stage of automatic and manual conversion of rules. Links in the rule graph are built automatically. This is implemented by building an adjacency matrix of the rule graph for each block based on the results of searching for a variable in a condition in the results of all other rules. This graph can be used to search for inconsistencies in rules, cycles in the structure of relationships between rules, or to optimize queries to the CDSS, taking into account active and inactive arrows, which show the fulfillment of various conditions. In general, the block of rules will have the form shown in Figure 3. The "Prev Rule" and "Next Rule" nodes indicate the rules in the previous and subsequent blocks. During the work of the decision-making process, it is necessary to take into account the arrows between the rules in different blocks.

To execute a rule block, it is necessary to find all the sources of the directed rule graph. To search for sources in a block in each block rule, all the parameters included in the conditions are sequentially checked to find those parameters that are the result of other conditions. Figure 4 shows the process of checking the rules within one block without considering the links between other blocks.

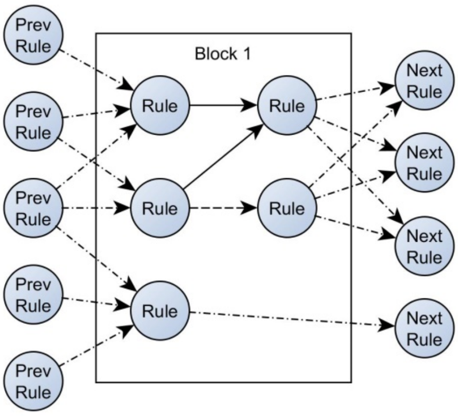

Figure 3. Generalized view of the block of rules

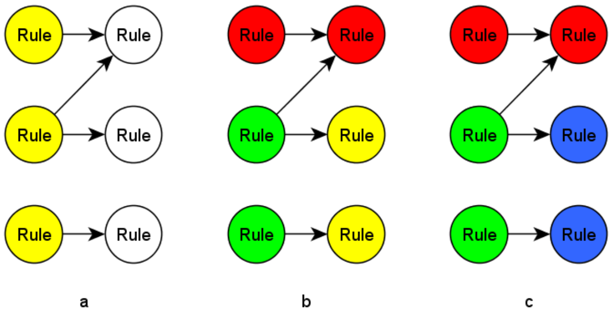

Figure 4. The process of checking rules within one block

The first stage is finding the sources in the graph. After finding the sources, Figure 4a, each rule is triggered on the patient's input, and when the rules are executed, subsequent rules in the graph are executed. Figure 4b shows an example of how 2 rules worked, but the first one did not work. Next, the remaining 2 rules in the graph are checked. Figure 4c shows that both rules worked, and they are marked in blue, since they are the result of executing a block of rules, i.e. sinks of two connected components of the graph.

The rules themselves should be able to change the array of facts about the patient. The rule predicate has the following form: rule (id, conclusion, input_data, output_data), where id is the identifier of the rule, conclusion is the conclusion (in the rule “if the condition then action” this is the action), input_data is the input array of facts, output_data is the output array of facts. As a result of the execution of this predicate, new facts can be added to the output array of facts, facts are deleted, or it can remain unchanged.

V. Comparison with Analogues

The results of the literature review showed that various algorithms and decision-making methods are used, such as: decision trees, random forest, fuzzy logic, clinical algorithms. These methods work very well with one task, when it is necessary to analyze a given patient and draw some kind of conclusion, for example, determining the severity of a patient in nursing algorithms. The application of such methods for the entire medical technological process becomes a complex and non-trivial task for both the developer and the user.

The advantage of the proposed solution is simplicity of implementation and very high visibility, which allows not only formalizing clinical guidelines and treatment algorithms, but also other, more specific medical sources, for example, different algorithms for determining mental state, etc. The advantage of this approach over decision trees is the ability to create an arbitrary structure of an acyclic graph, which will allow to reduce branching, and, accordingly, reduce the number of decision rules.

In this approach, the very representation of decision-making rules is a consequence of organizing and grouping the rules into meaningful blocks. Links in the graph are generated automatically, and cycles can occur in the graph. These loops are easy to find, but removing them requires the involvement of experts who could modify the rules so as not to alter the accuracy of the rule base.

The main disadvantage of this approach is the rigidly defined structure of the system, which uses only formalized rules of the "if-then" type. In real cases of diagnosis and treatment, there are often cases when it is impossible to formalize patient data, highlighting some values that would be used in logical conditions. This illustrates the contradiction indicated at the beginning of the work. The most common example is the analysis of electrocardiograms. For ECG analysis it is necessary to use various signal processing methods, for example, Wavelet transform.

A possible solution to the problem of the impossibility of using other methods may be to represent some blocks of rules in the form of a "black box" that uses various external programs to analyze complex data. An example of such a structure is shown in Figure 5.

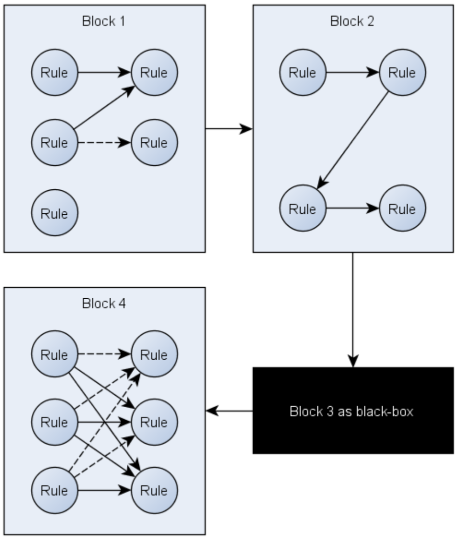

Figure 5. An example of using a black box

This block accepts the patient's data as input and gives a new result, supplementing the patient's data, for example, the ECG analyzer can add information about a possible arrhythmia or a possible latent myocardial infarction. By supplementing the structure of rules with external modules, it is possible to significantly expand the capabilities of the CDSS, by using computational capabilities, the implementation of which by means of CDSS is inefficient, which will significantly increase the quality of the medical process.

VI. An example of the application of the proposed approach

In this work, a subset of the algorithmic language was chosen as an intermediate format, namely, the “if-then” block to describe the decision-making rules. The language Prolog and a set of corresponding predicates compatible with the core of the CDSS were chosen as the language of the knowledge base of the CDSS.

As an example of a step-by-step conversion of the rule system, using one rule as an example, was selected a fragment of the text from the clinical guidelines on influenza in adults, shown in Figure 6.

- Рекомендуется госпитализировать пациентов [52, 63, 64, 95, 100, 102]:
  - с тяжелым и средне - тяжелым течением гриппа;
  - проживающих в общежитиях и в условиях неблагоприятных факторов жилой среды;
  - из группы риска развития неблагоприятного течения болезни: беременные на любом сроке беременности и в послеродовый период; лица с хроническими заболеваниями легких (бронхиальная астма, ХОБЛ и др.);лица с хроническими заболеваниями сердечно-сосудистой системы (пороки сердца, ГБ, ИБС с признаками сердечной недостаточности и др.); лица с нарушениями обмена веществ (сахарный диабет, ожирение 2 – 3 степени и др.); лица с хронической болезнью почек; лица с хроническими заболеваниями печени; лица, с определенными неврологическими состояниями (включая нейромускульные, нейрокогнитивные нарушения, эпилепсию); лица с гемоглобинопатиями; лица с первичными и вторичными иммунодефицитами (ВИЧ-инфекция, прием иммуносупрессорных медикаментов и т.п.); лица с злокачественными новообразованиями; лица в возрасте 65 лет и старше.

Figure 6. Part of the text from the clinical guidelines on influenza in adults in the chapter on treatment

A simplified version of the algorithmic language is used as an intermediate format, namely the “if-then” construction. As a condition, a logical structure is used, the variables of which are various symptoms of the patient, as well as information added in the process of working with the patient, for example, the severity of the disease, calculated at one of the stages. All symptoms are realized in the form of the construction: "symptom ("key)". Symptoms can be boolean as well as numerical or string. In the case of string or numeric values, comparison operators can be used in the condition, for example "symptom("saturation")<95".

This example shows a syntactic construction of the form: "Recommended "action" : "symptom" or "symptom" or ("symptom"or "symptom"... or "symptom")". Different risk group variations are indicated in brackets. Figure 7 shows a listing of an edited version of this rule, with corrected errors of automatic analysis. One unnecessary condition was removed, the grouping of conditions with the "AND" operation was removed and replaced by two conditions with the "OR" operation.

```
ЕСЛИ
    симптом    ("средне-тяжелое течение гриппа")
    ИЛИ симптом("тяжелое течение гриппа")
    ИЛИ пациент("проживание в общежитиях")
    ИЛИ пациент("проживание в условиях неблагоприятных факторов жилой среды")
    ИЛИ пациент("беременная")
    ИЛИ пациент("послеродовый период")
    ИЛИ пациент("хронические заболеваниями легких")
    ИЛИ пациент("хронические заболеваниями сердечно-сосудистой системы")
    ИЛИ пациент("нарушения обмена веществ")
    ИЛИ пациент("хроническая болезнь почек")
    ИЛИ пациент("хронические заболевания печени")
    ИЛИ пациент("неврологические состояния")
    ИЛИ пациент("гемоглобинопатия")
    ИЛИ пациент("первичными и вторичными иммунодефицитами ")
    ИЛИ пациент("злокачественные новообразования")
    ИЛИ пациент("возраст") > 65
ТО
    рекомендуется("госпитализировать пациента")
```

Figure 7. Listing of the revised adult flu rule

In this example, in the conditions there is a condition "moderate to severe flu". This fact is calculated by another rule, and its result is used in other, depending on it, rules. 18.1 An example of the converted rule into the Prolog language in Figure 8.

```
symptom(1, "средне-тяжелое течение гриппа").
symptom(2, "тяжелое течение гриппа").
symptom(3, "проживание в общежитиях").
symptom(4, "проживание в условиях неблагоприятных факторов жилой среды").
symptom(5, "беременная").
symptom(6, "послеродовый период").
symptom(7, "хронические заболеваниями легких").
symptom(8, "хронические заболеваниями сердечно-сосудистой системы").
symptom(9, "нарушения обмена веществ").
symptom(10, "хроническая болезнь почек").
symptom(11, "хронические заболевания печени").
symptom(12, "неврологические состояния").
symptom(13, "гемоглобинопатия").
symptom(14, "первичными и вторичными иммунодефицитами ").
symptom(15, "злокачественные новообразования").
symptom(16, "возраст").

reccomendation(1, "Рекомендуется: госпитализировать пациента").

rule(1, "Рекомендуется госпитализировать пациентов", FIn, FOut):-
    /* member(__, FIn),*/
    rule(1, FIn),
    append(FIn, ["Действие = госпитализировать пациентов"], FOut).
rule(1, _):-cs(S,1);cs(S,2);cs(S,3);cs(S,4);cs(S,5);cs(S,6);cs(S,7);
            cs(S,8);cs(S,9);cs(S,10);cs(S,11);cs(S,12);cs(S,13);cs(S,14);
            cs(S,15);cs(S,16).
```

Figure 8. An example of a rule in Prolog language

In this rule, symptom is a fact that is intended to store in the knowledge base various patient symptoms and their identifiers, in this case, the first ID. The assignments to be displayed are recorded using the recommendation predicate. The rule itself consists of two parts, the first is the processing of the patient's input data, changing them if necessary, and checking this rule in the knowledge base. This division is made to simplify testing and reduce the code base, since when combining predicate data, it would be necessary to generate check code many times.

At the beginning of the CDSS operation, the array of facts is initialized by adding facts from the patient's medical history. After that, the rules are sequentially run in each block, and the result of the check is displayed for each block. Due to the specifics of the implementation, the block can produce a negative result, i.e. not a single rule worked. In this case, the execution of the following blocks is not interrupted, but continues with unchanged patient data.

To test the operability of this approach to the structure of the rules, clinical guidelines for flu in adults from the Ministry of Health of the Russian Federation were formalized. The clinical guidelines were divided into 7 blocks: diagnosis, severity, clinical options, risk of disease, indications for laboratory diagnosis, differential diagnosis and treatment. Figure 9 shows an example of running the program in console mode for patient test data.

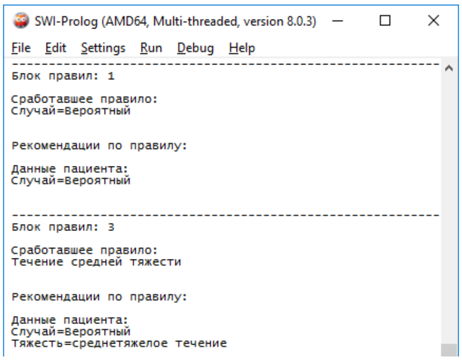

Figure 9. An example of the program operation based on the test rule base

The structure of rules in the form of a two-level dependency graph was developed, examples of various structures of the graph of rules were given. Using the rule of hospitalization of a patient from the clinical guidelines as an example, the process of step-by-step transformation of the rule into the format used in the proposed CDSS was shown, and the links in this rule with other rules were shown.

VII. Conclusion

The use of CDSS within the entire medical process is an urgent task of modern medicine. As a result of the analysis of works devoted to the use of CDSS in medicine, it was revealed that the systems that supported the entire medical workflow are quite few, and the existing systems in most do not have full coverage of decision support for the entire medical workflow, and therefore, it was concluded that the development such a system.

For the CDSS, which would support the entire medical workflow, a structure of rules was proposed for a comprehensive coverage of all stages, which is quite flexible and visual. The creation of rules using such a structure is proposed based on clinical guidelines for various diseases, which fully describe the entire medical workflow for specific diseases. An example was given of the step-by-step extraction of rules and text of clinical guidelines for flu in adults.

As it was mentioned in the previous chapter, the problem of pure application of the proposed approach is a rather large class of problems that cannot be formalized explicitly in the form of “if-then” rules. It is proposed to implement such tasks in the structure in the form of black boxes, which would process data in different ways, but return the result in a standard form.

Depending on the structure of the link of the block of rules, it is necessary to carefully plan the sequence of input conditions, for example, optimization of the "AND" or "OR" conditions in the rule based on statistical analysis of the most common symptoms in patients. These examples reflect the performance optimization of the CDSS kernel that uses the proposed rule structure.

The hypothesis about the creation of a universal system of rules for formalizing the rules in clinical guidelines was tested. To test the hypothesis, a prototype of an expert system in the Prolog language was created, clinical guidelines on flu in adults were formalized and divided into 7 blocks, which contain 33 different decision rules, which were tested on real cases and demonstrated satisfactory performance. Thus, the resulting structure of the rules can be used to formalize clinical guidelines for use in CDSS, but this structure requires improvements that will be carried out in future works on this topic.

Acknowledgment

This work was partly supported by RFBR grant № 19-07-00780.

References

[1] A. F.Sokol, R. V. Shurupova, “Information fatigue Syndrome: ways and methods of its prevention” in Psychology of development and education (Russian journal of applied research), no. 1, pp. 116-118, 2017. (in Russian)

[2] O. U. Atkov, “Medical Decision Support System”, in Vrach i informatsionnye tekhnologii, no. 6, pp. 67–75, 2013. (in Russian)

[3] I. V. Efimenko, V. F. Khoroshevsky, “Intelligent decision support systems in medicine: a retrospective review of the state of research and development and prospects.”, 2017. (in Russian)

[4] G. I.Nazarenko, G. S. Osipov, “Basics of the theory of medical technological processes. Part 1”, Moscow: Fizmatlit. p. 144, Dec. 2005. (in Russian)

[5] L. Moja, K.H. Kwag, T. Lytras, L. Bertizzolo, L. Brandt, V. Pecoraro, G. Rigon, A. Vaona, F. Ruggiero, M. Mangia, A. Iorio, "Effectiveness of computerized decision support systems linked to electronic health records: a systematic review and meta-analysis." in American journal of public health, vol.104, no. 12, pp. 12-22, Dec. 2014.

[6] M. J. Romano, R. S. Stafford, Electronic health records and clinical decision support systems: impact on national ambulatory care quality Archives of internal medicine, vol. 171, no. 10, pp. 897-903, May. 2011.

[7] A. A. Montgomery, T. Fahey, T. J. Peters, C. MacIntosh, D. J. Sharp “Evaluation of computer based clinical decision support system and risk chart for management of hypertension in primary care: randomised controlled trial” in Bmj, vol. 320, no. 7236, pp. 686-690, Mar. 2000.

[8] E. O. Kharbanda, S. E. Asche, A. R. Sinaiko, H. L. Ekstrom, J. D. Nordin, N. E. Sherwood, P. L. Fontaine, S. P. Dehmer, D. Appana, P. O’Connor, “Clinical decision support for recognition and management of hypertension: a randomized trial” in Pediatrics, vol. 141, no. 2, Feb. 2018.

[9] M. D. Soufi, T. Samad-Soltani, S. S. Vahdati, P. Rezaei-Hachesu, ”Decision support system for triage management: a hybrid approach using rule-based reasoning and fuzzy logic” in International journal of medical informatics, vol. 114, pp. 35-44, Jun. 2018.

[10] A. P. Yastremskij, A. I. Izvin, A. G. Sannikov, “Theoretical foundations of clinical diagnosis of various diseases of lororgan using information systems” in Rossijskaya otorinolaringologiya, vol. 1, no. 4, p. 74, 2016

[11] A. Richard, B. Mayag, F. Talbot, A. Tsouki`as, Y. Meinard “What does it mean to provide decision support to a responsible and competent expert?” in EURO journal on decision processes, pp. 1-32, Nov. 2020.

[12] J. Basilakis, N. H. Lovell, S. J. Redmond, B. G. Celler, “Design of a decision-support architecture for management of remotely monitored patients” in IEEE Transactions on Information Technology in Biomedicine, vol. 14, no. 5, pp. 1216-1226, Jul. 2010.

[13] E. Seto, K. J. Leonard, J. A. Cafazzo, J. Barnsley, C. Masino, H. J. “Ross Developing healthcare rulebased expert systems: case study of a heart failure telemonitoring system” in International journal of medical informatics, vol. 81, no. 8, pp. 556-565, Aug. 2012.

[14] V. Anand, P. G. Biondich, G. C. Liu, M. B. Rosenman, S. M. Downs, “Child health improvement through computer automation: the CHICA system” in Medinfo, pp. 187-191, Jan. 2004.

[15] O. Uzuner, I. Solti, E. Cadag, “Extracting medication information from clinical text” in Journal of the American Medical Informatics Association, vol. 17, no. 5, pp. 514-518, Sep. 2010.

[16] A. De Clercq, J. A. Blom, A. Hasman, H. H. M. Korsten, “GASTON: an architecture for the acquisition and execution of clinical guideline-application tasks” in Medical informatics and the Internet in medicine, vol. 25, no. 4, pp. 247-263, Jan. 2000.

[17] A. Wright, H. Goldberg, T. Hongsermeier, B. Middleton, “A description and functional taxonomy of rule-based decision support

content at a large integrated delivery network" in Journal of the American Medical Informatics Association, vol. 14, no.4, pp. 489-496, Jul. 2007.
[18] D. Rossille, J. F. Laurent, A. Burgun, "Modelling a decision-support system for oncology using rule-based and case-based reasoning methodologies" in International journal of medical informatics, vol. 74, no. 2-4, pp. 299-306, Mar. 2005.
[19] M. P. Hendriks, X. A. Verbeek, "Transformation of the National Breast Cancer guideline into datadriven clinical decision trees" in JCO clinical cancer informatics, vol. 3, pp. 1-14, May. 2019.
[20] G. Kong, D. L. Xu, X. Liu, J. B. Yang, "Applying a belief rule-base inference methodology to a guidelinebased clinical decision support system" in Expert Systems, vol. 26, no. 5, pp. 391-408, Nov. 2009.
[21] A. N. Nguyen, M. J. Lawley, D. P. Hansen, R. V. Bowman, B. E. Clarke, E. E. Duhig, S. Colquist, "Symbolic rule-based classification of lung cancer stages from free-text pathology reports" in Journal of the American Medical Informatics Association, vol. 17, no. 4, pp. 440-445, Jul. 2010.
[22] R. N. Shiffman, G. Michel, A. Essaihi, E. Thornquist, "Bridging the guideline implementation gap: a systematic, document-centered approach to guideline implementation" in Journal of the American Medical Informatics Association, vol. 11, no. 5, pp. 418-426, Sep. 2004.
[23] L. Duan, W. N. Street, E. Xu, "Healthcare information systems: data mining methods in the creation of a clinical recommender system" in Enterprise Information Systems, vol. 5, no. 2, pp. 169-181, May. 2011
[24] "GOST R 56034-2014 Clinical guidelines (treatment protocols). General provisions", Moscow: STANDARTINFORM, 2015. (in Russian)

Vafin Ruslan Rustamovich — postgraduate student of the Technical Cybernetics chair of USATU. Research interests: expert systems, natural language processing, decision making theory.

Nasyrov Rashit Vilyevich — PhD of Technical Sciences, PhD of Pharmaceutical Sciences, Associate Professor of the Technical Cybernetics chair of USATU, Scientific supervisor of the Laboratory for modeling biological and biotechnical systems. Research interests: decision-making under uncertainty, representation and processing of data and knowledge, simulation modeling, functional biomechanics, application of methods of mathematics and computer science in biomedical research.

Zulkarneev Rustem Khalitovich — Dr. Sci. of Medicine, professor of the Propaedeutics of Internal Diseases chair of the Bashkir State Medical University. Research interests: medical informatics, decision support systems, linguistic analysis of medical documents.