Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 35 (2014) 1426 – 1435

Procedia Computer Science

18th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems - KES2014

Self-learning bayesian networks in diagnosis

Petr Suchánek[a,*], Franciszek Marecki[b], Robert Bucki[c]

[a]Silesian University in Opava, School of Business Administration in Karviná, Univerzitní náměstí 1934/3, 73340 Karviná, Czech Republic
[b] *Wyższa Szkoła Biznesu w Dąbrowie Górniczej, ul. Cieplaka 1C, 41-300 Dąbrowa Górnicza, Poland*
[c] *Institute of Management and Information Technology, ul. Legionów 81, 43-300 Bielsko-Biała, Poland*

Abstract

The article presents the main bases of artificial intelligence, probabilistic diagnostic methods, development of the diagnostic database and diagnostic base of knowledge and Bayesian networks as a base of the diagnostic self-learning systems which are commonly used in medicine to recognize diseases on the basis of symptoms. Probabilistic models of diagnostic networks are based on the Bayesian formulas. These formulas let us determine probabilities of causes on the basis of probabilities of results. This is the reason why databases must be created and adequate probabilities determined. Results of this research are then analyzed by means of statistical methods.

Keywords: artificial intelligence; self-learning Bayesian networks; medical diagnostic; diagnostic systems; databases; simulation

1. Introduction

Intelligent information systems are more and more commonly used in diagnostics[1,2,3,4] , and others. The article presents the self-learning Bayesian networks implemented in medical diagnostics. The model of a network for medical diagnostics is shown however, it can be implemented in other disciplines as well. Probabilistic models of diagnostic networks are based on the Bayesian formulas. These formulas let us determine probabilities of causes on the basis of probabilities of results. This is the reason why databases must be created (e.g. records of patients, etc.) and adequate probabilities determined (e.g. diseases). Methods of probabilistic diagnostic can be used in medicine, psychology, economy, technology and other disciplines in which questioning, questionnaires or other experimental research is used[5,6]. Results of this research are then analyzed by means of statistical methods. Intelligent diagnostic

* Petr Suchánek. Tel.: +420-596398322; fax: +420-596312069.
E-mail address: suchanek@opf.slu.cz

doi:10.1016/j.procs.2014.08.200

methods require creating information systems which include: databases, bases of knowledge and concluding functions. Modification of a database (by updating data) and knowledge (by updating probabilities) results in a diagnostic system which becomes self-learning. This kind of learning consists of modifying parameters of a diagnostic system while changing the data stream of currently diagnosed people. Such systems can be implemented for diagnosing in other disciplines e.g. business, finances, stock markets, etc. Diagnostics is commonly used in medicine to recognize diseases on the basis of symptoms[7]. Before implementing the right therapy (treatment methods, medicaments, etc.) patient's disease must be determined. The diagnostic model is probabilistic because diseases are forecasted only with a certain probability (without certainty). It results from the specifics of the human body. Let us consider two ill persons as an example. They can have the same symptoms and different diseases or different symptoms and the same diseases. Therefore, knowledge has a statistical character in probabilistic diagnostic systems. So, extended databases (symptoms and diseases) and resulting from them bases of knowledge (probabilities at determined symptoms) are needed. Diagnostic networks are created during using a diagnostic procedure. This procedure consists of carrying out subsequent examinations (no all at once) as long as the reliable diagnosis of a disease is obtained. Moreover, at determined symptoms certain diseases are excluded (with high certainty). Medical examinations are usually time occupying and expensive (they require expensive equipment) which results in not carrying out all examinations without a reason. On the other hand, it is important to detect some diseases as soon as possible (e.g. cancer, etc.). Therefore, it is necessary to justify the need to carry out these time occupying and costly examinations. It is of a vital importance for increasing effectiveness of functioning the health system in each country. At the beginning an ill patient usually complains about certain symptoms occurring in him (or a doctor detects them). On this basis, certain diseases are diagnosed. If this diagnosis is not selective (i.e. showing with high certainty of a certain disease), a doctor recommends carrying out further examinations (e.g. blood analysis, X-ray, etc.). After obtaining these additional data, a more precise diagnosis is formulated. However, if even this diagnosis does not prove selective, the doctor recommends carrying out further examinations (CT, etc.). Such a procedure is carried out as long as the most reliable diagnosis is reached (defining a disease with high certainty).

2. Artificial intelligence

In a general case, diagnostics is one of practical methods of artificial intelligence. An expert (a doctor, a sociologist, an economist, an engineer, etc.) possesses natural intelligence (resulting from his life experience). However, the problem of passing this knowledge to the ones who replace him when he retires remains. So, diagnostic information systems based on professional experience should be available. Professional experience is the database which lets us conclude properly: What would happen if? Experience can be formalized and stored in information databases. Artificial intelligence is interpreted as a result of functioning specific information programs[8,9,10] (i.e. computer-based, micro-controller, mobile, satellite – GPS, etc.). Such programs work rationally (not optimally) in the way a man equipped with natural intelligence does. In many cases the computer simulation method allows us to answer quickly (without the need to be an expert) the question: What would happen if? However, in such a case, an adequate simulation model is of a vital importance. Creating an adequate model (i.e. useful for diagnostic purposes) is difficult. Therefore, scientific works are carried out in the field of: brain/mind modeling, stock as well financial market modeling[11], etc. Diagnostics requires specific knowledge. It is commonly assumed that the diagnosis is prepared by experts specializing in a given discipline. However, the diagnosis process carried out by experts is time occupying and costly. Therefore, computer-based diagnostic systems based on simulation and artificial intelligence replacing experts are built[12]. Artificial intelligence is associated with the human brain. In particular, to solve diagnostic problems artificial neuron networks are implemented[13]. Nevertheless, diagnosing depends on available information. For example, the use of computer tomography in medicine improves the effectiveness of diagnosing considerably. In a general case, supporting the human sense (e.g. seeing, hearing, smelling, etc.) lets us diagnose different phenomena e.g. diseases. Intelligence not only consists in joining information in one's brain but also getting required information. The widely understood diagnostic problem also concerns possibilities of performing certain complexes of operations (manufacturing, transport, service, etc.). This branch of artificial intelligence concerns the so-called agent systems in which there is a decisive agent as well as performing agents. The problem of agent systems is in the center of artificial intelligence interest now. The basic agent system has a two-level structure. There is a decisive agent at the higher level (a dispatcher, a coordinator,

a manager, etc.). Performing agents are at the lower level. There can be various logic relations between performing agents however, they are rather independent. The key feature of the agent system is data/information division between a decisive agent and performers. A managing agent gives orders to perform a task without defining parameters (e.g. time). Performers have to recognize (diagnose) the environment in which they perform tasks. Diagnostic systems suggest how to conclude about the state of certain objects. Moreover, this state should be verified. In the case of a good or bad diagnostic decision the database or the base of knowledge of the diagnostic system are corrected however, diagnostic conclusions change as a result. Therefore, it is assumed that diagnostic systems can learn i.e. adjust their parameters (e.g. probabilities). The need for self-learning diagnostic systems is obvious (e.g. financial markets). Moreover, development of medicine and pharmacology leads to the need for data updating about patients (diseases and symptoms).

3. Probabilistic diagnostic methods

Self-learning probabilistic diagnostic systems which can be implemented in medicine are presented hereby. Self-learning probabilistic diagnostic systems are based on the Bayesian formulas. This general approach towards creating self-learning diagnostic systems can also be used in other disciplines (e.g. in informatics to detect spam, while servicing cars to detect damaged sets, in banking to verify persons applying for a credit, etc. In a general case a diagnostic system consists of: the database, the knowledge base and the concluding generator. This is the reason why these elements of a probabilistic diagnostic system are discussed in detail. A probabilistic diagnostic method based on the Bayesian formulas is presented below. Let us assume there are N events B_n which exclude each other and are mutually independent. It is assumed that these events create the total system which means that probabilities of these events meet the condition $p(B_1)+,\ldots,+p(B_m)+,\ldots,+p(B_m) =1$. The event B_n is understood as occurrence of the n-th disease. It is assumed that diseases exclude each other and are mutually independent. At the same time it is assumed that $p(A_n)>0$, $n = 1,\ldots,N$. Let us assume that there are M independent events A_m, $m = 1,\ldots,M$ which do not exclude each other. These events are understood as occurring symptoms of a disease. Further, let us assume that probabilities of these events are given $p(A_m)>0$, m = $1,\ldots,M$. The diagnostic model uses the name "patient's profile" on the basis of symptoms of a disease. The patient's profile defines the conjunctions A of events A_m, $m = 1,\ldots,M$ where the probability of this conjunction is determined as (1):

$$p(A) = \prod_{m=1}^{M} p(A_m) \tag{1}$$

Events B_n and A are mutually dependent because the n-th disease comes into being by defined symptoms A_m, $m = 1,\ldots,M$. The known Bayesian formula is implemented for the conjunction of dependent events (2):

$$p(B_n \wedge A) = p(B_n) \cdot p(A / B_n) = p(A) \cdot p(B_n / A) \tag{2}$$

The conditional probability $p(B_n / A)$ is obtained from formula (2) (3):

$$p(B_n / A) = \frac{p(B_n) \cdot p(A / B_n)}{p(A)} \quad n = 1,\ldots, N \tag{3}$$

As it assumed that events A_m are not dependent so it can be written that (4):

$$p(A / B_n) = \prod_{m=1}^{M} p(A_m / B_n) \tag{4}$$

Substituting formulas (1) and (4) into (3) we obtain finally the following formula (5):

$$p(B_n / A) = \frac{p(B_n) \cdot \prod_{m=1}^{M} p(A_m / B_n)}{\prod_{m=1}^{M} p(A_m)} \tag{5}$$

The probability of the n-th disease (events B_n) can be determined from the Bayesian formula (5) on condition that the conjunctions of A symptoms comes into being if the following probabilities are known: a) absolute: $p(A_m)$ - occurring of the m-th disease, $m=1,\ldots,M$; b) absolute: $p(B_n)$ - occurring of the n-th symptom, $n=1,\ldots,N$; c) relative: $p(A_m / B_n)$ - occurring of the m-th disease, $m=1,\ldots,M$ at the n-th symptom, $n=1,\ldots,N$. The program consists of three loops. However, the following are required: a) the vectors of probabilities: $p(A_m)$, $m = 1,\ldots,M$ and $p(B_n)$, $n = 1,\ldots,N$; b) the matrix of probabilities: $P(A_m / B_n)$, $m = 1,\ldots,M$, $n = 1,\ldots,N$. The above probabilities form the base of knowledge which is determined with the use of the database.

4. The diagnostic database

Let us consider the basic case of medical diagnostic, without taking into account temporarily, the Bayesian nets. Let us also assume that a doctor has the patients' records in which there are K patients forming the database. With the pass of time the number K grows in practice. However, at the beginning let us consider the database for the constant number of patients K. Such a base is static in comparison with the dynamic database in which data change. In a specific case of the temporary database data can be subject to change in time. At the same time, there can be the constant number of records K e.g. updated ones (using e.g. the FIFO procedure). Such bases are used commonly in practice (e.g. to monitor share prices at the stock exchange, etc.). Also, the usage of temporary bases is justified in diagnostic databases because of development of medicine, changes of life conditions in the environment, etc. Let us assume that each patient P_k, $k = 1,\ldots,K$ was examined by the doctor in order to diagnose a disease to be treated. The doctor diagnoses a disease of subsequent patients on this basis. An experienced doctor (equipped with natural intelligence) can diagnose patients' disease precisely on the basis of symptoms. A patient should be treated properly in accordance with the diagnosed disease. The effect of such a diagnose and treatment process of patients is saved in the database as a result: a) diagnosed symptoms A_m, $m = 1,\ldots,M$; b) a diagnosed disease B_n which should be treated in a patient. From the formal point of view a database can be presented in the form of two tables (symptoms and illnesses) in the way presented further. Let us assume there is a record of patients (the database) where we can distinguish the first table consisting of M fields (columns) connected with symptoms marked by: $a_1,\ldots,a_m,\ldots,a_M$. So, from the mathematical/logical point of view, the occurrence of the symptom a_m is an event which can have different values. In a general case a symptom can: occur; not occur; not be examined; not be sure (only probable). In an analogical way the database can contain a table consisting of N fields (columns) connected with diseases which are marked as: $b_1,\ldots,b_n,\ldots,b_N$. It is assumed that a patient undergoes a treatment process which is connected with a certain disease detected in a diagnostic process. There are false diagnoses in practice but finally the patient is subject to the proper treatment process. So, from the mathematical/logical point of view, the occurrence of the disease b_n (which was treated) in a patient is an event with the value 1. Let us assume that the database (both tables with symptoms and diseases) has K records where the k-th record includes patient's data P_k, $k = 1,\ldots,K$. Data in the database refers to symptoms:

1) the binary matrix of symptoms $A=[a_{k,m}]$, $k = 1,\ldots,K$, $m = 1,\ldots,M$, where: $a_{k,m} = +1$ if the k-th patient has the m-th symptom; $a_{k,m} = -1$ if the k-th patient does not have the m-th symptom; $a_{k,m} = 0$ if the m-th symptom was not examined in the k-th patient. There is the value 1 in the k-th row of the matrix A, $k = 1,\ldots,K$ (6):

$$0 < \sum_{m=1}^{M} a(+1)_{k,m} \le M \tag{6}$$

At the same time let the element of the matrix A which equals 1 be marked $a(+1)$. It is assumed that at least one symptom was identified in each patient: $a_{k,m} = 1$ which means the patient was ill. Analogously, the sum of values 1 in the m-th column of the matrix A meets the requirements of (7):

$$0 \le \sum_{k=1}^{K} a(+1)_{k,m} \le K \tag{7}$$

as the same symptom can be seen in many patients.

2) the binary matrix of illnesses in the form: $B = [b_{k,n}]$, $k = 1,\ldots,K$, $n = 1,\ldots,N$, where: $b_{k,n} = 1$ if the k-th patient has the n-th illness; $b_{k,n} = 0$ if the k-th patient does not have the n-th disease. There is only one value 1 in the k-th row of the matrix B, $k = 1,\ldots,K$ (8):

$$\sum_{n=1}^{N} b_{k,m} = 1 \tag{8}$$

as it is assumed each patient has only one disease. The sum of values 1 in the n-th column of the matrix B meets the requirements of (9):

$$0 \le \sum_{k=1}^{K} b_{k,m} \le K - N \tag{9}$$

as there occurred N different diseases among K patients.

Analogously, the analysis of symptoms which were examined but did not occur in a patient can be carried out i.e. $a_{k,m} = -1$. It is important to emphasize that if a certain symptom does not occur, it lets us eliminate certain diseases. In a specific case, a patient should be healthy if they have symptoms of a disease. The probability of each disease should equal zero in the discussed diagnostic system. Simultaneously, the lack of information about occurrence of a symptom does not let us diagnose on the basis of this symptom. The base of knowledge can be created on the basis of the database defined in this way (i.e. probabilities of occurrence of symptoms and diseases). Such a database is required to diagnose diseases with the use of registered symptoms in case of the next patient (numbered as $K + 1$).

5. The diagnostic base of knowledge

An expert should have specialist knowledge. This knowledge results from data/information stored in databases. Data processing can lead to obtaining knowledge about an analyzed object or process. Diagnostic databases of knowledge are formed from the diagnostic databases. Diagnostic databases as well as diagnostic databases of knowledge are characterized by the type of information they consist of. Diagnostic databases include information about causes and effects of events concerning each object and process. On the other hand, diagnostic databases of knowledge include statistical information. In turn, the diagnostic knowledge base contains statistical information about the causes and consequences of a set of objects and processes. In the case of medical diagnosis, statistical information is treated as a probability. In a general case, a patient has a few symptoms and one disease (by definition). Thus, in the further analysis we assume that the symptoms are not mutually exclusive. If in a selected patient there are some symptoms at the same time, they form a conjunction from the mathematical/logical point of view. Symptoms which are not mutually exclusive can be either dependent or independent. For example, a patient's headache can be caused by temperature. However, sometimes there are cases of patients with a headache without temperature and temperature without a headache. For a particular patient it is difficult to determine whether the symptoms are dependent or independent. In case of acceptance (for simplicity) the fact that the symptoms are independent a diagnostic model becomes less complicated. In a general case, the patient falls ill with one disease, even though there are cases of simultaneous diseases. In the further analysis we assume that at the given data a patient suffers from one disease. There are characteristic symptoms for a specific disease. This information can be obtained from: from doctors, specialized medical books or the Internet. This gives a physician an opportunity to bring an accurate diagnosis, i.e. indication of a disease on the basis of identified symptoms. In general, however, there is a certain symptom occurring in various diseases (e.g. temperature, headache, etc.). For this reason, your doctor may have a problem with identifying a disease having the knowledge about given symptoms. This problem

results from the fact that human organisms differ and therefore: a) the same symptoms - patients may have various diseases; b) a variety of symptoms - patients may have the same disease.

For these reasons, medical diagnostics defines the probability of occurrence of a certain disease. These probabilities are derived from knowledge of the statistical analysis of data (contained in the files of symptoms and diseases). Typically, this is analysis for many patients, stretched in time, throughout several years of medical practice. Typically, a physician is responsible for a patient database (of symptoms and diseases) but not a knowledge base (i.e. the corresponding probabilities). The doctor's aggregated medical knowledge (in his natural memory) forms the so-called doctor's natural intelligence. The problem lies in the fact that an experienced doctor should pass his knowledge (professional experience) to a young doctor who has no experience (for example he has just graduated from a university). Otherwise, a young doctor will be gaining his professional experience (natural intelligence) for several years. In this case, diagnosis processes will last longer and will be more expensive. Let us assume that the database (matrixes A and B) was sorted out according to the numbers of patients' diseases. This enables us to determine the numbers of patients K_n who were diagnosed with the n-th disease, $n = 1,...,N$ $\sum_{n=1}^{N} K_n = K$

and at the same time the absolute probability can be calculated: $P(b_n)$, $n = 1,\ldots,N$ which is understood as the probability that the n-th disease will be detected (no matter what kind of symptoms occur) (10):

$$P(b_n) = \frac{K_n}{K} \quad n = 1,...,N \tag{10}$$

Analogously, it is possible to calculate the absolute probability $P(a_m)$, $m = 1,\ldots,M$ of the m-th symptom occurrence (no matter what kind of disease it is) (11):

$$P(a_m) = \frac{\sum_{k=1}^{K} a(+1)_{k,m}}{K} \tag{11}$$

so it is enough to sum up column elements which equal 1 in the matrix A in accordance with all rows for k, $k = 1,\ldots,K$. The absolute probabilities $P(a_m)$ and $P(b_n)$ form the statistical knowledge about patients from a certain environment (throughout the time of creating the database). This statistical knowledge is general intelligence which enables us to diagnose diseases effectively. Detailed intelligence requires determining conditional probabilities. Conditional probabilities $P(a_m / b_n)$, indicating that a patient with the disease b_n, $n = 1,\ldots,N$ had the symptom a_m, $m = 1,\ldots,M$, can be calculated with the use of the database, i.e. $a_m = +1$. Analogously, it is possible to determine the probability $Q(a_m / b_n)$ indicating that a patient with the disease b_n, $n = 1,\ldots,N$ did not have the symptom a_m, $m = 1,\ldots,M$, i.e. $a_m = -1$. Symptoms which were not examined have no value for diagnostics. As the database is sorted according to diseases $K_1,\ldots,K_n,\ldots,K_N$, elements $a_{k,m}$ are to be summed according to indexes k beginning with L_{n-1} to L_n where $L_0 = 0$ and $L_n = \sum_{i=1}^{i=K_n} K_i$. The relative probabilities $P(a_m / b_n)$ are calculated from the formulas (12 – a,b,c):

- for the disease b_1 (12-a)
- for the disease b_n (12-b)
- for the disease b_N (12-c)

$$\begin{array}{lll}
P(a_1 / b_1) = \frac{\sum_{k=1}^{k=K_1} a_{1,k}}{K_1} & P(a_1 / b_n) = \frac{\sum_{k=L_{n-1}}^{k=L_n} a_{1,k}}{K_n} & P(a_1 / b_N) = \frac{\sum_{k=1+L_{N-1}}^{k=L_N} a_{1,k}}{K_N} \\
\ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots \\
P(a_m / b_1) = \frac{\sum_{k=1}^{k=K_1} a_{m,k}}{K_1} & P(a_m / b_n) = \frac{\sum_{k=L_{n-1}}^{k=L_n} a_{m,k}}{K_n} & P(a_m / b_N) = \frac{\sum_{k=L_{N-1}}^{k=L_N} a_{m,k}}{K_N} \\
\ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots \\
P(a_M / b_1) = \frac{\sum_{k=1}^{k=K_1} a_{M,k}}{K_1} & P(a_M / b_n) = \frac{\sum_{k=L_{n-1}}^{k=L_n} a_{M,k}}{K_n} & P(a_M / b_N) = \frac{\sum_{k=L_{N-1}}^{k=L_N} a_{M,k}}{K_N}
\end{array} \tag{12}$$

The matrix of conditional probabilities $P(a_m / b_n)$ provides a detailed statistical knowledge base of a diagnostic system. It is possible to diagnose the disease effectively on the basis of absolute and conditional probabilities. Analogously, one can determine the knowledge base for the symptoms which were not found during testing.

6. Diagnosis

Let us assume that the profile X of the subsequent patient is given in the form of the vector of symptoms x_m, $x = 1,\ldots,M$. On the basis of this profile (examined symptoms) it is necessary to determine a disease of a patient. As shown in this paper the diagnostic system lets us conclude about the probabilities of patient's diseases: $b_1,\ldots,b_n,\ldots,b_N$ on the basis of examined symptoms $x_1,\ldots,x_m,\ldots,x_M$. The presented diagnostic model takes into account the following: $x_m = +1$ if the examined patient has the symptom a_m; $x_m = -1$ otherwise. Let x_i be symptoms which were examined and were not detected in a certain patient and x_i be symptoms which were examined and were detected in a certain patient. On the basis of data x_i it is possible to exclude the disease b_n for this certain patient if the condition below is met: $P(a_i / b_n) = 0$. The above condition indicates that none of K patients whose data are in the database did not fall ill with the disease b_n if the symptom a_i was not detected in them. Conclusion: A patient did not fall ill with the disease b_n because there is no detected symptom a_i in him. In this way, with the use of the data: $x_m = -1$ and the condition $P(a_i / b_n) = 0$ it is possible to eliminate diseases b_n. In the optimistic case, after such elimination there remains only one disease. However, if the result of such elimination leads to leaving more than one disease, the Bayesian formulas are used to determine the most probable disease. Because the symptoms and diseases are dependent, therefore the probability of dependent events must be determined from the Bayes Theorem $p(b_n \wedge Y) = p(b_n) \cdot p(Y / b_n) = p(Y) \cdot p(b_n / Y)$, where: Y – is a conjunction of symptoms that occurred in the examined patient, i.e. $x_m = -1$; b_n – are the diseases which have not been eliminated (in the previous calculation step). Thus, we receive from the above (13):

$$P(b_n / Y) = \frac{P(b_n) \cdot P(Y / b_n)}{P(Y)} \tag{13}$$

The left side of this equation is the probability of the disease b_n in the case of symptoms constituting the conjunction Y (i.e. the conjunction of symptoms for which $x_m = -1$) and this probability must be determined.

The absolute probability $P(b_n)$ is given on the right side of the above equation. It is stored in the knowledge database. Because it was assumed that symptoms are independent events, so the probability $P(Y)$ is calculated as the product of (14 – a,b):

- absolute probabilities $P(a_j)$, $j = 1,\ldots,J$ for the symptoms which occurred in a tested patient (i.e. $x_m = -1$) (14-a)
- similarly, conditional probabilities are calculated (14-b)

$$P(Y) = \prod_{j=1}^{j=J} P(a_j) \qquad P(Y / b_n) = \prod_{j=1}^{j=J} P(a_j / b_n) \tag{14}$$

The calculations are repeated for all diseases b_n which have not been eliminated by the previous procedure. The maximum value of the probability $P(b_n / Y)$ determines the most likely disease. If the value $P(b_n / Y)$ is much higher than for other diseases, it is assumed that the result of the diagnosis is the disease b_n. Otherwise, a doctor recommends a patient to perform additional tests.

7. Bayesian networks

In practice, the diagnostic system can use the Bayesian networks. Such networks are formed in the case where data about symptoms are not available at one moment. If, after the elimination of certain diseases more than one disease remain, a doctor may perform an additional study. As a result, the Bayesian formulas are implemented in

order to indicate one disease. From a theoretical point of view, the Bayesian networks are created then. The problem of diagnosis in medicine by means of the Bayesian method can be represented in the form of a network. Nodes of such a network are the states S which interpret data sets for studies of patients. Arcs (directed) mean the study of specific symptoms. Let us assume there are M symptoms in a certain field. Thus, a patient can be examined if he has the m-th symptom, $m = 1,\ldots,M$. In a general case there are 2^M states not considering the sequence of examination. Such a net has one initial state S^0 which represents a patient without examinations. Similarly, there is only one final state S^M in this network which represents a patient after all examinations of all M symptoms. Thus, in such a network stages can be distinguished (of carrying out the subsequent examinations): $e = 1,...,E$. From the theoretical point of $E = M$ however, in diagnostic practice $E \le M$. The number of states of the e-th stage equals 2^e, $e = 0,1,...,M$. In practice, in the medical diagnostics network queuing relationships are taken into account. It results from the logic sequence of testing. For example, examining the i-th symptom should precede the study of the j-th symptom. In practice there are also other reasons for the examining sequence constraints. For example, the subsequent j-th examination can be carried out only after the time τ_i or its cost equals c_i. Minimizing the time and cost of diagnosis, the examination sequence can be reduced. It is worth noting that the queuing restrictions are territorial (i.e. they may be different in different regions). In continuation, it is assumed that the states S in the discussed network include these queuing relationships. The state S^e, $e = 1,...,M$ interprets a collection of studies of certain patient's symptoms e. The result of each examination can be positive (i.e. a symptom occurs) or negative (i.e. a symptom does not occur). Therefore, the patient who had e specific examinations can be in 2^e states. In extreme cases: a) all examinations delivered a positive result; b) all examinations delivered a negative result. In this connection, it is necessary to distinguish the patient's state s^{ef} , $f = 1,...,2^e$ who had e examinations carried out on him and which delivered the f-th variant of results (positive or negative). To sum up, the Bayesian network consists of the states S^e for $e = 0,1,\ldots,M$; their number equals $L_e=2^e$ (one initial state and one final state). Each state S^e is a set of defined examinations on the basis of M symptoms. The study of symptoms can be positive or negative. Therefore, each state S^e is a set of states s^{ef} for $f = 1,...,2^e$ of the patient who had examinations carried out on him (depending on the obtained results). The number of states s^{ef} equals $L_{e,f} = 2^e$. Accordingly, the Bayesian diagnostic network has a very large size (the number of states). For this reason, the analysis (diagnosis) requires artificial intelligence methods (e.g. evolutionary algorithms). The problem of medical diagnosis consists in determining probabilities (at the given state s^{ef}): a) $P(b_n / X)$ of the fact that the patient has developed the n-th disease, $n = 1,\ldots, N_e$; b) $Q(b_n / X)$ of the fact that the patient does not have the n-th disease, $n = 1,\ldots,N_e$. If probabilities of the above are not selective (i.e. they do not indicate one disease) or do not exceed prescribed arbitrary limits g and d ($P > g$ or $Q > d$), it is necessary to carry out the next examination. The selection problem of the next j-th test can be solved by methods of artificial intelligence (such as an evolutionary algorithm). It is possible to simulate carrying out the j-th examination as well as the positive or negative results at the state s^{ef}. The choice of the j-th examination depends on the reduction degree of the diagnosis space. In particular, one disease should be indicated after the j-th study. If treatment of a diagnosed disease does not bring improvement of the patient's health state, the physician performs additional tests (or directs the patient to specialized studies such as lung X-ray, CT scan, blood analysis, etc.). On this basis, the first diagnosis may be affected and the patient can be treated for another disease by another professional.

8. The diagnostic self-learning system

The presented diagnostic system is static, i.e. the database does not change. In practice, the patient database changes over time as a result of medicine development, changes in living conditions, etc. Furthermore, the data in this database come from a selected community (geographically and temporally). These data may also be information obtained from doctors, from the Internet or medical publications. In the presented system, the database is emphasized on the basis of determined probabilities of a knowledge base. Each new patient provides their data about symptoms and a disease. These data can be introduced successively into the database, i.e. in a chronological order. As a consequence, the probabilities in the knowledge base under which the diagnosis is carried out, change. In particular, other diseases can be eliminated and another disease can be selected as the most probable. In view of these observations, we conclude that the diagnostic system should be self-learning. Data collected from each next patient should be introduced into the system. The database can contain the constant number of patients K. Therefore,

after introduction of a new patient data, the oldest data must be removed from the database (the FIFO principle). Self-learning diagnostic systems are intelligent systems. They are able to adapt to the randomly changing environment. Such systems can also be used in other fields such as psychology, economics, technology, etc.[14]

The self-learning system for medical diagnosis proposed in this paper implies that an interview with a patient about symptoms to infer about the disease is unavoidable. Such a system can also be made for the diagnosis of diseases on the basis of other data. For example, some diseases may be detected by means of blood analysis. Depending on the components (and their proportions) it is possible to determine the probability of certain diseases. However, if intervals of percentage shares of components in the blood are distinguished, the occurrence of certain events of shares is mutually exclusive (for the same component). For this reason, the Bayesian formulas change because the event A is an alternative, not a conjunction of events A_m. Modifying Bayesian formulas for such cases is not a significant problem.

A self-diagnostic system can be used in practical medical diagnosis. Revision of its effectiveness requires practical computer-based testing. In the information version it can be the system to be used in a single computer, the local network of computers or the Internet system. In case of a single computer or network, an interview with patients can be performed even by a nurse. Then a doctor, in a much shorter time, can make a diagnosis on the basis of calculated probabilities. In such a process a local knowledge base is created. Migration of patients between different health centers requires transferring data "to follow the patient". The ability to transfer data of a single patient leads to the concept of establishing a global (common to many health centers) knowledge base. In these databases it is essential to protect personal data but this kind of problem is not considered in this paper.

9. Conclusions

The probabilistic diagnostic system can also be used in other fields such as psychology, economics, repair processes, etc. Psychology similarly to medicine is a science based on experiments, interviews or questionnaires. For this reason, in the same way, psychological symptoms of patients can be observed and adequate therapies indicated. In economics (banking) solvency of natural or legal persons is diagnosed before granting the loan. To do this, the number of indicators (revenues, expenses, liabilities, etc.), which are symptoms, is checked. Moreover, repayment of the credit, its conversion, etc. is predicted. In practice, diagnostic systems are often used to predict the cause of device failure as repair usually involves costly disassembly. Hence, it is important to properly diagnose what went wrong in order to make sure that dismantling is needed. This procedure is commonly used in repair services. Diagnostic systems can also be used in selection of effective drugs or herbs. In this case, a disease can be recognized, and it is important to select drugs or herbs which prove to be effective. Thus, the diagnosis of therapy may also be considered. In this case, age, sex, physical characteristics of a patient (weight, height, etc.) are treated as symptoms. Therapies are associated with an implemented drug (e.g. due to price), the intensity of dosing the drug, a composition of various drugs, etc. In the case of an Internet system, answers: YES (1) or NO (0) may be given by the patient himself (e.g. whether he has a high temperature or a headache, etc.). In this way, an online self-learning diagnostic system can be treated as the so-called medical "zero contact". On the basis of an online diagnosis the patient can visit a doctor of the so-called "first contact" or a specialist. If a patient survey is available (via the Internet) for a doctor, he may introduce a modification to the database. Increasing access to an online self-learning diagnostic system will make the knowledge in the system more reliable and the health service more efficient.

Acknowledgements

The paper is one of the outcomes of the project “Innovation of Study Programs at Silesian University in Opava, School of Business Administration in Karviná” Nr. CZ.1.07/2.2.00/28.0017.

References

1. Cichocz P. *Systemy uczące się*. Warszawa : Wydawnictwa Naukowo-Techniczne; 2000.
2. Grewal A, Stephan DA. Diagnostics for personalized medicine: what will change in the era of large-scale genomics studies? *Personalized Medicine* 2013; 10(8): 835-848.

3. Hu XH, Cammann H, Meyer HA, Miller K, Jung K, Stephan C. Artificial neural networks and prostate cancer-tools for diagnosis and management. Nature Reviews Urology 2013; 10(3): 174-182.
4. Ifenthaler D, Pirnay-Dummer P, Seel NM. *Computer-Based Diagnostics and Systematic Analysis of Knowledge*. Springer; 2010.
5. Kasperski MJ. *Sztuczna inteligencja*. Gliwice: Wydawnictwo Helion; 2003.
6. Kychkin AV. Intelligent information and diagnostic system for examining blood vessels. *Journal of Computer and Systems Sciences International* 2013; 52(3): 439-448.
7. Marecki J, Marecki F. *Metody sztucznej inteligencji*. Bielsko-Biała: WSIZ; 2012.
8. Nguyen MN, Bao CY, Tew KL, Teddy SD, Li XL. Ensemble Based Real-Time Adaptive Classification System for Intelligent Sensing Machine Diagnostics. *IEEE Transactions on Reliability* 2012; 61(2): 303-313.
9. Nherera L, Marks D, Minhas R, Thorogood M, Humphries SE. Probabilistic cost-effectiveness analysis of cascade screening for familial hypercholesterolaemia using alternative diagnostic and identification strategies. *Heart* 2011; 97(14): 1175-1181.
10. Rezaei H, af Klint E, Kisten Y,van Vollenhoven RF. The Diagnostic Utility Of Musculoskeletal Ultrasound In Early Arthritis - a Probabilistic (Bayesian) Approach. *Arthritis and Rheumatism* 2013; 65(10): S835-S836.
11. Šperka R, Spišák M. Transaction Costs Influence on the Stability of Financial Market: Agent-based Simulation. *Business Economic Management* 2013; 14(1): s1-s12.
12. Warwick K. *Artificial Intelligence: The Basics*. London: Routledge; 2012.
13. Wei QF, Luo CS, Cao CZ, Guo Q. The intelligent diagnostic system of vegetable diseases based on a fuzzy neural network. *Mechatronics and Industrial Informatics* 2013, PTS 1-4, Book Series: Applied Mechanics and Materials 2013; p. 1907-1911.
14. Yang MT, Hu LS. Intelligent Fault Types Diagnostic System for Dissolved Gas Analysis of Oil-immersed Power Transformer. *IEEE Transactions on Dielectrics and Electrical Insulation* 2013; 20(6): 2317-2324.