Theme 8. Formal logic in solving the problems of diagnostics, prevention and treatment. Formalization and algorithmization of medical problems

The development of artificial intelligence methodology has been recognised as an important requirement in complex problem solving situations. Medical diagnosis is a particularly good example because of the complexity of the human mind and body and our limited and vague knowledge of how these function. This knowledge also varies with the expertise of the user. In a systematic approach to the acquisition of domain knowledge, the analysis of human physiology/psychology quickly produces large numbers of cause-effect relations at many interacting levels of both description and function. Necessarily, the relations are poor approximations of complex dynamic systems and some account has to be made for uncertainty at this level of description. Furthermore, the information available for searching this domain knowledge for a specific diagnosis is also usually vague (at least initially) in that evidence is indirect (reported symptoms or lack of them) and observations are incomplete and inaccurate due to the stochastic nature of mind/body psychology/physiology. The level of expertise of the clinician cannot be discounted in this process. Given that speed is also important in the diagnostic procedure, we should develop techniques in artificial intelligence that can support fast, reliable and accurate diagnosis with limited and vague information. It is important that experts are repeatedly able to classify the same situation in the same category.

Formalism is a description of something in formal mathematical or logical terms, the mathematical or logical structures of scientific arguments as distinguished from its subject manner. There is a formalism which expresses the idea of superposition.

Logic is a language for formalism and reasoning. It is a collection of rules we use when doing logical reasoning. Human reasoning has been observed over centuries from at least the times of the Greeks, and patterns appearing in reasoning have been extracted, abstracted, and streamlined.

In logic we are interested in true or false statements, and how the truth/falsehood of a statement can be determined from other statements. However, instead of dealing with individual specific statements, we are going to use symbols to represent arbitrary statements so that the results can be used in many similar but different cases. The formalization also promotes the clarity of thought and eliminates mistakes.

There are various types of logic such as logic of sentences (propositional logic), logic of objects (predicate logic), logic involving uncertainties, logic dealing with fuzziness, description logic.

Propositional logic is logic at the sentential level. The smallest unit we deal with in propositional logic is a sentence. We do not go inside individual sentences and analyze or discuss their meanings. We are going to be interested only in true or false sentences, and our major concern is whether or not the truth or falsehood of a certain sentence follows from those of a set of sentences, and if so, how.

<u>Statements, truth values and truth tables</u>

A ***statement*** is a declarative sentence that can be determined to be true or false. A statement differs from a sentence in that a sentence is only one formulation of a statement. There may be many sentences expressing the same statement. Examples of declarative sentences: “Patient has a disease”, “Doctor is right”. Such kinds of sentences are called ***propositions***. Examples of non-declarative sentences: “Look!” “What time is it?” They cannot be evaluated as true or false. Also "body tempreture is greater than 38^{o}C", where is body tempreture a variable representing a number, is not a proposition, because unless a specific value is given to body tempreture we cannot say whether it is true or false, nor do we know what body tempreture represents. If a proposition is true, then we say it has a ***truth value*** of "**true**"; if a proposition is false, its truth value is "**false**".For example, the statement “The patient with fever is ill” has truth value T.

Simple sentences which are true or false are basic propositions. Larger and more complex sentences are constructed from basic propositions by combining them with different

connectives. They are called ***compound*** statements. Thus propositions and connectives are the basic elements of propositional logic. Though there are many connectives, we are going to use the following **five basic connectives** here: ``**and'', ``or'', ``not'', ``if … then'' and `` … if and only if …''**. For example compound statements are: “If you finish your treatment then you can visit classes”, “I have read this topic and I understand the concept”.

In propositional logic, we often use letters, such as *p*, *q* and *r* to represent statements and the following symbols to represent the connectives (Table 8.1).

Table 8.1.

Connective	symbol	formal name
NOT	$\neg$	negation
AND	$\wedge$	Conjunction (logical multiplication)
OR	$\vee$	Disjunction (logical addition)
IF_THEN (or IMPLY)	$\rightarrow$	conditional
IF_AND_ONLY_IF	$\leftrightarrow$	biconditional

Note that the connective “or”in logic is used in the ***inclusive*** sense (not the exclusive sense as in English). Thus, the logical statement ”It is raining or the sun is shining” means it is raining, or the sun is shining or it is raining and the sun is shining.

Often we want to discuss properties/relations common to all propositions. In such a case rather than stating them for each individual proposition we use variables representing an arbitrary proposition and state properties/relations in terms of those variables. Those variables are called a **propositional variable**. Propositional variables are also considered a proposition and called **a proposition** since they represent a proposition. Hence they behave the same way as propositions. A proposition in general contains a number of variables. For example (P $\vee$ Q) contains variables P and Q each of which represents an arbitrary proposition. Thus a proposition takes different values depending on the values of the constituent variables. This relationship of the value of a proposition and those of its constituent variables can be represented by a table. It tabulates the value of a proposition for all possible values of its variables and it is called *a* ***truth table***.

First it is informally shown how complex propositions are constructed from simple ones. Then a more general way of constructing propositions is given. In everyday life we often combine propositions to form more complex propositions without paying much attention to them. For example combining "Grass is green", and "The sun is red" we say something like "Grass is green and the sun is red", "If the sun is red, grass is green", "The sun is red and the grass is not green" etc. Here "Grass is green", and "The sun is red" are propositions, and form them using connectives "and", "if... then ..." and "not" a little more complex propositions are formed. These new propositions can in turn be combined with other propositions to construct more complex propositions. They then can be combined to form even more complex propositions. This process of obtaining more and more complex propositions can be described more generally as follows:

Let X and Y represent arbitrary propositions. Then [$\neg$ X], [X $\wedge$Y], [X $\vee$Y], [X $\rightarrow$ Y], and [X $\leftrightarrow$Y] are propositions.

If *p* is the statement ``The wall is red'' and *q* is the statement ``The lamp is on'', then $p \vee q$ is the statement ``The wall is red or the lamp is on (or both)'' whereas $q \rightarrow p$is the statement ``If the lamp is on then the wall is red''. The statement $\neg p \wedge q$translates to ``The wall isn't red and the lamp is on''.

Statements given symbolically have easy translations into English but it should be noted that there are several ways to write a statement in English. For example, with the examples above, the statement $p \rightarrow q$directly translates as ``If the wall is red then the lamp is on''. It can also be stated as ``The wall is red only if the lamp is on'' or ``The lamp is on if the wall is red''. Similarly, $p \wedge \neg q$directly translates as ``The wall is red and the lamp is not on'' but it would be preferable to say ``The wall is red but the the lamp is off''.

p	$\neg p$
T	F
F	T

The truth value of a compound statement is determined from the truth values of its simple components under certain rules. For example, if *p* is a true statement then the truth value of $\neg p$ is F. Similarly, if *p* has truth value F, then the statement $\neg p$ has truth value T. These rules are summarized in the following truth table**.**

If *p* and *q* are statements, then the truth value of the statement $p \vee q$ is T except when both *p* and *q* have truth value F. The truth value of $p \wedge q$ is F except if both *p* and *q* are true. These and the truth values for the other connectives appear in the truth tables below.

p	q	$p \vee q$	$p \wedge q$	$p \to q$	$p \leftrightarrow q$
T	T	T	T	T	T
T	F	T	F	F	F
F	T	T	F	T	F
F	F	F	F	T	T

From these elementary truth tables, we can determine the truth value of more complicated statments. For example, what is the truth value of $p \wedge \neg q$ given that *p* and *q* are true? In this case, $\neg q$ has truth value F and from the second line of the tables above, we see the truth value of the compound statement is F. Had it been the case that *p* was false and *q* true, then again $\neg q$ would be false and from the fourth row of the above table we see that $p \wedge \neg q$ is a false statement. To consider all the possible truth values, we construct a truth table.

p	q	$p \wedge \neg q$
T	T	t F f
T	F	t T t
F	T	f F f
F	F	f F t

The lower case t and f were used to record truth values in intermediate steps. Note that while a truth table involving statements *p* and *q* has 4 rows to cover the possibility of each statement being true or false, if we have additional information about either statement this will reduce the number of rows in the truth table. If, for example, the statement *p* is known to be true, then in constructing the truth table of $\neg p \to q$ we will only have 2 rows. Truth tables involving *n* statements will have 2^n rows unless additional information about the truth values of some of these statements is known.

Logical reasoning is the process of drawing conclusions from premises using rules of inference. Here we are going to study reasoning with propositions. Later we are going to see reasoning with predicate logic, which allows us to reason about individual objects. However, inference rules of propositional logic are also applicable to predicate logic and reasoning with propositions is fundamental to reasoning with predicate logic. These inference rules are results of observations of human reasoning over centuries. Though there is nothing absolute about them, they have contributed significantly in the scientific and engineering progress that mankind have made. Today they are universally accepted as the rules of logical reasoning and they should be followed in our reasoning. Since inference rules are based on identities and implications, we are going to study them first. We start with three types of proposition which are used to define the meaning of "identity" and "implication".

Some propositions are always true regardless of the truth value of its component propositions. For example (P $\vee \neg$P) is always true regardless of the value of the proposition P. A proposition that is always true called a **tautology**. There are also propositions that are always false such as (P $\wedge \neg$P). Such a proposition is called a **contradiction**. A proposition that is neither a tautology nor a contradiction is called a **contingency**. For example (P $\vee$Q) is a contingency. These types of propositions play a crucial role in reasoning.

From the definitions (meaning) of connectives, a number of relations between propositions which are useful in reasoning can be derived.

We say that the statements *r* and *s* are *logically equivalent* if their truth tables are identical. For example the truth table of $\neg p \vee q$ is equivalent to $p \rightarrow q$. It is easily shown that the statements *r* and *s* are equivalent if and only if $r \leftrightarrow s$ is a tautology.

These identities are used in logical reasoning. In fact we use them in our daily life, often more than one at a time, without realizing it. If two propositions are logically equivalent, one can be substituted for the other in any proposition in which they occur without changing the logical value of the proposition. Below $\Leftrightarrow$ corresponds to $\leftrightarrow$ and it means that the equivalence is always true (a tautology), while $\leftrightarrow$ means the equivalence may be false in some cases, that is in general a contingency

The propositional logic is not powerful enough to represent all types of assertions that are used in computer science and mathematics, or to express certain types of relationship between propositions such as equivalence. For example, the assertion "x is greater than 1", where x is a variable, is not a proposition because you can not tell whether it is true or false unless you know the value of x. Thus the propositional logic can not deal with such sentences. However, such assertions appear quite often in our life and we want to do inferencing on those assertions. Also the pattern involved in the following logical equivalences can not be captured by the propositional logic: "Not all doctors are right" is equivalent to "Some doctors are not right ". "Not all drugs are expensive" is equivalent to "Some drugs are not expensive". Each of those propositions is treated independently of the others in propositional logic. For example, if P represents "Not all doctors are right " and Q represents "Some drugs are not expensive ", then there is no mechanism in propositional logic to find out whether or not P is equivalent to Q. Hence to be used in inferencing, each of these equivalences must be listed individually rather than dealing with a general formula that covers all these equivalences collectively and instantiating it as they become necessary, if only propositional logic is used. Thus we need more powerful logic to deal with these and other problems. The **predicate logic** is one of such logic. There is no other formalism which is simultaneously so well understood, so widely applicable, and so clear in its semantics. It must therefore play an important role in theorising about intelligent systems, and perhaps also in modelling them.

***Description logic** (DL)* is a family of formal knowledge representation languages. It is more expressive than propositional logic but has more efficient decision problems than first-order predicate logic. DL is used in artificial intelligence for formal reasoning on the concepts of an application domain. It is of particular importance in providing a logical formalism for ontologies and the Semantic Web. The most notable application outside information science is in bioinformatics where DL assists in the codification of medical knowledge.

Clinical judgment for the diagnosis and management of mans' diseases is an art. It can neither be acquired from textbooks alone, nor can it be taught, but has to be developed slowly through years of observation and experience. This is because unlike other professions, which thrive on calculations based on yes/no or present/absent, very little is clearly black and white in clinical medicine. Most clinical scenarios present in shades of gray. Instead of "present or absent", patients' symptoms are described using terms like "never, rarely, sometimes, often, most of the times, always, etc". Moreover, each specific symptom may also be graded as "mild, moderate or severe". This is compounded by the fact that most symptoms are experienced and described differently by patients and many symptoms may overlap in the same patient. Each individual patient may also have a multitude of characteristics other than the disease, rendering it unique in itself. Medical problems, therefore, cannot be generalized and analyzed using Aristotelian or binary logic, and an analytical program is desperately required which could integrate this complex network of problems and devise individualized solutions. Fuzzy logic is the nearest response to the call. It has the potential of combining human heuristics into computer-assisted decision making. Imagine combining the experience of five university professors with all the current literature and developing software that can calculate probabilities based on this, tailored specifically for each individual patient. Fuzzy logic can do all that. The concept was first introduced by Lotfi Zadeh. He defined fuzzy logic as "a class of objects with a continuum of grades of membership". It accounts for all the complexities and variations in patients and results in a statistical analysis which is appropriate for an "individual", unlike evidence-based medicine, which is applicable to a group of patients. It enables the scientific community to look into all shades of grey and determine the grade and severity of the disease. Fuzzy logic is a well-established concept in mathematics and engineering but its usefulness in medicine was not realized till the last decade. The use and applicability of fuzzy logic is accelerating at a significant pace in medical and scientific community.

Fuzzy logic is a multi-valued logic which was introduced by Zadeh in order to deal with vague and indecisive ideas. It has been described as an extension to the conventional Aristotelian and Boolean logic as it deals with "degrees of truth" rather than absolute values of "0 and 1" or "true/false". Fuzzy logic is not like computer software which understands only binary functions or concrete values like 1.5, 2.8, etc; instead, it is similar to human thinking and interpretation and gives meaning to expressions like "often", "smaller" and "higher". Fuzzy logic takes into account that real world is complex and there are uncertainties; everything cannot have absolute values and follow a linear function.

Characteristics of Fuzzy Logic. There are a few basic principles of fuzzy logic which were laid down by Zadeh in 1992:

- Exact reasoning is viewed as a limiting case of approximate reasoning.
- Everything is a matter of degree.
- Knowledge is interpreted as a collection of elastic, fuzzy constraints on a collection of variables.
- Inference is viewed as a process of propagation of elastic constraints.
- Any logical system can be "fuzzified".

Fuzzy Sets. A classical set of binary logic has "crisp" boundaries whereas fuzzy sets have fuzzy or imprecise boundaries. A fuzzy set consists of linguistic variables where values are words and not numerical. For example, intracranial pressure (ICP) can be defined as low, normal or high. Thus, ICP is a linguistic variable where the values have fuzzy margins and can overlap each other (Fig.8.1). The transition from one value to another is gradual and each value is given a membership function which represents the degree to which it belongs to that value.

Membership functions overlap each other as evident in (Fig. 8.1). Thus, a value for ICP can be both low and normal to a certain degree. Membership functions are not equivalent to probabilities. A membership value of low ICP does not signify that there is a certain probability of having low ICP or not; instead it is the degree to which it is a low ICP.

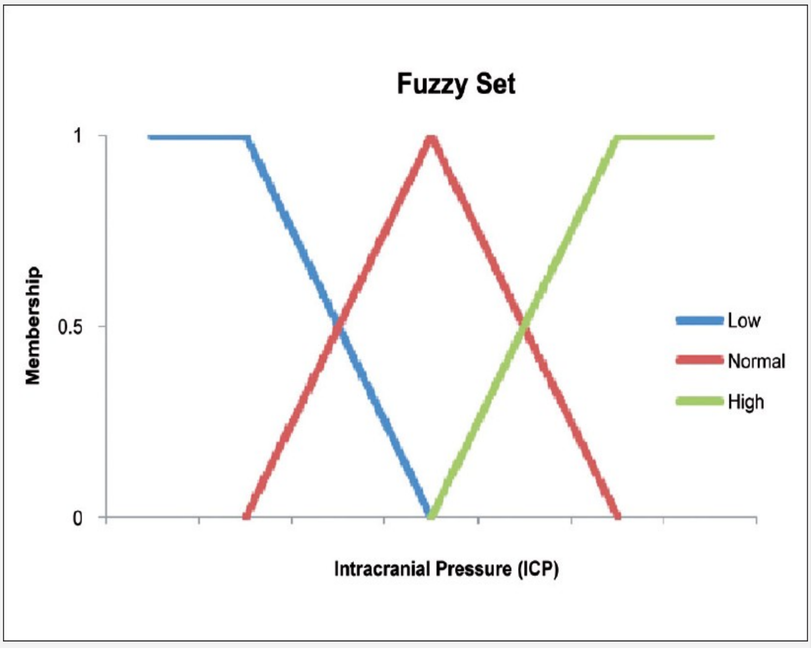

Fig. 8.1: Fuzzy sets: low, medium and high

Fuzzy Rules. Fuzzy rule is based on "if…then" rule and connects the different input and output fuzzy variables. It can be expressed as: if x is A then y is B where *A* is the antecedent and *B* is the consequent. Fuzzy rules are similar to common sense rules as they resemble human thinking and are based on human experience. For example, in order to control ICP in a patient with traumatic brain injury, sedation is often required but needs to be carefully monitored. A simple rule can be, "If the ICP is high, increase propofol infusion", or "If the ICP is low, stop propofol infusion". These rules are based on collective experience of specialists in the field as well as available literature. Thus, as more fuzzy rules and sets are obtained from various sources, uncertainties are potentially reduced.

Fuzzy Reasoning . Fuzzy reasoning is also called approximate reasoning and is the process of drawing conclusions from fuzzy sets and fuzzy rules.

Fuzzy Inference System. Fuzzy inference system (FIS) is a framework which is based on fuzzy sets, fuzzy rules and fuzzy reasoning. It has four main components including fuzzifier, rule base, inference engine and defuzzifier (Fig. 8.2). The fuzzifier creates fuzzy sets from "crisp" values like a fuzzy set for ICP. It will be divided into "low, normal and high" and a fuzzy set for propofol infusion will be divided into "stop, decrease and increase". Next, the fuzzy rules are formed based on these two input fuzzy sets: "If the ICP is low, stop propofol infusion", "If the ICP is normal decrease propofol infusion" and "If the ICP is high, increase propofol infusion". The inference engine applies all the fuzzy rules on the fuzzy sets to determine the resultant fuzzy output. If a "crisp" output value is required, the process of defuzzification converts the fuzzy output into a "crisp" output value by determining the center of mass of the combined, overlapping membership functions.

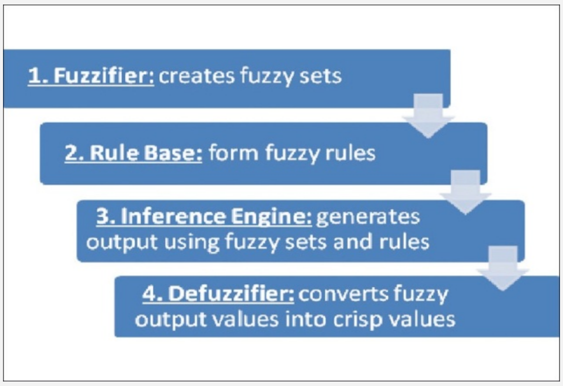

Figure 8.2. Components of fuzzy inference system

One of the most important uses of fuzzy logic is in drug delivery devices. Special fuzzy logic controllers have been designed for use in anesthesia and intensive care units. One of the studies used auditory evoked potential as a measure of depth of anesthesia and the fuzzy controller administered a certain amount of drug based on it. Similarly neuromuscular blocking agents are administered during surgery by monitoring the response at ulnar nerve; if the response at ulnar nerve is greater, more neuromuscular blocker is administered by the fuzzy controller.

Advantages and Disadvantages of Fuzzy Logic

Fuzzy logic is a solution complex problems in all fields of life, including medicine, as it resembles human reasoning and decision making. It looks into all shades of grey and answers uncertainties and ambiguities created by human language where everything cannot be described in precise and discrete terms. Fuzzy systems help define disease extent and severity and answer questions related to individual patients taking into account their risk factors and co-morbidities. On the other hand, it has a number of disadvantages too. It is tedious to develop fuzzy rules and membership functions and fuzzy outputs can be interpreted in a number of ways making analysis difficult. In addition, it requires lot of data and expertise to develop a fuzzy system. It does not give generalizable results and the program has to be run for each individual patient. Therefore, its clinical applicability and utilization is difficult without the availability of preprogrammed softwares for different pathologies and the basic training of clinicians to use these programs.

The application on fuzzy logic in medicine gained momentum in last two decades [when the usefulness of this technique was realized to correlate with the fuzzy nature of this field. In this era of technological advancements, most of the human workforce has been replaced by machines and robots and fuzzy logic is the means by which a system can be formulated, whereby machines can perform tasks by using rules similar to human reasoning and logic. All the disciplines of medicine have used fuzzy logic in some way. The various applications include: assessing the effectiveness of drugs, early detection of diabetic retinopathy and neuropathy, treatment of tropical diseases and diagnosis of diabetes, cardiac, renal and liver diseases where it has shown an accuracy of 79% in diagnosing diabetes and 97% in diagnosing dermatological diseases. In pulmonary medicine it has been used for chronic obstructive pulmonary disease, evaluating pulmonary function tests, as well as for ventilator support of patients in intensive care unit (ICU). In ICU setups, FIS has been shown to monitor patients, control blood pressure, provide adequate analgesia and anesthesia, assess ventilation requirements and even control ventilator settings, etc. FIS has also been shown to help in controlling tidal volume, and controlling the administration of neuromuscular blockade during surgery, as well as maintenance of the depth of anesthesia and optimum conditions for surgery. In oncology, detection of cancers, including lung, breast and prostate cancer, can be aided with fuzzy logic. It can also predict surgical outcomes and prognosis in malignancies and can be used to decide radiotherapy margins. It is also helpful in analyzing PET scan images for quantification of cancers. Fuzzy logic is not only applicable to clinical medicine, but it has been a useful statistical tool in basic sciences and bioinformatics as well. DNA sequencing, studying the complete genome and differences between polynucleotides and understanding various signaling pathways and cell signaling networks is possible by using fuzzy logic.

Formalization and algorithmization of medical problems

An algorithm is defined as "a step-by-step procedure for solving a problem or accomplishing some end especially by a computer". Algorithms provide a means to achieve specific goals. They are used extensively in both computer science and medical practice. To be of benefit they must be accurate, reliable, accessible and properly used.

We define a medical algorithm as any computation, formula, survey, or look-up table that is useful in healthcare. The purpose of a medical algorithm is to improve the delivery of medical care. A defining characteristic of algorithms is they are amenable to a programmatic representation. Data can be entered, processed according to formulas derived from the source material, and result in useful output.

The Medical Algorithm Project (MedAL), which stores peer-reviewed algorithms in an online database, contains more than 14,400. These tools can help physicians make diagnoses, choose treatments, calculate dosages, predict outcomes and monitor side effects. More are being developed every day.

Numerous algorithms are available in healthcare, and these may be used for diagnosis, management, monitoring or prognosis. Some *typical algorithms* in *Cardiology* include:

1. Calculation of the body mass index (BMI)
2. Calculation of anatomic areas and volumes from echocardiographic data
3. Application of the Framingham data for predicting the risk of cardiovascular disease
4. Preoperative risk assessment prior to cardiac or noncardiac surgery.
5. Identification of prognostic factors for short and long term prognosis following myocardial infarction
6. Evaluation of exercise test findings
7. Management of anticoagulation therapy in patients with atrial fibrillation
8. Determining the severity of disease and level of disability

Like their counterparts in mathematics, medical algorithms take myriad shapes. They can look like equations, scales, truth tables, checklists, scoring systems or decision trees. The simplest are performed with pen and paper, and the answers they provide may seem intuitive, something experienced physicians might come up with on their own, at least when dealing with familiar conditions. The widely used body mass index calculation, for instance, uses a straightforward ratio—mass in kilograms divided by the square of height in meters — to produce a number that physicians can use to see where a patient falls in a range from dangerously thin to morbidly obese.

Other clinical algorithms, however, are more complex and can help specialists keep up with a knowledge base that's expanding exponentially. These formulas are computerized and often sift huge amounts of data and alternative approaches before reaching conclusions. For example, the algorithms that drive automated external defibrillators analyze the pattern of a patient's heart rhythm to determine the number and strength of shocks required to restore normal functioning.

But simple or complicated, and despite their proliferation in textbooks, journals and, increasingly, electronic databases, most formal algorithms don't get used. Critics have written that algorithms lead physicians to interact with numbers, not patients, and has urged medicine to "give algorithms back to the mathematicians." But advocates argue that algorithms save time, money and lives — or would, if they were integrated into everyday practice.

Algorithms will have to be fully integrated into the everyday practice of health care. Every doctor would carry a smartphone or some other hand-held device that was directly connected to both an electronic medical records system and a carefully vetted database of algorithms. Then each time the doctor saw a patient, got back a lab result or had to deal with an adverse drug reaction, the device would query the database and choose the most relevant calculation. A doctor would only have to look down at the flashing device, then accept or override the result.

In almost all human activities we can discern three stages:

1) observation, 2) reasoning and 3) action.

In health care the same 3 stages can be seen in the so-called diagnostic-therapeutic cycle 1) Observation; 2) Diagnosis; 3) Therapy.

A patient tells his history, the clinician collects the data, comes to a conclusion and possibly even a diagnosis, and prescribes a therapy or carries out other treatment.

Diagnostics represents information processing in a system "doctor - patient", the purpose of which is building the most adequate model of a patient state.

The diagnostics process consists of three logically connected stages:

1) Clinicians collect the data about the state of the patient during a physical examinations; by laboratory tests or radiology and define the symptoms;

2) A stage of information processing, which includes: selection of the most essential symptoms, their comparison to a normal range according to gender, age, nationality, lifestyle indications etc.; compose patient's state symptoms into certain symptom-complexes (syndromes).

3) The diagnosis is made by comparison with the symptoms of other known diseases.

To make a diagnosis doctors have to follow certain rules, i.e. they are regulated by an algorithm. The diagnostic algorithm is a logic rules sequence; there the information about state of a patient is compared with a symptom–complex inherent to typical diseases.

Everyday medical practice contains many examples of probability. We often use words such as probably, unlikely, certainly, or almost certainly in all conversations with patients. We only rarely attach numbers to these terms, but computerized systems must use some numerical representation of likelihood in order to combine statements into conclusions. Probability is represented numerically by a number between 0 and 1. Statements with a probability of 0 are false. Statements with a probability of 1 are true. Most statements from real life fall somewhere in the middle. A probability of 0.5 or 50% are just as likely to be true as false. A round, opacified area seen in the lungs on a chest radiograph is probably pneumonia; one might assign a probability of 0.8, or 80%, (a 4 in 5 chance) to this statement. Based on the high probability of pneumonia, one might elect to treat this condition without performing further testing — a lung biopsy, perhaps — that would increase the probability of pneumonia to greater than 80%. We are accustomed to accepting the fact that our diagnoses has a certain probability of being wrong, so we counsel patients about what to do in the event (we might use the term "unlikely event") that things don't work out in the expected way.

Probabilities can be combined to yield new probabilities.

For example, the two statements: Pr(diabetes) = 0.6 and Pr(hypertension) = 0.3 mean that the probability of diabetes is 0.6, or 60%, (3 in 5 chance), and the probability of hypertension is 0.3, or 30%, (3 in 10 chance). We have not specified the clinical context of these statements, but suppose these probabilities applied to a particular population. Suppose further that the two conditions are independent; that is, the likelihood of patients having one disease is unaffected by whether they have the other (not always a safe assumption!). If we then want to know what the probability is of finding a patient in our specified population with both diseases, we simply multiply the two probabilities (0.6 and 0.3) to get 0.18, or 18%. If the two clinical conditions are not independent, (e.g., pulmonary emphysema and lung cancer) then we cannot combine the probabilities in such a simple, multiplicative manner. This is much like the AND function in Medline or the intersection function as applied to sets. The familiar "OR" function from our Medline program also has a mathematical meaning in combining probabilities. If we wanted to know how many patients in the above example had diabetes *or* hypertension (remember: this would also include those with both diseases in the usual mathematical sense of *or*), we would compute:

Pr (diabetes OR hypertension) = Pr(diabetes) + Pr(hypertension) – Pr (diabetes AND hypertension)

The last term in the above equation we already know to be 0.6 × 0.3 = 0.18,

so: **Pr(diabetes OR hypertension) = 0.6 + 0.3 - 0.18 = 0.72**.

Conditional probability is another type of probability often used in medicine. A conditional probability is the probability of an event (or the probability of the truth of a statement) *given the occurrence of another event* (or the truth of another statement).The most familiar case of conditional probability in medicine arises in the interpretation of diagnostic tests. For example, the probability of pneumonia given a round density on a chest radiograph is what we need to know in interpreting that diagnostic test if it is positive. In mathematical notation, this conditional probability is written this way: Pr (Pneumonia_Round Density on CXR). One reads this notation, "The probability of pneumonia given a round density on chest radiograph." This notation is convenient in the explanation of Bayes' rule, which is the cornerstone of the logic in many decision support systems.

Bayes' Rule

If we have a patient with jaundice, how likely is it that he has hepatitis? Written another way, we seek to learn: Pr(hepatitis_jaundice), which is read as "the probability of hepatitis given the presence of jaundice."

We may not have this probability at our fingertips, but we might be able to find a slightly different probability more easily: Pr (jaundice_hepatitis), which is, simply, the probability of jaundice given the presence of hepatitis. The latter probability could be found by studying a series of patients with proven hepatitis (it would be easy to get this data by looking up diagnosis codes in the medical records department) and computing the percentage of these patients who present with jaundice. However, this does not directly answer our original question. Bayes' rule allows us to compute the probability we *really* want—Pr(hepatitis_jaundice)—with the help of the more readily available number Pr(jaundice_hepatitis). Bayes' rule is simply this:

Pr (hepatitis_ jaundice) = Pr (hepatitis) × Pr (jaundice_hepatitis) / Pr (jaundice)

Notice that to solve this equation, we need not only Pr(jaundice_hepatitis), but Pr(hepatitis)—the probability of hepatitis independent of any given symptom — and Pr(jaundice) — the probability of jaundice independent of any particular disease. These two independent probabilities are called *prior probabilities*, since they are the probabilities prior to the consideration of other factors.

The derivation of Bayes' rule is very simple. We already know that the probability of any two events occurring simultaneously is simply the product of their individual probabilities. For example, the joint probability we already computed of diabetes and hypertension in a hypothetical population was:

Pr (diabetes AND hypertension) = Pr(diabetes)×Pr(hypertension) = 0.6 × 0.3 = 0.18.

We were free to multiply these together, because in our hypothetical population, the likelihood of one disease occurring in an individual was independent of the other. In other words:

Pr (hypertension) = Pr (hypertension_diabetes) and

Pr (diabetes) = Pr (diabetes_hypertension).

In this population, one's chance of having one disease is unaffected by the presence of the other disease.

In medicine, we are often faced with the question of the likelihood of two interrelated events occurring simultaneously in a patient. The case of a diagnostic test and the disease it is supposed to test for is a good example: what is the probability of an abnormal chest radiograph and pneumonia occurring in the same patient simultaneously? The question asks for this probability: **Pr (pneumonia AND abnormal CXR).**

Can't we simply find out what the incidence of pneumonia in the population is, and multiply it by the incidence of abnormal chest radiographs in the population? A moment's reflection should show that this simple calculation is not sufficient. For example, if the incidence of pneumonia is 1 in 1000, and the incidence of abnormal chest radiograph is 1 in 100, then the erroneous probability would be computed:

WRONG: Pr (pneumonia AND abnormal CXR) =0.001×0.01= 0.00001 = 0.001%

This does not fit with our clinical intuition very well, since we know that people with pneumonia tend to have abnormal chest films. Our intuition says that the probability of the two events occurring together should be pretty close to the probability of having pneumonia alone, since a majority of those patients will have abnormal chest films. What we *really* need to compute is this:

Pr (pneumonia AND abnormal CXR) = Pr(pneumonia) × Pr(abnormal CXR_pneumonia).

This is the probability of pneumonia multiplied by the probability of an abnormal chest radiograph given that pneumonia exists. If we take Pr(abnormal CXR_pneumonia) to be 90%, then the computation matches our intuition much better.

In general, for any two events A and B:

Pr(A AND B) = Pr(A) × Pr(B_A) and **Pr(B AND A) = Pr(B) × Pr(A_B).**

But since Pr (A AND B) must surely equal Pr(B AND A), we can say that the right-hand sides of the equations above are equal to each other:

Pr(A) × Pr(B_A) = Pr(B) × Pr(A_B)

Rearranging this equation, we have Bayes' rule:

Pr (A_B) = Pr (A) × Pr (B_A)/ Pr (B)

At an intuitive level, we use Bayes' rule when making seat-of-the-pants estimates of disease probability in patients. For example, if we designate hepatitis by A and jaundice by B, and there were an on-going epidemic of hepatitis (i.e., Pr (A) was high), then our index of suspicion for hepatitis in a jaundiced person would be increased. Likewise, if the likelihood of jaundice due to other causes was high (i.e., Pr(B) was high), then our estimation of the probability of hepatitis as a specific diagnosis would be lowered. Similarly, if jaundice were pathognomonic of hepatitis (i.e., Pr(A_B) was 1 or near to it), then our hepatitis diagnosis would be greatly increased. By using numerical estimates of the probability of diseases, findings, and conditional probabilities, Bayes' rule can help make medical decisions.

Algorithm of probabilistic diagnostic system

Use an algorithm to calculate a probability of disease.

The patient is characterized by set of symptoms, which are revealed on examination.

1) **Sample of probabilities of all symptoms for the likely diseases should be taken.** Copy probability for each patient's symptom from the initial table. For example, if symptoms, observed in the patient, correspond to table lines number 2, 7, 9, the set of symptoms consists of three symptoms (S2, S7 and S9). Probability S2 for the disease1 – PrS2_D1; Probability S2 for the disease2 – PrS2_D2; Probability S2 for the disease3 – PrS2_D3;

If there are 3 diseases (D1, D2, D3), three groups of numbers should appear:

	Disease1	Disease2	Disease3
Probability S2	Pr S2_D1	Pr S2_D2	Pr S2_D3
Probability S7	Pr S7_D1	Pr S7_D2	Pr S7_D3
Probability S9	Pr S9_D1	Pr S9_D2	Pr S9_D3

If there are a lot of symptoms and many likely diagnoses, as it happens in practice, it is impossible to accomplish this sample stage without using the computer facilities.

2) **Calculation of conditional probability of a symptom–complex for disease CPr_D.** The set of symptoms is called a symptom-complex of the given patient and designated as SC (symptom-complex) with probability CPr (conditional probability). Conditional probability (or symptom-complex probability) is a probability of having sets of symptoms in the patient the same time. Calculate CP according to the formula from the theory of probabilities. The conditional probability of a symptom-complex represents a product of probabilities of symptoms of the given symptom-complex in the given diagnosis. For example, for a symptom-complex out of 3 symptoms for a diagnose1: **CPr_D1 = Pr S2_D1 ×Pr S7_D1 ×Pr S9_D1** The number of obtained conditional probabilities is equal to the number of the diagnoses, considered in the system, (i.e. number of table columns).

3) Definition of prior probability of disease A

The prior probability of a some diagnosis is empirical frequency of observation of the given disease under some certain conditions. The prior probability is designated as A. It characterizes distribution of diseases in the given group of population. Such group can be a quota of the given hospital, given district, given city. It is referred to as the prior probability; it is already figured before deriving a symptom-complex. The value A is not a constant and depends on geographical, seasonal, epidemiological and other factors, which should be taken into account during making a diagnosis. For example, at a hospital 100 people were picked out at random, 70 of them had influenza. Therefore, the probability of influenza in all patients at the given hospital will amount to 70/100=0,7. When the influenza epidemic is liquidated, naturally the A for influenzas at this hospital will be different. The value of prior probability of the diagnosis is one of values, which requires monitoring during the operation of the diagnostic system.

4) Evaluation of standardized coefficient SC

Standardized coefficient represents complete probability of symptom-complex presence in all diseases. It represents the complete sum of paired products of conditional probabilities of a symptom-complex in the given diagnosis CP_D and prior probability of this diagnosis A :

SC = CPr_D1 × A1 + CPr_D2 × A2 + CPr_D3 × A3

The total number of items in the given sum is equal to the number of the diagnoses considered in the given system.

5) Evaluation of diagnosis probabilities at the given symptom-complex PrD (probability of diagnose)

This stage employs Bayes theorem (rule) (formula of hypothesis probability):

PrD1 = CPr_D1×A1/ SC

The number of probabilities of the diagnosis is equal to the number of the diagnoses of a system. Differently as a result of the given stage of work the system calculates probability of each of the available diagnoses.

6) Making a diagnosis

The stage is simple and is based on comparison of probabilities of diagnoses. The greatest value also indicates that diagnosis which is the most probable in the given symptom-complex.

In such a world, proponents say, algorithms would reach their potential to save time, money and lives because they would blend seamlessly into the background.

Control questions

1. What is an algorithm?
2. List advantages and disadvantages of using algorithms in everyday practice of health care.
3. Are there Limitations in using algorithms in medicine?
4. Explain the essence of conditional probability.
5. What is a prior probability?
6. When is Bayes' Rule used?
7. What is this standardized coefficient?
8. List the steps of a probabilistic diagnostic system.
9. What types of formal knowledge representation languages are used in medicine?
10. What is propositional logic?
11. When is used predicate logic?
12. List the basic principles of fuzzy logic
13. Explain the main components of Fuzzy inference system

References

1. Alan Rector, Jeremy Rogers Ontological Issues in using a Description Logic to Represent Medical Concepts: Experience from GALEN: Part 1 - Principles Available from: http://www.opengalen.org/download/Rector_Ontological_Issues_1.pdf Last accessed: 26 apr. 2012
2. Brown, J.S. and Burton, R.R. (1975). 'Multiple Representations of Knowledge for Tutorial Reasoning', in D.Bobrow and A.Collins (eds) Representation and Understanding, Academic Press.
3. Godil SS, Shamim MS, Enam SA, Qidwai U. Fuzzy logic: A "simple" solution for complexities in neurosciences?. Surg Neurol Int 2011;2:24 Available from: http://www.cs.bham.ac.uk/~axs/ Last accessed: 26 apr. 2012
4. M.A.Boden, Artificial Intelligence and Natural Man, Second edition 1986. MIT Press
5. Strengths and limitations of formal ontologies in the biomedical domain Available from: http://ontology.buffalo.edu/smith/articles/strengths_of_ontologies.pdf Last accessed: 26 apr. 2012
6. Patel, V. L., Arocha, J. F., & Zhang, J. Thinking and reasoning in medicine. In: Keith Holyoak: Cambridge Handbook of Thinking and Reasoning. Cambridge, UK: Cambridge University Press., 2004.