, , 1–19 ()

Fuzzy Temporal/Categorical Information in Diagnosis

JACQUES WAINER wainer@dcc.unicamp.br
Instituto de Computação, Universidade Estadual de Campinas
Campinas, SP 13803-970, Brazil

SANDRA SANDRI sandri@lac.inpe.br
LAC, INPE
São José dos Campos, SP 12201-970, Brazil

Abstract.
This paper proposes a way of incorporating fuzzy temporal reasoning within diagnostic reasoning. Disorders are described as an evolving set of necessary and possible manifestations. Ill-known moments in time, e.g. when a manifestation should start or end, are modeled by fuzzy intervals, which are also used to model the elapsed time between events, e.g. the beginning of a manifestation and its end. Patient information about the intensity and times in which manifestations started and ended are also modeled using fuzzy sets. The paper discusses many measures of consistency between the patient's data and the disorder model, and defines when the manifestations of the patient can be explained by a disorder. This work also discusses related issues such as the intensity of manifestations and the speed in which the disorder is evolving, given the patient's data, and how to use that information to make predictions about future and past events.

1. Introduction

Temporal information and temporal reasoning are important aspects of diagnostic reasoning [12, 8, 3, 19], specially in some domains, such as, the diagnostics of infectious diseases. For example, *Staphylococcus aureus* and short term *Bacillus cereus* are the only possible bacterial causes for nausea and vomiting within 1 to 6 hours from ingestion of contaminated matter. A patient with botulism (*Clostridium botulinum*) will only have those symptoms in 18 to 36 hours after the ingestion. In this case, temporal information is central for the correct diagnostic.

Another example, again within the food-borne diseases, is intoxication by ingestion of poisonous mushrooms from the species *Amanita phalloides, A. virosa*, and *A. verna* [14, Chap. 81]: it always causes abdominal cramps and nausea within 6 to 24 h from ingestion, which lasts for up to 24 h, followed by a period of 1 to 2 days of no symptoms, followed by hepatic and renal failure (which leads to death in 50% of the cases) (see Figure 1). Faced with a case in which the patient has ingested mushrooms, and felt abdominal cramps and nausea but does not show symptoms of renal and hepatic failure, one should not rule out intoxication by the *Amanita*

family without taking into consideration whether there has been enough time for those symptoms to develop. This is an example of temporal reasoning.

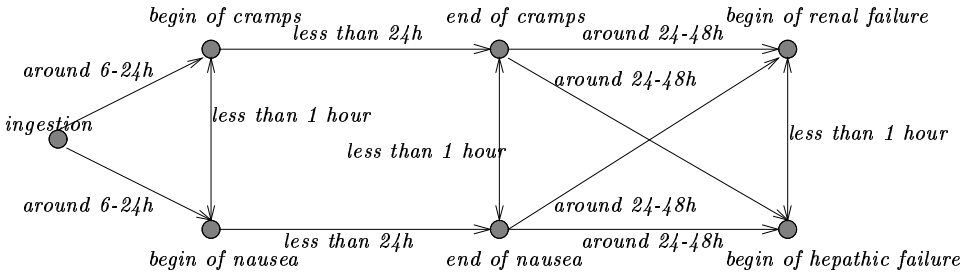

Figure 1. Temporal evolution of poisoning by *Amanita* mushrooms.

On the other hand, diagnostic problem solving has been an area of intense interest in Artificial Intelligence, and has generated many methodologies and models over the last two decades. There are many diagnostic models such as rule-based [2], set-based models [15, 13, 7], logic-based models [16, 11, 9], and case-based models [10].

But despite the importance of temporal aspects there are very few diagnostic models that incorporate time within its framework. [19] proposes an extension of Parsimonious Covering Theory (PCT) [15], a set cover model, so that information about the evolution of the manifestations, both the elapsed time between the beginning of the manifestations and their duration, can be taken into consideration. This paper extends that model so that instead of representing temporal information as intervals, it represents them as fuzzy intervals.

The usefulness of modeling temporal information as fuzzy intervals is clear when we consider the nature of the pieces of information found both in the temporal models of disorders and in the case information.

On the one hand, in most domains, the temporal constraints in the disorder model usually represent the accumulated knowledge of these intervals for a very large number of cases. Each of these temporal pieces of information may be modeled by a crisp interval, but then this interval might be either too small to contain spurious cases, or too large to provide useful information. Instead, modeling these temporal pieces of information in such a way as to allow one to distinguish between the *typical* cases from the only *possible* ones would be more useful.

In the medical domain, for instance, one frequently finds assessments such as "in disease d_1, symptom e_2 will follow symptom e_1 after *around 24 to 48 hours*". A physician confronted with the case of patient X for which symptom e_1 preceded symptom e_2 by 22 hours, would not completely disregard d_1 as a possible diagnostic. However, that would happen if *around 24 to 48 hours* would be modeled by the interval $[24, 48]$ in an automated system. Now, if the wider interval $[21, 52]$ hours would be used to model *around 24 to 48 hours* in d_1, and $[18, 26]$ hours would be

used to model the interval *around 20 to 24 hours* between e_1 and e_2 in d_2, the system would lose the means to state that d_2 is potentially a better diagnosis than d_1 for the case presented by X. In other words, one would lose the rich information that the case at hand is inside the typical cases of d_2 but only inside the possible cases in d_1.

However, this would not happen if fuzzy intervals would be used to model these temporal pieces of information. Moreover, a fuzzy interval is easily obtained by asking the expert to provide two nested intervals to account for the lapse of time between two events: one containing the interval of time between which the events typically occur, and another interval comprising all possible cases.

On the other hand, case information is usually tainted with vagueness. For instance, a patient is hardly ever capable of telling the precise moment in which a determined symptom started or ended. In this case, taking an imprecise interval to model a vague piece of information provided by the patient might have the effect of either ruling out a possible diagnosis, if that interval would be too small, or of unnecessarily increasing the number of possible diagnoses, if the interval would be too large. Here again, fuzzy intervals are more appropriate to model the temporal pieces of information than crisp intervals.

In this paper we will be interested in providing a model to answer the following questions:

- *when are the case information and the model for a disease inconsistent with each other from a temporal point of view?.* If we call the information on how a disease evolves as the model for a disease, as in the description of the evolution of symptoms for poisoning by the *Amanita* mushroom above, we are interested in knowing when the patient's progression of symptoms and the model are consistent or inconsistent with each other (and to what degree). For example, if the patient is exhibited nausea and abdominal cramps from 3 days and then 2 days after that showed signs of renal failure and loss of sensation in the limbs, can one state that this development of symptoms is consistent with the disease model?

- *when are the case information and the disease model categorically consistent, i.e. have all necessary symptoms in the model occurred (given that they have had enough time to occur)?.* Some symptoms must necessarily occur if the disease is allowed to develop. In particular, if the *Amanita* poisoning is not treated it will necessarily lead to the symptoms of renal and hepatic failure. If a necessary symptom does not occur (given that it has enough time to occur), then the case information and the disease model will be said to be categorically inconsistent with each other. For example, if the patient had abdominal cramps for one day, and after two days he had signs of renal failure, but no signs of hepatic failure, can one consider that the model of poisoning with *Amanita* and the case are categorically consistent?

- *when is the intensity of the symptoms in the case and the intensity of the symptoms as predicted by the disease model consistent with each other?.* Some symptoms have some form of intensity associated with it. Sometimes the intensity

is more precisely defined over some objective domain, for example, the patient fever was from 38.5 to 38.6 °C. Sometimes the intensity is defined over some subjective domain, for example, *severe* cramps, or *mild* nausea. If the patient reports suffering from mild nausea when the disease model specifies that the nausea will be strong, how much can one state that the model and the case patient are in agreement on the intensity of the symptom?

- *if we consider a particular disease as explaining all the patient symptoms, what else can we know about the other manifestations that the patient may have had, or will develop in the future.* For instance, if all the symptoms so far have occurred faster than expected, is it reasonable to expect that the symptoms that have not yet had the time to develop will also occur in a fast pace, and so immediate action should be taken by the physician?

The next section presents some basic concepts in fuzzy intervals, and define the model of a disease, and what is the case information available. Section 3 provides the answers to the questions raised above: when is the case temporally consistent with the model, when is the case categorically consistent with the model, and when is the model an explanation for the case. It will also include a subsection on intensity consistency. Section 4 discusses what other inferences can one perform given that it is assumed that a disease is the cause of a set of manifestations. Finally section 5 discusses the limits of the model proposed and future work.

2. Basic Definitions

2.1. Fuzzy Intervals

A (normalized) fuzzy set A in Θ [5], is characterized by a membership function $\mu_A : \Theta \rightarrow [0, 1]$, such that $\exists x \in \Theta, \mu_A(x) = 1$.

Let A and B be fuzzy sets in Θ, where Θ is a domain for which the operations $+$, $-$ and min are defined. The sum $A \oplus B$, the subtraction $A \ominus B$, the negation $-A$, and the intersection $A \cap B$ are respectively characterized by membership functions [5]:

$$\begin{aligned}
\mu_{A \oplus B}(z) &= sup_{\{(x,y)/z=x+y\}} min(\mu_A(x), \mu_B(y)) \\
\mu_{A \ominus B}(z) &= sup_{\{(x,y)/z=x-y\}} min(\mu_A(x), \mu_B(y)) \\
\mu_{-A}(z) &= \mu_A(-z) \\
\mu_{A \cap B}(z) &= min(\mu_A(z), \mu_B(z))
\end{aligned}$$

The height of non-normalized fuzzy set A is calculated as

$$h(A) = \sup_{x \in \Theta} \mu_A(x)$$

In this work, a fuzzy set A such that μ_A is convex will be called a fuzzy interval. An interval will be said to be positive if Θ is the real line, and $\forall x < 0, \mu(x) = 0$.

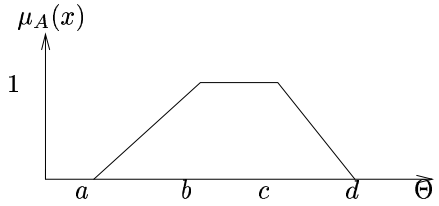

Figure 2. A trapezoidal fuzzy interval.

In some cases we will assume that the fuzzy interval is trapezoidal, as in Figure 2. In that case, the interval will be represented by a 4-tuple $\langle a, b, c, d\rangle$.

For a trapezoidal interval $A = \langle a_1, a_2, a_3, a_4\rangle$ the range $[a_2, a_3]$, where $\mu_A(x) = 1$, will be called the core of A. The range $[a_1, a_4]$, where $\mu_A(x) > 0$, will be called the support of A. An interval $\langle a, a, b, b\rangle$ will be said to be crisp and will be denoted by $\langle a, b\rangle$.

For two trapezoidal intervals $A = \langle a_1, a_2, a_3, a_4\rangle$, and $B = \langle b_1, b_2, b_3, b_4\rangle$, the $\oplus$ and $\ominus$ operations are simply $A \oplus B = \langle a_1 + b_1, a_2 + b_2, a_3 + b_3, a_4 + b_4\rangle$, and $A \ominus B = \langle a_1 - b_4, a_2 - b_3, a_3 - b_2, a_4 - b_1\rangle$.

Throughout this paper, we shall make use of four particular fuzzy intervals. Let θ be a moment in Θ. The fuzzy intervals describing the possibility of an event occurring *at any time*, *exactly at* θ, *after* θ, and *before* θ are respectively defined as:

$$\begin{aligned} I_{\text{anytime}} &= A \text{ such that } \forall x \in \Theta, \mu_A(x) = 1 \\ I_{=\theta} &= A \text{ such that } \mu_A(x) = 1 \text{ if } x = \theta, \text{ and } \mu_A(x) = 0 \text{ otherwise} \\ I_{\geq\theta} &= A \text{ such that } \forall x \in \Theta \text{ if } x \geq \theta, \mu_A(x) = 1, \text{ and } \mu_A(x) = 0 \text{ otherwise} \\ I_{\leq\theta} &= A \text{ such that } \forall x \in \Theta \text{ if } x \leq \theta, \mu_A(x) = 1, \text{ and } \mu_A(x) = 0 \text{ otherwise} \end{aligned}$$

We also use θ_0 to denote the present moment, and define $I_{beforenow} = I_{\leq\theta_0}$ and $I_{afternow} = I_{\geq\theta_0}$.

Finally, we will define that an interval A is tighter than an interval B (or, informally narrower) if $\mu_A(x) \leq \mu_B(x)$ for all $x \in \Theta$. If A and B are trapezoidal, then $A = \langle a_1, a_2, a_3, a_4\rangle$ is tighter than $B = \langle b_1, b_2, b_3, b_4\rangle$, iff $a_1 \geq b_1$, $a_2 \geq b_2$, $a_3 \leq b_3$, and $a_4 \leq b_4$.

2.2. Knowledge base of disorders

The knowledge base for a fuzzy temporal/categorical diagnostic problem is the information about disorders and how they evolve. The knowledge base is given by the tuple $\langle \Theta, D, M, N, P, V, T\rangle$ where:

- Θ is a time scale.
- D is the set of disorders.

- M is the set of manifestations.
- N is the necessary effects function that associates to each disorder d_l a set $M_L \subseteq M$ of manifestations that d_l *necessarily* causes. That is, if $N(d_1) = \{m_4, m_5, m_7\}$ then it is not possible to have the disorder d_1 without having eventually the symptoms m_4, m_5 and m_7.
- P is the possible effects function that associates to each disorder d_l a set $M_L \subseteq M$ of manifestations that d_l *may* cause.
- V associates to each disorder a set of (instantaneous) events. These events will be used to describe the evolution of the disorder. Among the events in $V(d_l)$ it must be included events that correspond to the beginning of all manifestations in $E(d_l)$. Furthermore, $V(d_l)$ can include events that correspond to the end of some of the manifestations in $E(d_l)$ and can also include other, non-observable events. For example, a common non-observable event in infectious diseases is the infection itself.
- $\mathcal{T}$ is a function that associates to *some* pairs of events $e_i, e_j \in V(d_l)$ a fuzzy temporal interval $\mathcal{T}(d_l)(e_i, e_j) = \pi$ which states that (according to the model for the disorder d_l) the elapsed time between the event represented by e_i and the event represented by e_j must be within the fuzzy temporal interval π (if e_i occurs before e_j then $\mathcal{T}(d_l)(e_i, e_j)$ is a positive fuzzy set).

We also define the derived function E, which yields the associated effects of a disorder, as $E(d) = N(d) \cup P(d)$.

Together, $V(d_l)$ and $\mathcal{T}(d_l)$ can be better understood in terms of a graph of events. The events in $V(d_l)$ are the nodes of the graph and if $\mathcal{T}(d_l)$ is defined for the pair of events (e_i, e_j) then there is a directed arc from e_j to e_i and the value in the arc is $\mathcal{T}(d_l)(e_i, e_j)$. We will call such interpretation as the temporal graph of the disorder.

Figure 3 is a simplified model of the temporal graph for poisoning by the *Amanita* mushroom. In this simplified model, we disregard the co-temporality of certain manifestations, and to facilitate the readability of the calculations, we use crisp fuzzy sets to model the temporal information in the example.

In Figure 3, the event e_0 is the ingestion of contaminated matter, m_1 is abdominal cramps, m_2 is nausea, m_3 is renal failure, and m_4 is hepatic failure. m_1^b and m_1^e refer to the beginning and end of the manifestation m_1. The intervals *around 6-24h*, *less than 24h* and *around 24-48h* are respectively modeled by $\langle 6, 24\rangle$, $\langle 0, 24\rangle$, $\langle 24, 48\rangle$. The interval *less than 1h* from Figure 1 (not represented in Figure 3) could be modeled by the set $\langle -1, 1\rangle$, which accommodates the double link between two manifestations.

2.3. Case information

For a particular diagnostic problem, one needs, besides the knowledge base about the disorders, a particular case. The case information should describe the manifestations that the patient is suffering and have suffered from, temporal information

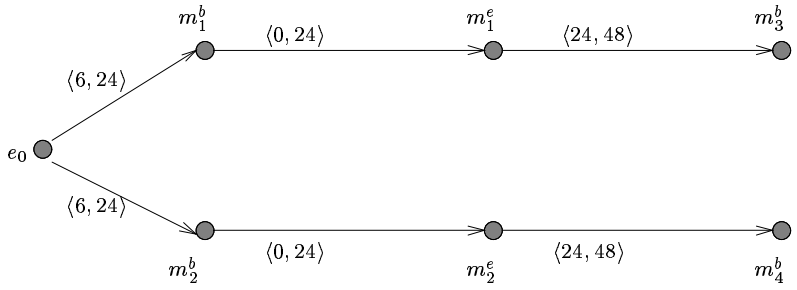

Figure 3. Temporal graph of the working example.

about when those symptoms started and ended, and information about manifestations that the diagnostician knows are not present in the patient.

Information about a given case is modeled by a tuple $Ca = \langle M^+, M^-, \text{EV}^+, \text{TIME}^+, \theta_0\rangle$, where

- M^+ is the set of manifestations known to be or to have been present in the case.
- M^- is the set of manifestations known to be absent from the case.
- EV^+ is a set of events for which one has temporal information. Among the events in EV^+ are the ones that represent the beginning of each manifestation in M^+. Events representing the end of the manifestations in M^+ may also belong to the set EV^+.
- TIME^+ is a function that associates to each event $e \in \text{EV}^+$ a fuzzy temporal interval that represents the possible moments in which that event happened.
- θ_0 is the moment of the diagnosis.

In our example, a piece of information such as "the patient had nausea (m_2), starting 24 hours before the consultation, which lasted for about 5 or 6 hours, and he is sure he did not have abdominal cramps (m_1)" would be modeled by $M^+ = \{m_2\}$, $M^- = \{m_1\}$ and $\text{EV}^+ = \{m_2^b, m_2^e\}$. If we consider that the consultation happened at time $\theta_0 = 120$, the temporal information could be translated as $\text{TIME}^+(m_2^b) = \langle 96, 96\rangle$ and $\text{TIME}^+(m_2^e) = \langle 101, 102\rangle$.

2.4. Minimal network

Finally we will make use of the concept of a minimal network [17]. Given a set of intervals among some events, the minimal network is way to compute the intervals between any two of those events, so that those computed intervals are as tight as possible.

The minimal network can be computed by the algorithm below, which is the Floyd-Warshall algorithm for all pairs shortest path [4, Chap. 26]. The minimal network will be computed for each disorder d_l. We assume that the events in $V(d_l)$ are arbitrarily numbered, and that $|V(d_l)| = n$. The algorithm computes the values t_{ij} with is the interval between events e_i and e_j in the minimal network for a particular disorder d_l.

1. **for** i = 1 **to** n **do**
2. **for** j = 1 **to** n **do**
3. **if** i = j **then** $t_{ii} = I_{=0}$
4. **else if** $\mathcal{T}(d_l)(e_i, e_j)$ is defined **then** $t_{ij} = \mathcal{T}(d_l)(e_i, e_j)$
5. **else if** $\mathcal{T}(d_l)(e_j, e_i)$ is defined **then** $t_{ij} = -\mathcal{T}(d_l)(e_j, e_i)$
6. **else** $t_{ij} = I_{\text{anytime}}$
7. **for** k = 1 **to** n **do**
8. **for** i = 1 **to** n **do**
9. **for** j = 1 **to** n **do**
10. $t_{ij} = t_{ij} \cap (t_{ik} \oplus t_{kj})$

We will define a function $\pi_l(e_i, e_j)$ which returns the value of t_{ij} in the minimal network for disorder d_l. If the disorder is clear from the context, we will not use the subscript l. In terms of the graph analogy of V and $\mathcal{T}$, the minimal network computes the transitive closure of the graph.

Figure 4 is part of the the minimal graph obtained for our example. The complete graph is such that for each two of events e_i and e_j there exist both (e_i, e_j) and (e_j, e_i); the value on a arc (e_i, e_j) not indicated in the figure is calculated as $\pi_l(e_i, e_j) = -\pi_l(e_j, e_i)$.

3. Degrees of consistency

3.1. Temporal consistency

In evaluating the temporal consistency between the case and the model, one needs to compare the elapsed time between the events in the case (the events in EV^+) and the corresponding fuzzy intervals as specified in the model.

We must compute the pairwise temporal distance between all events in EV^+. Given $e_i, e_j \in \text{EV}^+$ we compute

$$\text{DIST}^+(e_i, e_j) = \text{TIME}^+(e_j) \ominus \text{TIME}^+(e_i)$$

Thus $\text{DIST}^+(e_i, e_j)$ is the fuzzy temporal distance between the real occurrences of the events e_i and e_j. In order to verify how well these two events fit with the model of a particular disorder d_l we must compare $\text{DIST}^+(e_i, e_j)$ with $\pi_l(e_i, e_j)$ if both e_i and e_j belong to that disorder model.

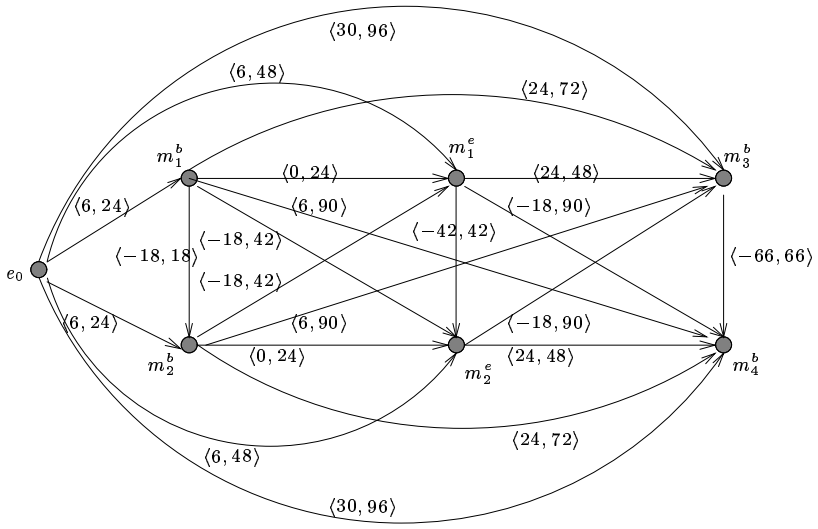

Figure 4. Part of the minimal graph of the working example with arcs labeled with $\pi_l(e_i, e_j)$.

In fact the degree of consistency of the pair of events e_i and e_j between the disorder model and the case information is the height of the intersection of $\text{DIST}^+(e_i, e_j)$ and $\pi_l(e_i, e_j)$.

Formally the temporal consistency degree of the disorder d_l is defined as:

$$\alpha(d_l) = \min_{e_i, e_j \in \text{EV}^+ \cap V(d_l)} [h(\text{DIST}^+(e_i, e_j) \cap \pi_l(e_i, e_j))]$$

In the remaining of the text, between two nodes e_i and e_j, only one arc shall to be taken into account in $\alpha(d_l)$, since $h(\text{DIST}^+(e_i, e_j) \cap \pi_l(e_i, e_j)) = h(\text{DIST}^+(e_j, e_i) \cap \pi_l(e_j, e_i))$.

EXAMPLE: Let us suppose that in our example we have $M^+ = \{m_1, m_2\}$, $\text{EV}^+ = \{m_1^e, m_2^e\}$, $\text{TIME}^+(m_1^e) = \langle 101, 102\rangle$ and $\text{TIME}^+(m_2^e) = \langle 57, 58\rangle$. Then we have $\text{DIST}^+(m_1^e, m_2^e) = \langle 43, 45\rangle$. From the minimal graph we have $\pi_l(m_1^e, m_2^e) = \langle -42, 42\rangle$ and consequently $\alpha(d_l) = h(\text{DIST}^+(m_1^e, m_2^e) \cap \pi_l(m_1^e, m_2^e)) = 0$. Therefore, the model and case are considered to be completely incompatible in temporal terms. Indeed, if we consider $\text{TIME}^+(m_1^e)$, the supposed ingestion of poisonous mushrooms would have occurred during the period $\langle 53, 96\rangle$, whereas considering $\text{TIME}^+(m_2^e)$, the ingestion would have occurred during the period $\langle 9, 52\rangle$. □

EXAMPLE: Let us suppose that $\text{TIME}^+(m_1^e) = \langle 100, 101, 102, 103\rangle$ and $\text{TIME}^+(m_2^e) = \langle 56, 57, 58, 59\rangle$. Then $\alpha(d_l) = h(\langle 41, 43, 45, 47\rangle \cap \langle -42, 42\rangle) = 0.5$, i.e. d_l would be considered a possible explanation to the manifestations in M^+. □

3.2. Categorical Consistency

Categorical consistency between model and case refers to the fact that a necessary manifestation of a disorder must happen, if the patient is suffering from that disorder. If the case does not have a manifestation then no disorder that considers that manifestation necessary can be a possible diagnostic, or be part of a possible diagnostic. But categorical inconsistency is tightly bound with temporal reasoning. In fact we can only state that a case is categorically inconsistent with the model if a necessary manifestation has not occurred and there has been enough time for it to happen.

One can say that a manifestation m_i has had enough time to occur in d_l if

- there exists an event e_j, which was supposed to happen after the start of m_i, and that event has already occurred;
- or there exists an event e_j, which was supposed to happen before the start of m_i, and that event did happen as expected, but the expected elapsed time between the event and the start of m_i has already expired.

Categorical consistency can be calculated as temporal consistency if we assume that all necessary manifestations that have not yet occurred will start sometime after the moment of consultation. If, because of either the two reasons above, there is other temporal information that states that this event should have already started, the temporal consistency index of the disorder will reflect it.

Thus, with the initialization

$$\forall m_i \in M^- \cap N(d_l), \mathrm{TIME}^+(m_i^b) = I_{\mathrm{afternow}}$$

the temporal consistency index $\alpha(d_l)$, will reflect both the temporal and the categorical consistency. We will call this combined temporal and categorical index as $\alpha_{ct}(d_l)$.

Example: For instance, let us suppose that in our example the consultation occurs at time $\theta_0 = 120$ and that we have $M^+ = \{m_2, m_4\}$, $M^- = \{m_1\}$, $\mathrm{EV}^+ = \{m_2^e, m_4^b\}$, with $\mathrm{TIME}^+(m_2^e) = \langle 91, 92\rangle$ and $\mathrm{TIME}^+(m_4^b) = \langle 114, 116\rangle$. Furthermore let us assume that the manifestation m_1 is necessary, that is, $m_1 \in N$.

Not considering the negative information about m_1, the temporal consistency would be $\alpha(d_l) = h(\mathrm{DIST}^+(m_2^e, m_4^b) \cap \pi_l(m_2^e, m_4^b)) = h(\langle 20, 25\rangle \cap \langle 24, 48\rangle) = 1$. In other words, without considering that m_1 did not yet happen, the model and case are considered to be completely compatible with each other.

Let us now take the information about m_1 into account. We make $\mathrm{TIME}^+(m_1^b) = I_{\mathrm{afternow}} = \langle 120, \infty\rangle$, and thus obtain

$$\alpha(d_l) \;=\; min \begin{pmatrix} h(\mathrm{DIST}^+(m_2^e, m_4^b) \cap \pi_l(m_2^e, m_4^b)), \\ h(\mathrm{DIST}^+(m_1^b, m_2^e) \cap \pi_l(m_1^b, m_2^e)), \\ h(\mathrm{DIST}^+(m_1^b, m_4^b) \cap \pi_l(m_1^b, m_4^b)) \end{pmatrix}$$

$$= \; min \begin{pmatrix} h(\langle 20, 25\rangle \cap \langle 24, 48\rangle), \\ h(\langle -\infty, -28\rangle \cap \langle -18, 42\rangle), \\ h(\langle -\infty, -4\rangle \cap \langle 6, 90\rangle) \end{pmatrix}$$
$$= \; min(1, 0, 0) = 0$$

Therefore, when we take into account the fact that m_1 has not yet occurred, the model and case are completely incompatible with each other. Indeed, reasoning backwards from the moment of occurrence of m_2^c, the moment of ingestion should have occurred in the interval $\langle -5, 62\rangle$, which would then make us expect m_1 to have already begun at some time inside the interval $\langle 1, 88\rangle$, i.e. at least 12 hours before the consultation. □

3.3. Intensity Consistency

In some diseases, it is important to quantify the intensity with which some of its manifestations occur. For instance, let us suppose a given disorder is characterized by strong fever at some time during its development; in this case, it is reasonable to suppose that that disorder will be the less plausible, the lower the temperature of the patient.

In order to provide information about the intensity of manifestations in relation to disorders, the knowledge base contains a function INT, which attributes to each node m of $E(d)$ a fuzzy set $\mathrm{INT}(m)$ describing the intensity with which that manifestation is expected to occur in d_l. Each fuzzy set $\mathrm{INT}(m)$ is defined on its particular domain $\Omega_{\mathrm{INT}(m)}$.

For manifestations m_i for which intensity is not relevant in d_l, $\mathrm{INT}(m_i)$ is constructed as $\forall x \in \Omega_{\mathrm{INT}(m)}$, $\mu_{\mathrm{INT}(m_i)}(x) = 1$. When the intensity can be quantified by a precise constant x^* in $\Omega_{\mathrm{INT}(m)}$, then we construct $\mathrm{INT}(m)$ as $\mu_{\mathrm{INT}(m)}(x) = 1$, if $x = x^*$, $\mu_{\mathrm{INT}(m)}(x) = 0$, otherwise.

In the same way, in order to provide information about the intensity of manifestations presented by a given patient, the case information contains a function INT^+, which attributes to each node $m_i \in M^+$ a fuzzy set $\mathrm{INT}^+(m_i)$ describing the intensity with which that manifestation occurred. Each fuzzy set $\mathrm{INT}^+(m_i)$ is defined on domain $\Omega_{\mathrm{INT}}(m_i)$.

The consistency of the intensity of a manifestation m_i, in relation to a disorder d_l, is quantified as follows:

$$\beta(m_i) = h(\mathrm{INT}(m_i) \cap \mathrm{INT}^+(m_i))$$

Finally, for a disorder d_l its intensity consistency is given by

$$\beta(d_l) = \inf_{m_i \in E(d_l)} \beta(m_i)$$

EXAMPLE: Let us suppose that manifestation m_1 in the *Amanita* poisoning disorder model is *severe* abdominal cramps. Let us also assume a domain for the intensity of cramps $\Omega_{\mathrm{INT}(m_1)} = [0, 10]$, and that severe cramps is modeled by

the fuzzy interval $\text{INT}(m_1) = \langle 7.5, 8.5, 10, 10\rangle$. If the patient claims to have *intense* cramps, modeled as the fuzzy interval $\text{INT}^+(m_1) = \langle 4, 5, 7, 8\rangle$, over the same domain $\Omega_{\text{INT}(m_1)}$, then the intensity consistency degree for cramps would be $\beta(m_1) = h(\langle 7.5, 8.5, 10, 10\rangle \cap \langle 4, 5, 7, 8\rangle = 0.25$. □

3.4. Diagnostic Explanation

In this paper we assume that every explanation, or better diagnostic explanation, is a single disorder that is temporal, categorical, and intensity consistent with the symptoms and explains all symptoms present in the case. Thus d_l is a diagnostic for the case $Ca = \langle M^+, M^-, \text{EV}^+, \text{TIME}^+, \theta_0, \text{INT}^+\rangle$, if

- $\alpha_{ct}(d_l) > 0$
- $\beta(d_l) > 0$
- for all $m_i \in M^+$ $m_i \in E(d_l)$.

4. Other informations about a disorder

Once one assumes that a particular disorder is the explanation for a set of symptoms, one can make further inferences about how this disorder will progress.

For the rest of this section we will assume that not only the disorder is an explanation for the set of symptoms, but also that it is 100% temporally consistent with it, that is, $\alpha(d_l) = 1$.

4.1. Revised elapsed time and revised event time

If the disease d_l is really what causes the events in EV^+, then it may be possible to reduce the uncertainties in the information about the elapsed time between some of the events in EV^+, since these elapsed times must also agree with the disorder model. We define the revised elapsed time between the events e_i and $e_j \in \text{EV}^+$, $\text{DIST}^r(e_i, e_j)$ as

$$\text{DIST}^r(e_i, e_j) = \text{DIST}^+(e_i, e_j) \cap \pi_l(e_i, e_j)$$

Given a particular event $\hat{e} \in \text{EV}^+$, which has the least uncertainty for its starting time, i.e. for which $\text{TIME}^+(\hat{e})$ is narrowest, we can anchor the other events on it. The revised event time for event $e \in \text{EV}^+$, based on $\hat{e}$ is defined as

$$\text{TIME}^r(e) = (\text{TIME}^+(\hat{e}) \oplus \text{DIST}^r(\hat{e}, e)) \cap \text{TIME}^+(e)$$

EXAMPLE: For instance, let us suppose that we have $M^+ = \{m_2, m_4\}$, $\text{EV}^+ = \{m_2^e, m_4^b\}$, with $\text{TIME}^+(m_2^e) = \langle 91, 92\rangle$ and $\text{TIME}^+(m_4^b) = \langle 114, 116\rangle$. Then we obtain $\text{DIST}^r(m_2^e, m_4^b) = \langle 22, 25\rangle \cap \langle 24, 48\rangle = \langle 24, 25\rangle$. We then use $\text{TIME}^+(m_2^e)$

(the narrowest time estimation) to anchor the estimated time of occurrence of $\text{TIME}^+(m_4^b)$ and obtain

$$\begin{aligned}\text{TIME}^r(m_4^b) &= (\text{DIST}^r(m_2^e, m_4^b) \oplus \text{TIME}^+(m_2^e)) \cap \text{TIME}^+(m_4^b) \\ &= (\langle 24, 25\rangle \oplus \langle 91, 92\rangle) \cap \langle 114, 116\rangle \\ &= \langle 115, 117\rangle \cap \langle 114, 116\rangle = \langle 115, 116\rangle\end{aligned}$$

□

4.2. Speed of development of a disorder

In the medical domain, the temporal constraints in the disorder model usually represent the accumulated knowledge of these intervals for a very large number of cases. As a consequence, considering d_l as the actual disorder, it is very likely that the revised interval $\text{DIST}^r(e_i, e_j)$ will be much more precise than the estimated interval $\pi_l(e_i, e_j)$. Using this information and considering the hypothesis of a regularity on the speed of the development of a disorder, one may make more precise predictions about events in the future, or unknown events in the past.

For example, let us suppose we have a case in which for all events $e_i, e_j \in \text{EV}^+$, $\text{DIST}^r(e_i, e_j)$ is always within the first fourth of the corresponding interval $\pi_l(e_i, e_j)$ of a disease d_l. Then if the disease model forecasts that a manifestation m will start anytime between the next 5 to 15 days, given the past history of how the disease is developing in this case, one can make the expect that, for this case, one can expect m to start within the next 5 and 7.5 days.

We will formally define the idea of "$\text{DIST}^r(e_i, e_j)$ is always within the first fourth of the corresponding interval $\pi_l(e_i, e_j)$" by defining a compression factor $\gamma(e_i, e_j)$ between $\pi_l(e_i, e_j)$ and $\text{DIST}^r(e_i, e_j)$. From now on, we will assume that the intervals are trapezoidal.

Given two events e_i and e_j, with $\pi_l(e_i, e_j) = \langle a, b, c, d\rangle$ and $\text{DIST}^r(e_i, e_j) = \langle a', b', c', d'\rangle$, we can compute the compression factor $\gamma(e_i, e_j)$ as a pair of pairs of coefficients:

$$\gamma(e_i, e_j) = \langle\langle \tau_1, \tau_2\rangle, \langle \tau_3, \tau_4\rangle\rangle = \langle\langle \frac{b'-b}{c-b}, \frac{c'-b}{c-b}\rangle, \langle \frac{a'-a}{d-a}, \frac{d'-a}{d-a}\rangle\rangle$$

for $a \neq d, b \neq c$. For $b = c$ we make $\langle \tau_1, \tau_2\rangle = \langle 0, 1\rangle$ and for $a = d$ we make $\langle\langle \tau_1, \tau_2\rangle, \langle \tau_3, \tau_4\rangle\rangle = \langle\langle 0, 1\rangle\langle 0, 1\rangle\rangle$. These coefficients are such that we always have $\tau_1 \leq \tau_2$ and $\tau_3 \leq \tau_4$.

The first pair of coefficients refer to the core of the intervals. $\frac{b'-b}{c-b}$ measures where within the core of $\pi_l(e_i, e_j)$ does the core of $\text{DIST}^r(e_i, e_j)$ starts. The second coefficient of the first pair $\frac{c'-b}{c-b}$ measures where within the core of $\pi_l(e_i, e_j)$ does the core of $\text{DIST}^r(e_i, e_j)$ ends. Thus if a case is developing within the first fourth of the (core) interval predicted by the model, the first pair of coefficients in the corresponding γ would be $\langle 0, 0.25\rangle$. That is the core of the case's interval starts exactly at the start of the core of predicted interval, and ends at 25% of the predicted interval.

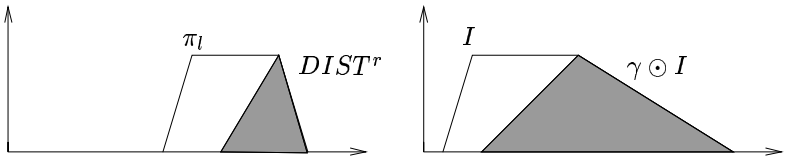

Figure 5. Speed of development of a manifestation.

The second pair of coefficients perform the same calculations, but in relation to the support of $\pi_l(e_i, e_j)$ in comparison to the support of $\text{DIST}^r(e_i, e_j)$.

The compression factor is way to transform a wider interval into a narrower one. If I is the interval $\langle a, b, c, d\rangle$ and γ is a compression factor $\langle\langle\tau_1, \tau_2\rangle, \langle\tau_3, \tau_4\rangle\rangle$, will define the composition operator $\odot$ as

$$\gamma \odot I = \langle a', b', c', d'\rangle$$

where

$$\begin{aligned} a' &= \min(a + \tau_3(d-a), b + \tau_1(c-b)), \\ b' &= b + \tau_1(c-b), \\ c' &= b + \tau_2(c-b), \\ d' &= \max(a + \tau_4(d-a), b + \tau_2(c-b)) \end{aligned}$$

The max and min operations are there to guarantee that the resulting interval will be trapezoidal. This operator is such that $\gamma \odot \pi_l(e_i, e_j) = \text{DIST}^r(e_i, e_j)$. The compression factor $\gamma^* = \langle\langle 0, 1\rangle\langle 0, 1\rangle\rangle$ is the neutral compression factor, i.e. for any interval I we have $\gamma^* \odot I = I$.

EXAMPLE: For instance, let us suppose we have $\pi_l(e_i, e_j) = \langle 10, 12, 16, 18\rangle$ and $\text{DIST}^r(e_i, e_j) = \langle 14, 16, 16, 18\rangle$. From $\pi_l(e_i, e_j)$ and $\text{DIST}^r(e_i, e_j)$ we obtain the compression factor $\gamma = \langle\langle 1, 1\rangle, \langle 1/2, 1\rangle\rangle$, which applied to interval $I = \langle 2, 6, 10, 20\rangle$ yields the compressed interval $\gamma \odot I = \langle 7, 10, 10, 20\rangle$ (see Figure 5). □

Two compression factors can be combined using the *cautions sum*. Given $\gamma_1 = \langle\langle x_1, x_2\rangle, \langle x_3, x_4\rangle\rangle$ and $\gamma_2 = \langle\langle y_1, y_2\rangle, \langle y_3, y_4\rangle\rangle$ we define the cautious sum $\gamma_1 \hat{+} \gamma_2$ as

$$\gamma_1 \hat{+} \gamma_2 = \langle\langle \min(x_1, y_1), \max(x_2, y_2)\rangle, \langle \min(x_3, y_3), \max(x_4, y_4)\rangle\rangle$$

The cautious sum creates a compression factor that combines its two arguments. For any interval I, $\gamma = \gamma_1 \hat{+} \gamma_2$ is such that $\forall x\ \mu_{\gamma \odot I}(x) \geq \mu_{\gamma_1 \odot I}$ and $\forall x\ \mu_{\gamma \odot I}(x) \geq \mu_{\gamma_2 \odot I}$. Moreover, for any compression factor γ' that satisfies these two properties we have $\forall x\ \mu_{\gamma' \odot I}(x) \geq \mu_{\gamma \odot I}(x)$. In other words, the cautious sum computes the best γ such that, if composed with I, will create an interval wider (but hopefully not too much wider) than both $\gamma_1 \odot I$ and $\gamma_2 \odot I$. Note that $\gamma_1 \hat{+} \gamma_2 = \gamma_1 \hat{+} \gamma_2$ and $\gamma_1 \hat{+} \gamma_1 = \gamma_1$, i.e. $\hat{+}$ is commutative and idempotent.

The compression factor $\gamma(d_l)$ for the disorder d_l is the cautious sum of all $\gamma(e_i, e_j)$ for all events $e_i, e_j \in \mathrm{EV}^+$:

$$\gamma(d_l) = \hat{\dagger}_{e_i,e_j \in \mathrm{EV}^+, h(\mathrm{DIST}^r(e_i,e_j) \cap I_{\geq 0})=1} \gamma(e_i, e_j)$$

This global compression factor is such that for all $e_i, e_j \in \mathrm{EV}^+$, for which the temporal interval between them is possibly positive, $\gamma(d_l) \odot \pi_l(e_i, e_j)$ is wider than, or includes, the corresponding $\mathrm{DIST}^r(e_i, e_j)$.

From $\gamma(d_l)$ one can obtain qualitative information that is important for a physician, for example. If $\gamma(d_l) = \langle\langle \tau_1, \tau_2\rangle, \langle \tau_3, \tau_4\rangle\rangle$ and $\max(\tau_1, \tau_2, \tau_3, \tau_4) \ll 0.5$ then it is clear that, given the information available the disorder is progressing much faster than "the average." In a medical situation, the disorder is acute. Similarly if $\min(\tau_1, \tau_2, \tau_3, \tau_4) \gg 0.5$ then the disorder is progressing much slower that "the average," or in medical terms, it is chronic. Here "average" is not to be understood as "typical" (across many cases) but as the central value of the interval.

4.3. Information about unknown manifestations

Let us now discuss how can we predict when unreported events may happen or should have happened. Predictions of the time of occurrence of past events are important to direct the physician to make pertinent questions to the patient, making it eventually possible to rule out a disorder upon the patient's answer. On the other hand, predictions of the time of occurrence of future events are important not only to eventually rule out disorders as time goes by, but also of allowing the physician to take sound preventive actions.

We distinguish two kinds of predictions, a sure one and a likely one. We will define the sure forecast time ($\mathrm{TIME}^s(e)$) for all events e in $V(d_l) - \mathrm{EV}^+$ anchored on an event $\hat{e} \in \mathrm{EV}^+$, as

$$\mathrm{TIME}^s(e) = \mathrm{TIME}^+(\hat{e}) \oplus \pi_l(\hat{e}, e)$$

The sure prediction is safe to take but may produce very imprecise results. With the information given by $\gamma(d_l)$, one may be able to make more precise forecast about unknown events, either in the past or in the future. We will define the likely forecast time ($\mathrm{TIME}^f(e)$) for all events e in $V(d_l) - \mathrm{EV}^+$ anchored on an event $\hat{e} \in \mathrm{EV}^+$, as

$$\mathrm{TIME}^f(e) = \mathrm{TIME}^+(\hat{e}) \oplus (\gamma(d_l) \odot \pi_l(\hat{e}, e)) \text{ for all } e \in V(d_l) - \mathrm{EV}^+$$

$\mathrm{TIME}^f(e)$ is the likely interval in which event e happened or will happen. Likely, because it uses the information of the case to make more precise predictions.

EXAMPLE: Let $M^+ = \{m_2\}$, $\mathrm{EV}^+ = \{m_2^b, m_2^e\}$, $\mathrm{TIME}^+(m_2^b) = \langle 95, 96, 96, 97\rangle$ and $\mathrm{TIME}^+(m_2^e) = \langle 100, 101, 102, 103\rangle$. Using m_2^b as anchor we would obtain, for example:

$$\begin{aligned}
\text{TIME}^s(m_1^b) &= \text{TIME}^+(m_2^b) \oplus \pi_l(m_2^b, m_1^b) \\
&= \langle 95, 96, 96, 97\rangle \oplus \langle -18, 18\rangle \\
&= \langle 77, 78, 114, 115\rangle \\
\text{TIME}^f(m_1^b) &= \text{TIME}^+(m_2^b) \oplus (\gamma(d_l) \odot \pi_l(m_2^b, m_1^b)) \\
&= \langle 95, 96, 96, 97\rangle \oplus (\langle\langle 5/24, 1/4\rangle\langle 1/8, 1/3\rangle\rangle \odot \langle -24, -6\rangle \\
&= \langle 95, 96, 96, 97\rangle \oplus \langle -13.5, -10.5, -9, -6\rangle \\
&= \langle 81.5, 85.5, 87, 91\rangle
\end{aligned}$$

□

It may be important to determine which necessary manifestations should have already occurred, in order to test for them. In the set-covering tradition of modeling diagnosis there are usually too many explanations for a set of symptoms, and since there is no concept of probabilities, it is not possible to select a more "probable" diagnostic. Thus, it may be important to perform tests that would allow one to remove a diagnostic from the set of hypotheses. That can only be accomplished by verifying that a necessary manifestation for that diagnostic is not present.

The necessary manifestations that must have already occurred are computed by:

$$\begin{aligned}
M^s_{\text{necpast}} = \{m_i |\ & m_i \in N(d_l) - M^+ \text{ and } h(\text{TIME}^s(m_i^b) \cap I_{\text{beforenow}}) > 0 \\
& \text{and } h(\text{TIME}^s(m_i^b) \cap I_{\text{afternow}}) = 0\}
\end{aligned}$$

If $h(\text{TIME}^s(m_i^b) \cap I_{\text{beforenow}}) > 0$, then $\text{TIME}^s(m_i^b)$ has some intersection with the past, so there are some times in the past in which m_i should have started. Furthermore if $h(\text{TIME}^s(m_i^b) \cap I_{\text{afternow}}) = 0$ then the interval $\text{TIME}^s(m_i^b)$ has no intersection with the future, and thus must be totally contained in the past.

If a manifestation in M^s_{necpast} is found not to be present in the case, the disorder d_l can be disregarded as categorically inconsistent with the new set of manifestations known to be present and absent from the case.

But if one is willing to use the more precise, but less sure forecast time, the set of necessary manifestations that are likely have already occurred are computed by:

$$\begin{aligned}
M^f_{\text{necpast}} = \{m_i |\ & m_i \in N(d_l) - M^+ \text{ and } h(\text{TIME}^f(m_i^b) \cap I_{\text{beforenow}}) > 0 \\
& \text{and } h(\text{TIME}^f(m_i^b) \cap I_{\text{afternow}}) = 0\}
\end{aligned}$$

Of course, if a manifestation in M^f_{necpast} is found to be absent from the case, one cannot make a categorical claim that the disorder is categorically inconsistent with the case. At most one can make the claim that the disorder and the case are likely to be inconsistent. We are not yet able to define a measure for this "likeliness".

EXAMPLE: Using the case information from the example above, assuming that $\theta_0 = 110$, and assuming that in the *Amanita* poisoning model $N = \{m_1\}$, we would have:

$$\begin{aligned}
M^s_{\text{necpast}} &= \{\} \\
M^f_{\text{necpast}} &= \{m_1\}
\end{aligned}$$

$M^s_{\text{necpast}} = \{\}$ because although $h(\text{TIME}^s(m^b_1) \cap I_{\leq 110}) = h(\langle 77, 78, 114, 115\rangle \cap I_{\leq 110}) = 1$, $h(\text{TIME}^s(m^b_1) \cap I_{\geq 110}) = h(\langle 77, 78, 114, 115\rangle \cap I_{\geq 110}) = 1 \neq 0$. On the other hand, $h(\text{TIME}^f(m^b_1) \cap I_{\leq 110}) = h(\langle 81.5, 85.5, 87, 91\rangle \cap I_{\leq 110}) = 1$ and $h(\text{TIME}^s(m^b_1) \cap I_{\geq 110}) = h(\langle 81.5, 85.5, 87, 91\rangle \cap I_{\geq 110}) = 0$. □

In other situations, it may be important to gather evidence in favor of a particular diagnostic. For that one needs to know which manifestations are likely to have already occurred, in order to test for them. The set of such manifestations is described as:

$$M^f_{\text{posspast}} = \{m_i | m_i \in E(d_l) - M^+ \text{ and } h(\text{TIME}^f(m^b_i) \cap I_{\text{beforenow}}) > 0\}$$

Again, one could define a similar set using the sure measure TIME^s. Note that $M^f_{\text{necpast}} \subseteq M^f_{\text{posspast}}$ and $M^s_{\text{necpast}} \subseteq M^s_{\text{posspast}}$.

5. Conclusions and future work

This work presented a model to include fuzzy temporal information, categorical information, and (fuzzy) intensity information within a diagnostic framework. We provided answers to the following questions: when is the temporal information in the case consistent with a disorder model, when is the case categorically consistent with the model, and how information about intensity can be included. Furthermore, we showed how to make forecasts about future manifestations, past manifestations that have not been tested for, and so on, based not only on what the disorder model predicts but also based on how fast the case is progressing.

In this paper we are not concerned on how the temporal information about the disorder model ($\mathcal{T}(d_l)$) is obtained. It is clear that in the disorder model, expressions like "a period of 1 to 2 days of no symptoms," as in the *Amanita* mushroom poisoning model, refer to the core of the interval. [1] proposes a method of automatically computing the support given information about the core of the interval.

Also, we are not concerned on how the temporal information about the case (TIME^+) is obtained. This is a more complex issue. Information about a single event in EV^+ can come from many sources, and not all of them will fully agree with each other. For example, a patient states that a symptom started more than two weeks ago, that it lasted for about a week, and that he is sure that the symptoms were over last Sunday, and that he remembers that the symptoms had not started yet on the 14th when he arrived from a trip. There are only two events of relevance in the statement, the beginning and the end of the symptom, but there are many intervals relating to each other and also to the time of consultation. All this information may not even be consistent; if the consultation is being held on the 20th, then there are conflicting information about when the symptom started: "after the 14th" and "more than two weeks ago", or before the 7th. [17] discusses this problem, and proposes a method to evaluate the consistency of the information and the most precise intervals for the occurrence of the events using minimal networks. Other researchers have also discussed similar issues [6].

It should also be pointed out that this paper provides little insight on how to use the various measures of consistency in a practical diagnostic situation. If the temporal/categorical and intensity degrees of consistency for a disorder are not all 1, and thus the case is not fully consistent with the model, how should those indices be used: should they be combined into a single index, which degree is more important than others, can one or the aggregation of these indices be used to classify competing diagnostic hypothesis. These questions and others will be addressed in an empirical, future step of this work.

The approach presented here yields only possibilistic compatibility degrees, but could be modified to obtain also entailment degrees, as in [17, 18]. Entailment degrees such as necessity measures, although considered to be too restrictive by the authors, may be useful to distinguish between two diagnostics that have the same possibilistic degree of compatibility in a given case.

This work extends a previous work by the authors [20]. In [20] it is assumed that the temporal graph of the disorder is a tree, and it contained only events corresponding to the beginning of the manifestations. In order to calculate the temporal consistency of the case and the model, the case information was propagated towards the root; any temporal inconsistency would result in conflicting intervals for the root.

In the future, we intend to exploit the possibility of having a set of disorders explaining the manifestations presented by the patient, rather than a single disorder, as addressed here. This issue is particularly interesting when a set of disorders have manifestations in common. In this case, the problem is that the indices defined above depend on which manifestations are attributed to each disorder, or in other words, which disorder is causing which manifestation. For example, a manifestation m_4 may be the reason the temporal consistency index of disease d_1 (in which m_4 is a possible manifestation) is very low, but if one could state that m_4 "is being caused" by a different disorder, say d_2, then the indices for d_1 would be better, as would the indices for explanation $\{d_1, d_2\}$. The consequences of this trading of inconsistency by new disorders, both in terms of the combinatorial explosion, and the applicability of the results in real diagnostic domains are not yet clear.

In the case that all disorders in an explanation do not have any common manifestations, it seems that the theory above could be generalized by calculating the consistency indices for each disorder and attributing global consistency indices to the explanation as the minimum of the disorder's indices.

We are also interested in allowing for "fuzzy" categorical information, making it possible to model pieces of information furnished by a medical expert such as "in disorder d_l, manifestation m_i is very likely to occur" or "in disorder d_l, m_i will seldom occur" In this case, N, P, will be substituted by possibility distributions. And finally we are interested in modeling uncertainty on whether a manifestation has occurred or not (which is different that modeling the imprecision in its intensity, as we did in this paper).

Acknowledgments

The authors would like to thank the reviewers for their helpful comments. The second author was partially funded by CNPq.

References

1. S. Barro, R. Marin, J. Mira, and A.R. Paton. A model and a language for the fuzzy representation and handling of time. *FSS*, 61:153–175, 1994.
2. B.G. Buchanan and E.H. Shortliffe. *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Addison-Wesley Publishing Company, 1984.
3. L. Console and P. Torasso. Temporal constraint satisfaction on causal models. *Information Sciences*, 68:1–32, 1993.
4. T. H. Cormen, C. E. Leiserson, and R. L. Rivest. *Introduction to Algorithms*. MIT Press, 1990.
5. D. Dubois and H. Prade. *Possibility Theory: an approach to computerized processing of uncertainty*. Plenum Press, 1988.
6. D. Dubois and H. Prade. Processing fuzzy temporal knowledge. *IEEE Trans. on S.M.C.*, 19(4), 1989.
7. D. Dubois and H. Prade. Fuzzy relation equations and causal reasoning. *Fuzzy Sets and Systems*, pages 119–134, 1995.
8. I. Hamlet and J. Hunter. A representation of time for medical expert systems. In J. Fox, M. Fieschi, and R. Engelbrecht, editors, *Lecture Notes in Med. Informatics*, volume 33, pages 112–119. Springer-Verlag, 1987.
9. K. Konolige. Adbuction versus clusure in causal theories. *Artificial Intelligence*, 53:255–272, 1992.
10. P. A. Koton. *Using Experience in Learning and Problem Solving*. PhD thesis, MIT Laboratory of Computer Science, MIT/LCS/TR441, 1988.
11. L.Console, D.T. Dupré, and P. Torasso. A theory of diagnosis for incomplete causal models. In *Proceedings of the 10th IJCAI*, pages 1311–1317, 1989.
12. W. Long. Reasoning about state from causation and time in a medical domain. In *Proc. of the AAAI 83*, 1983.
13. P. Lucas. Modelling interactions for diagnosis. In *Proceedings of CESA'96 IMACS Multiconference: Symposium on Modelling, Analysis and Simulation*, volume 1, pages 541–546, 1996.
14. G. L. Mandell, R. G. Douglas, and J. E. Bennett, editors. *Principles and practice of infectious diseases*. Churchill Livingstone, 4rd edition, 1995.
15. Y. Peng and J. A. Reggia. *Abductive Inference Models for Diagnostic Problem-Solving*. Springer-Verlag, 1990.
16. R. Reiter. A theory of diagnosis from first principles. *Artificial Intelligence*, 32(1):57–95, April 1987.
17. L. Vila and L. Godo. On fuzzy temporal constraint networks. *Mathware and Soft computing*, 3:315–334, 1994.
18. L. Vila and L. Godo. Possibilistic temporal reasoning on fuzzy temporal constraints. In C. S. Mellish, editor, *Proc. IJCAI'95*, volume 2, pages 1916–1922, 1995.
19. J. Wainer and A. Rezende. A temporal extension to the parsimonious covering theory. *Artificial Intelligence in Medicine*, 10:235–255, 1997.
20. J. Wainer and S. Sandri. A fuzzy temporal/categorical extension to the parsimonious covering theory. In *The Seventh Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU'98)*, Paris, 1998. To be published.