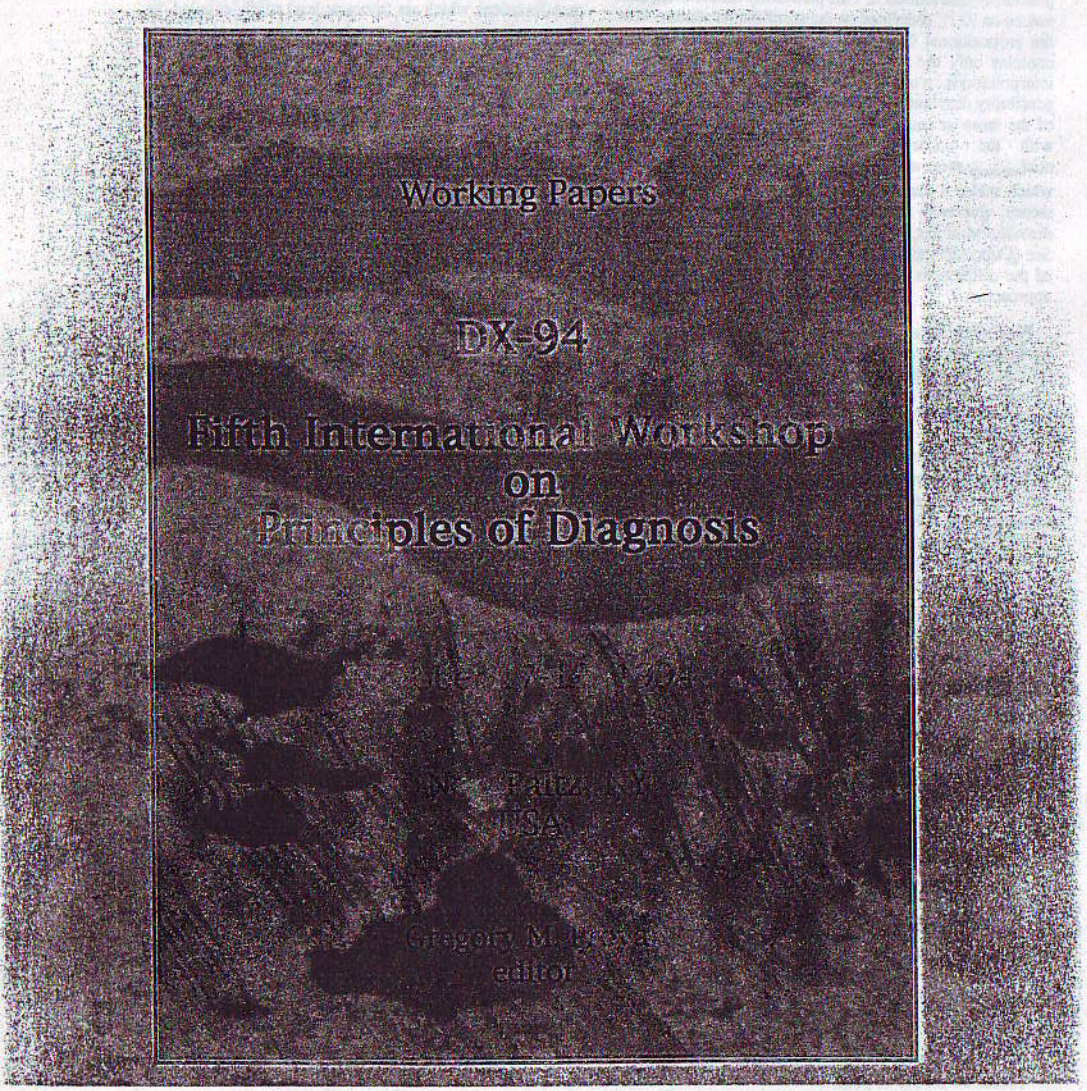
Qualitative and logical handling of uncertainty in a relational model for operational fault diagnosis

Didier Cayrac, Didier Dubois, Henri Prade



Qualitative and Logical Handling of Uncertainty in a Relational Model for Operational Fault Diagnosis

Didier Cayrac
Mission Optimization and Operations Assistance Dept.
MATRA MARCONI SPACE, 31, rue des Cosmonautes,
31077 Toulouse Cedex – FRANCE.

Didier Dubois, Henri Prade
IRIT, Institut de Recherche en Informatique de Toulouse,
C.N.R.S., Université Paul Sabatier, 118 route de Narbonne,
31062 Toulouse Cedex – FRANCE.

Abstract

Relational models for diagnosis are based on a direct description of the association between disorders and manifestations. The relational model we propose includes Reggia's model as a particular case, and allows a more expressive representation of the observations, and of the relation between disorders and manifestations. The model allows the distinction i) between manifestations which are certainly absent and those which are not (yet ?) observed, and ii) between manifestations which cannot be caused by a given disorder and manifestations for which we do not know if they can or cannot be caused by this disorder. This model uses a qualitative, non-probabilistic representation of uncertainty based on possibility theory and fuzzy sets. We also relate this possibilistic relational model to the logical approach to diagnosis, and outline an equivalent model using possibilistic logic. Using this model as a core, we propose extensions to exploit the knowledge of an indirect relation between disorders and manifestations, introducing intermediate entities such as (non directly observable) events. This approach is currently experimented in a satellite fault diagnosis application.

1. INTRODUCTION

Fault diagnosis is a problem that has been addressed by various kinds of techniques, depending on the type of failure to be detected or isolated, and on the available knowledge. Some diagnostic methods focus on a logical model of the system and perform abductive inference or consistency analysis e.g., (Marquis 1991). In a similar spirit, another class of approaches exploits causal information between failures and their symptoms; the causal knowledge is then represented as a Bayesian network (Pearl 1988), or more simply in terms of causal matrices (Peng and Reggia 1990). This paper is concerned by the latter type of methods, which are usually simple and require less costly information, having thus more appeal to operational applications. Its aim is to design a solution avoiding some of the drawbacks of Reggia's model, both at the representational and at the computational levels.

Matra Marconi Space (MMS) has been investigating and experimenting spacecraft diagnostic support systems for eight years, (Haziza 1990), and an operational system was installed in the Telecom 2 Satellite Control Center at the beginning of 1993 (Brenot et al. 1993). A progressive refinement of the models and of the reasoning paradigms (for instance to include the handling of uncertainty and time) is now being considered in the definition of a new generation of knowledge based systems, DIAMS-III (Cayrac et al. 1994).

Section 2 explains the rationale for the introduction of uncertainty handling and for the choice of possibility theory, on the basis of user's needs. In Section 3, we describe our model, which can be considered as extended Fault Mode Effect Analyses allowing the expression of uncertainty in the expected effects, and on its use in the diagnostic process along with uncertain observations. Section 4 relates it to the logical approach to diagnosis. On this basis, an extension allowing the exploitation of an indirect relation between disorders and manifestations is outlined in Section 5.

2. HANDLING UNCERTAINTY AND INCOMPLETENESS WITH THE POSSIBILITY THEORY

2.1. APPLICATION CONTEXT AND NEEDS

The knowledge exploited by the diagnostic support system and the information available on the current state of the satellite are incomplete and uncertain, due to the limitation of the knowledge acquisition investment, the limited observability of a satellite, and sometimes the complexity of telemetry analyses. Part of this imperfection can be abstracted by the experts, and expressed in the form of qualitative uncertainty. The information can be positive or negative, and possibly disjonctive. This qualitative expression of uncertainty led to the choice of the possibility theory, of which a brief overview follows in 2.2.

The system is a support tool for the identification of the fault(s). As an operational support system, its answers must be as reliable *and* as informative as possible given the available knowledge and observations. These conflicting requirements led to the joint use of consistency-based and abductive diagnostic approaches, their respective results

being sorted and presented to the user whith proper explanations. The system thus provides ordered sets of :

- the disorders that are (more or less) *consistent* with the symptoms observed. As pointed out by (Console, Torasso, 1991), these are the only "fully safe" solutions when the knowledge is incomplete.
- the disorders that are (more or less) *relevant* to the symptoms: disorders for which some predicted effects are indeed observed. Limitation due to incompleteness: a disorder will be discarded if none of its predicted effects is present.
- the disorders that (more or less) *cover* the symptoms: disorders that explain, i.e., predict, all the symptoms observed. Limitation due to incompleteness: a disorder will be discarded if some of its unpredicted effects are present.
- the disorders which are covers and for which *all predicted effects are indeed observed*. Limitation due to incompleteness: a disorder will be discarded if some of its predicted effects are absent.

These concepts, formalized in Section 3, can be straightforwardly extended to multiple faults. In this case, the set of relevant disorders provides the building blocks for the covers. The diagnostic process is iterative, and proper focusing must minimize the number of tests.

2.2. OVERVIEW OF POSSIBILITY THEORY

The basic idea of possibility theory, introduced by (Zadeh 1979) is to use fuzzy sets no longer to just represent the gradual aspect of vague concepts such as "large", but also to represent incomplete knowledge, tainted with imprecision and uncertainty. In this theory, uncertainty is described using dual possibility and necessity measures defined as follows :

A possibility measure Π defined on a propositional language and valued on [0,1] (or more generally on a totally ordered scale) satisfies the following axioms :

(i) $\Pi(\bot)=0$; $\Pi(\top)=1$
(ii) $\forall p, \forall q, \Pi(p \vee q)=\max(\Pi(p),\Pi(q))$

where $\bot$ and $\top$ denote respectively the ever-false proposition (contradiction) and the ever-true proposition (tautology). It should be emphasised that we only have $\Pi(p \wedge q) \leq \max(\Pi(p),\Pi(q))$ in the general case, since we may know that $p \wedge q$ is rather impossible (e.g., if $q=\neg p$, $p \wedge q$ is $\bot$, which is impossible) while p as well as q may remain somewhat possible when we are in a state of incomplete information. More generally, $\Pi(p \wedge q)$ is not a function of $\Pi(p)$ and $\Pi(q)$ only. This completely departs from fully truth functional multiple valued calculi like fuzzy logic, where degrees of truth of vague propositions are fully compositional with respect to *all* logical connectives.

A necessity measure is associated by duality with a possibility measure by:

$\forall p$, $N(p) = 1-\Pi(\neg p)$; it means that p is all the more certain that $\neg p$ is impossible.

(ii) is then equivalent to $\forall p, \forall q, N(p \wedge q) = \min(N(p),N(q))$, this means that for being certain of $p \wedge q$, we should be both certain of p and certain of q, and that the level of certainty of $p \wedge q$ is the smallest level of the ones attached to p and to q.

We have $N(p)>0 \Leftrightarrow \Pi(\neg p)<1 \Rightarrow \Pi(p)=1$ since $\max(\Pi(p),\Pi(\neg p))=\Pi(p \vee \neg p)=\Pi(\top)=1$; by duality with respect to Π, we only have $N(p \vee q) \geq \max(N(p),N(q))$. Indeed, we may be somewhat certain of the imprecise statement $p \vee q$ whithout being at all certain that p is true or that q is true.

We adopt the following conventions in possibility theory (see Figure 1):

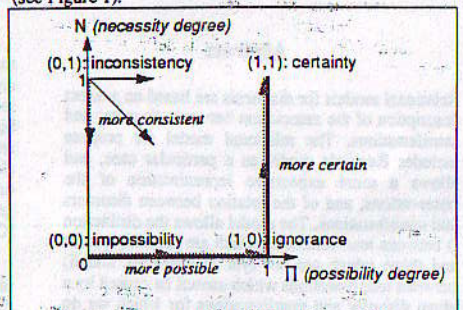


Figure 1. Conventions of Possibility Theory

- $N(p)=1$ means that, given the available knowledge, p is certainly true; conversely, if p is declared to be true we can consider it as certain.
- $1>N(p)>0$ means that p is somewhat certain and $\neg p$ not certain at all (since the axioms imply that $\forall p$, $\min(N(p),N(\neg p))=0$. $N(p)>0$ represents a tentative acceptance of p, to a degree $N(p)$; the last equality means that one cannot accept both p and $\neg p$.
- $\Pi(p) = \Pi(\neg p) = 1$ (equivalent to $N(p) = N(\neg p) = 0$) corresponds to the case of total ignorance; it expresses that, from the available knowledge, nothing enables us to say if p is rather true or rather false.
- $0<\Pi(p)<1$ (equivalent to $1>N(\neg p)>0$ means that p is somewhat impossible, i.e. that $\neg p$ is somewhat certain and p not certain at all. Again, it represents again the acceptance of $\neg p$.
- $\Pi(p)=0$ means that p is certainly false.

This constrasts with the probabilistic framework, which does not allow the distinction between possibility ($\Pi(A)$) and certainty ($N(A)$), nor between the certainty that A is false ($N(\overline{A})=1 \Leftrightarrow \Pi(A)=0$) and the absence of certainty that A is true ($N(A)=0 \Leftrightarrow \Pi(\overline{A})=0$). In probability theory, $P(\overline{A})=1-P(A)$, and therefore $P(\overline{A})=1 \Leftrightarrow P(A)=0$, while here $N(A)=0$ does not entail $N(\overline{A})=1$. In our context, this distinction between the notions of possibility and certainty is crucial. A disorder is "possible" as long as we cannot get evidence against it. It is certain when no other disorder (or

set of disorders if multiple faults are considered) is possible (this case never occurs if the knowledge is incomplete).

A possibility measure Π is generally associated with a possibility distribution π. This possibility distribution, also valued on [0,1], is defined on the set of interpretations Ω of the propositional language (for instance, if the language contains only the two propositions p, q, there are four interpretations : p true and q true, p false and q true...). This possibility distribution is supposed to encode the fuzzy set of the more or less plausible interpretations which agree with the available knowledge. We have $\Pi(p)=\max_{\omega \in mo(p)} \pi(\omega)$ where $mo(p)$ is the set of interpretations which make p true, i.e., the models of p. In possibility theory, ignorance is represented by a uniform possibility distribution of 1's, independently of the number of states. See (Dubois&Prade 1993a) for a more detailed discussion of the differences between probabilistic and possibilistic approaches.

Possibility theory offers a common framework for the modeling of uncertainty and imprecision in reasoning systems. The use of max and min operations, along with the complement to 1, is in agreement with the requirement of computational simplicity and with the rather rough and qualitative nature of the uncertainty that can be expressed in many real world applications. It should be noted that in possibility theory, the modelling of uncertainty remains qualitative. Indeed, we could use a finite completely ordered chain of levels of certainty ranging between 0 and 1, i.e. $\lambda_1 = 0 < \lambda_2 < \ldots < \lambda_n = 1$ instead of [0,1], with $\min(\lambda_i,\lambda_k) = \lambda_i$ and $\max(\lambda_i,\lambda_k) = \lambda_k$ if $i \leq k$, and $1 - \lambda_i = \lambda_{n+1-i}$. Multiple valued implications such as Gödel implication still make sense in a qualitative setting, since: $\lambda_i \rightarrow \lambda_k = 1$ if $\lambda_i \leq \lambda_k$, $\lambda_i \rightarrow \lambda_k = \lambda_k$ if $\lambda_i > \lambda_k$.

3. A POSSIBILISTIC RELATIONAL MODEL FOR DIAGNOSIS

In industry, knowledge of a direct relation between identified disorders and manifestations is often compiled at design time and known as "FMECA", Fault Mode Effects and Criticity Analyses. In Section 5, we will show how other kinds of knowledge, of a broader scope, can be exploited using the core diagnotic model proposed here.

3.1. REPRESENTATION OF INCOMPLETE, UNCERTAIN FAULT MODE EFFECTS

Incompleteness is inherent to FMECA. Our goal is to propose a flexible representation of the effects of fault modes that facilitates the expression of knowledge, down to the relevant level of detail, and that does not make any assumption about what is not said explicitly.

Let $\mathcal{D}$ be the set of possible Fault Modes, or *disorders* $\{d_1,\ldots,d_i,\ldots d_k\}$, and $\mathcal{M}$ a set of observable effects, or *manifestations* $\{m_1,\ldots,m_j,\ldots m_n\}$ related to these fault modes. In our model, disorders and manifestations are either present or absent.

For a given disorder d, we express its (more or less) certain manifestations, gathered in the fuzzy set $M(d)^+$, and those which are (more or less) impossible, gathered in the fuzzy set $M(d)^-$. Thus, the set $M(d)^+$ contains manifestations which are (more or less) surely entailed, or if we prefer caused, produced, by the presence of d alone. In terms of membership functions, $\mu_{M^+(d)}(m)=1$ means that m is always present when d alone is present. Conversely, the set $M(d)^-$ contains manifestations which (more or less) surely cannot be caused by d alone. E.g. $\mu_{M^-(d)}(m)=1$ means that m is never present when d alone is present. Complete ignorance regarding the relation between a disorder and a manifestation (we don't know whether m can be a consequence of d) is expressed by $\mu_{M(d)^+}(m)=\mu_{M(d)^-}(m)=0$. Intermediate membership degrees allow a gradation of the uncertainty. This model contrasts with early fuzzy set based approaches to fault diagnosis where the fuzzy relation describing causal knowledge represents the intensity of presence of symptoms (that are not binary) caused by disorders. Here, symptoms are binary and membership degrees represent uncertainty. Note that the fuzzy sets $M(d)^+$ and $M(d)^-$ are not possibility distributions because manifestations are clearly not mutually exclusive.

We require that $\forall d \in \mathcal{D}$, $M(d)^+ \cap M(d)^- = \emptyset$ (where the intersection is defined by the min operation), i.e. no manifestation can be a somewhat certain and somewhat impossible consequence of a disorder. In general, $M(d)^+ \cup M(d)^- \neq \mathcal{M}$, as knowledge about the effects is incomplete.

By complementation (defined by $\mu_{\overline{F}} = 1-\mu_F$), we obtain the fuzzy set $\overline{M(d)^-}$ of manifestations which are more or less *possibly* present when d is present. This corresponds to the usual duality between what is (more or less) certain , i.e. necessarily true (i.e., for an event A, $N(A)>0$), and what is (more or less) possibly true ($\Pi(A)>0$). Note that $M(d)^+ \subseteq \overline{M(d)^-}$ in the sense of fuzzy set inclusion (inequality between the membership functions). An even stronger inclusion holds. Since $M(d)^+ \cap M(d)^- = \emptyset$, the support of $M(d)^+$ (set of manifestations m_j such that $\mu_{M(d)^+}(m_j)>0$) is included in the core of $\overline{M(d)^-}$ (set of manifestations m_j such that $\mu_{\overline{M(d)^-}}(m_j)=1$); see Figure 2. This is in accordance with the fact that for crisp events A, we have $N(A)>0 \Leftrightarrow \Pi(\overline{A})<1 \Rightarrow \Pi(A)=1$ since one of A or $\overline{A}$, at least, should be completely possible.

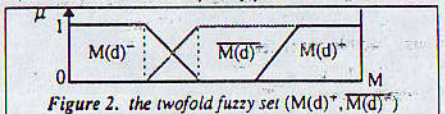


Figure 2. the twofold fuzzy set $(M(d)^+, \overline{M(d)^-})$ describes manifestations more or less certainly present ($M(d)^+$) when d is present more or less possibly present ($\overline{M(d)^-}$) when d is present

A pair of fuzzy sets (F,G) such that $F \cap G = \emptyset$ (i.e., $F \subseteq \overline{G}$) is called a twofold fuzzy set. Twofold fuzzy sets (F,G) have been introduced for modelling incompletely known sets, i.e. sets which characteristic function is ill-known. F represents the elements which more or less belong to this ill-known set and G represents the elements which more or less possibly belong to it. But $F \cup G$ might not cover the whole referential. Here, the pair $(M(d)^+, \overline{M(d)^-})$ defines a twofold fuzzy relation on $\mathcal{D} \times \mathcal{M}$. $\mu_{M(d)^+}(m)$ can thus be considered as the degree of necessity that d causes m, and $\mu_{\overline{M(d)}}(m)$ as the degree of possibility that d causes m.

This can be straightforwardly extended to multiple faults by working on subsets of $\mathcal{D}$ rather than with a disorder d supposed to be alone. In case several disorders can be jointly present (here we do not consider situations where disorder d_i followed by d_j has not the same effects in terms of manifestations as d_j followed by d_i), we have to define the set of manifestations produced by the presence of a pair of disorders (d_i,d_j) alone, and more generally by a tuple of disorders. Under the hypothesis tha effects can be added and do not interfere, we have in the crisp case

$M([di,dj])^+ = M(d_i)^+ \cup M(d_j)^+$ and

$M([d_i,d_j])^- = M(d_i)^- \cap M(d_j)^-$ (1)

In the fuzzy case :

$\mu_{M([di,dj])^+} = \max(\mu_{M(di)^+}, \mu_{M(dj)^+})$ and

$\mu_{M([di,dj])^-} = \min(\mu_{M(di)^-}, \mu_{M(dj)^-})$ (2)

If this hypothesis is not acceptable, a subset M(D) of entailed manifestations should be prescribed for each subset $D \subseteq \mathcal{D}$ of disorders which can be jointly present. In other words, we then work on a relation defined on $2^{\mathcal{D}} \times \mathcal{M}$ rather than on $\mathcal{D} \times \mathcal{M}$. If some disorders can never be jointly present, $2^{\mathcal{D}}$ should be replaced by the appropriate set $\mathcal{A}$ of associations of disorders which indeed make sense.

3.2. DIAGNOSTIC TECHNIQUES

A manifestation m is determined as being present or absent by the computation and qualification of the corresponding observable. M^+ and M^- are fuzzy sets of manifestations which are respectively (more or less certainly) present, and (more or less certainly) absent. $\mu_{M^+}(m)=1$ means that m is certainly present; $\mu_{M^-}(m)=1$ means that m is certainly absent; complete ignorance with respect to m (e.g., observable not yet tested) is represented by $\mu_{M^+}(m)=\mu_{M^-}(m)=0$. We impose $M^+ \cap M^- = \emptyset$. $(M^+, \overline{M^-})$ is thus a twofold fuzzy set containing manifestations more or less *certainly* present (M^+), and manifestations more or less *possibly* present ($\overline{M^-}$).

From a discrimination point of view, introducing graduality in the uncertainty of fault effects and in the characterisation of the observables allows a ranking of the solutions given by the system.

3.2.1. Consistency-based diagnosis

Given the sets M^+ and M^- at one point of the diagnostic process, the problem is to find the set of disorders $\hat{D}$ that are consistent with these symptoms. Especially if $\hat{D}$ is empty, we may be interested in the set $\hat{D}\hat{D}$ of subsets of disorders that are consistent with the symptoms (multiple fault case).

In the crisp case, ***a disorder is possible if it is consistent with the symptoms, i.e., if none of its expected consequences is absent, and none of its impossible consequences is present:*** :

$$\hat{D}_{crisp} = \{d \in \mathcal{D},\ M(d)^+ \cap M^- = \emptyset \text{ and } M(d)^- \cap M^+ = \emptyset\} \quad (3)$$

The extension of (3) to fuzzy sets of disorders can be done using the max-min consistency of two fuzzy sets. The consistency of F and G evaluates to what extent their intersection is not empty, i.e., the degree to which $F \cap G \neq \emptyset$. (Zadeh, 1979). It is defined by :

$$cons(F,G) = \max_u(\min(\mu_F(u), \mu_G(u))) \quad (4)$$

The degree to which $M(d)^+ \cap M^- = \emptyset$ is 1-cons($M(d)^+, M^-$). $\hat{D}$ can thus be defined by: $\forall d \in \mathcal{D}$,

$$\mu_{\hat{D}}(d) = \min(1-cons(M(d)^+, M^-); 1-cons(M(d)^-, M^+)) = 1-\max(cons(M(d)^+, M^-), cons(M(d)^-, M^+)) \quad (5)$$

$$= 1-\max_{i=1,n}(\min(\mu_{M(d)^+}(m_i), \mu_{M^-}(m_i)), \min(\mu_{M(d)^-}(m_i), \mu_{M^+}(m_i))) \quad (6)$$

(6) clearly expresses that *a disorder d is all the less a possible explanation as the fuzzy set of its more or less certain effects overlaps the fuzzy set of manifestations more or less certainly absent, or as the fuzzy set of its more or less impossible effects overlaps the fuzzy set of manifestations which are more or less certainly present.*

Let

$$Inconsistency(m_k,d) = \max(\min(\mu_{M(d)^+}(m_k), \mu_{M^-}(m_k)), \min(\mu_{M(d)^-}(m_k), \mu_{M^+}(m_k))) \quad (7)$$

represent the degree of inconsistency of the finding m_k with d. We have :

$$\mu_{\hat{D}}(d) = 1-\max_{i=1,n} Inconsistency(m_i,d) \quad (8)$$

Clearly, (5) straightforwardly extend to subsets D of disorders which altogether explain both M^+ and M^-, substituting D to the singleton {d}.

$$\forall D \in \mathcal{A} \subseteq 2^{\mathcal{D}},\ \mu_{\hat{DD}}(D) = \min(1-cons(M(D)^+, M^-), 1-cons(M(D)^-, M^+)) \quad (9)$$

Following the parsimony criterion, when $\hat{D} = \emptyset$, we have to look for two-element subsets D, and if none is found, for three-element subsets, and so on until a possible explanation is found.

3.2.2. Abductive diagnosis

The set $\hat{D}$ (respectively $\hat{DD}$) of possible causes of the symptoms observed contains all the disorders that are (more or less) consistent with the symptoms. It should be noted that irrelevant disorders, i.e., disorders which cannot be related to the available observations (e.g., a disorder which consequences are completely unknown) remain perfect candidates for a consistency based approach. It is therefore interesting to compute subsets of $\hat{D}$ (resp. $\hat{DD}$) containing disorders that are directly *relevant* to the current observations, or that *cover* the observations.

In the crisp case,*a disorder is relevant if some of its certain consequences are present manifestations or if some of its impossible consequences are absent* manifestations (the second part of this definition may of course be dropped if appropriate):

$$\hat{D}^*_{crisp} = \{d \in \hat{D}_{crisp},\ M(d)^+ \cap M^+ \neq \emptyset \text{ or } M(d)^- \cap M^- \neq \emptyset\} \quad (10)$$

This set eliminates the disorders that are not ruled out by the symptoms, but that are not suggested by them either.This can be easily extended to the fuzzy case using the consistency index. Namely,

$$\mu_{\hat{D}^*}(d) = \min(\mu_{\hat{D}}(d), \max(cons(M(d)^+, M^+), cons(M(d)^-, M^-))) \quad (11)$$

A more statistical view of these properties could be expressed through the use fuzzy cardinality ratios instead of consistency indexes; i.e., the extent to which $F \cap G$ is maximal can be defined as $\frac{|F \cap G|}{\min(|F|, |G|)}$. For instance,

$$\mu_{\hat{D}^{*'}}(d) = \frac{|M(d)^+ \cap M^+|}{\min(|M(d)^+|, |M^+|)} \quad (12)$$

is another relevance measure (evaluating to what extent d is a good*partial cover* of the more or less certainly present manifestations).

(11) and (12) express in different ways that *a disorder d is all the more relevant as the fuzzy set of its more or less certain effects overlaps the fuzzy set of manifestations more or less certainly present, (or as the fuzzy set of its more or less impossible effects overlaps the fuzzy set of manifestations which are more or less certainly absent).*

It is possible to refine further the set $\hat{D}^*$, for instance into the set $\hat{D}^{**}$ of covers. In the crisp case, ***a disorder is a cover if all present manifestations belong to its certain consequences, and (if appropriate), if the absent manifestations belong to its impossible consequences,*** i.e. to demand that the certain effects of d cover the present manifestations, and (if appropriate) that the impossible effects cover the absent manifestations.

$$\hat{D}^{**}_{crisp} = \{d \in \mathcal{D},\ M^+ \subseteq M(d)^+ \text{ and } M^- \subseteq M(d)^-\} \quad (13)$$

It can be extended to the fuzzy case by :

$$\mu_{\hat{D}^{**}}(d) = \min(\mu_{\hat{D}}(d), inc(M^+, M(d)^+), inc(M^-, M(d)^-)) \quad (14)$$

where inc(F,G) is a weak inclusion index requiring that $inc(F,G) = 1 \Leftrightarrow \mu_F \leq \mu_G$ for expressing the coverage property. It leads to take $inc(F,G) = \inf_m \mu_I(m) \rightarrow \mu_G(m)$ with $a \rightarrow b = 1$ if $a \leq b$ and $a \rightarrow b = b$ if $a > b$. The reason for this choice will be explicited in Section 4.

Here again, $\hat{D}^*$ and $\hat{D}^{**}$ can be easily generalised to sets $\hat{DD}^*$ and $\hat{DD}^{**}$, to take multiple faults into account, in a similar way as described in the previous section.

It should be noted that $\hat{D}^{**}$ is non-empty only when a rather complete knowledge of the effects of a given disorder is available. It should also be noted that as the information given to the system is incomplete, only $\hat{D}$ is completely "safe", i.e., an identified fault cannot be discarded from $\hat{D}$ because of missing information. Sets such as $\hat{D}^*$ and $\hat{D}^{**}$ should be used as additional focusing tools that help sorting the consistent disorders. A comparison with Peng and Reggia's approach (Peng and Reggia 1990), (Reggia et al. 1985) can be found in (Dubois and Prade 1993b). The main difference between the two approaches lies in the type of causal information that they capture. Our approach is more general because Peng and Reggia use only one type of causal information that corresponds to our $\overline{M(d)^-}$, and only positive observations, namely M^+. Moreover, they use a probabilistic treatment of uncertainty.

3.3. SIMPLE EXAMPLE (part 1)

This simplified example is provided as an illustration of the techniques proposed. A realistic example from DIAMS-III can be found in (Cayrac et al. 1994).

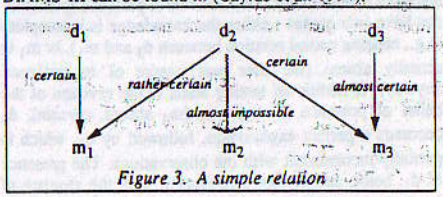


Figure 3. A simple relation

Let $\mathcal{D}=\{d_1,d_2,d_3\}$, $\mathcal{M}=\{m_1,m_2,m_3\}$; $\mu_{M^+(d1)}(m_1) = \mu_{M^+(d2)}(m_3)=1.0$, $\mu_{M^+(d2)}(m_1)=0.3$, $\mu_{M^-(d2)}(m_2)=0.7$, $\mu_{M^+(d3)}(m_3)=0.7$, all the other membership degrees are equal to 0.0. Figure 3 gives a graphical representation of this knowledge, using a simple qualitative scale for the membership degrees. We thus express for instance that m_1 is a certain consequence of d_1, while m_2 is an almost impossible consequence of d_2. The fact that for instance $\mu_{M^+(d1)}(m_3) = \mu_{M^-(d1)}(m_3)=0.0$ (no link on the graphical representation) does not mean that d_1 is independent from m_3. It only reflects our ignorance: m_3 may or may not be a consequence of d_1. In a similar way, we *do not know* the relationship between between d_3 and m_1, nor between d_1 or d_3 and m_2.

step	Observations	d_i	Possibility $\mu\hat{D}(d_i)$	Relevance $\mu\hat{D}^*(d_i)$	Cover $\mu\hat{D}^{**}(d_i)$
❶	m_1 present,	d_1	1.0	0.3	1.0
	rather certain:	d_2	1.0	0.3	1.0
	$\mu_{M^+}(m_1)$=0.7	d_3	1.0	0.0	0.0
❷	m_3 present,	d_1	1.0	0.3	0.0
	certain:	d_2	1.0	1.0	1.0
	$\mu_{M^+}(m_3)$=1.0	d_3	1.0	0.7	0.0
❸	m_2 present,	d_1	1.0	0.3	0.0
	almost certain:	d_2	0.3	0.3	0.3
	$\mu_{M^+}(m_2)$=0.7	d_3	1.0	0.7	0.0
❶	m_1 absent,	d_1	0.0	0.0	0.0
	certain:	d_2	0.3	0.3	0.3
	$\mu_{M^-}(m_1)$=1.0	d_3	1.0	1.0	1.0

Table 1. Partial trace of a sample diagnostic session

Let us consider a sample diagnostic session, of which Table 1 summarises the results. Initially, we consider that all disorders are completely possible (we don't express any prior preference). After the first symptom, (m_1 present, *rather certain*) is input, all three disorders remain completely possible. d_1 and d_2, being equally relevant and full covers, are the best candidate explanations. If the additonal criteria "all predicted effects present" was added, d_1 would be preferred. After (m_3 present, *certain*) is input, d_2 is the only cover, and is thus prefered, followed by d_3 and d_1, which are both partially relevant; (all disorders are still completely possible). Then, as the observation (m_2 present, *almost certain*) is inconsistent with the presence of d_2, this disorder is now *almost impossible* ; d_3 is preferred, as it is the most relevant of the fully possible hypotheses. However, it cannot explain the presence of m_1, and therefore is not a cover. This absence of perfect explanation can have two causes : either the knowledge is incomplete (e.g., missing causal relation between d_3 and m_1), or m_1 is actually absent (the user was unsure of its presence anyway). If additional testing leads to the revision of the belief of presence of m_1, e.g., (m_1 absent, *certain*), d_3 becomes a perfect explanation, followed by d_2, which is partially inconsistent with the observations. The presence of d_1, being completely inconsistent with the absence of m_1, is now ruled out.

3.4. FOCUSING THE DIAGNOSTIC PROCESS

As the diagnostic process we consider is iterative, it is useful to have advice with respect to the selection of the next observables to be tested. For this purpose, we propose measures of the utility of testing a given manifestation, in order to maximize the test impact on the evolution of the solutions.

From, (7) and (8), we can define the (negative) impact of the test of manifestation m_k on the possibility of the disorder d :

$$\text{Impact}(m_k,d) = \mu\hat{D}(d) - \mu\hat{D}^{new}(d)$$
$$= \max(\mu\hat{D}(d) + \text{Inconsistency}(m_k,d) - 1, 0) \quad (15)$$

taking into account (7). The impact of the test of m on the plausibility of d is all the greater as d was considered as highly plausible and the result of the test is incompatible with the presence of d.

We define the fuzzy sets $I_{m,D}^+$, $I_{m,D}^-$ of disorders, representing the maximal impact of the presence, respectively the absence of m, on the disorders of a subset D of $\mathcal{D}$; obtained respectively when $\mu_{M(d)^+}(m)=1$ and $\mu_{M(d)^-}(m)=1$. From (15) and (7) :

$$\mu_{I_{m,D}^+}(d) = \max(\mu\hat{D}(d) + \mu_{M(d)^-}(m) - 1, 0) \quad (16)$$

$$\mu_{I_{m,D}^-}(d) = \max(\mu\hat{D}(d) + \mu_{M(d)^+}(m) - 1, 0) \quad (17)$$

A measure of the discrimination power of the test of m with respect to the set D, which evaluates the extent to which the test will have a large impact on the possibility of the disorders in D, can then be defined from the fuzzy cardinalities $|I_{m,D}^+|$ and $|I_{m,D}^-|$ (defined as the sum over the membership degrees):

$$\text{MinImpact}(m,D) = \min(|I_{m,D}^+|, |I_{m,D}^-|) \quad (18)$$

In a set of manifestations having the same discrimination power, the maximal cumulated impact is a good selection criteria. Indeed, it corresponds to the best outcome of the test.

$$\text{MaxImpact}(m,D) = \max(|I_{m,D}^+|, |I_{m,D}^-|) \quad (19)$$

4. LINK WITH THE LOGICAL APPROACH TO DIAGNOSIS

In propositional logic, a so-called abductive explanation of a formula M (representing the observations) with respect to a theory $\mathcal{T}$ (e.g., in the simplest case, a collection of clauses $\neg d_i \vee m_j$, or being slightly more general, d_i can be replaced by a conjunction of literals d_k), is usually defined as a conjunction D of literals, such that $\mathcal{T}$, D $\models$ M, where D is also supposed to be consistent with $\mathcal{T}$. In the general case, the literals do not necessarily correspond to manifestations or to disorders only ("intermediary entities" are allowed, see next section), there is no restriction on the formulas in $\mathcal{T}$, nor on M, and the literals are not sorted. See for instance (Marquis 1991) for an extension of this definition to first order logic. The problem is then to compute (a) conjunction(s) D which is/are minimal (in the sense of subsumption) and non-trivial (i.e. not such that M$\models$D). Apart from this abductive view, we may also look for the disorders d which are ruled out by the positive or negative observations (M^+ and M^-), which are such that $\mathcal{T}$, M^+, $M^- \vdash \neg d$. This corresponds to the disorders which are not in $\hat{D}$ in the relational approach and also relates to consistency-based approaches since it comes down to check the inconsistency of $[\mathcal{T}, M^+, M^-, d]$. Methods based on the research of prime implicates or prime implicants have been proposed in the model based diagnosis literature, for instance by (Reiter 1987), (Reiter and de Kleer 1987), (Jackson 1991a). (see (Hamscher et al. 1992) for a collection of papers on the model-based approach).

More recently, (Jackson 1991b) has proposed to extend his prime implicate-based approach to possibilistic logic. In possibilistic logic (Dubois, Lang and Prade 1994), classical logic formulas are weighted by lower bounds of the necessity degree with which the formula is held for certainly true. Basically, d is then an abductive explanation of $N(m) \geq \alpha$ if $N(m) \geq \alpha$ can be derived from $\mathcal{T}$, now made of weighted clauses, together the weighted clause encoding $N(d) = 1$ (provided these clauses are consistent) using the extended resolution principle of possibilistic logic here written for clauses:

$$N(\neg p \vee q) \geq \alpha, N(p \vee r) \geq \beta \; ; N(q \vee r) \geq \min(\alpha,\beta) \quad (20)$$

We can translate the relational technique previously described into the following pattern of inference expressed in terms of a necessity measure N (using the possibilistic logic resolution rule): $N(d_i) \geq \gamma_i$ and $N(\neg d_i \vee m_j) \geq \alpha_{ij}$ entails $N(m_j) \geq \min(\gamma_i,\alpha_{ij})$. This can be applied $\forall i$, leading to the following lower bound for the necessity of m_j:

$$N(m_j) \geq \max_{i=1,n}\min(\gamma_i,\alpha_{ij}) \quad (21)$$

Let us now consider the following knowledge base made of uncertain implications $d_i \rightarrow m_j$ (defined as $\neg d_i \vee m_j$) between causes and effects, and of uncertain observations m_j, j = 1,k

$$N(\neg d_i \vee m_j) \geq \alpha_{ij}, \; i=1,m; j=1,n \; ; \; N(m_j) \geq \beta_j, \; j=1,k.$$

It corresponds to the case where $M^- = \emptyset$, $\forall d \; M(d)^- = \emptyset$, and $\mu_{M(di)^-}(m_j)=\alpha_{ij}$, which leads to $\hat{D} = \mathcal{D}$. Clearly, $\mu\hat{D}^{**}(d_i) = \min_j(\beta_j \rightarrow \alpha_{ij})$ (which justifies the particular implication involved in formula (14)) can be used as a plausibility weight for d_i, since it estimates to what extent d_i alone explains all the more or less certainly present manifestations. Indeed, $\gamma_i=\min_j(\beta_j \rightarrow \alpha_{ij})$ is the largest solution to the set of inequalitites derived from (21): $\beta_j \geq \max_{i=1,n}\min(\gamma_i,\alpha_{ij})$, j=1,k. Note that $\mu\hat{D}^{**}(d_i) = 1$ as soon as $\beta_j \leq \alpha_{ij}$, $\forall_j$, i.e. when $N(m_j) \geq \beta_j$, j = 1,k is a set of valid deductions from the weighted clauses $N(\neg d_i \vee m_j) \geq \alpha_{ij}$ and $N(d_i) = 1$.

If $\forall i$, $\mu\hat{D}^{**}(d_i) = 0$, we can consider to what extent a combination of causes $d_{i1},\ldots, d_{ir}$ can cause all the more or less certain observations, by the following expression $\mu\hat{D}^{**}([d_{i1}, \ldots, d_{ir}]) = \min_j[\beta_j \rightarrow \max(\alpha_{i1j},\ldots, \alpha_{irj})]$, starting with r = 2, i.e. looking for the possibility that the conjunction of two causes explains M^+. This corresponds to compute, using an ATMS language, the environments of the set of weighted formulas $N(m_j) \geq \beta_j$, j = 1,k, with r assumptions. Indeed, in the possibilistic Assumption-based Truth-Maintenance Systems (ATMS) introduced by (Dubois, Lang and Prade 1991) labels (which are weakly consistent, sound, complete and minimal) can be computed for a proposition m; such a label is made of a collection of weighted sets of assumptions. See (Benferhat et al. 1994) for applications to diagnosis problems.

To match the relational model previoulsy described, where it is possible that $M^- \neq \emptyset$ and $M(d)^- \neq \emptyset$, we should add negative information $N(\neg d_i \vee \neg m_j) \geq \lambda_{ij}$ and $N(\neg m_j) \geq \rho_j$.

$N(d_i) \geq \gamma_i$ and $N(\neg d_i \vee \neg m_j) \geq \lambda_{ij}$ entails $N(\neg m_j) \geq \min(\gamma_i,\lambda_{ij})$. We can deduce :

$$N(m_j) \geq \max_{i=1,n}\min(\gamma_i,\alpha_{ij}),\max_{i=1,n}\min(\gamma_i,\lambda_{ij}) \quad (22)$$

We then get, applying the resolution principle (20) to $N(\neg d_i \vee m_j) \geq \alpha_{ij}$, $N(\neg m_j) \geq \rho_j$, $N(\neg d_i \vee \neg m_j) \geq \lambda_{ij}$, $N(m_j) \geq \beta_j$ for all j :

$$N(\neg d_i) \geq \max[\max_j \min(\lambda_{ij}, \beta_j), \max_j \min(\alpha_{ij}, \rho_j)] \quad (23)$$

We recognize $\mu\hat{D}(d_i) \geq 1 - N(\neg d_i)$, where $\mu\hat{D}$ is given by (6) (with $\mu_{M(di)^+}(m_j) = \alpha_{ij}$, $\mu_{M(di)^-}(m_j) = \lambda_{ij}$, $\mu_{M^+(mj)} = \beta_j$, $\mu_{M^-(mj)} = \rho_j$). Thus, $\hat{D}$ indeed gathers the disorders which are somewhat possible, i.e. which are not more or less certainly ruled out by the positive or negative observations. We also find (14) again : $\mu\hat{D}^{**}(d_i) = \min(\min_j(\beta_j \rightarrow \alpha_{ij}), \min_j(\rho_j \rightarrow \lambda_{ij}))$, which estimates to what extent d_i alone explains all the more or less certainly present and all the more or less certainly absent manifestations, using the same line of reasoning as in the previous case.

Obviously, in the general case we should work with a possibilistic theory $\mathcal{T}$ allowing for weighted justifications of the form $N(\neg d_{i1} \vee \ldots \vee \neg d_{is(k)} \vee m_j) \geq \alpha_{kj}$, if the disorders interact in the manifestations they cause.

This brief discussion is an indication of the agreement between the relational approach (in the incomplete information case with graded uncertainty) and the possibilistic logic approach. We use this comparison as a basis of an extension of the model in the next section.

5. EXPLOITING AN INDIRECT RELATION BETWEEN MANIFESTATIONS AND DISORDERS

5.1. EXTENDED MODEL

Let us now consider an extension of the model, in which we allow the representation of indirect links between disorders and manifestations. In addition to disorders and manifestations, we consider intermediate entities called events, that are not directly observable. Let $\mathcal{E}=\{e_1,\ldots,e_i,\ldots e_p\}$ be the set of all the events used.

The knowledge is now expressed as a theory, $\mathcal{T}$ consisting of a collection of clauses in which disorders can only appear as causes. All clauses are weighted by a lower bound of a necessity degree (equal to 1 if the implication is certain). As previously, the (positive and negative) observations m or $\neg$m can also be weighted by a lower bound of a necessity degree.

5.2. COMPILATION OF THE DIRECT RELATION

From these definitions, it is possible to derive the direct relation between the disorders and the manifestations, characterised by $\mu_{M(di)^+}(m_j)$ and $\mu_{M(di)^-}(m_j)$, in the following way :

Let α_{ij} be maximal such that $N(m_j) \geq \alpha_{ij}$ can be derived from $\mathcal{C}$ and the weighted clause $N(d_i) = 1$ *(24)*

Let $\vdash$ denote the entailment in possibilistic logic (based on extended resolution).We will now show that :

$\mathcal{C} \cup \{N(d_i)=1\} \vdash N(m_j) \geq \alpha_{ij}$
$\Leftrightarrow \forall \rho_j, [\mathcal{C}, N(\neg m_j) \geq \rho_j] \vdash N(\neg d_i) \geq \min(\rho_j, \alpha_{ij})$ *(25)*

It means that if we are at least certain to the degree ρ_j that m_j is absent, we can deduce that d_i is only possible to the degree $1-\min(\rho_j,\alpha_{ij})$.

Proof :

Any possibilistic formula (p,α) has a truth value $\max(t(p),1-\alpha)$, where $t(p)$ is the classical truth value of p (i.e. 0 or 1). The truth value of a possibilistic knowledge base $\mathcal{C}=\{(p_i,\alpha_i),i=1,...\}$ is $t(\mathcal{C})=\min_{i=1,...}\max(t(p_i),1-\alpha_i)$. The degree of possibility that an interpretation of the language describes the actual world is equated to the truth value $t(\mathcal{C})$ computed for this interpretation. Hence any formula (p,α), and any set of possibilistic formulas $\mathcal{C}$ is viewed as a possibility distribution over interpretations of the language.

Let π be the possibility distribution associated to $\mathcal{C}$. Then the semantic entailment of (p,α) from $\mathcal{C}$ is defined as $\mathcal{C} \models (p,\alpha) \Leftrightarrow \pi \leq \max(t(p),1-\alpha)$. Due to the soundness and completeness of possibilistic logic (Dubois, Lang and Prade, 1994), (25) comes down to proving the equivalence between inequalities linking functions defined on the set of interpretations:

$\min(\pi,t(d_i)) \leq \max(1-\alpha_{ij},t(m_j)) \Leftrightarrow \forall \rho_j,\ \min(\pi,\max(1-t(m_j),1-\rho_j)) \leq \max(1-t(d_i),1-\min(\rho_j,\alpha_{ij}))$ *(26)*

We note that this equivalence is trivial when $t(d_i)=0$. Let us consider the case when $t(d_i)=1$.

As $t(m_j) \in \{0,1\}$ (the gradual uncertainty on m_j is expressed in ρ_j), let us consider both cases:

a/ $t(m_j)=0$ (26) reduces to: $\pi \leq 1-\alpha_{ij} \Leftrightarrow \forall \rho_j, \min(\pi,1) \leq 1-\min(\rho_j,\alpha_{ij})$ which is obvious

b/ $t(m_j)=1$ (26) reduces to: $\pi \leq 1 \Leftrightarrow \forall \rho_j, \min(\pi,1-\rho_j) \leq 1-\min(\rho_j,\alpha_{ij})$.

As $\pi \leq 1$ is always true, we must have : $\forall \rho_j, \min(\pi,1-\rho_j) \leq 1-\min(\rho_j,\alpha_{ij})$, which can be shown easily since $\forall \rho_j, \min(\pi,1-\rho_j) \leq \max(1-\alpha_{ij},1-\rho_j)$. ■

Let λ_{ij} be maximal such that $N(\neg m_j) \geq \lambda_{ij}$ can be derived from $\mathcal{C}$ and the weighted clause $N(d_i) = 1$ *(27)*

Identically, we can deduce
$\mathcal{C} \cup \{N(m_j) \geq \beta_j\} \vdash N(\neg d_i) \geq \min(\beta_j,\lambda_{ij})$ *(28)*

We find again formula (21): $N(\neg d_i) \geq \max[\max_j \min(\lambda_{ij}, \beta_j), \max_j \min(\alpha_{ij}, \rho_j)]$, equivalent to the definition of $\mu\Delta$, and we can therefore use $\mu_{M(di)^+}(m_j) = \alpha_{ij}$, $\mu_{M(di)^-}(m_j) = \lambda_{ij}$ to define the relation exploited by the diagnostic process.

5.3. RATIONALE FOR A COMPILED RELATION

The compilation of the direct relation between disorders and manifestations can be done either during the knowledge acquisition process, i.e. when the application is developed, or at runtime, for each diagnostic session. On the one hand, the off-line compilation of the relation allows the respect of real-time constraints, as the response time when a new symptom is added is almost constant. On the other hand, dynamic compilation is often necessary, since many causal links, or their strenght, depend on dynamic aspects such as satellite configuration. Even in this case, compilation of the direct relation is useful, since it allows the use of focusing tools for an efficient iterative discrimination, and explanation tools to justify the results proposed. It also allows a better reactivity of the system when new symptoms are input. Its drawback is of course the limitation of the explanations immediately available, which do not include the intermediate events. This is however acceptable in our application, since most relevant events are observable.

5.4. SIMPLE EXAMPLE (part 2)

In addition to the sets $\mathcal{D}$ and $\mathcal{M}$ defined in Section 3.3, let $\mathcal{E} = \{e_1,e_2,e_3,e_4\}$ be the set of intermediate events, and $\mathcal{C} = \{d_1 \rightarrow e_1$ (N,1.0), $d_2 \rightarrow e_1$ (N,0.3), $d_2 \rightarrow e_2$ (N,1.0), $d_3 \rightarrow e_4$ (N,0.7), $e_1 \rightarrow m_1$ (N,1.0), $e_2 \rightarrow e_3$ (N,0.7), $e_2 \rightarrow m_3$ (N,1.0), $e_4 \rightarrow m_3$ (N,0.7), $e_3 \rightarrow \neg m_2$ (N,1.0)$\}$, set of clauses weighted by a lower bound of a necessity degree, be the available diagnostic knowledge. Figure 3 illustrate this knowledge using the same graphical representation as previously.

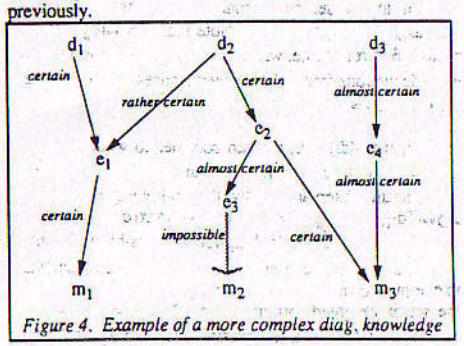


Figure 4. Example of a more complex diag. knowledge

Let us illustate the process described in the previous section. From $\{d_1$ (N,1.0)$\}$ and $\mathcal{C}$, and using the extended resolution principle of possibilistic logic, we can derive $\{m_1$ (N,1.0); we can therefore set $\mu_{M^+(d1)}(m_1)=1.0$. From $\{d_2$ (N,1.0)$\}$ and $\mathcal{C}$, we can derive $\{m_1$ (N,0.3), $\neg m_2$ (N,0.7), m_3 (N,1.0)$\}$; we set $\mu_{M^+(d2)}(m_1)=0.3$, $\mu_{M^-(d2)}(m_2)=0.7$ and $\mu_{M^+(d2)}(m_3)=1.0$, i.e., m_1 is a rather certain consequence of d_2, m_2 an almost impossible consequence, and m_3 a certain consequence. from $\{d_3$ (N,1.0)$\}$ and $\mathcal{C}$, we can derive $\{m_3$ (N,0.7); we set $\mu_{M^+(d3)}(m_3)=0.7$. All other membership degrees are set to 0.0. The compilation of the knowledge $\mathcal{C}$ illustrated on Figure 3 thus gives the direct relation described in the example of Section 3.3.

6. CONCLUSION

A possibilistic relational model was proposed for a greater expressivity of the knowledge and the observations. Consistency-based and abductive diagnosis algorithms using uncertain observations have been introduced, that allow a ranking of the disorders respectively in terms of their possibility and in terms of their relevance or coverage. The approach was related to the logical approach to diagnosis, and on this basis, a new extension of the model has been proposed to handle an indirect relation between a disorder and manifestations, using possibilistic logic. Undergoing experiments should prove that these techniques are ripe for introduction in real world knowledge based systems. It would also be particularly interesting to exploit temporal aspects in the chronologies of events, and to extend the model proposed to handle the functional models.

References

Benferhat, S., Dubois, D., Lang, J., and Prade, H. (1994), Hypothetical reasoning in possibilistic logic: basic notions and implementation issues, in Advances in Fuzzy Systems: Applications and Theory – Vol. I, Wang, P. Z., Loe, K. F., Eds., World Scientific Publ., Singapore,1994, 1-29.

Brenot J.M.,Caloud P., Valluy L., Gasquet A. (1993), "On the design and development choices to bring to operation a diagnostic expert system for the Telecom 2 satellite", Proc. Tooldiag Int. Conf. on Fault Diagnosis, 368-377,Toulouse, France.

Cayrac D., Dubois D., Haziza M., Prade H. (1994) "Possibility Theory in 'Fault Mode Effects Analyses – A Satellite Fault Diagnostic Application –'", Proc. of the 3rd IEEE Inter. Conf. on Fuzzy Systems (FUZZ'IEEE'94), June 1994. pp 1176-1181.

Console L., Torasso P. (1991), "A spectrum of logical definitions of model-based diagnosis", Computational Intelligence, vol.7 n.3, pp.133-141.

Dubois, D., Lang, J., and Prade, H. (1990), A possibilistic assumption-based truth maintenance system with uncertainty justifications, and its applications to belief revision, in Truth-Maintenance Systems (Proc. ECAI'90 Workshop), Martins, J. P., and Reinfrank, M., Eds., Lecture Notes in Computer Science, Vol. 515, Springer Verlag, 1991, 87-106.

Dubois D., Lang J., Prade H. (1994), "Automated Reasoning Using Possibilistic Logic : Semantics, Belief Revision and Variable Certainty Weights", IEEE Trans. on Knowledge and Data Engineering vol.6, 54-61, 1994.

Dubois D., Prade H. (1988) "Possibility Theory — An Approach to Computerized Processing of Uncertainty". Plenum Press, New York, 1988 (French editions, Masson, 1985 and 1987).

Dubois D., Prade H. (1993a) "Fuzzy sets and probability: Misunderstandings, bridges and gaps" Proc. of the 2nd IEEE Inter. Conf. on Fuzzy Systems (FUZZ'IEEE'93), April 1993, 1059-1068.

Dubois D., Prade H. (1993b) : "A fuzzy Relation-Based Extension of Reggia's Relational Model for Diagnosis". Proc. 9th "Uncertainty in AI" conference. pp.106-113, Washington, July 93. Heckerman D., Mamdani E., eds.

Haziza M. (1990) "Towards an Operational Fault Isolation Expert System for French Telecommunication Satellite Telecom 2": Proc. ESA Symposium "Ground data systems for spacecraft control", Darmstadt, FRG, pp. 385-392.

Hamscher W., Console L., de Kleer J. (Eds) (1992), "Readings in Model-Based Diagnosis", Morgan Kaufmann Publishers, San Mateo, CA, USA.

Jackson P. (1991), Prime implicates: their computation and use, in Proc. of the 11th Inter. Conf. on Expert Systems & Applications, Avignon, May 27-31, 1991, Vol. 1, 53-64.

Jackson P. (1991), "Possibilistic prime implicates and their use in abduction". Research Note, McDonnell Douglas Research Lab., St Louis, MO, USA.

Lee R.C.T. (1972), "Fuzzy Logic and the Resolution Principle", J. of the Association of Computing Machinery 19,109-119.

Marquis P. (1991), "Extending abduction from propositional to first-order logic", in Fundamentals of Artificial Intelligence Research (Proc. Inter. Workshop FAIR'91), Jorrand, Ph., and Kelemen, J., Eds.), Lecture Notes in Artificial Intelligence, Vol. 535, Springer Verlag, 141-155.

Pearl J. (1988) "Probabilistic Reasoning in Intelligent Systems : Networks of Plausible Inference", Morgan Kaufmann, CA, USA.

Peng Y., Reggia J.A. (1990) "Abductive inference models for diagnostic problem-solving", Springer Verlag, N.Y.C.

Reggia J.A., Nau D.S., Wang P.Y., Peng H. (1985) "A formal model of diagnostic inference". Information Sciences, 37, pp. 227-285.

Reiter R. (1987), A theory of diagnosis from first principles, Artificial Intelligence, 32, 57.

Reiter R. and De Kleer J. (1987), Foundations of assumption-based truth maintenance systems: preliminary report, in Proc. of the 6th National Conf. on Artificial Intelligence (AAAI'87), 183.

Zadeh L.A. (1979) "A theory of approximate reasoning". In : Machine Intelligence, Vol. 9.(J.E. Hayes, D. Michies, L.I. Mikulich, eds.), Halstead Press, NYC, pp. 149-194.