ELSEVIER

Fuzzy Sets and Systems 103 (1999) 127–152

FUZZY sets and systems

Independence concepts in possibility theory: Part I[1]

Luis M. de Campos*, Juan F. Huete

Departamento de Ciencias de la Computación e Inteligencia Artificial, Universidad de Granada, E-18071 Granada, Spain

Received May 1996; received in revised form April 1997

Abstract

The notion of independence is of great importance in any formalism for managing uncertainty, for both theoretical and practical reasons. In this paper we study the concept of independence in the framework of possibility theory. Our approach to defining conditional independence relationships is based on comparing conditional possibility measures. Different comparison criteria are presented, based on the ideas of 'not to modify', 'not to gain', and 'to obtain similar' information after conditioning. For each definition of independence considered, an axiomatic study has been carried out. Moreover, there are different operators to define conditional possibility measures, which are related to different views of possibility theory. Particularly, in the first part of the paper, we use Hisdal conditioning (whereas Dempster conditioning will be used in the second part). Finally, we study the marginal problem for possibility measures and, as an application, we show that it is possible to store large n-dimensional possibility distributions efficiently, using independence relationships among variables. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Independence; Possibility theory; Hisdal conditioning; Marginal problem

1. Introduction

Independence relationships, between events or variables, permit us to modularize the available knowledge about a given domain in such a way that, in order to perform inference tasks, we only need to consult relevant or dependent pieces of information. Dependence is a relationship stating a possible change in our current belief due to a specific change in our knowledge. When a variable X is considered independent of another variable Y, given the state of knowledge Z, then our belief about X does not change as a result of knowing additional information for Y. Thus, for example in belief network-based systems [24], the use of independence permits us to obtain efficient storage and updating of the information.

The concept of conditional independence (dependence) has been studied in depth mainly for probability measures [11,23,28]. However, there are some works about the same topic in other frameworks that handle uncertain information [4,5,9,27,30], and works that consider independence relationships in an abstract way [24,25,29]. In this paper, our aim is to study the concept of independence in Possibility theory. Different

* Corresponding author. E-mail: lci@decsai.ugr.es.

[1] This work has been supported by the DGICYT under Project n. PB92-0939.

PII: S0165-0114(97)00160-7

works about this topic have recently appeared: Benferhat et al. [1] approach this problem from a logical point of view; Dubois et al. [14] and Fariñas and Herzig [18] study independence between events and its application to default reasoning; Cooman and Kerre [10] propose different definitions of independence, considering events and variables; Fonck [19] studies conditional independence for variables; Campos, Huete and others have also studied possibilistic conditional independence for variables in previous works [2, 3, 7, 21].

In order to establish an independence relationship between variables, our approach is to compare the previous ('a priori') information with the information that we obtain after a new piece of information has been added ('a posteriori'). In this paper we consider possibility measures in a way similar to fuzzy sets, and therefore Hisdal [20] conditioning has been used (in the second part of this paper [6] we study the concept of independence using Dempster [12] conditioning for possibility measures). We present different comparison criteria, which give rise to different approaches to the concept of independence. Particularly, considering that possibility measures represent imprecise and uncertain knowledge, we try to include this fact in the comparison criteria. Moreover, in order to evaluate the different definitions of independence, a set of properties or axioms that seems reasonable to demand of any relationship that captures the notion of independence has been selected (the well-known *graphoid* axioms [24]), and our definitions are tested against this set. These axioms would also permit us to compare the given definitions with the definitions of independence obtained for other formalisms.

The paper is organized as follows: first, Possibility Theory is briefly summarized. In Section 3, we propose an intuitive approach to the concept of conditional independence and also review an axiomatic framework for this concept. In Section 4, we formalize the intuitive definitions of independence, and study the set of properties that each one verifies. In Section 5, we consider the marginal problem, i.e., how can we construct a possibility measure from a set of marginals and, as an application, we study the problem of how to store large possibility distributions efficiently using independence relationships between variables. Finally, Section 6 contains the concluding remarks.

2. Possibility measures

A fuzzy measure [31] permits us to consider problems where uncertainty appears as ambiguity, that is to say, problems where it is difficult to select one alternative among a set of possible ones. A possibility measure is a particular case of a fuzzy measure. Formally, let D_X be a reference set where a variable X takes its values, and A, B events in D_X. A fuzzy measure associates to each event, $A \subseteq D_X$, a value in $[0,1]$, denoted by $g(A)$, which expresses the belief in the occurrence of the event A. By convention, $g(A)$ increases as the confidence in the event A increases. In any case, the following requirements must be satisfied:

1. Limit values: $g(\emptyset) = 0$ and $g(D_X) = 1$.
2. Monotonicity condition: $\forall A, B \subseteq D_X$, if $A \subseteq B$, then $g(A) \leqslant g(B)$.

When we require other additional properties, we obtain more specific types of fuzzy measures, for instance upper and lower probabilities, evidence measures, probability measures or possibility measures. We will only review briefly the main concepts of possibility theory which are necessary for subsequent development. Several texts can be used for a more detailed treatment of possibility theory [15, 32].

Possibility measures can be obtained as a limit case for the monotonicity condition. That is, we can define a possibility measure, denoted by Π, by means of

1. $\Pi(\emptyset) = 0$ and $\Pi(D_X) = 1$.
2. $\forall A, B \subseteq D_X$ $\Pi(A \cup B) = \max\{\Pi(A), \Pi(B)\}$.

When the set D_X is finite, every possibility measure Π can be defined through the values on the singletons of D_X as follows:

$$\forall A \subseteq D_X, \quad \Pi(A) = \max\{\pi(x) \mid x \in A\},$$

where $\pi(x) = \Pi(\{x\})$, and π is a function from D_X into $[0, 1]$ called *possibility distribution*. This function is obviously normalized, i.e., there is x_0 in D_X such that $\pi(x_0) = 1$.

When our knowledge is fuzzy and imprecise, a possibility measure is a natural tool to formalize the uncertainty. If we have a variable X taking values in D_X, $\pi(x)$ can be interpreted as the degree for $x \in D_X$ to be the possible value that the variable takes. Therefore, $\Pi(A)$ is the possibility that the variable X takes any element in A as its value.

Associated to a possibility measure, there is a Necessity measure, its dual measure, defined by

$$N(A) = 1 - \Pi(\bar{A})$$

where $N(A)$ expresses the belief in the event A being necessarily certain, and says that an event is more certain the smaller the possibility is for its opposite event.

In order to study the concept of possibilistic independence, let us review the concepts of marginal possibility measures and conditional possibility measures.

Let X and Y be two variables taking their values in the finite sets $D_X = \{x_1, x_2, \ldots, x_n\}$ and $D_Y = \{y_1, y_2, \ldots, y_m\}$, respectively. Suppose that we have a joint possibility distribution over the cartesian product $D_X \times D_Y$. The concept of marginal possibility measure can be obtained as a particular case of the concept of marginal fuzzy measure, and leads to the following standard definition:

Definition 1. Given a bidimensional possibility measure $\Pi : \mathscr{P}(D_X \times D_Y) \to [0, 1]$, the marginal possibility measure on X, Π_X, (analogously on Y) is to be defined by means of

$$\Pi_X(A) = \Pi(A \times D_Y), \quad \forall A \subseteq D_X.$$

It is easy to see that the marginal measures are possibility measures too. Thus, the marginal possibility distribution on X (analogously on Y) can be defined as follows:

$$\forall x \in D_X, \quad \pi_X(x) = \Pi_X(\{x\}) = \Pi(\{x\} \times D_Y) = \max_{y \in D_Y} \pi(x, y). \tag{1}$$

However, the idea of conditional fuzzy measure is not clear enough [8], and there are several definitions for this concept. Therefore, we can also find different alternatives to define conditional possibility. For instance, when we consider possibility measures in a way closer to fuzzy sets, we can use Hisdal conditioning [20, 13], and if we consider possibility measures as particular cases of plausibility measures (consonant plausibility measures), we can use a particularization of Dempster conditioning [12, 26] for these measures.

In this paper, we focus on Hisdal conditioning. A similar study of the concept of independence using Dempster conditioning for possibility measures, is to be found in the second part of this paper [6].

Hisdal defined the conditional possibility measure, denoted by $\Pi_h(A \mid B)$, as the solution to the equation $\Pi(A, B) = \min\{\Pi(A \mid B), \Pi(B)\}$. This definition is based on Bayes's rule for probabilities, where the minimum operator replaces the product operator. The conditional measure can be obtained as the least specific solution of this equation:

Definition 2. Given a bidimensional possibility measure $\Pi : \mathscr{P}(D_X \times D_Y) \to [0, 1]$, the conditional possibility measure on X, $\Pi_h(. \mid B)$, given the event $B \subseteq D_Y$, is to be defined by means of

$$\forall A \subseteq D_X, \quad \Pi_h(A \mid B) = \begin{cases} \Pi(A, B) & \text{if } \Pi(A, B) < \Pi_Y(B), \\ 1 & \text{if } \Pi(A, B) = \Pi_Y(B). \end{cases}$$

Analogously, we can obtain the possibility distribution on X, conditioned to the event $[Y=y]$, denoted by $\pi_h(.\mid y)$, as

$$\forall x \in D_X, \quad \pi_h(x \mid y) = \begin{cases} \pi(x,y) & \text{if } \pi(x,y) < \pi_Y(y), \\ 1 & \text{if } \pi(x,y) = \pi_Y(y). \end{cases} \tag{2}$$

Obviously, the marginal and conditional possibility operators can be easily extended from the two-dimensional to the n-dimensional case. From now on, to simplify the notation of a marginal possibility distribution, we will drop the subindex, thus writing $\pi(x)$ and $\pi(y)$ instead of $\pi_X(x)$ and $\pi_Y(y)$, respectively.

3. Conditional independence: definitions and properties

In this section, we present an intuitive approach to the concept of conditional independence, and then, by considering independence relationships in an abstract way, we select a set of properties that seems reasonable to demand of any relationship attempting to capture the intuitive idea of independence. We shall denote by $I(X \mid Z \mid Y)$ the assertion '*X is independent of Y, given Z*', where X, Y, Z represent disjoint variables or subsets of variables. Given a knowledge state Z, a natural approach to the concept of conditional independence is to compare the previous knowledge about the variable X with the knowledge that we obtain after finding out a new piece of information about the variable Y. That is to say, given Z, compare the a priori and a posteriori information. Different comparison criteria may be considered, and therefore different definitions of independence can be obtained.

Perhaps, the most obvious way to define conditional independence, $I(X \mid Z \mid Y)$, is the following:

Definition 3 (*Not modifying the information*). Given any value of the variable Z, knowing the value that the variable Y takes does not modify our information about the values that variable X can take.

Considering that possibility theory permits us to represent uncertain and imprecise knowledge, it might be too strict to demand that our knowlegde does not change in any way after conditioning. Moreover, the problem gets worse if we note that the knowledge (i.e., the possibility values) must be drawn from a data set or from human judgments.

An alternative approach, where the concept of independence is relaxed, might be to assert the independence when we do not gain additional information after conditioning, although we may lose information, i.e., after conditioning we obtain information which is less precise than the original (the rationale is that independent events should not contribute additional information, an idea already suggested in [4, 5]):

Definition 4 (*Not gaining information*). Given any value of the variable Z, if we know the value that the variable Y takes, we do not gain additional information about the values of X.

Another approach to define independence, where the comparison operator is abstracted, is to consider that, before and after conditioning, we obtain a similar information:

Definition 5 (*Obtaining similar information*). Given any value of the variable Z, if we know the value that the variable Y takes, we obtain information about X similar to that prior to learning the value of Y.

It must be noted that the three definitions above are not equivalent. However, some relationships can be established among each other: the fulfilment of Definition 3 implies that Definitions 4 and 5 must also be verified, but the converse is not necessarily true. Anyway, the proposed definitions are rather abstract and, in order to obtain a procedural way of testing independence, in the next section we shall formalize the given comparison criteria, so that different particularizations for the concepts of not modifying, not gaining and obtaining similar information will be considered.

The definitions above represent, semantically, the concept of independence in different ways. Associated with this concept there is a set of properties that identifies the essence of the qualitative notion of independence (which should be common to most of the different formalisms of knowledge representation). These properties have been identified after studying of the concept of conditional independence in probability theory [11], the study of embedded multivalued dependence in database theory [17] and that of dependence graphs [24]. This set of properties, that we call axioms, may also be considered as a set of rules to infer new independence relationships from an initial set, and provides an appropiate way for comparing different formalizations of independence. These axioms are the following:

A1. *Trivial Independence*: $I(X \mid Z \mid \emptyset)$.
A2. *Symmetry*: $I(X \mid Z \mid Y) \Rightarrow I(Y \mid Z \mid X)$.
A3. *Decomposition*: $I(X \mid Z \mid Y \cup W) \Rightarrow I(X \mid Z \mid Y)$.
A4. *Weak Union*: $I(X \mid Z \mid Y \cup W) \Rightarrow I(X \mid Z \cup Y \mid W)$.
A5. *Contraction*: $I(X \mid Z \mid Y)$ and $I(X \mid Z \cup Y \mid W) \Rightarrow I(X \mid Z \mid Y \cup W)$.
A6. *Intersection*: $I(X \mid Z \cup W \mid Y)$ and $I(X \mid Z \cup Y \mid W) \Rightarrow I(X \mid Z \mid Y \cup W)$.

The intuitive interpretation of these axioms is as follows: Trivial Independence says that our knowledge does not change if we do not obtain any new piece of information. Symmetry asserts that in any state of knowledge Z, if Y tells us nothing new about X, then X tells us nothing new about Y. Decomposition establishes that if two combined pieces of information Y and W are considered irrelevant to X, then each separate piece of information is also irrelevant. Weak union asserts that learning the irrelevant information Y cannot help the irrelevant information W become relevant to X. Contraction states that if two pieces of information, X and W, are irrelevant to each other after knowing irrelevant information Y, then they were also irrelevant before knowing Y. Together, Weak union and Contraction mean that irrelevant information should not modify the nature of being relevant or irrelevant of other propositions in the system. Finally, Intersection asserts that if two combined items of information, Y and W, are relevant to X, then at least one of them is also relevant to X, when the other other is added to our previous state of knowledge, Z. A Dependency model is called *semi-graphoid* if it verifies the axioms A1–A5, and *graphoid* if it satisfies the axioms A1–A6. It is well known that the relationship of probabilistic conditional independence satisfies the properties A1–A5, and it also verifies A6 for probability distributions which are strictly positive.

4. Possibilistic conditional independence

In order to formalize the previous intuitive definitions of independence within the framework of possibility theory, we shall consider that X, Y and Z are disjoint subsets in a finite set of n variables, and that π is an n-dimensional joint possibility distribution on these variables. Any generic value that these variables can take on will be denoted by x, y, z, and particular instances for these variables will be denoted by subscripted or Greek letters.

Following the previous definitions of independence, the most natural way to define the possibilistic independence, $I(X \mid Z \mid Y)$, is to compare the a priori and a posteriori information: in any given knowledge state where we know the value of variable Z, say $Z = z$, the a priori information about X is given by the conditional possibility distribution $\pi_h(x \mid z)$. After knowing the value of variable Y, $Y = y$, the a posteriori distribution about X is represented by $\pi_h(x \mid yz)$. Therefore, different definitions of possibilistic independence can be obtained by considering different comparison operators $\odot$ between conditional possibilities, i.e., $\pi_h(x \mid z) \odot \pi_h(x \mid yz)$. The same idea has been developed in the probabilistic framework, where independence is tested using the equality operator as the comparison criterion, i.e., $P(x \mid yz) = P(x \mid z)$. A similar idea was used by Shenoy [27] in the more general framework of Valuation-Based Systems, by Studený [30] for different formalisms to represent the uncertainty, and by Fonck [19] for possibility measures.

The intuitive definitions of independence considered in the previous section may be formalized by considering different comparison operators, $\odot$, representing the notions of not modifying, not gaining, and obtaining similar information after conditioning. In the following subsections, we shall study these operators, and thereafter, for each one of them, we shall study the set of axioms that the corresponding definition of independence verifies.

4.1. The equality operator

The first idea is to assert the independence when we *do not modify* the information after conditioning. In this case, as the comparison operator we can use an equality relationship among conditional distributions:

Definition 6. (H1) *Not modifying the information.*

$$I(X \mid Z \mid Y) \Leftrightarrow \pi_h(x \mid yz) = \pi_h(x \mid z), \quad \forall x, y, z. \tag{3}$$

Using this definition, we obtain the following properties:

Proposition 1. *The independence relationship* H1 *satisfies the axioms* A1, A3–A6.

Proof. The proofs for the axioms A1 and A5 are immediate. Axiom A4 will follow directly from A3. So, we only prove the axioms A3 and A6.

Decomposition: $I(X \mid Z \mid Y \cup W) \Rightarrow I(X \mid Z \mid Y)$.

By definition, $I(X \mid Z \mid Y \cup W)$ is equivalent to $\pi_h(x \mid yzw) = \pi_h(x \mid z)$, $\forall xyzw$. We want to prove the equality $\pi_h(x \mid yz) = \pi_h(x \mid z)$, $\forall xyz$. Consider x, y, z fixed; we know that $\pi(xyz) = \max_{w \in D_W} \pi(xyzw)$. Let $\kappa \in D_W$ be the value where the maximum is reached, i.e., $\pi(xyz) = \pi(xyz\kappa)$. As $I(X \mid Z \mid Y \cup W)$ holds, we have $\pi_h(x \mid yz\kappa) = \pi_h(x \mid z)$. Now, we shall consider the different values that $\pi_h(x \mid yz\kappa)$ can take on:

(a) Suppose that $\pi_h(x \mid yz\kappa) = \pi(xyz\kappa)$, i.e., $\pi(xyz\kappa) < \pi(yz\kappa) \leqslant 1$.

Then $\pi_h(x \mid z) = \pi(xyz\kappa) < 1$, and therefore, we find that $\pi_h(x \mid z) = \pi(xz) < \pi(z)$, and so $\pi(xyz) = \pi(xyz\kappa) = \pi(xz) < \pi(z)$.

To obtain the desired result, it is suffice to check that $\pi(xyz) < \pi(yz)$ is true, because in this case we have $\pi_h(x \mid yz) = \pi(xyz) = \pi(xz) = \pi_h(x \mid z)$. But we know that $\pi(yzw) \leqslant \pi(yz)$, $\forall w$, and since $\pi(xyz\kappa) < \pi(yz\kappa)$, we find that $\pi(xyz) = \pi(xyz\kappa) < \pi(yz\kappa) \leqslant \pi(yz)$.

(b) Suppose that $\pi_h(x \mid yz\kappa) = 1$, i.e., $\pi(xyz\kappa) = \pi(yz\kappa)$.

Therefore, we deduce $\pi_h(x \mid z) = 1$. As $\pi_h(x \mid yzw) = \pi_h(x \mid z) = 1$, $\forall w$, then we have $\pi(xyzw) = \pi(yzw)$ $\forall w$, and we also obtain $\pi(xyz) = \pi(yz)$. So, $\pi_h(x \mid yz) = 1 = \pi_h(x \mid z)$.

Intersection: $I(X \mid Y \cup Z \mid W)$ and $I(X \mid Z \cup W \mid Y) \Rightarrow I(X \mid Z \mid Y \cup W)$.

Considering the left hand side of the implication, we know that $\pi_h(x \mid yzw) = \pi_h(x \mid yz) = \pi_h(x \mid zw)$ $\forall xyzw$. We have to prove that $\pi_h(x \mid yzw) = \pi_h(x \mid z)$ $\forall xyzw$.

Given x, z, let us consider any two values $y_0 \in D_Y$ and $w_0 \in D_W$. From the hypothesis we know that $\pi_h(x \mid y_0 z w_0) = \pi_h(x \mid y_0 z) = \pi_h(x \mid z w_0)$, and $\pi_h(x \mid yzw_0) = \pi_h(x \mid yz) = \pi_h(x \mid zw_0)$ $\forall y$. Therefore, we obtain $\pi_h(x \mid y_0 z w_0) = \pi_h(x \mid yz)$ $\forall y$. Let us study the two different cases that may appear:

(a) $\pi_h(x \mid y_0 z w_0) = 1$. In this case we have $\pi_h(x \mid yz) = 1$ $\forall y$, hence $\pi(xyz) = \pi(yz)$ $\forall y$, and $\pi(xz) = \pi(z)$. So, we have $\pi_h(x \mid z) = 1 = \pi_h(x \mid y_0 z w_0)$.

(b) $\pi_h(x \mid y_0 z w_0) < 1$. Then we have $1 > \pi(xy_0 z w_0) = \pi_h(x \mid y_0 z w_0) = \pi_h(x \mid yz) = \pi(xyz)$ $\forall y$, and $\pi(xyz) < \pi(yz)$ $\forall y$. So, on one hand we deduce $\pi(xz) < \pi(z)$, and therefore $\pi_h(x \mid z) = \pi(xz)$; on the other hand, from $\pi(xy_0 z w_0) = \pi(xyz)$ $\forall y$ we obtain $\pi(xy_0 z w_0) = \pi(xz)$. So, we conclude $\pi_h(x \mid z) = \pi(xz) = \pi(xy_0 z w_0) = \pi_h(x \mid y_0 z w_0)$.

Therefore, in any case we have $\pi_h(x \mid y_0 z w_0) = \pi_h(x \mid z)$. □

The only axiom excluded from the previous proposition is Symmetry. The following counterexample proves that this axiom does not hold in general. It shows that $I(X \mid \emptyset \mid Y) \not\Rightarrow I(Y \mid \emptyset \mid X)$, where X and Y are variables taking their values in the sets $D_X = \{x_1, x_2, x_3\}$ and $D_Y = \{y_1, y_2, y_3\}$, respectively:

xy	$\pi(xy)$
$x_1 y_1$	1.0
$x_1 y_2$	0.6
$x_1 y_3$	0.7
$x_2 y_1$	0.5
$x_2 y_2$	0.5
$x_2 y_3$	0.5
$x_3 y_1$	0.4
$x_3 y_2$	0.4
$x_3 y_3$	0.4

We can see that $\pi_h(x_1 \mid y) = \pi(x_1) = 1$, $\pi_h(x_2 \mid y) = \pi(x_2) = 0.5$ and $\pi_h(x_3 \mid y) = \pi(x_3) = 0.4$, i.e., $\pi_h(x \mid y) = \pi(x)$, $\forall xy$. So, we have $I(X \mid \emptyset \mid Y)$; however $\pi_h(y_2 \mid x_2) = 1 \neq 0.6 = \pi(y_2)$, and therefore $\neg I(Y \mid \emptyset \mid X)$.

As H1 is not symmetric, it might be interesting to build a symmetrical version of H1 by defining a new relationship I^s as follows:

$$I^s(X \mid Z \mid Y) \Leftrightarrow I(X \mid Z \mid Y) \text{ and } I(Y \mid Z \mid X). \tag{4}$$

In fact this is one of the definitions of independence proposed by Fonck [19], who proved that I^s verifies all the axioms A1–A6. However, we are going to show that this definition of independence is rather restrictive, so, unfortunately, we gain an additional property, symmetry, but at the expense of losing important representation capabilities. The following proposition proves an interesting characterization of the definition of independence H1, which is useful for our purposes:

Proposition 2. *The definition* H1 *is equivalent to*

$$I(X \mid Z \mid Y) \Leftrightarrow \pi(xyz) = \pi(xz) \wedge \pi(yz) \text{ and } (\pi(xz) = \pi(z) \text{ or } \pi(xz) < \pi(yz)), \ \forall x, y, z. \tag{5}$$

Proof. *Necessary condition*: we know that $\pi_h(x \mid yz) = \pi_h(x \mid z)$.

If $\pi_h(x \mid z) = \pi(xz)$ then $\pi(xz) < \pi(z) \leqslant 1$ and $\pi_h(x \mid yz) = \pi_h(x \mid z) = \pi(xz) < 1$. So, we have $\pi_h(x \mid yz) = \pi(xyz) < \pi(yz)$ and $\pi(xyz) = \pi(xz) < \pi(yz)$. Therefore in this case $\pi(xyz) = \pi(xz) \wedge \pi(yz)$ and $\pi(xz) < \pi(yz)$.

If $\pi_h(x \mid z) = 1$ then $\pi(xz) = \pi(z)$ and $\pi_h(x \mid yz) = \pi_h(x \mid z) = 1$. So, $\pi(xyz) = \pi(yz) \leqslant \pi(z) = \pi(xz)$, and therefore in this case $\pi(xyz) = \pi(xz) \wedge \pi(yz)$ and $\pi(xz) = \pi(z)$.

Sufficient condition: we know that $\pi(xyz) = \pi(xz) \wedge \pi(yz)$ and $\pi(xz) = \pi(z)$ or $\pi(xz) < \pi(yz)$.

If $\pi_h(x \mid z) = \pi(xz)$ then $\pi(xz) < \pi(z)$. Therefore, it has to be $\pi(xz) < \pi(yz)$. So, $\pi(xyz) = \pi(xz) < \pi(yz)$, and then $\pi_h(x \mid yz) = \pi(xyz) = \pi(xz) = \pi_h(x \mid z)$.

If $\pi_h(x \mid z) = 1$ then $\pi(xz) = \pi(z)$. So, $\pi(xyz) = \pi(xz) \wedge \pi(yz) = \pi(z) \wedge \pi(yz) = \pi(yz)$, and therefore $\pi_h(x \mid yz) = 1 = \pi_h(x \mid z)$. □

Fonck [19] proved the following characterization of the independence relationship defined in Eq. (4), i.e., the symmetrized version of H1, which can also be easily deduced from Proposition 2:

Proposition 3. *The independence relationship defined in Eq.* (4) *is equivalent to*

$$I^s(X \mid Z \mid Y) \Leftrightarrow \pi(xyz) = \pi(xz) \wedge \pi(yz) \text{ and } \pi(z) = \pi(xz) \vee \pi(yz) \quad \forall x, y, z. \tag{6}$$

Using the previous result, we can easily prove the following proposition:

Proposition 4. *If $I^s(X \mid Z \mid Y)$ then $\pi(x) = 1 \ \forall x \in D_X$ or $\pi(y) = 1 \ \forall y \in D_Y$.*

Proof. From Proposition 3, we know that $\pi(z) = \pi(xz) \vee \pi(yz) \ \forall xyz$. Then, $1 = \max_z \pi(z) = \max_z(\pi(xz) \vee \pi(yz)) = (\max_z \pi(xz)) \vee (\max_z \pi(yz)) = \pi(x) \vee \pi(y) \ \forall xy$.

Now, if we suppose that $\exists x_0 \in D_X$ such that $\pi(x_0) \neq 1$, then $1 = \pi(x_0) \vee \pi(y) \ \forall y$, and therefore $\pi(y) = 1 \ \forall y$. □

The previous proposition proves that the symmetrical version of H1 is a quite restrictive concept of independence, because it implies that, in order for two variables to be conditionally independent, the available information about at least one of these variables must be null.

4.2. The inclusion operator

In this subsection we study the independence of X from Y in the sense of not gaining information about X after conditioning to Y. This definition establishes that we do not gain information, but permits us to lose some information after conditioning. The loss of information has to be interpreted in the following sense: after conditioning, we obtain information which is similar to (or coherent with) the information we had before conditioning, but less precise. This situation cannot happen in the probabilistic case, because probability measures are always maximally precise. However, possibility measures can express different degrees of precision.

Let X be a variable taking its values in D_X, and A, B subsets of D_X. If $A \subseteq B$, then to state that the value of X is in A is more informative than asserting that the value of X is in B. An analogous reasoning can be applied when we consider possibility distributions (which, in a sense, can be seen as generalized sets, i.e., fuzzy sets): if the possibility distribution π is always less than or equal to the possibility distribution π', i.e., $\pi(x) \leqslant \pi'(x)$, $\forall x \in D_X$, then π' is less precise than π, or equivalently, π gives more information about the value of the variable X than π' does. This concept of a possibility distribution being more or less informative than another one is well expressed by the following definition of inclusion [15].

Definition 7. Let π and π' be two possibility distributions defined in the same reference set D_X. Then π' is said to be included in π (or π' is less informative than π) if and only if $\pi(x) \leqslant \pi'(x)$, $\forall x \in D_X$.

Using this inclusion relationship between possibility distributions, the definition of independence in the sense of not gaining information may be formally expressed in the following way:

Definition 8. (H2) *Not gaining information.*

$$I(X \mid Z \mid Y) \Leftrightarrow \pi_h(x \mid z) \leqslant \pi_h(x \mid yz), \quad \forall x, y, z. \tag{7}$$

The following proposition provides an interesting and revealing characterization of this type of independence, which does not use conditioning.

Proposition 5. *The definition* H2 *is equivalent to*

$$I(X \mid Z \mid Y) \Leftrightarrow \pi(xyz) = \pi(xz) \wedge \pi(yz), \quad \forall x, y, z. \tag{8}$$

Proof. *Necessary condition*: $\pi_h(x \mid yz) \geqslant \pi_h(x \mid z) \Rightarrow \pi(xyz) = \pi(xz) \wedge \pi(yz)$.

1. Suppose that $\pi_h(x \mid yz) = \pi(xyz)$. Then, we find that $\pi(xyz) < \pi(yz)$. We shall consider the two different alternatives for $\pi_h(x \mid z)$.
   (a) $\pi_h(x \mid z) = \pi(xz)$, i.e., $\pi(xz) < \pi(z)$: then we have $\pi(xyz) \geqslant \pi(xz)$, and as $\pi(xz) \geqslant \pi(xyz)$ is always true for possibilities, the only alternative is that $\pi(xz) = \pi(xyz) < \pi(yz)$, and therefore $\pi(xyz) = \pi(xz) \wedge \pi(yz)$.
   (b) $\pi_h(x \mid z) = 1$: in this case we have $\pi_h(x \mid yz) \geqslant \pi_h(x \mid z) = 1$, which contradicts the fact that $\pi_h(x \mid yz) = \pi(xyz) < \pi(yz) \leqslant 1$. So, this second alternative cannot occur.
2. Suppose that $\pi_h(x \mid yz) = 1$, so that $\pi(xyz) = \pi(yz)$. As $\pi(xyz) \leqslant \pi(xz)$, then $\pi(xyz) = \pi(yz) \leqslant \pi(xz)$, and therefore $\pi(xyz) = \pi(xz) \wedge \pi(yz)$.

Sufficient condition: $\pi(xyz) = \pi(xz) \wedge \pi(yz) \Rightarrow \pi_h(x \mid yz) \geqslant \pi_h(x \mid z)$.

1. Suppose that $\pi(xyz) = \pi(yz) \leqslant \pi(xz)$. Then $\pi_h(x \mid yz) = 1$, and therefore $\pi_h(x \mid yz) \geqslant \pi_h(x \mid z)$.
2. Suppose that $\pi(xyz) = \pi(xz) < \pi(yz) \leqslant \pi(z)$. Then $\pi_h(x \mid yz) = \pi(xyz) = \pi(xz)$ and $\pi_h(x \mid z) = \pi(xz)$, which implies $\pi_h(x \mid yz) = \pi_h(x \mid z)$. □

If we consider the particular case of marginal independence, i.e., when the conditioning subset of variables is empty, $Z = \emptyset$, from Eq. (8) we obtain

$$I(X \mid \emptyset \mid Y) \Leftrightarrow \pi(xy) = \pi(x) \wedge \pi(y).$$

This is the concept of non-interactivity proposed by Zadeh [32] for possibility measures or fuzzy sets. So, our concept of independence H2 is that we could call *conditional non-interactivity.*

There is another interesting property that relates the possibilistic independence H2 to probabilistic independence: when we consider probability theory, the relationship $I(X \mid Z \mid Y)$ is verified if and only if $P(x \mid yz) = P(x \mid z)$. This expression is equivalent to $P(xy \mid z) = P(x \mid z)P(y \mid z)$. A similar expression can be established considering possibility distributions instead of probability distributions, where the product operator is replaced by the minimum.

Proposition 6. *The definition* H2 *is equivalent to*

$$I(X \mid Z \mid Y) \Leftrightarrow \pi_h(xy \mid z) = \pi_h(x \mid z) \wedge \pi_h(y \mid z) \quad \forall x, y, z. \tag{9}$$

Proof. In order to prove the proposition we use the characterization of H2 given in Proposition 5.

Necessary condition: $\pi(xyz) = \pi(xz) \wedge \pi(yz) \Rightarrow \pi_h(xy \mid z) = \pi_h(x \mid z) \wedge \pi_h(y \mid z)$.

Two possible cases may be considered:

1. Suppose that $\pi(xyz) < \pi(z)$: as $\pi(xyz) = \pi(xz) \wedge \pi(yz)$, then at least one of $\pi(xz), \pi(yz)$ is less than $\pi(z)$ and equal to $\pi(xyz)$. Suppose that this relation holds for $\pi(xz)$, i.e., $\pi(xyz) = \pi(xz) < \pi(z)$, and $\pi(xz) \leqslant \pi(yz)$. In that case, $\pi_h(xy \mid z) = \pi_h(x \mid z)$. Moreover, $\pi_h(y \mid z) \geqslant \pi(yz) \geqslant \pi(xz) = \pi_h(x \mid z)$. Therefore $\pi_h(xy \mid z) = \pi_h(x \mid z) \wedge \pi_h(y \mid z)$.
2. Suppose that $\pi(xyz) = \pi(z)$: in that case, $\pi(xyz) = \pi(xz) = \pi(yz) = \pi(z)$. So, $\pi_h(xy \mid z) = \pi_h(x \mid z) = \pi_h(y \mid z) = 1$ and then $\pi_h(xy \mid z) = \pi_h(x \mid z) \wedge \pi(y \mid z)$.

Sufficient condition: $\pi(xy \mid z) = \pi_h(x \mid z) \wedge \pi(y \mid z) \Rightarrow \pi(xyz) = \pi(xz) \wedge \pi(yz)$.

Once again, we consider two different cases:
1. Suppose that $\pi(xyz) < \pi(z)$: Then, $\pi_h(xy\,|\,z) = \pi(xyz)$. Consider that $\pi_h(x\,|\,z)$ is equal to the minimum between $\pi_h(x\,|\,z)$ and $\pi_h(y\,|\,z)$, i.e., $\pi_h(xy\,|\,z) = \pi_h(x\,|\,z)$ (the case in which $\pi_h(xy\,|\,z) = \pi_h(y\,|\,z)$ is analogous). Then, we have $\pi_h(x\,|\,z) = \pi_h(xy\,|\,z) = \pi(xyz) < \pi(z) \leqslant 1$, so that $\pi_h(x\,|\,z) = \pi(xz)$ and therefore $\pi(xyz) = \pi(xz)$. Moreover, $\pi(yz) \geqslant \pi(xyz) = \pi(xz)$, hence $\pi(xyz) = \pi(xz) \wedge \pi(yz)$.
2. Suppose that $\pi(xyz) = \pi(z)$: then we have $\pi(xyz) = \pi(xz) = \pi(yz) = \pi(z)$, and so $\pi(xyz) = \pi(xz) \wedge \pi(yz)$. □

Now, let us study the axioms that the definition of independence H2 satisfies:

Proposition 7. *The independence relationship* H2 *satisfies the axioms* A1–A5.

Proof. The proof is very simple, using Proposition 5, so we omit it. □

We shall show that the Intersection axiom (A6) does not hold in general by using a counterexample. Let X, Y, Z, W be bivaluated variables with the following joint possibility distribution:

x_1yzw	$\pi(x_1yzw)$
$x_1y_1z_1w_1$	1.0
$x_1y_1z_1w_2$	0.8
$x_1y_1z_2w_1$	1.0
$x_1y_1z_2w_2$	1.0
$x_1y_2z_1w_1$	0.9
$x_1y_2z_1w_2$	1.0
$x_1y_2z_2w_1$	1.0
$x_1y_2z_2w_2$	1.0

x_2yzw	$\pi(x_2yzw)$
$x_2y_1z_1w_1$	0.9
$x_2y_1z_1w_2$	0.8
$x_2y_1z_2w_1$	1.0
$x_2y_1z_2w_2$	1.0
$x_2y_2z_1w_1$	0.9
$x_2y_2z_1w_2$	1.0
$x_2y_2z_2w_1$	1.0
$x_2y_2z_2w_2$	1.0

It may be seen that $\pi(xyzw) = \pi(xyz) \wedge \pi(yzw)$ and $\pi(xyzw) = \pi(xzw) \wedge \pi(yzw)$, $\forall xyzw$, i.e. $I(X\,|\,Z \cup Y\,|\,W)$ and $I(X\,|\,Z \cup W\,|\,Y)$. However, $\pi(x_2y_1z_1w_1) = 0.9 \neq 1 = \{\pi(x_2z_1) \wedge \pi(y_1z_1w_1)\}$, and therefore $I(X\,|\,Z\,|\,Y \cup W)$ does not hold.

4.3. Default conditioning

In the previous subsection we studied a concept of independence that allows a loss of precision after applying conditioning. Then, the following question naturally arises: *If, after conditioning, we lose information, would it be more convenient to keep the initial information*? This idea may be debatable, but it represents in some sense a default rule: If in a very specific context we do not have much information, then we can use the information available in a less specific context. In this subsection we discuss and formalize this idea by introducing a new conditioning operator, that we call *default conditioning*. The following example gives us an idea about the semantics of the default conditioning.

Example 1. Let us suppose that X represents the sentence 'Number of eggs an individual eats for breakfast'. The possible values are, for simplicity's sake, zero (0), one (1) and two or more (2+). Variable Y means 'Town where people are from', with values London (L) and Other (O). Variable Z means 'Nationality of people', with values British (B) and Non-British (NB). We have information about the breakfast habits for people in general, for British people and for Non-British people, in the form of possibility distributions $\pi(x)$, $\pi(x\,|\,B)$ and $\pi(x\,|\,NB)$, given below. However, we have no information about breakfast habits for people living in London (i.e., $\pi(x \mid L) = 1\ \forall x$).

x	0	1	2+
$\pi(x)$	1	1	0.8

x	0	1	2+
$\pi(x \mid B)$	0.7	1	1

x	0	1	2+
$\pi(x \mid NB)$	1	1	0.5

If we are told that Peter is a British man, then we update our general information, passing from $\pi(x)$ to $\pi(x \mid B)$. If, after discovering that Peter is from London, we should update again our belief about Peter's capacity for eating eggs (from $\pi(x \mid B)$ to $\pi(x \mid B, L)$); but in this case the updating produces the trivial conclusion: we do not know anything. So, we have obtained less precise information after conditioning to L. However, we could use the following reasoning: 'Knowing that Londoners are British, if we have no information about their particular customs, it might be reasonable to think that their habits are those that are usual for the British. Then, we can assign, by default, the possibility distribution given for the British'. It is not certain that the habits of people from London are the same as those of the British in general, but in the absence of specific information, we use the information available for British people.

Finally, consider that we focus the study on British people with a high level of cholesterol (C). In that case, as eggs are harmful in terms of cholesterol, the following possibility distribution could be associated

x	0	1	2+
$\pi(x \mid B, C)$	1	0.6	0.4

In that case, as after conditioning, we obtain more precise information (more exactly, information which is not less precise), the previous reasoning cannot be done, and we should use $\pi(x \mid B, C)$ instead of $\pi(x \mid B)$.

By generalizing this reasoning, we can state that, if after conditioning we obtain a less informative possibility distribution, we preserve the previous (a priori) more precise information. Otherwise, we use the a posteriori possibility distribution. Formally, this idea implies a change in the notion of conditioning, which may be called *Hisdal default conditioning*, denoted by $\pi_{\mathrm{h_c}}(.\mid.)$:

$$\pi_{\mathrm{h_c}}(x \mid y) = \begin{cases} \pi(x) & \text{if } \pi_{\mathrm{h}}(x \mid y) \geqslant \pi(x)\ \forall x, \\ \pi_{\mathrm{h}}(x \mid y) & \text{if } \exists x' \text{ such that } \pi_{\mathrm{h}}(x' \mid y) < \pi(x'). \end{cases} \tag{10}$$

Using the default conditioning $\pi_{\mathrm{h_c}}(.\mid.)$, a new independence relationship can be defined as follows:

Definition 9. (H3) *Default conditioning.*

$$I(X \mid Z \mid Y) \Leftrightarrow \pi_{\mathrm{h_c}}(x \mid yz) = \pi_{\mathrm{h_c}}(x \mid z), \quad \forall x, y, z. \tag{11}$$

Proposition 8. *The independence relationship* H3 *satisfies the axioms* A1 *and* A3–A5. *If the possibility distribution is strictly positive then axiom* A6 *also holds.*

Proof. The proof for axioms A1 and A5 is immediate. Axiom A4 can be obtained directly from A3. So, we shall only prove axioms A3 and A6.

Decomposition: $I(X \mid Z \mid Y \cup W) \Rightarrow I(X \mid Z \mid Y)$.

We know that $\pi_{\mathrm{h_c}}(x \mid yzw) = \pi_{\mathrm{h_c}}(x \mid z)\ \forall xyzw$.

Given any z, first let us suppose that $\pi_{\mathrm{h_c}}(.\mid z) = \pi_{\mathrm{h}}(.\mid z)$. In this case $\delta \in D_X$ has to exist such that $\pi_{\mathrm{h}}(\delta \mid z) < \pi(\delta)$. Then, $\forall yw$, $\pi_{\mathrm{h_c}}(\delta \mid yzw) = \pi_{\mathrm{h_c}}(\delta \mid z) = \pi_{\mathrm{h}}(\delta \mid z) < \pi(\delta)$, hence $\pi_{\mathrm{h_c}}(x \mid yzw) = \pi_{\mathrm{h}}(x \mid yzw)\ \forall xyw$.

Therefore we obtain $\pi_h(x \mid yzw) = \pi_h(x \mid z)$ $\forall xyw$. Now, using a reasoning similar to the one employed in Proposition 1 (Decomposition axiom), we can conclude that $\pi_h(x \mid yz) = \pi_h(x \mid z)$ $\forall xy$. Moreover, $\pi_h(\delta \mid yz) = \pi_h(\delta \mid z) < \pi(\delta)$, and therefore $\pi_{h_c}(. \mid yz) = \pi_{h_c}(. \mid z)$ $\forall y$.

Now, let us suppose that $\pi_{h_c}(. \mid z) = \pi(.)$. This means that $\pi_h(x \mid z) \geqslant \pi(x)$ $\forall x$. Then, $\pi_{h_c}(x \mid yzw) = \pi_{h_c}(x \mid z) = \pi(x)$ $\forall xyw$ and then $\pi_h(x \mid yzw) \geqslant \pi(x)$ $\forall xyw$. In this case, the same reasoning used in the proof of Proposition 5 allow us to assert that $\pi(xyzw) = \pi(x) \wedge \pi(yzw)$ $\forall xyw$. Then we obtain $\pi(xyz) = \pi(x) \wedge \pi(yz)$ $\forall xy$, and therefore $\pi_h(x \mid yz) \geqslant \pi(x)$ $\forall xy$, and $\pi_{h_c}(. \mid yz) = \pi(.) = \pi_{h_c}(. \mid z)$ $\forall y$.

So, in any case $\pi_{h_c}(. \mid yz) = \pi_{h_c}(. \mid z)$, hence $I(X \mid Z \mid Y)$.

Intersection: $I(X \mid Z \cup Y \mid W)$ & $I(X \mid Z \cup W \mid Y) \Rightarrow I(X \mid Z \mid Y \cup W)$.

We know that $\pi_{h_c}(x \mid yzw) = \pi_{h_c}(x \mid yz) = \pi_{h_c}(x \mid zw)$ $\forall xyzw$.

Given any z, first, let us suppose that for some y_0 and w_0 it is $\pi_{h_c}(. \mid y_0 z w_0) = \pi_h(. \mid y_0 z w_0)$. This means that $\exists \delta \in D_X$ such that $\pi_h(\delta \mid y_0 z w_0) < \pi(\delta)$. Then it can be easily deduced that $\pi_h(x \mid yzw) = \pi_h(x \mid yz) = \pi_h(x \mid zw) = \pi_{h_c}(x \mid yzw) = \pi_{h_c}(x \mid yz) = \pi_{h_c}(x \mid zw)$ $\forall xyw$. Then, a reasoning similar to the one used in the proof of Proposition 1 (Intersection axiom) permits us to assert that $\pi_h(x \mid yzw) = \pi_h(x \mid z)$ $\forall xyw$. Moreover, we have $\pi_h(\delta \mid z) = \pi_h(\delta \mid y_0 z w_0) < \pi(\delta)$, hence $\pi_{h_c}(. \mid z) = \pi_h(. \mid z)$, and then $\pi_{h_c}(x \mid z) = \pi_h(x \mid z) = \pi_h(x \mid yzw) = \pi_{h_c}(x \mid yzw)$ $\forall xyw$.

Second, let us suppose that $\forall yw$ $\pi_{h_c}(. \mid yzw) \neq \pi_h(. \mid yzw)$, i.e., $\forall yw$ $\pi_{h_c}(. \mid yzw) = \pi(.)$. Then we find that $\pi_h(x \mid yzw) \geqslant \pi(x)$ $\forall xyw$. Once again using a reasoning similar to the one employed in Proposition 5, this condition is equivalent to $\pi(xyzw) = \pi(x) \wedge \pi(yzw)$ $\forall xyw$, and therefore we obtain $\pi(xz) = \pi(x) \wedge \pi(z)$ $\forall x$, which in turn is equivalent to $\pi_h(x \mid z) \geqslant \pi(x)$ $\forall x$. So, $\pi_{h_c}(x \mid z) = \pi(x) = \pi_{h_c}(x \mid yzw)$ $\forall xyw$.

Therefore, in any case we obtain $\pi_{h_c}(. \mid yzw) = \pi_{h_c}(. \mid z)$, and thus $I(X \mid Z \mid Y \cup W)$. □

Finally, in order to show that definition H3 does not verify symmetry, we use the following counterexample, where the variable X takes its values in the set $\{x_1, x_2, x_3\}$, and Y and Z are bivaluated variables. Suppose that the joint possibility distribution for these three variables is the following:

x_1yz	$\pi(x_1yz)$
$x_1y_1z_1$	1
$x_1y_1z_2$	0.6
$x_1y_2z_1$	0.7
$x_1y_2z_2$	0.4

x_2yz	$\pi(x_2yz)$
$x_2y_1z_1$	0.7
$x_2y_1z_2$	0.6
$x_2y_2z_1$	0.7
$x_2y_2z_2$	0.4

x_3yz	$\pi(x_3yz)$
$x_3y_1z_1$	0.3
$x_3y_1z_2$	0.3
$x_3y_2z_1$	0.3
$x_3y_2z_2$	0.3

We can see that $\pi_{h_c}(x \mid yz) = \pi_{h_c}(x \mid z)$ $(= \pi(x))$, for all xyz, hence we obtain $I(X \mid Z \mid Y)$. However, $\pi_{h_c}(y_2 \mid z_2) = 0.4$ and $\pi_{h_c}(y_2 \mid x_3 z_2) = 1$. Therefore $\neg I(Y \mid Z \mid X)$.

The relationships between the three definitions of independence considered so far are as follows: as we might expect, the strictest definition is H1 (i.e., not modifying the information): if an independence relationship has been obtained with this definition, this independence relationship also holds using the other definitions. Moreover, the independence relationship based on the default conditioning (H3) is more restrictive than conditional non-interactivity (H2). The next proposition proves these results.

Proposition 9. *The different definitions of independence using Hisdal conditioning*, H1, H2 *and* H3, *satisfy*:

H1 $\Rightarrow$ H3 $\Rightarrow$ H2.

Proof. H1 $\Rightarrow$ H3: In this case the proof is immediate.

H3 $\Rightarrow$ H2: If $\pi_{h_c}(x \mid yz) = \pi_h(x \mid yz) = \pi_h(x \mid z) = \pi_{h_c}(x \mid z)$, it is obvious that $\pi_h(x \mid yz) \geqslant \pi_h(x \mid z)$.

On the other hand, if $\pi_{h_c}(x \mid yz) = \pi(x) = \pi_{h_c}(x \mid z)$, then we have $\forall x$, $\pi(xyz) = \pi(x) \wedge \pi(yz)$ and $\pi(xz) = \pi(x) \wedge \pi(z)$. Then, $\pi(xyz) = \pi(xz) \wedge \pi(yz)$, and therefore $\pi_h(x \mid yz) \geqslant \pi_h(x \mid z)$. □

The following counterexamples show that the reciprocal relationships do not hold. Using the previous joint possibility distribution we find that H3 $\not\Rightarrow$ H1, that is $I(X \mid Z \mid Y)$ holds using H3; however $\pi_h(x_2 \mid y_2 z_1) = 1 \neq 0.7 = \pi_h(x_2 \mid z_1)$. So, $I(X \mid Z \mid Y)$ does not hold when using H1.

Using the next distribution (X, Y, Z are bivaluated variables), we obtain H2 $\not\Rightarrow$ H3.

xyz	$\pi(xyz)$
$x_1 y_1 z_1$	0.2
$x_1 y_1 z_2$	0.6
$x_1 y_2 z_1$	1.0
$x_1 y_2 z_2$	0.8
$x_2 y_1 z_1$	0.2
$x_2 y_1 z_2$	0.6
$x_2 y_2 z_1$	0.5
$x_2 y_2 z_2$	0.7

We can see that $\pi(xyz) = \pi(xz) \wedge \pi(yz)$ for all xyz, i.e., $I(X \mid Z \mid Y)$ using H2, but, on the other hand, we find that $\pi_{h_c}(x_2 \mid y_1 z_1) = 0.7 \neq 0.5 = \pi_{h_c}(x_2 \mid z_1)$, and therefore $\neg I(X \mid Z \mid Y)$ using H3.

4.4. Similarity operators

Now, we shall consider the concept of independence using similarity relationships between conditional possibility distributions. The general idea is to assert the independence if the distributions obtained before and after conditioning are similar in some sense. More formally:

Definition 10. (H4) *Obtaining similar information.*

$$I(X \mid Z \mid Y) \Leftrightarrow \pi_h(. \mid yz) \simeq \pi_h(. \mid z) \quad \forall y, z, \tag{12}$$

where $\simeq$ is a similarity relationship defined on the set of possibility distributions for the variable X.

Our first purpose is to study, in an abstract way, the properties we have to require from the relationship $\simeq$, in order to guarantee that the associated independence relationship verifies some of the axioms. First, it is evident that transitivity of $\simeq$ guarantees Contraction (A5). If, in addition, $\simeq$ is symmetric, then it can easily be deduced that Decomposition (A3) is verified if and only if Weak Union (A4) is also verified. Moreover, in order to satisfy the basic coherence property of Trivial Independence (A1), it is obvious that $\simeq$ should be reflexive. Therefore, it seems that the equivalence relations are good candidates for defining independence through a similarity relationship $\simeq$ (observe that the equality operator can be considered as a very specific case of similarity operator, and therefore H1 is a particular case of the independence criterion H4). So, we must look for the additional properties that guarantee the fulfilment of Decomposition (A3) and Intersection (A6).

The first property that we require is that the equivalence relationship must be compatible with the maximum operator, in the following sense: if a possibility distribution is similar to each one of a family of possibility distributions, then it is also similar to their maximum.

Definition 11 (*Maximum property*). Given any family of possibility distributions $\{\pi_s(x)\}$, let π' be the possibility distribution obtained as $\pi'(x) = \max_s \pi_s(x)$, and let π be any other possibility distribution. Then, we say that a similarity relationship $\simeq$ verifies the maximum property if and only if

$$\pi_s \simeq \pi \quad \forall s \;\Rightarrow\; \pi' \simeq \pi.$$

This basic property alone does not guarantee the fulfilment of Decomposition. Therefore, other properties should be considered. The first one we propose establishes that the limit value (one) must be preserved, that is:

Definition 12 (*Limit property*). We say that a similarity relationship between possibility distributions, $\simeq$, preserves the limit value if and only if $\forall \pi_1, \pi_2$

$$\pi_1 \simeq \pi_2 \;\Rightarrow\; \forall x\; [\pi_1(x) = 1 \;\Leftrightarrow\; \pi_2(x) = 1].$$

The limit property asserts that two possibility distributions having different maximally possible values cannot be similar. Using these two properties we can prove the following result:

Proposition 10. *Given an equivalence relationship between possibility distributions, $\simeq$, a sufficient condition for the fulfilment of Decomposition is that $\simeq$ verifies the maximum and limit properties. Moreover, these properties also guarantee the fulfilment of intersection.*

Proof.

Decomposition: $I(X \mid Z \mid Y \cup W) \Rightarrow I(X \mid Z \mid Y)$.

We know that $\pi_h(. \mid yzw) \simeq \pi_h(. \mid z)\ \forall yzw$. If we could prove that $\forall yz\ \pi_h(. \mid yz) = \max_w \pi_h(. \mid yzw)$ then, using the maximum property we would obtain $\pi_h(. \mid yz) \simeq \pi_h(. \mid z)\ \forall yz$, and therefore $I(X \mid Z \mid Y)$. So, all we have to prove is that $\pi_h(x \mid yz) = \max_w \pi_h(x \mid yzw)$. Two different cases may be considered:

1. Suppose that $\max_w \pi_h(x \mid yzw) < 1$: in that case, we find that $\pi_h(x \mid yzw) = \pi(xyzw) < \pi(yzw) \leqslant 1\ \forall w$. Therefore we also obtain (by taking the maximum) $\max_w \pi_h(x \mid yzw) = \pi(xyz) < \pi(yz)$. So, $\pi_h(x \mid yz) = \pi(xyz) = \max_w \pi_h(x \mid yzw)$.
2. Suppose that $\max_w \pi_h(x \mid yzw) = 1$: then there exists $w_0 \in D_W$ such that $\pi_h(x \mid yzw_0) = 1$. As $\simeq$ is an equivalence relationship, then we know that $\pi_h(. \mid yzw_0) \simeq \pi_h(. \mid yzw)\ \forall w$. Moreover $\simeq$ preserves the limit value, so that $\forall w \in D_W$, $\pi_h(x \mid yzw) = 1$, i.e., $\pi(xyzw) = \pi(yzw)$, $\forall w$. Therefore $\pi(xyz) = \pi(yz)$ and then $\pi_h(x \mid yz) = 1 = \max_w \pi_h(x \mid yzw)$.

Intersection: $I(X \mid Y \cup Z \mid W)\ \&\ I(X \mid Z \cup W \mid Y) \Rightarrow I(X \mid Z \mid Y \cup W)$.

Using the symmetry and transitivity of $\simeq$ and the antecedent of the axiom we find that

$$\pi_h(. \mid yzw) \simeq \pi_h(. \mid yz) \simeq \pi_h(. \mid zw), \quad \forall yzw.$$

Particularly, $\pi_h(. \mid yz) \simeq \pi(. \mid zw)$, $\forall yzw$. We can demonstrate that $\pi_h(. \mid z) = \max_y \pi_h(. \mid yz)$ using a reasoning completely analogous to the one used to prove Decomposition. Therefore, by applying the maximum property, we obtain that $\pi_h(. \mid z) \simeq \pi_h(. \mid zw)$, $\forall zw$. Finally, as $\pi_h(. \mid zw) \simeq \pi_h(. \mid yzw)\ \forall yzw$, we can deduce $\pi_h(. \mid z) \simeq \pi_h(. \mid yzw)$, $\forall yzw$ by transitivity. □

The next corollary can be obtained directly from Proposition 10:

Corollary 1. *The independence relationship* H4, *where* $\simeq$ *is any equivalence relationship verifying the limit and maximum properties, satisfies axioms* A1 *and* A3–A6.

Given that any possibility distribution represents imprecise and uncertain knowledge, in some cases it might be considered too strict to require that the similarity relationship must preserve some specific value, particularly the limit value (for example, the two distributions π_1 and π_2 defined as $\pi_1(x_1)=\pi_1(x_2)=1$, $\pi_1(x_3)=0.3$, and $\pi_2(x_1)=1$, $\pi_2(x_2)=0.99$, $\pi_2(x_3)=0.3$, would not be similar). An alternative condition, that together with the maximum property, also guarantees that the independence defined using $\simeq$ has good properties, is the following:

Definition 13 (*Sandwich's property*). We say that a similarity relationship $\simeq$ satisfies the Sandwich's property if, for any three possibility distributions π, π_1, π_2 verifying $\pi_1(x)\leqslant\pi(x)\leqslant\pi_2(x)\ \forall x\in D_X$, then

$$\pi_1 \simeq \pi_2 \ \Rightarrow\ \pi_1 \simeq \pi \simeq \pi_2.$$

Using this property, we can prove the following result:

Proposition 11. *Given an equivalence relationship between possibility distributions,* $\simeq$, *a sufficient condition for the fulfilment of Decomposition is that* $\simeq$ *verifies the maximum and Sandwich's properties. Moreover, these properties also guarantee the fulfilment of intersection.*

Proof. First, given y and z, we shall prove the following previous result:

$$\exists \kappa\in D_W \text{ such that } \forall x,\ \pi_h(x\mid yz\kappa)\leqslant\pi_h(x\mid yz)\leqslant \max_w \pi_h(x\mid yzw). \tag{13}$$

In order to prove the left-hand side of the inequality, let us suppose that the relation does not hold, i.e., $\forall w\ \exists x_w$ such that $\pi_h(x_w\mid yzw)>\pi_h(x_w\mid yz)$. Then, we find that $\pi_h(x_w\mid yz)<\pi_h(x_w\mid yzw)\leqslant 1$, and so $\pi_h(x_w\mid yz)=\pi(x_w yz)<\pi(yz)$. Moreover, if it were $\pi_h(x_w\mid yzw)=\pi(x_w yzw)$ then we would find that $\pi_h(x_w\mid yz)=\pi(x_w yz)<\pi_h(x_w\mid yzw)=\pi(x_w yzw)$, which is not possible. The other option is that $\pi_h(x_w\mid yzw)=1$, i.e., $\pi(x_w yzw)=\pi(yzw)$, $\forall w$. But in this case we obtain $\pi(yzw)=\pi(x_w yzw)\leqslant\pi(x_w yz)<\pi(yz)$, $\forall w$, which is a contradiction too. So, $\exists\kappa\in D_W$ such that $\forall x$, $\pi_h(x\mid yz\kappa)\leqslant\pi(yz)$.

On the other hand, with a similar reasoning to the one used in Proposition 10, (Decomposition axiom), the right-hand side inequality, $\pi_h(x\mid yz)\leqslant \max_w \pi_h(x\mid yzw)\ \forall x$, can easily be deduced.

Now, we will consider the different axioms:

Decomposition: $I(X\mid Z\mid Y\cup W)\ \Rightarrow\ I(X\mid Z\mid Y)$.

We have $I(X\mid Z\mid Y\cup W)$, i.e., $\pi_h(.\mid yzw)\simeq\pi_h(.\mid z)$, $\forall yzw$. Given y and z, and using the previous result, we know that $\kappa\in D_W$ exists such that $\forall x$, $\pi_h(x\mid yz\kappa)\leqslant\pi_h(x\mid yz)\leqslant\max_w\pi_h(x\mid yzw)$. As $\simeq$ verifies the maximum property, we deduce $\max_w\pi_h(.\mid yzw)\simeq\pi_h(.\mid z)$, and by using transitivity we obtain $\max_w\pi_h(.\mid yzw)\simeq\pi_h(.\mid yz\kappa)$. Now, using the Sandwich property and transitivity, we find that $\pi_h(.\mid yz)\simeq\max_w\pi_h(.\mid yzw)\simeq\pi_h(.\mid z)$, and therefore $I(X\mid Z\mid Y)$.

Intersection: $I(X\mid Y\cup Z\mid W)\ \&\ I(X\mid Z\cup W\mid Y)\ \Rightarrow\ I(X\mid Z\mid Y\cup W)$.

From the hypothesis we deduce $\pi_h(.\mid yzw)\simeq\pi_h(.\mid yz)\simeq\pi_h(.\mid zw)$, $\forall yzw$. Particularly, $\pi_h(x\mid yz)\simeq\pi_h(x\mid zw)$, $\forall yzw$. Using a reasoning analogous to the one employed previously (Eq. (13)) it can be seen that, given any z, there exists $\kappa\in D_W$ such that $\forall x$, $\pi_h(x\mid z\kappa)\leqslant\pi_h(x\mid z)\leqslant\max_w\pi_h(x\mid zw)$. From $\pi_h(.\mid z\kappa)\simeq\pi_h(.\mid yz), \forall yz$, and using the maximum property, we find that $\max_w\pi_h(.\mid zw)\simeq\pi_h(.\mid z\kappa)\ \forall yz$. Now, by applying the Sandwich's property, we obtain $\pi_h(.|z\kappa) \simeq \pi_h(.|z) \simeq \max_w \pi_h(.|zw)$ $\forall z$. Therefore, we have $\pi_h(.|z) \simeq \pi_h(.|z\kappa) \simeq \pi_h(.|yz)$ $\forall yz$ and, taking into account that $\pi_h(.|yz) \simeq \pi_h(.|yzw)$ $\forall yzw$, we obtain $\pi_h(.|z) \simeq \pi_h(.|yzw), \forall yzw$ by using transitivity. Thus, $I(X|Z|Y \cup W)$ holds. □

The following result is a direct consequence of the previous proposition:

Corollary 2. *The independence relationship* H4, *where* $\simeq$ *is any equivalence relationship verifying the Sandwich's and maximum properties, satisfies axioms* A1 *and* A3–A6.

Now, let us see some examples of different similarity relationships $\simeq$ that can be used to define independence:

Isoordering: The idea is based on considering a possibility distribution as a formalism where uncertainty is represented as a preference among events. Thus, we can reason about events which are uncommon, unique or ill-known (for example because we have not enough statistical data), but it is possible to think that some events are more possible than others, even though we cannot assign precise numerical values to the possibility distribution. Therefore, the numerical values that we assign are not relevant but we are more interested in the relative ordering among events.

So, if we consider that a possibility distribution, essentially, establishes an ordering among the values that the variable that we are considering can take on, and the numbers (the possibility degrees) only have a secondary importance, then we can say that two possibility distributions are similar if they establish the same ordering. More formally, we can define the relation $\simeq$ by means of

$$\pi \simeq \pi' \Leftrightarrow \forall x, x' \in D_X \ [\pi(x) < \pi(x') \Leftrightarrow \pi'(x) < \pi'(x')].$$

Resemblance: Another option is to speak about similarity between distributions when the possibility degrees of each distribution, for each value, are alike. More concretely, we can discretize the $[0,1]$ interval and say that two distributions are similar if their respective discrete versions coincide. Formally, let m be any positive integer, and $\{\alpha_k\}_{k=0,\dots,m}$ be real numbers such that $\alpha_0 < \alpha_1 < \cdots < \alpha_m$, with $\alpha_0 = 0$ and $\alpha_m = 1$. If we denote $I_k = [\alpha_{k-1}, \alpha_k)$, $k = 1, \dots, m-1$, and $I_m = [\alpha_{m-1}, \alpha_m]$ then we define the relationship $\simeq$ by means of

$$\pi \simeq \pi' \Leftrightarrow \forall x \in D_X \quad \exists k \in \{1, \dots, m\} \text{ such that } \pi(x), \pi'(x) \in I_k.$$

An equivalent version of this definition can be obtained in terms of the α-cuts of the possibility distributions:

$$\pi \simeq \pi' \Leftrightarrow C(\pi, \alpha_k) = C(\pi', \alpha_k) \quad \forall k = 1, \dots, m-1,$$

where $C(\pi, \alpha) = \{x \in D_X \mid \pi(x) \geq \alpha\}$.

α_0-Equality: A third option is based on the following idea: consider a threshold α_0, and suppose that only from values greater than α_0 is it considered interesting to differentiate between the possibility degrees of two distributions; the values whose possibility degrees are below the threshold are not considered relevant. In terms of the α-cuts, this relationship $\simeq$ can be expressed as follows:

$$\pi \simeq \pi' \Leftrightarrow C(\pi, \alpha) = C(\pi', \alpha) \quad \forall \alpha \geq \alpha_0,$$

which is equivalent to

$$\pi \simeq \pi' \Leftrightarrow C(\pi, \alpha_0) = C(\pi', \alpha_0) \text{ and } \pi(x) = \pi'(x) \quad \forall x \in C(\pi, \alpha_0).$$

It can easily be proven that Isoordering, Resemblance and α_0-Equality are equivalence relationships and satisfy the Maximum property. Moreover, Isoordering and α_0-Equality verify the limit property, and

Resemblance and α_0-Equality verify the Sandwich's property. Therefore, from Corollaries 1 and 2, using any of these three similarity operators we can guarantee that axioms A1, and A3–A6 hold. However, the axiom of Symmetry does not hold. We prove this fact by means of the following counterexamples:

Isoordering

xy	$\pi(xy)$
x_1y_1	1.0
x_1y_2	0.9
x_2y_1	0.6
x_2y_2	0.8

Resemblance

xyz	$\pi(xyz)$
$x_1y_1z_1$	1.00
$x_1y_1z_2$	0.80
$x_1y_2z_1$	1.00
$x_1y_2z_2$	0.80
$x_2y_1z_1$	0.70
$x_2y_1z_2$	0.50
$x_2y_2z_1$	0.75
$x_2y_2z_2$	0.60

α_0-Equality

xy	$\pi(xy)$
x_1y_1	1.0
x_1y_2	0.6
x_1y_3	0.7
x_2y_1	0.5
x_2y_2	0.5
x_2y_3	0.5
x_3y_1	0.4
x_3y_2	0.4
x_3y_3	0.4

Isoordering: Consider two bivaluated variables X, Y, and consider the corresponding possibility distribution given in the table above. In this case, we have that $\pi(x_1)>\pi(x_2)$ and $\pi_h(x_1\mid y)>\pi_h(x_2\mid y)\ \forall y$, hence $I(X\mid\emptyset\mid Y)$. However, $\pi(y_1)>\pi(y_2)$ but $\pi_h(y_1\mid x_2)=0.6<\pi_h(y_2\mid x_2)=1$, and therefore $\neg I(Y\mid\emptyset\mid X)$.

Resemblance: In this case, consider three bivaluated variables X,Y,Z and the set of intervals $I_1=[0,0.5)$, $I_2=[0.5,0.7)$, $I_3=[0.7,0.8)$, $I_4=[0.8,1]$. We obtain $I(X\mid Z\mid Y)$, because $\pi_h(x_1\mid yz)$, $\pi_h(x_1\mid z)\in I_4\ \forall yz$, $\pi_h(x_2\mid yz_1)$, $\pi_h(x_2\mid z_1)\in I_3\ \forall y$, $\pi_h(x_2\mid yz_2)$, $\pi_h(x_2\mid z_2)\in I_2\ \forall y$. However, $\pi_h(y_1\mid z_1)\in I_4$ and $\pi_h(y_1\mid x_2z_1)\in I_3$, hence $\neg I(Y\mid Z\mid X)$.

α_0-Equality: We will use $\alpha_0=0.4$. Considering the corresponding possibility distribution, we obtain $\pi_h(x\mid y)=\pi(x)\ \forall xy$ and, on the other hand, we have $\pi_h(y_2\mid x_2)=1\neq 0.6=\pi(y_2)$. So, we get $I(X\mid\emptyset\mid Y)$ but $\neg I(Y\mid\emptyset\mid X)$.

We can also define independence using similarity relationships in a different way: instead of generalizing an equality between conditional possibility distributions ($\pi_h(x\mid yz)=\pi_h(x\mid z)$), as we have done above, we can use the concept of conditional non-interactivity, thus generalizing an equality between joint possibility distributions ($\pi(xyz)=\pi(xz)\wedge\pi(yz)$). In that case, the definition of independence becomes:

Definition 14. (H5) *Extended conditional non-interactivity.*

$$I(X\mid Z\mid Y)\ \Leftrightarrow\ \pi(xyz)\simeq\pi(xz)\wedge\pi(yz). \tag{14}$$

We have also studied the properties of $\simeq$ that guarantee the fulfilment of the axioms, for the definition of independence above. In this case $\simeq$ should be an equivalence relationship compatible with the marginalization and combination of possibility distributions (using the minimum as the combination operator), that is:

Proposition 12. *The independence relationship* H5, *where* $\simeq$ *is any similarity relationship between possibility distributions verifying*:
- $\simeq$ *is an equivalence relationship.*
- *If* $\pi(xy)\simeq\pi'(xy)$ *then* $\max_x\pi(xy)\simeq\max_x\pi'(xy)$.
- *If* $\pi_1(xz)\simeq\pi_1'(xz)$ *and* $\pi_2(yz)\simeq\pi_2'(yz)$ *then* $\pi_1(xz)\wedge\pi_2(yz)\simeq\pi_1'(xz)\wedge\pi_2'(yz)$,

satisfies axioms A1–A5.

Proof. The proof is quite simple, and we shall omit it. □

The next table summarizes the properties that each one of the previous definitions of independence satisfies. Therein, the symbol 'X' means that the corresponding property holds.

	A1	A2	A3	A4	A5	A6
H1 (Eq. (3))	X		X	X	X	X
H2 (Eq. (7))	X	X	X	X	X	
H3 (Eq. (11))	X		X	X	X	X
H4 (Eq. (12))	X		X	X	X	X
H5 (Eq. (14))	X	X	X	X	X	

Finally, observe that in order to use Hisdal conditioning, the only necessary operation is comparison. So, we could easily consider possibility distributions taking their values in sets different from the [0, 1] interval: it is suffice to use a set $(\mathscr{L}, \preccurlyeq)$, where

$$\mathscr{L} = \{L_0, L_1, \ldots, L_n\}$$

with $L_0 \preccurlyeq L_1 \preccurlyeq \cdots \preccurlyeq L_n$, that is to say, a totally ordered set (for example a set of linguistic labels) and then define possibility measures by means of

$$\Pi : \mathscr{P}(D_X) \rightarrow \mathscr{L}$$

verifying:
1. $\Pi(D_X) = L_n$,
2. $\Pi(A \cup B) = \vee_{\preccurlyeq} \{\Pi(A), \Pi(B)\}$, $\forall A, B \subseteq D_X$,

where $\vee_{\preccurlyeq}$ is the maximum (supremum) operator associated with the ordering $\preccurlyeq$. In these conditions, we can define the different concepts of independence and conditioning in exactly the same way as before, obtaining the same properties.

5. The marginal problem

When we have to manage uncertain knowledge, even using the expert knowledge, it is quite difficult to assess all the possibility values for a large set of variables (because the number of values grows exponentially with the number of variables). In that case, an alternative approach is to obtain these values for smaller subsets of variables, and use them to construct the joint measure in a reasonable way: this is the *marginal problem*, which will be studied in this section. Particularly, suppose that X, Y, Z are three disjoint subsets of variables, and that π_1 and π_2 are two possibility measures over XZ and YZ respectively. The problem is how to construct a joint possibility measure, π, over XYZ. A natural restriction should be that the marginalization of the joint measure over the initial subsets of variables preserves the original measures.

In order to construct the joint measure, a reasonable assumption is to consider that the variables involved satisfy some kind of independence condition, particularly a conditional independence relationship may be assumed. The most usual hypothesis is to assume that X and Y are independent given Z. So, one requirement for the joint distribution is that this independence relationship must be true, i.e., $I(X \mid Z \mid Y)$ holds for the distribution π. Therefore, the following requirements should be verified:
1. X and Y must be independent given Z, i.e. $I(X \mid Z \mid Y)$ holds for the distribution π.
2. The marginal measure on XZ must be preserved, i.e., $\pi(xz) = \max_y \pi(xyz) = \pi_1(xz)$.
3. The marginal measure on YZ must be preserved, i.e., $\pi(yz) = \max_x \pi(xyz) = \pi_2(yz)$.

In order to satisfy these requirements, we must impose on the two original possibility measures, π_1 and π_2, a basic compatibility condition: 'The marginal possibility measures over Z, $\pi_1(z)$ and $\pi_2(z)$ give the same information about the values that the variable Z can take on'. This compatibility relationship can be defined as follows:

Definition 15. Let π_1 and π_2 be two possibility measures defined on XZ and YZ, respectively. We say that π_1 and π_2 are compatible on Z, if and only if

$$\forall z \in D_Z, \quad \pi_1(z) = \pi_2(z).$$

Therefore, our first step is to check for the compatibility between π_1 and π_2 on the variable Z. Then, in the case of compatibility, we must fix a criterion that allows us to determine the joint distribution $\pi(xyz)$. As we have several definitions of independence, different approaches can be considered. First, we study the marginal problem using the independence criterion H1 ('not modifying the information'), denoted by $I_{\mathrm{H1}}(.\,|\,.\,|\,.)$. We shall see that the fulfilment of all the previous requirements cannot be achieved, and we must relax these requirements. Next, the independence criterion H2, denoted by $I_{\mathrm{H2}}(.\,|\,.\,|\,.)$, will be considered.

The following example shows that, when considering Hisdal conditioning, there exist cases where it is not possible to preserve the marginal measures and, at the same time, satisfy the criterion H1.

Example 2. Let X, Y, Z be bivaluated variables, π_1, π_2 the two possibility distributions shown below, defined on XZ and YZ, respectively, and let π be any possibility distribution obtained from π_1 and π_2.

xz	$\pi_1(xz)$
x_1z_1	0.4
x_1z_2	a_1
x_2z_1	1
x_2z_2	a_2

yz	$\pi_2(yz)$
y_1z_1	0.2
y_1z_2	b_1
y_2z_1	1
y_2z_2	b_2

In order to satisfy the compatibility of π_1 and π_2 on Z, we must impose that $\max\{a_1, a_2\} = \max\{b_1, b_2\}$. Moreover, suppose that π satisfies the independence criterion H1 and preserves the marginals.

In that case, we find that $\pi_1(x_1 \,|\, z_1) = \pi_1(x_1z_1) = 0.4$. As π preserves the marginal over XZ, then $\pi(x_1z_1) = \max_y \pi(x_1yz_1) = 0.4$ and $\pi(z_1) = 1$, which implies that $\pi(x_1 \,|\, z_1) = 0.4$. So, in order to satisfy the criterion H1, it has to be $\pi(x_1 \,|\, yz_1) = 0.4$, $\forall y$. This implies that $\forall y$, $0.4 = \pi(x_1 \,|\, yz_1) = \pi(x_1yz_1) < \pi(yz_1)$. But π also preserves the marginal over YZ, hence $\pi(yz_1) = \pi_2(yz_1)$. Therefore, we obtain $0.4 < \pi_2(yz_1)$ $\forall y$, which is obviously false, because $\pi_2(y_1z_1) = 0.2$. □

Therefore, when we use H1 as the independence criterion, we must relax the requirements about the marginals. We can use the fact that H1 is a non-symmetrical independence relationship, i.e., $I_{\mathrm{H1}}(X \,|\, Z \,|\, Y) \not\Rightarrow I_{\mathrm{H1}}(Y \,|\, Z \,|\, X)$. Then, if we consider that 'X is independent of Y, given Z', then we attempt to preserve only π_1 over XZ and, on the other hand, if we suppose that 'Y is independent of X, given Z', we attempt to preserve only π_2 over YZ.

From now on, suppose that $I_{\mathrm{H1}}(X \,|\, Z \,|\, Y)$ is considered (the case $I_{\mathrm{H1}}(Y \,|\, Z \,|\, X)$ is completely analogous). The most obvious way to preserve π_1 over XZ and the independence is to define π as $\pi(xyz) = \pi_1(xz)$, $\forall x, y, z$. But in this case we do not use the information given by π_2. We propose an alternative approach where this distribution is considered whenever it is possible. The result is that the independence criterion H1 is satisfied, the marginal distribution over XZ is preserved, and the marginal distribution over YZ is included in π_2:

Proposition 13. *Let π_1 and π_2 be two possibility distributions, defined on XZ and YZ respectively, and compatible on Z. Then, the joint possibility distribution π, defined by means of*

$$\pi(xyz) = \begin{cases} \pi_2(yz) & \text{if } \pi_1(\bar{x}z) < \pi_2(yz) < \pi_1(z), \\ \alpha_{xz} \in]\pi_1(\bar{x}z), \pi_1(z)] & \text{if } \pi_2(yz) \leqslant \pi_1(\bar{x}z) < \pi_1(z), \\ \pi_1(xz) & \text{if } \pi_1(\bar{x}z) = \pi_1(z) \text{ or } \pi_2(yz) = \pi_1(z), \end{cases}$$

where $\pi_1(\bar{x}z) = \max_{x' \neq x} \pi_1(x'z)$, satisfies:
1. *$I_{H1}(X \mid Z \mid Y)$, i.e., $\pi_h(x \mid yz) = \pi_h(x \mid z)$, $\forall xyz$.*
2. *$\pi(xz) = \pi_1(xz)$, $\forall xz$.*
3. *$\pi(yz) \geqslant \pi_2(yz)$, $\forall yz$.*

Proof. (2) $\pi(xz) = \pi_1(xz)$, $\forall xz$.

First, let us observe that if $\pi_1(\bar{x}z) < \pi_1(z)$ then $\pi_1(xz) = \pi_1(z)$. Therefore, from the definition of π it follows immediately that $\pi(xyz) \leqslant \pi_1(xz)$ $\forall xyz$. On the other hand, for all z $\exists y_z$ such that $\pi_2(y_z z) = \pi_2(z) = \pi_1(z)$. Therefore we have $\pi(xy_z z) = \pi_1(xz)$ and then $\pi(xz) = \max_y \pi(xyz) = \pi(xy_z z) = \pi_1(xz)$ $\forall xz$.

(3) $\pi(yz) \geqslant \pi_2(yz)$, $\forall yz$.

For all z, $\exists x_z$ such that $\pi(x_z z) = \max_x \pi(xz) = \pi(z)$. Then, from the definition of π it follows that $\pi(x_z yz) \geqslant \pi_2(yz)$ $\forall yz$. Therefore, $\pi(yz) = \max_x \pi(xyz) \geqslant \pi(x_z yz) \geqslant \pi_2(yz)$ $\forall yz$.

(1) $I(X \mid Z \mid Y)$, i.e., $\pi_h(x \mid yz) = \pi_h(x \mid z)$, $\forall xyz$.

Let us distinguish two possible cases, depending on the value that $\pi_h(x \mid z)$ takes:

- Suppose that $\pi(x \mid z) = \pi(xz) < \pi(z)$: we shall show that $\forall y$, $\pi(xyz) = \pi(xz)$ and $\pi(xyz) < \pi(yz)$, i.e., $\pi_h(x \mid yz) = \pi(xyz) = \pi(xz) = \pi_h(x \mid z)$ $\forall y$.
  As $\pi_1(xz) = \pi(xz) < \pi(z) = \pi_1(z)$, then $\pi_1(\bar{x}z) = \pi_1(z)$, and so $\pi(xyz) = \pi_1(xz) = \pi(xz)$ $\forall y$.
  On the other hand, from $\pi_1(xz) < \pi_1(z)$ it follows that $\exists x_z \neq x$ such that $\pi_1(x_z z) = \pi_1(z)$. Let us study the different possible values for $\pi(x_z yz)$:
  (i) If $\pi(x_z yz) = \pi_1(x_z z) = \pi_1(z)$, then $\pi(xyz) = \pi_1(xz) < \pi_1(z) = \pi_1(x_z z) = \pi(x_z yz) \leqslant \pi(yz)$, hence $\pi(xyz) < \pi(yz)$.
  (ii) If $\pi(x_z yz) = \pi_2(yz)$, then $\pi(xyz) = \pi_1(xz) \leqslant \max_{x' \neq x_z} \pi_1(x'z) = \pi_1(\overline{x_z}z) < \pi_2(yz) = \pi(x_z yz) \leqslant \pi(yz)$, hence $\pi(xyz) < \pi(yz)$.
  (iii) If $\pi(x_z yz) = \alpha_{x_z z}$, then $\pi(xyz) = \pi_1(xz) \leqslant \pi_1(\overline{x_z}z) < \alpha_{x_z z} = \pi(x_z yz) \leqslant \pi(yz)$, hence $\pi(xyz) < \pi(yz)$.
  Therefore, in any case we find that $\pi(xyz) < \pi(yz)$.
- Suppose that $\pi_h(x \mid z) = 1$, i.e., $\pi_1(xz) = \pi(xz) = \pi(z) = \pi_1(z)$: in this case, we must prove that $\forall y$, $\pi(xyz) = \pi(yz)$, i.e., $\pi(x \mid yz) = 1 = \pi(x \mid z)$ $\forall y$. Let us study the different possible values for $\pi(xyz)$:
  (i) If $\pi(xyz) = \pi_1(xz)$, then $\pi(xyz) = \pi_1(xz) = \pi_1(z) = \pi(z) \geqslant \pi(yz)$, and therefore $\pi(xyz) = \pi(yz)$.
  (ii) If $\pi(xyz) = \pi_2(yz)$, then it has to be $\pi_1(\bar{x}z) < \pi_2(yz) < \pi_1(z)$. So, we have $\pi_1(x'z) < \pi_2(yz) < \pi_1(z)$ $\forall x' \neq x$, and $\pi_1(\overline{x'}z) = \pi_1(z)$ $\forall x' \neq x$. Therefore, we deduce $\pi(x'yz) = \pi_1(x'z) < \pi_2(yz) = \pi(xyz)$ $\forall x' \neq x$, hence $\pi(xyz) = \max_{x'} \pi(x'yz) = \pi(yz)$.
  (iii) If $\pi(xyz) = \alpha_{xz}$, then it has to be $\pi_2(yz) \leqslant \pi_1(\bar{x}z) < \alpha_{xz} \leqslant \pi_1(z)$. Once again we obtain $\pi_1(x'z) < \alpha_{xz} < \pi_1(z)$ $\forall x' \neq x$ and $\pi_1(\overline{x'}z) = \pi_1(z)$ $\forall x' \neq x$. So, $\pi(x'yz) = \pi_1(x'z) < \alpha_{xz} = \pi(xyz)$ $\forall x' \neq x$, and $\pi(xyz) = \max_{x'} \pi(x'yz) = \pi(yz)$. □

Example 3. Let π_1 and π_2 be the possibility distributions, defined on XZ and YZ respectively, displayed in the tables below, where X, Y and Z are binary variables. Supposing that the conditional independence relationship $I(X \mid Z \mid Y)$ holds, we can construct, according to Proposition 13, the following joint possibility distribution π_{H1}

x z	$\pi_1(xz)$
x_1 z_1	0.25
x_1 z_2	0
x_2 z_1	0.5
x_2 z_2	1

y z	$\pi_2(yz)$
y_1 z_1	0.5
y_1 z_2	0.75
y_2 z_1	0
y_2 z_2	1

x z y	$\pi_{H1}(xyz)$
x_1 z_1 y_1	0.25
x_1 z_1 y_2	0.25
x_1 z_2 y_1	0
x_1 z_2 y_2	0
x_2 z_1 y_1	0.5
x_2 z_1 y_2	]0.25, 0.5]
x_2 z_2 y_1	0.75
x_2 z_2 y_2	1

□

Now, we shall consider the marginal problem using the criterion of independence H2 ('not gaining information'), $I_{H2}(X \mid Z \mid Y)$. This relationship is equivalent to $\pi(xyz) = \pi(xz) \wedge \pi(yz)$ $\forall xyz$. Therefore, a natural approach for the construction of the joint distribution is to use a similar scheme.

Proposition 14. *Let π_1 and π_2 be two possibility distributions defined on XZ and YZ, respectively, and compatible on Z. Then, the joint possibility distribution π, defined by means of*

$$\pi(xyz) = \pi_1(xz) \wedge \pi_2(yz), \quad \forall xyz$$

satisfies:
1. $I_{H2}(X \mid Z \mid Y)$, *i.e.*, $\pi_h(x \mid yz) \geqslant \pi_h(x \mid z)$, $\forall xyz$.
2. $\pi(xz) = \pi_1(xz)$, $\forall xz$.
3. $\pi(yz) = \pi_2(yz)$, $\forall yz$.

Proof. (2) $\pi(xz) = \max_y \pi(xyz) = \max_y\{\pi_1(xz) \wedge \pi_2(yz)\}$. Then $\pi(xz) = \pi_1(xz) \wedge \max_y \pi_2(yz) = \pi_1(xz) \wedge \pi_2(z) = \pi_1(xz) \wedge \pi_1(z) = \pi_1(xz)$.

(3) The proof is similar to the previous one.

(1) The proof is immediate, considering that the joint distribution preserves the marginal and the way in which the joint distribution was constructed. □

Example 4. Considering π_1 and π_2 as in the previous example, we can construct the following joint possibility distribution, π_{H2}, using H2 as the independence criterion:

x	z	y	$\pi_{H2}(xyz)$
x_1	z_1	y_1	0.25
x_1	z_1	y_2	0
x_1	z_2	y_1	0
x_1	z_2	y_2	0
x_2	z_1	y_1	0.5
x_2	z_1	y_2	0
x_2	z_2	y_1	0.75
x_2	z_2	y_2	1

Moreover, it can easily be proven that the joint measure obtained from Proposition 14 (π_{H2}) is always more informative than the joint measure obtained from Proposition 13 (π_{H1}), that is to say, $\pi_{H2} \leqslant \pi_{H1}$.

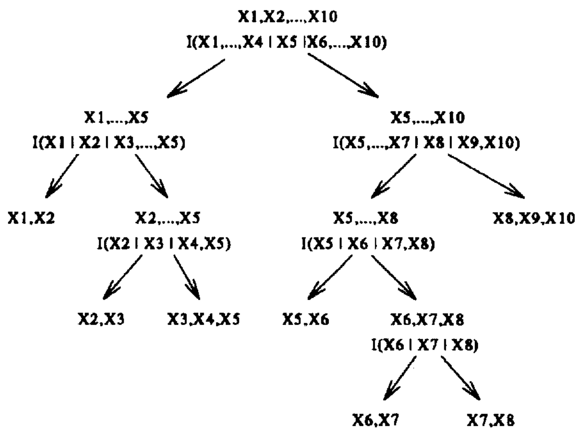

Fig. 1. Possibilistic Tree.

5.1. Application to the storage of large possibility distributions

An important question that we have to manage is the storage of the values of a joint distribution. A direct approach is to use a table of values. In this case, the storage requirements grow exponentially, and quickly become prohibitive when we have a great number of variables. In this section we present an approach which uses conditional independence in order to obtain a better memory request. Using independence relationships among variables we can factorize the joint possibility distribution in terms of its conditionally independent components.

Suppose that X, Y, Z are disjoint subsets of variables, and π is a joint possibility distribution over XYZ. Let us assume that $I_{H2}(X \mid Z \mid Y)$ holds, i.e., $\pi(xyz) = \pi(xz) \wedge \pi(yz)$, $\forall xyz$. Then, if we split the joint distribution into two components $\pi(xz)$ and $\pi(yz)$, it is possible to recover the original measure using the previous equation. This idea can easily be generalized to the n-dimensional case: suppose that we have n variables, $X_1, X_2, \ldots, X_n$, and let X_i be the variable (the same can be done for a set of variables) such that $I(X_1, \ldots, X_{i-1} \mid X_i \mid X_{i+1}, \ldots, X_n)$. In that case, we can split the initial distribution into two components, defined on $\{X_1, \ldots, X_{i-1}, X_i\}$ and $\{X_i, X_{i+1}, \ldots, X_n\}$. The same idea can be recursively applied to both subsets of variables, thus forming a tree structure where we only need to store the marginal possibility distributions on the leaf nodes. In order to recover the joint distribution, we can combine the marginal distributions in a bottom-up approach, using the minimum as the combination operator.

Example 5. Let $X_1, \ldots, X_{10}$ be bivaluated variables, and suppose that the independence relationships indicated in the nodes of the tree displayed in Fig. 1 hold. Then, we only need to store the following marginal distributions:

$$\pi(x_1,x_2),\ \pi(x_2,x_3),\ \pi(x_3,x_4,x_5),\ \pi(x_5,x_6),\ \pi(x_6,x_7),\ \pi(x_7,x_8),\ \pi(x_8,x_9,x_{10}).$$

Thus, we have to store 36 values instead of the $2^{10} = 1024$ values needed to store the complete joint distribution. Moreover, for example using the left subtree, the distribution $\pi(x_1,x_2,x_3,x_4,x_5)$ can be obtained, using the conditional independence relationship H2, by means of

$$\pi(x_1,x_2,x_3,x_4,x_5) = \pi(x_1,x_2) \wedge \pi(x_2,x_3,x_4,x_5),$$

where

$$\pi(x_2,x_3,x_4,x_5) = \pi(x_2,x_3) \wedge \pi(x_3,x_4,x_5).$$

```
Search ( S, node_n )
{
  X^n ← set of variables in node_n
  if (node_n ∈ Leaves ) then
    return max_{X^n\S} π(X^n);
  else
    S_L ← S ∩ L^n;
    S_R ← S ∩ R^n;
    if S_L = ∅
    then << search only in the right branch >>
      π_R ← Search(S_R ∪ (S ∩ K^n), rightchild_{node_n})
      return π_R
    else if S_R = ∅
        then << search only in the left branch >>
          π_L ←Search(S_L ∪ (S ∩ K^n), leftchild_{node_n})
          return π_L
        else << search in both branches >>
          π_L ← Search ( S_L ∪ K^n, leftchild_{node_n})
          π_R ← Search ( S_R ∪ K^n, rightchild_{node_n})
          return max_{K^n\S} π_L ∧ π_R
}
```

Fig. 2. Algorithm for finding the marginal distribution on the set of variables S.

Continuing in this way, the joint possibility distribution can be obtained using a sequence of joint steps. □

Proposition 15. *Let T be a possibilistic tree, $X_1,X_2,\dots,X_n$ variables (or sets of variables), and L_j, $j=1,\dots,m$ the leaves in T. Then, using the criterion of independence* H2, *the joint possibility distribution on $X_1,X_2,\dots,X_n$ may be obtained by means of*

$$\pi(x_1,x_2,\dots,x_n)=\pi_{L_1}\wedge\pi_{L_2}\wedge\cdots\wedge\pi_{L_m}=\bigwedge_{j=1}^{m}\pi_{L_j},$$

where π_{L_j} represents the marginal possibility distribution stored in the leaf L_j.

As we shall see, the independence properties in the tree can be used to obtain marginal possibility distributions without having to consider all the variables in the tree. In the previous example, if we are interested in obtaining the joint distribution for the subset of variables $\{X_2,X_5\}$, $\pi_{X_2X_5}$, we can avoid the construction of the complete joint distribution represented by the tree, i.e., we can marginalize the distribution $\pi(x_2,x_3,x_4,x_5)$ previously calculated on X_2X_5 and therefore, only some variables on the left subtree have been used.

Furthermore, these possibilistic structures can be viewed as a mechanism to perform possibilistic inference. We could obtain the complete joint distribution on all the variables, and after that update the knowledge by conditioning to the set of observed variables. This method is computationally unfeasible in practice. Then, a natural approach is to use the independence relationships represented in the tree in order to consult only relevant information. Therefore, more efficient inference algorithms can be obtained. This feature makes possibilistic reasoning similar to Bayesian Networks [24] techniques.

Let T be the set of variables that we are interested in, and let *Obs* be the set of observed variables. When *Obs* is not empty, an inference mechanism consists of calculating the marginal possibility measure on $T\cup Obs$, and then calculate the conditional measure $\pi_{T|Obs}$. The next algorithm, *Search* (Fig. 2), shows how to obtain a marginal possibility distribution over a generic subset of variables S. In that case, we use

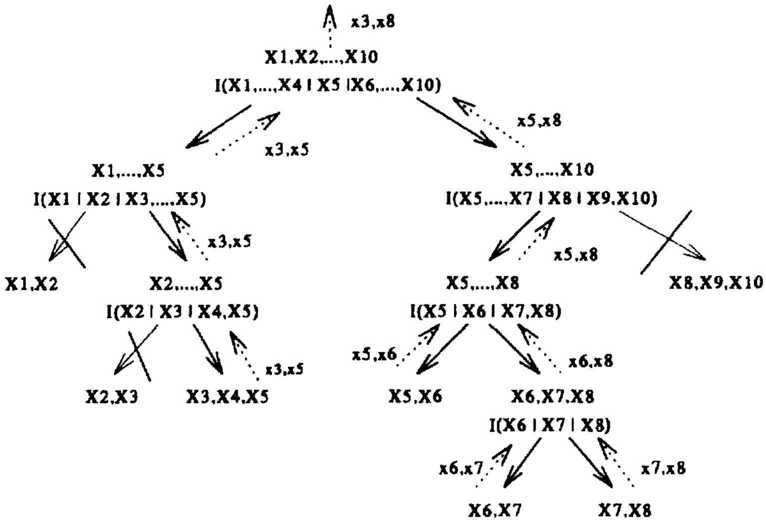

Fig. 3. Messages sent by each node to its parent to calculate $\pi_{X_3X_8}$.

the tree in order to split the subset S into two subsets, S_L, S_R, satisfying the independence relationship indicated in the nodes, i.e., $I(S_L \mid K \mid S_R)$. Then, we recursively search for the possibility values for $S_L \cup K$ and $S_R \cup K$ and after we construct the marginal measure over S by means of the combination operator, i.e., $\pi_S = \max_{K \setminus S} \pi_{S_L \cup K} \wedge \pi_{S_R \cup K}$. During this process, we are using the property of Decomposition. Initially, the algorithm must be called with *Search*$(S, root)$, and will return the marginal possibility distribution over S. We should note that in each level, the algorithm returns the possibility distribution over the variables that we are interested in by means of the marginalization operation.

In the algorithm we use the following notation:

S: The set of variables that we are interested in.
X^n: The set of variables stored in node n.
K^n: The set of variables that splits the node up into two conditionally independent subsets of variables, given K^n, i.e., $X^n = L^n \cup R^n \cup K^n$, and $I(L^n \mid K^n \mid R^n)$.

The following example illustrates how the algorithm works.

Example 6. Consider the previous possibilistic tree (Fig. 1). If we are interested in obtaining the marginal distribution over the set of variables $\{X_3, X_8\}$, then Fig. 3 indicates the possibility distribution that each node sends to its parent in the tree.

6. Concluding remarks

The concept of independence, as with other formalisms of uncertainty management, is also important in possibility theory. In this paper several alternative definitions of possibilistic independence have been proposed. Most of them are based on the concept of conditioning, and in this paper specially we use Hisdal conditioning [20] (in the second part of this paper [6], a similar study considering Dempster conditioning is carried out). The main differences between the different definitions come from the way in which we compare the a priori and a posteriori knowledge. We use the concepts of not modifying the information, not gaining additional information, and obtaining similar information. We have also studied the properties of our definitions of independence with respect to a well-known set of axioms [24], which try to capture the intuitive notion of independence. The given definitions always verify Trivial Independence, Decomposition, Weak Union and Contraction, whereas Symmetry and Intersection are more difficult to preserve, and only some definitions verify them.

Moreover, we have also studied the marginal problem, i.e., how to construct a joint piece of information from marginal measures, assuming a conditional independence criterion. As a result of this study we can conclude that it is possible to factorize a possibility distribution, and then to recover the original joint distribution by means of combination operations. Moreover, using independence relationships, we can construct any marginal measure without having to consider the complete set of variables. This process may be viewed as a particular case of possibilistic inference.

References

[1] S. Benferhat, D. Dubois, H. Prade, Expressing independence in a possibilistic framework and its application to default reasoning, in: A. Cohn (Ed.), 11th European Conf. on Artificial Intelligence, Wiley, New York, 1994, pp. 150–154.

[2] L.M. de Campos, Independence relationships in possibility theory and their application to learning belief networks, in: G. Della Riccia, R. Kruse, R. Viertl (Eds.), Mathematical and Statistical Methods in Artificial Intelligence, CISM Courses and Lectures, vol. 363, Springer, Wien, 1995, pp. 119–130.

[3] L.M. de Campos, J. Gebhardt, R. Kruse, Axiomatic treatment of possibilistic independence, in: C. Froidevaux, J. Kohlas (Eds.), Symbolic and Quantitative Approaches to Reasoning and Uncertainty, Lecture Notes in Artificial Intelligence, vol. 946, Springer, Berlin, 1995, pp. 77–88.

[4] L.M. de Campos, J.F. Huete, Independence concepts in upper and lower probabilities, in: B. Bouchon-Meunier, L. Valverde, R.R. Yager (Eds.), Uncertainty in Intelligent Systems, North-Holland, Amsterdam, 1993, pp. 49–59.

[5] L.M. de Campos, J.F. Huete, Learning non probabilistic belief networks, in: M. Clarke, R. Kruse, S. Moral (Eds.), Symbolic and Quantitative Approaches to Reasoning and Uncertainty, Lecture Notes in Computer Science, vol. 747, Springer, Berlin, 1993, pp. 57–64.

[6] L.M. de Campos, J.F. Huete, Independence concepts in possibility theory: Part II, Fuzzy Sets and Systems 103 (1999), to appear.

[7] L.M. de Campos, J.F. Huete, S. Moral, Possibilistic independence, in: Proc. 3rd European Congress on Intelligent Techniques and Soft Computing (EUFIT), vol. I, 1995, pp. 69–74.

[8] L.M. de Campos, M.T. Lamata, S. Moral, The concept of conditional fuzzy measure, Internat. J. Intelligent Systems 5 (1990) 237–246.

[9] L.M. de Campos, S. Moral, Independence concepts for convex sets of probabilities, in: P. Besnard, S. Hanks (Eds.), Proc. 11th Conf. on Uncertainty in Artificial Intelligence, Morgan and Kaufmann, San Mateo, 1995, pp. 108–115.

[10] G. de Cooman, E.E. Kerre, A new approach to possibilistic independence, in: Proc. 3rd IEEE Internat. Conf. on Fuzzy Systems, 1994, pp. 1446–1451.

[11] A.D. Dawid, Conditional independence in statistical theory, J. Roy. Statist. Soc. Ser. B 41 (1979) 1–31.

[12] A.P. Dempster, Upper and lower probabilities induced by a multivalued mapping, Ann. Math. Statist. 38 (1967) 325–339.

[13] D. Dubois, Belief structures, possibility theory, decomposable confidence measures on finite sets, Comput. Artif. Intell. 5 (1986) 403–417.

[14] D. Dubois, L. Farinas del Cerro, A. Herzig, H. Prade, An ordinal view of independence with applications to plausible reasoning, in: R. López de Mántaras, D. Poole (Eds.), Proc. 10th Conf. on Uncertainty in Artificial Intelligence, Morgan and Kaufmann, 1994, pp. 195–203.

[15] D. Dubois, H. Prade, Possibility Theory: An Approach to Computerized Processing of Uncertainty, Plenum Press, New York, 1988.

[16] D. Dubois, H. Prade, Inference in Possibilistic Hypergraphs, in: B. Bouchon-Meunier, R.R. Yager, L.A. Zadeh (Eds.), Uncertainty in Knowledge Bases, Lecture Notes in Computer Science, vol. 521, Springer, Berlin, 1990, pp. 250–259.

[17] R. Fagin, Multivalued dependencies and a new form for relational databases, ACM Trans. Database Systems 2 (1977) 262–278.

[18] L. Farinas del Cerro, A. Herzig, Possibility theory and independence, in: Proc. 5th Internat. Conf. on Information Processing and Management of Uncertainty in Knowledge Based Systems, IPMU, 1994, pp. 820–825.

[19] P. Fonck, Conditional independence in possibility theory, in: R. López de Mántaras, D. Poole (Eds.), Proc. 10th Conf. on Uncertainty in Artificial Intelligence, Morgan and Kaufmann, San Mateo, 1994, pp. 221–226.

[20] E. Hisdal, Conditional possibilities, independence and noninteraction, Fuzzy Sets and Systems 1 (1978) 283–297.

[21] J.F. Huete, Aprendizaje de Redes de Creencia mediante la Detección de Independencias: Modelos no Probabilisticos, Thesis, Universidad de Granada, Spain, 1995 (in Spanish).
[22] R. Kruse, J. Gebhardt, F. Klawonn, Foundations of Fuzzy Systems, Wiley, New York, 1994.
[23] S.L. Lauritzen, A.P. Dawid, B.N. Larsen, H.G. Leimer, Independence properties of directed Markov fields, Network 20 (1990) 491–505.
[24] J. Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan and Kaufmann, San Mateo, 1988.
[25] J. Pearl, D. Geiger, T. Verma, Conditional independence and its representation, Kybernetika 25 (1989) 33–34.
[26] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princeton, NJ, 1976.
[27] P.P. Shenoy, Conditional independence in uncertainty theories, in: D. Dubois, M.P. Wellman, B. D'Ambrosio, P. Smets (Eds.), Proc. 8th Conf. on Uncertainty in Artificial Intelligence, Morgan and Kauffmann, San Mateo, 1992, pp. 284–291.
[28] W. Spohn, Stochastic independence, causal independence and shieldability, J. Philos. Logic 9 (1980) 73–99.
[29] M: Studený, Attempts at axiomatic description of conditional independence, Kybernetika 25 (1989) 72–79.
[30] M. Studený, Formal properties of conditional independence in different calculi of A.I., in: M. Clarke, R. Kruse, S. Moral (Eds.), Symbolic and Quantitative Approaches to Reasoning and Uncertainty, Lecture Notes in Computer Science, vol. 747, Springer, Berlin, 1993, pp. 341–348.
[31] M. Sugeno, Theory of fuzzy integrals and its applications, Thesis, Tokyo Institute of Technology, Japan, 1974.
[32] L.A. Zadeh, Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets and Systems 1 (1978) 3–28.