Using of dialogue's clinical decision support system for treatment of lung cancer

Holovach Y Y

Ring Ukraine, 81 Antonovicha str., Kyiv 02000, Ukraine

Corresponding author e-mail: iurii.golovach@gmail.com

Abstract. The problem of construction of decision support systems today is one of the most advanced in the direction of medical software engineering. Systems for non-professional users have become extremely popular. But in the case when the life of a person may depend on the decisions, system developers should pay special attention to the issues of balanced decisions. The article is about the possibility of creating a system with pre-laid branches of the dialogue. The model of the dialogue's system to decision support for treatment of lung cancer is constructed and the principles of constructing its individual component modules are considered in detail. The results of the research were used in the construction of medical software. The developed system covered the entire diagnosis tree, which provided the procedure for interviewing patients and evaluating their responses. This approach to the development of the system has allowed its use not only by users with medical education, but also by patients. All patients need to do is follow the system's instructions for downloading files and entering (copying) test results.

1. Introduction

Indicators of the quality of diagnosing diseases and oncological diseases, as a special case, depend primarily on obtaining accurate data and information about the patient's health, as well as the interaction of the organism with the living environment, external technical and natural environment.
The doctor, when in contact with the patient, recognizes the images (visual, sound, speech, temperature, rhythm, etc.) on the basis of which he makes a diagnosis.

At such diagnosis the doctor finds out symptoms (signs) of a disease. Assessing the clinical manifestation of the symptom, determine its features such as severity, consistency (frequency of detection in this disease), resistance in different phases of the disease, variability, danger.

Implementation of information technology in medical diagnostics today is one of the most essential branches, because there is a great lack of specialists, and the list required for decision making, the parameters can reach more than a hundred.

This led to the creation of a wide range of medical information systems (MIS), which conditionally distinguish seven levels of generations/types (from automated medical record management systems to intelligent decision support systems in medicine) [1]. The success of treatment for patients with malignant tumors largely depends on the quality of the diagnosis.

Algorithm of detailed diagnostics and determination of the stage of the disease of lung cancer based on clinical and morphological classifications is presented in the [2]. The clinical classification (cTNM, clinical - "c") is based on data obtained from clinical, radiological, ultrasound, endoscopic, morphological and other appropriate examinations performed before treatment. A pathogistological

classification (pTNM) is based on clinical data from cTNM, which complement and refine as a result of surgical intervention and histological studies.

Grouping by stages is performed after evaluating T-, N-, M- and / or pT-, pN-, pM-categories. This information - TNM / pTNM is recorded in the medical documentation, fixed in the registry chancer. The clinical stage is the basis for choosing a treatment and assessing its efficacy. Pathogistological tumor study provides information to determine the prognosis of the disease.

Although this algorithm has been tested for several decades, some studies may still be out of date, but additional data is needed to improve the reliability of automatic diagnosis and treatment appointment.

2. Analyse of medical software

2.1. Analyze of medical software

More than 10 different software systems were considered, among which the following were highlighted (more detailed analysis is presented in the work [1,2]):

1) Oncology. Magazine PAS, Herzen;
2) Medicine Live. Doctor's Assistant (communication, analyzes, International Classification of Diseases (ICD), drugs);
3) Doctor's Guide: ICE, Registry of Medicines;
4) Doctor's Assistant;
5) First Aid.

The comparative analysis of software applications was decided to be conducted according to the following parameters which, according to experts, are most relevant in the diagnosis of malignant neoplasms: medical calculators, a reference book, PubMed, the International Classifier of Diseases, the TNM Handbook, medical news, TLMs, the possibility Attaching files with analysis results.

The analysis of these applications is summarized in Table. 1, in which the numbers in the columns correspond to the order of examination of the software applications, which is presented above, +/- is the presence / absence of the corresponding parameter.

Table 1. Results of comparative analysis of software applications.

Parameter	Application				
	1	2	3	4	5
Medical Calculators	-	+	+	+	-
Handbook of analyzes	-	+	+	+	+
PubMed	+	+	+	-	-
International Classification of Diseases	-	+	+	-	-
TNM Directory	+	-	+	-	-
Medical News	-	+	+	+	-
Medicines Register	-	+	+	-	-
Attaching files with analysis results	-	-	-	-	+

As you can see, almost all systems have drawbacks and do not provide sufficient list of necessary tools.

The new system, which is developed in cooperation with experts from the National Cancer Institute, takes into account these disadvantages, and for the greater distribution of the system, the possibility of multi-level usage of program system is provided not only by professionals, but also by ordinary citizens with the possibility of remote monitoring of treatment stages for patients (this function is necessary not only for the doctor, but also for close relatives).

2.2. Formalization of oncological disease's diagnostic problem

The human body is a complex, hierarchical, multi-level, self-organizing system, consisting of interconnected subsystems of various levels of subordination, which interact with the environment during the exchange of matter, energy and information. The optimal functioning of each subsystem of the body is determined by its final beneficial effect (FBE), the parameters of which are perceived by the receptor subsystems and transmitted to the central nervous system (CNS).

Moreover, at each *k*-th level of interaction between the internal and external environment, each *i*-th subsystem of the organism R_i (cardiovascular, respiration, nervous, immune, etc.) is in a certain *j*-th state s_{ij}^k of the multitude

$$S_i = \{s_{ij}^k\}; \quad k = \overline{0,14}; \; i = \overline{1,n_k}; \; j = \overline{1,n_i}$$

where n_k is the number of body subsystems at the *k*-th level of interaction with the environment;

n_i is the number of possible states of the *i*-th subsystem at the *k*-th level.

The set of states S_i is conventionally divided into subsets: S_{in} – norm, S_{ig} – borderline state and S_{ip} – pathology, that is:

$$S_i = S_{in} \cup S_{ig} \cup S_{ip}$$

The set of states of the whole organism *S* is determined by the set of states of all its subsystems

$$S = \{S_i\}, \; i = \overline{1,n_r}$$

where $n_r = \max_k n_k$

The system of diagnoses accepted in medicine D = {Di}, $i = \overline{1,n_d}$ is a reflection of the set of possible states of the body on the system of terms and definitions adopted at this stage of the development of medicine, stipulated by the accepted classifications and nomenclature of diseases $i = \overline{1,n_d}$. Thus, each diagnosis is a subset of the body conditions $i = \overline{1,n_d}$, and each state of the organism can reflect several levels of interaction of its subsystems. Diagnosis D_0 - "practically healthy" is defined as $i = \overline{1,n_d}$, and many other diagnoses $i = \overline{1,n_d}$.

With the development of medicine, the set of diagnoses expands by detailing some diagnoses and taking into account both lower levels of interaction (cellular, biomolecular, and below), and higher (level of the subtle etheric body and above). In this formulation, the problem of diagnostics using a computer medical diagnostic system is reduced to the problem of determining whether the current state of the organism or its individual subsystem belongs to one of the formalized states from the set of diagnoses $\{D_i\}$.

The initial data for any computer medical diagnostic system in the diagnosis of the i-th subsystem of the body is the system of diagnostic signs adopted in medical practice *X*.

$$X = \{x_0, \ldots, x_i, \ldots, x_m\},$$

which, to some extent, reflect the current j-th state of the i-th subsystem of the organism at the k-th level of interaction s_{ij}^k, that is, the set of states { s_{ij}^k } is mapped to many signs X.

Thus, the unknown current state of the body's subsystems s_{ij}^k is mapped both to the system of diagnoses $\{D_i\}$, and to the set of signs, while the task of the DSS is to determine the dependence X → $\{D_j\}$.

To this formalized representation of the body's subsystems, two remarks should be noted:

1) The human body constantly interacts with the factors of the external environment, and as a result of the interaction of the internal and external environment throughout the entire period of life (including the intrauterine period), taking into account the genetic development programs, the current state of all subsystems of the body is formed, for the diagnosis of which the developed in the work is proposed a formalized approach when examining a patient in normal clinical conditions. However, some environmental factors, such as stressful conditions, taking medications, narcotic substances, etc., not only form the state of the body's subsystems during prolonged exposure, but also at the time of

their impact lead to a jump-like transition to another energy level, which is characterized by a different a set of states.

2) With a specific implementation of the DSSM (Decision Support System in Medicine), not one, but several pathologies (and, accordingly, several borderline states) can be diagnosed. In this case, the diagnosed subsystem Ri has internal levels of the hierarchy and the subsets Sig and Sip are divided into smaller subsets according to the number of diagnosed pathologies (diseases).

Today, the problem of tumors in the human body (especially malignant) is one of the most advanced not only in the medical field, but also in the technical. The urgency of this problem is confirmed by the data of the National Cancer Registry of the Institute of Oncology of the Academy of Medical Sciences of Ukraine, according to which in our country up to 140 thousand people get cancer annually [6]. Of particular concern is the growing proportion of patients of working age.

The success of treatment of patients with malignant neoplasms largely depends on the quality of diagnosis. It is difficult to detect a tumor in the early stages, because it has many manifestations that are characteristic of inflammatory or chronic diseases. According to researchers, today only 41.1% of patients can detect a tumor in the first and second stages of the disease. The high percentage of neglect and mortality in the first year of pathology detection (over 41%) gives reason to think seriously about the quality of cancer diagnosis. Insufficient education of patients complicates the timely detection and, as a result, successful treatment of cancer.

Histogenetic classification of tumors. Malignant tumors are classified based on their origin ccording to the tissue from which they form:

1) cancer (carcinoma, malignant epithelioma) - is derived from epithelial tissue (has ecto- or endodermal origin);

2) sarcoma - comes from the connective tissue that occurs from the intermediate layer(mesoderm);

3) carcinosarcomas - originate from epithelial and connective tissue at the same time;

4) hemoblastosis (leukemia, malignant lymphoma) - a tumor of hematopoietic tissue;

5) aphids (carcinoids, paragangliomas, small cell lung cancer, medullary thyroid cancer, thymos, pheochromocytomas, chemodecomas) - originate from neuroendocrine cells;

6) tumors of the endothelium and mesothelium - tumors from blood and lymphatic vessels, synovial and serous membranes;

7) tumors of the nervous tissue;

8) tumors from embryonic remains;

9) trophoblastic tumors;

10) mixed tumors.

3. Types of source data and analysis of their processing features

The quality indicators of the medical system depend primarily on obtaining accurate data and information about the patient's health, as well as the interaction of the organism with the environment, external technical and natural environment [3].

The doctor in practice performs pattern recognition (visual, sound, taste, speech, temperature, rhythmic and arrhythmic, symmetrical and asymmetrical, psychological, genetic and others) in order to identify diagnostic signs (symptoms). Symptoms of the disease are the clinical criteria for the pathological state of the body, which are defined as statistically significant deviations of the body's vital signs from its normal values, the absence of natural fluctuations of indicators, the absence of certain signs or the appearance of some new phenomena or conditions not inherent in a healthy body.

In terms of diagnostic significance from the point of view of pathogenetic interpretation, non-specific (manifested in many diseases), specific (manifested in this class of diseases) and pathognomonic (characteristic strictly for one disease and no other).

Subjective symptoms include patient complaints about the pathological sensations of the body or the function of individual organs. Pathological changes and abnormalities that are manifested and actually recorded during a physical or laboratory-clinical examination of a patient are considered objective. Clearly consider the symptoms that can be detected by your own organs of sensation clinician. Hidden symptoms manifest during clinical, laboratory or instrumental studies.

In medicine, there are thousands of symptoms that allow the diagnosis of syndromes and diseases. The division of clinical medicine into areas (cardiology, pulmonology, gastroenterology, endocrinology, rheumatology, etc.) significantly reduces the lists of symptoms, but the list of new paraclinical methods (laboratory, biochemical, electro-and magneto-physiological, ultrasound, phonographic, etc.) working with signals, images, their parameters and combinations, is steadily growing. By metrological assessment, diagnostic signs differ [4]:

1) Quantitative or numerical signs are measured in a certain scale and are expressed by numbers with a certain accuracy of measurement (results of instrumental studies and signs obtained as a result of processing biosignals and images).

2) Qualitative, rank or ball are used to express medical terms and concepts that do not have numerical values (for example, the severity of the condition) and are measured on a scale of order.

3) Binary or dichotomous, take only two values ("0" or "1", "YES" or "NO") and are used to fix the presence or absence of some symptom in formalized documents.

4) Classification or nominal (for example, gender, profession, blood type) are signs measured in the scale of names.

For each of the considered types of attributes, their own research methods are used, although many algorithms developed for one type adapt to other types [4].

4. Dialogue's system to decision support

The system must be able to perform complex actions, for example, it must be able to choose a sequence of scenarios where each element of the script will be executed only under certain conditions. We will call such complex actions strategies.

Training of such systems is allowed by generating random sequences of actions and memorizing "useful" ones.

Consider the implementation of such strategies. Each strategy can be represented by an association. In this case, some associations are transferred to the interaction mode (IM), while others are transferred to the exposure mode (EM). If the system needs to first perform an x SF_1 action, and then SF_2, then the $F_1 | \Rightarrow F_2$ association is used, operating in IM.

The execution of such a sequence is provided by selecting the governing vertices $\eta_{12} = \left[F_1 | \Rightarrow F_2\right]$.

Consider a more difficult case (Figure 1), where SF_2 will start after SF_1 only with input Sx_1, than allocated control vertices β:

$$\beta \overline{\overline{Df}} \left[x_1 \Rightarrow \left[F_1 | \Rightarrow F_2\right]\right].$$

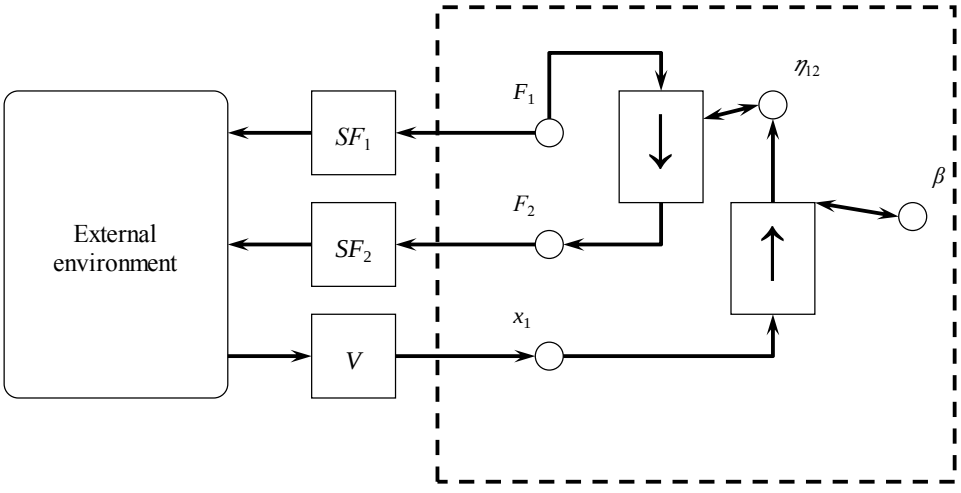

Figure 1. Implementation of dialog association.

The expediency of comparative representation of expert knowledge is due not only to the natural hierarchy of objects, but also the necessarity to take into account the parameters of the state as knowledge accumulation on the object. In addition, usage of the hierarchy principle allows you to simplify the rules and significantly reduce their numbe [6, 7].

The program realization of output tree presented on Figure 2.

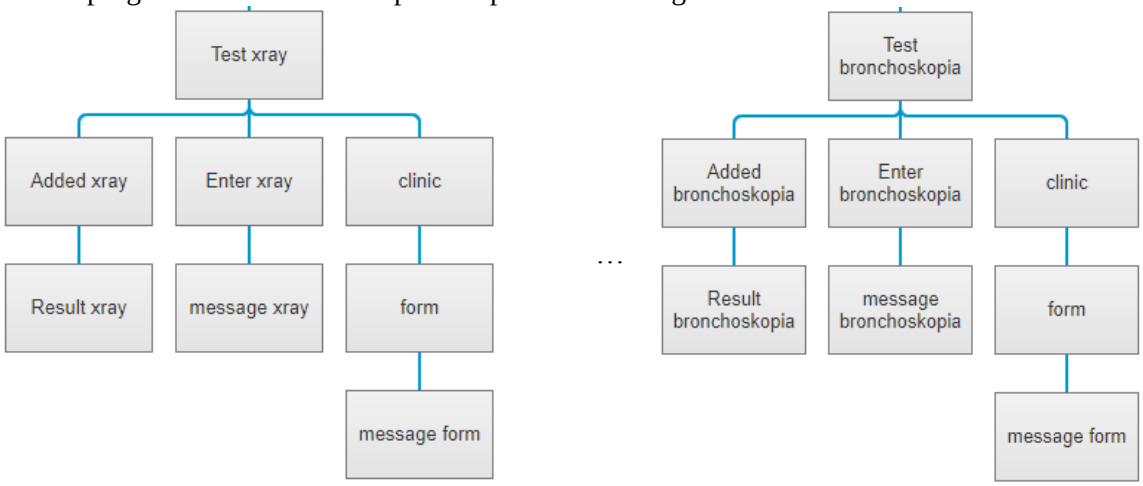

Figure 2. Some forms of application that display the necessary diagnostic procedures and regulate data entry to the system.

Conclusions

As a result of solving the tasks set in the a, the appropriateness of the chosen path in the organization of the work of the program support system for decision-making in the treatment of lungs' cancer was shown. On the basis of the presented data it is envisaged to obtain a set of new scientific results that will allow to consistently substantiate the methods of determining the necessary diagnosis in patients' examination and effective treatment.

Based on the review of existing software solutions for decision support in the diagnosis of lungs' cancer, the feasibility of introducing a new software system that will use modern approaches to work with all groups of users of the system has been proven. Construction of this system is not possible in the development of a mathematical model that describes the relationship of treatment methods and necessary diagnostic procedures with patient parameters. The complexity of obtaining a mathematical model is due to its dimensionality, multicriteria and uncertainty in a number of states, and today it is solved through the usage of expert opinions or a priori protocols for diagnosis and treatment.

Each of the approaches goes through practically the same stages, only in a different order (the number of stages and the depth of their disclosure also depends on the chosen methodology of designing the hardware and software systems - from the classic waterfall model to the new-fashioned agile technologies).

With a top-down approach, more attention is paid to the idea of global use of the system, with a bottom-up approach to individual functions and available resources. At the same time, each of the approaches determines the importance of individual tasks at each stage, up to their possible total depreciation.

References

[1] Golovach Y 2018 Formalization of the diagnostic problem of cancer in automated systems *"AVIATION IN THE XXI-st CENTURY" – "Safety in Aviation and Space Technologies"* (Kyiv: NAU) http://conference.nau.edu.ua/index.php/Congress/Congress2018/paper/view/5619

[2] Artamonov Y B and Golovach Y Y 2018 Pidhodi do organizaciï roboti programnogo kompleksu pidtrimki priinyattya rishen' pri likuvanni raku legeniv *Bulletin of the Engineering Academy of Ukraine*. #1 (Kyiv: NAU) pp 128-134 (in Ukranian)

[3] Vesnenko A I, Popov A A and Pronenko M I 2002 Topo-tipologiya strukturyi rozvernutogo klinicheskogo diagnoza v sovremennyih meditsinskih informatsionnyih sistemah i tehnologiyah *Kibernetika i sistemnyiy analiz* #6 (Kyiv: V M Glushkov Cybernetics Institute) pp 143-154

[4] Tosun A B, Kandemir M, Sokmensuer C and Gunduz-Demir C 2009 Object-oriented texture analysis for the unsupervised segmentation of biopsy images for cancer detection *Pattern Recognition* Volume 42 Issue 6 pp 1104-1112 https://doi.org/10.1016/j.patcog.2008.07.007

[5] 2018 Bulletin of National Cancer Registry of Ukraine 2018 URL: http://www.ncru.inf.ua/publications/BULL_19/index_e.htm

[6] Artamonov E B and Zholdakov O O 2010 Concept of creating a software environment for automated text manipulation *Scientific journal "Proceedings of the National Aviation University"* (Kyiv: NAU) Vol 3 (44) pp 111-115

[7] Artamonov Y B and Panf'orov O V (2016) Formuvannya adaptivnih dinamichnih scenariïv u komp'yuternih sistemah navchannya *Tehnologichnii audit ta rezervi virobnictva* (Kharkiv: NTU «KhPI») vol. № 6/1(32). pp. 66-71 DOI: 10.15587/2312-8372.2016.86342 (in Ukranian)