УДК 004.318

MODEL OF KNOWLEDGE BASE FOR FORECASTING DEVELOPMENT OF COMPLICATIONS IN HEART ATTACK OF MYOCARDIUM

Burdaev V.P. / Бурдаев В.П.
Ph.D. of Physico-mathematical Sciences, as.prof. / к.ф.-м.н., доц.
ORCID: 0000-0001-9848-9059
Simon Kuznets Kharkiv National University of Economics, Kharkiv, Nauky avenue, 9-A, 61166

Abstract. *The article considers the knowledge base model of a hierarchical functional system for predicting the development of complications of myocardial infarction, based on the concept of knowledge base bundle. The knowledge base model helps the medical staff to make a decision in the diagnosis of complications after myocardial infarction. For decision making, the Bayes formula is used based on the clinical data of patients after myocardial infarction. The knowledge base model is implemented using the "KARKAS" system - toolkit to create knowledge base models. The following are given: decision-making scenario, examples of rules, knowledge base frames and the type of hierarchical functional system of the knowledge base model. The expert system can be consulted online using the @Ribs_karkas_bot chat bot.*

Key words: *hierarchical functional system, knowledge base, expert system, Bayes formula.*

Introduction.

Methods for constructing knowledge base models are often based, albeit on imprecise, but generally objective information about an object. However, situations are possible when, when building models of knowledge bases, the information obtained from an expert is of decisive importance, usually of a qualitative nature. They reflect the substantive features of the object being studied and are formulated in a natural language. The description of the object in this case is fuzzy.

It should also be noted that in this case the coefficient of certainty is assigned not only to the rules, but also to the facts of the knowledge base.

One of the problems that arise when creating expert systems is to account for the inaccuracy and unreliability of any information. The first expert systems introduced such a concept as robustness of the system, that is, the ability to produce a plausible conclusion from incomplete, not accurate information.

In addition, there is the task of assessing the degree of unreliability of logical conclusions and recommendations with inaccurate initial information and information received as a result of a dialogue with the user.

Among the diseases, diseases of the cardiovascular system take the first place, especially complications of myocardial infarction. Therefore, the problem arises on the basis of the first manifestations of the disease before the start of medical intervention to try to predict the occurrence of complications of myocardial infarction.

The purpose of the given work is dedicated to development a knowledge base model for predicting complications after myocardial infarction using tools for creating knowledge models of the "KARKAS" system.

Inference engine based on Bayes formula for expert system.

Consider one of the methods of reasoning in the presence of uncertainty, based on the Bayes formula from the theory of probability, implemented in the "KARKAS" system inference engine [1, 2].

Suppose there is some hypothesis H and its a priori probability that the hypothesis H is true. This probability $P(H)$ is set at the very beginning (a priori probability) of consultation with the expert system.

Further, it is assumed that there is some evidence E relating to this hypothesis, then on the basis of this information it is possible to clarify the a priori probability of the truth of the hypothesis H (a posteriori probability). According to the Bayes formula, we have:

$$P(H \mid E) = P(E \mid H) \times \mathrm{P(H)} / (\mathrm{P(E \mid H)} \times \mathrm{P(H)} + \mathrm{P(E \mid \overline{H})} \times \mathrm{P(\overline{H})}),$$

where $P(H)$ – a priori probability of a hypothesis H in the absence of any evidence; $P(H \mid E)$ – a posteriori probability of a hypothesis H in the presence of evidence E; $P(\overline{\mathrm{H}}) = 1 - \mathrm{P(H)}$ – the probability of the event not occurring H; $P(H \mid E), P(E \mid H)$ − conditional probabilities.

Using the Bayesian formula, it is possible to accumulate information from various sources in order to confirm or not confirm a certain hypothesis.

The algorithm inference engine of the "KARKAS" system using the Bayes formula is implemented as follows:

Step 1. Let be a prior probability $P(H)$, which is stored in the database of facts.

Step 2. Choose a evidence E.

Step 3. Calculate $P(H|E)$.

Step 4. Replace $P(H|E)$ in place $P(H)$.

Thus, obtaining a new evidence leads to a new update (increase or decrease) of probability $P(H)$. Each time the current value of this probability will be considered a priori for applying the Bayes formula.

Knowledge base model for the "INFARKT" expert system.

With the help of the "KARKAS" system, the "INFARKT" expert system was created, which can carry out consultations both offline and online using the Telegram @Ribs_karkas_bot chat bot [3]. In fig. 1 is a screenshot of the start of a consultation using a chat bot.

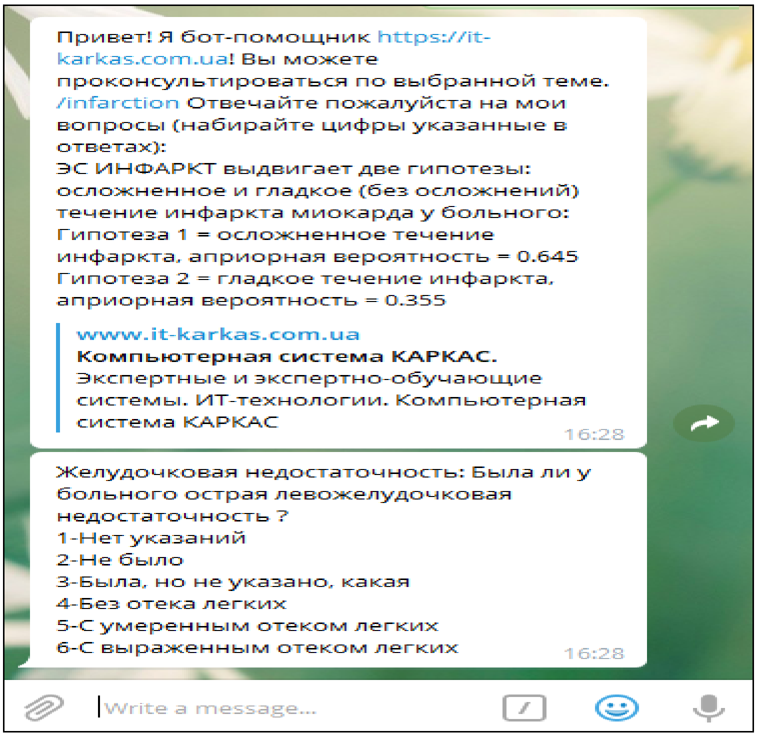

Fig. 1. Online consultation using the Telegram @Ribs_karkas_bot chat bot

Step 1. Initially, the consultation of the "KARKAS" inference engine system puts forward two hypotheses: complicated and smooth (without complications) myocardial infarction in a patient. The a priori probability of the first hypothesis is 0.645, and the second is 0.345 (fig. 1).

Step 2. Then, in a process of consultation with a doctor or nurse, the inference engine calculates the chances of these hypotheses based on the patient's symptoms and history (Fig. 2).

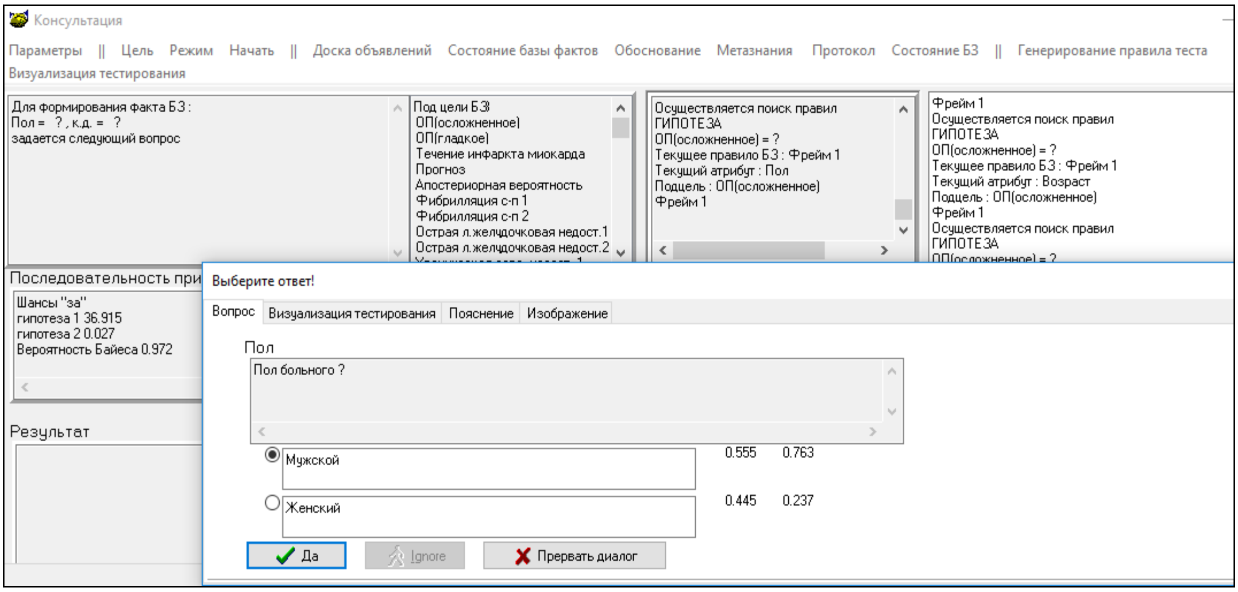

Fig. 2. Selection of symptom indicator (offline mode)

Chances and probabilities are related by the following formula:

$$O(H) = P(H) \backslash (1 - P(H)),$$

so for the first hypothesis: $O(H_1) = 1.816,$ and for the second hypothesis: $O(H_2) = 0.550.$

Step 3. In fig. 3 shows a screenshot of the operation of the inference engine in the third step of the consultation.

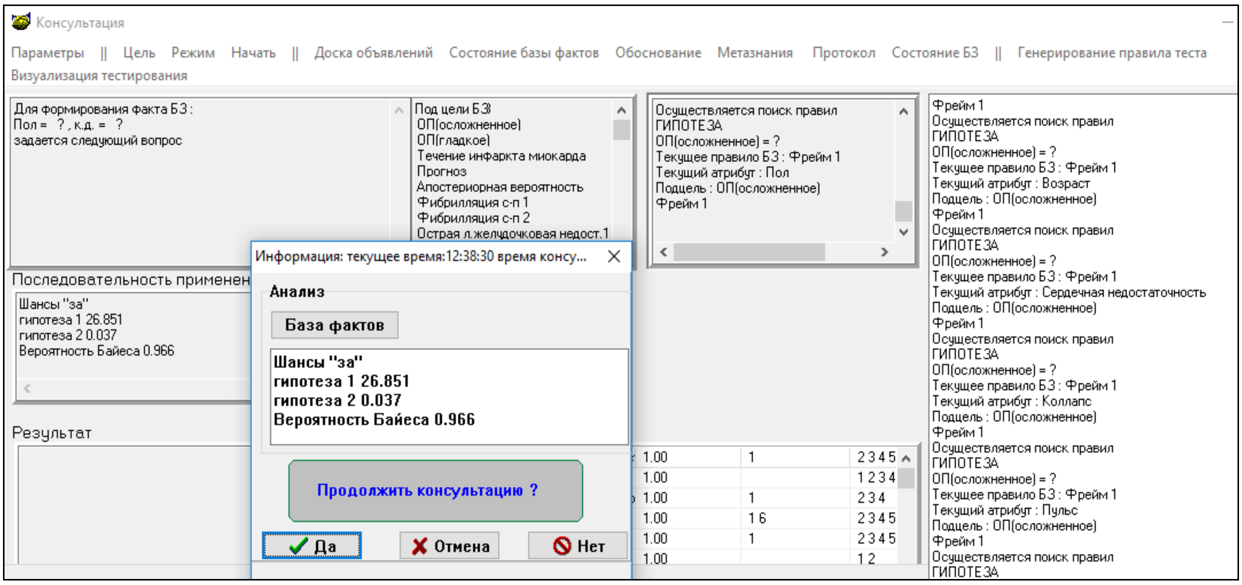

Fig. 3. Calculation of the chances of hypotheses (offline mode)

If the first hypothesis is accepted, then the analysis of the calculated posterior probability begins, which serves as the basis for diagnosing the degree of development of the corresponding complication of a heart attack. In the third step of the consultation, the posterior probability for the first hypothesis is 0.966.

To predict complications of myocardial infarction based on data from the initial phase of the acute period, a start-period is used. Start-period is defined as the time from the first manifestations of the disease to the start of medical intervention. For example, start-period 1 is 24 hours, start-period 2 is 48 hours. The rules specifying the diagnosis use these two start-periods.

If the second hypothesis is accepted, then the inference engine ends up and a message is issued that there are no complications after a heart attack.

Step 4. In fig. 4 and fig. 5 shows the types of results of consultations with the "INFARKT" expert system, respectively: offline and online.

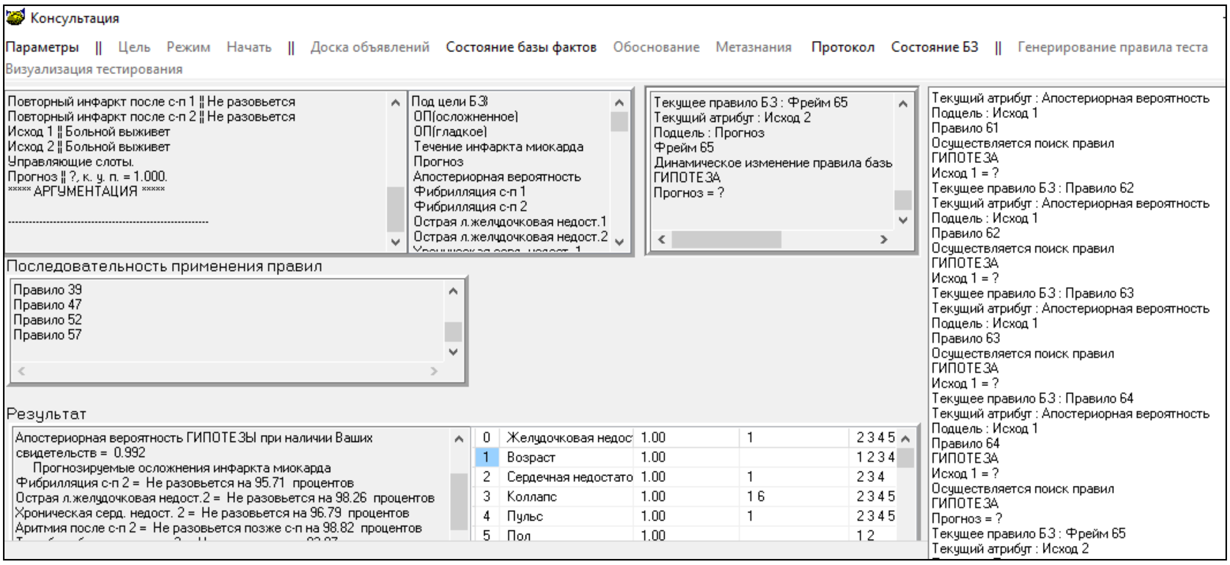

Fig. 4. Type of result of consultation with the expert system "INFARKT" (offline mode)

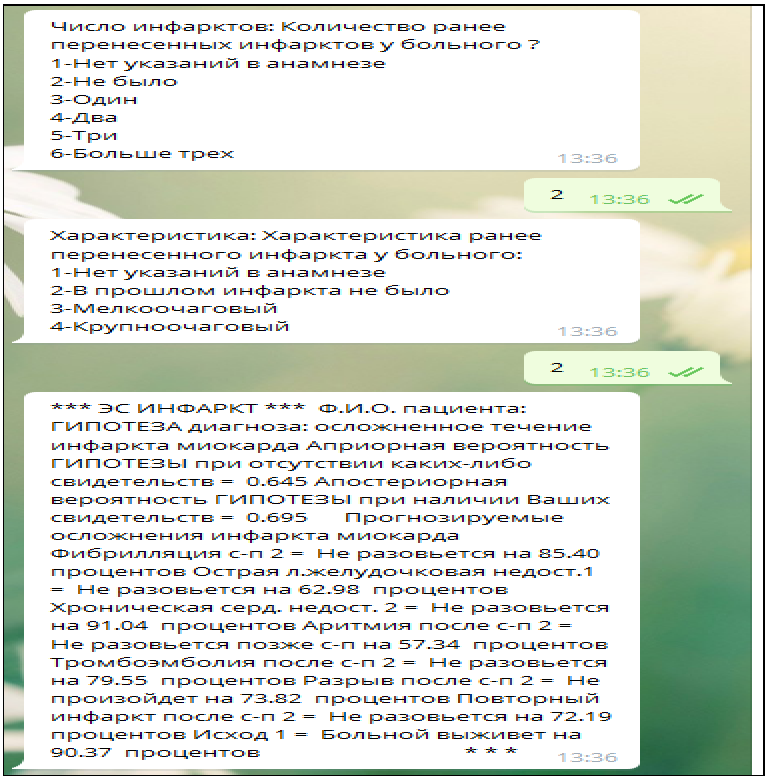

Fig. 5. Type of result of consultation with the "INFARKT" expert system (online mode)

Format of rules, frames and hierarchical functional system of the knowledge base model of the expert system "INFARKT"

In fig. 6 presents the values of a priori probabilities of the symptom "Hypertension" with respect to hypotheses: a smooth and complicated course of myocardial infarction.

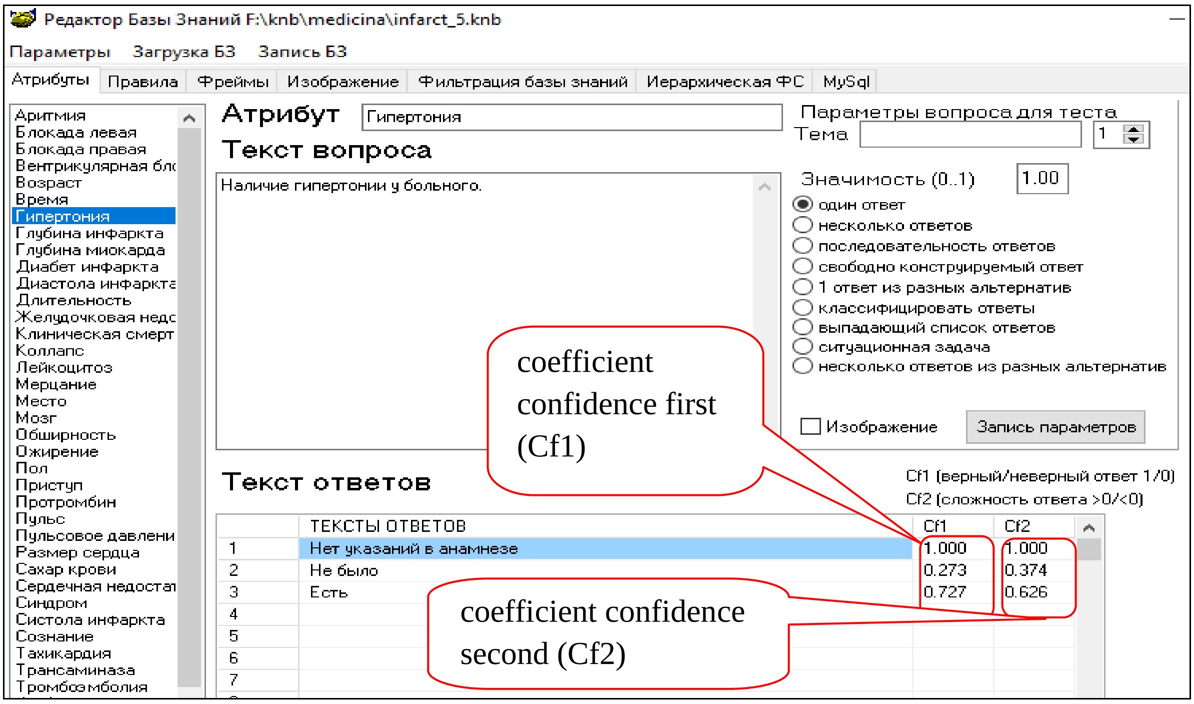

Fig. 6. A priori probabilities of the symptom (hypertension) for two hypotheses (offline mode)

The attribute list contains 28 symptoms of the disease. For express diagnostics, 7 of the most informative were selected using statistical methods, which are presented in the following order of decreasing informativeness: ventricular failure; age; heart failure; collapse; pulse; paul; the number of heart attacks; a characteristic of a previous heart attack (Fig. 7).

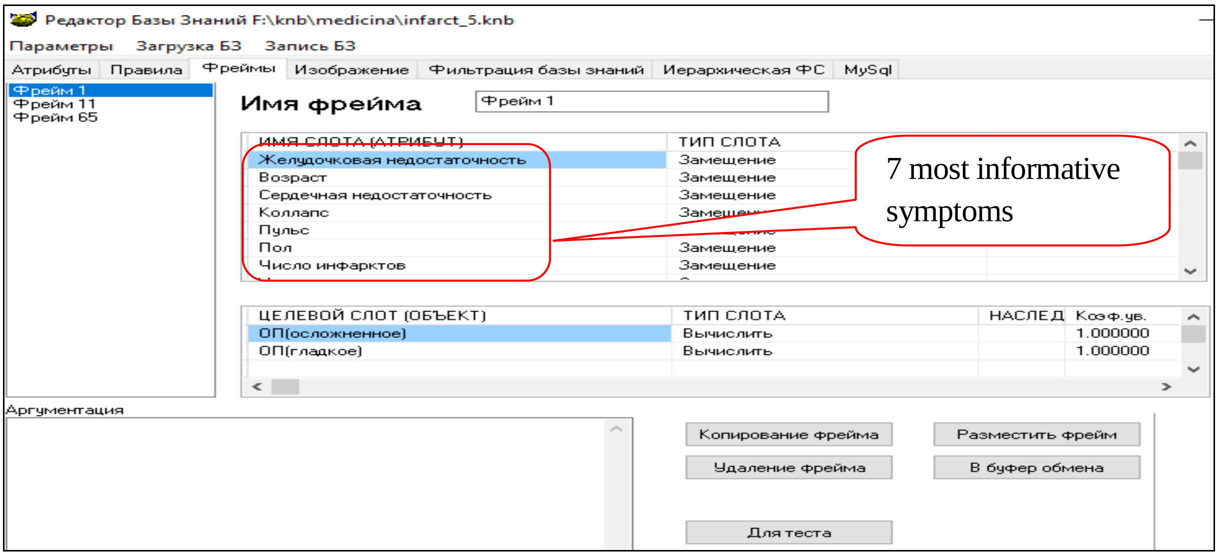

Fig. 7. The frame with the seven most informative simpots (offline mode)

For each question attribute (symptom) and its corresponding value (answer to the question):

coefficient confidence first (Cf1) means a prior probability for the first hypothesis (complicated course of myocardial infarction in a patient);

coefficient confidence second (Cf2) means a prior probability for the second hypothesis (smooth course of myocardial infarction in a patient).

When the user answers the question about the presence of a symptom, the inference engine forms the fact: attribute = value, Cf1, Cf2, where the coefficients of confidence are recalculated according to the Bayes formula. The calculation of confidence ratios was based on clinical data from [4].

The format of the attribute record for the knowledge base.

Attribute Name: Ventricular insufficiency.

Did the patient have acute left ventricular failure?

Answers:

1. There is no indication, Cf1 = 1.000, Cf2 = 1.000.
2. Was not, Cf1 = 0.649, Cf2 = 0.906.
3. It was, but not indicated what, Cf1 = 0.094, Cf2 = 0.020.
4. Without pulmonary edema, Cf1 = 0.063, Cf2 = 0.037.
5. With moderate pulmonary edema, Cf1 = 0.059, Cf2 = 0.016.
6. With severe pulmonary edema, Cf1 = 0.135, Cf2 = 0.021.

Example frame and knowledge base rules.

Frame 1.

Slot name | Slot type | Inheritance

Ventricular failure | Substitution |

Age | Substitution |

Heart failure | Substitution |

Collapse | Substitution |

Pulse | Substitution |

Paul | Substitution |

The number of heart attacks | Substitution |

Characteristics | Substitution |

Goal slot

Chance forecasting (complicated) | Calculate.

Chance forecasting (smooth) | Calculate.

Rule 2. A & B #.

IF

A Chance forecasting (complicated)> 1

B Chance forecasting (smooth) <1

THEN

The course of myocardial infarction = Complicated.

Rule 10. A & B #.

IF

A Chance forecasting (complicated) <1

B Chance forecasting (smooth) = 1

THEN

Forecasting = smooth myocardial infarction.

In fig. 8 shows a view of the hierarchical functional system for determining the value of the consultation goal – "Forecast".

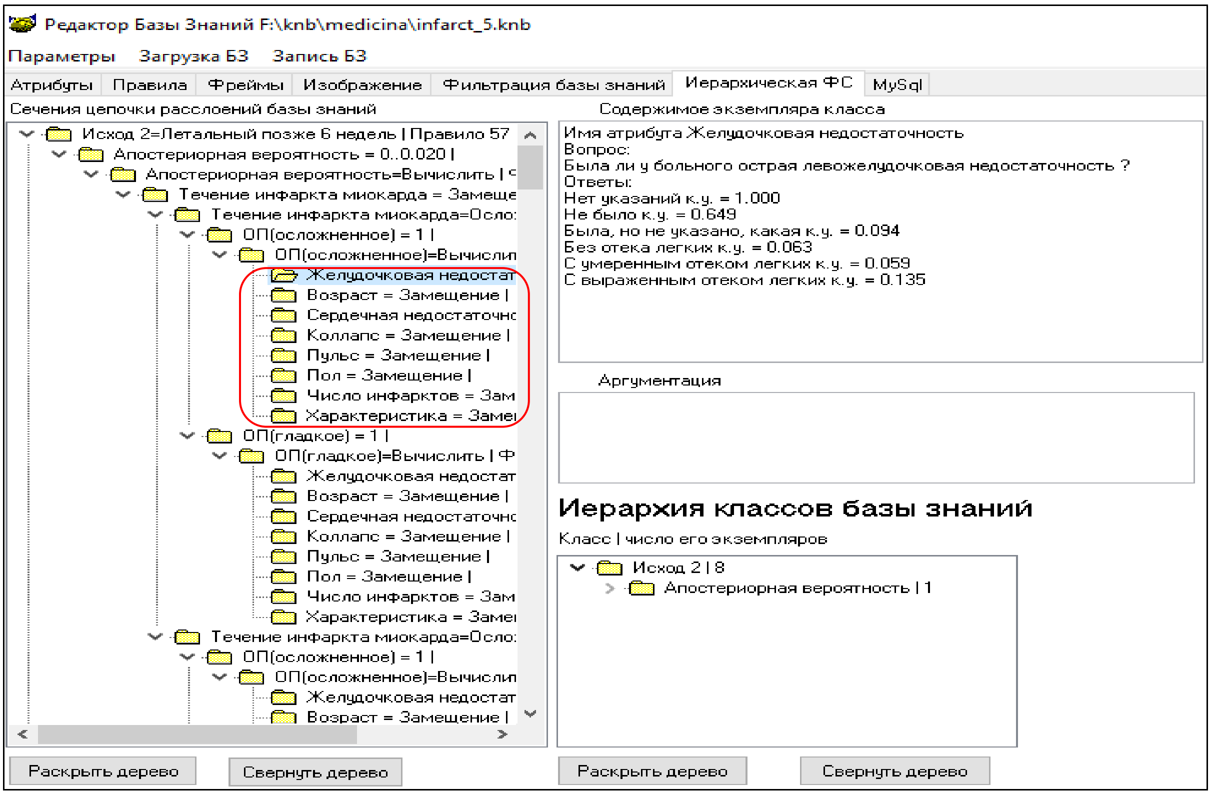

Fig. 8. Type of hierarchical functional system for determining the value of the consultation goal – "Forecast" (off-line)

Rule 12. A #.

IF

A A posteriori probability [] 0..0.009

THEN

Start-period 1 (fibrillation) = Will develop later than 6 weeks.

Start-period 2 (fibrillation) = Will develop later than 6 weeks.

Rule 30. A #.

IF

A A posteriori probability [] 0.737..1

THEN

Chronic heart not enough 1 = Will not develop.

Chronic heart not enough 2 = Will not develop.

Rule 41. A #.

IF

A A posteriori probability [] 0.015..0.017

THEN

Thromboembolism after the start-period 1 = Will develop in 1 day, later the start-period.

Thromboembolism after the start-period 2 = Will develop at 2 days.

Rule 64. A #.

IF

A A posteriori probability [] 0.662..1

THEN

Outcome 1 = The patient will survive.

Outcome 2 = The patient will survive.

Frame 65.

Slot name | Slot type | Inheritance

Start-period 1 (fibrillation) | Substitution | n

Start-period 2 (fibrillation) | Substitution | n

Acute l.ventricular netsu.1 | Substitution | n

Acute l.ventricular lack.2 | Substitution | n

Chronic heart not enough 1 | Substitution | n

Chronic heart not enough 2 | Substitution | n

Arrhythmia after the start-period 1 | Substitution | n

Arrhythmia after the start-period 2 | Substitution | n

Thromboembolism after the start-period 1 | Substitution | n

Thromboembolism after the start-period 2 | Substitution | n

Gap after the start-period 1 | Substitution | n

Gap after the start-period 2 | Substitution | n

Repeated heart attack after a start-period 1 | Substitution | n

Repeated infarction after a start-period 2 | Substitution | n

Exodus 1 | Substitution | n

Exodus 2 | Substitution | n

Goal slot

Forecast | ?

Conclusions.

A knowledge base model is proposed for predicting complications after myocardial infarction based on a hierarchical functional system using the Bayes formula.

Based on the viewed knowledge base model, a prototype of the "INFARKT" expert system was created, which helps doctors diagnose infarction patients, assess their condition and forecast the development of the following complications of myocardial infarction: fibrillation, acute left ventricular failure, chronic heart failure, arrhythmia, thromboembolism, rupture myocardium, repeated heart attack.

The "INFARKT" expert system can provide advice both offline and online using the @Ribs_karkas_bot chat bot.

A trial version of the "KARKAS" system is available at https://it-karkas.com.ua (offline mode).

The @Ribs_karkas_bot chat bot (Telegram messenger) is available online: to do this, send a message using the chat site https://it-karkas.com.ua to activate it.

Consultation with the "INFARKT" expert system in case of 28 symptoms takes 10 minutes, and the option of express consultation (7 symptoms) takes 3 minutes.

When using chat bot consultation can be carried out mobile using a smartphone.

Literature:

1. Бурдаев В.П. Модель функциональной системы динамической предметной области. // Искусственный интеллект. Наука і освіта ІПШІ МОН України і НАН України - 2011. - №3. - С.439 - 448.

2**.** Бурдаєв В.П. Моделі баз знань. Харків, Україна: ХНЕУ, 2010.

3. Бурдаєв В.П. Інтегрування месенджерів з системою "КАРКАС". // Тези доповідей: міжнарод. наук.-практ. конф. Проблеми і перспективи розвитку ІТ-індустрії, Харків, 2018, с.7.

4. Штейн Л. Б. Опыт прогнозирования в медицине с помощью ЭВМ. / Под ред. В. М. Ахутина, В. А. Якубовича. – Л.: Изд. Ленингр. ун-та, 1987. – 146 с.

References:

1. Burdaev, V., 2011. Model` funkcy`onal`noj sy`stemы dy`namy`cheskoj predmetnoj oblasty` [Model of the functional system of the dynamic domain] in Y`skusstvennыj y`ntellekt [Artificial Intelligence], issue 3, pp. 439 - 448.

2**.** Burdaev, V. 2010. Modeli baz znan` [Knowledge base models]. Kharkiv: Simon Kuznets Kharkiv National University of Economics.

3. Burdaev, V., 2018. Integruvannya mesendzheriv z sy`stemoyu "KARKAS" [Integration of messengers with the "KARKAS" system] in Problemy` i perspekty`vy` rozvy`tku IT-industriyi [Problems and Prospects of IT Industry Development]: Mizhnarod. nauk.-prakt. konf. [International sci. pract. conf.], Kharkiv, p.7.

4. Shtein, L., 1987. Opыt prognozy`rovany`ya v medy`cy`ne s pomoshh`yu ЭVM [The experience of forecasting in medicine with the help of computers]. Leningrad : Ed. Leningrad University.

Аннотация.

В работе рассматривается модель базы знаний иерархической функциональной системы для прогноза развития осложнений при инфаркте миокарда, на основе понятия расслоения базы знаний. Модель базы знаний позволяет помочь медицинскому персоналу принять решение при диагностике осложнений после инфаркта миокарда. Для принятия решения используется формула Байеса на основе клинических данных пациентов, перенесших инфаркт миокарда. Модель базы знаний реализована с помощью системы "КАРКАС" – инструментария для создания моделей баз знаний. Приведены: сценарий принятия решения, примеры правил, фреймов базы знаний и вид иерархической функциональной системы модели базы знаний. Консультацию с экспертной системой можно осуществлять онлайн режиме с помощью чат-бота @Ribs_karkas_bot.

Ключевые слова: *иерархическая функциональная система, база знаний, экспертная система, формула Байеса.*

Article submitted: February 26, 2019