Conditional Independence Relations in Possibility Theory

Jiřina Vejnarová*
Laboratory of Intelligent Systems, Prague
vejnar@vse.cz

Abstract

The aim of this paper is to survey and briefly discuss various rules of conditioning proposed in the framework of possibility theory as well as various conditional independence relations suggested for these rules. These conditioning rules and conditional independence relations are confronted with formal properties of conditional independence. Special attention is paid to the conditioning rule based on measure-theoretical approach [3]. It is argued that this way of conditioning and the related conditional independence notion [12] not only generalize some of presented rules and conditional independence relations, but also their properties correspond to those possessed by stochastic conditional independence.

Keywords. Possibility measure, possibility distribution, conditioning rule, natural extension, conditional possibility distribution, possibilistic conditional independence, formal properties of conditional independence.

1 Introduction

Probability theory had been the only mathematical tool at our disposal for uncertainty quantification and processing for three centuries, and therefore many important theoretical as well as practical achievements have been obtained in this field. However, during the last thirty years some new mathematical tools alternative to probability theory have emerged. Their aims are to treat either the cases, when the nature of uncertainty in question does not meet the demands requested by probability theory, or the cases in which probabilistic approaches are based on too strong and hardly assurable (or even verifiable) conditions. Nevertheless, probability theory has always served as a source of inspiration for the development of these non-probabilistic calculi and they have been continually confronted with probability theory and mathematical statistics from various points of view. The topic of this paper is a good example of this fact.

Conditioning belongs to the most important features of any model of uncertainty and therefore it has been studied within possibility theory from its very beginning. In possibility theory, in contrary to the probabilistic framework, various rules were proposed to define conditional possibility measures (or distributions) from joint ones. But there exist no criteria along which these rules can be compared.

Conditional independence notion, on the other hand, is connected mainly with the application of uncertainty theories (namely probability theory) to artificial intelligence. Complexity of practical problems that are in the center of interest of artificial intelligence results usually in necessity to model the field of application with the help of a great number of variables, more precisely hundreds or thousands rather than tens. Processing distributions of such dimensionality would not be possible without some tools allowing to reduce demands on computer memory. Conditional independence, which belongs to such tools, allows to express these multidimensional distributions by means of lowdimensional ones and therefore to decrease substantially demands on the computer memory.

We will start with the notion of stochastic conditional independence. Supposing X, Y and Z are random variables with a joint probability distribution P we say that X is *conditionally independent* of Y given Z with respect to P and write $I_P(X, Y|Z)$ if the equality

$$P(x, y, z) \cdot P_Z(z) = P_{XZ}(x, z) \cdot P_{YZ}(y, z) \quad (1)$$

(where P_{XZ}, P_{YZ}, P_Z denote corresponding marginal distributions) holds for every value x, y, z of the variables X, Y, Z. It means that in every situation when the value of Z is known the values of X and Y are

*The work on this paper was supported by the grants MŠMT no. VS96008, GA ČR no. 201/98/1487 and GA AV ČR no. A1075801.

completely unrelated (from the stochastic point of view). The distribution P is usually fixed and therefore omitted from the notation.

There exist many equivalent definitions of stochastic conditional independence, e.g.

$$P_{X|YZ}(x|y,z) = P_{X|Z}(x|z), \quad (2)$$

but this definition may be used only in the situation when $P_{YZ}(y,z)$ is positive[1]. Nevertheless, possibilistic counterparts of this very equality are often taken for definition of possibilistic conditional independence as will be seen later.

Among the properties satisfied by the ternary relation $I(X,Y|Z)$ the following are of principal importance:

(A1) $I(X,Y|Z) \rightarrow I(Y,X|Z)$ symmetry,

(A2) $I(X,YZ|W) \rightarrow I(X,Z|W)$ decomposition,

(A3) $I(X,YZ|W) \rightarrow I(X,Y|ZW)$ weak union,

(A4) $[I(X,Y|ZW) \wedge I(X,Z|W)] \rightarrow I(X,YZ|W)$ contraction,

(A5) $[I(X,Y|ZW) \wedge I(X,Z|YW)] \rightarrow I(X,YZ|W)$ intersection.

It is well known (see e.g. [11]), that $I(X,Y|Z)$ defined by (1) or (2) satisfies (A1) – (A4). If P is strictly positive, then also (A5) is satisfied.

The properties (A1) – (A5) may be thought of as purely formal properties of the notion of conditional independence *without any connection to probability theory* (see e.g. [10]). In the mentioned book a few examples from different areas of mathematics as well as a nice example concerning reading books can be found.

In the present paper we will deal with some definitions of conditional possibility distributions and conditional independence and summarize which properties are satisfied by those independence relations. These properties, evidently, depend not only on the definition of the conditioning rule but also on the definition of the conditional independence relation, and therefore for one rule of conditioning we can obtain several independence relations possessing different properties.

[1] Let us note that we have adopted Kolmogorov axiomatic probability theory. In de Finetti's approach, (2) is fundamental equality and (1) holds only for positive probabilities.

2 Possibility Measures and Distributions

Let $\mathbf{X}$ be a finite set called *universe of discourse* which is supposed to contain at least two elements. A *possibility measure* Π is a mapping from the power set $\mathcal{P}(\mathbf{X})$ of $\mathbf{X}$ to the real unit interval $[0,1]$ satisfying the following requirement: for any family $\{A_j, j \in J\}$ of elements of $\mathcal{P}(\mathbf{X})$

$$\Pi(\bigcup_{j\in J} A_j) = \max_{j\in J} \Pi(A_j).^2$$

For any $A \in \mathcal{P}(\mathbf{X})$, $\Pi(A)$ is called the *possibility of* A. Π is called *normal* iff $\Pi(\mathbf{X}) = 1$.

For any Π there exists a mapping $\pi : \mathbf{X} \rightarrow [0,1]$ such that for any $A \in \mathcal{P}(\mathbf{X})$, $\Pi(A) = \max_{x\in A} \pi(x)$. π is called a *distribution* of Π. This function is a possibilistic counterpart of a density function in probability theory. It is evident, that Π is normal iff there exists at least one $x \in \mathbf{X}$ such that $\pi(x) = 1$.

Now, let us consider an arbitrary possibility measure Π defined on a product universe of discourse $\mathbf{X} \times \mathbf{Y}$. The *marginal possibility measure* is then defined by the equality

$$\Pi_X(A) = \Pi(A \times \mathbf{Y})$$

for any $A \subset \mathbf{X}$ and the *marginal possibility distribution* by the corresponding expression

$$\pi_X(x) = \max_{y\in\mathbf{Y}} \pi(x,y) \quad (3)$$

for any $x \in \mathbf{X}$. In what follows, we will omit the subscript if there are no doubts, which marginal we have in mind.

Normality of joint possibility measure implies normality of its marginals. Because normality seems to be quite a natural requirement, we will always suppose in the sequel that the joint possibility distribution is normal and discuss whether the conditional distribution is also normal.

3 Conditioning and Independence: an Overview

In this section we will overview various definitions of conditional possibility distributions[3] and conditional independence. Their properties are summarized in Table 1 at the end of this section.

[2] max operator must be replaced by sup operator if $\mathbf{X}$ is not supposed to be finite.

[3] This is not the complete list of these definitions; in [5] many other conditioning rules can be found, but those rules have not been studied from the viewpoint of conditional independence.

3.1 Zadeh's Conditioning Rule and Noninteractivity Notion

Zadeh's conditioning rule [14] is very simple and consists in setting conditional possibility distributions equal to the joint ones:

$$\pi_Z(x|y) = \pi(x, y) \tag{4}$$

for all $(x, y) \in \mathbf{X} \times \mathbf{Y}$. As mentioned in [5], this conditioning rule has one great disadvantage — it produces conditional possibility distributions which are not normal, whenever the marginal $\pi(y) < 1$.

First attempt to incorporate independence notion into possibility theory was also done by Zadeh. He called two variables[4] X and Y *noninteractive* iff for any $(x, y) \in \mathbf{X} \times \mathbf{Y}$

$$\pi(x, y) = \min(\pi(x), \pi(y)).$$

This notion can be generalized as follows (see e.g. [8]): the variables X and Y are *conditionally noninteractive given* Z ($I_N(X, Y|Z)$) iff for any $(x, y, z) \in \mathbf{X} \times \mathbf{Y} \times \mathbf{Z}$

$$\pi(x, y|z) = \min(\pi(x|z), \pi(y|z)). \tag{5}$$

If we take into account Zadeh's definition of conditional possibility distributions (4), we can rewrite this equation into

$$\pi(x, y, z) = \min(\pi(x, z), \pi(y, z)).$$

3.2 Hisdal's Conditioning Rule and Related Independence Notions

Hisdal [9] proposed to define conditional possibility distributions as a solution of the equation

$$\pi(x, y) = \min(\pi(y), \pi(x|y)), x \in \mathbf{X}, y \in \mathbf{Y}. \tag{6}$$

This equation, unfortunately, does not have unique solution (the only exception is if $\pi(y) = 1$ for all $y \in \mathbf{Y}$). All its solutions are given by

$$\pi_H(x|y) \in \begin{cases} \{\pi(x, y)\} & \text{if } \pi(x, y) < \pi(y), \\ [\pi(x, y), 1] & \text{if } \pi(x, y) = \pi(y). \end{cases}$$

An arbitrary solution need not be normal (e.g. Zadeh's conditioning rule is a solution of (6)) and therefore some additional condition must be required in order to the normality is obtained. One of the possibilities, proposed by Dubois and Prade [7], is to take the *greatest* solution of the equation[5] (or the least specific), i.e.

$$\pi_{DP}(x|y) = \begin{cases} \pi(x, y) & \text{if } \pi(x, y) < \pi(y), \\ 1 & \text{if } \pi(x, y) = \pi(y). \end{cases}$$

In the same paper Hisdal defined the possibilistic independence in the following way: variable X is *possibilistically independent* of the variable Y iff for any $(x, y) \in \mathbf{X} \times \mathbf{Y}$

$$\pi_H(x|y) = \pi(x). \tag{7}$$

Then she moreover showed that independence implies noninteractivity but the reverse is not valid.

However, Hisdal's approach exhibited one substantial difficulty already noticed in [3]. Since the conditional possibility distribution is not defined uniquely, her definition of independence is of little sense. The modification with the greatest solution, on the other hand, is unique, and therefore it is sensible to take a generalization of (7) for the definition of conditional independence $I(X, Y|Z)$, i.e.

$$\pi_{DP}(x|y, z) = \pi_{DP}(x|z). \tag{8}$$

Let us note that this definition is possibilistic counterpart of (2).

The asymmetry is often considered to be an unpleasant property and therefore Fonck [8] defined conditional independence in the following way:

$$I_S(X, Y|Z) \Leftrightarrow I(X, Y|Z) \text{ and } I(Y, X|Z). \tag{9}$$

The notion of conditional independence defined by I_S is, however, rather restrictive (as mentioned in [1]), since $I_S(X, Y|Z)$ implies either $\pi(x) = 1$ for all $x \in \mathbf{X}$ or $\pi(y) = 1$ for all $y \in \mathbf{Y}$.

It is possible to argue, as suggested in [1], that even the definition (8) is too restrictive as it requires equality between two conditional distributions. The alternative idea is the following: the supplementary knowledge of the value y cannot improve our knowledge of x given z, but can *deteriorate* it, i.e. some information can be lost. Therefore De Campos et al. [1] defined

$$I_I(X, Y|Z) \Leftrightarrow \pi_{DP}(x|y, z) \geq \pi_{DP}(x|z) \tag{10}$$

for all x, y, z.

In [1] still another possibility, how to weaken the definition of conditional possibility is suggested. We can

[4] For the purpose of this section we may adopt informal definition of variable (used e.g. in [14, 9]) as an abstract object that can assume values in certain universe.

[5] Let us note that Zadeh's conditioning rule is the *smallest* solution of the equation.

replace the equality in (8) by equivalence relation $\approx$ compatible with using the minimum (or Gödel's) t-norm (see Subsection 4.1) as the combination operator of possibility distribution (more precisely

$$\begin{aligned}&\pi_{DP}(x|y,z) \approx \pi_{DP}(x|z) \Leftrightarrow \\ &\quad \Leftrightarrow \min(\pi_{DP}(x|y,z), \pi(y,z)) = \\ &\qquad = \min(\pi_{DP}(x|z), \pi(y,z))\end{aligned}$$

for all x, y, z) and define

$$I_M(X,Y|Z) \Leftrightarrow \pi_{DP}(x|y,z) \approx \pi_{DP}(x|z) \qquad (11)$$

for all x, y, z.

De Campos et al. [1] proposed also the following modification π_{DP_C} of the conditioning rule π_{DP} suggested by Dubois and Prade

$$\pi_{DP_C}(x|y) = \begin{cases} \pi(x) & \text{if } \pi_{DP}(x|y) \geq \pi(x) \\ & \text{holds for all } x, \\ \pi_{DP}(x|y) & \text{if there exists } x' \\ & \text{such that} \\ & \pi_{DP}(x'|y) < \pi(x'). \end{cases}$$

The interpretation of this definition is the following: if the conditional possibility distribution is "worse" (i.e. less specific) than the unconditional one, it is better to use the latter one in order not to loose information. Evidently, this rule produces normal conditional distributions.

The new definition of conditional independence is then (analogously to (8))

$$I_C(X,Y|Z) \Leftrightarrow \pi_{DP_C}(x|y,z) = \pi_{DP_C}(x|z) \qquad (12)$$

for all x, y, z. Despite the fact that π_{DP_C} is more restrictive than π_{DP}, both the conditional independence relations $I(X,Y|Z)$ and $I_C(X,Y|Z)$ meet the same system of properties (A1) – (A4) (cf. Table 1).

3.3 Dempster's Conditioning Rule and Related Independence Notions

In his seminal paper [6] on upper and lower probabilities induced by multivalued mappings Dempster suggested a rule for conditioning upper probabilities. Since possibility measures are upper probabilities induced by multivalued mappings of a special type, Dempster's rule can also be used for defining conditional possibility distributions. Applying Dempster's rule, we obtain:

$$\pi_D(x|y) = \begin{cases} \frac{\pi(x,y)}{\pi(y)} & \text{if } \pi(y) > 0, \\ 1 & \text{if } \pi(y) = 0, \end{cases}$$

i.e. if $\pi(y) = 0$ Dempster's rule chooses the least specific value $\pi_D(x|y) = 1$ and therefore conditional possibility distributions are evidently normal.

In this case, conditional independence $I_D(X,Y|Z)$ is usually defined analogously to (8) if

$$\pi_D(x|y,z) = \pi_D(x|z) \qquad (13)$$

for all values x, y, z such that $\pi(y,z) > 0$.

Using the same arguments as in the case of π_{DP} (see [1]), we can weaken this definition by setting

$$I_{DI}(X,Y|Z) \Leftrightarrow \pi_D(x|y,z) \geq \pi_D(x|z) \qquad (14)$$

for all x, y, z such that $\pi(y,z) > 0$. The properties of this conditioning rule are, nevertheless, somewhat surprising (cf. Table 1).

This fact led De Campos et al. to a modification π_{D_C} of π_D proposed in [1] in the following way:

$$\pi_{D_C}(x|y) = \begin{cases} \pi(x) & \text{if } \pi(x,y) \geq \pi(x)\cdot\pi(y) \\ & \text{holds for all } x, \\ \pi_D(x|y) & \text{if there exists } x' \\ & \text{such that} \\ & \pi(x',y) < \pi(x')\cdot\pi(y). \end{cases}$$

Again, this conditioning rule is more restrictive than Dempster's one (cf. π_{DP_C} and π_{DP} in Subsection 3.2), i.e. the resulting conditional possibility distribution is more specific, nevertheless normal. However, in this case the modification influences the properties possessed by the conditional independence relation $I(X,Y|Z)$ defined by

$$I_{DC}(X,Y|Z) \Leftrightarrow \pi_{D_C}(x|y,z) = \pi_{D_C}(x|z) \qquad (15)$$

for all x, y, z.

3.4 "Łukasziewicz' " Conditioning Rule and Independence Notion

Fonck [8] studied also properties of conditional independence based on the Lukasziewicz' t-norm (see Subsection 4.1). The conditioning rule is then

$$\pi_L(x|y) = \pi(x,y) \Leftrightarrow \pi(y) + 1$$

for all $x \in \mathbf{X}$ and $y \in \mathbf{Y}$, and it is again normal. She defined conditional independence analogously to (8) by

$$I_L(X,Y|Z) \Leftrightarrow \pi_L(x|y,z) = \pi_L(x|z) \qquad (16)$$

for all $x \in \mathbf{X}, y \in \mathbf{Y}$ and $z \in \mathbf{Z}$.

This conditioning rule, however, seems to be somewhat strange. If we consider (x,y) such that $\pi(x,y) = 0$, i.e. impossible combination of events, and $\pi(y) = 0.5$, then $\pi_L(x|y) = 0.5$, which is at least contraintuitive.

3.5 Properties of Conditional Independence

We have presented a great number of conditioning rules and even more conditional independence relations. The following table summarizes their properties (CR is conditioning rule, CI is conditional independence and $*$ means that the property is satisfied).

CR	CI	A1	A2	A3	A4	A5
probability theory		$*$	$*$	$*$	$*$	
π_Z	I_N	$*$	$*$	$*$	$*$	
π_{DP}	I		$*$	$*$	$*$	$*$
	I_S	$*$	$*$	$*$	$*$	$*$
	I_I	$*$	$*$	$*$	$*$	
	I_M	$*$	$*$	$*$	$*$	
π_{DP_C}	I_C		$*$	$*$	$*$	$*$
π_D	I_D	$*$	$*$	$*$	$*$	
	I_{DI}	$*$	$*$		$*$	
π_{D_C}	I_{DC}		$*$	$*$	$*$	$*$
π_L	I_L	$*$	$*$	$*$	$*$	$*$

Table 1: Properties of conditional independence.

Which conditioning rule and independence relation should be chosen? Since properties (A2) – (A4) can be viewed as pure formal properties of the notion of irrelevance (as argued in [10]), we can exclude De Campos relation (14). Nevertheless, the question remains: what is more suitable to require — symmetry or intersection (or both)?

4 Measure-theoretical Approach

These problems may be avoided, at least partially, if we adopt de Cooman's measure-theoretical approach [2, 3, 4]. In order to be able to do it, we need a few definitions concerning the t-norms. In next three subsections we will follow above mentioned de Cooman's papers, but for the purpose of this paper we decided to give up generality in favour of simplicity.

4.1 Triangular Norms

A *triangular norm* (or a *t-norm*) T is a binary operator on $[0,1]$ (i.e. $T:[0,1]^2 \to [0,1]$) satisfying the following three conditions:

(i) *boundary conditions:* for any $a \in [0,1]$

$$T(1,a) = a, \qquad T(0,a) = 0;$$

(ii) *isotonicity:* for any $a_1, a_2, b_1, b_2 \in [0,1]$ such that $a_1 \leq a_2, b_1 \leq b_2$

$$T(a_1,b_1) \leq T(a_2,b_2);$$

(iii) *associativity and commutativity:* for any $a,b,c \in [0,1]$

$$\begin{aligned} T(T(a,b),c) &= T(a,T(b,c)), \\ T(a,b) &= T(b,a). \end{aligned}$$

A t-norm T is called *continuous*, if T is a continuous function.

There exist three distinguished continuous t-norms (all of them already mentioned in this paper):

(i) *Gödel's t-norm:* $T(a,b) = \min(a,b)$;

(ii) *product t-norm:* $T(a,b) = a \cdot b$;

(iii) *Lukasziewicz' t-norm:* $T(a,b) = \max(0, a+b \Leftrightarrow 1)$.

4.2 Conditional Possibility Distributions

A mapping $h: \mathbf{X} \to [0,1]$ is called *fuzzy variable*. The set of fuzzy variables on $\mathbf{X}$ will be denoted by $\mathcal{G}(\mathbf{X})$.

Let T be a t-norm on $[0,1]$. For any possibilistic measure Π on $\mathbf{X}$ with distribution π we define the following binary relation on $\mathcal{G}(\mathbf{X})$. For h_1 and h_2 in $\mathcal{G}(\mathbf{X})$ we say that h_1 and h_2 are (Π,T)-*equal almost everywhere* (and write $h_1 \overset{(\Pi,T)}{=} h_2$) iff for any $x \in X$

$$T(h_1(x),\pi(x)) = T(h_2(x),\pi(x)).$$

This notion is very important for the the definition of *conditional possibility distribution* which is defined as *any* solution of the equation

$$\pi_{XY}(x,y) = T(\pi_Y(y), \pi_{X|Y}(x|y)), \qquad (17)$$

for any $(x,y) \in \mathbf{X} \times \mathbf{Y}$. The solution of this equation is not unique, but the ambiguity vanishes when almost everywhere equality is considered. We are able to obtain a representant of these conditional possibility distributions (if T is a continuous t-norm) by taking the residual

$$\pi_{X|Y}(x|\cdot) \overset{(\Pi_Y,T)}{=} \pi_{XY}(x,\cdot)\Delta_T \pi_Y(\cdot), \qquad (18)$$

that is defined as the greatest solution of the equation (17).

Let us remark, that if we use product t-norm, we will obtain Dempster's rule of conditioning, Lukasziewicz' t-norm corresponds to "Lukasziewicz' " rule of conditioning, Gödel's t-norm leads to Hisdal's rule of conditioning and the choice of Gödel's t-norm together with (18) gives the modification of Hisdal's rule proposed by Dubois and Prade.

4.3 Independence

Regarding the independence notions presented in the preceding section, the independence in [4] is defined in substantially different way. De Cooman considered two variables X and Y *possibilistically* T*-independent* iff for any $F_X \in X^{-1}(\mathcal{P}(\mathbf{X}))$, $F_Y \in Y^{-1}(\mathcal{P}(\mathbf{Y}))$,

$$\Pi(F_X \cap F_Y) = T(\Pi(F_X), \Pi(F_Y)),$$
$$\Pi(F_X \cap F_Y^C) = T(\Pi(F_X), \Pi(F_Y^C)),$$
$$\Pi(F_X^C \cap F_Y) = T(\Pi(F_X^C), \Pi(F_Y)),$$
$$\Pi(F_X^C \cap F_Y^C) = T(\Pi(F_X^C), \Pi(F_Y^C)).$$

From this definition it immediately follows that the independence notion is parametrized by T. This fact was not mentioned in Zadeh's and Hisdal's works since they used only one t-norm, Gödel's t-norm. The analogy holds also for independence connected with Dempster's and "Łukasziewicz' " conditioning rules. However, in these cases it is not so apparent as in Zadeh's approach (but analogous to Hisdal's one); the definitions of independence are, in fact the same. The t-norms, however, are "hidden" in the conditioning rules.

De Cooman's definition, moreover, reveals the relation between independence of variables and events. This problem (although very interesting) is behind the scope of this paper, nevertheless it is thoroughly studied in [4].

What is more important, from the viewpoint of this paper, is the following theorem which is an immediate consequence of Proposition 2.6. in [4].

Theorem 1 *Let us assume that t-norm T is continuous. Then the following propositions are equivalent.*

(i) X and Y are T-independent.

(ii) For any $x \in \mathbf{X}$ and $y \in \mathbf{Y}$

$$\pi_{XY}(x, y) = T(\pi_X(x), \pi_Y(y)).$$

(iii) For any $x \in \mathbf{X}$ and $y \in \mathbf{Y}$

$$\begin{aligned} T(\pi_X(x), \pi_Y(y)) &= T(\pi_{X|Y}(x|y), \pi_Y(y)) = \\ &= T(\pi_{Y|X}(y|x), \pi_X(x)). \end{aligned}$$

This theorem shows that the notion of independence defined by de Cooman is equivalent (for $T = \min$) to Zadeh's notion of noninteractivity and, in a sense, also to Hisdal's notion of independence — if the equality sign in (7) is substituted by almost everywhere equality.

4.4 Conditional Independence

In light of these facts we defined in [12] the conditional possibilistic independence in the following way. Variables X and Y are *possibilistically conditionally* T*-independent* given Z ($I_T(X, Y|Z)$ iff for any pair $(x, y) \in \mathbf{X} \times \mathbf{Y}$

$$\pi_{XY|Z}(x, y|\cdot) \overset{(\Pi_Z, T)}{=} T(\pi_{X|Z}(x|\cdot), \pi_{Y|Z}(y|\cdot)). \quad (19)$$

Let us stress again that we do not deal with the pointwise equality, but with the *almost everywhere equality* in contrast to the conditional noninteractivity (5). The following theorem is a "conditional counterpart" of Theorem 1.

Theorem 2 *Let us assume that t-norm T is continuous. Then the following propositions are equivalent.*

(i) X and Y are T-independent given Z.

(ii) For any $x \in \mathbf{X}$, $y \in \mathbf{Y}$ and $z \in \mathbf{Z}$

$$\begin{aligned} &T(\pi_{X|YZ}(x|y, z), \pi_{YZ}(y, z)) = \\ &\quad = T(\pi_{X|Z}(x|z), \pi_{YZ}(y, z)). \end{aligned}$$

Proof. Let (i) be satisfied. Then

$$\begin{aligned} &T(\pi_{X|YZ}(x|y, z), \pi_{YZ}(y, z)) = \\ &\quad = T(\pi_{X|YZ}(x|y, z), T(\pi_{Y|Z}(y|z), \pi_Z(z))) = \\ &\quad = T(\pi_{XY|Z}(x, y|z), \pi_Z(z)) = \\ &\quad = T(T(\pi_{X|Z}(x|z), \pi_{Y|Z}(y|z)), \pi_Z(z)) \\ &\quad = T(\pi_{X|Z}(x|z), \pi_{YZ}(y, z)), \end{aligned}$$

where we used only (17), (19) and associativity of a t-norm.

Let (ii) hold. Then (19) is equivalent to the following equality

$$\begin{aligned} &T(\pi_{XY|Z}(x, y|z), \pi_Z(z)) = \\ &\quad = T(T(\pi_{X|YZ}(x|y, z), \pi_{Y|Z}(y|z)), \pi_Z(z)) = \\ &\quad = T(\pi_{X|YZ}(x|y, z), \pi_{YZ}(y, z)) = \\ &\quad = T(\pi_{X|Z}(x|z), \pi_{YZ}(y, z)) = \\ &\quad = T(\pi_{X|Z}(x|z), T(\pi_{Y|Z}(y|z), \pi_Z(z))) = \\ &\quad = T(T(\pi_{X|Z}(x|z), \pi_{Y|Z}(y|z)), \pi_Z(z)), \end{aligned}$$

where we used only (17), (ii) and associativity of a t-norm. □

Theorem 2 unifies the notions of conditional noninteractivity (5) and Hisdal's definition of conditional independence (8) and also (13) and (16) (if we substitute Gödel's t-norm in (5) by product and Łukasziewicz'

t-norm, respectively) in such a sense, that *pointwise* equalities are substituted by *almost everywhere* ones.

It should also be mentioned that one particular type of the conditional independence $I_T(X,Y|Z)$ has been proposed in [1] for Gödel's t-norm (see also $I_M(X,Y|Z)$ in Subsection 3.2).

Theorem 3 *For any continuous t-norm T relation $I_T(X,Y|Z)$ satisfies (A1) – (A4).*

Proof.

(A1) Symmetry immediately follows from commutativity of a t-norm.

(A2) Let

$$\pi_{XYZ|W}(x,y,z|\cdot) \overset{(\Pi_W,T)}{=} \overset{(\Pi_W,T)}{=} T(\pi_{X|W}(x|\cdot), \pi_{YZ|W}(y,z|\cdot)), \quad (20)$$

then

$$\begin{aligned}
&\pi_{XZ|W}(x,z|\cdot) = \\
&= \max_{y\in\mathbf{Y}} \pi_{XYZ|W}(x,y,z|\cdot) \overset{(\Pi_W,T)}{=} \\
&\overset{(\Pi_W,T)}{=} \max_{y\in\mathbf{Y}} T(\pi_{X|W}(x|\cdot), \pi_{YZ|W}(y,z|\cdot)) = \\
&= T(\pi_{X|W}(x|\cdot), \max_{y\in\mathbf{Y}} \pi_{YZ|W}(y,z|\cdot)) = \\
&= T(\pi_{X|W}(x|\cdot), \pi_{Z|W}(z|\cdot))
\end{aligned}$$

due to isotonicity of a t-norm.

(A3) Let (20) be satisfied, we want to prove that also

$$\pi_{XY|ZW}(x,y|\cdot,\cdot) \overset{(\Pi_{ZW},T)}{=} \overset{(\Pi_{ZW},T)}{=} T(\pi_{X|ZW}(x|\cdot,\cdot), \pi_{Y|ZW}(y|\cdot,\cdot)), (21)$$

which is equivalent to the following equality

$$\begin{aligned}
&T(\pi_{XY|ZW}(x,y|z,w), \pi_{ZW}(z,w)) = \\
&= T(\pi_{XY|ZW}(x,y|z,w), \\
&\qquad T(\pi_{Z|W}(z|w), \pi_W(w))) = \\
&= T(\pi_{XYZ|W}(x,y,z|w), \pi_W(w)) = \\
&= T(T(\pi_{X|W}(x|w), \pi_{YZ|W}(y,z|w)), \\
&\qquad \pi_W(w)) = \\
&= T(\pi_{X|W}(x|w), T(\pi_{YZ|W}(y,z|w), \\
&\qquad \pi_W(w))) = \\
&= T(\pi_{X|W}(x|w), T(\pi_{Y|ZW}(y|z,w), \\
&\qquad \pi_{ZW}(z,w))) = \\
&= T(T(\pi_{X|W}(x|w), \pi_{Y|ZW}(y|z,w)), \\
&\qquad \pi_{ZW}(z,w)) = \\
&= T(\pi_{Y|ZW}(y|z,w), T(\pi_{X|W}(x|w), \\
&\qquad \pi_{ZW}(z,w))) = \\
&= T(\pi_{Y|ZW}(y|z,w), T(\pi_{X|ZW}(x|z,w), \\
&\qquad \pi_{ZW}(z,w))) = \\
&= T(T(\pi_{X|ZW}(x|z,w), \pi_{Y|ZW}(y|z,w)), \\
&\qquad \pi_{ZW}(z,w)),
\end{aligned}$$

where we used (17), (20), associativity and commutativity of a t-norm and Theorem 2.

(A4) Let (20) and

$$\pi_{XZ|W}(x,z|\cdot) \overset{(\Pi_W,T)}{=} T(\pi_{X|W}(x|\cdot), \pi_{Z|W}(z|\cdot))$$

be satisfied, then

$$\begin{aligned}
&T(\pi_{XYZ|W}(x,y,z|w), \pi_W(w)) = \\
&= T(T(\pi_{XY|ZW}(x,y|z,w), \pi_{Z|W}(z|w)), \\
&\qquad \pi_W(w)) = \\
&= T(T(T(\pi_{X|ZW}(x|z,w), \pi_{Y|ZW}(y|z,w)), \\
&\qquad \pi_{Z|W}(z|w)), \pi_W(w)) = \\
&= T(T(\pi_{X|ZW}(x|z,w), \pi_{Y|ZW}(y|z,w)), \\
&\qquad T(\pi_{Z|W}(z|w), \pi_W(w))) = \\
&= T(T(\pi_{Y|ZW}(y|z,w), \pi_{X|ZW}(x|z,w)), \\
&\qquad \pi_{ZW}(z,w)) = \\
&= T(\pi_{Y|ZW}(y|z,w), \\
&\qquad T(\pi_{X|ZW}(x|z,w), \pi_{ZW}(z,w))) = \\
&= T(\pi_{Y|ZW}(y|z,w), \\
&\qquad T(\pi_{X|W}(x|w), \pi_{ZW}(z,w))) = \\
&= T(T(\pi_{X|W}(x|w), \pi_{Y|ZW}(y|z,w)), \\
&\qquad T(\pi_{Z|W}(z|w), \pi_W(w))) = \\
&= T(T(\pi_{X|W}(x|w), \pi_{YZ|W}(y,z|w)), \pi_W(w)),
\end{aligned}$$

where we used only (17), (20), associativity of a t-norm and Theorem 2. □

Property (A5) is not fulfilled, in general, which is obvious from the following example.

Example 1 Let $\mathbf{X} = \mathbf{Y} = \mathbf{Z} = \{0,1\}$ and

$$\pi_{XYZ}(x,y,z) = \begin{cases} 1 & \text{if } x = y = z, \\ 0 & \text{else.} \end{cases}$$

Then

$$\pi_{XY}(x,y) = \begin{cases} 1 & \text{if } x = y, \\ 0 & \text{else,} \end{cases}$$

$$\pi_{XZ}(x,z) = \begin{cases} 1 & \text{if } x = z, \\ 0 & \text{else,} \end{cases}$$

$$\pi_{YZ}(y,z) = \begin{cases} 1 & \text{if } y = z, \\ 0 & \text{else,} \end{cases}$$

and

$$\pi_Y \equiv \pi_Z \equiv 1.$$

Then, for *any* t-norm,[6]

$$\begin{aligned}\pi_{XY|Z}(x,y|z) &= T(\pi_{X|Z}(x|z), \pi_{Y|Z}(y|z)),\\ \pi_{XZ|Y}(x,z|y) &= T(\pi_{X|Y}(x|y), \pi_{Z|Y}(z|y)),\end{aligned}$$

for any $(x,y,z) \in \mathbf{X} \times \mathbf{Y} \times \mathbf{Z}$, but e.g.

$$\pi_{XYZ}(1,0,0) \neq T(\pi_X(1), \pi_{YZ}(0,0)),$$

i.e. $I(X,Y|Z)$ and $I(X,Z|Y)$ hold, but $I(X,YZ|\emptyset)$ does not. $\diamond$

Therefore we can conclude: *There exists no t-norm T such that $I_T(X,Y|Z)$ satisfies (A1)–(A5) for arbitrary possibility distribution.*

This fact perfectly corresponds to the properties of probabilistic conditional independence. In probability theory (A5) need not be satisfied if the probability distribution is not strictly positive. In this case the conditional probability distributions need not be defined uniquely. In possibility theory this nonuniqueness is caused by the use of t-norms. If we adopt the axiomatic approach presented in this section, fullfilness of (A5) depends on the choice of a t-norm and on properties of possibility distribution in question. For example, if we choose product t-norm, (A5) is always satisfied by strictly positive possibility distributions as expressed by Theorem 4.

Lemma 1 *Let $\pi(x,y,z)$ be strictly positive. Then the following statements are equivalent.*

(i) Variables X and Y are conditionally product-independent given Z.

(ii) Joint distribution of X, Y and Z has a form

$$\pi(x,y,z) = \rho_1(x,z) \cdot \rho_2(y,z).$$

Proof. Let (i) be satisfied. Then

$$\pi(x,y,z) = \pi(x|z) \cdot \pi(y|z) \cdot \pi(z)$$

and (ii) is obviously fullfilled (e.g. $\rho_1(x,z) = \pi(x,z)$ and $\rho_2(y,z) = \frac{\pi(y,z)}{\pi(z)}$).

Let (ii) be satisfied. Then

$$\begin{aligned}&\pi_{XY|Z}(x,y,z) =\\ &\quad= \frac{\rho_1(x,z)\cdot\rho_2(y,z)}{\rho_1(z)\cdot\rho_2(z)} =\\ &\quad= \frac{\rho_1(x,z)\cdot\rho_2(z)}{\rho_1(z)\cdot\rho_2(z)} \cdot \frac{\rho_1(z)\cdot\rho_2(y,z)}{\rho_1(z)\cdot\rho_2(z)} =\\ &\quad= \pi_{X|Z}(x|z)\cdot\pi_{Y|Z}(y|z),\end{aligned}$$

i.e. (i) is satisfied. $\square$

[6]Let us note, that the following equalities are pointwise, since $\pi_Y \equiv \pi_Z \equiv 1$.

Theorem 4 *Let T be product t-norm and π be strictly positive possibility distribution. Then also (A5) is satisfied.*

Proof. Let $I(X,Y|ZW)$ and $I(X,Z|YW)$ be satisfied. It means that (due to Lemma 1) π has a form

$$\begin{aligned}&\pi_{XYZW}(x,y,z,w) =\\ &\quad= \rho_1(x,z,w)\cdot\rho_2(y,z,w) =\\ &\quad= \sigma_1(x,y,w)\cdot\sigma_2(y,z,w).\end{aligned}$$

Thus, we have for all z

$$\sigma_1(x,y,w) = \frac{\rho_1(x,z,w)\cdot\rho_2(y,z,w)}{\sigma_2(y,z,w)}.$$

Choosing a fixed $z = z_0$ we have

$$\sigma_1(x,y,w) = f(x,w)\cdot g(y,w)$$

where

$$f(x,w) = \rho_1(x,z_0,w)$$

and

$$g(y,w) = \frac{\rho_2(y,z_0,w)}{\sigma_2(y,z_0,w)}.$$

Therefore

$$\pi_{XYZW}(x,y,z,w) = f(x,w)\cdot g(y,w)\cdot\sigma_2(y,z,w)$$

and hence $I(X,YZ|W)$ (again due to Lemma 1) as desired. $\square$

Although we do not have analogical results for other t-norms, we conjecture that for any of them (at least for any "reasonable" one) there exists a class of distributions satisfying (A5).

5 Conclusions

We have overviewed a great deal of conditioning rules and conditional independence relations that have been introduced in possibility theory. Special attention was paid to the measure-theoretical approach to conditioning and independence presented in Section 4. As we have already mentioned, solutions of the equation (17) produce almost all of previously introduced conditioning rules, perhaps with the exception of the modifications of Hisdal's and Dempster's rules proposed by De Campos et al. Similarly, adopting conditional independence notion (19), we will obtain most of presented conditional independence relations. Properties of this measure-theoretical approach to conditioning and conditional independence correspond to those possessed by stochastic conditional independence. On the other hand, the choice of a t-norm should not be arbitrary, as mentioned in the case of "Lukasziewicz' " conditioning rule.

Still, there exist some conditioning rules suggested by various authors we have not mentioned, but most of them look rather "artificial". The only exception, as far as we know, are conditioning rules based on behavioural interpretation of possibility theory [5]. Unfortunately, the detailed study of this approach is beyond the scope of this paper as the author is not so much familiar with this approach.

The aim of this paper was rather to unify different approaches to conditioning and conditional independence then to discriminate between them. From the practical point of view, the latter task seems to be more important. Since conditioning and independence are very closely connected, any of them will determine the other. One possibility is to justify conditioning rule as suggested in [5], another way is to find satisfactory justification for one independence notion and to derive conditioning rule according to this notion.

References

[1] L. M. De Campos, J. F. Huete and S. Moral. Possibilistic independence, in: *Proceedings of EUFIT 95* (Third European Congress on Intelligent Techniques and Soft Computing), 1:69–73, Verlag Mainz and Wisseschaftverlag, Aachen, 1995.

[2] G. de Cooman. Possibility theory I: The measure- and integral-theoretic groundwork, *Int. J. General Systems,* 25:291–323, 1997.

[3] G. de Cooman. Possibility theory II: Conditional possibility, *Int. J. General Systems,* 25:325–351, 1997.

[4] G. de Cooman. Possibility theory III: Possibilistic independence, *Int. J. General Systems,* 25:353–371, 1997.

[5] G. de Cooman and P. Walley. Coherence rules for defining conditional possibility, to appear in *Internat. J. of Approximate Reasoning.*

[6] A. Dempster. Upper and lower probabilities induced by multivalued mappings, *Ann. Math. Statist.,* 38:325–339, 1967.

[7] D. Dubois and H. Prade. *Possibility theory.* Plenum Press, New York, 1998.

[8] P. Fonck. Conditional independence in possibility theory, in: R. L. de Mantaras, P.Poole (eds.) *Proceedings of 10th conference UAI,* 221–226, Morgan Kaufman, San Francisco, 1994.

[9] E. Hisdal. Conditional possibilities independence and noninteraction, *Fuzzy Sets and Systems,* 1:299-309, 1978.

[10] S. L. Lauritzen. *Graphical models*, Oxford University Press, 1996.

[11] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, San Mateo, 1988.

[12] J. Vejnarová. Possibilistic independence and operators of composition of possibility measures, in: M. Hušková, J. Á. Víšek, P. Lachout (eds.) *Prague Stochastics'98,* 575–580, JČMF, Prague, 1998.

[13] P. Walley. Measures of uncertainty in expert systems, *Artificial Intelligence,* 83:1–58, 1996.

[14] L. A. Zadeh. Fuzzy sets as a basis for theory of possibility, *Fuzzy Sets and Systems,* 1:3–28, 1978.