Obrad Kasum, Aleksandar Perović and Aleksandar Jovanović

***Abstract*—In this talk we will give a short survey of weighted logics and their application in artificial intelligence. The emphasis will be on propositional real-valued probability logics, the main axiomatization issues, their relation with classical, modal logics and theoretical computer science, as well as their application in various expert systems.**

***Index Terms*—Fuzzy logic, probabilistic logics.**

I. Introduction

IN broader sense, weighted logics can be understood as a tool for reasoning with incomplete or imprecise information (knowledge), where the uncertainty of the premises is expressed by qualitative or quantitative statements. Arguably, the most important types of weighted logics are possibility and necessity logics, fuzzy logics and probability logics. The common feature of all this logics is the fact that all of them allow more than two truth values. In the standard form the domain of truth values is the real unit interval $[0,1]$. Though the syntax is far from standardized (practically each research team has its own notation), weighted logics utilize syntactical means to express degrees of truthfulness of certain sentences, as well as qualitative statements. For example, qualitative statements like “the seasonal flue is the probable cause for the patient’s fever”, or “the HPV is more probable cause for the observed cervical cancer than exposure to gamma radiation” have very simple mathematical representation in various fuzzy and probability logics. Similarly, quantitative statements like the incidence of monozygotic twinning is about 3/1000”, or in “around 30% cases of the coronary thrombosis are caused by smoking” can also be quite simply and naturally coded within the framework of weighted logics.

Since the late sixties, probability theory has found application in development of various medical expert systems. Bayesian analysis, which is essentially an optimal path finding through a graph called Bayesian network, has been (and still is) successfully applied in so called sequential diagnostics, when the large amount of reliable relevant data is available. The graph (network) represents our knowledge about connections between studied medical entities (symptoms, signs, diseases); the Bayes formula is applied in order to find the path (connection) with maximal conditional probability. Moreover, a priori and conditional probabilities were used to define a number of measures designed specifically to handle uncertainty, vague notions and imprecise knowledge. Some of those measures were implemented in MYCIN in the early seventies [96]. The success of MYCIN has initiated construction of rule based expert systems in various fields.

However, expert systems with the large number of rules (some of them like CADIAG-2 have more than 10000) are designed without any proper knowledge of mathematical logic. As an unfortunate consequence, most of them are turned to be inconsistent. On the other hand, the emergence of theoretical computer science as a new scientific discipline has led to discovery that the completeness techniques from mathematical logic are the only known methods for proving correctness of hardware and software. Consequently, mathematical logic has become a theoretical foundation of artificial intelligence.

The last three decades has brought a rapid development of various formal logics that can describe plethora of AI settings. Arguably, the most significant among those are fuzzy logics [3,4,6,9], possibilistic logics [1-3] and probability logics. Our focus will be solely on the probability logics, since it is our field of expertise.

II. Probability logics

Mathematical representation of probabilistic reasoning extends basic logical language (that involves propositional connectives and universal and existential quantifiers) with probabilistic operators and probabilistic quantifiers. Though the roots of probability logic can be traced at least to Leibnitz, the modern era of probability logic has started with the work of Jerome Keisler [11-13] throughout the seventies and the mid eighties of the XX century. It is worth mentioning that Bayesian analysis (application of Bayes formula in determination of optimal diagnostic/therapy strategies) has been successfully applied in early clinical decision support systems specialized in sequential diagnostics in the late sixties (see for instance [7]).

This work is partially supported by Serbian Ministry of Education and Science through grants.

Obrad Kasum is a student at School of Mathematics in Belgrade and a member of Group for Intelligent Systems at Faculty of Mathematics in Belgrade.

Aleksandar Perović is associated professor at Faculty of Transport and Traffic Engineering, University of Belgrade. He is also a member of Group for Intelligent Systems at Faculty of Mathematics in Belgrade (e-mail: pera@sf.bg.ac.rs).

Aleksandar Jovanović is associated professor at Faculty of Mathematics, University of Belgrade. He is also a member of Group for Intelligent Systems at Faculty of Mathematics in Belgrade (e-mail: aljosha.jovanovich@gmail.com).

The modal representation of probability, i.e. introduction of modal like probability operators in classical reasoning, deeply motivated by intensive application in various expert systems, was initiated by Nils Nilsson [14,15] in the mid eighties and early nineties. A major breakthrough along these lines was made by Ronald Fagin, Joseph Halpern and Nimrod Megiddo [4], especially in terms of decidability and computational complexity (so called small model theorems). Though the introduced syntax was not modal per se, it was very similar to it and the developed modal probability semantics has become the standard one.

The first probability logic with unary modal probability operator was introduced by Miodrag Rašković in the early nineties [20]. Soon after, a rather rapid development of the subject has followed. We shall track a selection of the existing research in the field of probability logic: primarily our own contributions, "seasoned" with certain papers (books) that are closely related to our work.

To begin with, an extensive study of finitely additive probability measures was given in [19]. Historical development, various boundaries of probability functions and many other important concepts regarding sentential probability logics are given in [8]. An extensive study of uncertainty and its connection with probability was given in [10]. Various formalizations of probability with variety of scopes - simple probabilities, higher order (nesting of probability operators) probabilities, conditional probabilities, representation of default reasoning etc. are presented in past three decades by numerous research groups. Usually, axiomatizations involve wide range of probabilistic distributions, i.e. there are very few restrictions on sematical functions other than the basic condition that they have to be finitely additive probabilities. Then, as probabilities are generally not truth-functional, the best one can do is to calculate bounds on probabilities of conclusions starting from probabilities of assumptions [8].

One of the main proof-theoretical problems is providing an axiom system that would be strongly complete in the sense that every consistent theory has a model. This problem originates from the inherent non-compactness of so called non restricted real-valued probability logics. Namely, in such formalisms it is possible to define an inconsistent infinite set of formulas, every finite subset of which is consistent. For example, one such theory is given by

$$\left\{ P_{>0}\alpha \right\} \cup \left\{ P_{<\frac{1}{n}}\alpha : n \text{ is a positive integer} \right\}.$$

As it was pointed in [16,22], there is an unpleasant consequence of finitary axiomatization in that case: there exist unsatisfiable sets of formulas that are consistent with respect to the assumed finite axiomatic system (since all finite subsets are consistent and deductions are finite sequences). Another important theoretical problem is related to the decidability issue.

LPP_2 logic. We shall briefly describe the LPP_2 probability logic. Detailed exposition can be found in [17]. It is an extension of the classical propositional logic with probability operators of the form $P_{\geq s}$, where s can be any rational number between 0 and 1 (including both of them).
The initial syntactical layer is formed of classical propositional formulas; they will be denoted by α, β and γ, indexed or primed if necessary. Basic probability formulas are expressions of the form

$$P_{\geq s}\alpha .$$

The intended meaning of $P_{\geq s}\alpha$ is rather obvious: the probability of α is at least s. Finally, complex probability formulas are formed from the basic ones by application of logical connectives: negation (denoted by $\neg$) and implication (denoted by $\rightarrow$). Probability formulas will be denoted by A, B and C, indexed or primed if necessary. Some standard abbreviations (e.g. formal introduction of conjunction, disjunction and equivalence) are defined in the usual way:

- $A \wedge B =_{def} \neg(A \rightarrow \neg B)$;
- $A \vee B =_{def} \neg A \rightarrow B$;
- $A \leftrightarrow B =_{def} (A \rightarrow B) \wedge (B \rightarrow A)$;
- $P_{\leq s}\alpha =_{def} P_{\geq s}\neg\alpha$;
- $P_{>s}\alpha =_{def} P_{\geq s}\alpha \wedge \neg P_{\leq s}\alpha$;
- $P_{<s}\alpha =_{def} P_{\leq s}\alpha \wedge \neg P_{\geq s}\alpha$;
- $P_{=s}\alpha =_{def} P_{\geq s}\alpha \wedge P_{\leq s}\alpha$.

Due to the modal nature of probability operators, the standard probabilistic semantics is defined on so called probability Kripke structures. A probability Kripke structure is any triple (W, H, μ) with the following properties:

1. W is a nonempty subset of the set of all classical evaluations $\{0,1\}^{Var}$;
2. H is an algebra of sets (it is nonempty and closed under intersection, union and complement) that contains all sets of the form $[\alpha]$. Here by $[\alpha]$ we have denoted the set of all evaluations satisfying α;
3. $\mu : H \longrightarrow [0,1]$ is a finitely additive probability measure.

The satisfiability relation $\vDash$ is defined in the following way:

- $(W, H, \mu) \vDash \alpha$ iff $[\alpha] = W$;
- $(W, H, \mu) \vDash P_{\geq s}\alpha$ iff $\mu([\alpha]) \geq s$;

- $(W,H,\mu) \vDash \neg A$ iff $(W,H,\mu) \nvDash A$;
- $(W,H,\mu) \vDash A \to B$ iff either $(W,H,\mu) \nvDash A$, or $(W,H,\mu) \vDash A$ and $(W,H,\mu) \vDash B$.

The axioms of the LPP_2 logic are the following ten schemata:

- Ax1: $\alpha \to (\beta \to \alpha)$;
- Ax2: $(\alpha \to (\beta \to \gamma)) \to ((\alpha \to \beta) \to (\alpha \to \gamma))$;
- Ax3: $(\neg\beta \to \neg\alpha) \to (\alpha \to \beta)$;
- Ax4: $A \to (B \to A)$;
- Ax5: $(A \to (B \to C)) \to ((A \to B) \to (A \to C))$;
- Ax6: $(\neg B \to \neg A) \to (A \to B)$;
- Ax7: $P_{\geq 0}\alpha$;
- Ax8: $P_{\geq s}\alpha \to P_{>r}\alpha$ for all $r < s$;
- Ax9: $P_{\geq s}\alpha \wedge P_{\geq r}\beta \wedge P_{=0}(\alpha \wedge \beta) \to P_{\geq \min(r+s,\ 1)}(\alpha \vee \beta)$;
- Ax10: $P_{\leq s}\alpha \wedge P_{\leq r}\beta \to P_{\leq \min(r+s,\ 1)}(\alpha \vee \beta)$.

The inference rules of the LPP_2 logic are the following one:

- Modus ponens for classical formulas: from α and $\alpha \to \beta$ infer β;
- Modus ponens for probability formulas: from A and $A \to B$ infer B;
- Necessitation: from α infer $P_{=1}\alpha$;
- Archimedean rule: from the set of premises $\{A \to P_{\geq r}\alpha : r < s\}$ infer $A \to P_{\geq s}\alpha$.

The notion of deduction differs from the classical one only in the length of the inference: since Archimedean rule has countably many premises, the length of the inference can be any countable successor ordinal.

Intuitively, Archimedean rule should be understood in the following way: if the probability of α is infinitely close to the rational number s, then it must be equal to s. As a consequence, "problematic" finitely satisfiable but unsatisfiable theories such as previously mentioned theory

$$\{\ P_{>0}\alpha\ \} \cup \left\{P_{<\frac{1}{n}}\alpha : n \text{ is a positive integer}\right\}$$

become inconsistent in LPP_2. The proof of the strong completeness theorem for LPP_2 logic can be found in [17].

Decidability and complexity. Any potential or actual application of weighted logics in artificial intelligence is closely related to the satisfiability problem and related computational complexity estimation. Here we shall outline the satisfiability procedure for the LPP_2-formulas and give its exact complexity. So, let $A \in For_P$. Recall that an atom a of A is a formula of the form $\pm p_1 \wedge \cdots \wedge \ \pm p_n$, where $\pm p_i$ is either p_i or $\neg p_i$, and $p_1, \ldots, p_n$ are all primitive propositions appearing in A. For example, if A is the formula $P_{\geq 0.9}(p \vee q)$, then its atoms are $p \wedge q$, $p \wedge \neg q$, $\neg p \wedge q$ and $\neg p \wedge \neg q$.

Note that atoms are pairwise disjoint. Hence, for any probability measure μ and any pair of atoms a_i and a_j ($a_i \neq a_j$) we have that

$$\mu(a_i \vee a_j) = \mu(a_i) + \mu(a_j).$$

As a next step we can equivalently transform the given formula A into its complete disjunctive normal form

$$DNF(A) = \bigvee_{i=1}^{m} \bigwedge_{j=1}^{k_i} X^{i,j}(p_1,\ \ldots,\ p_n),$$

where:

- $X^{i,j}$ is one of probability operators $P_{\geq s_{i,j}}$ and $P_{<s_{i,j}}$;
- $X^{i,j}(p_1, \ldots, p_n)$ denotes the fact that the propositional formula which is in the complete disjunctive normal form, i.e. the propositional formula is a disjunction of the atoms of A.

Example. The complete disjoint normal form of the formula A defined by

$$P_{<0.1}p \vee P_{\geq 0.8}q$$

is disjunction of the following four formulas:

- $P_{<0.1}\big((p \wedge q) \vee (p \wedge \neg q)\big) \wedge P_{\geq 0.8}\big((p \wedge q) \vee (\neg p \wedge q)\big)$;
- $P_{<0.1}\big((p \wedge q) \vee (p \wedge \neg q)\big) \wedge P_{<0.8}\big((p \wedge q) \vee (\neg p \wedge q)\big)$;
- $P_{\geq 0.1}\big((p \wedge q) \vee (p \wedge \neg q)\big) \wedge P_{\geq 0.8}\big((p \wedge q) \vee (\neg p \wedge q)\big)$;
- $P_{\geq 0.1}\big((p \wedge q) \vee (p \wedge \neg q)\big) \wedge P_{<0.8}\big((p \wedge q) \vee (\neg p \wedge q)\big)$. □

The logic LPP_2 is decidable, i.e. the satisfiability and validity of LPP_2-formulas is algorithmically solvable. Firstly we will outline satisfiablilty algorithm in general case, then we will apply it on the previous example.

The first step is to transform the given For_P-formula A to its complete disjunctive normal form

$$DNF(A) = \bigvee_{i=1}^{m} \bigwedge_{j=1}^{k_i} X^{i,j}(p_1, \ldots, p_n).$$

So, satisfiability of A is equivalently reduced to the satisfiability of $DNF(A)$. Thus, A is satisfiable iff at least one disjunct from $DNF(A)$ is satisfiable. Let the measure of the atom a_i be denoted by y_i. We use an expression of the form $a_t \in X(p_1, \ldots, p_n)$ to denote that the atom a_t appears in the propositional part of $X(p_1, \ldots, p_n)$.

Furthermore, a disjunct $D = \bigwedge_{j=1}^{k} X^j(p_1, \ldots, p_n)$ from $DNF(A)$ is satisfiable iff the following system of linear equalities and inequalities is satisfiable:

$$y_1 + \cdots + y_{2^n} = 1$$

$$y_1 \geq 0$$

$$\vdots$$

$$y_{2^n} \geq 0$$

$$\sum_{a_t \in X^1(p_1, \ldots, p_n) \in D} y_t \geq^{X^1} s_1$$

$$\vdots$$

$$\sum_{a_t \in X^k(p_1, \ldots, p_n) \in D} y_t \geq^{X^k} s_k,$$

where $\geq^{X^i} = \geq$ if $X^i = P_{\geq s_i}$, otherwise $\geq^{X^i} = <$.

Since the satisfiability of A is reduced to the linear systems solving problem, the satisfiability problem for LPP_2-logic is decidable. Finally, since A is valid iff $\neg A$ is unsatisfiable, the validity problem is also decidable.

Back to the previous example: the atoms of the given probability formula $A = P_{<0.1}p \vee P_{\geq 0.8}q$ are $p \wedge q$, $p \wedge \neg q$, $\neg p \wedge q$ and $\neg p \wedge \neg q$. By $y_1, \ldots, y_4$ we will denote their unknown probabilities. The first disjunct in $DNF(A)$ generates the following system:

$$y_1 + y_2 + y_3 + y_4 = 1$$

$$y_1 \geq 0$$

$$y_2 \geq 0$$

$$y_3 \geq 0$$

$$y_4 \geq 0$$

$$y_1 + y_2 < 0.1$$

$$y_1 + y_3 \geq 0.8.$$

One solution of this system is given by $y_1 = y_2 = 0$, $y_3 = 0.8$ and $y_4 = 0.2$, so the formula A is satisfiable.

Concerning complexity estimation of the decision procedure, we shall show that it is NP-complete. Indeed, the lower bound follows from the complexity of the same problem for classical propositional logic. The upper bound is a consequence of the NP-complexity of the satisfiability problem for linear weight formulas from [4].

III. CONCLUSION

Since the late sixties, probability logics have been used in construction of various expert systems and decision support systems in medicine and other areas. Development of fuzzy sets and systems in mid-sixties and the corresponding rapid development in engineering community (fuzzy controllers) have led to extensive development of the corresponding mathematics and its integration in automated reasoning and various problem solutions in the field of artificial intelligence. In the mid-eighties Didier Dubois developed possibilistic logic and together with Henri Prade formed a research team that have integrated this new type of logic as another relevant and useful logic for modeling uncertainty.

Today weighted logics represent a thriving research area for mathematicians, theoretical computer scientists and various engineers.

REFERENCES

1. D. Dubois, J. Lang, H. Prade. Possibilistic logic. In: Handbook of Logic in Artificial Intelligence and Logic Programming, Vol. 3, ed. D. M. Gabbay et al., Oxford University Press (1994), pp. 439-513.
2. D. Dubois, H. Prade. Possiblistic logic: a retrospective and prospective view. Fuzzy sets and systems 144: 3-23, 2004.
3. D. Dubois, L. Godo, H. Prade. Weighted logic for artificial intelligence: an introductory discussion. In ECAI-2012 workshop Weighted logic for artificial intelligence (eds L. Godo, H. Prade): 1-7, 2012.
4. R. Fagin, J. Halpern and N. Megiddo. A logic for reasoning about probabilities. Information and Computation 87(1-2): 78-128, 1990.
5. L. Godo, F. Esteva, P. Hájek. Reasoning about probability using fuzzy logic. Neural network world 10(5): 811-824, 2000.
6. L. Godo, E. Marchioni. Coherent conditional probability in a fuzzy logic setting. Logic journal of the IGPL 14(3): 457-481, 2006.
7. A. Gorry, O. Barnett. Sequential Diagnosis by Computer. The Journal of the American Medical Association 205(12): 849-854, 1968.
8. T. Hailperin. Sentential Probability Logic, Origins, Development, Current Status, and Technical Applications. Lehigh University Press, 1996.
9. P. Hajék. Metamathematics of Fuzzy Logic. Kluwer 1998.
10. J. Halpern. Reasoning about uncertainty. The MIT Press, 2003.
11. H.J. Keisler. Model theory for infinitary logic. North-Holland 1971.
12. H. J. Keisler. Hyperfinite model theory. In: Logic Colloquim 76. (eds. R. O. Gandy, J. M. E. Hyland): 5-110. North-Holland 1977.

13. H. J. Keisler. Probability quantifiers. In: Model Theoretic Logics, (eds. J. Barwise and S. Feferman): 509-556. Springer 1985.
14. N. Nilsson. Probabilistic logic. Artificial Intelligence, 28: 71-87, 1986.
15. N. Nilsson. Probabilistic logic revisited. Artificial Intelligence, 59: 39-42, 1993.
16. Z. Ognjanović, M. Rašković. Some first-order probability logics. Theoretical Computer Science 247(1-2): 191-212, 2000.
17. Z. Ognjanović, M. Rašković, Z. Marković. Probability logics. In Logic in computer science (ed Z. Ognjanović): 35-111, Matematički Institut SANU, 2009, ISBN 978-86-80593-40-1.
18. A. Perović, Z. Ognjanović, Miodrag Rašković, D. G. Radojević: Finitely additive probability measures on classical propositional formulas definable by Gödel's t-norm and product t-norm. Fuzzy Sets and Systems 169(1): 65-90, 2011.
19. K. P. S. Bhaskara Rao, M. Bhaskara Rao. Theory of charges. Academic press 1983.
20. M. Rašković. Classical logic with some probability operators. Publications de l'Institut Mathematique, n.s. 53(67): 1-3, 1993.
21. E. Shortliffe, B. Buchanan. A model of inexact reasoning in medicine. Mathematical Biosciences 23: 351-379, 1975.
22. W. van der Hoeck. Some consideration on the logics $P_F D$. Journal of Applied Non-Classical Logics 7(3): 287-307, 1997.