Abstract Diagnostics Based on Uncertain Temporal Scenarios

V. Terziyan[1], V. Ryabov[2]
[1]Department of Mathematical Information Technology,
[2]Department of Computer Science and Information Systems,
University of Jyväskylä, FIN-40351, Jyväskylä
Finland
E-mails: {vagan,vlad}it@jyu.fi

Abstract

Temporal representation and reasoning has been applied so far to many areas of AI, including temporal diagnostics, where identification of temporal patterns plays an important role. Temporal scenarios of different device failures, their causes, and symptoms preceding the failure provide important additional information for making decisions concerning operation of the device. In this paper we propose an application of algebra of uncertain temporal relations to diagnostic problems. We represent uncertain temporal relations within a temporal 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 a network and a scenario.

1 Introduction

Temporal diagnostics is one important area of application of temporal representation and reasoning formalisms. It includes medical and industrial diagnostics, diagnostics in field device management, etc. Automated diagnostic reasoning about technical systems was overviewed in [1]. In that paper, the focus was on two main issues: (1) any serious, general solution has to address fundamental AI problems; (2) the nature of the task constrains these problems such that solutions become possible. In overall, the field of diagnostics was underlined as an important environment for assessing the actual utility of existing AI methods.

In [2] two generations of diagnostic systems are characterized and their advantages and limitations are discussed. There is a widespread opinion that first generation diagnostic systems work quite efficiently, but are unreliable and incomplete, whereas model-based systems are complete and robust, but suffer from complexity of models and intractability of the incorporated algorithms. This is thought to happen because the first generation diagnostic systems use heuristics whilst the model-based ones do not, and combining the two types of systems may solve the problem. In his paper, Struss argues that the tension between heuristics and model-based reasoning is a non-problem, because model-based diagnosis involves certain heuristics and also requires them. A principled integration that is clearly and formally grounded on model-based diagnosis and does not require essential changes in the implementation is proposed.

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 [3]. 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.

Temporal reasoning can also be used in medical diagnostics. For example, in [4] it was applied to the Heart Disease Program (HDP). Temporal constraints (relationships) are used together with probabilistic formalism in order to model the processes of cardiovascular reasoning accurately. The HDP has temporal constraints on the causal relations specified in the knowledge base and temporal properties on the patient input provided by the user. In overall, that paper discusses the issues and solutions that have been developed for temporal reasoning integrated with a pseudo-Bayesian probabilistic network in this challenging domain for diagnosis.

Another example of use of temporal diagnostics in medical domain is [5]. The example of hepatitis B was considered to describe a model-based framework for complex temporal behavior. The concept of abstract observations was introduced as an abstraction from observations at time points into assumptions over time intervals. This leads to a more intuitive representation and makes diagnosis independent of the number of actual observations and the granularity of time.

In [6] a template system is described that uses fuzzy set theory to provide a consistent mechanism of accounting for uncertainty in the existence of events, as well as vagueness in their starting times and duration. Fuzzy set theory allows the creation of fuzzy templates from linguistic rules. The fuzzy template system that is introduced in this paper can accommodate multiple time signals, relative or absolute trends, and obviates the need to also design a regression formula for pattern matching. The target application for the fuzzy template system was anesthesia monitoring.

In this paper we use Allen's interval algebra [7] to represent temporal uncertainty in abstract temporal diagnosis applications. Uncertain temporal relations are represented within a scenario graph using probabilities of the basic relations that can hold between two temporal primitives. In this paper we also propose:

(1) to generate 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 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. In Section 3 we define the measure of the distance between two uncertain temporal relations. Reasoning operations are discussed in Section 4. In Section 5 we show how to generate an uncertain temporal scenario combining a number of networks of temporal relations. In Section 6 we show how to compare the relational network representing the situation to be diagnosed with known scenarios using the special measure of the distance between a network and a scenario. Section 7 presents the discussion including the case from medical diagnostics area. Finally, Section 8 presents conclusions.

2 Representation of uncertain relations

Let us define the following time ontology. Time is linear, unbounded in both directions, and the time line is directed from the past to the future. Time model used is discrete meaning that the time axis is considered as a sequence of discrete temporal elements. Temporal points are the main ontological primitives isomorphic to natural numbers, i.e. each temporal point has a unique successor. Throughout this paper we will denote temporal points with small non-bold letters, i.e. a, b. Let us also denote a relation between two temporal points with a small bold letter $\mathbf{r}$ with subscript indicating the primitives, i.e. $\mathbf{r}_{a,b}$ is a temporal relation between points a and b.

Axiom 1. 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 $\alpha \in \mathbf{A}$.

A temporal interval is represented as a pair of temporal points denoting the start and the end of this interval, where the starting point is always before the endpoint. Intervals are denoted as capital non-bold letter, i.e. *A*, *B*. The relation between two intervals is denoted with a capital letter **R**, i.e. $\mathbf{R}_{A,B}$. There are thirteen basic Allen's relations [7] that can hold between two temporal intervals. The set of these relations is denoted as **X**={eq,b,bi,d,di,o,oi,m,mi,s,si,f,fi}. We will refer to an element of this set as $\chi \in \mathbf{X}$. A relation between two intervals could also be represented as a conjunction of the four relations between the endpoints of these intervals, as it was shown, for example, in [8]. Considering all consistent (according to the definition of interval) pairs of these four values and keeping in mind Axiom 1, one can easily see that these thirteen relations are the only relations that can hold between two intervals.

Ryabov and Puuronen in [9] 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. In this paper will use that representation. The probability of a basic temporal relation between two primitives is further denoted using letter "e" with a superscript indicating this basic relation and a subscript indicating these temporal primitives. For example, $\mathbf{r}_{a,b}\{e^{\alpha} | \alpha \in A\}$ is the uncertain relation between temporal points *a* and *b*, including the probabilities $e^{<}_{a,b}$, $e^{=}_{a,b}$, and $e^{>}_{a,b}$. An uncertain relation $\mathbf{R}_{A,B}\{e^{\chi} | \chi \in \mathbf{X}\}$ between intervals *A* and *B* includes thirteen probabilities of Allen's relations. The sum of all probability values within **r** or **R** is equal to 1. When $\exists \mathbf{e}^{\chi}_{A,B}=1$ within $\mathbf{R}_{A,B}\{e^{\chi} | \chi \in \mathbf{X}\}$ we call such $\mathbf{R}_{A,B}$ 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 suppose that the relations $\mathbf{r}_{(a,b)_1}\{e^{\alpha} | \alpha \in A\}$ and $\mathbf{r}_{(a,b)_2}\{e^{\alpha} | \alpha \in A\}$ are *equal* if and only if $\mathbf{e}^{<}_{(a,b)_1}=\mathbf{e}^{<}_{(a,b)_2}$, $\mathbf{e}^{=}_{(a,b)_1}=\mathbf{e}^{=}_{(a,b)_2}$, and $\mathbf{e}^{>}_{(a,b)_1}=\mathbf{e}^{>}_{(a,b)_2}$. Otherwise, relations $\mathbf{r}_{(a,b)_1}$ and $\mathbf{r}_{(a,b)_2}$ are *unequal*. In a similar way we will reason about interval relations.

Let us suppose that two temporal points p_1 and p_2 have been randomly selected on the time axis defined in terms of our time ontology. In this case, let us further assume that the probabilities of $p_1 < p_2$, $p_1 = p_2$, and $p_1 > p_2$ are $\mathbf{e}^{<}_{p_1,p_2}$, $\mathbf{e}^{=}_{p_1,p_2}$, and $\mathbf{e}^{>}_{p_1,p_2}$ correspondingly. We believe that the values of these probabilities vary depending on the particular application domain. For example, if no other information is available except for the mentioned above, it would be natural to guess that the value $\mathbf{e}^{=}_{p_1,p_2}$ will tend to zero, whereas the other two values will tend to 0.5. Let us define these three probability values in the interpretation given above as the *domain probability values* for point relations, and let us denote them as $\mathbf{e}^{<}_{\mathbf{D}}$, $\mathbf{e}^{=}_{\mathbf{D}}$, and $\mathbf{e}^{>}_{\mathbf{D}}$. In this case, the general uncertain point relation $\mathbf{r}\{\mathbf{e}^{<}_{\mathbf{D}},\mathbf{e}^{=}_{\mathbf{D}},\mathbf{e}^{>}_{\mathbf{D}}\}$ is called a *totally free distribution* (TFD) of probabilities of the basic point relations. In a similar way let us define a TFD for probabilities of the basic interval relations, i.e. $\mathbf{R}\{\mathbf{e}^{\chi}_{\mathbf{D}} | \chi \in \mathbf{X}\}$.

3 Distance between two uncertain relations

One approach to estimate the distance between two temporal relations was proposed in [10]. 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 $\mathbf{R}_{A,B}$ and $\mathbf{R}_{C,D}$ (Figure 1).

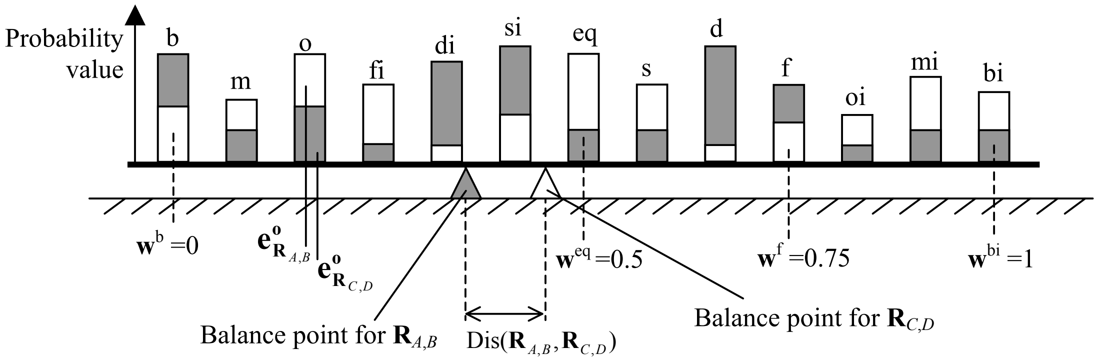


Figure 1. Physical interpretation of the measure of the distance between uncertain temporal relations

The probabilities of Allen's relations are represented as rectangles in Figure 1 with gray fill for $\mathbf{R}_{A,B}$ and with white fill for $\mathbf{R}_{C,D}$. 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.

Let us suppose that Allen's relations distributed on a virtual lath in Figure 1 have weights, denoted as $\mathbf{w}^b$, $\mathbf{w}^m$,…, and $\mathbf{w}^{bi}$. To simplify the notation of formulas further in the text, let us introduce the set of the weights as $W=\{\mathbf{w}^b,\mathbf{w}^m,\mathbf{w}^o,\mathbf{w}^{fi},\mathbf{w}^{di},\mathbf{w}^{si},\mathbf{w}^{eq},\mathbf{w}^s,\mathbf{w}^d,\mathbf{w}^f,\mathbf{w}^{oi},\mathbf{w}^{mi},\mathbf{w}^{bi}\}$ with a strict order of the relations within this set. We will refer to an element of the set W as w_i, where the subscript $i=\overline{0,12}$ stands for the number of the particular relation within the set W, in a way that w_0 is $\mathbf{w}^b$, w_6 is $\mathbf{w}^{eq}$, and w_{12} is $\mathbf{w}^{bi}$. The values of the weights are defined according to the formula $w_i=\frac{i}{12}$. In this way, for example, $\mathbf{w}^b$=0, $\mathbf{w}^{eq}$=0.5, and $\mathbf{w}^{bi}$=1.

The value of the balance point for the relation $\mathbf{R}_{A,B}$ is denoted as Bal($\mathbf{R}_{A,B}$) and is calculated as a sum of the multiplications of the weight of each Allen's relation on the corresponding probability value of this relation within $\mathbf{R}_{A,B}$, divided onto the sum of all probability values within $\mathbf{R}_{A,B}$ as in formula (1):

$$\mathrm{Bal}(\mathbf{R}_{A,B})=\frac{\sum_{i=0}^{12} w_i e_{A,B}^{\chi_i}}{\sum_{i=0}^{12} e_{A,B}^{\chi_i}}=\sum_{i=0}^{12} w_i e_{A,B}^{\chi_i}\,. \quad (1)$$

Taking into account that the lower part of the indicated division in (1) represents the sum of all probability values within the uncertain relation $\mathbf{R}_{A,B}\{e^{\chi}|\,\chi\in\mathbf{X}\}$, which equals to 1 according to the definition in Section 2, we obtained the final formula for the value of the balance point. The distance between two uncertain relations $\mathbf{R}_{A,B}$ and $\mathbf{R}_{C,D}$ is calculated as a module of the mathematical difference between their balance point values, as in formula (2):

$$\mathrm{Dis}\,(\mathrm{R}_{A,B},\mathrm{R}_{C,D}) = \left|\mathrm{Bal}(\mathrm{R}_{A,B}) - \mathrm{Bal}(\mathrm{R}_{C,D})\right|. \tag{2}$$

In a similar way, we derive the formula for the distance between uncertain point relations [10]. The value of the Dis function belongs to the interval [0,1]. When Dis ($\mathrm{R}_{A,B}$,$\mathrm{R}_{C,D}$) is equal to 0, this means that the values of the balance points Bal($\mathbf{R}_{A,B}$) and Bal($\mathbf{R}_{C,D}$) are equal, and this in its own turns suggests that the uncertain relations $\mathbf{R}_{A,B}$ and $\mathbf{R}_{C,D}$ are equal as it was defined in Section 2. The maximum value of the function Dis is 1, meaning that the relations $\mathbf{R}_{A,B}$ and $\mathbf{R}_{C,D}$ are maximally different. The examples of totally different relations are "<" and ">" for points and "before" and "after" for intervals.

4 Reasoning operations

In this section we briefly overview the reasoning mechanism including inversion, composition, and addition operations, proposed in [9]. 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.

Definition (unary inversion for interval relations). The operation of inversion (~) derives the relation $\mathbf{R}_{B,A}$ when the relation $\mathbf{R}_{A,B}$ is defined, and $\mathbf{R}_{B,A} = \widetilde{\mathbf{R}}_{A,B}$. We suppose that the probability values $\mathbf{e}^{\chi}_{A,B}$, where $\chi\in\mathbf{X}$, are known. In this case, the probability values $\mathbf{e}^{\chi}_{B,A}$ are calculated according to the inversion table for Allen's interval relations [7], i.e. $\mathbf{e}^{\mathrm{oi}}_{B,A} = \mathbf{e}^{\mathrm{o}}_{A,B}$.

Definition (composition for point relations). The operation of composition (⊗) derives the relation $\mathbf{r}_{a,c}$, when the relations $\mathbf{r}_{a,b}$ and $\mathbf{r}_{b,c}$ are defined, and $\mathbf{r}_{a,c}=\mathbf{r}_{a,b}\otimes\mathbf{r}_{b,c}$. We suppose that the probability values $\mathbf{e}^{\alpha}_{a,b}$, where $\alpha\in$ A, and $\mathbf{e}^{\alpha}_{b,c}$ are known. In this case, the probability values $\mathbf{e}^{\alpha}_{a,c}$ are calculated using the composition table [11] according to the formulas (3)-(5):

$$\mathbf{e}^{<}_{a,c} = \mathbf{e}^{<}_{a,b}\mathbf{e}^{<}_{b,c} + \mathbf{e}^{<}_{a,b}\mathbf{e}^{=}_{b,c} + \mathbf{e}^{=}_{a,b}\mathbf{e}^{<}_{b,c} + \mathbf{e}^{<}_{a,b}\mathbf{e}^{>}_{b,c}\mathbf{e}^{<}_{\mathbf{D}} + \mathbf{e}^{>}_{a,b}\mathbf{e}^{<}_{b,c}\mathbf{e}^{<}_{\mathbf{D}}, \tag{3}$$

$$\mathbf{e}^{=}_{a,c} = \mathbf{e}^{=}_{a,b}\mathbf{e}^{=}_{b,c} + \mathbf{e}^{<}_{a,b}\mathbf{e}^{>}_{b,c}\mathbf{e}^{=}_{\mathbf{D}} + \mathbf{e}^{>}_{a,b}\mathbf{e}^{<}_{b,c}\mathbf{e}^{=}_{\mathbf{D}}, \tag{4}$$

$$\mathbf{e}^{>}_{a,c} = \mathbf{e}^{>}_{a,b}\mathbf{e}^{>}_{b,c} + \mathbf{e}^{>}_{a,b}\mathbf{e}^{=}_{b,c} + \mathbf{e}^{=}_{a,b}\mathbf{e}^{>}_{b,c} + \mathbf{e}^{<}_{a,b}\mathbf{e}^{>}_{b,c}\mathbf{e}^{>}_{\mathbf{D}} + \mathbf{e}^{>}_{a,b}\mathbf{e}^{<}_{b,c}\mathbf{e}^{>}_{\mathbf{D}}. \tag{5}$$

In formulas (3)-(5) we used the values $\mathbf{e}^{<}_{\mathbf{D}}$, $\mathbf{e}^{=}_{\mathbf{D}}$, and $\mathbf{e}^{>}_{\mathbf{D}}$ in two situations: when "<" is composed with ">", and ">" with "<". According to the composition table by Vilain an Kautz [11] in those situations the resulting relation is "?" (unknown relation). Therefore, the probability $\mathbf{e}^{<}_{a,b}\mathbf{e}^{>}_{b,c}$ needs to be distributed between $\mathbf{e}^{<}_{a,c}$, $\mathbf{e}^{=}_{a,c}$, and $\mathbf{e}^{>}_{a,c}$ according to the TFD, i.e. $\mathbf{e}^{<}_{ab}\mathbf{e}^{>}_{bc}\mathbf{e}^{<}_{\mathbf{D}}$ contributes to $\mathbf{e}^{<}_{a,c}$, $\mathbf{e}^{<}_{ab}\mathbf{e}^{>}_{bc}\mathbf{e}^{=}_{\mathbf{D}}$ to $\mathbf{e}^{=}_{a,c}$, and $\mathbf{e}^{<}_{ab}\mathbf{e}^{>}_{bc}\mathbf{e}^{>}_{\mathbf{D}}$ to $\mathbf{e}^{>}_{a,c}$.

Definition (composition for interval relations). The operation of composition (⊗) derives the relation $\mathbf{R}_{A,C}$, when the relations $\mathbf{R}_{A,B}$ and $\mathbf{R}_{B,C}$ are defined, and $\mathbf{R}_{A,C}=\mathbf{R}_{A,B}\otimes\mathbf{R}_{B,C}$. We suppose that the probability values $\mathbf{e}^{\chi}_{A,B}$ and $\mathbf{e}^{\chi}_{B,C}$, where $\chi\in\mathbf{X}$, are known. In this case, the probability values $\mathbf{e}^{\chi}_{A,C}$ are calculated using the composition table [7] according to the algorithm in Figure 2.

1. $\mathbf{e}^{\chi}_{A,C}=0$, where $\chi\in\mathbf{X}$;
2. **for** i=1 **to** 13 **do**
3. **for** j=1 **to** 13 **do**
4. **for** k=1 **to** m **do** $\mathbf{e}^{\chi_k}_{A,C}=\mathbf{e}^{\chi_k}_{A,C}+\mathbf{e}^{\chi_k}_{\mathbf{D}}\mathbf{e}^{\chi_i}_{A,B}\,\mathbf{e}^{\chi_j}_{B,C}$
5. // where $\chi_k\in\{\chi_1,\chi_2,\ldots,\chi_m\}$; $\mathbf{e}^{\chi_k}_{\mathbf{D}}\in\left\{\mathbf{e}^{\chi_1}_{\mathbf{D}},\mathbf{e}^{\chi_2}_{\mathbf{D}},\ldots,\mathbf{e}^{\chi_m}_{\mathbf{D}}\right\}$; $\chi_i,\chi_j\in\mathbf{X}$;

Figure 2. Algorithm for the composition of uncertain interval relations

In the algorithm in Figure 2 we consider all possible combinations of the probability values from $\mathbf{R}_{A,B}$ and $\mathbf{R}_{B,C}$ similarly we have done for the composition of point relations. The set $\{\chi_1,\chi_2,\ldots,\chi_m\}$ is a subset of $\mathbf{X}$, and includes possible basic relations that can hold between A and C when χ_i and χ_j, both belong to $\mathbf{X}$, are composed. For example, the result of the composition of "before" and "during" is {b,d,o,m,s}. After that, we distribute the probability $\mathbf{e}^{\mathrm{b}}_{A,B}\mathbf{e}^{\mathrm{d}}_{B,C}$ between the values from the set $\{\chi_1,\chi_2,\ldots,\chi_m\}$ according to the TFD values $\left\{\mathbf{e}^{\chi_1}_{\mathbf{D}},\mathbf{e}^{\chi_2}_{\mathbf{D}},\ldots,\mathbf{e}^{\chi_m}_{\mathbf{D}}\right\}$, which are supposed to be defined for each possible resulting subset in the composition table [7].

Definition (multiple addition for interval relations). The operation of addition (⊕) derives the relation $\mathbf{R}_{A,B}$ summing up two or more uncertain relations $\mathbf{R}_{(A,B)_1},\mathbf{R}_{(A,B)_2},\ldots,\mathbf{R}_{(A,B)_\mathbf{n}}$, and $\mathbf{R}_{A,B}=\oplus(\mathbf{R}_{(A,B)_1},\mathbf{R}_{(A,B)_2},\ldots,\mathbf{R}_{(A,B)_\mathbf{n}})$. We suppose that the values $\mathbf{e}^{\chi}_{(A,B)_1},\mathbf{e}^{\chi}_{(A,B)_2},\ldots,\mathbf{e}^{\chi}_{(A,B)_\mathbf{n}}$ where $\chi\in\mathbf{X}$, are known. In this case, the probability values $\mathbf{e}^{\chi}_{A,B}$ are calculated using the formula (6):

$$\mathbf{e}^{\chi}_{A,B}=\frac{\mathbf{e}^{\chi}}{\sum\limits_{\chi\in\mathbf{X}}\mathbf{e}^{\chi}},\quad \mathbf{e}^{\chi}=\frac{\prod\limits_{\mathrm{j}=1}^{\mathbf{n}}\mathbf{e}^{\chi}_{(A,B)_\mathrm{j}}}{\sum\limits_{\mathrm{j}=1}^{\mathbf{n}}\mathbf{e}^{\chi}_{(A,B)_\mathrm{j}}}. \tag{6}$$

The formula (6) is applied to each basic relation $\chi\in\mathbf{X}$ within the operands consequently, meaning that first χ is "before" and we calculate the probability of "before" in $\mathbf{R}_{A,B}$, then χ is "meets", and so on. The physical interpretation of the proposed formula could be related to the formula of parallel resistance from the electrical division of physics. Finally, the obtained probability value $\mathbf{e}^{\chi}_{A,B}$ is neither smaller that the minimal one among the corresponding probability values within the operands, nor bigger that the maximum one.

5 Generation of uncertain temporal scenarios

Let us represent a network $N(V,\Psi)$ of binary uncertain temporal relations as a directed graph, the nodes of which represent some symptoms and the arcs represent temporal relations between these events. We represent such a graph as a set V of n variables $\{v_1,v_2,\ldots,v_n\}$ and the relations between these variables are represented as $\mathbf{r}_{v_i,v_j}\{\mathbf{e}^{\alpha} | \alpha \in \mathbf{A}\}$ or $\mathbf{R}_{v_i,v_j}\{\mathbf{e}^{\chi} | \chi \in \mathbf{X}\}$, where $v_i, v_j \in V$. The set of all uncertain temporal relations for the network N is denoted as Ψ.

Let us consider k networks $N_1(V,\Psi_1)$, $N_2(V,\Psi_2)$,…, $N_k(V,\Psi_k)$ of uncertain temporal relations. The set of nodes $V=\{v_1,v_2,\ldots,v_n\}$ is the same for each network. The sets of uncertain temporal relations Ψ_1, Ψ_2,…, Ψ_k are defined for each network. These sets of relations are such that an element included in one set is not necessarily included in other sets, for example, a relation $\mathbf{r}_{\mathbf{b,c}} \in \Psi_1$, but $\mathbf{r}_{\mathbf{b,c}} \notin \Psi_2$. We suppose that an uncertain temporal scenario $S(V,\Psi_s)$ is a network of uncertain temporal relations defined by the set of nodes V, and the set of relations $\Psi_s=\Psi_1 \cup \Psi_2 \cup \ldots \cup \Psi_k$. The relations within Ψ_s are obtained using the multiple operation of addition of the corresponding relations between the same variables from all the sets $\Psi_1,\Psi_2,\ldots,\Psi_k$ according to the algorithm (the notation for point relations is used) in Figures 3, 4, and 5.

1. **for** i=1 **to** n **do**
2. **for** j=i+1 **to** n **do**
3. **if** $(\exists\, \mathbf{r}_{v_i,v_j} \in \forall(\Psi_1,\ldots,\Psi_k))$ **then** {
4. **for** g=1 **to** n **do**
5. **if** $(\mathbf{r}_{v_i,v_j} \notin \Psi_g)$ **then** *Derive_Relation*$(\mathbf{r}_{v_i,v_j}, \Psi_g)$
6. $(\mathbf{r}_{v_i,v_j} \in \Psi_s) = \oplus\, (\mathbf{r}_{v_i,v_j} \in \Psi_t)$, where $t=\overline{1,k}$ }

Figure 3. The main algortihm for generation of temporal scenarios

1. **procedure** *Derive_Relation* $(\mathbf{r}_{v_i,v_j}, \Psi_g)$ {
2. $V' = \{v'_1, v'_2, .., v'_k\}$, $V' \subseteq V$ // The set of nodes derived using Dijkstra algorithm and // representing the shortest path between v_i and v_j in Ψ_g as $\mathbf{v_i} \rightarrow v'_1 \rightarrow \ldots \rightarrow v'_k \rightarrow \mathbf{v_j}$;
3. $\Omega=\{\mathbf{r}_1,\ldots,\mathbf{r}_n\}$; // the sequence of relations, such as $\mathbf{r}_1 = \mathbf{r}_{v_i,v'_1}$, …, $\mathbf{r}_n = \mathbf{r}_{v'_k,v_j}$.
4. **if** $\Omega=\{\varnothing\}$ then $(\mathbf{r}_{v_i,v_j} \in \Psi_g)$=TUR **else** $(\mathbf{r}_{v_i,v_j} \in \Psi_g)$ = *Compose_Sequence* (Ω); }

Figure 4. Procedure *Derive_Relation*

1. **function** *Compose_Sequence* ($\Omega=\{\mathbf{r}_1,\ldots,\mathbf{r}_n\}$– ordered set of relations, and $n=|\Omega|$) {
2. **if** n=1 **then return** $\mathbf{r} \in \Omega$; // the function returns $\mathbf{r}$ when it is the only relation in Ω
3. k=1; $\Omega' =\{\varnothing\}$;
4. for i=1 to mod(n/2) do {
5. $\mathbf{r}'_i = \mathbf{r}_k \otimes \mathbf{r}_{k+1}$; // $\mathbf{r}_k$, $\mathbf{r}_{k+1} \in \Omega$
6. $R' = R' \cup \{\mathbf{r}'_i\}$ // the set Ω' is ordered according to the value of the subindex "i".
7. k=k+2; }
8. if $n \neq i \times 2$ then $\Omega' = \Omega' \cup \{\mathbf{r}_n\}$ // $\mathbf{r}_n \in \Omega$
9. *Compose_Sequence* (Ω'); }

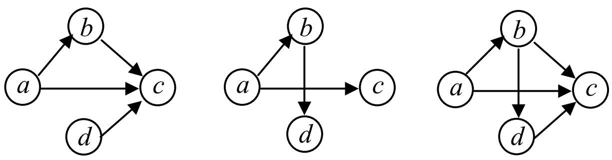
Figure 5. Recursive function *Compose_Sequence*

In the main algorithm, presented in Figure 3, we have two nested loops within which we basically control the key condition (line 3): if there exists a relation $\mathbf{r}_{v_i,v_j}$ in at least one of the networks to be composed. In the case this condition is satisfied, we try to derive (the cycle at lines 4 and 5) the desired relation in those networks where it is not implicitly present. After that, we use the operation of multiple addition to sum up the present and the derived relations $\mathbf{r}_{v_i,v_j}$ from all the networks, obtaining in this way the relation $\mathbf{r}_{v_i,v_j}$ for the resulting scenario graph. In the case the condition in line 3 is not satisfied (the relation is absent in all the networks), we proceed with the next iteration of the cycle.

The procedure *Derive_Relation*, presented in Figure 4, is intended for obtaining the desired relation $\mathbf{r}_{v_i,v_j}$ in the particular network $N_g(V,\Psi_g)$. Within this procedure, we first obtain the set of nodes representing the shortest path between the nodes v_i and v_j in N_g using Dijsktra algorithm (line 2). In the interests of space, we do not discuss in detail this process. An interested reader can easily find this algorithm in, for example, [12]. We represent this set of nodes as a sequence of relations $\Omega = \{\mathbf{r}_1,\ldots,\mathbf{r}_n\}$ between these nodes in a way, that $\mathbf{r}_1 = \mathbf{r}_{v_i,v'_l}$, …, $\mathbf{r}_n = \mathbf{r}_{v'_k,v_j}$ (line 3). After that, if $\Omega=\{\varnothing\}$, meaning that there is no a path between v_i and v_j in N_g, we assign the desired relation the value TUR (line 4). In the opposite case, we call the function *Compose_Sequence*, where the sequence Ω is a parameter of this function.

Within the body of the function *Compose_Sequence*, we first check the number of relations within the sequence Ω. In the case there exist only one element in this set, we return it to the procedure *Derive_Relation* as the desired result. Otherwise, we divide the elements into pairs, like $\mathbf{r}_1$ and $\mathbf{r}_2$, $\mathbf{r}_3$ and $\mathbf{r}_4$, …, until $\mathbf{r}_{n-1}$ and $\mathbf{r}_n$ if n is even, and until $\mathbf{r}_{n-2}$ and $\mathbf{r}_{n-1}$ if n is odd. Using the operation of composition we combine the obtained pairs and achieve the new sequence set Ω'. In the case n is odd we need to add the final element $\mathbf{r}_n$ to Ω'. The set Ω' has at most twice less elements plus one element compared to Ω. Finally, we call the function *Compose_Sequence* together with the sequence set Ω' as a new parameter. In this way, we will reach the situation when there would be only one element in the sequence set, and at this point we return to the procedure *Compose_Sequence*.

Let us consider an example of generating uncertain temporal scenario in Figures 6 a, b, and c.



a) Network N_1 b) Network N_2 c) Scenario S

Figure 6. Two networks and the generated scenario

For the network $N_1(V,\Psi_1)$ in Figure 5a the set of temporal points is $V=\{a,b,c,d\}$ and the set of temporal relations between these points $\Psi_1=\{\mathrm{r}_{a,b}, \mathrm{r}_{b,c}, \mathrm{r}_{a,c}, \mathrm{r}_{d,c}\}$. Network $N_2(V,\Psi_2)$ in Figure 5b is defined by the set V and $\Psi_2=\{\mathrm{r}_{a,b}, \mathrm{r}_{a,c}, \mathrm{r}_{b,d}\}$. The uncertain temporal scenario $S(V,\Psi_s)$, presented in Figure 5c, is defined by the set V and the set of relations $\Psi= \Psi_1 \cup \Psi_2=\{\mathrm{r}_{a,b}, \mathrm{r}_{b,c}, \mathrm{r}_{b,d}, \mathrm{r}_{a,c}, \mathrm{r}_{d,c}\}$.

6 Diagnostics through recognition of temporal scenarios

The basic idea of our diagnostics is in comparison of the network, representing the case to be diagnosed, with known temporal scenarios. To perform the comparison we use the measure of the distance between two uncertain relations. We also propose one way to numerically estimate the distance between the network representing the case to be diagnosed and a scenario, taking into account and combining, using the proposed in this paper formula, the calculated distances between separate temporal relations within the scenario and the network.

Firstly, let us show how to compare the particular network with the particular scenario. Let us suppose that the relational network $N_I(V,\Psi_1)$ represents the situation to be diagnosed. The set of nodes is $V=\{n_1, n_2,.., n_k\}$ and the set of relations between them is Ψ_1. Let us also suppose that the uncertain temporal scenario $S(V,\Psi_s)$ describes the particular situation. We suppose that the sets Ψ_1 and Ψ_s are equal at the symbolic level of representation of relations, i.e. $\{\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_m\}$, 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 $\{\mathbf{e}^{<}_{a,b}=0.5, \mathbf{e}^{=}_{a,b}=0.5, \mathbf{e}^{>}_{a,b}=0\}$ within Ψ_1 and as $\{\mathbf{e}^{<}_{a,b}=0, \mathbf{e}^{=}_{a,b}=1, \mathbf{e}^{>}_{a,b}=0\}$ within Ψ_s.

Let us assume that not all the relations with the set Ψ_s are equally important for the particular scenario. For example, it often happens that the importance of the relation between two particular symptoms prevails over the others. We propose to assign the temporal relations within each scenario with numerical weight values denoted as W_i, where $i=\overline{1,m}$, and $\Psi_s=\{\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_m\}$. Each value W_i represents the weight of the corresponding uncertain temporal relation $\mathbf{r}_i$ from Ψ_s. We do not propose in this paper any mean to define these values, supposing that it heavily depends on the specific of the application domain and subjective reasoning of the person performing control over the diagnostic system.

To estimate the distance between N_I and the scenario S we propose the formula (7):

$$\mathrm{Dis}(N,S)=\frac{\sum_{i=1}^{m} W_i\,\mathrm{Dis}(r_i\in\Psi_1, r_i\in\Psi_s)}{\sum_{i=1}^{m} W_i}, \qquad (7)$$

We need to find the distances between the same relations (e.g., r_1) taken from Ψ_1 and Ψ_s using formula (2) proposed in Section 3. Then each obtained distance value should be multiplied on the weight (e.g., W_1) of this relation in S. After that, we find the sum of all such multiplications, presented in the upper part of the indicated division in formula (7). The lower part of formula (7) includes the sum of all the weights among the scenario S.

In practice, the sets Ψ_1 and Ψ_s are initially different. Therefore, before we can calculate the distance value Dis(N,S) we should include the additional relations within Ψ_1 (if needed) in the following way. If a relation included in Ψ_s is absent in Ψ_1, we try to derive it using the algorithm similar to the one presented in Figures 4 and 5.

In the situation, when we have several temporal scenarios S_1, S_2,.., S_n, we can estimate the distances between the network representing the situation to be diagnosed and each of the defined scenarios. In this way, we will obtain the values Dis(N_1,S_1),.., Dis(N_1,S_n). Those scenarios, that are the closest to the network to be diagnosed, represent the most probable diagnoses for the situation observed. This corresponds to the minimal values of the calculated Dis functions. The derived values can also be represented as percentage values of similarity of the network *N* with every scenario.

7 Discussion: medical diagnostics case

Let us imagine that a patient has turned to a therapist when his unknown illness was already in a quite neglected form. In this case, the therapist asks about the dynamics of symptoms and often patients cannot describe it precisely. In such a way, the therapist has to deal with the uncertain temporal scenario of the disease's dynamics. He also needs to estimate the probabilities of possible diagnoses in this case. The approach proposed in this paper could be used to formalize and process this situation. In a scenario graph we represent the symptoms as vertices (temporal events) similarly as it was discussed regarding industrial diagnostics, and the arcs between the vertices include the labels standing for the uncertain temporal relations between these events.

A brief overview of research efforts in designing and developing time-oriented systems in medicine during the past decade was presented in [13]. Other authors, for example in [14], underlined that the ability to reason with time-oriented data is central to the practice of medicine. In [15] the crucial role of the temporal-reasoning and temporal-maintenance tasks for modern medical information and decision support systems was shown. Monitoring clinical variables over time often provides information driving medical decision-making, e.g. [16]. The examples of diagnostic systems incorporating temporal dimension are [17], [6], and [3].

8 Conclusions

In this paper we proposed an application of uncertain temporal relations algebra to abstract diagnostics. A network of uncertain temporal relations describes a particular course of events with the set of symptoms and temporal relationships between them. We have shown 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 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 datasets are considered as one of the directions for further research.

Acknowledgements

The authors would like to thank InBCT TEKES Project of Academy of Finland (Agora Center, University of Jyväskylä), which had financially supported in part this research. This work was also supported by the grant (#55626) of the Academy of Finland.

References

[1] P. Struss, Knowledge-Based Diagnosis - An Important Challenge and Touchstone for AI, *Proceedings of the 10th European Conference on Artificial Intelligence (ECAI-92)*, B. Neumann (Ed.), John Wiley & Sons, 1992.

[2] P. Struss, An Amalgamation of Model-Based and Heuristic Reasoning for Diagnosis, *Industrial Applications of Knowledge-Based Diagnosis*, G. Guida, et al. (Eds.), Elsevier, 311- 329, 1992.

[3] W. Nejdl & J. Gamper, Model-Based Diagnosis with Qualitative Temporal Uncertainty, *Proceedings of the 10th Conference on Uncertainty in AI*, Seattle, *USA*, 432-439, 1994.

[4] W. Long, Temporal Reasoning for Diagnosis in a Causal Probabilistic Knowledge Base, *Artificial Intelligence in Medicine,* Vol. 8, 193-215, 1996.

[5] J. Gamper & W. Nejdl, Abstract Temporal Diagnosis in Medical Domains, *Artificial Intelligence in Medicine*, Vol. 10, No. 3, 209-234, 1997.

[6] A. Lowe, R. Jones & M. Harrison, Temporal Pattern Matching Using Fuzzy Templates*,* *Journal of Intelligent Information Systems*, Vol. 13, No. 1-2, 27-45, 1999.

[7] J. Allen, Maintaining Knowledge about Temporal Intervals, *Communications of the ACM*, Vol. 26, No. 11, 832-843, 1983.

[8] V. Ryabov, Estimating Uncertain Relations between Indeterminate Temporal Points and Intervals, *Proceedings of the 8-th International Symposium on Temporal Representation and Reasoning (TIME'01)*, IEEE Computer Society Press, 69-74, 2001.

[9] V. Ryabov & S. Puuronen, Probabilistic Reasoning about Uncertain Relations between Temporal Points, *Proceedings of the 8-th International Symposium on Temporal Representation and Reasoning (TIME'01)*, IEEE Computer Society Press, 35-40, 2001.

[10] H. Kaikova & S. Puuronen, Reasoning Temporal Sequence from Multiple Temporal Sequences, *Computational Intelligence for Modeling, Control & Automation: Intelligent Image Processing, Data Analysis & Information Retrieval*, M. Mohammadian (Ed.), IOS Press, 215-220, 1999.

[11] N. Vilain & H. Kautz, Constraint propagation algorithms for temporal reasoning. In Proceedings of the Fifth National Conference of the American Association for Artificial Intelligence, AAAI Press, 377-382, 1986.

[12] T. Cormen (Ed.), C. Leiserson & R. Rivest, *Introduction to Algorithms*, The MIT Press, 2-d Edition, Massachusetts, USA, 2001.

[13] C. Combi & Y. Shahar, Temporal Reasoning and Temporal Data Maintenance in Medicine: Issues and Challenges, *Computers in Biology and Medicine*, Vol. 27, No. 5, 353-368, 1997.

[14] J. H. Nguyen, Y. Shahar, S. W. Tu, A. K. Das & M. A. Musen, Integration of Temporal Reasoning and Temporal-Data Maintenance into a Reusable Database Mediator to Answer Abstract, Time-Oriented Queries: The Tzolkin System, *Journal of Intelligent Information Systems*, Vol. 13, No. 1-2, 121-145, 1999.

[15] Y. Shahar, Dimension of Time in Illness: An Objective View, *Annals of Internal Medicine*, Vol. 132, No. 1, 45-53, 2000.

[16] L. Console, A. J. Rivolin, D. T. Dupr'e & P. Torasso, Integration of Causal and Temporal Reasoning in Diagnostic Problem Solving, *Proceedings of the 9th International Workshop on Expert Systems and Their Applications*, 1989.

[17] M. Dojat & C. Sayettat, A Realistic Model for Temporal Reasoning in Real-Time Patient Monitoring, *Applied Artificial Intelligence*, Vol. 10, 121-143, 1996.