The use of a production approach in the development of decision support systems for medical purposes

D V Gruzenkin¹, O S Novikov¹, A V Sukhanova¹, V V Kukartsev¹,², V S Tynchenko¹,² and A S Mikhalev¹

¹Siberian Federal University, 79 Svobodny pr., Krasnoyarsk, 660041, Russian Federation
²Reshetnev Siberian State University of Science and Technology, 31 Krasnoyarsky Rabochy Av., Krasnoyarsk, 660037, Russian Federation

E-mail: headeynagai@gmail.com

Abstract. This article discusses the application of the production approach to the implementation of decision support systems, the input for which is a number of signs, and the output is a number of response actions. The general solution using the Cartesian product of sets is described. In addition, block diagrams and diagrams are given for visualization of theoretical material, and an example of an existing medical system implemented on the basis of the described solution is also considered. The relevance of traditional approaches to the development of a decision support system is proved along with the growing popularity of artificial neural networks.

1. Introduction

Today, the popularity and relevance of artificial intelligence (AI) using in all kinds of areas from the economy to medicine, is no longer in doubt. However, it is most commonly accepted to associate AI with artificial neural networks than with some other varieties of it [1]. Undoubtedly, with the use of neural networks, a wide range of non-formalizable and hard-formalizable problems can be solved. The so-called decision support systems (DSS) do not lose their relevance - on the contrary, they have been widely used relatively recently, despite the fact that the concept of DSS was formulated in the 70s of the last centuries [2].

In a general sense, a decision support system consists of a set of software components and is aimed at highly efficient resolution of a certain range of problem situations. Moreover, such a system is aimed at reducing the information load on a person, frees him from routine work and thorough study of a large amount of data, while at the same time minimizing the degree of potential errors in making any decisions, especially in critical situations requiring a quick response. In other words, the existence of such systems can greatly simplify and automate familiar processes [3].

For the implementation of such systems, many approaches have been developed, distinguished by what model of knowledge representation underlies. In the general case, these models are divided into logical and heuristic.

In logical models, as a rule, first-order predicate calculus is used, supplemented by several heuristic strategies. The use of predicate calculus was especially intensified after the creation of powerful derivation search procedures, namely, the resolution method and the inverse method [4]. In systems

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Published under licence by IOP Publishing Ltd
implemented by these methods, for a given system of premises, there is a fixed system of inference rules.

In heuristic models, more complex rule constructions are used, allowing taking into account heuristic information about the specifics of the problem environment, often expressed in the form of semantic structures [5]. Such models are commonly referred to as a network, a frame, a production, and an object-oriented [6].

The presented traditional models of knowledge representation for decision support systems have a number of advantages over systems that use a trained artificial neural network (ANN) instead of a knowledge base. Often, for training ANNs, a sufficient amount of “good” input data is required, whereby sufficient we mean a number greater than one hundred (this is the minimum), and “good” are data that are as similar as possible to the system with which to work during its productive use [7 - 9]. Examples from the training set should fully cover all possible input data variants and the corresponding system responses [10]. Selection of input data for training a neural network is almost the most difficult stage of its implementation. Moreover, the development of a decision-making system based on ANN is more resource-intensive in comparison with traditional approaches.

Based on the foregoing, we can conclude that the creation of DSS using traditional models of knowledge representation is still relevant. At the moment, based on the solution described in the article, a DSS for medical purposes has already been developed, intended for making a diagnosis based on the results of the examination and determining the treatment plan. However, the approach presented in this article is universal and can find its application in various subject areas where it is necessary to diagnose something by a set of signs and determine a set of further actions. For example, such a DSS can be used to diagnose malfunctions in various technical and biological systems.

2. Decision principle

A productive approach to the development of decision support systems is based on the rules, which can be represented in the form “If (condition), then (action)”, where the condition means a sample to search the knowledge base, and the action means actions or operators executed in case the fulfillment of condition [6]. These actions may be intermediate and act as conditions for the execution of subsequent actions, or they may terminate the system. This model of knowledge representation is clear, modular, easy to adjust, and the inference mechanism is quite simple. Despite its simplicity, such a model can be used to solve quite complex problems. That is why the production approach was chosen to create DSS.

Let us consider in general form the following system. There is a certain analyzed object with a set of primary features that it can transmit to the input of the system (in this case, both a person and an inanimate object can be analyzed). In addition, there is a person who directly works with the system (hereinafter - the Specialist). After the Specialist receives a preliminary decision or a set of decisions based on the results of the examination and determining the treatment plan. However, the approach presented in this article is universal and can find its application in various subject areas where it is necessary to diagnose something by a set of signs and determine a set of further actions. For example, such a DSS can be used to diagnose malfunctions in various technical and biological systems.

The logical conclusion of the decision-making system is based on facts and rules, as mentioned above. When implementing such a system, each potential solution is associated with a set of features, one feature may be included in feature sets for different solutions.

Moreover, in the course of inference, the following types of rules are used (depending on the solution): the decision is confirmed if at least one of the characteristics set in accordance with it is present in the analyzed object; the decision is confirmed if all the attributes put in accordance with it are present in the analyzed object.
In the process of defining a set of solutions, the features of the analyzed object are correlated with the decisions in accordance with the rules from the knowledge base. In order to provide the most complete understanding of how the system functions, a block diagram of the algorithm of the system’s operation in general form is given (Figure 2). From the point of view of the business process, the system looks as shown in Figure 3.

As testing of a system implemented on the basis of the described approach shows, using data from open sources, this method of decision making is quite accurate. However, the described DSS is intended not only to determine a set of decisions, but also to determine a general plan of further actions based on this set of decisions.

3. Defining a plan for further action
A production approach is also used to define an action plan. Each action plan consists of elements of an action plan (separate instructions for various actions). The same element may be present in several action plans. Moreover, the characteristics (for example, duration, quantity) of this element may be different for different solutions, but may be completely identical. Each decision is associated with an action plan, i.e. set of elements (instructions for action). The rule used is: “if the decision is confirmed, then all components of its action plan should be prescribed to the object of analysis”.

However, when the system determines a set of decisions, repeating elements may occur in the action plans, which forces to display information with duplication. A situation may also arise when the elements of different plans are similar, but have different parameters: for example, when considering on a medical system, it could be the dosage (if it is a medicine), duration and other parameters. In this case, ambiguity occurs, which can lead, for example, to an overdose of the patient. Such a definition of a treatment plan is unacceptable [11]. Therefore, it is not possible to use the production approach in its pure form, without modifications.
4. Modification of the production approach

Such a modification for defining an action plan can be the use of a production approach in conjunction with set theory. After defining a set of solutions, there are two sets: solutions and actions. It is necessary to match the elements of the action plan with the elements of the set of decisions, but in such a way that there are no repetitions, and also highlight the more significant elements of the action plan if they are repeated at the same time for several decisions from the set of specific ones, but differ in characteristics.

Suppose that the patient was given several diagnoses by the system, treatment plans for two of which require the patient to use the same drug, but in different dosages. Then the system must make a choice in favour of that element of the plan, the dosage of which is higher, but in no case should not summarize and duplicate these elements.
This task could be accomplished with the initial distribution of action plan elements by groups and their further ranking within these groups takes place. An example of such a group is, for example, “Taking antibiotics”. Taking the same antibiotic depending on the disease can last a different amount of time. Then, the longer the duration, the higher the rank of the element within the group, however, the ranking takes place according to a number of parameters, since taking antibiotics also implies some specific drugs and their dosage. The ranking of treatment elements is performed by the Specialist in compiling the knowledge base. In the particular case, this means that in the case of repeating elements for several action plans, the choice will be made in favour of the one whose rank is higher.

Thus, after determining the subsets (groups) of the set of elements of the action plan, the most significant ones are identified, while the insignificant elements are not taken into account at all, that is, they can actually be excluded from the set. As a result, only unique and significant elements of the action plan remain that fully correspond to a specific set of decisions, which will eliminate duplication of elements, in particular, the possibility of overdose or poor-quality treatment is excluded in the medical system. In addition to the above, upon receipt of two sets with unique (non-repeating) elements, it is necessary to relate them to each other. For this, in set theory, there is a Cartesian product that has been applied in database theory, represented by the JOIN operator in SQL. In simple terms, there is a comparison (connection) of the elements of two database tables that fall under the general condition.

As mentioned above, systems of this kind can be used to carry out any diagnostics. In this case, the user of this system should be given an action plan to eliminate or minimize the shortcomings found during the diagnosis. Moreover, this plan should be logical and understandable to the contractor. Redundancy, as well as incompleteness of the action plan, is unacceptable in order to avoid inefficient use of resources to eliminate or minimize the most complete list of identified deficiencies. For this, it is necessary to get rid of possible duplicates and similar actions with various parameters. Correlation of the elements of the action plan with the diagnostic results is necessary to ensure that the performer understands these actions of the entire plan.

Of course, such a description of the selected solution is greatly simplified in order to gain understanding by the end reader. In fact, everything is more complicated, but it is successfully implemented using modern tools, in particular, the SQL language. It allows us joining tables literally in several rows, thereby producing a Cartesian product with the simultaneous indication of the necessary conditions, which directly proves that there is no need to use ANNs for such a difficult task.
5. System implementation example

For the software implementation, the Python programming language and the Django framework were chosen. This choice is due to many factors, the main of which is the simplicity of building the database architecture and access to the information stored in it. Tables are defined by classes (models) that describe the structure of the stored data, including field types, their settings and behaviour. Also, this toolkit eliminates the need to write huge SQL queries to display the necessary information, just use the Django Object-Relation Mapper (ORM) - a technology that connects databases with the concepts of object-oriented programming. In other words, it is a layer between the database and the code that represents the data as objects. ORM has a fairly simple syntax and has a large number of methods for manipulating data: saving/updating / deleting, accessing related tables, filtering, sorting, and much more. This approach greatly simplifies working with the database, because it does not require extensive knowledge of SQL.

The database structure of the diagnosis module of the implemented DSS is presented in figure 4.

![Diagram of database structure](image)

Figure 4. Data structure.

This structure allows us flexibly combining data and more accurately diagnose. There are basic models: Patient - patient information; Diagnosis - possible diagnoses; Sample - current state of the patient; Symptom - possible symptoms; Value - symptom values; MedicalDisposalItem - elements of a treatment plan. Using intermediate tables, data is correlated with each other. The SampleSymptomValue model allows us storing data about the symptoms that are currently observed in the patient. Using the DiagnosisSymptomValue model, symptoms and their meanings are correlated with the diagnosis. Based on these models, the system determines possible diagnoses for the patient, which vary depending on changes in SampleSymptomValue. Symptoms are determined on the basis of many factors: the patient filling out a medical and dental questionnaire, doctor's supervision, test results and additional studies, and others. The MedicalDisposalItemDiagnosis model allows us correlating diagnoses with the elements of treatment plans that are necessary to confirm this diagnosis.
6. Conclusion

In conclusion, it is worth noting that the application of a production approach to solving diagnostic problems is still advisable. Since software systems for decision-making support of this kind can be relatively quickly and easily implemented, due to the simplicity of the production approach itself, their effectiveness is not lower than that of well-trained neural networks (of course, if the set of possible decisions is deterministic.

Moreover, the application of the production approach in conjunction with the theory of sets allows not only correctly drawing a conclusion about the state of the analyzed object on the basis of its features, but also to decide on further actions with this object. Moreover, operations with sets are defined in the standards of many modern programming languages (for example, in Python and Julia), including, and especially in SQL. That allows us simply applying set theory in the software implementation of DSS together with almost any decision-making approach, including production.

Thus, using a combination of the production approach and set theory, a decision-making support system in medicine was implemented, with the help of which a set of diagnoses for the patient is made and a treatment plan corresponding to this set of diagnoses is determined. The developed system was tested on real data in one of the clinics of Krasnoyarsk, where it showed its effectiveness in decision-making in combination with speed. In the future, the application of set theory to determine possible options for action based on a certain set of states of any object can be applied to other tasks. For example, to determine the direction of motion of an unmanned aerial vehicle after recognizing the image that arose in front of it. Based on the foregoing, we can conclude that the application of classical decision-making and information processing methods is still relevant and appropriate when developing decision support systems.

References
[1] Turban E 1988 Decision support and expert systems: Managerial perspectives (New York: Macmillan)
[2] Marakas G M 2003 Decision support systems in the 21st century (Vol. 134) (Upper Saddle River, NJ: Prentice Hall)
[3] Yam R C M, Tse P W, Li L and Tu P 2001 Intelligent predictive decision support system for condition-based maintenance The International Journal of Advanced Manufacturing Technology 17(5) 383-91
[4] Morozova V A and Pautov V I 2017 Knowledge representation in expert systems: Tutorial (Yekaterinburg: Ural University Press)
[5] Beleites C 2012 Sample size planning for classification models Analytica Chimica Acta 760 25-33
[6] Anzai Y 1979 Knowledge-based problem solving by a labelled production system In: 6th int. joint conf. on Artificial intelligence IJCAI'79 22-4
[7] Bukhtoyarov V V, Tynchenko V S, Petrovskiy E A, Bashmur K A, Kukartsev V V and Bukhtoyarova N A 2019 Neural network controller identification for refining process J. Phys.: Conf. Ser. 1399 044095
[8] Milov A V, Tynchenko V S, Kukartsev V V, Tynchenko V V and Bukhtoyarov V V 2018 Use of artificial neural networks to correct non-standard errors of measuring instruments when creating integral joints J. Phys.: Conf. Ser. 1118 012037
[9] Tynchenko V S, Tynchenko V V, Bukhtoyarov V V, Kukartsev V V, Kukartsev V A and Ereemeev D V 2019 Application of Kohonen self-organizing maps to the analysis of enterprises’ employees certification results IOP Conf. Ser.: Mater. Sci. Eng. 537 042010
[10] Yan H, Jiang Y, Zheng J, Peng C and Li Q 2006 A multilayer perceptron-based medical decision support system for heart disease diagnosis Expert Systems with Applications 30(2) 272-81
[11] Harkevich D A 2010 Pharmacology textbook for universities (Moscow: GEOTAR-Media)