# Irrelevance and Independence Axioms in Quasi-Bayesian Theory

Fabio Gagliardi Cozman

Escola Politécnica, Universidade de São Paulo,
Av. Prof. Mello Moraes, 2231, Cidade Universitária, São Paulo, SP - Brazil,
fgcozman@usp.br

**Abstract.** This paper investigates Walley's concepts of irrelevance and independence, as applied to the theory of closed convex sets of probability measures. Walley's concepts are analyzed from the perspective of axioms for conditional independence (the so-called semi-graphoid axioms). Two new results are demonstrated in discrete models: first, Walley's concept of irrelevance is an asymmetric semi-graphoid; second, Walley's concept of independence is an incomplete semi-graphoid. These results are the basis for an understanding of irrelevance and independence in connection to the theory of closed convex sets of probability measures, a theory that has received attention as a powerful representation for uncertainty in beliefs and preferences.

## 1 Introduction

The purpose of this paper is to investigate technical properties of Walley's concepts of irrelevance and independence. These concepts, connected to the theory of closed convex sets of probability measures, have several intuitive properties, but have not received a complete treatment in the literature.

Closed convex sets of probability have been used to represent imprecision in probability models or to represent divergences in group decision-making [10, 12]; currently, the main practical application of these models is the field of robust Bayesian Statistics [2, 8, 14]. Under various names, such as imprecise probability theory [12] or Quasi-Bayesian theory [7], closed convex sets of probability measures have received extensive technical analysis. In particular, several concepts of independence in Quasi-Bayesian models have emerged [3, 6]. Recently, intuitive definitions of irrelevance and independence were proposed by Walley [12, Chapter 9], definitions that can be explained in a direct manner from basic axioms of Quasi-Bayesian theory.

The purpose of this paper is to investigate Walley's concepts of irrelevance and independence from the perspective of the so-called semi-graphoid axioms. These axioms are meant to capture the subtle idea behind "independence." In fact, many important concepts of independence, like probabilistic independence and graphical independence, are semi-graphoids.

Two new results are demonstrated in this paper in connection to discrete Quasi-Bayesian models: first, Walley's irrelevance defines an asymmetric semi-graphoid; second, Walley's independence defines an incomplete semi-graphoid.

These results are the basis for a thorough understanding of irrelevance and independence concepts in the theory of closed convex sets of probability measures.

## 2 Semi-graphoid Axioms

Take random variables $X$, $Y$, $Z$ and $W$ on a probability space $(\Omega, \mathcal{F})$. Random variables considered in this paper are discrete and functions of random variables are bounded (values of all functions are finite). In other words, a random variable $X$ induces a countable partition on an underlying outcome space, and any function $f(X)$ considered in this paper is a finite constant on each element of the partition induced by $X$.

As summarized by Dawid [4], a ternary relation $X \perp\!\!\!\perp Y \,|\, Z$ is a *semi-graphoid* when it captures the notion that "once the value of $Z$ has been specified, any information about $Y$ is irrelevant to uncertainty about $X$." The semi-graphoid axioms are postulates that capture the essential properties of this abstract notion. Note that Dawid ignores possible differences between concepts of irrelevance and independence.

The axioms for $\perp\!\!\!\perp$ are:

**A1** If $X \perp\!\!\!\perp Y \,|\, Z$ then $Y \perp\!\!\!\perp X \,|\, Z$.
**A2** $X \perp\!\!\!\perp Y \,|\, X$.
**A3** If $X \perp\!\!\!\perp Y \,|\, Z$ and $W = g(Y)$ (a bounded function of $Y$), then $X \perp\!\!\!\perp W \,|\, Z$.
**A4** If $X \perp\!\!\!\perp Y \,|\, Z$ and $W = g(Y)$ (a bounded function of $Y$), then $X \perp\!\!\!\perp Y \,|\, (W, Z)$.
**A5** If $X \perp\!\!\!\perp Y \,|\, Z$ and $X \perp\!\!\!\perp W \,|\, (Y, Z)$, then $X \perp\!\!\!\perp (Y, W) \,|\, Z$.

Call a structure that satisfies only A2, A3, A4 and A5 an *asymmetric semi-graphoid*. Call a structure that satisfies only A1, A2, A3 and A4 an *incomplete semi-graphoid*.

## 3 Quasi-Bayesian Theory

A closed convex set of probability measures is called a *credal set*; existence of credal sets is derived from axioms about preferences [7]. Convexity of a set $K$ means that, given any two measures $P_1$ and $P_2$ in $K$, the convex combination $\alpha P_1 + (1-\alpha)P_2$ (for $\alpha \in [0,1]$) is also in $K$. To simplify terminology, sets of probability densities (defined whenever possible) are also called credal sets. A credal set defined by a set of densities $p(X)$ is indicated by $K(X)$.

Given a random variable $X$, lower and upper expectations for a function $f(X)$ are respectively defined as:

$$\underline{E}[f(X)] = \min_{p(X) \in K(X)} E_{p(X)}[f(X)]\,, \quad \overline{E}[f(X)] = \max_{p(X) \in K(X)} E_{p(X)}[f(X)]\,,$$

where $E_{p(X)}[f(X)]$ is the standard expectation of the function $f(X)$. There is a one-to-one correspondence between lower (or upper) expectations and credal sets: a collection of lower (or upper) expectations for all arbitrary bounded functions $f(X)$ defines a credal set $K(X)$. Lower expectations can be obtained from upper expectations through the expression $\underline{E}[f(X)] = -\overline{E}[-f(X)]$. Bounds on the probability of any event $A$ can be obtained by taking the lower expectation of the indicator function $I_A(X)$, which is one if $X \in A$ and zero otherwise. For a credal set $K(X)$, the function $\underline{P}(A) = \min_{p(X) \in K(X)} P(A) = \min_{p(X) \in K(X)} E_{p(X)}[I_A(X)]$ is called a *lower envelope* and the function $\overline{P}(A) = \max_{p(X) \in K(X)} P(A) = \max_{p(X) \in K(X)} E_{p(X)}[I_A(X)]$ is called a *upper envelope*.

Closed convex sets of conditional probability measures are employed to represent conditional beliefs. The *conditional credal set* $K(X|A)$ is defined by conditional densities $p(X|A)$.

Quasi-Bayesian inference is performed by applying Bayes rule to each measure in a jont credal set. A *posterior credal set* is the union of all posterior measures obtained in this process. Consequently, for a bounded function $f(X)$ and an event $A$ with positive lower envelope:

$$\underline{E}[f(X)|A] = \min_{p(X) \in K(X)} \frac{E_{p(X)}[f(X) I_A(X)]}{E_{p(X)}[I_A(X)]}. \tag{1}$$

If the event $A$ has lower envelope equal to zero, the posterior credal set $K(X|A)$ is taken *by convention* to contain all probability densities $p(X|A)$. For an event $A$ with positive lower envelope, $\underline{E}[f(X)|A]$ is the unique solution of the following equation in $\lambda$, called *generalized Bayes rule* by Walley [12, Section 6.4]:

$$\underline{E}[(f(X) - \lambda) I_A(X)] = 0,$$

where $I_A(X)$ is the indicator function of the event $A$ (to simplify notation, $I(X)$ denotes the indicator function for a particular event defined by a value of $X$). In Walley's theory, the generalized Bayes rule is derived from basic principles of coherent behavior, but the rule can be derived directly from Expression (1). The generalized Bayes rule has a number of straightforward properties; the next lemma proves one such property that is important for the main results in this paper.

**Lemma 1.** *If $\underline{P}(W)$ and $\underline{P}(X)$ are larger than zero, the value of $\lambda$ that satisfies the equation $\underline{E}[(g(Y) - \lambda) I(W)|X] = 0$ is equal to $\underline{E}[g(Y)|W, X]$ for an arbitrary bounded function $g(Y)$.*

There are two possible proofs for this result, one directly derived from the concept of convex sets of measures and standard Bayes rule, and the other derived from the generalized Bayes rule.

*Proof.* Using Expression (1):

$$\underline{E}[g(Y)|W, X] = \min_{p(Y,W|X) \in K(Y,W|X)} \frac{E_{p(Y,W|X)}[g(Y) I(W)|X]}{E_{p(Y,W|X)}[I(W)|X]}.$$

Consequently:

$$\min_{p(Y,W|X) \in K(Y,W|X)} \frac{E_{p(Y,W|X)}[g(Y) I(W) - \underline{E}[g(Y)|W, X]\, I(W)|X]}{E_{p(Y,W|X)}[I(W)|X]} = 0.$$

This equation has a solution when the numerator is zero, so an equivalent equation is obtained:

$$\min_{p(Y|W,X)\in K(Y|W,X)} E_{p(Y,W|X)}[(g(Y)-\lambda)I(W)|X] = \underline{E}[(g(Y)-\lambda)I(W)|X] = 0,$$

where $\lambda$ is the unique solution and is consequently equal to $\underline{E}[g(Y)|W,X]$.

*Proof.* Define $\gamma(\lambda) \equiv \underline{E}[(g(Y)-\lambda)I(W)|X]$. By the generalized Bayes rule, the value of $\gamma(\lambda)$ is the solution of the equation

$$\underline{E}[((g(Y)-\lambda)I(W)-\gamma(\lambda))\, I(X)] = 0.$$

Because $\underline{E}[(g(Y)-\lambda)I(W)|X]$ is strictly decreasing with $\lambda$ and has positive and negative values, there must be $\lambda^*_X$ such that $\gamma(\lambda^*_X) = 0$. For this value of $\lambda$,

$$\underline{E}[((g(Y)-\lambda^*_X)I(W)-0)\, I(X)] = 0.$$

Consequently, for $\lambda = \lambda^*_X$:

$$\underline{E}[(g(Y)-\lambda^*_X)I(W)I(X)] = 0.$$

So, $\lambda^*_X$ is the unique solution of the equation $\underline{E}[(g(Y)-\lambda^*_X)I(W,X)] = 0$, and $\lambda^*_X = \underline{E}[g(Y)|W,X]$.

The concept of independence was not defined in the original formulation of Quasi-Bayesian theory by Giron and Rios [7]; indeed, there is considerable controversy regarding judgements of independence in Quasi-Bayesian models [1,3,6,9,13]. The results presented in this paper adopt Walley's definition of independence because this formulation can be reduced to preference relations, following the Quasi-Bayesian philosophy [12].

Walley's concepts can be defined as follows. Consider $X$, $Y$ and $Z$, and the joint credal set $K(X,Y,Z)$. Two conditional credal sets $K(X|Y,Z)$ and $K(X|Y)$ are said *equal* if any density in $K(X|Y',Z')$ (for any fixed $Y=Y'$, $Z=Z'$) also belongs to $K(X|Z')$, and vice-versa.

**Definition 1.** *Variable $Y$ is* irrelevant *to $X$ given $Z$ if $K(X|Z)$ is equal to $K(X|Y,Z)$, regardless of the value of $Z$.*

**Definition 2.** *Variables $X$ and $Y$ are* independent *given $Z$ if $X$ is irrelevant to $Y$ given $Z$ and $Y$ is irrelevant to $X$ given $Z$.*

Note that $Z$ can be empty; in this case the irrelevance and independence concepts are not "conditional" on any variable.

Walley's original definition of irrelevance is stated in terms of lower expectations. The definitions are equivalent and stem from the equivalence between lower expectations and convex sets of measures [12, Chapter 3]; the lower expectation approach is convenient in several proofs.

**Definition 3.** *Variable $Y$ is* irrelevant *to $X$ given $Z$ if $\underline{E}[f(X)|Y,Z]$ is equal to $\underline{E}[f(X)|Z]$ for any bounded function $f(X)$, regardless of the value of $Z$.*

## 4 Irrelevance and the Semi-graphoid Axioms

The symbol $\perp\!\!\!\perp$ is described by Dawid as meaning that "any information about $Y$ is irrelevant to uncertainy about $X$" [4]. This description is actually close to Walley's concept of irrelevance, as it does not depend on any symmetry axiom like A1. In fact, the next theorem shows that Walley's concept of irrelevance satisfies all semi-graphoid axioms except A1.

**Theorem 1.** *Walley's concept of irrelevance is an asymmetric semi-graphoid.*

*Proof.* Use $f(X)$ to denote an arbitrary bounded function. The axioms A2-A5 (understood in terms of Walley's irrelevance) can be rephrased using lower expectations:

**A2'** $\underline{E}[f(X)|Y,X] = \underline{E}[f(X)|X]$.
**A3'** If $\underline{E}[f(X)|Y,Z] = \underline{E}[f(X)|Z]$ and $W = g(Y)$ (a bounded function of $Y$), then $\underline{E}[f(X)|W,Z] = \underline{E}[f(X)|Z]$.
**A4'** If $\underline{E}[f(X)|Y,Z] = \underline{E}[f(X)|Z]$ and $W = g(Y)$ (a bounded function of $Y$), then $\underline{E}[f(X)|Y,W,Z] = \underline{E}[f(X)|W,Z]$.
**A5'** If $\underline{E}[f(X)|Y,Z] = \underline{E}[f(X)|Z]$ and $\underline{E}[f(X)|W,Y,Z] = \underline{E}[f(X)|Y,Z]$, then $\underline{E}[f(X)|W,Y,Z] = \underline{E}[f(X)|Z]$.

These statements can now be proved:

**A2'** Note that

$$\underline{E}[f(X)|X,Y] = \min_{p(X|X,Y) \in K(X|X,Y)} E_{p(X|X,Y)}[f(X)|X,Y] = f(X),$$

and also

$$\underline{E}[f(X)|X] = \min_{p(X|X) \in K(X|X)} E_{p(X|X)}[f(X)|X] = f(X).$$

**A3'** The lower expectation $\underline{E}[f(X)|Z]$ is equal to

$$\min_{p(X,Y|Z) \in K(X,Y|Z)} E_{p(Y|Z)}\left[E_{p(X|Y,Z)}[f(X)|Y,Z]\,|Z\right].$$

Consider a probability density $p^*(X,Y,Z)$ such that $E_{p^*(X,Y,Z)}[f(X)|Y,Z] = \underline{E}[f(X)|Z]$ regardless of $Y$. Such a $p^*(X|Y,Z)$ must exist, because the minimum value of $E_{p(X|Y,Z)}[f(X)|Y,Z]$ is $\underline{E}[f(X)|Z]$ for every value of $Y$ (irrelevance of $Y$ to $X$ given $Z$) and the minimum of $E_{p(X|Z)}[f(X)|Z]$ is also $\underline{E}[f(X)|Z]$. If no $p^*(X,Y,Z)$ existed, then the minimum of the expectation $E_{p(Y|Z)}\left[E_{p(X|Y,Z)}[f(X)|Y,Z]\,|Z\right]$ would be larger than the minimum of $E_{p(X|Y,Z)}[f(X)|Y,Z]$ (equal to $\underline{E}[f(X)|Y,Z]$). To simplify notation, expectations calculated with respect to $p^*(X,Y,Z)$ have a $*$ superscript.
Now compute $E^*[f(X)|W,Z]$:

$$E^*[f(X)|W,Z] = \frac{E^*\left[I_{\{W=g(Y)\}}(Y)E^*[f(X)|Y,Z]\,|Z\right]}{E^*\left[I_{\{W=g(Y)\}}(Y)|Z\right]}$$

$$= \frac{\underline{E}[f(X)|Z]\,E^*\left[I_{\{W=g(Y)\}}(Y)|Z\right]}{E^*\left[I_{\{W=g(Y)\}}(Y)|Z\right]} = \underline{E}[f(X)|Z]\,.$$

Note that the expectation produced by $p^*(X, Y, Z)$ is actually a lower bound because

$$\begin{aligned} &\underline{E}[f(X)|W, Z] = \\ &\qquad \min E_{p(Y|W,Z)}\left[I_{\{W=g(Y)\}}(Y) E_{p(X|Y,W,Z)}[f(X)|Y, W, Z]\,|W, Z\right] \\ &\qquad \geq E_{p(Y|W,Z)}\left[I_{\{W=g(Y)\}}(Y) \underline{E}[f(X)|Y, Z]\,|W, Z\right] \\ &\qquad = E_{p(Y|W,Z)}\left[I_{\{W=g(Y)\}}(Y) \underline{E}[f(X)|Z]\,|W, Z\right] = \underline{E}[f(X)|Z]\,. \end{aligned}$$

Consequently, $\underline{E}[f(X)|W, Z] = \underline{E}[f(X)|Z]$.

**A4'** By the generalized Bayes rule, $\underline{E}[f(X)|Y, W, Z]$ is the value $\mu^*$ that satisfies

$$\underline{E}[(f(X) - \mu)\, I(YWZ)] = 0.$$

Because $I(YWZ) = I(YZ)$, the value $\mu^*$ is also the solution of

$$\underline{E}[(f(X) - \mu)\, I(YZ)] = 0.$$

Consequently, $\mu^* = \underline{E}[f(X)|Y, Z]$. Because $\underline{E}[f(X)|Y, Z] = \underline{E}[f(X)|Z]$ by hypothesis, $\underline{E}[f(X)|Y, W, Z] = \underline{E}[f(X)|Z] = \underline{E}[f(X)|W, Z]$ (by Axiom A3').

**A5'** By transitivity: $\underline{E}[f(X)|W, Y, Z] = \underline{E}[f(X)|Y, Z] = \underline{E}[f(X)|Y]$.

## 5 Independence and the Semi-graphoid Axioms

Walley's concept of independence can conceivably satisfy all semi-graphoid axioms, as it certainly displays the symmetry property required by A1. However, axiom A5 is not satisfied by Walley's concept of independence. Denote Walley's relations of irrelevance and independence respectively by $\perp\!\!\!\perp_R$ and $\perp\!\!\!\perp_N$. The following property is not true:

**A5"** If $X \perp\!\!\!\perp_N Y \,|\, Z$ and $X \perp\!\!\!\perp_N W \,|\, (Y, Z)$, then $(Y, W) \perp\!\!\!\perp_R X \,|\, Z$.

An example demonstrates the failure of A5". Consider three binary variables $X$, $Y$ and $W$ (in this example the variable $Z$ is omitted). Suppose that $K(X, Y, W)$ is the convex hull of three joint densities $p_1(X, Y, W)$, $p_2(X, Y, W)$ and $p_3(X, Y, W)$ displayed in Table 1. This credal set satisfies $X \perp\!\!\!\perp_N Y$ and $X \perp\!\!\!\perp_N W \,|\, Y$ because:

$$\begin{aligned} &K(X_0|Y_0) = K(X_0|Y_1) = K(X_0) \quad (p(X_0) \in [0.2, 0.3]) \\ &K(Y_0|X_0) = K(Y_0|X_1) = K(Y_0) \quad (p(Y_0) \in [0.3, 0.4]) \\ &K(X_0|W_0, Y_0) = K(X_0|W_1, Y_0) = K(X_0|W_0, Y_1) = K(X_0|W_1, Y_1) = K(X_0) \\ &K(W_0|X_0, Y_0) = K(W_0|X_1, Y_0) = K(W_0|Y_0) \quad (p(W_0|Y_0) \in [0.1, 0.2]) \\ &K(W_0|X_0, Y_1) = K(W_0|X_1, Y_1) = K(W_0|Y_1) \quad (p(W_0|Y_1) \in [0.4, 0.8]). \end{aligned}$$

But for $h(Y, W) = I_{\{Y=Y_0, W=W_0\}}(Y, W)$, $\underline{E}[h(Y, W)|X_0] = \underline{P}(Y_0, W_0|X_0) = 0.0372$ and $\underline{E}[h(Y, W)] = \underline{P}(Y_0, W_0) = 0.04$. Consequently, it is not true that $(Y, W) \perp\!\!\!\perp_R X$.

Despite the failure of $A5''$, all other semi-graphoid axioms are satisfied by Walley's independence.

| $W$ | $X$ | $Y$ | $p_1(X,Y,W)$ | $p_2(X,Y,W)$ | $p_3(X,Y,W)$ |
|---|---|---|---|---|---|
| $W_0$ | $X_0$ | $Y_0$ | 0.008 | 0.018 | 0.0093 |
| $W_1$ | $X_0$ | $Y_0$ | 0.072 | 0.072 | 0.0757 |
| $W_0$ | $X_1$ | $Y_0$ | 0.032 | 0.042 | 0.037 |
| $W_1$ | $X_1$ | $Y_0$ | 0.288 | 0.168 | 0.228 |
| $W_0$ | $X_0$ | $Y_1$ | 0.096 | 0.084 | 0.09 |
| $W_1$ | $X_0$ | $Y_1$ | 0.024 | 0.126 | 0.075 |
| $W_0$ | $X_1$ | $Y_1$ | 0.384 | 0.196 | 0.290 |
| $W_1$ | $X_1$ | $Y_1$ | 0.096 | 0.294 | 0.195 |

**Table 1.** Joint densities that define a credal set.

**Theorem 2.** *Walley's concept of irrelevance is an incomplete semi-graphoid.*

*Proof.* Axiom A1 is immediate as Walley's independence is just a symmetrization of Walley's concept of irrelevance. Take $f(Y)$ and $g(W)$ to denote arbitrary bounded functions. To demonstrate axioms A2-A4 (understood in terms of Walley's independence), only the following facts must be proved in addition to statements already proved in Theorem 1:

**A2"** $\underline{E}[f(Y)|X,X] = \underline{E}[f(Y)|X]$.
**A3"** If $\underline{E}[f(Y)|X,Z] = \underline{E}[f(Y)|Z]$ and $W = g(Y)$, then $\underline{E}[f(W)|X,Z] = \underline{E}[f(W)|Z]$.
**A4"** If $\underline{E}[f(Y)|X,Z] = \underline{E}[f(Y)|Z]$ and $W = g(Y)$, then $\underline{E}[f(Y)|X,W,Z] = \underline{E}[f(Y)|W,Z]$.

Axiom A2" is immediate. The other ones can be demonstrated as follows:

**A3"** Because $\underline{E}[f(W)|X,Z]$ is equal to $\underline{E}[f(g(Y))|X,Z]$ and $\underline{E}[f(g(Y))|X,Z] = \underline{E}[f(g(Y))|Z]$ by hypothesis, $\underline{E}[f(W)|X,Z] = \underline{E}[f(W)|Z]$.
**A4"** By Lemma 1, $\underline{E}[f(Y)|X,W,Z]$ is the value of $\lambda^*$ that solves the equation

$$\underline{E}[(f(Y)-\lambda)\, I(W)|X,Z] = 0.$$

By hypothesis, $\underline{E}[(f(Y)-\lambda)\, I(g(Y))|X,Z] = \underline{E}[(f(Y)-\lambda)\, I(g(Y))|Z]$, so $\lambda^*$ is also the solution of the equation

$$\underline{E}[(f(Y)-\lambda)\, I(W)|Z] = 0.$$

Consequently, $\lambda^* = \underline{E}[f(Y)|X,W,Z] = \underline{E}[f(Y)|W,Z]$.

## 6 Conclusion

The properties of Walley's concept of irrelevance and independence depart in various ways from the properties of standard probabilistic independence. The most obvious difference is that irrelevance and independence are not identical in Walley's framework. This distinction is not commonly employed in statistical parlance; for example, both Dawid [4] and Pearl [11] justify the semi-graphoid axioms using the words irrelevance and independence almost interchangeably. Dawid even introduces the properties of the semi-graphoids indicating that "we can rephrase these as assertions of *irrelevance*" [4].

But concepts of irrelevance and independence do not have identical meaning in our quotidian discourse. The concept of irrelevance is associated with an asymmetric notion; if an object has no bearing upon some situation, we are content to say that the object is irrelevant to the situation, without further analysis of the reciprocal relation. Quite differently, the concept of independence seems to be intrinsically tied to a notion of symmetry, as two objects or situations are independent only when they are irrelevant to each other.

Such considerations of symmetry are crucial to appreciate the results in this paper. The failure of A1 for Walley's irrelevance and the failure of A5" for Walley's independence are quite satisfactory under this perspective.

Axiom A1 is simply a symmetry requirement, so it must fail for irrelevance. And it must work for independence, as it does in Walley's framework.

The failure of Axiom A5" may seem strange at first, because A5 is taken as a very reasonable requirement of independence. But all that is reasonable about A5 is contained in the irrelevance property A5'. In fact, Dawid [4] and Pearl [11] defend only A5' when they justify Axiom A5. To quote Dawid:

> P5 [Axiom A5 in the current paper] says that if, knowing $Z$, $Y$ is irrelevant to $X$; and also, knowing both $Y$ and $Z$, $W$ is irrelevant to $X$; then, knowing $Z$, $(Y, W)$ must be jointly irrelevant to $X$. However we understand "irrelevance", this seems to be a desirable property.

Property A5' is certainly desirable when applied to irrelevance, but it is hard to see why its converse A5" must be valid for independence. Suppose that, knowing $Z$, $X$ is irrelevant to $Y$; and also, knowing both $Y$ and $Z$, $X$ is irrelevant to $W$; then, is it desirable that, knowing $Z$, $X$ must be irrelevant to $(Y, W)$? Why would it be impossible for $X$ be irrelevant to each of $Y$ and $W$ taken separately, but still $X$ be relevant to $Y$ and $W$ taken together? Property A5" only seems reasonable when preconceived notions of symmetry are applied to A5'.

These considerations indicate that Walley's framework offers an appropriate path for the study of irrelevance and independence. The present paper contributes with an initial step in this path, and suggests that asymmetric and incomplete semi-graphoids are quite important structures that deserve special attention.

## Acknowledgements

Thanks to Milan Studený by asking whether Walley's concepts of irrelevance and independence satisfy the semi-graphoid axioms, and for clarifying several issues on this problem. The author was partially supported by CNPq through grant 300183/98-4.

## References


1. J. Berger and E. Moreno. Bayesian robustness in bidimensional model : Prior independence. *Journal of Statistical Planning and Inference*, 40:161–176, 1994.
2. J. O. Berger. Robust Bayesian analysis: Sensitivity to the prior. *Journal of Statistical Planning and Inference*, 25:303–328, 1990.
3. L. Chrisman. Independence with lower and upper probabilities. *XII Uncertainty in Artificial Intelligence Conference*, pages 169–177, 1996.
4. A. P. Dawid. Conditional Independence. To appear, *Encyclopedia of Statistical Sciences, Update Volume 3*, Wiley, New York, 1999.
5. A. P. Dawid and M. Studený. Conditional Products: An Alternative Approach to Conditional Independence. *Uncertainty99: The Seventh International Workshop on Artificial Intelligence and Statistics*, Fort Lauderdale, Florida, 1999.
6. L. de Campos and S. Moral. Independence concepts for convex sets of probabilities. *XI Uncertainty in Artificial Intelligence*, pages 108–115, 1995.
7. F. J. Giron and S. Rios. Quasi-Bayesian behaviour: A more realistic approach to decision making? In J. M. Bernardo, J. H. DeGroot, D. V. Lindley, and A. F. M. Smith, editors, *Bayesian Statistics*, pages 17–38. University Press, Valencia, Spain, 1980.
8. J. B. Kadane. *Robustness of Bayesian Analyses*, volume 4 of *Studies in Bayesian econometrics*. Elsevier Science Pub. Co., New York, 1984.
9. H. E. Kyburg and M. Pittarelli. Set-based Bayesianism. *IEEE Transactions on Systems, Man and Cybernetics A*, 26(3):324–339, May 1996.
10. I. Levi. *The Enterprise of Knowledge*. The MIT Press, Cambridge, Massachusetts, 1980.
11. J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kauffman, San Mateo, CA, 1988.
12. P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, New York, 1991.
13. P. Walley and T. L. Fine. Towards a frequentist theory of upper and lower probability. *The Annals of Statistics*, 10(3):741–761, 1982.
14. L. Wasserman. Recent methodological advances in robust Bayesian inference. In J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, editors, *Bayesian Statistics 4*, pages 483–502. Oxford University Press, 1992.