Copula as a Bridge Between Probability Theory and Fuzzy Logic

Germano Resconi and Boris Kovalerchuk

Abstract This work shows how dependence in many-valued logic and probability theory can be fused into one concept by using copulas and marginal probabilities. It also shows that the t-norm concept used in fuzzy logic is covered by this approach. This leads to a more general statement that axiomatic probability theory covers logic structure of fuzzy logic. This paper shows the benefits of using structures that go beyond the simple concepts of classical logic and set theory for the modeling of dependences.

1 Introduction

Modern approaches for modeling uncertainty include Probability Theory (PT), Many-Valued Logic (MVL), Fuzzy Logic (FL), and others. PT is more connected with the classical set theory, classical propositional logic, and FL is more connected with fuzzy sets and MVL. The link of the concept of probability with the classical set theory and classical propositional logic is relatively straightforward for the case of a *single variable* that is based on the use of set operations and logic propositions. In contrast, modeling joint probability for *multiple variables* is more challenging because it involves complex *relations* and *dependencies* between variables. One of the goals of this work is to clarify the *logical structure* of a joint probability for multiple variables. Another goal is to *clarify links* between different logics and joint probability for multiple variables.

Differences between concepts of probability, fuzzy sets, possibility and other uncertainty measures have been studied for a long time [1]. Recently several new attempts have been made to connect the concepts of probability, possibility, and

G. Resconi (✉)
Department of Mathematics and Physics, Catholic University, 25121 Brescia, Italy
e-mail: resconi@speedyposta.it

B. Kovalerchuk
Department of Computer Science, Central Washington University,
Ellensburg, WA 98926-7520, USA
e-mail: borisk@cwu.edu

V. Kreinovich (ed.), *Uncertainty Modeling*, Studies in Computational
Intelligence 683, DOI 10.1007/978-3-319-51052-1_15

fuzzy logic into a single theory. Dubois, Prade and others [2–7] discuss conditional probability and subjective probability as a mixed concept that includes both fuzzy and probability concepts.

In [8–11] we proposed the Agent-based Uncertainty Theory (AUT) with a model for fuzzy logic and many-valued logic that involves agents and conflicts among agents. Later we extended AUT, providing a more formal description of uncertainty as the conflict process among agents with the introduction of a: new concept denoted as an *active set* [12]. This agent-based approach to uncertainty has a connection with the recent works in the theory of team semantics [13], and dependence/independence logic [13, 14].

The *dependence logic* [13–21] assumes that the logic truth value of a particular proposition *depends* on the truth value of other propositions. Dependence logic is based on *team semantics*, in which the truth of formulas is evaluated in sets of assignments provided by a team, instead of single assignments, which is common in the classical propositional logic, which is *truth functional*. The basic idea of dependence logic is that certain properties of sets *cannot be expressed* merely in terms of properties satisfied by each set individually. This situation is more complex than that which is considered in the classical logic approach. It is closer to the agent-based uncertainty theory [8], and the active sets [12] that deal with the sets of assignments, generated by teams of agents, which can conflict and produce uncertainty of the truth value, leading to a many-valued logic. Dependence logic is a special type of the first order predicate logic. While the goal to go beyond truth-functionality is quite old in fuzzy logic [22] and the dependence logic is not the first attempt outside of the fuzzy logic, the major issue is that to reach this goal in practice we need more data than truth-functional algorithms use. The current survey of attempts to go beyond truth-functionality is presented in [23].

In PT and statistics, a *copula* is a multivariate probability distribution function for "normalized" variables [24]. In other words, copula is a multivariate composition of marginal probabilities. The copula values represent degrees of dependence among values of these variables when the variables are "cleared" from differences in scales and frequencies of their occurrence. This "clearing" is done by "normalizing" arguments of copula to uniform distributions in [0, 1] interval.

Quite often the probabilistic approach is applied to study the frequency of independent phenomena. In the case of dependent variables, we cannot derive a joint probability $p(x_1, x_2)$ as a product of independent probabilities, $p(x_1)p(x_2)$ and must use the multidimensional probability distribution with dependent valuables. The common technique for modeling it is a Bayesian network. In the Bayesian approach, the evidence about the true state of the world is expressed in terms of degrees of belief in the form of Bayesian conditional probabilities. These probabilities can be causal or just correlational.

In the probability theory, the conditional probability is the main element to express the dependence or inseparability of the two states x_1 and x_2. The joint probability $p(x_1, x_2, \ldots, x_n)$ is represented via multiple conditional probabilities to express the dependence between variables. The *copula approach* introduces a *single* function $c(u_1, u_2)$ denoted as *density of copula* [25] as a way to model the *dependence* or *inseparability* of the variables with the following property in the case of two variables. The copula allows representing the joint probability $p(x_1, x_2)$ as a combination (product) of single dependent part $c(u_1, u_2)$ and independent parts: probabilities $p_1(x_1)$ and $p_2(x_2)$.

It is important to note that probabilities $p_1(x_1)$ and $p_2(x_2)$ belong to two different pdfs which is expressed by their indexes in p_1 and p_2. Below to simplify notation we will omit these indexes in similar formulas.

The investigation of copulas and their applications is a rather recent subject of mathematics. From one point of view, copulas are functions that join or 'couple' one-dimensional distribution functions u_1 and u_2 and the corresponding joint distribution function.

Copulas are of interest not only in probability theory and statistics, but also in many other fields requiring the aggregation of incoming data, such as multi-criteria decision making [26], and probabilistic metric spaces [27, 28]. Associative copulas are special continuous triangular norms (t-norms for short, [29, 30]). They are studied in several domains such as many-valued logic [31], fuzzy logic and sets [32], agent-based uncertainty theory and active sets, along with t-conorms to model the uncertainty.

2 Copula: Notation, Definitions, Properties, and Examples

2.1 Notation, Definitions and Properties

Definitions and properties below are based on [24, 33–36].

A *joint probability distribution* is

$$
\begin{aligned}
&p(x_1 x_2, \ldots x_n) = \\
&p(x_1)p(x_2|x_1)p(x_3|x_1, x_2)\ldots p(x_n|x_1, x_2, \ldots, x_{n-1}).
\end{aligned}
$$

A function $c(u_1, u_2)$ is a *density of copula* for $p(x_1, x_2)$ if

$$p(x_1, x_2) = c(u_1 u_2)p(x_1)p(x_2) = p(x_1)p(x_2|x_1)$$

where u_1 and u_2 are *inverse functions*,

$$
\begin{aligned}
\frac{du_1(x_1)}{dx_1} &= p(x_1), u_1(x_1) = \int_{-\infty}^{x_1} p(s)ds, \\
\frac{du_2(x_2)}{dx_2} &= p(x_2), u_2(x_2) = \int_{-\infty}^{x_2} p(r)dr
\end{aligned}
$$

To simplify notation of $u_1(x_1), u_2(x_2), \ldots, u_n(x_n)$ sometimes we will write u_1, $u_2, \ldots, u_n$ omitting arguments.

A *cumulative function* $C(u_1(x_1), u_2(x_2))$ with inverse functions u_i as arguments is called a *copula* and its derivative $c(u_1, u_2)$ is called a *density of copula*:

$$C(u_1(x_1), u_2(x_2)) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} c(u_1(s), u_2(r)p(s)p(r)dsdr =$$
$$= \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} c(u_1(s), u_2(r)du_1(s)du_2(r)$$
$$c(u_1, u_2) = \tfrac{\partial^2 C(u_1,u_2)}{\partial u_1 \partial u_2}$$

Copula properties for 2-D case:

$$\begin{array}{c} p(x_1)p(x_2|x_1) = \dfrac{\partial^2 C(u_1, u_2)}{\partial u_1 \partial u_2} p(x_1)p(x_2) \\ p(x_2|x_1) = \dfrac{\partial^2 C(u_1, u_2)}{\partial u_1 \partial u_2} p(x_2) \end{array},$$

Copula properties for 3-D case:

Joint density function:

$$p(x_1, x_2, x_n) = p(x_1)p(x_2|x_1)p(x_3|x_1, x_2) =$$
$$c(u_1, u_2, u_3)p(x_1)p(x_2)p(x_3)$$

Copula, copula density, joint and conditional probabilities.

Below in $c(u_1, u_2, u_3)$ we omitted arguments s_1, s_2, s_3 of u_1, u_2, u_3 for simplicity.

$$C(u_1(x_1), u_2(x_2), u_3(x_3)) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\int_{-\infty}^{x_3} p(s_1, s_2, s_3)ds_1ds_2ds_3 =$$
$$\int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\int_{-\infty}^{x_3} c(u_1, u_2, u_3)p(s_1)p(s_2)p(s_3)ds_1ds_2ds_3 =$$
$$\int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\int_{-\infty}^{x_3} p(s_1)p(s_2|s_1)p(s_3|s_1, s_2)ds_1ds_2ds_3$$

$$p(x_1) = \frac{du_1}{dx_1}, p(x_2) = \frac{du_2}{dx_2}, p(x_3) = \frac{du_3}{dx_2}$$

$$C(u_1,u_2,u_3)=\int\limits_{-\infty}^{x_1}\int\limits_{-\infty}^{x_2}\int\limits_{-\infty}^{x_3} c(u_1,u_2,u_3)p(s_1)p(s_2)p(s_3)ds_1ds_2ds_3$$

$$c(u_1,u_2,u_3)=\frac{\partial^3 C(u_1,u_2,u_3)}{\partial u_1\partial u_2\partial u_3}$$

$$p(x_1,x_2,x_n)=p(x_1)p(x_2|x_1)p(x_3|x_1,x_2)=$$
$$\frac{\partial^3 C(u_1,u_2,u_3)}{\partial u_1\partial u_2\partial u_3}p(x_1)p(x_2)p(x_3)$$
$$p(x_2|x_1)p(x_3|x_1,x_2)=\frac{\partial^2 C(u_1,u_2,u_3)}{\partial u_1\partial u_2}p(x_2)p(x_3|x_1,x_2)=$$
$$\frac{\partial^3 C(u_1,u_2,u_3)}{\partial u_1\partial u_2\partial u_3}p(x_2)p(x_3)$$

$$\frac{\partial^2 C(u_1,u_2,u_3)}{\partial u_1\partial u_2}p(x_3|x_1,x_2)=\frac{\partial^3 C(u_1,u_2,u_3)}{\partial u_1\partial u_2\partial u_3}p(x_3)$$
$$p(x_3|x_1,x_2)=\frac{\partial^3 C(u_1,u_2,u_3)}{\partial u_1\partial u_2\partial u_3}\frac{1}{\frac{\partial^2 C(u_1,u_2,u_3)}{\partial u_1\partial u_2}}p(x_3)$$

Copula properties and definitions for general n-D case:

$$p(x_n|x_1,x_2,\ldots,x_{n-1})=$$
$$=\frac{\partial^n C(u_1,u_2,\ldots,u_n)}{\partial u_1,\ldots,\partial u_n}\frac{1}{\frac{\partial^{n-1} C(u_1,u_2,\ldots,u_n)}{\partial u_1,\ldots,\partial u_{n-1}}}p(x_n)$$

Conditional copula:

$$c(x_n|x_1,x_2,\ldots,x_{n-1})=$$
$$=\frac{\partial^n C(u_1,u_2,\ldots,u_n)}{\partial u_1,\ldots,\partial u_n}\frac{1}{\frac{\partial^{n-1} C(u_1,u_2,\ldots,u_n)}{\partial u_1,\ldots,\partial u_{n-1}}}$$

Density of copula:

$$\frac{\partial^n C(u_1,u_2,\ldots,u_n)}{\partial u_1,\ldots,\partial u_n}=c(u_1,u_2,\ldots,u_n)$$

2.2 Examples

When $u(x)$ is a marginal probability $F_{\mathrm{i}}(x), u(x) = F_i(x)$ and u is uniformly distributed, then the inverse function $x(u)$ is not uniformly distributed, but has values concentrated in the central part as in the Gaussian distribution. The inverse process is represented graphically in Figs. 1 and 2.

Consider another example where a joint probability density function p is defined in the two dimensional interval $[0, 2]\times[0, 1]$ as follows,

$$p(x, y) = \frac{x + y}{3}$$

Then the cumulative function in this interval is

$$C(x, y) = \int_{-\infty}^{x}\int_{-\infty}^{y} p(s, r)dsdr = \int_{-\infty}^{x}\int_{-\infty}^{y} \frac{s + r}{3}dsdr = \frac{xy(x + y)}{6} \tag{1}$$

Next we change the reference (x, y) into the reference (u_1, u_2). For

$$p(x_1) = \frac{du_1}{dx_1}, \quad p(x_2) = \frac{du_2}{dx_2}$$

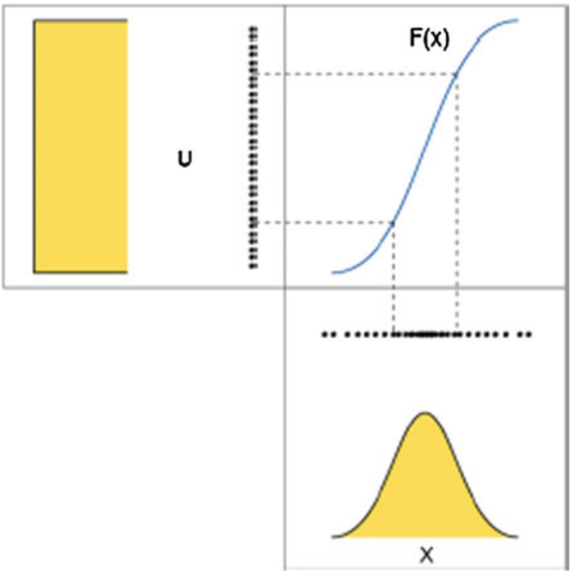

Fig. 1 Relation between marginal probability $F_{\mathrm{i}}(x)$ and the random variable x

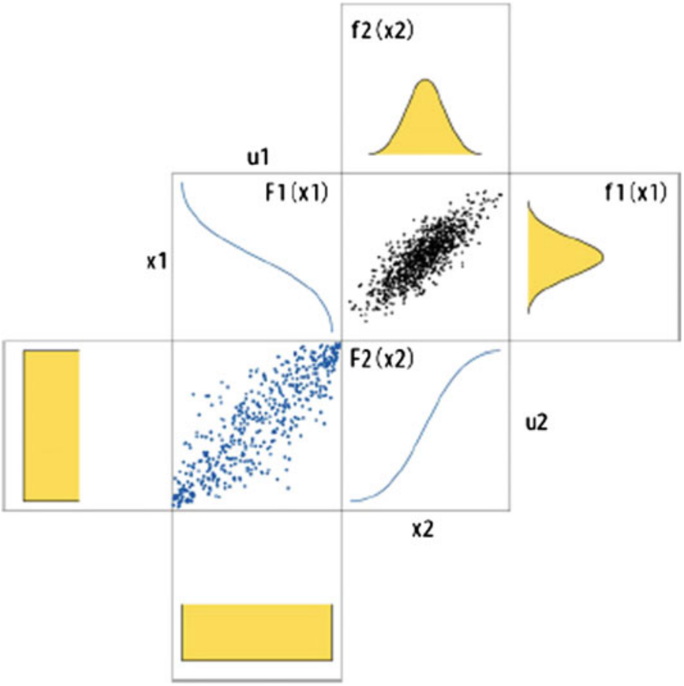

Fig. 2 Symmetric joint probability and copula

and use the marginal probabilities

$$u_1(x) = \int_{-\infty}^{x} p(s)ds, u_2(y) = \int_{-\infty}^{y} p(r)dr$$

to get

$$u_1(x) = C(x, 1) = \frac{x(x+1)}{6}, \; x \in [0, 2]$$
$$u_2 = C(2, y) = \frac{y(y+2)}{6} \; y \in [0, 1]$$

This allows us to compute the inverse function to identify the variables x and y as functions of the marginal functions u_1 and u_2:

$$x(u_1) = \frac{\sqrt{1+24u_1} - 1}{2} \quad y(u_2) = \sqrt{1+3u_2} - 1$$

Then these values are used to compute the copula C in function (1),

$$C(u_1, u_2) = \frac{\sqrt{1+24u_1}-1}{2}(\sqrt{1+3u_2}-1)\frac{\dfrac{\sqrt{1+24u_1}-1}{2}+(\sqrt{1+3u_2}-1)}{6}$$
$$= \frac{\sqrt{1+24u_1}-1}{2}(\sqrt{1+3u_2}-1)\frac{\dfrac{\sqrt{1+24u_1}-1}{2}+(\sqrt{1+3u_2}-1)}{6} \quad (2)$$

3 Approach

3.1 Meaning of Copulas

The main idea of the copula approach is *transforming a multivariate probability distribution* $F(x_1, x_2, \ldots, x_n)$ into a "normalized" multivariate distribution called a *copula*. The normalization includes transforming each of its arguments x_i into the [0, 1] interval and transforming the marginal distributions $F_i(x_i)$ into uniform distributions. The main result of the copula approach (Sklar's Theorem [24, 35, 36]) is the basis for this. It splits modeling the marginal distributions $F_i(x)$ from the modeling of the *dependence structure* that is presented by a normalized multivariate distribution (copula).

What is benefit of splitting? The copula is invariant under a strictly increasing transformation. Thus variability of copulas is less than the variability of n-D probability distributions. Therefore the selection of the copula can be simpler than the selection of a generic n-D probability distribution.

Next building a multivariate probability distribution, using the copula, is relatively simple in two steps [37, 38]:

(1) Selecting the univariate margins by transforming each of the one-factor distributions to be a uniform by setting $u_i = F_i(x_i)$ where the random variable u_i is from [0, 1].
(2) Selecting a copula connecting them by using the formula (3).

$$C(u_1, u_2, \ldots, u_n) = F^{-1}(F_1^{-1}(u_1), F_2^{-1}(u_2), \ldots, F_n^{-1}(u_n)) \quad (3)$$

These selections are done by using available data and a collection of parametric models for the margins $F_i(x_i)$ and the copula to fit the margins and the copula parameters with that data.

Next the property $\dfrac{\partial^n C(u_1, u_2, \ldots, u_n)}{\partial u_1, \ldots, \partial u_n} = c(u_1, u_2, \ldots, u_n)$ is used to produce pdf

$$p(x_1, x_2, \ldots, x_n) = c(u_1, u_2, \ldots, u_n)p(x_1)p(x_2)\ldots p(x_n)$$

where c is a copula density. Another positive aspect of using copulas is the abilities to generate a significant number of samples of n-D vectors that satisfy the copula probability distribution. After a copula in a "good" analytical form is approximated from the limited given data, a copula distribution can be simulated to generate more samples of data distributed according to it.

This is a known practical value of copulas. Does it ensure the quality/value of the produced multivariate distribution? It is noted in [24] that copulas have limited capabilities for constructing *useful* and *well–understood* multivariate pdfs, and much less for multivariate stochastic processes, quoting [39]: "Religious Copularians have *unshakable faith* in the value of transforming a multivariate distribution to its copula". These authors expect future techniques that will go beyond copulas. Arguments of heated discussions on copulas can be found in [40]. The question "which copula to use?" has no obvious answer [24]. This discussion reminds us a similar discussion in the fuzzy logic community on justification of t-norms in fuzzy logic.

3.2 Concept of NTF-Logical Dependence

Dependences traditionally are modeled by two distinct types of techniques based on the probability theory or logic (classical and/or many-valued). It is commonly assumed that dependencies are either stochastic or non-stochastic. Respectively it is assumed that if the dependence is not stochastic then it must be modeled by methods based on the classical and many-valued logic not by methods based on the probability theory.

This is an unfortunate and *too narrow* interpretation of probabilistic methods. In fact Kolmogorov's measures (K-measures) from the probability theory allow modeling both stochastic and non-stochastic "logical" dependencies (both classical and multi-valued). The basis for this is simply considering these measures as general measures of dependence, not as the frequency-based valuations of dependence. For instance, we can measure the dependence between two circles by computing a ratio S/T, where S is the area of their overlap and T is the total area of two circles. This ratio satisfies all of Kolmogorov's axioms of the probability theory, which is a "deterministic" method to measure the dependence in this example without any frequency.

Moreover, many dependencies are a mixture of stochastic and non-stochastic components. These dependencies also can be modeled by K-measures. Below we show this. Consider a dependency that is a mixture of stochastic and non-stochastic components. We start from a frequency-based joint probability distribution $F(x_1, \ldots, x_n)$ and *extract* from it a *dependence* that is not a *stochastic* dependence, but is a "*logical",* "deterministic" dependence. How do we define a logical dependence?

We will call dependence between variables $x_1, \ldots, x_n$ a Non-Truth-Functional *Logical Dependence* (***NTF-logical dependence***) if it is derived from F, but *does not* depend on the frequencies of occurrence of $x_1, \ldots, x_n$.

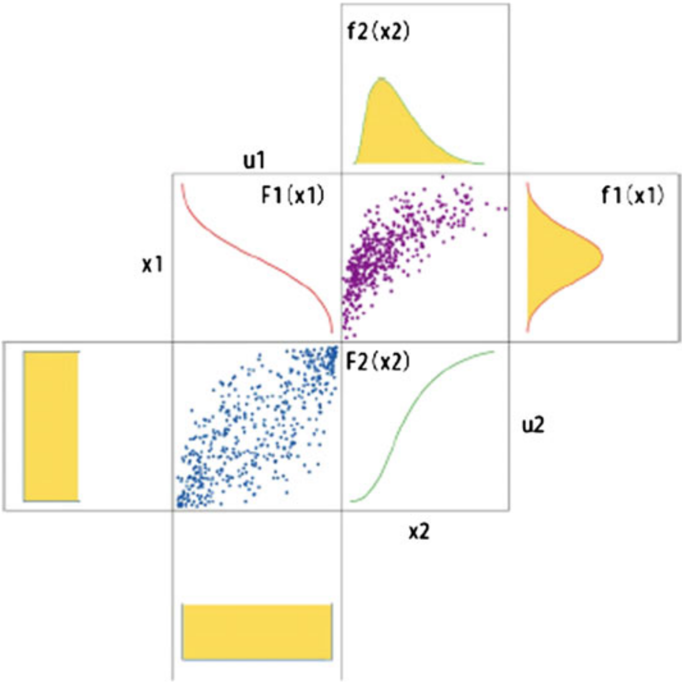

Fig. 3 Asymmetric joint probability and symmetric copula

How do we capture this logical dependence? We use a copula approach outlined above. First we convert (*normalize*) all marginal distributions $F_i(x_i)$ of F to uniform distributions $u_1, \ldots, u_n$ in [0, 1]. This will "eliminate" the dependence from frequencies of variables and differences in scales of variables. Next we build a copula: dependence C between $u_1, \ldots, u_n$, C: $\{(u_1, \ldots, u_n)\} \rightarrow [0, 1]$.

Figure 3 illustrates the difference between two dependencies. The first one is captured by F and the second one is captured by C. Copula C tells us that the logical dependence between x_1 and x_2 is symmetric, but differences in frequencies make the dependence asymmetric that is captured by F.

A similar idea is used in *machine learning* and *data mining* for discovering dependencies in data with *imbalanced classes*. For instance, in cancer diagnostics we can get many thousands of benign training cases and fortunately a much less number of malignant training cases. The actual discovery of the dependency (diagnostic rule) is conducted by selecting a similar number of training cases from both classes [41] which can be done by "normalizing" the marginal distributions of training data to the uniform distributions in each class in a coordinated way.

If F is not based on the frequencies then the interpretation of the marginal distributions is not that obvious. We still can formally convert marginal distributions $F_i(x_i)$ to uniform distributions and analyze two distributions F and C. In this case both distributions can be considered as logical dependencies. Here the copula will indicate dependence where some properties F are eliminated (lost). The usefulness of such elimination must be justified in each particular task. In essence we should show that eliminated properties are not about the mutual dependence of the variables, but are about other aspects of the variables, or even completely irrelevant, insignificant or "noise".

Note that the defined logical dependence is not expressed in the traditional logic form, e.g., as a propositional or first order logic expression, but as a copula, that is a specialized probability distribution. In fact, the copula corresponds to a set of logical expressions

$$\{(x_1 = a_1)\&(x_2 = a_2)\&\ldots\&(x_\mathrm{n} = a_\mathrm{n}) \text{ with probability } c(x_1, x_2, \ldots, x_\mathrm{n})\}$$

where c is a copula density.

The *interpretation* of copula as NTF-dependence is one of the key new points of this work showing how dependence in many-valued logic and probability theory can be fused in one concept by using copulas and marginal probabilities to clarify the relations between PT, FL and MVL. The next sections elaborate this approach.

3.3 Copula and Conditional Probability

Conditional probabilities capture some aspects of dependence between variables that include both frequency-based and logic based aspects. The relations between copula density and conditional probability are as follows as has been shown in Sect. 2:

$$\begin{array}{ll} p(x_2|x_1) = c(u_1, u_2)p(x_2) & p(x_1|x_2) = c(u_1, u_2)p(x_1) \\ p(x_2|x_1)/p(x_2) = c(u_1, u_2) & p(x_1|x_2)/p(x_1) = c(u_1, u_2) \end{array}$$

This shows that copula density as a "frequency-free" dependence can be obtained from frequency-dependent conditional probabilities $p(x_1|x_2)$ by "normalizing" it. We also can see that

$$p(x_2|x_1) \le \mathrm{c}(u_1, u_2)$$
$$p(x_1|x_2) \le \mathrm{c}(u_1, u_2)$$

because $p(x_2) \le 1$. In other words, frequency-free measure of dependence is no less than max of conditional probabilities:

$$\max(p(x_1|x_2), p(x_2|x_1)) \le \mathrm{c}(u_1, u_2)$$

In a logical form copula density represents a measure $c(u_1, u_2)$ of dependence for the conjunction,

$$(x_1 = a_1) \;\&\; (x_2 = a_2)$$

In contrast the conditional probabilities $p(x_1|x_2)$ and $p(x_2|x_1)$ represent measures of dependence for the implications:

$$(x_1 = a_1) \Rightarrow (x_2 = a_2), \qquad (x_2 = a_2) \Rightarrow (x_1 = a_1).$$

In a general case the relation between conditional probability and copula involves conditional density of copula as shown in Sect. 2.

$$p(x_{\mathrm{n}}|x_1, \ldots, x_{\mathrm{n}-1}) = c(x_{\mathrm{n}}|x_1, \ldots, x_{\mathrm{n}-1})\, p(x_{\mathrm{n}})$$

This conditional density of copula also represents a frequency-free dependence measure.

3.4 Copulas, t-norms and Fuzzy Logic

Many copulas are t-norms [42], which are the basis of many-valued logics and fuzzy logic [31, 32]. Many-valued logics, classical logic, and fuzzy logic are *truth-functional*: "…the truth of a compound sentence is determined by the truth values of its component sentences (and so remains unaffected, when one of its component sentences is replaced by another sentence with the same truth value)" [31]. Only in very specific and simple for modeling cases, probabilities and copulas are made truth-functional. This is a major difference from t-norms that are truth-functional. The impact of the truth functionality on fuzzy logic is shown in the Bellman-Giertz theorem [43]. Probability theory, fuzzy logic and many-values logics all used t-norms, but not in the same way.

The concept of copula and t-norm are not identical. Not every copula is a t-norm [42] and only t-norms that satisfy the 1-Lipschitz condition are copulas [44]. Equating t-norms with copulas would be equating truth-functional fuzzy logic with probability theory that is not truth-functional. A joint probability density $p(x_1, x_2)$ in general cannot be made a function of only $p(x_1)$ and $p(x_2)$. A copula density or a conditional probability is required in addition due to properties $p(x_1, x_2) = c(x_1, x_2) \quad p(x_1)p(x_1)$ and $p(x_1, x_2) = p(x_1|x_2)p(x_2)$. In this sense the *probability theory is more general than fuzzy logic*, because it covers both truth-functional and non-truth functional dependences. Thus, in some sense fuzzy logic with t-norms represents the "*logic of independence*". On the other side, the copula gives a probabilistic interpretation of the fuzzy logic t-norms as well as interpretation of fuzzy logic theory as a many-valued logic.

Postulating a t-norm is another important aspect of the differences in the usage of copulas in FL and PT. In FL and possibility theory, a t-norm, e.g., $\min(x_1, x_2)$ is often postulated as a *"universal"* one for multiple data sets, without the empirical justification for specific data.

In PT according to Sklar's Theorem [24, 35, 36], the copula $C(\mathrm{u}_1, \ldots, \mathrm{u}_{\mathrm{n}})$ is *specific/unique* for each probability distribution F with uniform marginals. If this copula happened to be a t-norm, then it is a unique t-norm, and is also *specific* for that probability distribution.

This important difference was also pointed out in [45]: "In contrast to fuzzy techniques, where the same "and"-operation (t-norm) is applied for combining all pieces of information, the vine copulas allow the use of different "and"-operations (copulas) to combine information about different variables." Note that the copula approach not only allows different "and"-operations but it will actually get different "and"-operations when underlying joint probability distributions differ. Only for identical distributions the same "and"-operations can be expected.

The t-norm $\min(x_1, x_2)$ is a dominant t-norm in FL and possibility theory. This t-norm is known in the copula theory as a *comonotonicity copula* [46]. It satisfies a property

$$\mathrm{P}(\mathrm{U}_1 \le \mathrm{u}_1, \ldots, \mathrm{U}_\mathrm{d} \le \mathrm{u}_\mathrm{d}) = \mathrm{P}(\mathrm{U} \le \min\{\mathrm{u}_1, \ldots, \mathrm{u}_\mathrm{d}\}) = \min\{\mathrm{u}_1, \ldots, \mathrm{u}_\mathrm{d}\}, \mathrm{u}_1, \ldots, \mathrm{u}_\mathrm{d} \in [0, 1]$$

where a *single* random uniformly distributed variable $\mathrm{U} \sim \mathrm{U}[0, 1]$ produced the random vector $(U_1, \ldots, U_\mathrm{d}) \in [0, 1]^\mathrm{d}$. The variables of the comonotonicity copula are called *comonotonic* or *perfectly positively dependent variables* [47].

3.5 Forms of Dependence Based on Common Cause and Direct Interaction

Forms of dependence. Dependence can take multiple forms. One of them is dependence with *common source/cause*, i.e., a third object S is a cause of the dependence between A and B. We will call it as *third-party dependence*. In dancing music, S serves as the third party that synchronizes dancers A and B. In physics such a type of dependence is often apparent in synchronization of events, e.g., in laser photon synchronization the source for the photon synchronization is interaction of the photons in a crystal.

Technically the dependencies between A and B in these cases can be discovered by using the correlation methods, which includes computing the copulas and the conditional probabilities $p(A|B)$, $p(B|A)$. However, correlation methods do not uncover the *type of dependency*, namely *third-party* dependence or dependence as a result of *direct local interaction* of objects A and B. The last one is a main type of laws in the classical physics.

Historically correlation based dependencies (e.g., Kepler's laws) were augmented later by more causal laws (e.g., Newton's laws). It was done via accumulation of more knowledge in the physics area and much less by more sophisticated mathematical methods for discovering the causal dependencies. This is an apparent now in constructing Bayesian networks (BN) where causal links are mostly built manually by experts in the respective field. Links build by discovering correlations in data are still may or may not be causal.

4 Physics and Expanding Uncertainty Modeling Approaches

This section provides physics arguments for expanding uncertainty modeling approaches beyond current probability theory and classical and non-classical logics to deal with dependent/related evens under uncertainty. After that we propose a way of expanding uncertainty modeling approaches. The first argument for expansion is from the area of duality of particles and waves and the second one is from the area of Bell's inequality [48, 49]. The dependencies studied in the classical physics can be described using the classical logic and the classical set theory.

Logic operates with true/false values assigned to states of particles. To establish the true state of a particle without involving fields it is sufficient to know the state of the individual particle without considering all the others particles. In this sense particles are independent one from the others. In this sense the classical logic and the associated set theory are the conceptual instruments to study classical physics. In quantum physics with the "entanglement" the states of all particles are "glued" in one global entity where any particle is sensitive or dependent to all the others. In this sense the information of any individual particle (local information) is not sufficient to know the state of the particle. We must know the state of all the other particles (global information) to know the state of one particle. The physical phenomena that are studied in the quantum physics have more complex (non-local) dependencies where one physical phenomenon is under the influence of *all* the other phenomena both local and non-local. These dependencies have both stochastic probabilistic component (that can be expressed by a joint probability) and a logic component that we express by copula.

4.1 Beyond Propositional Logic and Probability Theory

Particles and waves are classically considered as distinct and incompatible concepts because particles are localized and waves are not [50]. In other words, particles interact locally or have *local not global dependence*, but waves interact *non-locally* possibly with all particles of its media [51]. The classical propositional logic and set theory have no tools to deal with this issue and new mechanisms are needed to deal with it. The global dependence (interaction with non-local elements of the media) is a property of the structure of the media (whole system). Feynman suggested a concept of the negative probability that later was developed by Suppes and others [51–54] to address such issues.

Limitations of probability theory. Kolmogorov's axioms of PT deal only with *one type of dependence*: dependence of sets elementary events in the form of their *overlap* that is presence of the common elements in subsets. This is reflected in the formula for probability of the union $A \cup B$:

$$p(A \cup B) = p(A) + p(B) - p(A \cap B).$$

where dependence is captured by computing $p(A \cap B)$ and subtracting it to get $p(A \cup B)$.

While capturing overlap is a huge *advantage* of PT over truth-functional logics (including fuzzy logic and possibility theory) that do not take into account overlap $A \cap B$. However, axioms of the probability theory do not include other dependencies. In fact, PT assumes *independence* of each elementary event from other elementary events, requiring their overlap to be empty, $\mathrm{e_a} \cap \mathrm{e_b} = \emptyset$.

What are the options to deal with the restriction of the probability theory in the tasks where more dependencies must be taken into account?

We have two options:

- The first option is to *choose other elementary events* that will be really independent.
- The second one is *incorporating relations* into a new theory.

The first option has difficulty in quantum physics. Particles are dependent and cannot serve as independent elementary events. This creates difficulties for building a classical probabilistic theory of particles. Next, at a particular time, we simply may have no other candidates for elementary events that would be more elementary than particles.

Generalization of Probability The previous analysis shows that we need to explore the second option: adding relations between elementary events and their sets. Below we propose a generalization of the probability for dependent events that we call **relation-aware probability**.

Let $\{R_i^+(A, B)\}$ be a set of "positive relations" and $\{R_j^-(A, B)\}$ be a set of "negative relations" between sets A and B. Positive relations increase probability $p(A \cup B)$ and negative relations decreases probability $p(A \cup B)$.

Next we introduce probability for relations $\{p(R_i^+(A, B))\}$, $\{p(R_j^-(A, B))\}$ and define the **relation-aware probability** of the union:

$$p\,(A \cup B) = p\,(A) + p\,(B) + \sum_{i=1}^{n} p(R_i^+\,(A, B)) - \sum_{j}^{m} p(R_j^-(A, B))$$

The classical formula

$$p\,(A \cup B) = p\,(A) + p\,(B) - p(A \cap B)$$

is a special case of this probability with the empty set of positive relations and with a single negative relation, $R\,(A, B) \equiv A \cap B \neq \emptyset$ that is intersection of A and B is not empty. A negative relation can work as "annihilator". A positive relation between A and B can work as a catalyst to increase A or B. For instance, we can get

$$p\,(A \cup B) = p\,(A) + p\,(B) + 0.2p\,(A) + 0.1p(B) - p(A \cap B).$$

As a result, the relation-aware probability space can differ from Kolmogorov's probability space having the more complex additivity axioms.

Another potential approach is associated with changing the classical logic into logics that allow conflicts, e.g., fuzzy logic, agent model of uncertainty [8], and active sets [12]. This will make the paradox more *apparent* due to abilities to embed the conflict via gradual measures of uncertainty such as known in fuzzy logic, and many-valued logic.

4.2 Beyond the Set Theory: Bell Inequality

Consider sets S_1-S_8 formed by three circles A, B and C as shown in Fig. 4, e.g., $S_1 = A \cap B \cap C$ and $S_8 = A^C \cap B^C \cap C^C$ for complements of A, B and C. These sets have the following classical set theory properties that include the *Bell's inequality*

$$
\begin{aligned}
&(A \cap B^C) \cup (B \cap C^C) = S_1 \cup S_7 \cup S_4 \cup S_6 \\
&A \cap C^C = S_7 \cup S_4 \\
&A \cap C^C \subseteq (A \cap B^C) \cup (B \cap C^C) \\
&\left|A \cap C^C\right| \leq \left|(A \cap B^C) \cup (B \cap C^C)\right|
\end{aligned}
$$

Next consider the *correlated* (*dependent*) events, for instance we can have an event with the property A, and another event with a negated property A^C. This is a type of dependence that takes place for particles. Consider an event with property $A \cap C^C$ that is with both properties A and C^C at the same time. We cannot measure the two properties by using one instrument at the same time, but we can use the correlation to measure the second property, if the two properties are correlated. We can also view an event with property $A \cap C^C$ as two events: event e_1 with property A and e_2 with the property C in the opposite state (negated). The number of pairs of events (e_1, e_2) is the same as the number of events with the superposition of A and C^C, $A \cap C^C$. In [49] d'Espagnat explains the connection between the set theory

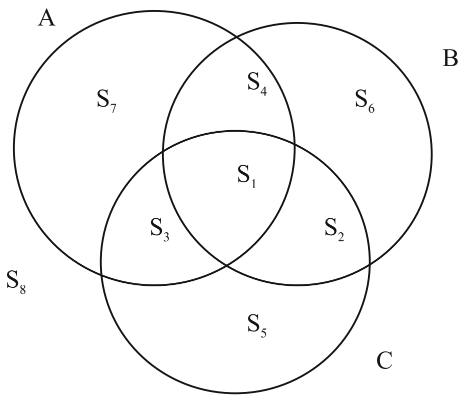

Fig. 4 Set theory intersections or elements

and Bell's inequality. It is known that the Bell's inequality that gives us the reality condition is violated [48].

How to reconcile the logic Bell's inequality and its violation in physics? The Bell inequality is based on the classical set theory that is consistent with the classical logic.

The set theory assumes *empty overlap* (as a form of *independence*) of elementary units. For dependent events the Bell's inequality does not hold. Thus the logic of dependence (logic of dependent events) must differ from the logic of independence (logic of independent events) to resolve this inconsistence. It means that we should use a theory that is *beyond* the classical set theory or at least use the classical set theory differently. The dependence logic [14, 15] is one of the attempts to deal with these issues.

4.3 Dependence and Independence in the Double Slit Experiment

The double slit experiment involves an electron gun that emits electrons through a double slit to the observation screen [55, 56]. Figure 5 [51] shows theoretical result of the double slit experiment when only the set theory is used to combine events: one event e_1 for one slid and another event e_2 for the second slid. In this set-theoretical approach it is assumed that events e_1 and e_2 are *elementary events* that do not overlap, $(e_1 \cap e_2) = \emptyset$, independent). In this case, the probability $p\,(e_1 \cap e_2) = 0$ and the probability that either one of these two events will occur is $p\,(e_1 \cup e_2) = p\,(e_1) + p(e_2)$. The actual distribution differs from distribution shown in Fig. 5 [51].

As an attempt to overcome this dependence difficulty and related non-monotonicity in quantum mechanics the use of *upper probabilities*, with the axiom subadditivity, has been proposed [51–54].

Below we explore whether the *copula* approach and *relation-aware probability* can be helpful in making this dependence *evident* and *explicit*. The relation-aware probability is helpful, because it explicitly uses the relations that are behind the dependences. This is its advantage relative to negative probabilities. The axioms of negative probability have no such components.

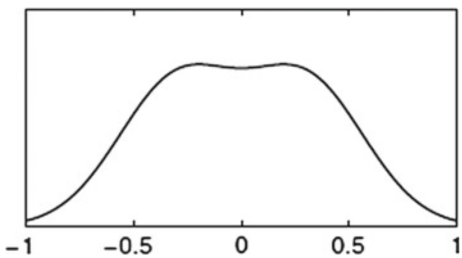

Fig. 5 [51]. Distribution of independent particles (events)

The value of the copula approach for this task needs more explorations. Copula allows representing the actual distribution F as a composition that includes the copula. The actual distribution is [51]

$$p(\alpha_1, \alpha_2) = k\cos^2(\alpha_1 - \alpha_2)$$

Respectively its copula is

$$C(F(\alpha_1), F(\alpha_2)) = k\sin^2[A \cdot B - C \cdot D] = C(u_1, u_2) \tag{4}$$

$$A = \frac{1}{2}\left(\frac{\frac{\pi}{2} \pm \sqrt{(\frac{\pi}{2})^2 + \frac{16}{\pi}\arccos(\frac{1}{k}\sqrt{F_1(\alpha_1)})}}{2}\right)^2$$

$$B = \frac{\frac{\pi}{2} \mp \sqrt{(\frac{\pi}{2})^2 + \frac{16}{\pi}\arccos(\frac{1}{k}\sqrt{F_2(\alpha_2)})}}{2}$$

$$C = \left(\frac{\frac{\pi}{2} \mp \sqrt{(\frac{\pi}{2})^2 + \frac{16}{\pi}\arccos(\frac{1}{k}\sqrt{F_2(\alpha_2)})}}{2}\right)^2$$

$$D = \frac{\frac{\pi}{2} \pm \sqrt{(\frac{\pi}{2})^2 + \frac{16}{\pi}\arccos(\frac{1}{k}\sqrt{F_1(\alpha_1)})}}{2}$$

This copula represents the dependence with the eliminated impact of marginal distributions via conversion of them into the uniform distribution. In this way the copula has the meaning of *graduate dependence between variables* resembling the many-valued logic, and t-norm compositions of variables. This is a valuable property. Thus while the copula is still a probability distribution it also represents a gradual degree of truth and falseness. The copula (5) is tabulated as follows:

$$\mathrm{M1} = \begin{pmatrix} 1 & 0.846 & 0.7 & 0.561 & 0.43 & 0.307 & 0.192 & 0.088 & 0 \\ 0.846 & 1 & 0.969 & 0.899 & 0.805 & 0.692 & 0.56 & 0.401 & 0.125 \\ 0.7 & 0.969 & 1 & 0.979 & 0.923 & 0.839 & 0.725 & 0.572 & 0.25 \\ 0.561 & 0.899 & 0.979 & 1 & 0.982 & 0.93 & 0.843 & 0.708 & 0.375 \\ 0.43 & 0.805 & 0.923 & 0.982 & 1 & 0.982 & 0.927 & 0.82 & 0.5 \\ 0.307 & 0.692 & 0.839 & 0.93 & 0.982 & 1 & 0.98 & 0.909 & 0.625 \\ 0.192 & 0.56 & 0.725 & 0.843 & 0.927 & 0.98 & 1 & 0.973 & 0.75 \\ 0.088 & 0.401 & 0.572 & 0.708 & 0.82 & 0.909 & 0.973 & 1 & 0.875 \\ 0 & 0.125 & 0.25 & 0.375 & 0.5 & 0.625 & 0.75 & 0.875 & 1 \end{pmatrix}.$$

It has the same form in its extreme values as classical logic equivalence relation ($\equiv$).

$$y = x_1 \equiv x_2 = \begin{bmatrix} & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

This link is also evident from [42, p. 182]:

$$\mathrm{C(x, 0) = C(0, x) = 0, C(x, 1) = C(1, x) = x}$$

which is true for all x from [0, 1] including x = 1.

5 Conclusion and Future Work

This paper had shown the need for a theory *beyond* the classical set theory, classical logic, probability theory, and fuzzy logic to deal with dependent events including the physics paradox of particles and waves and the contradiction associated with the Bell inequality.

It is shown that the *copula* approach and *relation-aware probability* approach are promising approaches to go beyond the limitations of known approaches of PT, FL, MVL, and the classical set theory. Specifically *copula* and *relation-aware probability* approaches allow making dependences between variables in multidimensional probability distributions more *evident* with abilities to model the dependencies and relations between multiple variables more e*xplisit*.

It is pointed out that the concept of copula is *more general* than the concept of t-norm. Therefore the probability theory in some sense is *more general* than fuzzy logic, because it covers both truth-functional and non-truth functional dependences.

The *third-party dependence, direct local interaction* and causal dependence are analyzed concluding that discovering causal dependencies is far beyond current approaches in PT, FL, and MVL.

We introduced the concept of a Non-Truth-Functional *Logical Dependence* (***NTF-logical dependence***) derived as a copula from multidimensional probability distribution that does not depend on the frequencies of occurrence of variables. The interpretation of copula as NTF-dependence is one of the key of new points of this work. It shows how dependence in many-valued logic and probability theory can be fused in one concept by using the copulas and marginal probabilities to clarify the relations between PT, FL, and MVL to give a "frequency-free" dependence measure. It shows how, starting from probability and specifically from conditional probability, we introduce "logic" dependence.

We conclude that presence of t-norms in both probability theory (in copulas), and in fuzzy logic, creates a new opportunity to join them using t-norm copulas as a bridge.

Future directions are associated with linking the proposed approaches with dependence logic. Dependence logic proposed in 2007 [14] was later expanded to the contexts of modal, intuitionistic and probabilistic logic. Its goal is modeling the dependencies and interaction in dynamical scenarios. It is a type of first-order logic, and a fragment of second-order logic. While it claims applications in a wide range of domains from quantum mechanics to social choice theory, it still is in earlier stages of applications.

Another future opportunity to build the more realistic uncertainty modeling approaches is associated with the involvement of agents. Agents and teams of agents often are a source of input data, and the values of probabilities, copulas, and t-norms. The situation, when these agents are in a total agreement, can be modeled as a *total dependence* of their valuations, graduated or discrete (True/False). When a team of agents is in a conflict with each other in the valuation of the same attribute, they provide conflicting degrees of truth/falseness, which require adequate modeling approaches [8]. Active sets [12] that represent sets of agents can be a bridge between the many-valued logic, and fuzzy sets for providing the interpretable input data (valuations).

References

1. B. Kovalerchuk, and D.Shapiro, "On the relation of the probability theory and the fuzzy sets theory, Foundation Computers and Artificial Intelligence, Vol. 7, pp. 385–396, 1988.
2. D. Dubois, and H. Prade "Fuzzy sets, probability and measurement," Europ. J. of Operations Research, Vol. 40, pp. 135–154, 1989.
3. D.I. Blockley, "Fuzziness and probability: a discussion of Gaines' axioms," Civ. Engng Syst., Vol. 2, pp. 195–200, 1985.
4. P. Cheeseman, "Probabilistic versus fuzzy reasoning," in Uncertainty in Artificial Intelligence 1, L. Kanal, and J. Lemmer, Eds., North-Holland, Amsterdam, 1988, pp. 85–102
5. B.R. Gaines, "Fuzzy and probability uncertainty logics," Information and Control, Vol. 38, pp. 154–169, 1978.
6. G.J. Klir, and B. Parviz, "Probability-possibility transformations: a comparison," Int. J. of General Systems, Vol. 21, pp. 291–310, 1992.
7. B. Kosko, "Fuzziness vs. probability," Int. J. of General Systems, Vol. 17, pp. 211–240, 1990.
8. G. Resconi, B. Kovalerchuk: Agents' model of uncertainty. Knowl. Inf. Syst. 18(2): 213–229, 2009.
9. G. Resconi, B. Kovalerchuk: Agents in neural uncertainty. IJCNN 2009, pp. 2649–2656
10. G. Resconi, B. Kovalerchuk: Fusion in agent-based uncertainty theory and neural image of uncertainty. IJCNN 2008: pp. 3538–3544, 2007
11. G. Resconi, B. Kovalerchuk: Explanatory Model for the Break of Logic Equivalence by Irrational Agents in Elkan's Paradox. EUROCAST 2007: 26–33
12. G. Resconi, C. J. Hinde: Introduction to Active Sets and Unification. T. Computational Collective Intelligence 8: 1–36, 2012.
13. Graadel, E., Vaananen, J.: Dependence and independence. Studia Logica, 101(2), pp. 399–410, 2013.
14. J.Vaananen, Dependence Logic - A New Approach to Independence Friendly Logic, Cambridge University Press, 2007.
15. J. Ebbing, A. Hella, J-S. Meier, J. Muller, J. Virtema and H. Vollmer, Extended modal dependence logic, in: WoLLIC, 2013, pp. 126–137.
16. J. Ebbing, P. Lohmann, Complexity of model checking for modal dependence logic, in: M. Bieliková, G. Friedrich, G. Gottlob, S. Katzenbeisser and G. Turán, editors, SOFSEM, LNCS 7147, 2012, pp. 226–237.

17. J. Ebbing, P. Lohmann and F. Yang, Model checking for modal intuitionistic dependence logic, in: G. Bezhanishvili, S. Lobner, V. Marra and F. Richter, eds., Logic, Language, and Computation, Lecture Notes in Computer Science 7758, Springer, 2013 pp. 231–256.
18. P. Galliani, The dynamification of modal dependence logic, Journal of Logic, Language and Information 22 (2013), pp. 269–295.
19. J. Hintikka, and G. Sandu, Informational independence as a semantical phenomenon, in: Logic, methodology and philosophy of science, VIII (Moscow, 1987), Stud. Logic Found. Math. 126, North-Holland, Amsterdam, 1989 pp. 571–589.
20. J.-S. Müller, and H. Vollmer, Model checking for modal dependence logic: An approach through post's lattice, in: L. Libkin, U. Kohlenbach and R. Queiroz, editors, Logic, Language, Information, and Computation, Lecture Notes in Computer Science 8071, 2013 pp. 238–250.
21. F. Yang, "On Extensions and Variants of Dependence Logic," Ph.D. thesis, University of Helsinki, 2014.
22. I.R. Goodman, Fuzzy sets as equivalence classes of random sets, In: Yager, R., et al. (eds.), Fuzzy Sets and Possibility Theory, Oxford, UK: Pergamon Press, 327–432, 1982.
23. H. T. Nguyen, V. Kreinovich, "How to Fully Represent Expert Information about Imprecise Properties in a Computer System–Random Sets, Fuzzy Sets, and Beyond: An Overview", International Journal of General Systems, 2014, 43(5–6), pp. 586–609.
24. P. Embrechts, F.Lindskog, A. McNeil: Modelling dependence with copulas and applications to risk management. In *Handbook of heavy tailed distributions in finance*, Rachev ST, ed. Elsevier/North-Holland, Amsterdam, 2003.
25. P. Jaworski, F. Durante, W. K. Härdle, T. Rychlik (Eds), Copula Theory and Its Applications, Lecture Notes in Statistics, Springer, 2010.
26. J. C. Fodor and M. Roubens: Fuzzy Preference Modelling and Multicriteria Decision Support, Kluwer, Dordrecht 1994.
27. O. Hadžić, E. Pap: Fixed Point Theory in Probabilistic Metric Spaces. Kluwer, Dordrecht 2001.
28. B. Schweizer, A. Sklar: Probabilistic Metric Spaces. North-Holland, New York, 1983.
29. E. Klement, R. Mesiar and E. Pap: Triangular Norms. Kluwer, Dordrecht 2000.
30. E. Klement; R. Mesiar; E. Pap, Invariant copulas, Kybernetika, Vol. 38 2002, No. 3, 275–286.
31. Many-Valued Logic, Stanford Encyclopedia of Philosophy, 2015, http://plato.stanford.edu/entries/logic-manyvalued/#TNorBasSys
32. P. Hájek, Metamathematics of Fuzzy Logic, Dordrecht: Kluwer, 1998.
33. K. Aas, C. Czado, A. Frigessi, and H. Bakken. Pair-copula constructions of multiple dependencies. Insurance: Mathematics and Economics, 44:182–198, 2009.
34. R. Accioly and F. Chiyoshi. Modeling dependence with copulas: a useful tool for field development decision process. Journal of Petroleum Science and Engineering, 44:83–91, 2004.
35. A. Sklar. Fonctions de repartition a n dimensions et leurs marges. Publications de l'Institut de Statistique de L'Universite de Paris, 8:229–231, 1959.
36. A. Sklar: Random variables, joint distribution functions, and copulas. Kybernetika 9, 1973, 449–460.
37. J.-F. Mai, M. Scherer, "Simulating Copulas: Stochastic Models, Sampling Algorithms, and Applications", World Scientific, 2012, http://www.worldscientific.com/doi/suppl/10.1142/p842/suppl_file/p842_chap01.pdf
38. M.S. Tenney, "Introduction to Copulas", Enterprise Risk Management Symposium, 2003, Casualty Actuarial Society, http://www.casact.org/education/rcm/2003/ERMHandouts/tenney1.pdf
39. C. Klüppelberg, C. and S. Resnick, The Pareto copula, aggregation of risk and the Emperor's socks. Preprint, Technical University of Munich, 2009.
40. T. Mikosch, Copulas: tales and facts. Extremes, Vol. 9, Issue 1, pp. 3–20, Springer, 2006.
41. B. Kovalerchuk, E. Vityaev, J. Ruiz, Consistent and Complete Data and "Expert" Mining in Medicine, In: Medical Data Mining and Knowledge Discovery, Springer, 2001, pp. 238–280.
42. G. Beliakov, A. Pradera and T. Calvo, Aggregation Functions: A Guide for Practitioners, Springer, Heidelberg, Berlin, New York, 2007.

43. R. Bellman, M. Giertz. “On the Analytic Formalism of the Theory of Fuzzy Sets.” Information Sciences 5 (1973): 149–156.
44. M. Navara, Triangular norms and conorms, Scholarpedia, 2(3):2398, 2007. http://www.scholarpedia.org/article/Triangular_norms_and_conorms
45. S. Sriboonchitta, J. Liu, V. Kreinovich, and H.T. Nguyen, “Vine Copulas as a Way to Describe and Analyze Multi-Variate Dependence in Econometrics: Computational Motivation and Comparison with Bayesian Networks and Fuzzy Approaches”, In: V.-N. Huynh, V. Kreinovich, and S. Sriboonchitta eds.), Modeling Dependence in Econometrics, Springer, Berlin, 2014, pp. 169–187.
46. E. Jouini and C. Napp, Conditional comonotonicity, Decisions in Economics and Finance, 27 (2):153–166, 2004.
47. A. Patton, Applications of Copula Theory in Financial Econometrics, Dissertation, University of California, San Diego, 2002.
48. J. Bell, “On the Einstein Podolsky Rosen Paradox”. Physics 1 (3): 195–200, 1964
49. D’Espagnat Conceptual Foundations of Quantum Mechanics, 2nd ed. Addison Wesley
50. R. Feynman, R., Leighton, and M. Sands, The Feynman lectures on physics: Mainly mechanics, radiation, and heat, volume 3. Basic Books, NY, 2011.
51. J.A. de Barros, G. Oas and P. Supper, Negative probabilities and counterfactual reasoning on the double – slit experiment, [quant-ph] 16 Dec 2014 arXiv:1412.4888v1.
52. P. Suppes, M. Zanotti, Existence of hidden variables having only upper probabilities Foundations of Physics 21 (12):1479–1499 (1991)
53. J. A. de Barros, P. Suppes, “Probabilistic Inequalities and Upper Probabilities in Quantum Mechanical Entanglement,” Manuscrito, v. 33, pp. 55–71 (2010).
54. P. Suppes, S. Hartmann, Entanglement, upper probabilities and decoherence in quantum mechanics, In: M. Suaráz et al (ed.), EPSA Philosophical Issues in the Sciences: Launch of the European Philosophy of Science Association, Springer, pp. 93–103, 2010
55. P. Feynman, QED: The Strange Theory of Light and Matter. Princeton University Press, 1988.
56. Double-slit experiment, Wikipedia, http://en.wikipedia.org/wiki/Double-slit_experiment, 2015.