From Acceptance Relations to Causality Ascription in a Belief Function Framework

Imen Boukhris[1,2], **Zied Elouedi**[1] and **Salem Benferhat**[2]

[1]LARODEC, Université de Tunis, ISG Tunis, Tunisie
[2]CRIL, Université d'Artois, Faculté Jean Perrin, France
imen.boukhris@hotmail.com, zied.elouedi@gmx.fr, benferhat@cril.univ-artois.fr

Abstract

From a generic set of uncertain pieces of information about the normal course of things and a temporal sequence of reported facts, an intelligent artifact should be able to identify causally related events and distinguish between factors that facilitate or justify the occurrence of events from other facts. In this paper, we propose a model for causality and facilitation ascriptions when the background knowledge on n-ary variables is represented under the belief function framework. In order to ascribe causality, an agent has to judge if an event is accepted or rejected. We propose different definitions of acceptance and rejection to specify the strength of the causal link in the context of n-ary variables. Even accepted, an event can be confirmed or attenuated which leads to different forms of causality ascriptions. This paper shows that our model can treat these cases in the framework of belief functions.

Introduction

Causality (Shoham 1990; Shafer 1976; Pearl 2000; Halpern and Pearl 2001; Halpern 2008) plays an important role in our understanding of the world. In fact, causal perception is important in several applications as for the diagnosis of the potential causes from observed effects, to induce causal laws from observations as well as for ascribing causes of some abnormal events (Geffner 1990; Benferhat et al. 2008). Though at first glance, it seems obvious what causation is, there are difficulties to define causally linked events.

On the basis of available background knowledge about the normal behavior of things and in front of a chain of reported events, an intelligent artifact should be able to ascribe causality. Note that this task of identification of causal links between events, is different from the standard diagnosis (Peng and Reggia 1990) and simulation problems (Giunchiglia et al. 2004). Thus, it seems clear that the identification of causal links depends on the choice of the framework of representation for the agent's knowledge.

A common sense understanding of the world tells us that we have to deal with imperfect knowledge where an event may be present in the absence of its cause or absent in its presence which means that causality may exist without determinism.

In this paper, we are interested in ascribing causality when the agent's background knowledge is expressed in the belief function formalism (Shafer 1976; Smets 1998) which is a mathematical framework extending probability and possibility theory. It quantifies beliefs (Shafer 1976) in a more flexible way than probability theory does. In particular, it is appropriate to represent ignorance situations (Yager 1983; Wakker 2000).

Despite the expressive power of the belief function theory to model different forms of uncertainty to deal with the poor knowledge of the world that agents usually possess including full knowledge, partial ignorance, total ignorance and even probabilistic knowledge, no model exists for ascribing causality with the belief function formalism.

In this paper, we propose a model to identify causal links between events when the background knowledge is uncertain and expressed in terms of belief functions. Our model overcomes the limitation of qualitative models (Bonnefon et al. 2006; 2008) for ascribing causality in which the representation of events is restrained to binary variables.

The causality-ascribing agent needs to use the concepts of acceptance (Dubois, Fargier, and Prade 2004), rejection and ignorance to affirm that an event involving n-ary variables truly causes another one. In fact, a belief is tentatively accepted (resp. rejected) and may be revised by the arrival of new information. In this work, we propose several definitions of these notions specifying the strength of the causal link, suitable for binary and non-binary variables. We define facilitation ascription and justification and distinguish them from causation under the belief function framework. The concepts of attenuation and confirmation are also introduced.

The rest of the paper is organized as follows: In Section 2, we recall the basic definitions of the causality-ascribing agent's background knowledge language namely the belief function theory. We then define the concepts of ignorance, acceptance and rejection for the case of binary and non-binary variables in Section 3. Section 4 presents our model for causality ascription under the belief function framework and sets its properties. We also define its related notions namely, facilitation, justification, confirmation and attenuation. Section 5 concludes the paper.

Belief Function Theory

In the following, we recall some of the basics of belief function theory. More details can be found in (Smets 1998).

Notations

We will use the following notations:

- Let $U = \{A, B, \ldots\}$ denotes a set of variables.
- $\Theta_A = \{a_1, \ldots, a_n\}$ is a finite set denoting the domain relative to the variable A.
- $X, Y, Z, ..$ denote disjoint subsets of variables from U.
- $\Theta_X = \times_{A \in X} \Theta_A = \{x_1, x_2, \ldots, x_n\}$ represents the cartesian product of the domains of the variables in X.
- $\Theta = \times_{A \in U} \Theta_A$ denotes the frame of discernment, which is the cartesian product of all variable domains in U interpreted as a set of possibilities, exactly one of them corresponds to the truth. It is not possible to have several true propositions simultaneously.

Basic belief assignment

In the belief function theory (Shafer 1976), beliefs are expressed on propositions belonging to the powerset of Θ. The basic belief assignment (bba), denoted by m^Θ, is a mapping from 2^Θ to $[0, 1]$ such that any proposition is associated with a real number belonging to $[0, 1]$ where the sum over all subsets is 1. When there is no ambiguity, m^Θ will be shortened m.

$$\sum_{A \subseteq \Theta} m^\Theta(A) = 1. \tag{1}$$

$m(A)$ is a basic belief mass (*bbm*) assigned to $A \subseteq \Theta$. It represents the part of belief *exactly* committed to the event A. The subsets of Θ such that $m(A) > 0$ are called focal elements. The union of all focal elements is called its core. Shafer (Shafer 1976) initially did not consider the empty set as a focal element. The mass function respecting this constraint is called *normalized*:

$$m(\emptyset) = 0. \tag{2}$$

Complete ignorance corresponds to $m(\Theta) = 1$ (vacuous belief function), and a certain knowledge corresponds to allocating the whole bbm to one singleton of Θ.

Plausibility function

The bba m can be equivalently represented by a plausibility function pl: $2^\Theta \to [0, 1]$, defined as:

$$pl(A) = \sum_{B \cap A \neq \emptyset} m(B) \ \text{ and } pl(\emptyset) = 0. \tag{3}$$

$pl(A)$ quantifies the maximum amount of belief that could be given to a subset A of Θ.

Dempster's rule of combination

The belief function theory provides tools to take into account the dynamics of belief states. In fact, the combined effect of two distinct sources is computed with the Dempster's rule of combination (Smets 1990). It is defined as the orthogonal sum of two bba's m_1 and m_2, whose focal elements are all the possible intersections between pairs of focal elements of m_1 and m_2 respectively.

$$m_1 \oplus m_2(A) = K^{-1} \sum_{B \cap C = A} m_1(B).m_2(C), \forall B, C \subseteq \Theta \tag{4}$$

where the normalization factor: $K = 1 - \sum_{B \cap C = \emptyset} m_1(B).m_2(C)$.

Disjunctive rule of combination

The disjunctive rule of combination defined as follows is used when we only know that at least one piece of evidence holds:

$$m_1 \textcircled{\cup} m_2(A) = \sum_{B \cup C = A} m_1(B).m_2(C), \forall B, C \subseteq \Theta \tag{5}$$

Conditioning

Conditioning consists in revising the agent belief originally defined on A, following the arrival of a new more precise information $B \subseteq A$. Indeed, the mass that was specifically allocated to A is transferred to $A \cap B$ using Dempster's rule of conditioning. $pl(.|B)$ denotes the conditional plausibility function obtained after revising the corresponding m with a new evidence B (where $pl(B) > 0$) and defined as (Smets 1991):

$$pl(A|B) = \frac{pl(A \cap B)}{pl(B)}. \tag{6}$$

Accepted Events with Non-Binary Variables under a Belief Function Framework

In order to infer an unknown causal relation, the agent has to reason from three components:

- his non-causal background knowledge about the *normal* course of the world expressed here in the belief function formalism which is a general and appropriate framework to model the world imperfection.
- a set of observations occurring in his environment, $O = \{f_1, \ldots, f_n\}$. We define an observation as a subset of the frame of discernment $f_i \subseteq \Theta$.
- an event that contradicts its judgment about the normal course of things: an abnormal event $e_i \subseteq \Theta$. This event is involved in the causal process. In fact, the agent will ascribe the causes of this abnormal event.

Hence, causal perception consists in determining among observed events f_i's those that are causes of the abnormal event e_i. The causality-ascribing agent's goal is to identify unknown causal links among several exclusive events. For that purpose, we define abnormal events as *a partition* representing exhaustive and mutually exclusive events of the cartesian product of the domain of some n-ary variables.

The set of possible events is denoted by $E = \{e_1, e_2, \ldots, e_n\}$ verifying these properties:

$$\begin{cases} Exhaustibility : e_1 \cup e_2 \cup \ldots \cup e_n = \Theta, \\ Exclusivity : \forall i, j \quad e_i \cap e_j = \emptyset. \end{cases}$$

The complement of e_i w.r.t. Θ_E, denoted by $\overline{e_i}$ is defined as:

$$\overline{e_i} = \bigcup_{e_j \in E, e_j \neq e_i} e_j$$

An abnormal event can be an atomic event represented any instance of some n-ary variable $A_i = a_{ij}$. In this case $E = \{[a_{i1}], \ldots, [a_{in}]\}$ where $[a_{ij}]$ is a set of all elements $\theta \in \Theta$ such that A_i in θ has the value a_{ij}.

The agent will identify from the set of observed events f_i, the ones that causes the occurrence of e_i. When there is no ambiguity, all events including abnormal ones will be denoted by e_i.

In an uncertain environment, the agent's judgments are subject to changes. In fact, his set of beliefs are only tentatively accepted (resp. rejected) and may be revised in front of the arrival of a new piece of information. This has a direct impact on causal reasoning and precisely on ascribing causality. Hence, as argued in (Bonnefon et al. 2008) the definition of acceptance is essential to identify causal links between related events. The concepts of acceptance can be defined for binary variables (propositional variables) and also n-nary variables. The event e_i has different possible status namely being, accepted, rejected or ignored. The following subsections provide detailed definitions of these concepts.

a) Acceptance

An event is considered as accepted if it is likely enough to be considered as it holds. In belief function theory, we propose to distinguish between three forms of acceptance as explained in the following. Hence, an event e_i can be either very strongly, strongly or weakly accepted (see Figure 1).

Very Strong Acceptance An event is very strongly accepted if the confidence in this event is strictly greater than the confidence in its complement according to Θ_E: $pl(e_i) > pl(\overline{e_i})$.

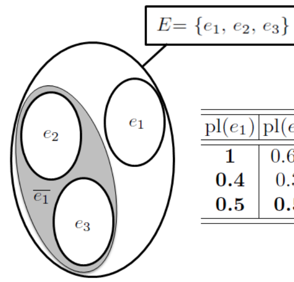
Strong Acceptance When an event is *exclusively* in $Argmax_{e_j \in \Theta_E} pl(e_j)$, it is said to be strongly accepted. Formally, an event e_i is strongly accepted if:

- $pl(e_i) \in Argmax_{e_j \in \Theta_E}(pl(e_j))$
- $\forall k \neq i, \nexists e_k$ where $\text{pl}(e_k) \in Argmax_{e_j \in \Theta_E}(\text{pl}(e_j))$.

Weak Acceptance An event is considered as weakly accepted, if it has the highest plausibility but there are also *other events* in $Argmax_{e_j \in \Theta_E}$ $\text{pl}(e_j)$, formally:

- $pl(e_i) \in Argmax_{e_j \in \Theta_E}(pl(e_j))$
- $\forall k \neq i, \exists e_k \neq e_i$ where $\text{pl}(e_k) \in Argmax_{e_j \in \Theta_E}(\text{pl}(e_j))$.

There exist other events weakly accepted aside from e_i. We consider that the level of acceptance of e_i depends on the number of these events by dividing the value of *Argmax* by the cardinality of events sharing this value.



$\text{pl}(e_1)$	$\text{pl}(e_2)$	$\text{pl}(e_3)$	$\text{pl}(\overline{e_1})$	States of e_1
1	0.64	0.24	**0.73**	very strong accept
0.4	0.3	0.3	0.6	strong accept
0.5	**0.5**	0.1	0.6	weak accept

Figure 1: Forms of acceptance

b) Ignorance

Let $\Theta_E = \{e_i\}$ be a set of events. When an agent has the same confidence in an event e_i and in its complement, the event is perceived as ignored. Formally, $pl(e_i) = pl(\overline{e_i})$.

c) Rejection

By symmetry, different levels of rejection can be defined:

Very Strong Rejection An event is very strongly rejected if the plausibility of this event is strictly less than the plausibility of its complement. Formally, $pl(e_i) < pl(\overline{e_i})$.

Strong Rejection An event is strongly rejected if:

- $pl(e_i) \in Argmin_{e_j \in \Theta_E}(pl(e_j))$,
- $\forall k \neq i, \nexists e_k$ where $pl(e_k) \in Argmax_{e_j \in \Theta_E}(pl(e_j))$.

Weak Rejection We suppose that an event is considered as weakly rejected if:

- $pl(e_i) \in Argmin_{e_j \in \Theta_E}(pl(e_j))$,
- $\forall k \neq i, \exists e_k \neq e_i$ where $pl(e_k) \in Argmin_{e_j \in \Theta_E}(pl(e_j))$.

Here also we consider that the level of acceptance of e_i depends on the number of these events by dividing the value of *Argmin* by the cardinality of events sharing this value.

The following definitions provide a characterization of different concepts of acceptance and rejection introduced above:

Definition 1 *An event $e_i \subseteq \Theta$ is perceived as:*

- *very strongly accepted if $pl(e_i) > pl(\overline{e_i})$*

- *strongly accepted if:*
  - $pl(e_i) \in Argmax_{e_j \in \Theta_E}(pl(e_j))$,
  - $\forall k \neq i$, $\nexists e_k$ *where* $pl(e_k) \in Argmax_{e_j \in \Theta_E}(pl(e_j))$.
- *weakly accepted if:*
  - $pl(e_i) \in Argmax_{e_j \in \Theta_E}(pl(e_j))$
  - $\forall k \neq i$, $\exists e_k$, *where* $pl(e_k) \in Argmax_{e_j \in \Theta_E}(pl(e_j))$.
- *very strongly rejected if* $pl(e_i) < pl(\overline{e_i})$
- *strongly rejected if:*
  - $pl(e_i) \in Argmin_{e_j \in \Theta_E}(pl(e_j))$,
  - $\forall k \neq i$, $\nexists e_k$ *where* $pl(e_k) \in Argmax_{e_j \in \Theta_E}(pl(e_j))$.
- *weakly rejected if:*
  - $pl(e_i) \in Argmin_{e_j \in \Theta_E}(pl(e_j))$,
  - $\forall k \neq i$, $\exists e_k$ *where* $pl(e_k) \in Argmin_{e_j \in \Theta_E}(pl(e_j))$.

If we restrict our definitions to the special case of binary partitions, i.e. $\Theta_E = \{e_1, \overline{e_i}\}$, we show that all definitions of acceptance (resp. rejection) are equivalent.

Proposition 1 *If* $\Theta_E = \{e_1, \overline{e_i}\}$, *then:*

- e_i *is weakly accepted iff* e_i *is strongly accepted iff* e_i *is very strongly accepted iff* $pl(e_i) > pl(\overline{e_i})$.
- e_i *is weakly rejected iff* e_i *is strongly rejected iff* e_i *is very strongly rejected iff* $pl(e_i) < pl(\overline{e_i})$.

d) Conditional Acceptance and Conditional Rejection

The concepts of acceptance, rejection described in previous sections can be extended in order to take into account a given context. Suppose that the context is defined on Θ_F (e.g. f_j), then if an event h_j defined on Θ_H and consistent with f_j takes place, then the context will be defined on $\Theta_F \times \Theta_H$ namely, (f_j, h_j). In this subsection, we introduce different forms of conditional acceptance and rejection.

Belief Function Conditional Acceptance

Very Strong Conditional Acceptance After the occurrence of the event f_j, e_i becomes very strongly accepted if: $pl(e_i|f_j) > pl(\overline{e_i}|f_j)$ and thus $pl(e_i, f_j) > pl(\overline{e_i}, f_j)$.

Strong Conditional Acceptance After the occurrence of the event f_j, e_i becomes strongly accepted if:

- $pl(e_i|f_j) \in \text{Argmax}(pl(e_k|f_j))$;
- $\forall e_j \neq e_i$, $pl(e_j|f_j) \notin \text{Argmax}(pl(e_k|f_j))$;

Weak Conditional Acceptance After the occurrence of the event f_j, e_i becomes very strongly accepted if:

- $pl(e_i|f_j) \in \text{Argmax}(pl(e_k|f_j))$;
- $\exists e_j \neq e_i$, $pl(e_j|f_j) \in \text{Argmax}(pl(e_k|f_j))$

Belief Function Conditional Rejection

Very Strong Conditional Rejection After the occurrence of the event f_j, e_i becomes very strongly rejected if: $pl(e_i|f_j) < pl(\overline{e_i}|f_j)$ which is equivalent to $pl(e_i, f_j) < pl(\overline{e_i}, f_j)$.

Strong Conditional Rejection After the occurrence of the event f_j, e_i becomes strongly rejected if:

- $pl(e_i|f_j) \in \text{Argmin}(pl(e_k|f_j))$;
- $\forall e_j \neq e_i$, $pl(e_i|f_j) \notin \text{Argmin}(pl(e_k|f_j))$

Weak Conditional Rejection After the occurrence of the event f_j, e_i becomes weakly rejected if:

- $pl(e_i|f_j) \in \text{Argmin}(pl(e_k|f_j))$;
- $\exists\ e_j \neq e_i$, $pl(e_j|f_j) \in \text{Argmin}(pl(e_k|f_j))$;

Note that the events f_i, defined here as a context, are linked to the observed events, also denoted by f_i and described in Section 2. In fact, the context of an event is the set of the circumstances and conditions surrounding it. Therefore, it is a set of observed events occurring in the normal course of things.

Ascribing Causality from Belief Function Background Knowledge

In this section, we propose a model for identifying causal links between reported events when the background knowledge is expressed under the belief function framework. We also explain how to ascribe facilitation and distinguish causation from justification. Finally, we show that confirmation and attenuation make sense under the belief function framework. Notice that here definitions are restricted to very strong acceptance (resp. very strong rejection) and consequently to very strong causes, facilitations and justifications.

Causality Ascription

As we mentioned in the introduction there is no consensus about the definition of causality. In fact, it can be seen as a regular association (Hume 2006), a counterfactual dependence (Lewis 2004) or a probability raising (Eells 1991), etc. Causes are necessary related to the occurrence of their effects. A causal link defines a higher belief of effects when a cause takes place. Thus, if a cause does not arise then the plausibility of the occurrence of the effect will decrease. Assume that we have a sequence of events, then:

Definition 2 ***Belief function causality ascription:*** *If an event* e_i *is very strongly rejected and after observing an event* e_j *it becomes very strongly accepted,* e_j *is said to be a* ***very strong cause*** *of* e_i*, namely*

$$pl(\overline{e_i}, e_j) < pl(e_j, e_i) \leq\ pl(e_i) < 1. \quad (7)$$

Example 1 *Assume that an agent has in his disposal some information expressed with the belief function formalism: "generally, airlines do not delay their flights, except when there is a high quantity of ash clouds". The quantity of ash clouds after the eruption of a volcano is represented with* Θ_A*={high (h)* $> 4mg/m^3$*, medium (m)* $> 2mg/m^3$ *and* $< 4mg/m^3$*, low (l)* $< 2mg/m^3$*}. Flights delay is represented with* Θ_F*={yes (y), no (n)}. Initial knowledge is uncertain and quantified as follows:*
- $m^F(\{y\}) = 0.2$*,* $m^F(\{n\}) = 0.7$*,* $m^F(\Theta_F) = 0.1$*. Thus,* $pl^F(\{y\}) = 0.3 < pl^F(\{n\}) = 0.8$*.*
- $m^A(\{h\}) = 0.1$*,* $m^A(\{m\}) = 0.6$*,* $m^A(\{l\}) = 0.2$*,* $m^A(\Theta_A) = 0.1$*, accordingly* $pl^A(\{h\}) = 0.2$*,* $pl^A(\{m\}) = 0.7$*,* $pl^A(\{l\}) = 0.3$*.*
- In context of hight quantity of ash clouds, companies are more likely to delay their flights. $m^A(\{y\}|\{h\}) = 0.8$*,* $m^A(\{n\}|\{h\}) = 0.1$*,* $m^A(\Theta_F|\{h\}) = 0.1$*. Thus,* $pl^A(\{y\}|\{h\}) = 0.9$*,* $pl^A(\{n\}|\{h\}) = 0.2$*.*

From this background knowledge and the occurrence of the event: "The eruption of the Eyjafjöll volcano creates high quantity of ash clouds. Afterwards, all European countries delayed their flights", the agent should be able to identify if it exists a causal link between ash clouds and flights delay.
We notice that pl($\{y\}$*) is an* ***abnormal event*** *that is* ***very strongly rejected*** *since* $pl(\{y\}) = 0.3 < pl(\{n\}) = 0.8$*.*
In context of high ash clouds, it becomes ***very strongly accepted*** *(i.e.* $pl(\{(y,h)\}) > pl(\{(n,h)\})$*). Hence, the high quantity of ash clouds can be seen as a* ***very strong cause*** *of flights delay.*

Causality Ascription Properties

This section presents some desirable properties regarding ascribing causality. Indeed, these properties help the intelligent artifact to identify causal relations. In fact, an agent will identify abnormal events occurring in the normal course of things as causes. An agent should be aware when ascribing causality because generally an event cannot at the same time be causing an event and be caused by this same event. Besides, the transitive nature of causation is satisfied only if events are reported sequentially in the time. When an agent knows that two causal links share a common cause, he can affirm that this common event is a cause of either both of them or one of them. This relation is satisfied within the stability property. This paper does not consider the situation where the presence of some correlated relations is explained by some hidden variables.

Abnormal Causes The first property concerns abnormal causes. It says that in most cases the cause that provokes the belief change of the event from an abnormal event to a normal one is an abnormal event (Hilton and Slugoski 1986). This property is true in our model.

Proposition 2 *If* $e_i \in \Theta_{E_i}$ *is very strongly rejected (i.e.* $pl(e_i) < pl(\overline{e_i})$*) and after the occurrence of* $e_k \in \Theta_{E_k}$*, it becomes very strongly accepted (i.e.* $pl(e_i|e_k) > pl(\overline{e_i}|e_k)$*) then it is likely that* e_k *is also a very strongly rejected event (i.e.* $pl(e_k) < pl(\overline{e_k})$*).*

Example 2 *Let us continue with the background knowledge presented in Example 1. A high quantity of ash clouds in considered as a cause of flights delay and it is very strongly rejected since* $pl(\{h\}) = m(\{h\}) + m(\Theta_A) = 0.2 < pl(\{l,m\}) = m(\{l\}) + m(\{m\}) + m(\Theta_A) = 0.9$*.*

Asymmetric Causality In a given context, an event cannot be at the same time a very strong cause and an effect of a second event. Namely given three events $e_i \in \Theta_{E_i}$, $e_j \in \Theta_{E_j}$ and $e_k \in \Theta_{E_k}$, if e_i is a very strong cause of e_j in context e_k then e_j cannot be simultaneously an effect and a very strong cause of e_i in context e_k.

e_i very strongly causes e_j in context e_k, if and only if the conditions in Table 1 are satisfied:

Table 1: e_i very strongly causes e_j in context e_k

events	time t	time $t+1$
context: e_k	✓	✓
occurrence of e_i	−	✓
status of e_j	strongly rejected $pl(\overline{e_j},e_k) > pl(e_j, e_k)$	strongly accepted $pl(e_i,e_j,e_k) > pl(e_i,\overline{e_j},e_k)$

Assuming that causality is symmetric, e_j should be considered a very strong cause of e_i in context e_k.

This relation is possible in only one case, namely the two causal relations should not be simultaneous. Saying that e_i very strongly causes e_j in context e_k, supposes that the conditions in Table 2 are satisfied:

Table 2: e_i very strongly causes e_j in context e_k

events	time t	time $t+1$
context: e_k	✓	✓
occurrence of e_j	−	✓
status of e_i	strongly rejected $pl(\overline{e_i},e_k) > pl(e_i, e_k)$	strongly accepted $pl(e_j,e_i,e_k) > pl(e_j,\overline{e_i},e_k)$

As shown here, it is inconsistent to say that at time $t+1$ e_j causes e_i in context e_k and at the same time e_i causes e_j in context e_k.

Proposition 3 *Let us consider three events* $e_i \in \Theta_{E_i}$*,* $e_j \in \Theta_{E_j}$ *and* $e_k \in \Theta_{E_k}$*, if* e_i *is a very strong cause of* e_j *in context* e_k *(i.e.* $pl(e_i,e_j,e_k) > pl(e_i,\overline{e_j},e_k)$*)then* e_j *cannot be simultaneously an effect and a very* *strong cause of e_i in context e_k (i.e. $pl(e_i,e_j,e_k)> pl(\overline{e_i},e_j,e_k)$).*

Example 3 *Given the two binary variables insomnia (I) and anxiety (A), insomnia can be the cause of anxiety and later anxiety may cause insomnia, but if the two events occur at the same time it is impossible to have I causes A and A causes I. In fact, as explained before to have A causes I, this means that I is abnormal thus, at time t, $pl(\{\overline{I}\})> pl(\{I\})$. If this person is anxious at time $t+1$, then $pl(\{(\overline{I}, A)\})< pl(\{(I,A)\})$. To have I causing A, A should be observed before the observation of I which contradicts with what is exposed above since A causes I if A is observed after I.*

Transitivity Causality is generally not transitive (Pearl 2000; Bonnefon, Dubois, and Prade 2008). However if events in the chain are temporally sequenced then it becomes transitive (Lewis 2004). It means that under the normal course of things, if e_i is a very strong cause of e_j in context e_k and e_j is a very strong cause of e_l in context e_k; and e_l is reported after e_i then e_i is considered as a very strong cause of e_l in context e_k.

Proposition 4 *Let e_i, e_j and e_k three events. Assume that:*

1. *e_i is a very strong cause of e_j in context e_k (i.e. $pl(e_j,e_k)< pl(\overline{e_j}, e_k)$ and $pl(e_j,e_k,e_i)> pl(\overline{e_j},e_k,e_i)$);*
2. *e_j is a very strong cause of e_l in context e_k (i.e. $pl(e_l,e_k)< pl(\overline{e_l},e_k)$ and $pl(e_l,e_k,e_j) > pl(\overline{e_l},e_k,e_j)$);*
3. *e_l is reported after e_i. Then, e_i is a very strong cause of e_l in context e_k (i.e. $pl(e_i,e_k,e_l) > pl(e_i,e_k,\overline{e_l})$).*

Example 4 *From the following observation: "An eruption of a volcano (Θ_E={yes, no}) sends thousands of tonnes of volcanic ash into the sky ($A = \{h\}$). The ash cloud is threatening to be a very strong cause of air travel chaos ($F = \{y\}$) across many countries", it is possible to conclude that volcano eruption very strongly causes travel disruption only if the eruption is considered a salient very strong cause of ash clouds ($pl(\{l,m\}) > pl(\{h\})$ and $pl(\{h\}|\{yes\}) > pl(\{l,m\}|\{yes\})$) and airlines delay their flights after the eruption of the volcano.*

Stability The last property corresponds to the stability of causality with respect to conjunction and disjunction of causes and effects. Conjunctions are here represented with intersections and disjunctions with unions. Three relations are verified: the first relation deals with causal relations sharing a same effect. It is more plausible that this common event is caused by only one of the causes (i.e. their disjunction) than by all causes simultaneously (i.e. their conjunction). The second one deals with causal relations sharing a same cause. This event can be considered as a common very strong cause of the disjunction of the effects involved in the causal relations. The third relation concerns also causal relations sharing a common cause. This event can be considered as a common very strong cause of the conjunction of the effects involved in the causal relations. More formally:

Proposition 5 *Let e_i, e_j and e_k be three events.*

- *If e_j very strongly causes e_i (i.e. $pl(e_i \mid e_j)> pl(\overline{e_i} \mid e_j)$) and e_k very strongly causes e_i (i.e. $pl(e_i \mid e_k)> pl(\overline{e_i} \mid e_k)$) then $e_j \cup e_k$ very strongly causes e_i (i.e. $pl(e_i \mid e_j \cup e_k)> pl(\overline{e_i} \mid e_j \cup e_k) = pl(e_i| e_j) \textcircled{\cup} pl(e_i| e_k) > pl(\overline{e_i}| e_j) \textcircled{\cup} pl(\overline{e_i}| e_k)$).*
- *If e_i very strongly causes e_j (i.e. $pl(e_j, e_i)> pl(\overline{e_j}, e_i)$) and e_i very strongly causes e_k (i.e. $pl(e_k, e_i)> pl(\overline{e_k}, e_i)$) then e_i very strongly causes $e_j \cup e_k$ (i.e. $pl(e_j \cup e_k \mid e_i)> pl(\overline{(e_j \cup e_k)}|e_i)$).*
- *If e_i very strongly causes e_j (i.e. $pl(e_j, e_i)> pl(\overline{e_j}, e_i)$) and e_i very strongly causes e_k (i.e. $pl(e_j, e_i)> pl(\overline{e_j}, e_i)$) then e_i very strongly causes $e_j \cap e_k$ (i.e. $pl(e_j \cap e_k|e_i)> pl(\overline{e_j \cap e_k}| e_k)$).*

Example 5 *If an agent ascribes that a virus (v) causes fever (f): $pl(\{f\}) < pl(\{\overline{f}\})$ and $pl(\{(v,f)\})> pl(\{v,\overline{f}\})$ and also the overexposure to the sun (o) causes fever, $pl(\{(o,f)\})>pl(\{(o,\overline{f})\})$. If he observes in the summer, a person suffering from fever, he may conclude that it is the effect of either a virus or the overexposure to the sun but less plausibly that he catches a virus and at the same time was overexposed to the sun:*
- $m(\{f\}|\{o\}) = 0.6$, $m(\{\overline{f}\}|\{o\}) = 0.3$, $m(\{f,\overline{f}\}|\{o\}) = 0.1$; $pl(\{f\}|\{o\}) = 0.7$;
- $m(\{f\}|\{v\}) = 0.7$, $m(\{\overline{f}\}|\{v\}) = 0.1$, $m(\{f,\overline{f}\}|\{v\}) = 0.2$; $pl(\{f\}|\{v\}) = 0.9$.
As shown in Table 3, the virus or the overexposure to the sun causes fever with a plausibility of 0.95 which is less than having fever with only one of them: 0.58. Note that the plausibility of suffering from a virus and at the same time to have been overexposed to the sun is only 0.87.

Table 3: $m(.|\{o\}) \textcircled{\cup}\, m(.|\{v\})$ and its corresponding *pl*

	|{v}		|{o}		(.|{o}) $\textcircled{\cup}$ (.|{v})	
	m	pl	m	pl	m	pl
{f}	0.7	0.9	0.6	0.7	0.42	0.97
{$\overline{f}$}	0.1	0.3	0.3	0.4	0.03	0.58
{f,$\overline{f}$}	0.2	1	0.1	1	0.55	1

Facilitation Ascription

This notion is very related to causality. An agent deals with facilitation when he is cautious in his causal interpretation of the sequence of events. In fact, as for identifying causal links, to ascribe facilitation the agent starts not believing in the occurrence of an event under the normal course of things and by observing another event, he changes his beliefs afterwards. However, this change consists to not believe in the event neither in its complement instead of accepting it as it is the case for causality.

Definition 3 ***Belief function facilitation ascription:*** *If an event e_j is very strongly rejected and after* *observing an event e_i it becomes ignored then e_i is said to very strongly facilitate the occurrence of e_j. Namely,*

$$0 < pl(e_i, e_j) = pl(\overline{e_i}, e_j) \leq pl(\overline{e_j}) < pl(e_j) \quad (8)$$

Example 6 *If an agent has at his disposal some background information about flight delay: $pl(\{y\}) = 0.3 < pl(\{n\}) = 0.8$ (a very strongly rejected event). After noticing that some airlines strikes, ($S = \{yes\}$), the agent revises his beliefs and the event $F = \{y\}$ becomes ignored (i.e. $pl(\{(y, yes)\}) = pl(\{(n, yes)\})$). Accordingly, he will perceive $S = \{yes\}$ as very strongly facilitating the occurrence of $F = \{y\}$ ($F = \{y\}$ is unsurprising, but not expected to the agent).*

Justification

If an agent judges that the occurrence of an event e_i gave reason to expect the occurrence of e_j, we deal with justification. e_i caused the agent to start believing e_j, and that it should not be surprised of having e_j reported afterwards.

Definition 4 ***Belief function justification:*** *Given a sequence of events, e_j is said to very strongly justify e_i, if e_i is ignored and becomes very strongly accepted after the observation of e_j. Namely,*

$$pl(\overline{e_i}, e_j) < pl(e_j) \leq pl(e_i) = pl(\overline{e_i}). \quad (9)$$

Example 7 *Let $\Theta_W = \{cold, hot, warm\}$. Assume that an agent ignores if the weather is cold $pl(\{cold\}) = pl(\{warm, hot\}) = 0.5$. After observing the event: many persons are wearing coats, $CL = \{coat\}$, the agent very strongly accepts $W = \{cold\}$. $pl(\{(cold, coat)\})=0.8 > pl(\{(hot, coat), (warm, coat)\}) = 0.1$. In this case, he concludes that $CL = \{coat\}$ very strongly justifies the cold weather.*

Attenuation and Confirmation

Within the qualitative models (Bonnefon et al. 2006), if an event is held as accepted, then after the observation of a second event, it only may remain accepted or becomes rejected. Thus, attenuation and confirmation do not make sense. In qualitative possibilistic framework (Benferhat and Smaoui 2008), the authors have shown that no distinction is made between weak independence and the case of confirmation and that the concept of attenuation cannot also be defined within that model. In the quantitative belief function framework acceptance, rejection can be confirmed or attenuated upon observing a new event.

Definition 5 ***Belief function confirmation:*** *An event e_j is said to confirm another event e_i if the plausibility of observing e_i after observing e_j is greater than the plausibility of observing e_i alone. Namely,*

$$pl(e_i) \cdot pl(\overline{e_j}) < pl(e_i, e_j) < pl(e_j) \quad (10)$$

Definition 6 ***Belief function attenuation:*** *e_j is said to attenuate e_i if the plausibility of observing e_i after observing e_j is smaller than the plausibility of observing e_i alone. Namely,*

$$pl(e_i, e_j) < pl(e_i) \cdot pl(e_j) < pl(e_j) \quad (11)$$

Example 8 *Suppose that an agent initial knowledge about the weather is:($pl(\{cold\}) = 0.8$, $pl(\{hot\}) = 0.1$ and $pl(\{cold, hot\}) = 1$). If later, he observes many people eating ice creams, $I = \{yes\}$, then his beliefs according to this new information are updated: $pl(\{(cold, yes)\}) = 0.4 < pl(\{cold\}) = 0.8$. He identifies eating ice creams as attenuating his belief about the cold weather.*

Conclusion

In this paper, we have proposed a causal model that an agent will use to identify causal links between events in a chain. Since his background knowledge is expressed with belief functions, our model handles both propositional and n-ary variables. We have made a distinction between events related in a causal way and those when facilitation or justification are involved according to the definitions of acceptance and rejection that we have proposed. Confirmation and attenuation of acceptance make sense in our model which is not the case for the qualitative model based on the logical rules.

As future works, we intend the inclusion of other definitions of acceptance to ascribe causality. We plan also to let our model able to identify causal links when external events are proceeded on the system using of the DO operator (Pearl 2000; Boukhris, Elouedi, and Benferhat 2011). Doing this way, will help to better identify causal links because these interventions will allow the distinction between correlation and causation.

References

Benferhat, S., and Smaoui, S. 2008. Quantitative possibilistic networks: Handling interventions. In *Mexican International Conference on Artificial Intelligence*, 720–731.

Benferhat, S.; Bonnefon, J.; Chassy, P.; Da Silva Neves, R. and Dubois, D.; Dupin de Saint Cyr - Bannay, F.; Kayser, D.; Nouioua, F.; Nouioua-Boutouhami, S.; Prade, H.; and Smaoui, S. 2008. A comparative study of six formal models of causal ascription. In *International Conference on Scalable Uncertainty Management*, 47–62.

Bonnefon, J.; Da Silva Neves, R.; Dubois, D.; and Prade, H. 2006. Background default knowledge and causality ascriptions. *European Conference on Artificial Intelligence* 11–15.

Bonnefon, J.; Da Silva Neves, R.; Dubois, D.; and Prade, H. 2008. Predicting causality ascriptions from background knowledge: model and experimental validation. *Int. J. Approx. Reasoning* 48(3):752–765.

Bonnefon, J.-F.; Dubois, D.; and Prade, H. 2008. Transitive observation-based causation, saliency, and the markov condition. In *International Conference on Scalable Uncertainty Management*, 78–91.

Boukhris, I.; Elouedi, Z.; and Benferhat, S. 2011. Modeling interventions using belief causal networks. In *International Florida Artificial Intelligence Research Society Conference, FLAIRS'11*, 602–607. AAAI Press.

Dubois, D.; Fargier, H.; and Prade, H. 2004. Ordinal and probabilistic representations of acceptance. *J. Artif. Intell. Res. (JAIR)* 22:23–56.

Eells, E. 1991. *Probabilistic Causality.* Cambridge: Cambridge University Press.

Geffner, H. 1990. Causal theories for nonmonotonic reasoning. In *National Conference on Artificial Intelligence. Boston, Massachusetts*, 524–530.

Giunchiglia, E.; Lee, J.; McCain, N.; Lifschitz, V.; and Turner, H. 2004. Non-monotonic causal theories. *Artificial Intelligence* 153:49–104.

Halpern, J., and Pearl, J. 2001. Causes and explanations: A structurel model approach. In J. Breese, D. K., ed., *the 17th Conf. on Uncertainty in Articficial Intelligence*, 194–202.

Halpern, J. Y. 2008. Defaults and normality in causal structures. In Brewka, G., and Lang, J., eds., *KR*, 198–208. AAAI Press.

Hilton, D., and Slugoski, B. 1986. Knowledge-based causal attribution: The abnormal conditions focus mode. *Psychological Review* 93:75–88.

Hume, D. 2006. *An Enquiry Concerning Human Understanding.* Digireads.com.

Lewis, D. 2004. *Causation as Influence.* Collins, Hall, and Paul. 75–106.

Pearl, J. 2000. *Causality: Models, Reasonning and Inference.* Cambridge University Press.

Peng, Y., and Reggia, J. A. 1990. *Abductive Inference Models for Diagnostic Problem-Solving.* Symbolic Computation. Spring-Verlag.

Shafer, G. 1976. *A mathematical theory of evidence.* Princeton university Press.

Shoham, Y. 1990. Nonmonotonic reasoning and causation. *Cognitive Science* 14(2):213–252.

Smets, P. 1990. The combination of evidence in the transferable belief model. *IEEE Pattern analysis and Machine Intelligence* 12(5):447–458.

Smets, P. 1991. About updating. In *Uncertainty in Artificial Intelligence*, 378–385.

Smets, P. 1998. *The transferable belief model for quantified belief representation*, volume 1. Kluwer Academic Publisher. 267–301.

Wakker, P. 2000. Dempster belief functions are based on the principle of complete ignorance. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 8(3):271–284.

Yager, R. R. 1983. Hedging in the combination of evidence. *Journal of Information and Optimization Science* 4:73–81.