Analysis of mechanisms of using the concept of probabilities to model plausible reasoning in technical diagnostics

E S Litvinov¹, A V Daneev¹, D A Stepanenko², A S Stepanenko²

¹ Irkutsk State University of Railway Engineering, 15, Chernyshevskogo ave., Irkutsk, 664074, Russia
² East-Siberian Branch of the Russian Academy of Justice, 23A, Ivana Franko ave., Irkutsk, 664074, Russia

E-mail: njrdev@gmail.com

Abstract. Automated systems of support and decision making (ASDS), using knowledge bases, containing expert opinions of leading specialists in a certain type of expertise, can render great assistance in decision-making. Such intellectualized systems can be effective, in particular, in the work of a criminal investigator and conduct forensic examinations. A criminal expert is obliged not only to report his conclusion, he must justify it and at the same time, so that the investigator or the court could understand and assess whether this conclusion is correct. In order to address issues related to the improvement of forensic analysis, it is necessary to have, above all, appropriate information support. The following paper offered a number of specialized applied problems in the field of technical diagnostics, the received results can be used for intellectual support of substantiation and decision-making also concerning questions of investigation of separate kinds of crimes or carrying out of other kinds of forensic analyses.

1. Introduction

Intellectualized technical means using knowledge bases, content of expert opinions of leading specialists in a certain type of expertise can render great assistance in decision-making. Such intellectualized systems can be effective, in particular, in the work of a criminal investigator and conduct forensic examinations. At investigation of cases often there are questions for which the special knowledge is necessary.

Forensic examination is a study of materials (circumstances) aimed at the identification or resolution of special issues, which are not identified, the study of evidence, substances, or other diverse materials of the case (both criminal and civil). This research shall be conducted by a knowledgeable person (expert).

The expert is obliged not only to report his conclusion on the issue of interest to the investigator or court, he must justify it, and at the same time, so that the investigator or court can understand and assess whether this conclusion is correct. Automated Support and Decision-making Systems (ASDS) can provide essential assistance in this regard.

The expert version is an expert scientific assumption based on factual materials and special knowledge, explaining the essence, properties and origin of the fact under study. Versions are the stages of scientific expert research on the way to solving the questions posed to the expert. They determine the content, nature and direction of the expert's research work. Versions require a deep logical analysis from
an expert based on special knowledge of each fact individually and in their totality and their relationship. In some cases, automated databases of expert knowledge would be a good help.

In order to address issues related to the improvement of forensic analysis, it is necessary to have, above all, appropriate information support. There is no doubt that an effective means of improving this area is the use of new information technologies, which until now had a local character and had a limited positive effect.

This is largely due to the lack of a general concept of informatization, necessary data, scientifically based models and methods of preparing and making decisions on the organization of activities, modern technology, as well as progressive methods of organizing and conducting diagnostic work in the conditions of applying new information technologies.

The manual offers a number of specialized applied tasks in the field of technical diagnostics and provides the following results:

- Mechanisms of conclusion of typical situations on the issues of forensic diagnostics in the knowledge bases of operative-advisory expert systems have been developed;
- The method of training and adaptation of decision support system in the field of forensic diagnostics is proposed;
- Modal logic was chosen as a mechanism for modeling plausible reasoning in the selected subject area as the most acceptable mechanism for modeling plausible reasoning in conducting forensic examinations;
- The example of using the output mechanism studied in the manual allows us to count on its successful use in the practice of designing real databases of OSES knowledge (operational and advisory expert systems) of diagnostic tasks in the selected subject area;
- Researches on definition of limits of possibilities of experts in problems of ordinal classification are made. Recommendations on using a hierarchy of criteria, identifying groups of similar criteria, defining concepts characteristic of these groups, building classification hierarchies are proposed.

The obtained results can be used for intellectual support of substantiation and decision-making also on the issues of investigation of certain types of crimes, conducting other types of forensic examinations.

2. Analysis of mechanisms of using the concept of probabilities to model plausible reasoning in technical diagnostics

The obtained results can be used for intellectual support of substantiation and decision-making also on the issues of investigation of certain types of crimes, conducting other types of forensic examinations.

The field of logic that explores plausible reasoning, namely inductive and probabilistic logic, has received a new powerful impetus for development in connection with its applications to artificial intelligence. The interest to inductive and probabilistic problems on the part of specialists is caused by the fact that the majority of human reasoning (i.e., in which new, nontrivial, even if not 100% reliable conclusions are obtained) are, in fact, inductive. Therefore, in connection with work in the field of artificial intelligence, in the broadest sense, inductive problems acquire great importance: the extraction of general laws from disparate data; behavior under conditions of uncertainty, i.e., the ability to operate with inaccurate probability rules and data; conclusions based on evidence of pros and cons*. Research in this field is of great applied importance for building expert systems where one usually has to deal with a lot of data, inaccurate information and rules of approximate, intuitive character.

Various means are used to present plausible reasoning. These are non-monotonous logic, fuzzy logic, other logical systems that use numerical values of degrees of plausibility, modal logic. It seems reasonable to describe in detail the methods most related to the traditional problems of inductive logic.

Obviously, the Bacon-Mill methods as a model for real scientific activity are too much simplification. They can provide some justification for expert conclusions, but such rules are not enough to automate the process of hypothesis making. To strictly follow these rules, one should either consider an infinite number of cases or accept a judgment (which in its turn cannot be based on anything...
but induction) about the range of possible causes of this phenomenon of intellect. It would be very good to teach the robot to learn from experience and generalize facts by passing on at least some of the skills that people have. Undoubtedly, it would correspond to creation of mini logic of opening which deals always with the final information, with the limited sphere of reality, but, nevertheless, simulating some processes of human inductive reasoning. The question consists only in whether these processes can be automated.

The results obtained recently answer this question in the affirmative. Programs have been created that can find regularities in the database, make generalizations, build new concepts and propose hypotheses.

Numerous notions of the degree of fact support have been developed that use the function of probabilistic measure, but differ from simple conditional probability. Their definitions are given in the book [Kaiberg, 1998]. The similarity of the functions suggested by different authors is that the degree of fact support is always proportional to the difference between the conditional probability of a given hypothesis and its initial probability: \( P(E/H) - P(E) \) or its initial probability when the hypothesis is rejected: \( P(E/N) - P(E/\neg H) \). Moreover, it is inversely proportional to the initial probability of evidence \( E \) - the less empirical facts are expected to be encountered, the less their initial probability, the more important their presence is.

Usually an expert, including in the forensic field, has to deal with incomplete information. Therefore, the system must also learn to reason by making hypotheses and comparing the degree of their validity. The expert system deals not only with data on specific facts, but also with knowledge. This knowledge, i.e. rules related to professional experience (tasks of expert diagnostics in forensics), which can be expressed, for example, in the form of product rules. Normally, rules and information about facts are attributed to reliability characteristics expressed in numbers, for example: "If A, then - with reliability degree x-B". The output mechanism must find logical implications from the information available and determine the degree of reliability. The user should be able to find out on the basis of which rules this or that conclusion is made, so usually an explanatory subroutine is included in the system, which in response to a request gives a list of rules used.

For a number of reasons, the degree of confirmation is not equated with the conditional probability \( P(N/E_1\&\ldots\&E_p) \). One of them is that a fact similar to the paradoxes of confirmation has been revealed. When working with experts who were asked to assess the reliability levels for the rules, it turns out that the identification of the reliability level with the probability does not suit experts for some intuitive reasons. The experts refused to attribute even a small degree of reliability on the basis of \( E_1, E_2, E_3 \) to \( H \) denial, because this evidence \( H \). This makes it necessary to choose another function to represent the degree of reliability more consistent with intuition. Namely, it is advisable to introduce two functions: the MB confidence measure reflecting the presence of evidence in favour of the hypothesis and the MD mistrust measure to present refuting evidence. The number that is put in accordance with each rule, the degree of reliability of \( CF \) is defined as the difference in values of these functions. Let the reliability levels for all rules be known now and let the data be available to apply several rules. How to calculate the resulting degree of reliability (summarize the certificates)? It is impossible to do it using only theoretical and probabilistic calculations. For \( P(H/E_1) \) and \( P(N/E_2) \) it is impossible to calculate \( P(N/E_1\&E_2) \), for \( P(N/E) \) and \( P(E/X) \) it is impossible to find \( P(N/X) \). Therefore, it is also impossible to calculate the corresponding reliability levels using only MB and MD definitions. Therefore, it is necessary to introduce approximate estimation methods. These methods have been constructed as methods of finding an approximate value of reliability when the probability required to calculate the exact value is unknown. Similarly, for complex hypotheses, fuzzy logic methods are used that also give only an approximate value:

\[
MB(H_1\&H_2,E) = \min(MB(H_1,E), MB(H_2,E))
\]

\[
MD(H_1 \&H_2,E) = \max(MD(H_1, E), MD(H_2, E))
\]
MB(H1 H2 ,E) = max(MB(H1E), MB (H2, E))

MD(H1 H2 ,E) = min(MD(H1E), MD (H2,E))

A parcel in a rule may, in turn, only have a certain degree of reliability. Then the following rule can be used to calculate the reliability of the conclusion:

\[ MB(H,E) = MB'(H,E) \max(0, CF(E,X)) \]  \hspace{1cm} (2)

\[ MD(H,E) = MD'(H,E) \max(0, CP(E,X)) \]

where X is all available data; MB’ and MD’ is a measure of trust (mistrust) in case you know that E is true.

Evidence in favor of the chosen methods of calculation of reliability degrees could be the successful work of the expert system. But, although the system gives successful recommendations, it is shown that they will not change if the values of reliability characteristics are changed within 0.2. Thus, the question of validity of using approximate methods remains.

It is very difficult to avoid the use of approximate methods to model plausibility reasoning. With respect to probability, logical strings do not have the property which in logic is called truth functionality, i.e. the probability of a complex statement cannot be calculated by the probabilities of its components. For example, the probability A&B cannot be determined by the probabilities A and B; one must know more probability A provided that B is true, or some other measure of dependence A and B.

But maybe if you know the probabilities of the original statements and the degree of their relationship to each other, you can calculate the probability of any statement? To do this, the degree of dependence must have a property of truthful functionality, i.e., it must be possible to calculate with(A,C)(where with(A,C) is the dependence of A and C) and with(B,C). The conditional probability does not possess this property. It is proved that there is no function expressing the dependences of the two statements that possesses such a property [Kaidberg, 1997]. Whatever is the initial information about probabilities, at some stage it will turn out that it is impossible to calculate the degree of reliability of the statement of interest using only the laws of probability theory.

A serious argument is reliance on assumptions. It's proven that if the "standard" assumptions

1) Hypotheses H1, ...Hp n >2 are mutually exclusive and exhaustive,
2) for any E1 E2,Hi (1 t n) P(E1&E2/Hi) = P(E1/Hi) P(E2/Hi) add one more:
3) for any E1 E2 Hi (1 t n) P(E1&E2/Hi) = P(E1/Hi) P(E 2/Hi)

It follows from the totality of the assertions 1...3 that for each Hi there is only one Ej, changing its probability. It is clear that this situation simply cannot exist in a normal intellectual system. However, Assumption 3 is sometimes accepted. If there is sufficient statistical information for the area in which the system will be applied, including the mutual dependence of symptoms on each diagnosis, etc., the Bayesian theorem will give the best results.

The conclusions about the limits of the probability of statements are proven in probability theory by the initial information about the probabilities of other statements. Therefore, the system deals not with point values but with interval values of probabilities.

But maybe if you know the probabilities of the original statements and the degree of their relationship to each other, you can calculate the probability of any statement? To do this, the degree of dependence must have a property of truthful functionality, i.e., it must be possible to calculate with(A,C)(where with(A,C) is the dependence of A and C) and with(B,C). The conditional probability does not possess this property. It is proved that there is no function expressing the dependences of the two statements that possesses such a property [Kaidberg, 1997]. Whatever is the initial information about probabilities, at some stage it will turn out that it is impossible to calculate the degree of reliability of the statement of interest using only the laws of probability theory.

A serious argument is reliance on assumptions. It's proven that if the "standard" assumptions

1) Hypotheses H1, ...Hp n >2 are mutually exclusive and exhaustive,
2) for any E1 E2,Hi (1 t n) P(E1&E2/Hi) = P(E1/Hi) P(E2/Hi) add one more:
3) for any E1 E2 Hi (1 t n) P(E1&E2/Hi) = P(E1/Hi) P(E 2/Hi)

It follows from the totality of the assertions 1...3 that for each Hi there is only one Ej, changing its probability. It is clear that this situation simply cannot exist in a normal intellectual system. However, Assumption 3 is sometimes accepted. If there is sufficient statistical information for the area in which the system will be applied, including the mutual dependence of symptoms on each diagnosis, etc., the Bayesian theorem will give the best results.

The conclusions about the limits of the probability of statements are proven in probability theory by the initial information about the probabilities of other statements. Therefore, the system deals not with point values but with interval values of probabilities.

In the general case, we know that probability A lies in the range \{e(A), p(A)\}, i.e. probability A is equal to at least e(A), and probability A is equal to at least 1 - p(A). The advantages of this approach are obvious: unlike the Bayesian one, it allows to consider evidence for and against A separately. Although the result of the probability boundary statements is less certain, it is more reliable. If there is indeed a relationship of independence or incompatibility between statements, this can be explicitly considered and used in probability calculations. In this case, the result is not only more definite, but also completely reliable. It is possible to enter information about probabilities of any set of statements into the system and to investigate the consequences for the whole system: to reveal inconsistencies, to clarify the limits of probability for any statement (there is no division into evidence and hypothesis).

The expression A characterizes two values: t(A) and f(A), where P(A)t(A) and P(A)f(A). Each
expression has trivial limits: \( t(A) = 0 \) and \( f(A) = 0 \). As we get new information, the interval in which probability \( A \) can lie becomes everything. This is expressed in increasing values of \( t(A) \) and \( f(A) \). The information about statement \( A \) is agreed if \( t(A) + f(A) = 1 \).

Type relations are used:

- \( A \) indicates in favor of \( S \) with degree \( X \) \( P(S/A) \leq X \);
- And there is a conjunction of independent \( (S_9...Sn) \): \( A \& G_3 G \) and for all \( ij \) \( P(S.t&\&S_j) = P(St) \) \( P(Sj) \), etc.

Let us suppose that we know the meaning for the boundaries of a statement. All relations in which it is involved should be considered, and perhaps the meanings for the boundaries of other statements should be increased so that the rules of probability redistribution are followed. They represent inequalities which are interpreted as product rules:

IF: the previous value of the border on the left side is less than the value on the right side,
THEN: The border should be increased to this new value,
For example, \( A \) testifies in favor of \( S \) with degree \( X \):

\[
\begin{align*}
t(S) & = t(A)X \\
t(A) & = 1 - f(S) \left(1 - f(A)X\right) \\
\end{align*}
\]

Unfortunately, if the rules are deduced from the interpretation of relations, then the opposite is wrong. The values for the boundaries derived from the rules will always be correct, but sometimes weaker than that derived from the interpretation of relationships.

The aforementioned methods were mainly applied in practice, many of which do not have sufficient theoretical substantiation and are “case-by-case” in nature. We will now focus on approaches to the presentation of plausible reasoning in artificial intelligence in terms of theoretical logic.

The article [Nielson, 1997] suggests a probabilistic logic to solve the problem of probabilistic sequencing: to determine the probability (or boundaries, in which the probability lies) of some proposals on the probabilities of others. Let the set of sentences \( \{S_j,...Sn\} \) be given to these sentences. \( k (k \geq n) \) truth values can be assigned to these sentences in different ways (we mean only coordinated adsorption). Each such assignment can be identified with a possible world (more precisely, with the class of equivalence of possible worlds relative to this assignment).

Let the probabilities of all proposals, except \( Sn \). Let's denote them as \( 1 ... n-1 \). For any solvable logic, a matrix can be built that specifies all possible coordinated adsorption of \( S_1,..., St \) proposals. Let's call this matrix \( V \). Each column of this matrix corresponds to one of possible truth assignment for the given sentences \( v_{ij} \) is equal to \( 1 \), if at \( j \)-assignment \( Si \) gets the value "truth", and \( 0 \) otherwise:

\[
V = \begin{bmatrix} v_{11} & v_{12} & \ldots & v_{1k} \\
v_{21} & v_{22} & \ldots & v_{2k} \\
\vdots & \vdots & \ddots & \vdots \\
v_{n1} & v_{n2} & \ldots & v_{nk} \end{bmatrix}
\]

Let the \( k \)-dimensional vector of column \( P \) be given, its \( J \) component is \( p_J \) - probability of \( J \)-offering. Probabilities of proposals are given by vector \( P \): \( \Pi = V*P \) (1.5)

Thus, the probability \( i \) of the \( Si \) sentence (i-component \( P \)) is equal to the sum of the probabilities of those ascribed to which \( Si \) is true.

Let the probabilities \( S1...Sn-1 \) be known now. It is necessary to determine the probability \( Sn \). Let us introduce \( T \) (tautology) as the first element of the set of sentences. The first line \( V \) will consist of one unit. Let the last line \( V \) contain the truth values of \( Sn \). The conditional probability values are given for all sentences except \( Sn \); the tautology probability is \( 1 \). Let's consider the matrix \( V^T \) (without the last line) and vector \( P^T \) (without \( Sn \) probability). Having solved the equation \( P^T = V^T*P \) and found \( P \), it is easy to determine the probability \( Sn \) - it is equal to the product of the last line \( V \) in \( P \). Usually there are many possible solutions for \( P \), and for probability \( Sn \) we can only specify the boundaries in which it should lie.

The article [Kaidberg, 1998] gives several methods that simplify the solution of the resulting system of equations. But all of them become unsuitable when a large number of proposals are considered.
simultaneously. In this case calculations become practically impossible. For this case (large matrices) we also give several variants of approximate calculations the accuracy of which depends on available computational resources.

Nilson’s first work containing a description of his probabilistic logic appeared in 1984, and in 1985 A. Banda, another well-known specialist in artificial intelligence, offered his solution to the problems associated with probabilistic reasoning in expert systems. It consists in making the ligaments of probabilistic logic truth-functional. For this purpose, as a measure of reliability of the formula, A. Bundy does not introduce numerical values of probabilities but sets of points or elementary events for which this proposal is true. He referred to such sets as incidents, or degrees of prevalence of sentences.

Let W be a set of points (elementary events), exhaustive and mutually exclusive.

For computational reasons, the finite set W is considered; i(A) - (incident A) - a subset of W including all those points where A is true.

\[
i(T) = W \\
I(F) = \\
i(A) = W / i(A) \\
i(A&B) = i(A) & i(B) \\
i(AB) = i(A) & i(B) \\
i(xA\theta) i(A\theta) & i(\thetaA),
\]

where the e is a tautonic that contains no variables free in \( A \);

T and F are tautology and contra-vocabulary, respectively. Bonds are thus truth-functional, quantors are not. Upper and lower boundaries are given for quantors.

Probability A - P(A) - equals the sum of probabilities of the points constituting i(A). The article [Mitchie, 1999] considers the simplest case when all the points are equal. Then probability A equals the relation of the number of points in i(A) to the number of points in W.

The mechanism of output under conditions of uncertainty should include the rule of determination of reliability level B by reliability levels A and AB. A. The gang names this property by analogy with the truth functionality as evidentiary functionality. It takes place in a weakened form, i.e. as a result of several applications of modusropes we obtain different values for the lower boundary B, then we apply the multiple-theoretical property

\[
i1 (B) & i2 (B) i1 (B),
\]

where i(B) is incident B; i1 and i2 are the lower limits. This method is similar to the summation of evidence in MICINE, but in contrast to the MICINE rule, which is ad hoc in nature, is based on the theory of sets. Since the ligaments are truth-functional, the lower boundary is defined much more precisely than in the logic that uses probabilities.

The article [Mitchie, 1999] gives a mechanism for deriving the logic of incidents that deal with the upper and lower limits. It is complete for propositional logic. This algorithm is implemented by the author on the Prologue. In addition, it provides a mechanism to detect inconsistencies in the entered data. The method proposed by Nilson requires much less computation in comparison with other methods.

The output rules allow to go from one attribute, say, F, to another - G, which gives narrower limits for Scr(A) and inf (A) - the highest and lowest values of incident A. Here is an example of a rule for the conjuncture:

\[
\text{ANDLsupg} = \{\text{supf}(A&B) \setminus \text{inff} (B)\} \text{ supf}(A)
\]

Modal logic is increasingly used in artificial intelligence, including for the presentation of plausible
reasoning. First of all, this applies to those cases when the concept of probability is of sufficient quality: "likely", "more likely", "less likely", etc. That is why, with respect to expert systems of the management type, it seems optimal to construct a logic with the operator I: LA reads approximately as "A is a plausible hypothesis concerning available knowledge". There are no quantitative estimations of degrees of plausibility, but their difference can be transferred due to iteration of operators: LLA is weaker than LA, and in general, Ln, where n is the number of operators before A, is weaker than if m<n. For this logic, an enabling procedure is given to determine whether some statement from the set of other statements follows or whether some statement is true. It is possible to use modal logic not only in qualitative concepts. For example, it is natural to use logic with operators "the probability...is more or equal to x" or "the probability...is in the range from x to y" when checking the entered statements about the probability of consistency. It is the modal logic that seems to be the most acceptable as a mechanism for modeling plausible reasoning in the chosen subject area.

References
[1] Closs K 2007 Data Mining. A knowledge discovery approach. Springer. 606
[2] Kaidberg G 1998 Probability logic and probabilistic induction. Progress. 375
[3] Kaidberg G 1999 Probabilistic Inductive Logic Programming ILP. 31
[4] Cielen D, Meysman A, Mohamed A 2017 Introducing Data Science: Big Data, Machine Learning, and more, using Python tools. St-Petersburg. 119-121
[5] Steinbach M, Kumar V, Pang-Ning T 2016 Introduction to Data Mining. London. 71-74, 125-129
[6] Casey D, Burrell P, Sumner N 2018 Decision Support Systems in Policing. London. 2-3
[7] Rzaev R R 2016 Decision Making Support Information System. Systems and means of informatics. St-Petersburg. 182–198
[8] Benmimoune L 2015 Ontology-based Medical Decision Support System to Enhance Chronic Patients' Lifestyle within E-care Telemonitoring Platform. ICIMTH. 279–282
[9] Salamon J S 2017 Using Goal Modeling and OntoUML for Reengineering the Good Relations Ontology. ONTOBRAS. 91–102
[10] Rospocher M, Serafini L 2012 An ontological framework for decision support. Springer, Berlin, Heidelberg. 239–254
[11] Golovnin O K, Romanova E A 2018 Precedential decision support system for administrative offenses. Samara University. 1(31) 44-45
[12] Mizumoto M, "Fuzzy reasoning with a fuzzy conditional proposition "if ... then ... else ..." Pergamon Press 1982. 211
[13] Smarandache F, New types of Neutrosophic Set/Logic/Probability, Neutrosophic Over-/Under-/Off-Set, Neutrosophic Refined Set, and their Extension to Plithogenic Set/Logic/Probability, with Applications 2019. 45