Application of Uncertain Temporal Relations Algebra to Diagnostic Problems

VLADIMIR RYABOV¹, VAGAN TERZIYAN²
¹Department of Computer Science and Information Systems,
²Department of Mathematical Information Technology,
University of Jyväskylä,
P.O.Box 35, 40351, Jyväskylä,
FINLAND
{vlad,vagan}@it.jyu.fi

Abstract: - Temporal representation and reasoning has been applied so far to many areas of AI, including medical applications. Identification of temporal patterns plays an important role in medical diagnostics. Temporal scenarios of different diseases, therapy protocols, and other temporal graphs provide important additional information for medical decisions. In this paper we propose an application of algebra of uncertain temporal relations to diagnostic problems. We represent uncertain temporal relations within a medical scenario graph using the probabilities of the basic relations that can hold between two temporal primitives. Also in the paper we show how: (a) to generate temporal scenarios by integrating appropriate relational networks from already diagnosed cases; (b) to classify a new case using the measure of the distance between network and scenario.

Key-Words: - Temporal Diagnostics, Temporal Scenario, Temporal Relation, Uncertainty, Probability

1 Introduction
Reasoning with time-oriented data is central to the practice of medicine. Monitoring clinical variables over time often provides information that drives medical decision-making (e.g., clinical diagnosis and therapy planning). In some medical diagnostic applications, e.g. [3] and [6], temporal information about the occurrence of symptoms is vital for correct diagnostics and some medical expert systems.

A brief overview of research efforts in designing and developing time-oriented systems in medicine during the past decade is presented in [2]. In this overview the two main future research directions are emphasized: temporal reasoning supporting various temporal inference tasks, and temporal data maintenance dealing with storage and retrieval of data with heterogeneous temporal dimensions.

The temporal aspect is especially important in automated decision support to patient care over significant time periods. The crucial role of temporal-reasoning and maintenance tasks for modern medical information and decision support systems was also underlined in [11]. While temporal reasoning involves mainly intelligent analysis of time-oriented clinical data, temporal maintenance focuses more on effective data storage and retrieval.

A reasoning model, presented in [5], is based on two basic cognitive mechanisms: aggregation of similar observed situations and forgetting of non-relevant information. Managing clinical therapy by refinement of domain-independent temporal entities and inferences was also shown in that paper.

In planning a long-term therapy it is important to analyze whether a new drug or drug combination is able to improve the prognosis [8]. Appropriate therapy protocols often randomize patients of a particular risk group and assign them to different, competitive therapy branches. A knowledge-based system in this domain has to deal with different types of temporal constraints and relationships such as exact and inexact quantitative and qualitative relations. The therapy is based not only on the drugs themselves and their dosages, but also on the complex and sophisticated temporal interaction structure of these drugs. Therefore, the special attention has to be paid to the adequate representation and processing of temporal data.

One way to combine fuzzy temporal reasoning within diagnostic reasoning was proposed in [12]. 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 time when manifestations started and ended are also modeled using fuzzy sets.
A framework for model-based diagnosis of dynamic systems by using and expressing temporal uncertainty in the form of qualitative Allen's interval relations is described in [9]. That approach is based on a logical framework extended by qualitative and quantitative temporal constraints. It was also shown there how to describe behavioral models, how to use abstract observations and how to compute abstract temporal diagnoses.

In this paper we also use Allen's interval algebra [1] to represent temporal uncertainty in abstract temporal diagnosis applications. Uncertain temporal relations are represented within a medical scenario graph using probabilities of the basic relations that can hold between two temporal primitives. In this paper we also propose:

(1) to generate temporal scenarios as temporal graphs with uncertain temporal relations of some diseases by integrating appropriate relational networks from already diagnosed cases;
(2) to classify a new case using the special measure of the distance between network and scenario.

The following text is organized as follows. In Section 2 we introduce the basic concepts used throughout the paper. Reasoning operations are discussed in Section 3. In Section 4 we show how to generate uncertain temporal scenario combining a number of networks of temporal relations. In Section 5 we show how to compare a relational network with known scenarios using the special measure of the distance between network and scenario. Finally, Section 6 presents conclusions.

2 Basic concepts

In this section we will consider the notation used in the paper, the representation of uncertain relations, and the notion of the distance between two relations.

Let us denote temporal points with small non-bold letters, i.e. a, b, and temporal intervals with capital non-bold letters, i.e. A, B. Let us denote a relation between two temporal points with a small bold letter r with subscript including the primitives, i.e. \( r_{a,b} \) is a temporal relation between points a and b. The relation between two intervals is denoted with a capital letter R. There are three basic temporal relations that can hold between two points: “before” (\(<\) ), “at the same time” (\(=\) ), and “after” (\(>\) ). Let us define a set of these relations as \( A = \{<,=,>\} \).

We will refer to an element of this set as \( \alpha \in A \).

There are thirteen Allen’s interval relations [1] \( X = \{eq,bi,di,oi,mi,si,fi,fi\} \). We will refer to an element of this set as \( \chi \in X \).

Ryabov and Puuronen in [10] proposed to represent an uncertain relation between two temporal primitives as a set of probabilities of all basic relations that can hold between these primitives. For example, \( r_{a,b}\{e_{\alpha} \ (\alpha \in A)\} \) is the uncertain relation between temporal points a and b, including the probabilities \( e_{a,b} <, e_{a,b} =, \) and \( e_{a,b} > \). The probability of a basic temporal relation between two primitives is further denoted using letter “e” with a superscript indicating the basic relation and a subscript indicating the temporal primitives.

An uncertain relation \( R_{a,b}\{e_{\alpha} \ (\alpha \in X)\} \) between intervals A and B includes thirteen probabilities of Allen’s relations. The sum of all probability values of the basic relations within r or \( R \) is equal to 1. When \( \exists e_{\alpha} \ (\alpha \in X) \) we call such \( R_{a,b} \) a totally certain relation (TCR). Allen’s interval relations are the examples of TCRs. When all the probability values within r or \( R \) are equal we call such relation a totally uncertain relation (TUR).

Let us further assume that the relations \( r_{(a,b)} \{e_{\alpha} \ (\alpha \in A)\} \) and \( r_{(a,b)} \{e_{\alpha} \ (\alpha \in A)\} \) are equal if and only if \( e_{a,b}^{<} = e_{(a,b)}^{<} \), \( e_{a,b}^{=} = e_{(a,b)}^{=} \), and \( e_{a,b}^{>} = e_{(a,b)}^{>} \). Otherwise, relations \( r_{(a,b)} \) and \( r_{(a,b)} \) are unequal. In a similar way we will reason about interval relations.

The domain probability values are the probabilities of the basic relations between two primitives in the situation, when we know nothing about the relation between them in the given domain area. The sum of these probability values is equal to 1, and they are distinguished with subscript including letter D, i.e., \( e_{a,b}^D \), \( e_{a,b}^D \), and \( e_{a,b}^D \).

The distance (denoted as \( d \)) between two uncertain temporal relations is a variable belonging to the interval [0,1]. When \( d=0 \) the uncertain relations compared are equal, and when \( d=1 \) the relations are totally different. The examples of totally different relations are the point relations “<” and “>”. One approach to estimate the value of \( d \) for uncertain relations between temporal points was proposed in [7]. In its physical interpretation the approach is based on the assumption that the two relations to be compared are distributed on the virtual lath, and where the basic relations within the uncertain ones are assumed to be physical objects. We extend that approach to be able to estimate the distance between the interval relations.

Let us consider as an example two uncertain interval relations \( R_{a,b} \) and \( R_{C,D} \) (Figure 1).
The probabilities of Allen’s relations are represented as rectangles in Figure 1 with gray fill for $R_{A,B}$ and with white fill for $R_{C,D}$. For every relation we find out the balance point, which in physical interpretation is a moment of mass for the physical objects distributed on the lath. We assume that the distance between two neighbor objects on the lath is equal for all neighbor pairs. The module of the mathematical difference between the values of the balance points for these two relations is the value of the distance between these relations.

The virtual lath from Figure 1 is being marked as a closed interval $[0,1]$ with weights for the interval relations defined as $w^{0}=0$, $w^{0.083}=0.5$, $w^{0.167}=1$. The value of the balance point, i.e., for $R_{A,B}$, is calculated by formula (1):

$$\text{Bal}(R_{A,B}) = \frac{1}{12} \sum_{i=0}^{12} i \times e^{x}_{A,B} = \frac{1}{12} \sum_{i=0}^{12} i \times e^{x}_{A,B}.$$  \hspace{1cm} (1)

The sum in the upper part of the indicated division in formula (1) includes the weights of the relations, obtained by multiplication of the constant $1/12$ on the variable $i = 0, 12$, correspondingly multiplied on the particular probability values of Allen’s relations from $R_{A,B}$. The lower part of the indicated division in formula (1) equals to 1 since it represents the sum of the probabilities of all Allen’s relations within $R_{A,B}$.

The distance between the uncertain relations $R_{A,B}$ and $R_{C,D}$ is calculated by formula (2):

$$d_{R_{A,B}, R_{C,D}} = |\text{Bal}(R_{A,B}) - \text{Bal}(R_{C,D})|.$$  \hspace{1cm} (2)

In a similar way, we can derive the formula for the distance between uncertain point relations [7].

### 3 Reasoning mechanism

In this section we briefly overview the reasoning mechanism including inversion, composition, and addition operations, proposed in [10]. The definitions for inversion and addition are presented using the notation for interval relations, and except for this, there is no difference between them and the corresponding definitions for point relations.

The operation of inversion ($\sim$) derives the relation $R_{B,A}$ when the relation $R_{A,B}$ is known.

**Definition** (Unary operation of inversion for interval relations). Let us suppose that the uncertain relations between temporal intervals $A$ and $B$ and between $B$ and $A$ are defined as $R_{A,B} \{ \chi \in X \}$ and $R_{B,A} \{ \chi \in X \}$ correspondingly, the probability values $e^{X}$ within $R_{A,B}$ are known, and $R_{B,A} = \sim R_{A,B}$. In this case, the probability values $e^{X}_{B,A}$ are calculated according to the inversion table for Allen’s interval relations [1]. For example, $A$ “overlaps” $B$ with the probability 0.1 then $B$ is “overlapped-by” $A$ with the same probability, i.e. $e^{X}_{B,A} = e^{X}_{A,B}$.

The operation of composition ($\otimes$) derives the relation $r_{a,c}$ when there exist relations $r_{a,b}$ and $r_{b,c}$.

**Definition** (Binary operation of composition for point relations). Let us suppose that the uncertain relations between temporal points $a$ and $b$, $b$ and $c$, and between $a$ and $c$, are defined as $r_{a,b} \{ e_{a,b} | \alpha \in A \}$, $r_{b,c} \{ e_{b,c} | \alpha \in A \}$, and $r_{a,c} \{ e_{a,c} | \alpha \in A \}$ correspondingly, the probability values within $r_{a,b}$ and $r_{b,c}$ are known, and $r_{a,c} = r_{a,b} \otimes r_{b,c}$. In this case, the probability values $e_{a,c}$ are calculated using the formulas (3)-(5):

$$e^{X}_{a,c} = e_{a,b}^{X} e_{b,c}^{X} + e_{a,b}^{X} e_{b,c}^{X} + e_{a,b}^{X} e_{b,c}^{X} + e_{a,b}^{X} e_{b,c}^{X}.$$  \hspace{1cm} (3)

$$e^{X}_{a,c} = e_{a,b}^{X} e_{b,c}^{X} + e_{a,b}^{X} e_{b,c}^{X} + e_{a,b}^{X} e_{b,c}^{X} + e_{a,b}^{X} e_{b,c}^{X}.$$  \hspace{1cm} (4)

$$e^{X}_{a,c} = e_{a,b}^{X} e_{b,c}^{X} + e_{a,b}^{X} e_{b,c}^{X} + e_{a,b}^{X} e_{b,c}^{X} + e_{a,b}^{X} e_{b,c}^{X}.$$  \hspace{1cm} (5)

**Definition** (Binary operation of composition for interval relations). Let us suppose that uncertain
relations between temporal intervals \(A\) and \(B\) and \(C\), and \(A\) and \(C\) are defined as \(R_{A,B}\{e^\delta \in X\}\), \(R_{B,C}\{e^\delta \in X\}\), and \(R_{A,C}\{e^\delta \in X\}\) correspondingly, the probability values for \(R_{A,B}\) and \(R_{B,C}\) are known, and \(R_{A,C}=R_{A,B} \oplus R_{B,C}\). In this case, the probability values \(e^\chi_{A,C}\) are calculated using the composition table for interval relations [1] according to the algorithm in Figure 2.

1. \(e^\chi_{A,C}=0\), where \(\chi \in X\);
2. for \(i=1\) to \(13\) do
3. for \(j=1\) to \(13\) do
4. \(\begin{align*}
&\text{begin} \\
&\quad X=\{\chi_1, \chi_2, \ldots, \chi_m\} \\
&\quad \text{for } k=1 \text{ to } m \text{ do } e^\chi_{A,C}=e^\chi_{A,C} + \frac{1}{m} e^\chi_{A,B} e^\chi_{B,C} \\
&\quad // \text{ where } \chi_1 \in X, \chi_2 \in X; \\
&\quad \text{end.}
\end{align*}\)

Figure 2. Composition of uncertain interval relations

In the algorithm in Figure 2 we consider all possible combinations of the probability values from \(R_{A,B}\) and \(R_{B,C}\). As a result of composition for every considered combination we compose a set \(X\) of Allen’s relations (line 5 of the algorithm), which are possible between \(A\) and \(C\). For example, for the combination of values \(e^b_{A,B}\) and \(e^d_{B,C}\), \(X=\{b,d,o,m,s\}\). After that, we distribute the joint probability \(e^b_{A,B} e^d_{B,C}\) between the probability values of the relations from \(X\) within \(R_{A,C}\) (line 6). In this way, the values \(e^\chi_{A,C}\) are accumulated and after consideration of all combinations of \(e^b_{A,B}\) and \(e^d_{B,C}\), we achieve the probability values for \(R_{A,C}\).

The operation of addition (\(\oplus\)) combines two uncertain interval relations \(R_{(A,B)}_h\) and \(R_{(A,B)}_i\) into a single relation \(R_{(A,B)}\).

**Definition** (Binary operation of addition for interval relations). Let us suppose that two uncertain relations between temporal intervals \(A\) and \(B\) are defined as \(R_{(A,B)}\{e^\delta \in X\}\) and \(R_{(A,B)}\{e^\delta \in X\}\), the probability values \(e^\chi_{(A,B)}\) and \(e^\chi_{(A,B)}\) are known, and \(R_{A,B}=R_{(A,B)_h} \oplus R_{(A,B)_i}\). In this case, the values \(e^\chi_{A,B}\) are calculated by formula (6):

\[
e^\chi_{A,B} = \sum_{\chi \in X} e^\chi \cdot e^\chi = \frac{e^\chi_{(A,B)_h} e^\chi_{(A,B)_i}}{e^\chi_{(A,B)_h} + e^\chi_{(A,B)_i}}. \tag{6}
\]

The obtained probability value \(e^\chi_{A,B}\) is neither smaller that the minimum of corresponding values \(e^\chi_{(A,B)_h}\) and \(e^\chi_{(A,B)_i}\) nor bigger that the maximum one.

4 Generating uncertain scenarios

In this section we propose one way to generate uncertain temporal scenarios using the reasoning mechanism defined in Section 3.

Let us represent a network of binary uncertain temporal relations as a directed graph, the nodes of which represent some events and the arcs represent temporal relations between these events. Formally, we represent such a graph as a set \(V\) of \(n\) variables \(v_1, v_2, \ldots, v_n\) and binary uncertain relations between these variables represented as \(r_{v_i, v_j}\{e^\chi \in X\}\), where \(v_i, v_j \in V\), if the variables are temporal points and as \(R_{v_i, v_j}\{e^\chi \in X\}\), where \(v_i, v_j \in V\), if the variables are temporal intervals.

**Definition** (Multiple operation of addition for point relations). Let us suppose that \(k\) uncertain interval relations between temporal intervals \(A\) and \(B\) are defined as \(R_{(A,B)_1}, R_{(A,B)_2}, \ldots, R_{(A,B)_k}\), the probability values \(e^\chi_{(A,B)_1}\), \(e^\chi_{(A,B)_2}\), and \(e^\chi_{(A,B)_k}\) are known, and \(R_{A,B}=\oplus (R_{(A,B)_1}, R_{(A,B)_2}, \ldots, R_{(A,B)_k})\). In this case, the probability values \(e^\chi_{A,B}\) are calculated by formula (7):

\[
e^\chi_{A,B} = \sum_{\chi \in X} e^\chi \cdot e^\chi = \frac{\prod_{i=1}^{n} e^\chi_{(A,B)_i}}{\sum_{j=1}^{n} e^\chi_{(A,B)_j}}, \text{ where } \chi \in X. \tag{7}
\]

Except for the notation, the definition of multiple operation of addition for point relations is similar.

Let us consider \(k\) networks \(N_1, N_2, \ldots, N_k\) of uncertain temporal relations defined by the set of nodes \(V=\{v_1, v_2, \ldots, v_n\}\), which is the same for each network, and the sets of uncertain temporal relations \(R_{1, R_{2}, \ldots, R_{k}}\) given for each network. These sets of relations are such that an element of one set is not necessarily included in other sets, for example, a relation \(r_{A,B}\in R_{1}\), but \(r_{A,B}\notin R_{2}\).

We suppose that an uncertain temporal scenario is a network of uncertain temporal relations defined by the set of nodes \(V=\{v_1, v_2, \ldots, v_n\}\), the set of relations \(R=R_1\cup R_2\cup \ldots \cup R_k\), where the relations within \(R\) are obtained by multiple operation of addition of the corresponding relations between the
same variables from all the sets $R_1$, $R_2$, ..., and $R_k$ according to the algorithm in Figure 3.

1. for $i=1$ to $n$ do
2. for $j=i+1$ to $n$ do
3. if $(\exists r_{i,j} \in R_i)$ or ... or $(\exists r_{i,j} \in R_k)$ then
4. begin
5. for $g=1$ to $n$ do
6. if not $(\exists r_{i,j} \in R_g)$ then Reasoning($r_{i,j}$, $R_g$)
7. // if “Reasoning” procedure fails to derive the desired relation then $(r_{i,j} \in R_g)$=TUR
8. $(r_{i,j} \in R) = \Theta (r_{i,j} \in R_i), \text{ where } t=1..k$
9. end
10. else go to line 2
Figure 3. Generating uncertain temporal scenario

Within the procedure “Reasoning” in line 6 of the algorithm from Figure 3, we obtain the set $V' = \{v'_1, v'_2, ..., v'_k\}$, where $V' \subseteq V$, which is a set of nodes derived using Dejkstra algorithm and representing the shortest path in the graph from $v_i$ to $v_j$ as $v_i \rightarrow v'_j \rightarrow v'_2 \rightarrow ... \rightarrow v'_k \rightarrow v_j$. There is at least one element in the set $V'$, otherwise the relation between $v_i$ and $v_j$ is present in the set $R_y$. After that, using the operation of composition we derive the desired relation $r_{i,j}$ as it is shown, for instance, in Figures 4 a, b, and c.

Figure 4. Deriving the unknown uncertain relation

The complexity of the main body of the algorithm in Figure 3 is $O(n^2)$, where $n$ is the number of nodes in the graph. Within the procedure “Reasoning” we apply Dejkstra algorithm, the complexity of which basically depends on its programming realization. In the worst case it is also $O(n^2)$ according to [4].

Let us consider an example of generating uncertain temporal scenario in Figures 5 a, b, and c.

![Networks and generated scenario](image)

Figure 5. Two networks and generated scenario

Network $N_f$ in Figure 5a is defined by the set of temporal points $V=\{a, b, c, d\}$ and the set of temporal relations between these points $R_f=\{r_{a,b}, r_{b,c}, r_{a,c}, r_{d,c}\}$. Network $N_2$ in Figure 5b is defined by the set $V=\{a, b, c, d\}$ and $R_2=\{r_{a,b}, r_{a,c}, r_{b,d}\}$. An uncertain temporal scenario $S$, presented in Figure 5c, is defined by the set of nodes $V=\{a, b, c, d\}$ and the set of relations $R=R_f \cup R_2=\{r_{a,b}, r_{b,c}, r_{b,d}, r_{a,c}, r_{d,c}\}$.

5 Temporal scenario recognition

Let us suppose that a relational network $N$ is defined by the set of nodes $V=\{n_1, n_2, ..., n_k\}$ and a set of relations between them $R_n$. Let us also suppose that an uncertain temporal scenario $S$ is defined by the same set of nodes $V=\{n_1, n_2, ..., n_k\}$ and a set of relations between them $R_s$. We suppose that the sets $R_n$ and $R_s$ are equal at the symbolic level of representation of relations, for example, both sets can include the relations $r_{a,b}, r_{b,c}, r_{b,d}, r_{a,c}$, and $r_{d,c}$. At the same time, each of these relations is defined by the set of probability measures for the basic relations that can hold between two particular temporal primitives. For instance, the relation $r_{a,b}$ can be defined as $\{e_{a,b}^{<}=0.5, e_{a,b}^{=} = 0.5, e_{a,b}^{>}=0\}$ within $R_n$ and as $\{e_{a,b}^{<}=0, e_{a,b}^{=} = 1, e_{a,b}^{>}=0\}$ within $R_s$.

The distance between the relational network $N$ and the scenario $S$ is calculated by formula (8):

$$D_{N,S} = \frac{\sum_{i=1}^{m} w_i d_i}{\sum_{i=1}^{m} w_i},$$

(8)

where $w_i$ - is the weight of $i$-th relation in the scenario $S$, $d_i$ - is the distance between two $i$-th relations from $R_n$ and from $R_s$.

In practice, the relations within $R_n$ and $R_s$ are initially different, like it is for example, in Figure 9a and 9c for the network $N_f$ and temporal scenario $S$. Therefore, before we calculate the distance value $D$ we should include the additional relations within $R_n$ (if needed) in the following way. If a relation, which is present within set $R_n$ is absent within set $R_s$ we try
to derive it within the network $N$ using the algorithm like in Figure 4. If this procedure fails then we assign the value TUR for this relation.

In Figure 6 there are $n$ uncertain temporal scenarios $S_1, S_2, \ldots, S_n$ and relational network $N$.

![Figure 6. Conceptual schema for scenario recognition](image)

In many situations it is necessary to know to which temporal scenario the network $N$ belongs, or if it is impossible to know then how close to every scenario the network is. Using the measure of distance between a temporal scenario and a relational network we can calculate the distances $D_{N,S_1}, D_{N,S_2}, \ldots, D_{N,S_n}$ between $N$ and every temporal scenario, as it is shown in Figure 6. The derived values can be represented as percentage values of similarity of the network $N$ with every scenario.

### 6 Conclusions

In this paper we proposed an application of uncertain temporal relations algebra to abstract medical diagnostic. A network of uncertain temporal relations describes a particular clinical course with the set of symptoms and temporal relationships between them. We showed how to generate a temporal scenario combining a number of networks having the same set of symptoms. During further diagnostic a relational network describing a particular clinical course can be compared with a number of scenarios using the formal criteria of distance between network and scenario.

Some experiments with implementation of the proposed mechanism using artificial settings have proved that the formalism is reasonable. Experiments with real medical datasets are considered as one of the directions for further research.

### References:


