Conditional Independence and Markov Properties in Possibility Theory

Jitina Vejnarova
Laboratory for Intelligent Systems
University of Economics, Prague
and
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic

Abstract

Conditional independence and Markov properties are powerful tools allowing expression of multidimensional probability distributions by means of low-dimensional ones. As multidimensional possibilistic models have been studied for several years, the demand for analogous tools in possibility theory seems to be quite natural. This paper is intended to be a promotion of de Cooman’s measure-theoretic approach to possibility theory, as this approach allows us to find analogies to many important results obtained in probabilistic framework. First we recall semigraphoid properties of conditional possibilistic independence, parameterized by a continuous t-norm, and find sufficient conditions for a class of Archimedean t-norms to have the graphoid property. Then we introduce Markov properties and factorization of possibility distributions (again parameterized by a continuous t-norm) and find the relationships between them. These results are accompanied by a number of counterexamples, which show that the assumptions of specific theorems are substantial.

1 INTRODUCTION

Conditional independence and Markov properties are fundamental notions of graphical modeling; therefore they are strongly connected with the application of probability theory to artificial intelligence. Complexity of practical problems that are of primary interest in the field of artificial intelligence usually results in the necessity to construct models with the aid 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 us to reduce demands on computer memory. Conditional independence and Markov properties, which are among such tools, allow expression of these multidimensional distributions by means of low-dimensional ones, and therefore to substantially decrease demands on computer memory.

For three centuries, probability theory was the only mathematical tool at our disposal for uncertainty quantification and processing. As a result, many important theoretical and practical advances have been achieved in this field. However, during the last thirty years some new mathematical tools have emerged as alternatives to probability theory. They are used in situations whose nature of uncertainty does not meet the requirements of probability theory, or those in which probabilistic approaches employ criteria that are too strict. Nevertheless, probability theory has always served as a source of inspiration for the development of these nonprobabilistic calculi and these calculi have been continually confronted with probability theory and mathematical statistics from various points of view.

Good examples of this fact include the numerous papers studying conditional independence in various calculi (de Campos and Huete 1999) (Fonck 1994) (Shenoy 1994) (Spohn 1980). With this paper, we will continue our efforts in the area of possibility theory (Vejnarova 1999), attempting to unify the conditional independence notion. We will follow de Cooman’s measure-theoretic approach (de Cooman 1997) to possibility theory, but for purposes of this paper we have chosen to forego generality in favour of simplicity.

2 BASIC TERMINOLOGY

In this section we will provide a brief overview of basic notions and results from (de Cooman 1997) and in Section 2.1 also from (de Baets et al. 1999) that are necessary for understanding the sections that follow.

2.1 TRIANGULAR NORMS

A triangular norm (or a t-norm) \( T \) is a binary operator on \([0,1]\) (i.e. \( T : [0,1]^2 \rightarrow [0,1] \)) satisfying the following three conditions:
(i) boundary conditions: for any \( a \in [0, 1] \)
\[
T(1, a) = a, \quad T(0, a) = 0;
\]
(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 \( a, b, c \in [0, 1] \)
\[
T(T(a, b), c) = T(a, T(b, c)), \quad T(a, b) = T(b, a).
\]

A t-norm \( T \) is called continuous, if \( T \) is a continuous function. In this case commutativity is just a consequence of the remaining conditions.

There exist three distinct continuous t-norms, which will be studied in this paper:

(i) G"odel's t-norm: \( T_G(a, b) = \min(a, b) \);
(ii) product t-norm: \( T_p(a, b) = a \cdot b \);
(iii) Lukasiewicz' t-norm: \( T_L(a, b) = \max(0, a + b - 1) \).

Because of its associativity, any t-norm \( T \) can be extended to an \( n \)-ary operator \( T^n : [0, 1]^n \to [0, 1] \), namely, in the following way
\[
T(T^{n-1}(x_1, \ldots, x_{n-1}), x_n) = T^n(x_1, \ldots, x_n).
\]

A t-norm \( T \) is called Archimedean iff, for any \( (x, y) \in (0, 1)^2 \), there exists \( n \in \mathbb{N} \) such that
\[
T^n(x, \ldots, x) < y.
\]

Let us note that both product and Lukasiewicz' t-norms are Archimedean, but G"odel's is not.

A t-norm \( T \) is called strict iff it is continuous and, for any \( x \in (0, 1), T(x, \cdot) \) is strictly increasing. A t-norm \( T \) is called nilpotent iff it is continuous, Archimedean and not strict, i.e., any continuous Archimedean t-norm is either strict or nilpotent. \( \varphi \)-transform of \( T \) is a t-norm \( T_\varphi \) defined by
\[
T_\varphi(x, y) = \varphi^{-1}(T(\varphi(x), \varphi(y))).
\]

It can easily be seen that product t-norm is strict and Lukasiewicz' t-norm is nilpotent. Moreover, the following proposition (see e.g. (de Baets et al. 1999)) holds true:

**Proposition 1** (i) A t-norm \( T \) is strict if and only if there exists a \([0, 1]\)-automorphism \( \varphi \) such that \( T \) is the \( \varphi \)-transform of the product t-norm.

(ii) A t-norm \( T \) is nilpotent if and only if there exists a \([0, 1]\)-automorphism \( \psi \) such that \( T \) is the \( \psi \)-transform of Lukasiewicz' t-norm.

---

\( \)\(^1\)G"odel's t-norm is sometimes also called minimum t-norm. But we prefer the first name, since its residual (see the last paragraph of this subsection) coincides with G"odel's implication well-known from fuzzy logic.
2.2.1 Almost everywhere equality

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

Let \( T \) be a \( t \)-norm on \([0,1]\). For any possibility measure \( \Pi \) on \( X \) with distribution \( \pi \), we define the following binary relation on \( G(X) \). For \( h_1 \) and \( h_2 \) in \( G(X) \) we say that \( h_1 \) and \( h_2 \) are \((\Pi, T)\)-equal almost everywhere (and write \( h_1 \sim_{(\Pi,T)} h_2 \)) iff for any \( x \in X \)

\[
T(h_1(x), \pi(x)) = T(h_2(x), \pi(x)).
\]

The strength of this equality is dependent not only on the possibility measure (as in probabilistic framework) but also on the choice of \( t \)-norm. For example, for product \( t \)-norm it means that the equivalent functions may differ only on a set \( E \) such that \( \Pi(E) = 0 \), while for Gödel’s one the classes of equivalence are much wider. For more details see (de Cooman 1997).

This notion is very important for the definition of conditional possibility distribution as well as for that of (conditional) independence.

2.2.2 Conditioning

Conditional possibility distribution is defined as any solution of the equation

\[
\pi_{XY}(x, y) = T(\pi_Y(y), \pi_{X\mid Y}(x|y)),
\]

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

\[
\pi_{X\mid Y}(x|\cdot) \overset{(\Pi_Y,T)}{=} \pi_{XY}(x, \cdot) \Delta_T \pi_Y(\cdot),
\]

which is defined as the greatest solution of the equation (4) (cf. also Section 2.1).

We note that if we use product \( t \)-norm, we will obtain Dempster’s rule of conditioning (Dempster 1967). Lukasiewicz’ \( t \)-norm corresponds to “Lukasiewicz’” rule of conditioning\(^2\) (Fonck 1994), Gödel’s \( t \)-norm leads to Hisdal’s rule of conditioning (Hisdal 1978), and the choice of Gödel’s \( t \)-norm together with (5) gives the modification of Hisdal’s rule proposed by (Dubois and Prade 1988).

2.2.3 Independence

De Cooman considered two variables \( X \) and \( Y \) possibilistically \( T \)-independent iff for any \( F_X \in X^{-1}(\Pi(X)) \), \( F_Y \in Y^{-1}(\Pi(Y)) \) and any \( G_X \in \{ F_X, F_X^c \} \), \( G_Y \in \{ F_Y, F_Y^c \} \)

\[
\Pi(G_X \cap G_Y) = T(\Pi(G_X), \Pi(G_Y)).
\]

\(^2\)This conditioning rule is somewhat questionable: from impossible combination of events we can obtain that one of them conditioned by the other is “somewhat possible”; for more discussion see (Vejnarová 1999).

From this definition it immediately follows that the independence notion is parameterized 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. Furthermore, de Cooman’s definition reveals the relationship between independence of variables and events; for more details see (de Cooman 1997).

What is more important, from the viewpoint taken in this paper, is the following theorem which is an immediate consequence of Proposition 2.6. in (de Cooman 1997).

**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 \( x \in X \) and \( y \in Y \)

\[
\pi_{XY}(x, y) = T(\pi_X(x), \pi_Y(y))
\]

(iii) For any \( x \in X \) and \( y \in Y \)

\[
T(\pi_X(x), \pi_Y(y)) = T(\pi_{X\mid Y}(x|y), \pi_Y(y)) = T(\pi_{Y\mid X}(y|x), \pi_X(x)).
\]

This theorem shows that the notion of independence defined by de Cooman is equivalent (for \( T = \text{min} \)) to Zadeh’s notion of noninteractivity (Zadeh 1978) and, in a sense, also to Hisdal’s notion of independence (Hisdal 1978) — if the equality sign in her definition is substituted by almost everywhere equality.

3 CONDITIONAL INDEPENDENCE

Among the properties satisfied by the ternary relation \( I(X, Y \mid Z) \) of independence, regardless the framework in question, the following are of principal importance:

(A1) \( I(X, Y \mid Z) \to I(Y, X \mid Z) \) symmetry,

(A2) \( I(X, Y \mid W) \to I(X, Z \mid W) \) decomposition,

(A3) \( I(X, Y \mid W) \to I(X, Y \mid ZW) \) weak union,

(A4) \([I(X, Y \mid ZW) \land I(X, Z \mid W)] \to I(X, Y \mid ZW) \) contraction,

(A5) \([I(X, Y \mid ZW) \land I(X, Z \mid YW)] \to I(X, Y \mid ZW) \) intersection.

3.1 CONDITIONAL \( T \)-INDEPENDENCE

In light of the above-mentioned facts (cf. Theorem 1), in (Vejnarová 1998) we defined the conditional possibilistic independence in the following way. Variables \( X \) and \( Y \) are possibilistically conditionally \( T \)-independent given \( Z \) \( (I_{P}(X, Y \mid Z)) \) iff for any pair \((x, y) \in X \times Y\)

\[
\pi_{XY | Z}(x, y|\cdot) \overset{(\Pi_Z,T)}{=} T(\pi_X(x|\cdot), \pi_Y(y|\cdot)).
\]
Let us stress again that we do not deal with pointwise equality, but with almost everywhere equality in contrast to the conditional noninteractivity (Fonck, 1994). The following theorem proven in (Vejnarová, 1999) 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 \( x \in X \), \( y \in Y \) and \( z \in Z \)

\[
T(\pi_{X\mid Y\mid Z}(x, y, z), \pi_{Y\mid Z}(y, z)) = T(\pi_{X\mid Z}(x, z), \pi_{Y\mid Z}(y, z)).
\]

Theorem 2 unifies the notions of conditional noninteractivity (Fonck 1994) and various notions of conditional independence, which can be found in (de Campos and Huete 1995), (Fonck 1994) (if we use specific t-norms). It should also be mentioned that one particular type of conditional independence \( I_T(X, Y \mid Z) \) has been proposed in (de Campos and Huete 1995) for Gödel's t-norm.

### 3.2 SEMI-GRAPHOID PROPERTIES OF CONDITIONAL \( T \)-INDEPENDENCE

In (Vejnarová 1999) we studied semi-graphoid properties of the relation \( I_T(X, Y \mid Z) \) and proved

**Theorem 3** For any continuous t-norm \( T \), the relation \( I_T(X, Y \mid Z) \) satisfies (A1)-(A4).

Property (A5) is not fulfilled in general, which is obvious from Example 1. This example was previously published in (Vejnarová 1999), but it seems useful to have it reiterated here, as we will see later.

**Example 1** Let \( X = Y = Z = \{0, 1\} \) and

\[
\pi_{X\mid Y\mid Z}(x, y, z) = \begin{cases} 1 & \text{if } x = y = z, \\ 0 & \text{else}. \end{cases}
\]

Then

\[
\pi_{X\mid Y}(x, y) = \begin{cases} 1 & \text{if } x = y, \\ 0 & \text{else}, \end{cases} \quad \pi_{X\mid Z}(x, z) = \begin{cases} 1 & \text{if } x = z, \\ 0 & \text{else}, \end{cases} \quad \pi_{Y\mid Z}(y, z) = \begin{cases} 1 & \text{if } y = z, \\ 0 & \text{else}, \end{cases}
\]

and

\[
\pi_Y \equiv \pi_Z \equiv 1.
\]

Then, for any continuous t-norm,\(^3\)

\[
\pi_{X\mid Y\mid Z}(x, y, z) = T(\pi_{X\mid Z}(x, z), \pi_{Y\mid Z}(y, z)), \quad \pi_{X\mid Y\mid z}(x, z) = T(\pi_{X\mid Y}(x, y), \pi_{Z\mid Y}(z, y)),
\]

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

\[
\pi_{X\mid Y\mid z}(1, 0, 0) \neq T(\pi_X(1), \pi_{Y\mid Z}(0, 0)),
\]

i.e., \( I_T(X, Y \mid Z) \) and \( I_T(X, Z \mid Y) \) hold, but \( I_T(X, Y \mid Z) \) does not.

Therefore we concluded:

**Proposition 2** There exists no t-norm \( T \) such that \( I_T(X, Y \mid Z) \) satisfies (A1)-(A5) for arbitrary possibility distribution.

### 3.3 GRAPHOID PROPERTY FOR ARCHIMEDEAN T-NORMS

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 non-uniqueness is caused by the use of t-norms. If we adopt the axiomatic approach presented in this paper, satisfaction of (A5) depends on the choice of a t-norm and on the properties of the possibility distribution in question. If we choose an Archimedian t-norm, (A5) is always satisfied by strictly positive possibility distributions, as expressed by Theorem 4.

**Lemma 1** Let \( T \) be a strict t-norm and \( \pi(x, y, z) \) be strictly positive. Then the following statements are equivalent.

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

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

\[
\pi(x, y, z) = \varphi^{-1}(\rho_1(x, z) \cdot \rho_2(y, z))
\]

for some \( \rho_1 \) and \( \rho_2 \), such that \( \rho_1(x, z) \cdot \rho_2(y, z) \in [0, 1] \) for all \((x, y, z) \in X \times Y \times Z\), and \([0, 1]\)-automorphism \( \varphi \), such that \( T \) is \( \varphi \)-transform of product t-norm.

**Proof.** Let (i) be satisfied. Then

\[
\pi(x, y, z) = T(T(\pi(x, z), \pi(y, z)), \pi(z)) = T(\pi(x, z), \pi(y, z)) = \varphi^{-1}(\varphi(\pi(x, z)) \cdot \varphi(\pi(y, z)))
\]

by (1) and (i) in Proposition 1, and therefore (ii) is obviously fulfilled (e.g. \( \rho_1(x, z) = \varphi(\pi(x, z)) \) and \( \rho_2(y, z) = \varphi(\pi(y, z)) \)).

Let (ii) be satisfied. Then by (2), (1) and (i) in Proposition 1

\[
\pi_{X\mid Y\mid Z}(x, y, z) = \varphi^{-1}(\rho_1(x, z) \cdot \rho_2(y, z)) \Delta_T \varphi^{-1}(\rho_1(z) \cdot \rho_2(z)) = \varphi^{-1}\left(\frac{\rho_1(x, z) \cdot \rho_2(y, z)}{\rho_1(z) \cdot \rho_2(z)}\right)
\]
Let $T$ be an Archimedean t-norm and $\pi(x, y, z)$ be strictly positive possibility distribution. Then (A5) is also satisfied.

Proof. Let $I_{T}(X, Y|ZW)$ and $I_{T}(X, Z|YW)$ be satisfied. Due to the fact that t-norm $T$ is Archimedean, $\pi$ has either the form

$$\pi_{XYZW}(x, y, z, w) = \varphi^{-1}(\rho_1(x, z) \cdot \rho_2(y, z))$$

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

$$g(y, w) = \rho_2(y, z_0, w)$$

Therefore

$$\pi_{XYZW}(x, y, z, w) = \varphi^{-1}(f(x, w) \cdot g(y, w) \cdot \sigma_2(y, z, w))$$

and $f(x, w) \cdot g(y, w) \cdot \sigma_2(y, z, w) \in [0, 1]$ for all $(x, y, z, w) \in X \times Y \times Z \times W$, where both $\rho_1(x, z, w)$ and $\rho_2(y, z, w)$ are conditionally $T$-independent.

The proof in the second case is completely analogous. We have that for all $z$

$$\sigma_1(x, y, w) = \rho_1(x, z, w) \cdot \rho_2(y, z, w) - \sigma_2(y, z, w)$$

Choosing a fixed $z = z_0$ we have

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

where

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

and

$$g(y, w) = \rho_2(y, z_0, w)$$

Therefore

$$\pi_{XYZW}(x, y, z, w) = \varphi^{-1}(f(x, w) + g(y, w) + \sigma_2(y, z, w))$$

and $f(x, w) + g(y, w) + \sigma_2(y, z, w) \in [0, 1]$ for all $(x, y, z, w) \in X \times Y \times Z \times W$, (the arguments are
the same as above), and hence \( I_T(X, Y \mid Z \mid W) \) (again due to Lemma 2) as desired. \( \square \)

Since Gödel's \( t \)-norm is not Archimedean, Theorem 4 does not hold true for it, as can be seen from the following example:

**Example 2** Let \( X = Y = Z = \{0, 1\} \) and

\[
\pi_{X \mid Y \mid Z}(x, y, z) = \begin{cases} 1 & \text{if } x = y = z, \\ \frac{1}{2} & \text{else.} \end{cases}
\]

Then analogously to Example 1

\[
\pi_{X \mid Y \mid Z}(x, y, z) = \min(\pi_{X \mid Z}(x \mid z), \pi_{Y \mid Z}(y \mid z)), \\
\pi_{X \mid Z \mid Y}(x, z \mid y) = \min(\pi_{X \mid Y}(x \mid y), \pi_{Z \mid Y}(z \mid y)),
\]

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

\[
\pi_{X \mid Y \mid Z}(1, 0, 0) \neq \min \pi_X(1), \pi_Y(0), 0),
\]

i.e. \( I_G(X, Y \mid Z) \) and \( I_G(X, Z \mid Y) \) hold, but \( I_G(X, Y \mid Z \mid W) \) does not. \( \square \)

## 4 MARKOV PROPERTIES AND FACTORIZATION

Before introduction of Markov properties and factorization, let us present a few necessary notions from graph theory.

A **graph** is a pair \( G = (V, E) \), where \( V \) is a finite set of vertices and the set of edges \( E \) is a subset of the set \( V \times V \) of (unordered) pairs of distinct vertices. A subset of the vertex set \( A \subseteq V \) induces a subgraph \( G_A = (A, E_A) \), where the edge set \( E_A = E \cap (A \times A) \) is obtained from \( G \) by keeping edges with both end points in \( A \).

A graph is **complete** if all vertices are joined by a line. A subset is complete if it induces a complete subgraph. A maximal (with respect to set inclusion) complete subset is called **clique**.

If there is a line between \( a \in V \) and \( b \in V \), \( a \) and \( b \) are said to be **adjacent**, otherwise they are **non-adjacent**. The **boundary** \( \text{bd}(A) \) of a subset \( A \subseteq V \) of vertices is the set of vertices in \( V \setminus A \) that are adjacent to vertices in \( A \). The **closure** of \( A \) is \( \text{cl}(A) = A \cup \text{bd}(A) \).

If there is a **path** from \( a \) to \( b \) (a sequence \( a = a_0, a_1, \ldots, a_n = b \) of distinct vertices such that \( (a_{i-1}, a_i) \in E \) for all \( i = 1, \ldots, n \)) we say that \( a \) and \( b \) are in the same **connectivity component**. A subset \( S \subseteq V \) is called an \((a, b)\)-**separator** if all paths from \( a \) to \( b \) intersect \( S \). The subset \( S \) **separates** \( A \) from \( B \) if it is an \((a, b)\)-**separator** for every \( a \in A \) and \( b \in B \).

### 4.1 MARKOV PROPERTIES

Now, let us consider the conditional independence in a special situation: we have a graph \( G = (V, E) \) and a finite collection of variables \((X_i)_{i \in V}\) taking their values in \((X_i)_{i \in V}\). For \( A \subseteq V \) we define

\[
X_A = \bigcup_{i \in A} X_i
\]

and \( X = X_V \). Here we will use, for simplicity sake, the symbol \( I_T(A, B \mid C) \) instead of \( I_T(X_A, X_B \mid X_C) \).

With any undirected graph \( G = (V, E) \) and a collection of variables \((X_i)_{i \in V}\) we can (analogous to probability theory) associate three different Markov properties. A possibility measure is said to obey

(P) the **pairwise Markov property**, relative to \( G \) and \( T \) if, for any pair \((i, j)\) of non-adjacent vertices,

\[
I_T(i, j)\{V \setminus \{i, j\}\};
\]

(L) the **local Markov property**, relative to \( G \) and \( T \) if, for any vertex \( i \in V \),

\[
I_T(i, V \setminus \text{cl}(i) \cup \text{bd}(i));
\]

(G) the **global Markov property**, relative to \( G \) and \( T \) if, for any triple \((A, B, S)\) of disjoint subsets of \( V \) such that \( S \) separates \( A \) from \( B \) in \( G \),

\[
I_T(A, B \mid S).
\]

The global Markov property (G) gives the general criterion for deciding whether two groups of variables \( A \) and \( B \) are conditionally independent, given a third group of variables \( S \). It is the strongest property, as can be seen from Theorem 5.

The following two theorems (Theorem 5 and Theorem 6) are presented without proofs. Their proofs can be found in (Lauritzen 1996), since they completely depend on semi-graphoid and graphoid properties, respectively, and not on the distributions in question.

**Theorem 5** For any undirected graph \( G \) and any possibility distribution on \( X \) it holds true that

\[
(G) \quad \implies \quad (L) \quad \implies \quad (P).
\]

These three properties are really different (analogous to probabilistic framework), as can be seen from the following two examples, inspired by Example 3.6 and Example 3.5 from (Lauritzen 1996).

**Example 3** Let \( X = Y = Z = \{0, 1\} \) and the joint distribution of three bivariate variables \( X, Y \) and \( Z \) on the graph

\[
\begin{array}{ccc}
X & \xrightarrow{a} & Y \xrightarrow{b} & Z \\
\end{array}
\]

be defined as in Example 1. From the results obtained in that example, we can see that this possibility distribution satisfies the pairwise Markov property (P), since \( I_T(X, Z \mid Y) \), but does not satisfy the local Markov property (L), since not \( I_T(X, Y \mid Z) \).

\( \square \)
Example 4 Let $U = W = X = Y = Z = \{0, 1\}$ and the joint possibility distribution of five bivariate variables be defined on the graph $G$

\[
\begin{array}{c}
U \quad W \quad X \quad Y \quad Z
\end{array}
\]

as follows

\[
\pi(u, w, x, y, z) = \begin{cases} 
1 & \text{if } u = w = 1 \& x = y = z = 0, \\
0 & \text{or } u = w = x = 0 \& y = z = 1, \\
0 & \text{else.}
\end{cases}
\]

It is easy (but somewhat time-consuming) to verify that this distribution satisfies (L) (for any continuous t-norm $T$) but, since

\[
T(\pi_{UW|X}(0, 0|0), \pi_{YZ|X}(0, 0|0)) = 1 \neq
\end{equation}

\[

\neq 0 = \pi_{UWYZ|X}(0, 0, 0, 0),
\]

it does not satisfy (G) (again for arbitrary continuous t-norm $T$).

Note that it was necessary to use five variables in Example 4, since $(G) \iff (L)$ up to $|V| = 4$. To see this, let us consider three disjoint subsets of $V$: $A, B, S$, such that $S$ separates $A$ from $B$. Let us assume that each of these subsets is nonempty (otherwise the equivalence is trivial). Then at least one of $A$ and $B$ is a singleton $\{i\}$, and therefore $bd\{i\} \subset S$. Then $(G)$ follows from $(L)$ using weak union $(A3)$.

Theorem 6 If a possibility distribution on $X$ is such that $(A5)$ holds true for disjoint subsets $A, B, C, D$ substituted for $X, Y, Z, W$ then

\[
(G) \iff (L) \iff (P).
\]

Due to the results obtained in the previous section (Theorem 4), we immediately get the following

Corollary 1 Let $T$ be an Archimedean t-norm and $\pi$ a strictly positive possibility distribution. Then

\[
(G) \iff (L) \iff (P).
\]

4.2 Factorization

Factorization is a property even more strict than the global Markov property $(G)$. We will say that a possibility distribution $\pi$ factorizes (has a property $(F)$) with respect to $G$ and $T$, if, for all complete subsets $A \subset V$, there exist non-negative functions $\psi_A$ of $x_A$ such that $\pi$ has the form

\[
\pi(x) = T^{\mathcal{A}}(\psi_{A_1}(x_{A_1}), \ldots, \psi_{A_{|\mathcal{A}|}}(x_{A_{|\mathcal{A}|}})),
\]

where $\mathcal{A}$ denotes the set of all complete subsets of $V$. The functions $\psi_A$ are not uniquely determined (in general), since they can be “multiplied” in several ways.

Without loss of generality we can assume (cf. (Lauritzen 1996)) that only cliques (maximum complete subgraphs) appear as the sets, i. e., that

\[
\pi(x) = T^{(C)}(\psi_{C_1}(x_{C_1}), \ldots, \psi_{C_{|C|}}(x_{C_{|C|}})),
\]

where $C$ denotes the set of all cliques of $G$.

Theorem 7 Let $T$ be an Archimedean t-norm. Then, for any undirected graph $G$ and any possibility distribution $\pi$ on $X$,

\[
(F) \iff (G).
\]

Proof. Let $A, B, S$ be three disjoint subsets of $V$ such that $S$ separates $A$ from $B$. Let $A$ denote the connectivity component in subgraph of $G$ induced by $V \setminus S$ containing $A$ and let $B = V \setminus (A \cup S)$. Since $A$ and $B$ are separated by $S$, their elements must be in different connectivity components of the subgraph induced by $V \setminus S$, and any clique of $G$ is a subset of either $A \cup S$ or $B \cup S$. Let us denote by $CA$ the set of cliques contained in $A \cup S$. Using (8), we get

\[
\pi(x) = T^{(C)}(\psi_{C_1}(x_{C_1}), \ldots, \psi_{C_{|C|}}(x_{C_{|C|}})) =
\end{equation}

\[
= T(\psi_{C_1}(x_{C_1}), \ldots, \psi_{C_{|C|}}(x_{C_{|C|}})),
\]

\[
T^{(C)}(\psi_{C_1}(x_{C_1}), \ldots, \psi_{C_{|C|}}(x_{C_{|C|}})) =
\end{equation}

\[
= T(\psi_{C_1}(x_{C_1}), \ldots, \psi_{C_{|C|}}(x_{C_{|C|}})).
\]

Since $T$ is Archimedean, we will obtain, due to Proposition 1 and Lemma 1 or Lemma 2, that $I_T(A, B|S)$. Application of decomposition (A2) gives $I_T(A, B|S)$ and the second one (due to symmetry $(A1)$) $I_T(A, B|S)$ as desired.

The reverse of this implication is not valid, as can be seen from the following example:

Example 5 Let $X, Y, Z$ and $W$ be four bivariate distributions defined on the graph with the joint possibility distribution $\pi$ defined as follows: any of the following combinations has possibility equal to 1, the remaining combinations are equal to 0:

\[
\begin{array}{c}
(0, 0, 0, 0) & (0, 0, 0, 1) & (0, 0, 1, 1) & (0, 1, 1, 1) & (1, 0, 0, 0) & (1, 1, 0, 0) & (1, 1, 1, 1)
\end{array}
\]

From this definition we immediately see that $\pi_{XZ} \equiv \pi_{YZ} \equiv 1$ and therefore the conditional distributions equal unconditional ones (for any continuous t-norm). Therefore we have

\[
\pi_{XZ|YW}(0, 0|0, 0) = 1 = T(\pi_{XZ|YW}(0|0, 0), \pi_{Y|Z|W}(0|0, 0))
\]

\[
\pi_{XZ|YW}(1, 0|0, 0) = 1 = T(\pi_{XZ|YW}(1|0, 0), \pi_{Y|Z|W}(0|0, 0))
\]

\[
\pi_{XZ|YW}(0, 1|0, 0) = 0 = T(\pi_{XZ|YW}(0|0, 0), \pi_{Y|Z|W}(1|0, 0))
\]

\[
\pi_{XZ|YW}(1, 1|0, 0) = 0 = T(\pi_{XZ|YW}(1|0, 0), \pi_{Y|Z|W}(1|0, 0))
\]

and similarly for the other values of $Y$ and $W$. We will also get similar results for $\pi_{Y|W|XZ}$; therefore, we can see that the distribution $\pi$ satisfies the global Markov property $(G)$ with respect any continuous t norm $T$ and the chordless 4-cycle:
But $\pi$ does not factorize with respect to this graph. To see it, let us assume that $\pi$ factorizes. Then

$$1 = \pi(0,0,0,0) = T^4(\phi_{XY}(0,0), \phi_{YZ}(0,0), \phi_{ZW}(0,0), \phi_{WX}(0,0)),$$

but also

$$0 = \pi(0,1,0,0) = T^4(\phi_{XY}(0,1), \phi_{YZ}(1,0), \phi_{ZW}(0,0), \phi_{WX}(0,0)).$$

From this we have

$$T(\phi_{XY}(0,1), \phi_{YZ}(1,0)) = 0.$$

Since

$$1 = \pi(1,1,0,0) = T^4(\phi_{XY}(1,1), \phi_{YZ}(1,0), \phi_{ZW}(0,0), \phi_{WX}(0,1)),$$

it is evident that $\phi_{YZ}(1,0) = 1$, therefore $\phi_{XY}(1,0) = 0$. But then

$$1 = \pi(0,1,1,1) \neq T^4(\phi_{XY}(0,1), \phi_{YZ}(1,1), \phi_{ZW}(1,1), \phi_{WX}(1,0)) = 0;$$

therefore, it is evident that $\pi$ cannot factorize. ◇

5 CONCLUSIONS

The goal of this paper was to promote de Cooman’s measure-theoretic approach to possibility theory by demonstrating the parallels between the probabilistic and possibilistic approaches. We have shown that for the wide class of Archimedean $t$-norms, the conditional $T$-independence has the same properties as the conditional independence in probability theory. We also introduced Markov properties of possibility measures and demonstrated that the relationships between them are completely analogous with those of probability measures. And finally, after introducing the $T$-factorization, we proved that it implies the global Markov property for Archimedean $t$-norms.

There are still some open problems, which should be solved in the near future. The first one, perhaps most important, concerns Gödel’s $t$-norm. Gödel’s $t$-norm is the “classical” one in possibility theory, but many theorems presented in this paper do not hold true for it, since it is not Archimedean. Hence, the task is to find analogous theorems for Gödel’s $t$-norm. The second one is to find an analogy to the Clifford-Hammersley theorem for possibility measures, i.e., to find conditions under which

$$(F) \iff (G) \iff (L) \iff (P).$$

The usefulness of such a theorem is obvious, since it would enable us to determine whether or not a possibility distribution factorizes with respect to a given graph, simply by checking the pairwise Markov property.

Acknowledgement

The research was financially supported by the grants MŠMT no. VS96008, GA ČR no. 201/98/1487 and KONTAKT ME 200/1998.

References

B. de Baets, E. Tsiorkova and R. Mesiar (1999). Conditioning in possibility theory with strict order norms. Fuzzy Sets and Systems 106:221–229.

L. M. de Campos and J. F. Huete (1999). Independence concepts in possibility theory: part 1, part 2. Fuzzy Sets and Systems 103:127–152, 487–505.

G. de Cooman (1997). Possibility theory I – III. Int. J. General Systems, 25:291–371.

A. Dempster (1967). Upper and lower probabilities induced by multivalued mappings. Ann. Math. Statist., 38:325–339.

D. Dubois and H. Prade (1988). Possibility theory. Plenum Press, New York.

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

E. Hisdal (1978). Conditional possibilities independence and noninteraction, Fuzzy Sets and Systems, 1:299–309.

S. L. Lauritzen (1996). Graphical models, Oxford University Press.

P. A. Meyer (1967). Markov Processes, Springer, LNM 26.

P. P. Shenoy (1994). Conditional independence in valuation-based systems, Int. J. Approximate Reasoning, 10:203–234.

W. Spohn (1980). Stochastic independence, causal independence and shieldability. J. Phil. Logic, 9:73–99.

J. Vejnarová (1998). Possibilistic independence and operators of composition of possibility measures. In: M. Hušková, J. A. Víšek, P. Lachout (eds.), Prague Stochastics'98, JČMF, Prague, 575–580.

J. Vejnarová (1999). Conditional independence relations in possibility theory. In: G. de Cooman, F. Cozman, S. Moral, P. Walley (eds.), ISIPTA'99, Universiteit Gent, 343–351.

L. A. Zadeh (1978). Fuzzy sets as a basis for theory of possibility. Fuzzy Sets and Systems, 1:3–28.