ABOUT ONE IMPLEMENTATION OF AN EXPERT SYSTEM ON THE ANDROID PLATFORM

ABOUT ONE IMPLEMENTATION OF AN EXPERT SYSTEM ON THE ANDROID PLATFORM

Burdaev V.P. / Бурдаєв В.П.
c.ph.math.s., as.prof. / к.ф.м.н., доц.
ORCID: 0000-0001-9848-9059
National Technical University "Kharkiv Polytechnic Institute",
Kharkiv, Kyrpychova 2, 61002

Abstract. The article presents the development of an expert system for decision support on the Android platform. To create a knowledge base, a medical subject area was chosen using the example of determining the risk of coronary heart disease in a practically healthy person. The main and additional risk factors for coronary heart disease were identified, which determined the hierarchical structure of the knowledge base. The knowledge base structure contains knowledge clusters so that the user can consult with the expert system depending on the choice of clusters. The knowledge base was implemented using the KARKAS shell, which has tools for creating expert systems: a visual knowledge base editor, a knowledge base parser, a hierarchical functional system for forming a knowledge base structure based on rules and frames, and various modules for implementing logical conclusions.

Key words: knowledge base, expert system, mobile application, Android and Embarcadero platforms.

Introduction.

In Expert systems (ES) in the field of artificial intelligence are one of the first successes of its application [1 - 3]. They are intended to assist the user in making decisions in informal subject areas. The main modules of the expert system are the knowledge base and the inference engine, which gives advantages over neural networks, since it allows you to get the reasoning behind the decision. The traditional ways to implement an inference engine are backwards, forwards, mixed, Bayesian chains of reasoning. Due to the asynchrony of the Android operating environment, the inference engine algorithm of a mobile application mainly uses direct and Bayesian chains of reasoning. To resolve conflicts in the choice of rules and frames during the consultation, various strategies are used:

• rules are selected only once;
• rules with a higher priority will be selected first (priority is set by the expert);
• the rules are selected first if they have more complex conditions than the same ones;
• the order of the rules was fixed when they were filtered, and others [4].

The global market for mobile applications for Android is constantly growing. Here are some of the benefits of the mobile app:

• mobile applications are more convenient for communicating with users than chatbots and websites, for example, they provide access to native smartphone functions (for example, a video camera, GPS navigator, voice recognition
functions);

- mobile applications increase customer loyalty (high conversion), as the smartphone is always with the client (for example, they allow you to make timely decisions);
- mobile applications effectively influence users by sending push notifications.

Taking into account current trends, the development of mobile applications that perform the functions of expert systems is an urgent problem. Since there is a constant need to adapt them for use on mobile devices.

Various frameworks (shells) are used to create mobile expert systems: Flutter, Ionic, React Native, Android Native, Xamarin and others.

The shell for creating knowledge bases KARKAS was created using the FireMonkey (FMX) cross-platform framework from Embarcadero, which is part of the RAD Studio development environment and is designed to create user interfaces. The framework allows you to use not only vector graphics, but also the native capabilities of mobile devices. In addition, there is another great feature of the framework, which is that the application code can be compiled into native code to run on different platforms: Windows, Android and iOS. Thus, using the FMX framework, you can quickly create prototypes of mobile applications for various mobile devices [5,6].

Using the KARKAS system, chatbots were developed: @Ribs_karkas_bot, @es_test_karkas_bot, @es_economy_karkas_bot, @es_info_tech_karkas_bot, which allows you to conduct online consultations with users and test students' knowledge [7 - 9].

One of the main advantages of mobile applications in customer service is that interlocutors are free to ask questions that they would not ask a support representative or company manager.

Taking into account current trends, the development of mobile applications that perform the functions of expert systems is an urgent problem. Since there is a constant need to adapt them for use on mobile devices.

Using the KARKAS system, a knowledge base was developed for the ES_RFCHD mobile application (an expert system for determining the risk of coronary heart disease). The main function of this mobile application is to provide the user with control over the level of risk factors for coronary heart disease.

The paper [10] considers a mobile expert system implemented on the Android platform. The system uses a rule-based knowledge base that provides diagnostics and medical advice for ten common diseases in Nigeria. The knowledge base contains five tables that contain information about symptoms, diagnoses, causes of diseases, and preventive recommendations. The system uses a direct chain of reasoning.

This article [11] is devoted to the development of a decision support expert system model for the Android platform without serious programming in java programming. The knowledge base of the system for this knowledge area was formed by inductive learning methods based on examples from the WEKA data mining system, and the system was implemented using the Expert2Go shell and the e2gDroid Lite Expert shell for mobile devices. The knowledge base contains a total of 28 rules and 7 classes.

The article [12] considers the approach of creating a mobile application of an expert system using a linear sequential software development model or also known as a
waterfall model. The development phases in this model include the process of knowledge analysis, design, coding or implementation, and testing.

This paper [13] presents a mobile expert system that uses knowledge representations based on rules and a knowledge modeling method using a decision tree. To implement this system, the Java programming language was used, and the user interface was designed and developed using XML, which fits into modern smartphones.

The article [14] considers the implementation of an expert system capable of diagnosing diabetes mellitus. The ExpeRT system uses Bayesian method and MySQL database for early diagnosis of diabetes mellitus, which allows it to be a good information carrier for storage, education and detection of diabetes mellitus.

The work [15] analyzes an expert system aimed at diagnosing various types of infectious diseases and their complications. The constructed system uses an inference engine based on the inverse reasoning method to recognize infectious diseases such as pharyngitis, diphtheria and tuberculosis. In addition, the system uses a fuzzy inference system to determine the severity of the disease.

The article [16] presents the development of a medical expert system based on fuzzy sets for assessing the degree of coronary artery disease in patients with coronary heart disease. The inference engine uses the fuzzy set method to solve problems of medical diagnostics, in particular, in assessing the degree of anatomical lesions of the coronary arteries in patients with various forms of coronary heart disease.

The article [17] presents an expert system for diagnosing mood disorders, which is designed for an Android application. JavaScript and PHP are used as a programming language, Cordova as a framework for creating mobile applications, PHP 7.2.1 as a web server, MySQL 5.0.21 as a database. The inference engine uses Bayes' theorem. This expert system can produce a process for diagnosing mood disorders with some data about mood disorders in the form of symptomatic data, disease data and its probabilistic significance.

The work [18] considers an expert system that helps patients with cholesterol diagnose the disease. The application uses a direct chain of reasoning inference mechanism when making decisions. The system is integrated with databases and programming languages such as PHP-MySQL.

The article [19] proposes a model of a mobile medical application for Android for diagnosing breast cancer and daily health forecasting. The user interface is implemented using Android/Java programming, and the disease diagnosis part is developed using disease symptom data and using direct reasoning chain.

The work [20] presents an expert system for diagnosing facial skin diseases using the direct chain of reasoning method and the certainty factor. The developed expert system for diagnosing facial skin diseases can diagnose 10 types of diseases.

In [10 - 20] methods, algorithms and procedures for their implementation for expert systems are described.

**Formulating the Purpose of the Article.**

Analysis of the construction of a knowledge base for determining the risk of coronary heart disease in humans and its implementation in an expert system on the Android platform.
Main text

Consider the structure of the knowledge base for determining the risk of coronary heart disease (CHD) of the ES_RFCHD expert system. The relevance of its development is explained by:

1) firstly, the evolutionary transition from CHD therapy to CHD prevention; from population prevention to individual prevention;
2) secondly, the desire of a person to independently acquire knowledge on the risk assessment of coronary heart disease;
3) thirdly, the possibility of self-monitoring of changes in the risk of coronary heart disease to make a decision to seek medical advice with a high degree of it.

The appointment of ES is a preventive consultation of the patient on the assessment of the risk of coronary heart disease.

The scope of ES is various medical enterprises: dispensaries, clinics, medical units.

The purpose of the ES is to model the decision-making on the risk of coronary heart disease in a practically healthy person.

Initial data: medical tests.

Expected results (list of possible consultation goal values):
- individual prevention of coronary heart disease;
- self-acquisition of knowledge about the degree of risk of coronary heart disease;
- self-control of the risk level of coronary heart disease;
- receive appropriate recommendations to reduce the risk of coronary heart disease.

Subject area identification. A risk factor is understood as the social, biological and economic status of a person, his behavior patterns and conditions that contribute to the occurrence of coronary heart disease.

According to the results of expert studies, it was established that the activity of an outpatient doctor should be aimed at identifying only the main risk factors for coronary heart disease, namely: hypercholesterolemia, arterial hypertension, smoking, overweight, poor nutrition and alcohol. The reasonableness of this recommendation is due to the fact that the main function of the doctor is to diagnose the disease itself, and not its prenosological conditions. The practical priority of prenosological diagnostics contributed to the fact that the expert group expanded the number of CHD risk factors to 13: hypercholesterolemia (high cholesterol), hypertension (high blood pressure), smoking, physical inactivity, overweight, coronary behavior, stress, lack of social and psychological support, poor nutrition, diabetes, unfavorable heredity, inadequate rest, alcohol.

Thus, to assess the risk of coronary heart disease in a practically healthy person, 13 leading factors were selected according to the opinions of WHO experts.

In addition, the mobile application provides an analysis of additional risk factors: the physical condition of a person, a symptom of shortness of breath, and others.

Hierarchy of CHD risk factors.

The described risk factors play an unequal role in the occurrence of CHD. To streamline them, we use the "Delphi" method, which initially provides for the isolated presentation by experts of their opinions and their further adjustment based on the acquaintance of each expert with the opinions of other experts until the assessment of
the consistency of opinions of experts stabilizes.

This method differs from other methods of group interaction of experts in three features: anonymity, the use of the results of the previous round of the survey, and the statistical characteristics of the group response.

Anonymity is manifested in the fact that the members of the expert group do not interact with each other and, therefore, can change their personal opinion without losing their reputation. In this case, any assessment is accepted, regardless of who its author is. The condition of anonymity is necessary so that a supporting or contradicting opinion does not affect the judgment of an individual expert.

Using the results of the previous round has the value that each expert is informed about the average opinion of the group and about the opinion of the typical and extraordinary experts. This achieves the goal of presenting all points of view, allowing the expert to concentrate on the next round of the survey. The reasons for the change in assessments, as research results show, are actually in response to the point of view of their fellow experts.

The statistical characteristic of a group response is an assessment of the consistency of expert opinions in the form of a coefficient of agreement. The agreement coefficient can be reliable or not reliable and allows, depending on the result of the calculation, to increase or decrease the number of rounds of the survey.

The advantage of the "Delphi" method is the use of feedback during the survey, which significantly increases the objectivity of expert assessments.

To determine the hierarchy of CHD risk factors, the following method is proposed, which includes 6 stages:

Stage 1. Development of the "Questionnaire for the selection of priority risk factors for coronary heart disease";
Stage 2. Selection and questioning of experts. Evaluation of the consistency of expert opinions;
Stage 3. Construction of "Matrix of shares of alternative opinions of experts";
Stage 4. Standardization of "Matrix of shares of alternative opinions of experts";
Stage 5 Approximation of the "Matrix of shares of alternative opinions of experts";
Stage 6 Determination of the hierarchy of risk factors for coronary heart disease.

The final hierarchy of CHD risk factors according to the above method (Table 1).

<table>
<thead>
<tr>
<th>Code the factor</th>
<th>Risk factor for CHD</th>
<th>Hierarchy coefficient</th>
<th>Ranking place</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Hypertension</td>
<td>0.126</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>High-cholesterol</td>
<td>0.121</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>Diabets</td>
<td>0.113</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>Smoking</td>
<td>0.099</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>Heredity</td>
<td>0.097</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>Type A behavior</td>
<td>0.087</td>
<td>6 - 7</td>
</tr>
<tr>
<td>7</td>
<td>Stress</td>
<td>0.087</td>
<td>6 - 7</td>
</tr>
<tr>
<td>8</td>
<td>Body-weight</td>
<td>0.079</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 1 - Hierarchy of CHD risk factors
Figure 1 shows screenshots of CHD risk factors in the ES_RFCHD expert system.

![Figure 1 - Risk Factors for CHD](image)

### Table I: Risk Factor for CHD

<table>
<thead>
<tr>
<th>Code the factor</th>
<th>Risk factor for CHD</th>
<th>Hierarchy coefficient</th>
<th>Ranking place</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Hypodynamia</td>
<td>0.059</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>Mental-health</td>
<td>0.045</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>Poor-diet</td>
<td>0.041</td>
<td>11</td>
</tr>
<tr>
<td>12</td>
<td>Inadequate rest</td>
<td>0.031</td>
<td>12</td>
</tr>
<tr>
<td>13</td>
<td>Alcohol</td>
<td>0.015</td>
<td>13</td>
</tr>
</tbody>
</table>

The obtained values of the coefficients of the hierarchy became the basis for calculating the coefficients of certainty of knowledge facts and determining its filtering in a hierarchical functional system. The subject area classes with the level of their hierarchy are presented (Table 2).

Figure 2 shows a view of a hierarchical functional system (HFS) for diagnosing the risk of coronary heart disease in the KARKAS system in the knowledge base editing mode.

A functional system (FS) is a system formed to achieve a given useful result...
(target function) in the process of its functioning. Its backbone factor is a specific result [4].

Table 2 – Knowledge base classes

<table>
<thead>
<tr>
<th>Class</th>
<th>Number of instances of classes</th>
<th>Level of the class hierarchy</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHD risk</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Hypertension</td>
<td>17</td>
<td>2</td>
</tr>
<tr>
<td>High-cholesterol</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Diabets</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Smoking</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Heredity</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Type A behavior</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Stress</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Body-weight</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>Hypodynamia</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Mental-health</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Poor-diet</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Inadequate rest</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Alcohol</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

Figure 2 - Type of HFS for determining coronary heart disease (knowledge base editing mode)
HFS is characterized by the following properties:

- **connectivity** - a chain of knowledge base connections
- **complexity** - hierarchy of levels of local knowledge bases
- **stability** (adaptive behavior of the system) - the structure of the FS digraph does not change with vertical violations of the rules.

To build knowledge base filtering, it is enough to specify a chain of HFS rules to achieve the main goal. The number of local knowledge bases corresponds to the levels of the hierarchical FS.

The physical model of the knowledge base stores class instances, objects, object attribute values and logical relationships between classes, objects.

**Fine-tuning the knowledge base**

As a result of the conceptualization of the subject area, CHD risk factors have been identified, each of which is described by its own set of attributes.

Knowledge characterizing risk factors is presented in the form of clusters. A knowledge cluster is understood as a group of rules that allow displaying knowledge regarding the degree of risk of an individual for a particular CHD factor. The entire knowledge base is divided into 13 clusters. Each cluster has its own risk factor.

The division of knowledge into clusters and their filtering in the knowledge base made it possible to control the conduct of the consultation and greatly facilitated experiments with knowledge modifications.

Three types of attributes are used to interpret cluster knowledge:
- question;
- intermediate (temporary);
- representative (sub-target);
- target (for example, the target attribute could be the attribute "CHD risk").

Knowledge base clusters have a hierarchical structure. At the lower level of the cluster, there are rules that establish a correspondence between question and intermediate attributes and their values. The conditions of these rules are formed from question attributes, the values of which are determined by asking questions to the user.

The next level contains rules that establish relationships between sub-target attributes.

At the top level of the hierarchy, there are rules that set values for the target attribute.

Knowledge base facts can be formed as a result of answers to questions (stating facts), as well as when rules are followed (resulting facts). Knowledge base facts confidence coefficients are calculated for each cluster according to its own method, based on the cluster hierarchy coefficient.

In the ES_RFCHD system, the coefficients of certainty of the rules are put down by the expert before consultation. But the knowledge base of the system contains clusters in which the coefficients of certainty of the rules (c.c.r) are calculated when the user is asked. For this, the coefficients of confidence (c.c.) of ascertaining facts are used.

In the system, the coefficients of factor of the rules are calculated in various ways. Consider the bisection method:

$$\beta_1 = 0, \beta_i = (i - 1) I_k / ((s - 1) 100), (i = 2, \ldots, s - 1),$$
where $\beta_i$ is the value of the confidence coefficient of the i-th rule, $l_k$ is the value of the hierarchy coefficient of the k-th cluster, s is the number of cluster rules containing a representative attribute.

When $\beta_1 = 0$, the representative attribute in the rule takes the value "Not a risk" and c.c.r. is not affixed, in other cases - "Risk" with the corresponding value of c.c. equal to $\beta_i$ ($i = 2, ..., s - 1$).

The procedure for calculating the coefficients of certainty of the rules using the scoring method is that first the sum of points is calculated based on the user's answers to questions, and then the bisection method is applied.

As an example, consider the structure of cluster number 6 ("type A behavior"). This cluster contains three static rules and one dynamic rule. Let us show how the calculation of the confidence coefficients of static rules is performed.

According to the results of the psychological test, if the total score is 29 or less, then the behavior is type A and is classified as a risk of coronary heart disease. The sum of points from 30 to 41 characterizes an intermediate type of behavior (between A and B). If the total score is 42 or more, then the behavior is type B and is classified as not a risk of CHD.

This situation is described by the following rules, whose confidence coefficients are calculated using the bisection method. The coefficient of the cluster hierarchy with number 6 ("type A behavior") is 8.7%.

**Rule 37. A#.**
IF A Score coronary behavior [] 16 ... 29 THEN type A behavior = Risk, c.c. r. = 0.087.

**Rule 38. A#.**
IF A Score coronary behavior [] 30 ... 41 THEN type A behavior = Risk, c.c. r. = 0.086.

**Rule 39. A#.**
IF A Score coronary behavior [] 42 ... 80 THEN type A behavior = No risk, c.c. r. = 1.000.

For example, if as a result of answers to questions on assessing coronary behavior, the total score was 22, then rule 39 will be fulfilled and the resulting fact will be formed:

\[
\text{type A behavior} = \text{Risk, c.c.} = 0.087.
\]

So, with a sum of 22 points, coronary behavior corresponds to a risk of coronary heart disease, equal to 8.7%.

The knowledge of the cluster is described by sixteen question attributes that form a dynamic rule. The latter changes depending on the user's answers to questions. The dynamic rule serves here to determine the sum of the coronary behavior scores. It is represented by frame 36, which has sixteen slots, each of which is a daemon.

**Frame 36.**

<table>
<thead>
<tr>
<th>Slot name</th>
<th>Slot type</th>
<th>Inheritance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labor Day</td>
<td>Sum</td>
<td></td>
</tr>
<tr>
<td>Problems</td>
<td>Sum</td>
<td></td>
</tr>
</tbody>
</table>
Cases | Sum |
Food | Sum |
Thought | Sum |
Game | Sum |
Mobility | Sum |
Energy | Sum |
Employment | Sum |
Efficiency | Sum |
Sensitivity | Sum |
Haste | Sum |
Distraction | Sum |
Attitude to work | Sum |
Life | Sum |
Intensity | Sum |
Target slot.
Score coronary behavior | Define.

The final rule of the knowledge base dynamically changes during the consultation process and is intended to determine the integral assessment of the risk of coronary heart disease and before the consultation is presented as follows:

Frame 73.
Slot name | Slot type | Inheritance
Pressure | Replacement | n
Cholesterol | Replacement | n
Diabetes | Replacement | n
Hereditary | Replacement | n
stress | Replacement | n
Smoking | Replacement | n
Behavior | Replacement | n
Body weight | Replacement | n
Movement | Replacement | n
Support | Replacement | n
Nutrition | Replacement | n
Alcohol | Replacement | n
Recreation | Replacement | n
Target slot.
Risk of CHD | Calculate.

Figure 3 shows screenshots of the details of the user's consultation with the CHD of the ES_RFCHD expert system.

The frame contains the representative attributes of the clusters that have slot status. The target attribute "Risk of CHD" is a demon and serves to calculate the coefficient of confidence of the received fact. The value of c.c. is determined by the summation of c.c. facts that involve representative attributes.
Summary and conclusions.

The ES_RFCHD mobile expert system has been developed for the Android platform, which is available on Google Play.

A knowledge base was built to determine the risk factor for coronary heart disease.

The knowledge base contains rules and frames for representing knowledge. Facts and heuristics can be stored in a knowledge base. In addition to these, meta-knowledge is used to guide decision making. The knowledge base contains both static and dynamic rules that change their structure.

Frames have the following characteristics:

• inheritance of slot values;
• managing daemon attributes.

Inheritance avoids duplication of information. Grouping the knowledge base into clusters allows you to:

• perform flexible inference by activating clusters during consultation;
• test the knowledge base by clusters, that is, check for consistency and completeness of knowledge.
The knowledge base of the system is constantly being modified and updated with new rules. Further design of the system is aimed at developing top-priority preventive recommendations to reduce the risk of coronary artery disease.

References:
7. V. P. Burdaev, On one approach to the formation of the user interface with the expert system, Modern engineering and innovative technologies, Published by: Sergeieva & Co Karlsruhe, Germany, 21.1 (2022): 97–108.
15. Istiadi, Emma Budi Sulistiari, Rudy Joegijantoro, Anik Vega Vitianingsih, Affi Nizar Suksmawati, "Mamdani fuzzy expert system for online learning to diagnose..."


Анотація. У статті представлено розробку експертної системи підтримки прийняття рішень на платформі Android. Для створення бази знань було обрано медичну тематику на прикладі визначення ризику ішемічної хвороби серця у практично здорової людини. Виявлено основні та додаткові фактори ризику ІХС, що визначило ієрархічну структуру бази знань. Структура бази знань містить кластери знань, щоб користувач міг консультуватися з експертною системою в залежності від вибору класерів. Базу знань реалізовано за допомогою оболонки КАРКАС, яка містить інструменти для створення експертних систем: візуальний редактор бази знань, аналізатор бази знань, ієрархічну функціональну систему формування структури бази знань на основі правил і фреймів, різноманітні модулі. для здійснення логічних висновків.

Ключові слова: база знань, експертна система, мобільний додаток, платформи Android та Embarcadero.