Possibilistic Explanation

Sara Boutouhami and Aïcha Mokhtari

Institut d'Informatique, USTHB, BP 32, EL Alia, Algers, Algeria
e-mail: s_boutouhami@yahoo.fr , mokhtari_aissani@yahoo.fr

Abstract

The philosophy literature has been struggling with the problem of defining causality. There has been extensive discussion about it. Hume taught that talk of causation was metaphysics and should be consigned to the flames (D. Hume, 1748). We present a comparison between two approaches that model this notion: The first one is the normative approach which is based on an interventionist conception of the cause (D. Kayser, A. Mokhtari, 1998). The second one is the structural-model approach that gives a new definition of actual causes, using structural equations to model counterfactuals (J. Y.Halpern, J. Pearl, 2005). Both approaches propose a modelling of the explanation using the notion of causation. Halpern and Pearl define the notion of partial explanation, so that every explanation will not be considered equally good. In order to make difference between them, the authors add a probability to the set of possible contexts (J. Y. Halpern, J. Pearl, 2005). Our purpose in this paper is to modify this definition, rather than to use a probability (quantitative modelling). We suggest to affect a degree of possibility (a more qualitative modelling) which is nearer to the human way of reasoning (P. Prade, D. Dubois, 1994). We propose to give a stratification of all possible partial explanations for a given request of an agent. The explanations in the first stratum are more possible than those belonging to the other strata.

Keywords: Causality, possibility, query, agent.

1. INTRODUCTION

Causation is a deeply intuitive and familiar relation, gripped powerfully by common sense, or so it seems. But as is typical in philosophy, deep intuitive familiarity has not led to any philosophical account of causation that is at once clean, precise, and widely agreed upon. It is safe to say that none has yet succeeded. It is also safe to say that the effort put into their development has yielded a wealth of insights into causation (N. Hall, L. A., Paul, 2004). A spring of difficulties seems to be that the notion of causality is bound to other ideas like that of explanation, of responsibility or to the problems of diagnosis where, from observations possibly imprecise and uncertain, one searches by abduction the plausible cause(s) of a given situation. This makes it's comprehension difficult (J. Y. Halpern, J. Pearl, 2005).

The paper is organized as follows. We discuss briefly in the section 2, fundamentals of the normative method of causality (A. Mokhtari, 1997). We present in section 3, the structure-based causal models, the notion of weak cause, the basic idea is

to extend the basic notion of counterfactual dependency to allow "contingent dependency" (J. Y. Halpern, J. Pearl, 2005).

The two definitions yield a plausible and elegant account of causation that handles well many of examples which have caused problems for other definitions like pre-emption and double pre-emption. They agree also with the idea that causality is not transitive. Section 4 is devoted to a comparison between the two approaches.

Thereafter, we present the definition of the partial explanation and the explanatory puissance proposed in (J. Y. Halpern, J. Pearl, 2005). The explanation proposed in the normative approach is qualitative, which is not the case in the structural approach. In order to give a qualitative character to this later, we propose to affect a degree of possibility to the definition advocated by Halpern and Pearl and then we carry out a qualitative reasoning by using the possibilistic logic based on the operation of minimum (min-based), a background of the possibilistic logic is given in this section, (P. Prade, D. Dubois, J, 1994). Finally, in section 6, we conclude and we give some perspectives of this work.

2. NORMATIVE METHOD OF CAUSALITY

Normative method of causality (D.Kayser, A. Mokhtari, 1998), (A. Mokhtari, 1997), extended in (M. Khelfellah, A. Mokhtari, 2001) is based on an interventionist concept of causality where an agent has the choice to perform or not an action (free will). It is based also, on the principle which stipulates that "an action may cause one or many effects".

There are two main ideas to represent: the continuity of time and the openness of future and past. The former corresponds to the idea that things (actions, events, or others) do not happen discontinuously, the latter to the idea of non-determinism of past and future from the perspective of any particular point of time. We introduce hereafter some definitions.

Time has been explicitly defined by mean of tamed states. A *tamed state* is characterized by a subset of propositions and a date. A tamed state is a universe snapshot where the propositions of the point are true at a date of the snapshot.

Let T be the set of all tamed states, a *date* of a tamed state is defined by the following function:

$$date: \ T \rightarrow R$$

where $date(t)=d$ (noted d_t) means that d is the date tamed state t.

A *time line* is a set of tamed states which are in bijection with a set of dates. It represents a possible evolution of the universe, intended as the complete knowledge of the truth value of 'interesting' propositions.

A time point of this succession is supposed to respect the general principle: '*there is no effect without a cause'*. It is then the result of a *cause-effect* relation defined as a *causal rule.*

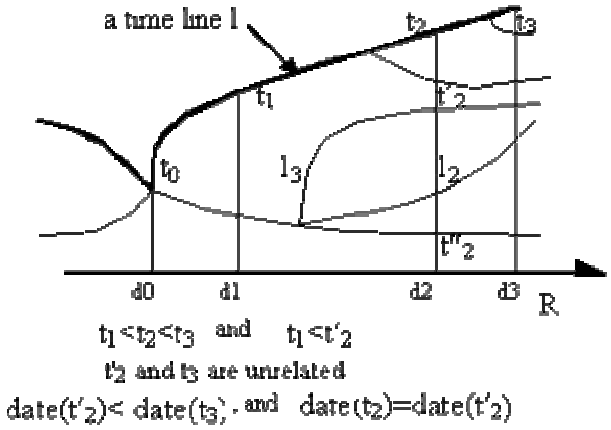

Fig 1 : Branching structure.

Let L be the set of all time lines, the tamed states of a time line are totally ordered by the precedence relation noted "≤", where $t_1 \leq t_2$ means that t_2 doesn't precede t_1 whenever $t_1 \leq t_2$ we have $d_{t1} \leq d_{t2}$.
The precedence relation expresses the principle: "no effect precedes its cause".

The representation of the "free will" notion requires a structure of time with a branching in the future. A branching in the past is also required to examine different courses of events leading to the same situation.

Two time lines l_1 *and* l_2 *coincide* until a tamed state t (noted *coincide*(l_1,l_2,t)) iff:

$$\forall\, t' \leq t,\ t' \in l_1 \equiv t \in l_2.$$

The set of preferred *time lines* for a tamed state t (noted $Lp(l,d_t)$) is defined as a function :

$$Lp : L \times R \rightarrow 2^L$$

Such that : $\forall\, l'\ (l' \in Lp(l,d_t) \supset coincide(l,l',t))$.

Among the timed states, some particular ones, called *choice points*, are intended to represent states of affairs where an agent can take decision of performing or not an action. The decision of the agent is represented as a *time line* splitting into two families of futures, these two families being composed of time lines sharing the same states up to the choice point.

The proposed language is defined on two levels:

1. The first level represents *static information*. It is a plain propositional language in which: P is a set of propositions we are interested in, A is a set of actions, E is a set of effects (facts, events…), with $A \cap E = \varnothing$ and $P = A \cup E$.
   A first level formula is either an action formula, or an effect formula.
   Let *FOR(A)* (respectively *FOR(E)*) be the set of *action* (respectively *effect*) formulas.
   An *effect literal* is either an effect, or an effect negation. The set of effect literals is noted *LIT(E)*. It is defined as:

$$LIT(E) = E \cup \{ \neg e : e \in E \}.$$

2. The second level expresses *dynamic information* represented by formulas of the form: *v(p,l,d),* which means that the formula *p* is true in time line *l* at date *d*.
   Causality is expressed by "normal causality" operator, noted "$\Rightarrow$". $a \Rightarrow e[\Delta]$ expresses that action *a* normally implies effect e in the delay Δ, unless there is an occurrence of an event inhibiting the effect *e*. The formulation of such notion needs non-monotonic reasoning which is expressed by means of action norm and inhibiting events.

Action norm is defined as the set of propositions which must normally be true in order to perform the action. Formally, the norm is defined as a function:

$$Norm : A \rightarrow 2^{FOR(E)}$$

where *norm(a)* contains the qualifications of action *a* (propositions which are true unless otherwise specified when an agent considers to perform *a*).

The set of the external events that can inhibit the effect of the action *a* is defined as a function:

$$Inhibit: LIT(E) \times A \rightarrow 2^{LIT(E)}.$$

where $e \in ihibit(e,a)$ iff *e'* inhibits the effect *e* of the action *a*.

The language of the normative method has certain predicates that were introduced to deal with the r*amification problem* (M.Khelfallah, A. Mokhtari, 2001):

- *occ*(*e,l,d*), with the intended meaning that the effect *e* has been generated in time line *l* at the date *d*. We have : $\forall$ *e,l,d : occ(e,l,d)*$\supset$ v(e,l,d).
- *Persist(e,*δ*),* where *e LIT(E)* and $\delta \in R$, means that the effect *e* normally persists during the delay δ.
- Another operator called "indirect implication", denoted "$\rightarrow$". "$(e_1, e_2) \rightarrow e[\Delta]$" expresses that the occurrence of the effect e_1, in a situation in which formula e_2 is true, indirectly implies the occurrence of the effect *e* within a delay Δ.
  These effects are generally delayed ones, i.e., their occurrence is not necessarily instantaneous.

The normative method tackles concurrent actions without any extension. The only condition to guarantee the consistency of the obtained results is the *non existence*

of actions with opposite effects executed simultaneously in the same time line. It provides an intuitively correct analysis of the main problems encountered in the AI literature: the explanation problem, the prediction problem and the ramification problem using a simple temporal ontology (for more details see (D. Kayser, A.Mokhtari, 1998)).

3. STRUCTURAL APPROACH

In a recent paper (J. Y. Halpern, J. Pearl, 2005)(J. Y. Halpern, J. Pearl, 2000), Halpern and Pearl propose a definition of cause (*actual cau*se) within the framework of structural causal models. Specifically, they express stories as a structural causal model (or more accurately, a causal world), and then provide a definition for when one event causes another, given this model of the story. The main idea is that a candidate C is an actual cause of an effect E when C and E have both occurred, and there exists some c*ounterfactual contingency W* under which E is *counterfactually dependent* on C.

Halpern and Pearl define their notion of causation within the language of structural models. Essentially, structural models are a system of equations over a set of random variables. We can divide the variables into two sets: endogenous (each of which have exactly one structural equation that determines their value) and exogenous (whose values are determined by factors outside the model, and thus have no corresponding equation). First we establish some preliminaries. We will generally use upper-case letters (e.g. X,Y) to represent random variables, and the lower-case correspondent (e.g. x, y) to represent a particular value of that variable. Dom(X) will denote the domain of a random variable X. We will use bold-face upper-case letters to represent a set of random Variables (e.g.**X,Y**). The lower-case correspondent (e.g. **x,y**) will represent a value assignment for the corresponding set. Formally, *a signature S* is a tuple ***(U,V,R)***, where ***U*** is a set of exogenous variables, *V* is a set of endogenous variables, and *R* associates with every variable $Y \in \boldsymbol{U} \cup \boldsymbol{V}$ a nonempty set *R(Y)* of possible values for *Y* (that is, the set of values over which Y ranges).

A *causal model* (or *structural model*) over signature *S* is a tuple *M*=(*S,F*), where *F associates with each variables* $X \in \boldsymbol{V}$ a function denoted F_X such that $F_X : (\times_{u \in U} R(\boldsymbol{U})) \times (\times_{Y \in V-\{X\}} R(Y)) \rightarrow R(X)$. F_X determines the values of *X* given the values of all the other variables in $\boldsymbol{U} \cup \boldsymbol{V}$. Causal models can be depicted as a *causal diagra*m: a directed graph whose nodes correspond to the variables in *V* with an edge from *X* to *Y* if F_Y depends on the value of *X*. Given a causal model *M=(S,F)*, a (possibly empty) vector ***X*** of variable in V, and vectors **x** and **u** of values for the variables in ***X*** and *U*, respectively, we can define a new causal model denoted $M_{\mathbf{X} \leftarrow \mathbf{x}}$ over the signature $S_{\boldsymbol{X}} = (\boldsymbol{U}, \boldsymbol{V}-\boldsymbol{X}, R_{|\boldsymbol{V}-\boldsymbol{X}})$. $M_{\mathbf{X} \leftarrow \mathbf{x}}$ is called a *submodel* of M by (J. Pearl, 2000), $R_{|\boldsymbol{V}-\boldsymbol{X}}$ is the restriction of *R* to the variables in $\boldsymbol{V}-\boldsymbol{X}$. Intuitively, this is the causal model that results when the variables in ***X*** are set to **x** by some external action that effects only the variables in ***X***. Formally $M_{\mathbf{X} \leftarrow \mathbf{x}} = (S_{\boldsymbol{X}}, F^{\mathbf{X} \leftarrow \mathbf{x}})$, where $F_Y^{\mathbf{X} \leftarrow \mathbf{x}}$ is obtained from F_Y by setting the values of the variables in ***X*** to ***x***.

3.1 Syntax and Semantics:

To make the definition of actual causality precise, it is helpful to have a logic with a formal syntax. Given a signature S=***(U,V,R)***, a formula of the form X=x, for $X \in \mathbf{V}$ and

$x \in R(X)$, is called *primitive event*. A basic *causal formula* (over *S)* is one of the form $[Y_1 \leftarrow y_1 \ldots Y_k \leftarrow y_k]\varphi$, where

- φ is a Boolean combination of primitive events,
- $Y_1, \ldots, Y_k$ are distinct variables in $\boldsymbol{V}$,
- $y_i \in R(Y_i)$.

Such formula is abbreviated as $[\mathbf{Y} \leftarrow \mathbf{y}]\varphi$. A *basic causal formula* is a Boolean combination of basic formulas. A causal formula ψ is true or false in a causal, given a context. We write $(M, \mathbf{u}) \models \psi$ if ψ is true in the causal model M given the context **u**.

Equipped with this background, we can now proceed to Halpern and Pearl's definition of actual cause.

Definition 3.2

Let M=(**U,V,F**), be a causal model. Let $\mathbf{X} \subseteq \mathbf{V}$, $\mathbf{X}=\mathbf{x}$ i*s an actual cause of* φ *if* the following three conditions hold (J.Y. Halpern, J. Pearl, 2001(a)):

(AC1): $(M,\mathbf{u}) \models \mathbf{X}=\mathbf{x} \wedge \varphi$.(that is, both **X=x** and φ are true in the actual world).

(AC2): There exists a partition (**Z,W)** of **V** with $\boldsymbol{X} \subseteq \mathbf{V}$, $\boldsymbol{W} \subseteq \mathbf{V} \backslash \mathbf{X}$ *and* some setting $(\mathbf{x}', \mathbf{w}')$ of the variables in (**X,W**) such that if $(M,\mathbf{u}) \models Z=z^*$ for $Z \in \mathbf{Z}$, then both of the following conditions hold :

- a) $(M,\mathbf{u}) \models [\mathbf{X} \leftarrow \mathbf{x}', \mathbf{W} \leftarrow \mathbf{w}']\neg\varphi$. In worlds, changing (**X,W)** from (**x,w**) to($\mathbf{x}', \mathbf{w}'$) changes φ from true to false.
- b) $(M,\mathbf{u}) \models [\mathbf{X} \leftarrow \mathbf{x}, \mathbf{W} \leftarrow \mathbf{w}', \mathbf{Z}' \leftarrow \mathbf{z}^*]\varphi$ for all subsets $\mathbf{Z}'$ of **Z.** In, words, setting **W** to $\mathbf{w}'$ should have no effects on φ as long as **X** is kept at its current value **x**, event if all the variables in an arbitrary subset of **Z** are set to their original values in the context **u**.

(AC3): ***X is minimal; no subset of X satisfies conditions AC1 and AC2.*** Intuitively, **x** is an actual cause of φ if (AC1) **x** and φ are the "actual values" and (AC2) under some counterfactual contingency **w,** the value of φ is dependent on **X**, such that setting **X** to its actual value will ensure that φ maintains its "actual value," even if we force all other variables in the model back to their "actual values." (AC3) is a simple minimality condition.

3.3. Explanation

Halpern and Pearl propose a new definition of the *explanation,* using structural equations to model counterfactuals. The definition is based on the notion of actual cause. Essentially, an explanation can be a fact that is not known to be certain but, if found to be true, would constitute an actual cause of the fact to be explained, regardless of the agent's initial uncertainty. An explanation is relative to the agent's epistemic

state, in that case, one way of describing an agent's state is by simply describing the set of the context the agent considers possible (J.Y. Halpern, J. Pearl, 2001(b)).

Definition 3.4:

(*Explanation*) given a structural model M, **X=x** is *an explanation of φ relative to a set K of contexts* if the following conditions hold (J. Y. Halpern, J. Pearl, 2005)(J. Y. Halpern, J. Pearl, 2000):

EX1: (M,**u**) $\models$φ for each **u**∈K. (that is, φ must hold in all contexts the agent considers possible the agent considers what he is trying to explain as an established fact).

EX2: **X=x** is a *weak cause* (without the minimal condition) of φ in (M,**u**) for each **u**∈K such that (M,**u**) $\models$**X=x**.

EX3: **X** is minimal; no subset of **X** satisfies EX2.

EX4: (M,**u**) $\models \neg$(**X=x**) for some **u**∈K and (M,**u'**) $\models$**X=x** for some **u'**∈K. (This just says that the agent considers a context possible where the explanation is false, so the explanation is not known to start with, and considers a context possible where the explanation is true, so that it is not vacuous).

Example 3.4.1:

Suppose two arsonists lit matches in different parts of a dry forest, and both cause trees to start burning. Assume now either match by itself suffices to burn down the whole forest. We may model such a scenario in the structural-model framework as follows (see Fig 2 below):

We assume two binary background variables U_1 and U_2, which determines the motivation and the state of mind of the two arsonists, where U_i is 1 iff arsonist i intends to start a fire. We then have three binary variables A_1, A_2 and B which describe the observable situation, where A_i is 1 iff. arsonist i drops the match, and B is 1 iff the whole forest burns down. The causal dependencies between these variables are expressed by functions, which say that the value of Ai is given by the value of U_i, and that B is 1 iff either A_1 or A_2 is 1.

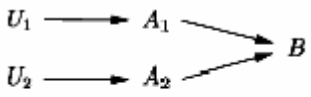

Fig 2: Causal graph

Causes and explanations for events, such as B =1 (the whole forest burns down), are defined by considering the values of variables in the above model and certain hypothetical variants. For example, arsonist 1 starting a fire is a (weak and an actual) cause of the whole forest burning down under every possible context in which arsonist 1 intends to start a fire. Moreover, arsonist 1 starting a fire is an explanation of the whole

forest burning down relative to the set of all possible contexts in which either arsonist intends to start a fire.

For more examples and extensive background on structural causal models, we refer especially to (J.Y. Halpern, J. Pearl, 2001(a)) (J.Y. Halpern, J. Pearl, 2001(b)).

4. COMPARISON BETWEEN THE TWO APPROACHES

The two approaches yield a plausible and elegant account of causality that handles well examples which have caused problems for other definitions like pre-emption. However, theses approaches don't use the same reasoning about causality. Consequenly, if we apply them to an example, we may not have the same result. We propose in this section a comparison that arises the common points (the cases which are handled by the two) and the different points.

4.1 Common points:

- "An explanation is relative to an agent's epistemic state". What counts as an explanation for one agent may not count for another agent", this need is expressed in the structural approach by the set of contexts that the agent considers possible. The normative method allows also this point, by the *norm notion*. The norm is relative to the agent and to the contexts of the causal history. Example (A. Mokhtari, 1997), Serge, big smoker, is just returned in his apartment; he attempts to smoke a cigarette. To this instant an explosion destroys the apartment. It is there had a gas leak". What that will be interpreted by the insurance companies, as having caused the destructive explosion of the apartment" It is not the event «light a cigarette", as well as it is a necessary condition to provoke the explosion. In fact, it is normal that serge smokes at his apartment. On the other hand if serge takes a walk in a petrochemical factory where it is natural to find gas in suspension in the atmosphere, and that he lights a cigarette, the cause of the explosion that will follow will be attributed to the event "light a cigarette". This event is not here a legitimate description of the natural course of the things.

- "The relation of causality is not transitive", the two definitions agree on this property. Halpern and Pearl give an example (see (J.Y. Halpern, J. Pearl, 2005)) to show that causality is not transitive, according to their definitions. For the normative approach, the transitivity is excluded in the definition of the relation of causality: "actions are the only elements that can change world state".
- The notion of action is well formulated in the normative approach by the notion of choice point. The two approaches agree on the idea that only action can changes the state of world. In the structural approach, only endogenous variables can be part of a cause.

- "The *pre-emption*": the two approaches deal with the pre-emption (where there are two potential causes of an event, one of which preempts the other). In (J.Y. Halpern, J. Pearl, 2005) a story is taken from (N. Hall, 2004) and the author give an adequate modelisation of the example. The normative approach can handle the pre-

emption cases using the notion of norm : when the first action is realised, it violates the norm (pre-condition) of the second action (the effect of an action must be true).

- "The choice of the type of variables is difficult": as in any approach, there is no formalism that enables to choose the types of the variables. This is a task involving experts in the domain and some heuristics to make appropriate choices. In (A. Mokhtari, 1997) the partition of P into actions A and events E, it is not always obvious to decide precisely what proposition falls into what category. The same case for (J.Y. Halpern, J. Pearl, 2005) it is not easy to make the choice for a given variable if she is endogenous or exogenous see (J.Y. Halpern, J. Pearl, 2005) show the importance of the choice of variable (adding an endogenous variable corresponding to an exogenous variable can result in there being an explanation when there was none before).

- The two methods can not deal with the disjunctive causes. In (J.Y. Halpern, J. Pearl, 2005) the only reasonable definition of "A$\vee$B causes C" seems to be that "A causes C or B causes C". In (A. Mokhtari, 1997) the problem is the disjunction of norm.

- "Inexplicable events are possible"; the two approaches allow inexplicable events.

- "The causality by production is transitive": In (J.Y. Halpern, J. Pearl, 2005), the condition AC2(b) captures some of the features of production (A forced B to happen, even if W=w'). For the normative approach, this is done by applying the function closure for ramification and by exploiting the static and causal rules (only action can modify the world state) (for more details see (M. Khelfellah, A. Mokhtari. 2001)).

- The two approaches use a non-monotonic reasoning. In (A. Mokhtari, 1997) the possibility that an unexpected event prevents *e* from remaining true, the notion of persistence and the notion of norm. In (J.Y. Halpern, J. Pearl, 2005) is that a cause is relative to a given context in a given causal model, which means that the cause change from one context to another one.

4. 2. Difference

- The time plays a crucial role in the perception of the causality, and yet it has no explicit representation in structural models. (M. Hopkins, J. Pearl, 2003). While in the normative approach the time is represented explicitly as a parameter in the predicates. And when we look for the cause or the explanation, the search is based on a base of facts and effects ordered on their realisation time.

- The normative approach allows the representation of the temporal aspect with the capacity to handle:

- The duration, the realisation of an action takes time,
- The delay, i. e., effects are not always immediate,
- The effects of concurrent actions (the only condition is the non existence of actions with opposite effects executed simultaneously).

- The problem with using structural causal models is that the language of structural models is simply not expressive enough to capture certain intricate relationships that are important in causal reasoning. Within the structural models framework, we are only dealing with values assignments to random variables and hence cannot express the following, among others (M. Hopkins, J. Pearl, 2003):
  - A distinction between a condition and a transition,
  - A distinction between the presence and the absence of an event,
  - Time and temporal relationships.

Halpern and Pearl propose a sophisticated definition for actual causality based on structural causal models, however although this definition works on many previously problematic examples, it still does not fit with intuition on all examples. Some of these difficulties can be traced to the limited expressiveness of the structural model formulation (M. Hopkins, J. Pearl, 2003). The explanation proposed in this approach unlike the normative approach is not qualitative. To handle this problem, we propose an improvement of this definition in the next section.

5. POSSIBILISTIC EXPLANATION

The explanation proposed in the latter section is not qualitative, in order to give a qualitative character to this later, we propose to affect a degree of possibility (a more qualitative modelling) which is nearer to the human reasoning (H. Prad, D. Dubois, 1988). We propose a new definition of explanation using the possibilistic logic, as a tool for ordering the set of possible explanations. The agent's epistemic state will be represented by describing the set of the interpretations that the agent considers possible. We propose a stratification of all possible partial explanations, this stratification reflects a hierarchy of priority between partial explanations.

We give a background on the possibilistic logic, for more details see (L.A.Zadah, 1978) (D.Dubois, J.Lang, H.Prade, 1994).

5.1 Possibilistic logic

The possibilistic logic based on possibility theory (L.A.Zadah, 1978) offers an uncertainty modelling framework for dealing with possibility distributions representing fuzzy information. It is for it, agent believes and the preferences are expressed by means of possibilistic logic.

Let L a finite propositional language and Ω be the set of all propositional interpretations. Let φ, Ψ,… be propositional formulas. $\omega \models \varphi$ means that ω is a model of φ.

A possibility distribution π is a mapping from a set of interpretation Ω into the unit interval [0, 1]. $\pi(\omega)$ represents the degree of compatibility of the interpretation ω with available pieces of information (D.Dubois, J.Lang, H.Prade, 1994).
By convention:

- ✓ $\pi(\omega) = 0$ means that ω is impossible to be the real world,
- ✓ $\pi(\omega) = 1$ means that ω is totally possible to be the real world,
- ✓ $\pi(\omega) > \pi(\omega')$ means that ω is preferred candidate to ω' for being the real world.

A possibility distribution π is said to be normalized if there exists an interpretation ω such that $\pi(\omega) = 1$. Given a possibility distribution π, two dual measures are defined :

- ✓ The possibility measure of a formula φ, defined by :

$$\prod(\varphi) = \max \{\pi(\omega) : \omega \models \varphi \text{ and } \omega \in \Omega\}$$

which evaluates the extent to which φ is consistent with the available beliefs expressed by π.

- ✓ The necessity measure of a formula φ, defined by :

$$N(\varphi) = 1 - \prod(\neg\varphi) = \min \{1 - \pi(\omega) : \omega \models \neg\varphi\}$$

which evaluates the extent to which φ is entailed by the available beliefs.

A possibilistic knowledge base $\sum$ is a set of weighted formulas :

$$\sum = \{(\varphi_i, \alpha_i) : i = 1, \ldots n\}$$

where φ_i is a propositional formula and $\alpha_i \in [0,1]$ which represents the certainty level of φ_i.

Each piece of information (φ_i, α_i) of possibilistic knowledge base can be viewed as a constraint which restricts the possibility degrees of possible interpretations (D.Dubois, J.Lang, H.Prade, 1994).

If an interpretation ω satisfies φ_i then its possibility degree is equal to 1, otherwise it is equal to $1 - \alpha_i$.

More formally, the possibility distribution associated with a weights formula (φ_i, α_i) is $\forall\omega \in \Omega$ [11] :

$$\pi_{(\varphi i,\ \alpha i)}(\omega) = \begin{cases} 1 - \alpha_i & \text{if } \omega \models \neg\varphi_I \\ 1 & \text{otherwise} \end{cases}$$

More generally, the possibility distribution associated with a possibilistic knowledge base $\sum$ is the result of combining possibility distributions associated with each weighted formula (φ_i, α_i) of $\sum$, namely $\forall\omega\in\Omega$ (D.Dubois, J.Lang, H.Prade, 1994) :

$$\pi_{\sum}(\omega) = \square \{\pi_{(\varphi i,\ \alpha i)}(\omega)\ (\varphi_i, \alpha_i) \in \sum\}$$

where $\square$ is in general either equal to the minimum (min) operator (in standard possibilistic logic), or the product operator (*).

In the rest of the paper, we only focus on the case where $\square$ = min. The possibilistic base $\sum$ is then called qualitative possibilistic knowledge base.

Explanation definition 5.2

Let π be a distribution of possibility i.e., a mapping from a set of interpretations Ω that the agent considers possible into the interval [0,1]. Let ω be an interpretation that the agent considers possible ($\omega\in\Omega$). Given a structural model M, **X=x** is *an explanation of* φ *relative to a set* Ω *of possible interpretations* if the following conditions hold:

EX1': $(M, \omega) \models \varphi$ for each $\omega\in\Omega$. (that is, φ must be satisfied in all interpretation the agent considers possible).

EX2': **X=x** is a *weak cause* of φ in (M, ω) for each $\omega\in\Omega$ such that $(M, \omega) \models$**X=x**.

EX3': **X** is minimal; no subset of **X** satisfies EX2.

EX4': $(M, \omega) \models \neg$(**X=x**) for some $\omega\in\Omega$ and $(M, \omega') \models$**X=x** for some $\omega'\in\omega$.

Not all explanations are considered equally good. Some explanations are most plausible than others. We propose to define the goodness of an explanation by introducing a degree of possibility (include priority levels between explanations).

The measure of possibility of an explanation is given by: Π(**X=x**)=max $\{\pi(\omega)$: $\omega \models$ **X=x**, $\omega\in\Omega\}$.

Example 5.2.1:

Suppose I see that Victoria is tanned and I seek an explanation. Suppose that the causal model includes variables for "Victoria took a vacation", "It is sunny in the Canary Islands", "Victoria went to a tanning". The set of Ω includes interpretations for all settings of these variables compatible with Victoria being tanned. Note that, in particular, there is an interpretation where Victoria both went to the Canaries (and didn't get tanned there, since it wasn't sunny) and to a tanning salon. Victoria taking a vacation is not an explanation (relative to Ω), since there is an interpretation where Victoria went to the Canary Islands but it was not sunny, and the actual cause of her tan is the tanning salon, not the vacation. However, intuitively it is "almost" satisfied, since

it is satisfied (by every interpretation in Ω), in witch Victoria goes to the Canaries. "Victoria went to the Canary Islands" is *partial* explanation of "Victoria being tanned". There is a situation where we can't find a complete explanation (it is inexplicable). In the next section we give our defionition of the goodness of a partial explanation.

Definition 5.3

Let $\Omega_{X=x,\varphi}$ be the largest subset Ω' of Ω such that **X=x** is an explanation of φ relative to $\Omega_{X=x,\varphi}$ (it consists of all interpretations in Ω except the those where **X=x** is true but is not a weak cause of φ) (S. Boutouhami, A. Mokhtari, 2005).
$\Omega'=\Omega-\{\omega: \omega\in\Omega \mid \omega \models \mathbf{X=x}, \omega \models \varphi$ and **X=x** is not a weak cause of φ $\}$.

- **X=x** is a *partial explanation of φ with the goodness* $\Pi(\Omega_{X=x,\varphi} \mid \mathbf{X=x})=\max\{\pi(\omega) : \omega \models \mathbf{X=x}, \omega \in \Omega'\}$.
- **X=x** is a α*-partial explanation of φ* relative to π and Ω, if Ω′ exists and $\Pi(\Omega_{X=x,\varphi} \mid \mathbf{X=x})\geq\alpha$.
- **X=x** is an *partial explanation* of φ relatively to π and Ω iff **X=x** is a α*-partial explanation* of φ, and $\alpha \geq 0$.

Partial explanations will be ordered, in a set of strates $S_{\alpha 1}\cup\ldots\cup S_{\alpha n}$ for a given request (see fig 3).

- In the strate $S_{\alpha 1}$ we will have complete explanations if there exists,
- X=x is in the strate $S_{\alpha i,}$ if $\Pi(\Omega_{X=x,\varphi} \mid \mathbf{X=x})=\alpha i$,
- Let X=x be a partial explanation in the strate $S_{\alpha i}$, and Y=y a partial explanation in the strate $S_{\alpha j}$. X=x is a partial explanation more plaussible than the partial explanation Y=y If $\Pi(\Omega_{X=x,\varphi} \mid \mathbf{X=x})=\alpha i > \Pi(\Omega_{Y=y,\varphi} \mid \mathbf{X=x})=\alpha j$.

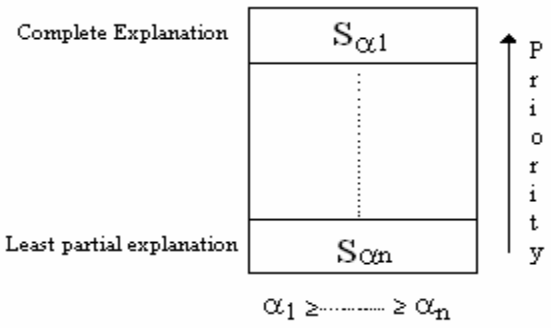

Fig 3 : A stratification of partial explanations

Example 5.3.1:

Suppose that two arsonists drop lit matches in different parts of a dry forest, and both cause trees to start burning. Consider the Disjunctive scenario; either match by itself suffices to burn down the whole forest. There are two possible explanations in the disjunctive scenario: arsonist 1 (A1) did it or arsonist 2 (A2) did it. Suppose that there is a variable O corresponding to the presence of oxygen. Suppose that there is an interpretation in which there is no oxygen (very unlikely). In that case, the presence of oxygen becomes a partial explanation of the fire. Nevertheless, it is an explanation with very little *explanatory power.*

Let π^- a distribution of possibility on the set Ω^- of interpretations (includes Ω). π^- represents the agent's prior possibility before the explanandum φ is observed or discovered. Thus, π is the conditioning π^- on φ and Ω consists of the interpretation in Ω^- that satisfy φ. $\pi = \pi^-(. \mid \varphi)$.

The usual definition of a conditional distribution of possibility is:

$$\pi(\omega \mid \varphi) = \begin{cases} 1 \text{ if } \Pi(\varphi)=\pi(\omega), \\ \pi^-(\varphi) \text{ if } = \pi^- < (\omega)\Pi(\varphi) \text{ and } \omega \mid \neq \varphi, \\ 0 \text{ else.} \end{cases}$$

Conditioner with φ consists on a revision of degree of possibility associated to different interpretations, after having the certain information φ. (φ is a certain information, so interpretation that falsifies φ, are impossible) (S. Benferhat, S. Lagrue, O. Papipi, 2003).

We propose the measure of *explanatory power* of **X=x** to be $\prod^-(\Omega_{X=x,\varphi} \mid X=x) = \max \{\pi^-(\omega) : \omega \models X=x, \omega \in \Omega' \}$.

In the precedent example, the presence of oxygen is an explanation with a very low explanatory power, while the arsonist 1 has a high explanatory power.

5.4 Algorithm of Generation of Strata

The main idea of our algorithme is to provide a set of choice of ordred partial explanations for a given request of the agent.

Let φ be request for which the agent seeks an explanation, let V be the set of endogenous variable, let $X \subseteq V / Y_i$, $Y_i \in \varphi$ the set of possible variable my formulate the explanation. For all subset X' of X, decide if there exist an attibution of values wich make it a partial explanantion, if is the case then compute $\prod(\Omega_{X'=x',\varphi} \mid X'=x')$. Ones that done, seek if there exist a strate wich will belong to, then add it to this latter ; If there isn't, create a new strate wich will have as an element this partial explanation, finally insert the new strate in the apropriate order with the existent strates. At the end of that algorithm we will have all partial explanations, this strusture facilitate the task of searching an explanation when we have a new consideration of the agent, an adaptation with the evolution of the agent beliefs.

S : the set of all partial explanation;
V : the set of endogenous variables.
φ : the fact to be explained.
Dom(X): the set of possible attributions of X.

Input : {S=∅, **V**, φ, Ω, Dom(X)}

begin

X=V-{Yj}/ ∀ Yj, Yj is a variable in φ.
for all X′⊆X **do**
begin
- Decide if there exist x′∈Dom(X), such that X′=x′ is a α-partial explanation of φ relative to Ω.
- **if** X′=x′ a α-partial explanation **then**
begin
Compute $\prod(\Omega_{X'=x',\varphi}|X'=x')$; Let =αi=$\prod(\Omega_{X'=x',\varphi}|X'=x')$
if the strate $S_{\alpha i}$ exists **then**
Add {X′=x′} to the strate $S_{\alpha i}$.
else
begin
Create a new strate $S_{\alpha i}$
Add {X′=x′} to the strate $S_{\alpha i}$.
Insert the strate $S_{\alpha i}$ in the good order.
S=S∪ $S_{\alpha i}$.
end.
end.
end.
end.

Out put {S= $\cup S\alpha_i$}.

The problem is to compute the set of all Subset of X, wich are partial explanation of φ, the problem can be reduced to the problem of computing the set of all partial explnation, computing the set of all valid formulas among a *Quantified Boolean formula (QBF)= ∃B ∀C ∃D f(B,C,D)*, where "*∃B ∀C ∃D y*" is a reduction of guessing some X′⊆X and x′∈Dom(x), and deciding whether X′=x′ is α-partial explanation.

6. CONCLUSIONS

In this paper, we have given a comparison between two approaches that model the notion of causality. We have proposed the use of the possibilistic logic which provides a priority level between the explanations; we prefer the use of possibility instead of probability because on one hand, possibility reflects better the human reasoning, which is rather qualitative than quantitative.

We have proposed a stratification of all partial explanations for a given request. This stratification facilitates the task of searching a new explanation when we have a new consideration of the agent (an evolution of the agent beliefs).
As an immediate extension to this work, we analyse of the complexity of the reasoning process to find the stratification of the all partial possibilistic explanations

7. REFERENCES

[1] S. Benferhat, S. Lagrue, O. Papini, 2003. A possibilistic handling of partially ordered information, UAI 2003, pp 29-36.

[2] S. Boutouhami, A. Mokhtari, 2005. Possibilictic explanation, Proceeding of the Information & Communication Technologies International Symposium ICTI'S 05, Tetuan Morocco, pp 7-13, June 2005.

[3] N. Hall, L. A., Paul, Causation and counterefactuals, Edited by John Collins, Ned Hall and L. A. Paul.,(Eds) Cloth/June 2004.

[4] J.Y. Halpern, J. Pearl, Causes and Explanations: "A Structural-model Approach". To appear in British Journal for Philosophy of Science, 2005.

[5] J.Y. Halpern, J. Pearl, Causes and explanations: A structural-model approach, Part I: Causes, in: Proceedings of the Seventeenth Conference in Uncertainty in Artificial Intelligence (UAI-01), Seattle, WA, pp. 194–202, 2001.

[6] J.Y. Halpern, J. Pearl, Causes and explanations: A structural-model approach, Part II: Explanations, in: Proceedings of IJCAI-2001, pp. 27–34, Seattle, WA, 2001.

[7] J.Y. Halpern, J. Pearl, Causes and explanations: A structural-model approach, Technical Report R-266, UCLA Cognitive Systems Lab., 2000.

[8] M. Hopkins, J. Pearl, Clarifying the Usage of structural models for Commonsense Causal Reasoning. In Proceedings of the AAAI Spring Symposium on Logical Formalizations of Commonsense Reasoning, 2003.

[9] Hume, D.,1748, An Enquiry concerning Human Understanding. Reprinted Open Court Press, LaSalle, IL, 1958.

[10] D. Kayser, A. Mokhtari, Time in a causal theory. AMAI. Journal,Vol 22, pp 117-138, 1998.

[11] M. Khelfallah, A. Mokhtari, Ramification in the Normative Method of Causality, ECSQARUS.,pp 704-713. In LNCS n° ,Toulouse, France, September 2001

[12] A. Mokhtari, Action-Based Causal Reasoning, Applied Intelligence, The International Journal of Artificial Intelligence, Neural Networks, and Complexes Problem-Solving Technologies. Volume 7, Number 2, , pp 99-112 April 1997.

[13] J. Pearl. Causality Models, Reasoning, and Inference. New York: Cambridge University Press, 2000.

[14] H. Prade, D. Dubois, J. Lang, Possibilistic logic, In Handbook of Logic in Artificial Intelligence and Logic Programming, (D Gabbay, Oxford University Press, pp 439-513) , 1994.

[15] H. Prade, D. Dubois. Possibility theory: An approach to computerized, processing of Uncertainty. Plenum Press, New York, 1988.

[16] D.Dubois, J.Lang, H.Prade – Possibilistic logic – In handbook of logic in Artificial Intelligence and logic programming, (D.Gabbay et al., eds, 3, Oxford University Press : pages 439-513, 1994).

[17] L. A. Zadah- Fuzzy sets as a basis for a theory of possibility – Fuzzy Sets and Systems, 1, 3-28, 1978.