# Uncertain Relations between Indeterminate Temporal Intervals

Vladimir Ryabov
researcher, M.Sc.
Department of Computer Science and Information Systems, University of Jyväskylä,
P.O.Box 35, FIN-40351 Jyväskylä, Finland
+380 14 2602479
vlad@jytko.jyu.fi

## ABSTRACT

Time is important in modeling dynamic aspects of the world, and particularly in the temporal databases field. Many research articles about temporal databases assume that complete and accurate temporal information is available. However, in many real applications temporal information is imperfect and we need to find some way of handling it. The two main commonly used temporal ontological primitives are point and interval. Often a temporal point is indeterminate, which means that an interval of possible values for the point is given, and the probability mass function for that interval can be defined. An indeterminate temporal interval is represented as a pair of indeterminate points denoting the start and the end of the interval.

Sometimes there is a need to know relations between temporal intervals, as for example, in query processing. When intervals are indeterminate, it is almost impossible to derive a certain relation between them. This paper presents an approach to estimate the uncertain relation between two indeterminate temporal intervals by calculating the probabilities of Allen's relations. The relation between two indeterminate intervals is represented using four relations between their indeterminate endpoints. Using the information about the endpoints, we derive four uncertain relations between them, and then calculate the probabilities of Allen's relations by the proposed formulas. We present an example of the behavior of the proposed mechanism regarding calculation complexity using different input data. We also consider an example illustrating a possible application of the approach.

ADVANCES IN DATA MANAGEMENT 2000
Krithi Ramamritham, T.M. Vijayaraman (Editors)
Tata McGraw-Hill Publishing Company Ltd.


## 1. INTRODUCTION

Time is important in modeling dynamic aspects of the world. Even though representation and reasoning about temporal information has already achieved significant results to some extent, there still exist topics which require and deserve further research attention. Temporal formalisms are applied, for example, in natural language understanding, planning, process control, and temporal databases, i.e. in the areas, where the time course of events plays an important role.

In temporal databases each event is associated with a timestamp that indicates when the event has occurred. Many research articles about temporal databases assume that complete and accurate temporal information is available. Generally, the proposed approaches give little or no support for situations in which temporal imperfection exists. However, in many real applications temporal information is imperfect and we need to find some way of handling it.

One kind of imperfect information is indeterminacy, which means that it is known only approximately when a particular event happened, for example, we may know that "it happened during last week" or "between 8 a.m. and 9 a.m.". Indeterminacy can arise from different sources, such as suggested in [6]:

- granularity mismatch (when an event is measured in one granularity, and is recorded in a system with a finer granularity);

- clock measurements (every clock measurement has some imprecision);

- dating techniques (many dating techniques are inherently imprecise);

- unknown or imprecise event times (in general, occurrence times could be unknown or imprecise);

- uncertainty in planning (projected completion dates are often inexactly specified).

Motro [8] suggests that imperfect information can also result from unreliable sources, such as faulty sensors, input errors, or the inappropriate choice of representation. Kwan et al. [7] mention a number of sources of uncertainty and incompleteness in databases used in scientific applications. For example, some data is recorded statistically and so is inherently uncertain, whilst other data is deliberately made uncertain for reasons of security, and some data can not be measured accurately, due perhaps to some mechanical effect.

Point and interval are the two main temporal ontological primitives proposed in the literature. For an indeterminate point a closed interval of possible values is defined together with the probability mass function defining the probabilities of these values. An indeterminate interval is represented as a pair of indeterminate endpoints denoting the start and the end of the interval.

Often there is a need to know the relation between two temporal intervals, as for example, in query processing. The relation between two temporal intervals can be represented in two main ways: using Allen's interval relations [1], and using four relations between the endpoints of the intervals. The representation of the relation between two temporal intervals using Allen's relations is desirable in many applications and it is more expressive compared to the representation using the relations between the endpoints. Often the information about Allen's relations between two intervals is not readily available, however, the information about the endpoints of these intervals is almost always present, but in many situations in real applications this information is indeterminate. In this paper we propose one way to derive Allen's relations from the information about the endpoints of two indeterminate intervals.

The structure of the paper is the following. In Section 2 we present the main concepts used in the paper. The notion of the probability mass function and its sources are discussed in Section 3. Section 4 presents an approach to estimate uncertain relations between two indeterminate temporal points. We compose the probabilities of Allen's relations between two indeterminate intervals using the four uncertain relations between their endpoints in Section 5. Section 6 presents an example of the behavior of the mechanism regarding calculation complexity using different input data. Section 7 includes an example of the estimation of Allen's relations in the temporal database. The discussion about the related research is included in Section 8, and, finally, in Section 9 we make conclusions.

## 2. MAIN CONCEPTS

In this section we present the main concepts used in the paper.

The ontology of time used in this paper is below defined similarly to the one proposed in [6]. The various models of time that have been proposed in the literature are often classified as discrete, dense, and continuous models. We use the discrete model, which is commonly used in the temporal database research field. Temporal points, as the main ontological primitives, are isomorphic to natural numbers, i.e. there is the notion that every point has a unique successor. The time between two points is known as a *temporal interval*. A c*hronon* is an indivisible time interval of some fixed duration. A time line is represented by a sequence of chronons of identical duration. We do not specify the particular chronon size, but let it vary depending on the application. A temporal point is *determinate* if it is exactly known during which particular chronon it is located. Often it is not known exactly, but an interval of chronons, during which this point can be found, is given.

**Definition 1.** An *indeterminate* temporal point **a** is a temporal point such that $\mathbf{a} \in [\mathbf{a^l},\mathbf{a^u}]$, where $\mathbf{a^l}$ (lower bound) is the first chronon of the interval $[\mathbf{a^l},\mathbf{a^u}]$, $\mathbf{a^u}$ (upper bound) is the last chronon, $\mathbf{a^l} \leq \mathbf{a^u}$, and it is attached with a probability mass function (p.m.f.) **f**(**a**).

**Definition 2.** Let an *uncertain relation* between two indeterminate temporal points **a** and **b** be represented by a vector $(\mathbf{e^<},\mathbf{e^=},\mathbf{e^>})$, where the value $\mathbf{e^<}$ is the probability that $\mathbf{a}<\mathbf{b}$, the value $\mathbf{e^=}$ is the probability that $\mathbf{a}=\mathbf{b}$, and the value $\mathbf{e^>}$ is the probability that $\mathbf{a}>\mathbf{b}$. The sum of $\mathbf{e^<}$, $\mathbf{e^=}$, and $\mathbf{e^>}$ is equal to 1, since these values represent all the possible basic relations between points **a** and **b**.

**Definition 3.** Let an *indeterminate* temporal interval **A** be defined as a pair of indeterminate temporal points **s** and **e**, specifying the start and the end of the interval **A**. The starting point **s** from the interval $[\mathbf{s^l},\mathbf{s^u}]$ should be before the end point **e** which belongs to the interval $[\mathbf{e^l},\mathbf{e^u}]$, so that the endpoints **s** and **e** do not overlap, i.e. $\mathbf{s^u}<\mathbf{e^l}$.

There are two main approaches to represent the relation between two temporal intervals. One approach is to use the thirteen interval relations proposed by Allen [1]: "equals" (eq), "before" (b), "after" (bi), "meets" (m), "met-by" (mi), "during" (d), "contains" (di), "overlaps" (o), "overlapped-by" (oi), "starts" (s), "started-by" (si), "finishes" (f), and "finished-by" (fi) (Figure 1).

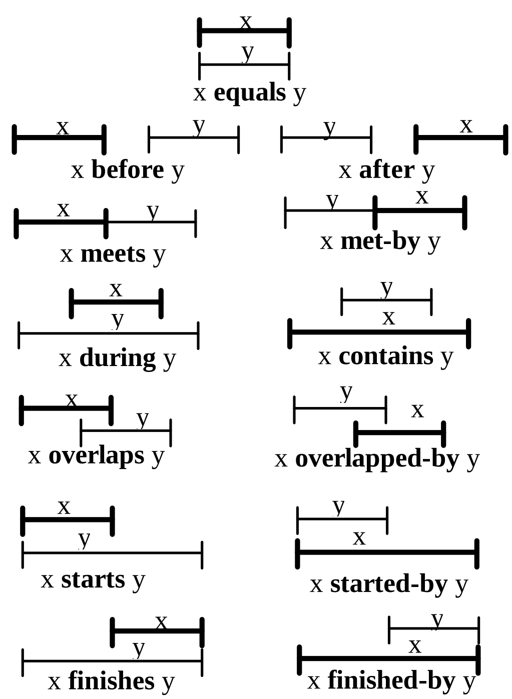


**Figure 1. Allen's interval relations.**

Another approach suggests that the relation between two temporal intervals can be represented using four relations between the endpoints of these intervals. Figure 2 presents the uncertain relation between two indeterminate intervals $\mathbf{A}[\mathbf{s_1},\mathbf{e_1}]$ and $\mathbf{B}[\mathbf{s_2},\mathbf{e_2}]$ using the relations $\mathbf{r_{11}}$, $\mathbf{r_{12}}$, $\mathbf{r_{21}}$, and $\mathbf{r_{22}}$, which can take the values "<", "=", and ">".

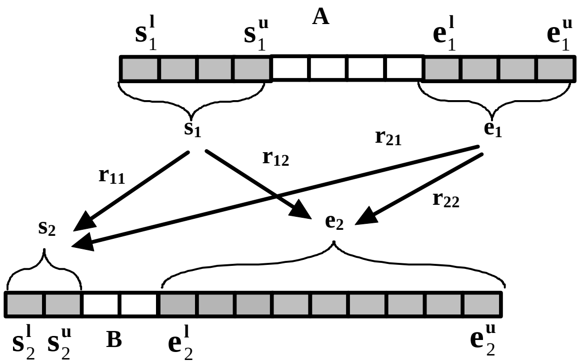


**Figure 2. The uncertain relation between indeterminate intervals A [s₁,e₁] and B [s₂,e₂].**

It is convenient to represent the relation between two temporal intervals using the matrix $\hat{A} = \begin{bmatrix} \mathbf{r}_{11} & \mathbf{r}_{12} \\ \mathbf{r}_{21} & \mathbf{r}_{22} \end{bmatrix}_{\mathbf{A},\mathbf{B}}$, where $\mathbf{r_{11}}$, $\mathbf{r_{12}}$, $\mathbf{r_{21}}$, and $\mathbf{r_{22}}$ are the relations between the endpoints of the intervals. The uncertain relation between two indeterminate intervals **A** and **B** is represented as a matrix $\hat{A}$, where the uncertain relations $\mathbf{r_{11}}$, $\mathbf{r_{12}}$, $\mathbf{r_{21}}$, and $\mathbf{r_{22}}$ are represented as in Definition 2:

$$\hat{A} = \begin{bmatrix} (\mathbf{e}^{<},\mathbf{e}^{=},\mathbf{e}^{>})_{\mathbf{r}_{11}} & (\mathbf{e}^{<},\mathbf{e}^{=},\mathbf{e}^{>})_{\mathbf{r}_{12}} \\ (\mathbf{e}^{<},\mathbf{e}^{=},\mathbf{e}^{>})_{\mathbf{r}_{22}} & (\mathbf{e}^{<},\mathbf{e}^{=},\mathbf{e}^{>})_{\mathbf{r}_{22}} \end{bmatrix}_{\mathbf{A},\mathbf{B}}.$$

In the next section we briefly discuss the notion of the probability mass function and its sources.

## 3. PROBABILITY MASS FUNCTION

In this section we present a discussion about the notion of the probability mass function, which is mainly drawn from the article by Dyreson and Snodgrass [6], but heavily relates to the topic of this paper and is essential for further understanding of the material.

In many applications the chronons within the interval for an indeterminate point may not be equally probable. It is reasonable to take into account the probabilities of these chronons by defining the probability mass function.

In some application domains the middle value from the interval may have the highest probability. For example, if it is known, that the temperature outside is in the interval [+10,+30] degrees, then it is natural to guess that the most probable value of the temperature is about +20. In other domains the chronons inside the interval may have equal probabilities, and then the interval is considered as a set of consequent equally probable values.

The probability mass function (p.m.f.) $\mathbf{f}(\mathbf{a})$ for an indeterminate point $\mathbf{a}$, which belongs to the interval $[\mathbf{a}_1^{\mathbf{l}}, \mathbf{a}_1^{\mathbf{u}}]$, defines the probabilities of the chronons within this interval so, that $\sum_{\mathbf{a}=\mathbf{a}^{\mathbf{l}}}^{\mathbf{a}^{\mathbf{u}}} \mathbf{f}(\mathbf{a}) = 1$. The requirement that the sum of all the probabilities of the chronons is equal to 1 results from the definition of time ontology, according to which, a temporal point occurs exactly during one particular chronon. We do not consider situations when the probabilities defined by the p.m.f. are joint or dependent.

Figures 3 and 4 present two examples of the p.m.f.s for the endpoints $\mathbf{s_1}$ and $\mathbf{e_2}$ from Figure 2.

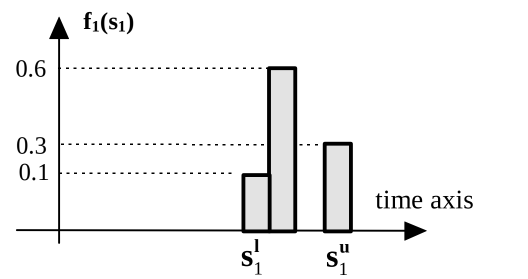


**Figure 3. P.m.f. for $\mathbf{s_1}$.**

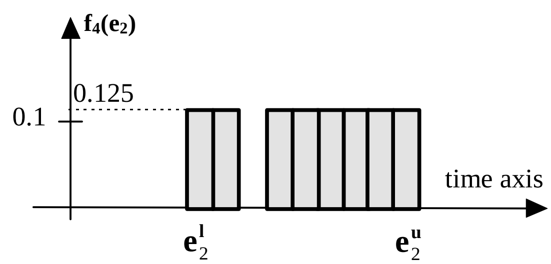


**Figure 4. P.m.f. for $\mathbf{e_2}$.**

In the interval for the point $\mathbf{s_1}$ all chronons have different probabilities with the maximum one in the middle of the interval. In the interval for the point $\mathbf{e_2}$ almost all chronons are equally probable.

We assume that a p.m.f. is given when an indeterminate point is created. Generally, the p.m.f. stems from the sources of indeterminacy, such as granularity mismatch, dating and measurements techniques, etc. When the granularity mismatch is the source of indeterminacy the uniform distribution is a useful assumption. For example, if an event is known in the granularity of one hour then in a system with the granularity of one second it is indeterminate, and we have no reason to favor one second over another. Some measurement techniques or instruments can have fixed trends in measurements, for example, the normal distribution of a variable. In some situations, the analysis of past data can provide a hint to defining the p.m.f. For example, we may know that a particular type of event in a particular situation tends to occur during the last chronons of the interval.

Several other means of determining the p.m.f. were suggested by Dey and Sarkar [3]. Also Dyreson and Snodgrass [6] point out that in some cases a user may not know the underlying mass function because that information is unavailable. In such cases the distribution can be specified as missing, which represents a complete lack of knowledge about the distribution. In our approach we suppose that the distribution is already and totally known. In the case when the distribution is not specified, one of the above mentioned means of defining the p.m.f. can be applied.

In the next section we propose formulas for the probabilities of the basic relations between two indeterminate temporal points.

## 4. UNCERTAIN RELATIONS BETWEEN INDETERMINATE TEMPORAL POINTS

In this section we present an approach to estimate the uncertain relation between two indeterminate temporal points by calculating probabilities of the basic relations between them. The discussion is based around two indeterminate points $\mathbf{s}_1$ and $\mathbf{e}_2$ from the intervals $[\mathbf{s}_1^{\mathbf{l}}, \mathbf{s}_1^{\mathbf{u}}]$ and $[\mathbf{e}_2^{\mathbf{l}}, \mathbf{e}_2^{\mathbf{u}}]$ correspondingly (Figure 2). We also take into account the p.m.f.s for these points $\mathbf{f}_1(\mathbf{s}_1)$ and $\mathbf{f}_4(\mathbf{e}_2)$, which are defined by Figures 3 and 4.

One approach to estimate the uncertain relation between $\mathbf{s}_1$ and $\mathbf{e}_2$ is a simple comparison of all possible values of $\mathbf{s}_1$ and $\mathbf{e}_2$. This approach is very time consuming in the case when the intervals $[\mathbf{s}_1^{\mathbf{l}}, \mathbf{s}_1^{\mathbf{u}}]$ and $[\mathbf{e}_2^{\mathbf{l}}, \mathbf{e}_2^{\mathbf{u}}]$ include many chronons. A more advanced way is to compare only those chronons that are located inside the common part of the intervals. We distinguish between two main cases: 1) the intervals $[\mathbf{s}_1^{\mathbf{l}}, \mathbf{s}_1^{\mathbf{u}}]$ and $[\mathbf{e}_2^{\mathbf{l}}, \mathbf{e}_2^{\mathbf{u}}]$ do not overlap, and 2) the intervals overlap. When the intervals do not overlap and, hence, have no common chronons, the relation between them can easily be estimated by comparing the values of the endpoints $\mathbf{s}_1^{\mathbf{u}}$, $\mathbf{e}_2^{\mathbf{l}}$, $\mathbf{e}_2^{\mathbf{u}}$, and $\mathbf{s}_1^{\mathbf{l}}$. When $\mathbf{s}_1^{\mathbf{u}} < \mathbf{e}_2^{\mathbf{l}}$ the relation between $\mathbf{s}_1$ and $\mathbf{e}_2$ is "<". When $\mathbf{e}_2^{\mathbf{u}} < \mathbf{s}_1^{\mathbf{l}}$ the relation between $\mathbf{s}_1$ and $\mathbf{e}_2$ is ">".

In the second case the intervals overlap, which means they have at least one common chronon. The common chronons are included in the common part interval, the endpoints of which are $\mathbf{a}$ (starting point) and $\mathbf{b}$ (end point), where $\mathbf{a} = \mathbf{max}\left(\mathbf{s}_1^{\mathbf{l}}, \mathbf{e}_2^{\mathbf{l}}\right)$ and $\mathbf{b} = \mathbf{min}\left(\mathbf{s}_1^{\mathbf{u}}, \mathbf{e}_2^{\mathbf{u}}\right)$. Let us consider the composition of the formula for the probability $\mathbf{e}^{<}$ (Figure 5).

Figure 5 includes the items that are used to compose the probability $\mathbf{e}^{<}$. The intervals $[\mathbf{s}_1^{\mathbf{l}}, \mathbf{s}_1^{\mathbf{u}}]$ and $[\mathbf{e}_2^{\mathbf{l}}, \mathbf{e}_2^{\mathbf{u}}]$ are divided by filled brackets into subintervals, which distinguish the common part of the intervals. For each pair of possible values of $\mathbf{s}_1$ and $\mathbf{e}_2$ taken from the intervals $[\mathbf{s}_1^{\mathbf{l}}, \mathbf{s}_1^{\mathbf{u}}]$ and $[\mathbf{e}_2^{\mathbf{l}}, \mathbf{e}_2^{\mathbf{u}}]$ correspondingly, we find out the joint probability of the pair as $\mathbf{f}_1(\mathbf{s}_1)\times\mathbf{f}_4(\mathbf{e}_2)$.

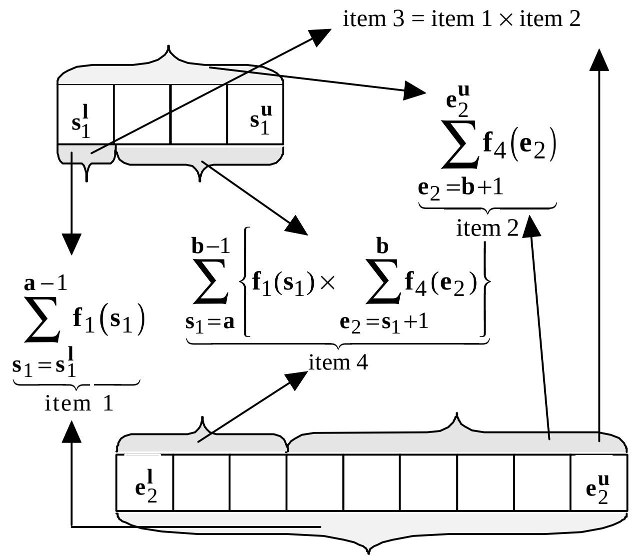


**Figure 5. Items of the probability $\mathbf{e}^{<}$.**

Item 1 represents the probability of the pairs of $\mathbf{s}_1$ and $\mathbf{e}_2$, where $\mathbf{s}_1$ is taken from the subinterval $[\mathbf{s}_1^{\mathbf{l}}, \mathbf{a}\text{-}1]$, and $\mathbf{e}_2$ belongs to the interval $[\mathbf{e}_2^{\mathbf{l}}, \mathbf{e}_2^{\mathbf{u}}]$. Since $\sum_{\mathbf{e}_2=\mathbf{e}_2^{\mathbf{l}}}^{\mathbf{e}_2^{\mathbf{u}}} \mathbf{f}_4(\mathbf{e}_2) = 1$, item 1 is represented only by a sum of all the values of $\mathbf{s}_1$ within the subinterval $[\mathbf{s}_1^{\mathbf{l}}, \mathbf{a}\text{-}1]$, i.e. item 1 = $\sum_{\mathbf{s}_1=\mathbf{s}_1^{\mathbf{l}}}^{\mathbf{a}-1} \mathbf{f}_1(\mathbf{s}_1)$. Item 2 is obtained similarly. Some combinations of the values of $\mathbf{s}_1$ and $\mathbf{e}_2$ within the closed subintervals $[\mathbf{s}_1^{\mathbf{l}}, \mathbf{a}\text{-}1]$ and $[\mathbf{b}+1, \mathbf{e}_2^{\mathbf{u}}]$ are represented twice across items 1 and 2. The sum of probabilities of these combinations, which is represented by item 3, needs to be subtracted from $\mathbf{e}^{<}$.

Next, we consider the common part $[\mathbf{a},\mathbf{b}]$ of the intervals $[\mathbf{s}_1^{\mathbf{l}}, \mathbf{s}_1^{\mathbf{u}}]$ and $[\mathbf{e}_2^{\mathbf{l}}, \mathbf{e}_2^{\mathbf{u}}]$. Here we follow the trivial approach, i.e. a simple comparison of all possible combinations of values of $\mathbf{s}_1$ and $\mathbf{e}_2$ within $[\mathbf{a},\mathbf{b}]$, since no useful heuristics can be applied. Item 4 is composed of the combinations that support the relation "<" from the common part.

Finally, the probability $\mathbf{e}^{<}$ is composed of all four items from Figure 5:

$$\mathbf{e}^{<} = \underbrace{\sum_{\mathbf{s}_1=\mathbf{s}_1^{\mathbf{l}}}^{\mathbf{a}-1} \mathbf{f}_1(\mathbf{s}_1)}_{\text{item 1}} + \underbrace{\sum_{\mathbf{e}_2=\mathbf{b}+1}^{\mathbf{e}_2^{\mathbf{u}}} \mathbf{f}_4(\mathbf{e}_2)}_{\text{item 2}} - \underbrace{\sum_{\mathbf{s}_1=\mathbf{s}_1^{\mathbf{l}}}^{\mathbf{a}-1} \left\{ \mathbf{f}_1(\mathbf{s}_1) \times \sum_{\mathbf{e}_2=\mathbf{b}+1}^{\mathbf{e}_2^{\mathbf{u}}} \mathbf{f}_4(\mathbf{e}_2) \right\}}_{\text{item 3}} +$$

$$+ \underbrace{\sum_{\mathbf{s}_1=\mathbf{a}}^{\mathbf{b}-1} \left\{ \mathbf{f}_1(\mathbf{s}_1) \times \sum_{\mathbf{e}_2=\mathbf{s}_1+1}^{\mathbf{b}} \mathbf{f}_4(\mathbf{e}_2) \right\}}_{\text{item 4}}.$$

Similarly we can obtain the formula for $\mathbf{e}^{>}$:

$$\mathbf{e}^{>} = \sum_{\mathbf{e}_2=\mathbf{e}_2^{\mathbf{l}}}^{\mathbf{a}-1} \mathbf{f}_4(\mathbf{e}_2) + \sum_{\mathbf{s}_1=\mathbf{b}+1}^{\mathbf{s}_1^{\mathbf{u}}} \mathbf{f}_1(\mathbf{s}_1) - \sum_{\mathbf{e}_2=\mathbf{e}_2^{\mathbf{l}}}^{\mathbf{a}-1} \left\{ \mathbf{f}_4(\mathbf{e}_2) \times \sum_{\mathbf{s}_1=\mathbf{b}+1}^{\mathbf{s}_1^{\mathbf{u}}} \mathbf{f}_1(\mathbf{s}_1) \right\} +$$

$$+ \sum_{\mathbf{e}_2=\mathbf{a}}^{\mathbf{b}-1} \left\{ \mathbf{f}_4(\mathbf{e}_2) \times \sum_{\mathbf{s}_1=\mathbf{e}_2+1}^{\mathbf{b}} \mathbf{f}_1(\mathbf{s}_1) \right\};$$

Composing the probability

$$\mathbf{e}^{=} = \sum_{\mathbf{s}_1=\mathbf{a}}^{\mathbf{b}} \left\{ \mathbf{f}_1(\mathbf{s}_1) \times \mathbf{f}_4(\mathbf{e}_2) \right\}$$

we take into account only those pairs of values of $\mathbf{s_1}$ and $\mathbf{e_2}$ that belong to the common part [**a**,**b**], since only these values can be equal. Using the proposed formulas we can estimate the uncertain relation between any two indeterminate temporal points.

# 5. PROBABILITIES OF ALLEN'S RELATIONS BETWEEN INDETERMINATE INTERVALS

In this section we propose formulas for the probabilities of Allen's relations between two indeterminate temporal intervals. The formulas are composed using the notion of conditional probability, where the basic events are the possible values of the relations between the endpoints of the intervals.

The relations $\mathbf{r_{11}}$, $\mathbf{r_{12}}$, $\mathbf{r_{21}}$, and $\mathbf{r_{22}}$ can take the values "<","=", and ">". Totally there exist 81 combinations of different values of these four relations, among which only 13 combinations correspond to valid Allen's relations, because the rest of the combinations contradict the definition of the temporal interval. This means that we are dealing with a system of dependent relations. Let us define a set of events $W_1$={ $\mathbf{r}_{11}^{<}, \mathbf{r}_{11}^{=}, \mathbf{r}_{11}^{>}, \mathbf{r}_{12}^{<}, \mathbf{r}_{12}^{=}, \mathbf{r}_{12}^{>}, \mathbf{r}_{21}^{<}, \mathbf{r}_{21}^{=}, \mathbf{r}_{21}^{>}, \mathbf{r}_{22}^{<}, \mathbf{r}_{22}^{=}, \mathbf{r}_{22}^{>}$ }, where each event represents the situation when one of the four relations between the endpoints takes some particular value. For each event from the set $W_1$ we define the possible values of the relations $\mathbf{r_{11}}$, $\mathbf{r_{12}}$, $\mathbf{r_{21}}$, and $\mathbf{r_{22}}$ (Table 1).

Twelve columns of Table 1 represent 12 events from the set $W_1$. For each case we define possible values of the relations $\mathbf{r_{11}}$, $\mathbf{r_{12}}$, $\mathbf{r_{21}}$, and $\mathbf{r_{22}}$. For example, in the case "$\mathbf{r}_{11}^{<}$" the values of the relations $\mathbf{r_{12}}$, $\mathbf{r_{21}}$, and $\mathbf{r_{22}}$ are "<", "?", and "?" correspondingly. The question mark "?" means that the value of the relation is a disjunction of the three basic relations. The dependency cases are derived from the definition of temporal interval.

**Table 1. Dependencies between $\mathbf{r_{11}}$, $\mathbf{r_{12}}$, $\mathbf{r_{21}}$ and $\mathbf{r_{22}}$**

| | $\mathbf{r}_{11}^{<}$ | $\mathbf{r}_{11}^{=}$ | $\mathbf{r}_{11}^{>}$ | $\mathbf{r}_{12}^{<}$ | $\mathbf{r}_{12}^{=}$ | $\mathbf{r}_{12}^{>}$ | $\mathbf{r}_{21}^{<}$ | $\mathbf{r}_{21}^{=}$ | $\mathbf{r}_{21}^{>}$ | $\mathbf{r}_{22}^{<}$ | $\mathbf{r}_{22}^{=}$ | $\mathbf{r}_{22}^{>}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mathbf{r_{11}}$ | | | | ? | > | > | < | < | ? | ? | ? | ? |
| $\mathbf{r_{12}}$ | < | < | ? | | | | < | < | ? | < | < | ? |
| $\mathbf{r_{21}}$ | ? | > | > | ? | > | > | | | | ? | > | > |
| $\mathbf{r_{22}}$ | ? | ? | ? | ? | > | > | < | < | ? | | | |

The important concept which is used further in the text is the notion of conditional probability. The conditional probability P(A|BCD) denotes the probability of the event A, calculated under the assumption that the events B, C, and D occurred. If the event A does not depend on any of these events, then the conditional probability P(A|BCD) is transformed into an ordinary probability P(A) of the event A.

To calculate the probabilities of Allen's relations we need to know the probabilities of the relations between the endpoints. Since we are dealing with a system of dependent relations, the probability of the relation between two endpoints should take into account the possible values of three other endpoint relations. In other words, the probability of each of the relations $\mathbf{r_{11}}$, $\mathbf{r_{12}}$, $\mathbf{r_{21}}$, and $\mathbf{r_{22}}$ is the conditional probability of this relation under some particular values of the other three relations. These particular values are the combinations that correspond to Allen's interval relations.

Now, let us compose the formula for the probability P(eq) of the Allen's relation "equals". The probability of this relation is a multiplication of the four conditional probabilities of the relations $\mathbf{r_{11}}$, $\mathbf{r_{12}}$, $\mathbf{r_{21}}$, and $\mathbf{r_{22}}$ when the events $\mathbf{r}_{11}^{=}$, $\mathbf{r}_{12}^{<}$, $\mathbf{r}_{21}^{>}$, and $\mathbf{r}_{22}^{=}$ from the set $W_1$ take place simultaneously:

$$\text{P(eq)} = \mathbf{P}\left(\mathbf{r}_{11}^{=} \middle| \mathbf{r}_{12}^{<}\mathbf{r}_{21}^{>}\mathbf{r}_{22}^{=}\right) \cdot \mathbf{P}\left(\mathbf{r}_{12}^{<} \middle| \mathbf{r}_{11}^{=}\mathbf{r}_{21}^{>}\mathbf{r}_{22}^{=}\right) \cdot$$

$$\cdot \mathbf{P}\left(\mathbf{r}_{21}^{>} \middle| \mathbf{r}_{11}^{=}\mathbf{r}_{12}^{<}\mathbf{r}_{22}^{=}\right) \cdot \mathbf{P}\left(\mathbf{r}_{22}^{=} \middle| \mathbf{r}_{11}^{=}\mathbf{r}_{12}^{<}\mathbf{r}_{21}^{>}\right) = \mathbf{e}_{11}^{=} \cdot \mathbf{e}_{22}^{=}.$$

According to Table 1, the value of the relation $\mathbf{r_{11}}$ does not depend on the events $\mathbf{r}_{12}^{<}$, $\mathbf{r}_{21}^{>}$, and $\mathbf{r}_{22}^{=}$. This means that the conditional probability $\mathbf{P}\left(\mathbf{r}_{11}^{=} \middle| \mathbf{r}_{12}^{<}\mathbf{r}_{21}^{>}\mathbf{r}_{22}^{=}\right)$ is transformed into an ordinary probability $\mathbf{P}\left(\mathbf{r}_{11}^{=}\right)$, which is equal to the probability value $\mathbf{e}_{11}^{=}$ from the matrix $\hat{A}$. Similarly, $\mathbf{P}\left(\mathbf{r}_{22}^{=} \middle| \mathbf{r}_{11}^{=}\mathbf{r}_{12}^{<}\mathbf{r}_{21}^{>}\right) = \mathbf{e}_{22}^{=}$. In the conditional probability $\mathbf{P}\left(\mathbf{r}_{12}^{<} \middle| \mathbf{r}_{11}^{=}\mathbf{r}_{21}^{>}\mathbf{r}_{22}^{=}\right)$ the value of the relation $\mathbf{r}_{12}$ depends on the event $\mathbf{r}_{11}^{=}$, and can only take the value "<". Therefore, the probability $\mathbf{P}\left(\mathbf{r}_{12}^{<} \middle| \mathbf{r}_{11}^{=}\mathbf{r}_{21}^{>}\mathbf{r}_{22}^{=}\right)$ is equal to 1. Similarly, the conditional probability $\mathbf{P}\left(\mathbf{r}_{21}^{>} \middle| \mathbf{r}_{11}^{=}\mathbf{r}_{12}^{<}\mathbf{r}_{22}^{=}\right)$ is also equal to 1. In a similar way we can compose the probabilities of other Allen's relations (Figure 6).

$$\mathrm{P(eq)}=\mathbf{e}_{11}^{=}\ \mathbf{e}_{22}^{=};$$

| | |
|---|---|
| $\mathrm{P(b)}=\mathbf{e}_{11}^{<}\ \mathbf{e}_{21}^{<};$ | $\mathrm{P(bi)}=\mathbf{e}_{11}^{>}\ \mathbf{e}_{12}^{>};$ |
| $\mathrm{P(m)}=\mathbf{e}_{11}^{<}\ \mathbf{e}_{21}^{=};$ | $\mathrm{P(mi)}=\mathbf{e}_{11}^{>}\ \mathbf{e}_{12}^{=};$ |
| $\mathrm{P(d)}=\mathbf{e}_{11}^{>}\ \mathbf{e}_{12}^{<}\ \mathbf{e}_{22}^{<};$ | $\mathrm{P(di)}=\mathbf{e}_{11}^{<}\ \mathbf{e}_{21}^{>}\ \mathbf{e}_{22}^{>};$ |
| $\mathrm{P(o)}=\mathbf{e}_{11}^{<}\ \mathbf{e}_{21}^{>}\ \mathbf{e}_{22}^{<};$ | $\mathrm{P(oi)}=\mathbf{e}_{11}^{>}\ \mathbf{e}_{12}^{<}\ \mathbf{e}_{22}^{>};$ |
| $\mathrm{P(s)}=\mathbf{e}_{11}^{=}\ \mathbf{e}_{22}^{<};$ | $\mathrm{P(si)}=\mathbf{e}_{11}^{=}\ \mathbf{e}_{22}^{>};$ |
| $\mathrm{P(f)}=\mathbf{e}_{11}^{>}\ \mathbf{e}_{12}^{<}\ \mathbf{e}_{22}^{=};$ | $\mathrm{P(fi)}=\mathbf{e}_{11}^{<}\ \mathbf{e}_{21}^{>}\ \mathbf{e}_{22}^{=}.$ |

**Figure 6. Probabilities of Allen's relations.**

Further we prove that the sum of the above probabilities is equal to 1, since Allen's relations are the only thirteen basic relations that can hold between two intervals.

Let the set of all possible basic events for Allen's relations between intervals be $\Omega_2$ ={eq,b,bi,m,mi,o,oi,d,di,s,si,f,fi}, where each element of the set defines the situation when a particular relation holds between the intervals. According to Definition 2 (Section 2), the sum of the probability values $\mathbf{e}$ from the vectors from the matrix Â is equal to 1:

$$\mathbf{e}_{11}^{<}+\mathbf{e}_{11}^{=}+\mathbf{e}_{11}^{>}=1,\ \mathbf{e}_{12}^{<}+\mathbf{e}_{12}^{=}+\mathbf{e}_{12}^{>}=1,\ \mathbf{e}_{21}^{<}+\mathbf{e}_{21}^{=}+\mathbf{e}_{21}^{>}=1,$$
$$\text{and }\ \mathbf{e}_{22}^{<}+\mathbf{e}_{22}^{=}+\mathbf{e}_{22}^{>}=1.$$

The sum of the probabilities of Allen's relations is transformed by taking into account the above equations:

$$\sum_{\mathbf{a}\in\Omega_2}\mathbf{P(a)}=\mathbf{e}_{11}^{=}\ \mathbf{e}_{22}^{=}+\mathbf{e}_{11}^{<}\ \mathbf{e}_{21}^{<}+\mathbf{e}_{11}^{>}\ \mathbf{e}_{12}^{>}+\mathbf{e}_{11}^{>}\ \mathbf{e}_{12}^{<}\ \mathbf{e}_{22}^{<}+$$
$$+\mathbf{e}_{11}^{<}\ \mathbf{e}_{21}^{>}\ \mathbf{e}_{22}^{>}+\mathbf{e}_{11}^{<}\ \mathbf{e}_{21}^{>}\ \mathbf{e}_{22}^{<}+\mathbf{e}_{11}^{>}\ \mathbf{e}_{12}^{<}\ \mathbf{e}_{22}^{>}+\mathbf{e}_{11}^{<}\ \mathbf{e}_{21}^{=}+$$
$$+\mathbf{e}_{11}^{>}\ \mathbf{e}_{12}^{=}+\mathbf{e}_{11}^{=}\ \mathbf{e}_{22}^{<}+\mathbf{e}_{11}^{=}\ \mathbf{e}_{22}^{>}+\mathbf{e}_{11}^{>}\ \mathbf{e}_{12}^{<}\ \mathbf{e}_{22}^{=}+\mathbf{e}_{11}^{<}\ \mathbf{e}_{21}^{>}\ \mathbf{e}_{22}^{=}=$$
$$=\mathbf{e}_{11}^{=}(\mathbf{e}_{22}^{=}+\mathbf{e}_{22}^{<}+\mathbf{e}_{22}^{>})+\mathbf{e}_{11}^{>}\ \mathbf{e}_{12}^{<}(\mathbf{e}_{22}^{<}+\mathbf{e}_{22}^{>}+\mathbf{e}_{22}^{=})+$$
$$+\mathbf{e}_{11}^{<}\ \mathbf{e}_{21}^{>}(\mathbf{e}_{22}^{>}+\mathbf{e}_{22}^{<}+\mathbf{e}_{22}^{=})+\mathbf{e}_{11}^{<}\ \mathbf{e}_{21}^{<}+\mathbf{e}_{11}^{>}\ \mathbf{e}_{12}^{>}+\mathbf{e}_{11}^{<}\ \mathbf{e}_{21}^{=}+$$
$$+\mathbf{e}_{11}^{>}\ \mathbf{e}_{12}^{=}=\mathbf{e}_{11}^{=}+\mathbf{e}_{11}^{>}\ \mathbf{e}_{12}^{<}+\mathbf{e}_{11}^{<}\ \mathbf{e}_{21}^{>}+\mathbf{e}_{11}^{<}\ \mathbf{e}_{21}^{<}+\mathbf{e}_{11}^{>}\ \mathbf{e}_{12}^{>}+$$
$$+\mathbf{e}_{11}^{<}\ \mathbf{e}_{21}^{=}+\mathbf{e}_{11}^{>}\ \mathbf{e}_{12}^{=}=\mathbf{e}_{11}^{=}+\mathbf{e}_{11}^{>}(\mathbf{e}_{12}^{<}+\mathbf{e}_{12}^{>}+\mathbf{e}_{12}^{=})+$$
$$+\mathbf{e}_{11}^{<}(\mathbf{e}_{21}^{>}+\mathbf{e}_{21}^{<}+\mathbf{e}_{21}^{=})=\mathbf{e}_{11}^{=}+\mathbf{e}_{11}^{>}+\mathbf{e}_{11}^{<}=1.\ \bullet$$

In Section 7 we will consider an example of using the proposed formulas when estimating Allen's relations between indeterminate temporal intervals in the temporal database. Before that, let us consider an example of the behavior of the proposed estimation mechanism regarding computational complexity, which we present in the next section.

## 6. COMPUTATIONAL COMPLEXITY

In this section we present the results of the simulation of the proposed estimation mechanism. The time required to calculated the probabilities of Allen's relations was measured under different input data. We study how the number of chronons within the intervals for the endpoints and the overlapping of these intervals affect the calculation time.

The calculation time mainly depends on the computational power of the processor that is used, but it also depends on the input data. Our estimation mechanism uses the values of the endpoints of the intervals $[\mathbf{s}_1^{\mathbf{l}},\mathbf{s}_1^{\mathbf{u}}]$, $[\mathbf{e}_1^{\mathbf{l}},\mathbf{e}_1^{\mathbf{u}}]$, $[\mathbf{s}_2^{\mathbf{l}},\mathbf{s}_2^{\mathbf{u}}]$, $[\mathbf{e}_2^{\mathbf{l}},\mathbf{e}_2^{\mathbf{u}}]$, and the p.m.f.s for these intervals as input data. The algorithm of estimation is so that, firstly, the four relations $\mathbf{r}_{11}$, $\mathbf{r}_{12}$, $\mathbf{r}_{21}$, and $\mathbf{r}_{22}$ between the endpoints $\mathbf{s}_1$, $\mathbf{e}_1$, $\mathbf{s}_2$, and $\mathbf{e}_2$ are estimated, and then the probabilities of Allen's relations are calculated using the values from the obtained vectors. The second part of the estimation does not depend on input data and is always performed in a fixed amount of time. Hence, the estimation of the four relations between the endpoints is the main time consuming operation, which depends on the input data. The time required to estimate each of the four relations depends on two factors:

- The number of chronons within the intervals $[\mathbf{s}_1^{\mathbf{l}},\mathbf{s}_1^{\mathbf{u}}]$, $[\mathbf{e}_1^{\mathbf{l}},\mathbf{e}_1^{\mathbf{u}}]$, $[\mathbf{s}_2^{\mathbf{l}},\mathbf{s}_2^{\mathbf{u}}]$, $[\mathbf{e}_2^{\mathbf{l}},\mathbf{e}_2^{\mathbf{u}}]$. Generally, the more chronons there are within these intervals, the more time is needed to estimate the relations between the endpoints, and hence between the indeterminate intervals.
- The number of common chronons within the intervals for the endpoints $\mathbf{s}_1$ and $\mathbf{s}_2$, $\mathbf{s}_1$ and $\mathbf{e}_2$, $\mathbf{e}_1$ and $\mathbf{s}_2$, and $\mathbf{e}_1$ and $\mathbf{e}_2$. This factor stems from the composition of formulas presented in Section 4, according to which we calculate the joint probability only for those possible values of the endpoints that are located inside the common part of the intervals. This means that the more common chronons two intervals for the endpoints have, the more time is needed to calculate the probabilities of the basic relations between them.

The mechanism for estimation of Allen's relations between indeterminate temporal intervals was implemented using the "C++ Builder 4.0" programming environment. Using the developed software several tests were conducted, where the time required to calculate the probabilities of Allen's relations was measured. A desktop computer Pentium II 330 MHz with 64 Mb of memory was used to measure and collect the results of the simulation.

Four series of tests with 10 cases in each series were conducted. The series are distinguished by the number of chronons the intervals for the endpoints: 100, 1 000, 5 000, and 10 000. For simplicity we assumed that all four intervals for the four endpoints include the same number of chronons. The different cases represent 10 different situations of overlapping of the intervals for the endpoints, which are distinguished by the percentage of common chronons in these intervals. These percentages take the values from 10 to 100. So, the case "50" in the series "100" means that the intervals for the endpoints consist of 100 chronons each and overlap with 50% of the common chronons. For example, the indeterminate intervals A[1-100, 501-600] and B[51-150, 551-650] satisfy this condition. The relations between them is estimated in about 160 milliseconds. The calculation time for all series and cases are presented in Table 2.

**Table 2. Time (sec.) required to estimate Allen's relations**

| | 100 | 1 000 | 5 000 | 10 000 |
|---|---|---|---|---|
| 10 | 0,06 | 0,55 | 13,7 | 50,15 |
| 20 | 0,06 | 1,87 | 57,7 | 202,29 |
| 30 | 0,06 | 4,67 | 110,79 | 456,76 |
| 40 | 0,11 | 7,96 | 193,78 | 798,73 |
| 50 | 0,16 | 12,69 | 326,59 | 1266,37 |
| 60 | 0,22 | 18,04 | 418,75 | 1822,26 |
| 70 | 0,27 | 24,93 | 565,24 | 2477,74 |
| 80 | 0,33 | 32,18 | 742,41 | 3241,76 |
| 90 | 0,44 | 40,92 | 940,88 | 4099,15 |
| 100 | 0,5 | 50,75 | 1143,71 | 4657,51 |

Figure 7 illustrates the obtained results and presents the surface graph along three dimensions: percentage of common chronons, number of chronons, and the calculation time (sec.).

It is quite obvious that the percentage of common chronons is the most influencing factor on the calculation time. The obtained value for the series "100" is about 835 times less than that for the series "10000" in case "10", at the same time the difference between these series in case "100" is about 9315 times.

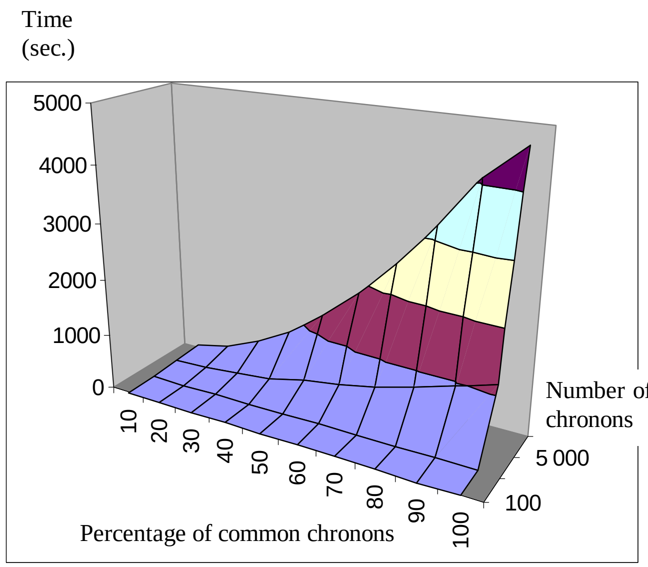


**Figure 7. Calculation time depends on the number and the percentage of common chronons.**

The described above simulation does not pretend to be an exhaustive study of the computational complexity of the estimation algorithm, but it provides some hints to understanding the behavior of the proposed mechanism.

## 7. EXAMPLE

This section presents an example of the estimation of Allen's relations between indeterminate temporal intervals in the temporal database. Let us consider a temporal database from a tube manufacturing plant and a database from a warehouse, which supplies some materials to the plant. Table 3 presents a fragment of the database from the plant with a granularity of 1 day.

**Table 3. Production of tubes by the plant**

| Series of tubes | Production period | Defective tubes (%) |
|---|---|---|
| 1020 | 1~2 Jan — 5~13 Jan | 11 |
| 1021 | 7~8 Jan — 19~21 Jan | 2 |
| 1022 | 9~11 Jan — 22~30 Jan | 3 |

Each produced series of tubes has a unique identity number included in the column "Series". The "Production period" column includes the indeterminate temporal intervals that define the production periods for series. For example, the production of the series 1020 started between the 1-st and 2-th of January, and ended between the 5-th and 13-th of January. The column "Defective tubes" includes the percentage of defective tubes for each produced series.

The warehouse delivers steel blanks to the plant, that are used in production. It keeps a database with delivery records, fragments of which are presented in Table 4.

**Table 4. Delivery of steel blanks by the warehouse**

| Series of blanks | Delivery period |
|---|---|
| 100 | 4~7 Jan —12~15 Jan |
| 101 | 10~11 Feb —25~26 Feb |
| 102 | 1~2 Feb —10~14 Feb |

Each series of steel blanks has a unique identity number included in the column "Series of blanks". The "Delivery period" column includes the indeterminate temporal intervals that define the delivery dates for each series of blanks. The starting point of an interval defines the date when the series was sent from the warehouse. The end point of the interval defines the date when the series of blanks arrived at the plant.

When the percentage of defective tubes for a series is more than 10 we are interested in which series of steel blanks could be used in the production of this series of tubes. To answer the question we need to estimate the temporal relation between the production of tubes date and the delivery of blanks date. We suppose that blanks could be used in production if they have arrived at the plant before or on the last day of the production. The temporal relations "after", "overlapped-by", "met-by", and "finishes" satisfy this assumption. Using the approach proposed in this paper we can calculate the probabilities of these Allen's relations. The sum of the calculated probabilities is the probability that the particular series of blanks could be used in the production of the particular series of tubes.

The series of tubes 1020 from Table 3 has a percentage of defective tubes equal to 11. Let us estimate, for example, the relation between the production of the tubes 1020 and the delivery of the series of blanks 100 supplied by the warehouse. The relation between the indeterminate intervals **A**[4∼7,12∼15] and **B**[1∼2,5∼13] defining the delivery period for the blanks and the production period for the tubes respectively is represented in Figure 2 (Section 2).

The values of the p.m.f.s for the indeterminate points $\mathbf{s_1}$ (the starting point of **A**) and $\mathbf{e_2}$ (the end point of **B**) are presented at Figures 3 and 4 (Section 3). Let us suppose in this example that the values of the p.m.f.s for the point $\mathbf{e_1}$ (the end point of **A**) are the same as for $\mathbf{s_1}$. We also suppose that two chronons included in the period of indeterminacy for the point $\mathbf{s_2}$ (the starting point of **B**) are equally probable, which means that the values of the p.m.f. $\mathbf{f_3}(\mathbf{s_2})$ are: $\mathbf{f_3}(1)=0.5$ and $\mathbf{f_3}(2)=0.5$.

Using the formulas for the probabilities of the basic relations between two indeterminate points we derive the relations $\mathbf{r_{11}}$, $\mathbf{r_{12}}$, $\mathbf{r_{21}}$, and $\mathbf{r_{22}}$, the relational matrix

$$\hat{A} = \begin{bmatrix} (0,0,1)_{\mathbf{r}_{11}} & (0.8500,\ 0.0750,\ 0.0750)_{\mathbf{r}_{12}} \\ (0,0,1)_{\mathbf{r}_{21}} & (0.0125,\ 0.0875,\ 0.9)_{\mathbf{r}_{22}} \end{bmatrix}_{\mathbf{A},\mathbf{B}},$$

and the probabilities of the Allen's relations: P(bi)=0.0750, P(mi)=0.0750, P(oi)=0.7650, P(d)=0.0106, P(f)=0.0744. The probabilities of all other Allen's relations are equal to 0. The sum of the probabilities P(bi), P(oi), P(mi), and P(f) is equal to 0.989375, which means that the steel blanks 100 were very probably used in the production of the tubes 1020.

## 8. RELATED RESEARCH

Many published research articles deal with imperfect information. Various approaches to this problem are mentioned in the bibliography on uncertainty management by Dyreson [5], in the surveys by Parsons [9] and by Parsons and Hunter [10], although not many of them consider temporal imperfection. Formalisms intended for dealing with imperfection are often distinguished as symbolic and numerical. Among the numerical approaches the most well known are probability theory, Dempster-Shafer's theory of evidence [13], possibility theory [4], and certainty factors [14]. In this paper we adopted a probabilistic approach.

Representation and reasoning with uncertain temporal relations between points was discussed by van Beek [17] and by van Beek and Cohen [16]. They represented an uncertain temporal relation as a disjunction of the three basic relations, similar to ourselves. Also algorithms for reasoning with an uncertain relation were proposed and their complexity studied. However, no numerical means for estimating temporal relation was provided, and temporal points were supposed to be determinate only, which rarely happens in practical applications.

The probabilistic representation of uncertain relations was studied also by Ryabov et al. in [12], where an algebra for reasoning with uncertain relations was proposed. The algebra includes three operations: negation, composition, and addition, which make it possible to derive unknown relations in a relational network combining already known uncertain relations. That approach can be used together with the one proposed in this paper, where we stressed the deriving of the uncertain relations between indeterminate temporal points and intervals.

Barbara et al. [2] have proposed a Probabilistic Data Model (PDM) intended to represent in a database entities whose properties can not be deterministically classified. The approach, however, is applied to relational databases and does not discuss explicitly the management of imperfect temporal information. That paper focuses exclusively on discrete probability distribution functions, but the authors claim that the approach can be extended to continuous probabilities.

Dyreson and Snodgrass [6] proposed a mechanism supporting valid-time indeterminacy in temporal databases, which can be seen as an extension of PDM. They represent indeterminate temporal points similarly to ourselves in this paper, although their main stress was on the development of a query language.

Our paper concentrates on the description of different situations when estimating uncertain relations between indeterminate temporal points and intervals. The proposed mechanism can be used in a query language that supports temporal indeterminacy using probabilities, for example, TSQL2 [15], and hence we did not conceive as a goal the development of a new query language.

# 9. CONCLUSIONS

In this paper we proposed one way to estimate relations between indeterminate temporal intervals. Using the information about the intervals for the indeterminate endpoints and p.m.f.s for those intervals we derived the relational matrix and then used it to calculate the probabilities of Allen's relations. The approach assumes full knowledge about the distributions inside the intervals of values for the indeterminate endpoints, which does not happen very often in real applications. In the case when the distribution is unknown, the uniform distribution is a useful assumption, because we have no reason to favor one chronon over another. As one direction for further research we consider the development of some formal means for specifying the distribution using available indirect information about it.

The probabilistic approach that was adopted in the paper is actually one of the means for handling uncertainty, as well as possibility theory, Dempster-Shafer theory, and numerous logical approaches. The method for handling uncertainty was selected reflecting our goals of having numerical measures of uncertainty, and a solid mathematical background behind the method. The probabilistic approach has also a close relation to statistics which potentially can be used as one of the means of defining the p.m.f. by analyzing the past data.

Generally, the proposed representation can be applied to continuous time model. Certainly there are some domains where the continuous time model is more natural, but more applications are those, where the discrete representation is used. The discrete case is also given more attention in the temporal databases area, one of the important application areas for temporal representation and reasoning.

# 10. ACKNOWLEDGMENTS

This work was supported by COMAS graduate school at the University of Jyväskylä.

# 11. REFERENCES

[1] Allen, J., "Maintaining Knowledge about Temporal Intervals", *Communications of the ACM,* (26), 832-843 (1983).

[2] Barbara, D., H. Garcia-Molina, and D. Porter. "The Management of Probabilistic Data", *IEEE Transactions on Knowledge and Data Engineering,* 4(5), 487-502 (1992).

[3] Dey, D., and S. Sarkar, "A Probabilistic Relational Model and Algebra", *ACM Transactions on Database Systems,* 21(3), 339-369 (1996).

[4] Dubois, D., and H. Prade, *Possibility Theory: An Approach to the Computerized Processing of Uncertainty*, Plenum Press, New York (1988).

[5] Dyreson, C., "A Bibliography on Uncertain Management in Information Systems", In: Motro, A., ed., *Uncertainty Management in Information Systems: From Needs to Solutions,* Kluwer Academic Publishers, 415-458 (1997).

[6] Dyreson, C., and R. Snodgrass, "Supporting Valid-Time Indeterminacy", *ACM Transactions on Database Systems,* 23(1), 1-57 (1998).

[7] Kwan, S., F. Olken, and D. Rotem, "Uncertain, Incomplete and Inconsistent Data in Scientific and Statistical Databases", *Proc. Second Workshop on Uncertainty Management and Information Systems: From Needs to Solutions,* Catalina, USA (1993).

[8] Motro, A., "Sources of Uncertainty in Information Systems", *Proc. Second Workshop on Uncertainty Management and Information Systems: From Needs to Solutions*, Catalina, USA (1993).

[9] Parsons, S., "Current Approaches to Handling Imperfect Information in Data and Knowledge Bases", *IEEE Transactions on Knowledge and Data Engineering,* 8(3), 353-372 (1996).

[10] Parsons, S., and A. Hunter, "A Review of Uncertainty Formalisms", In: Hunter, A., and S. Parsons, eds., *Applications of Uncertainty Formalisms. Lecture Notes in Artificial Intelligence*, 1455, Springer-Verlag, 8-37 (1998).

[11] Ryabov, V., "Probabilities of Allen's Relations between Indeterminate Temporal Intervals", *Proc. 8-th Finnish Artificial Intelligence Conference,* Helsinki, Finland, 61-68 (2000).

[12] Ryabov, V., S. Puuronen, and V. Terziyan, "Representation and Reasoning with Uncertain Temporal Relations", *Proc. Twelfth International Florida AI Research Society Conference,* AAAI Press, Orlando, Florida, USA, 449-453 (1999).

[13] Shafer, G., "A Mathematical Theory of Evidence", Princeton University Press, USA (1976).

[14] Shortliffe, E., "Computer-Based Medical Consultations: MYCIN", Elsevier, New York (1976).

[15] Snodgrass, R., I. Ahn, G. Ariav, D. Batory, J. Clifford, C. Dyreson, R. Elmasri, F. Grandi, C. Jensen, W. Käfer, N. Kline, K. Kulkarni, T. Cliff Leung, N. Lorentzos, J. Roddick, A. Segev, M. Soo, M. Suryanarayana, and S. Spirada, *The TSQL2 Temporal Query Language*, Kluwer Academic Publishers (1995).

[16] van Beek, P., "Temporal Query Processing with Indefinite Information", *Artificial Intelligence in Medicine,* 3, 325-339 (1991).

[17] van Beek, P., and R. Cohen, "Exact and Approximate Reasoning about Temporal Relations". *Computational Intelligence,* (6), 132-144 (1990).