

Kamil Matoušek and Zdeněk Kouba

*The Gerstner Laboratory, Department of Cybernetics
Faculty of Electrical Engineering, Czech Technical University in Prague
Technická 2, CZ-166 27 Prague 6, Czech Republic
{matousek,kouba}@labe.felk.cvut.cz*

This paper describes an intelligent approach to represent information including uncertainly specified time periods by software agents in medical environment. Temporal reasoning capabilities are based on uncertainty ontology modeling and generalized Allen's interval relations. An example of temporal reasoning capabilities in data annotated by means of the temporal ontology is given.

1. INTRODUCTION

Current multi-agent medical systems range from hospital theatre scheduling systems to intelligent decision support systems. A survey of such existing systems can be found in (Foster, 2005). In order to perform their diagnostic and other useful tasks, the typical activities often include examination of patient's history for previous medical treatments, or gathering medical data together from different hospitals and other data sources.

Records of patients collected by medical specialists represent a significant source of data to be included in a digital knowledge base utilized by the agents. For knowledge sharing, ontology is a typical means. In practice, most current patient records are in form of text files containing some structured data such as body weight, blood pressure etc., but a significant part of the report consists of notes made by the physician.

In order to use the data for agent-based decision support among different medical laboratories and hospitals, it is necessary to convert as much information as possible into a structured form with added semantics. This can be achieved by annotation of medical records using ontologies and subsequent knowledge extraction. To capture as much detail from the patients' records as possible, several ontologies are necessary.

As an example, ontology describing the structure of medical records in the Institute of Biology and Medical Genetics at Charles University in Prague is being developed. OCML modeling language (Motta, 1999) is used for defining inference rules in the ontology modeling process. One of its parts, the *task ontology*, captures

the structure of records of examinations relevant for patients with neurofibromatosis. In addition, domain ontologies have to be used, e.g. *ontology of diseases*, *ontology of drugs* etc. in order to enrich the defined tasks with formalized concepts of the medical domain. In case of inherited diseases also an ontology describing *family relations* from genetic point of view was necessary.

It has been discovered that in addition to medical domain ontologies, an ontology capturing the imprecise and uncertain information contained in the medical records could significantly improve representational capabilities.

Existing medical standards introduce e.g. *temporal concept* in UMLS (Unified Medical Language system) as “a concept which pertains to time or duration”, or HL7 standard includes the relation *temporally related to*. However, these descriptive elements are on a very high abstraction level and neglect the interesting detailed semantics of the specific terms.

2. CATEGORIES OF TEMPORAL STATEMENTS

Patients are located in space and time, while living in contexts and periods of their individual lives. Tracking changes of patients' health during time has always been a method that promised answering questions of professional physicians.

Concerning individual expressions of time, there is a wide range of precise, imprecise or vague, uncertain dating, which cause difficulties and inaccuracy in their use. Some reasons of inaccuracy of objects' dating may be e.g. the following:

- Data is not available (i.e. no written resources),
- it is a subjective information given by the patient,
- processes that lasted for a longer time are referred to as a single time point (e.g. feeling a pain).

In the domain of time, statements like “in childhood”, “soon after the surgery” or “during the treatment in hospital XY” are corresponding examples. Our contribution is an attempt to deal with uncertainty in temporal assertions. The goal is to suggest a suitable and effective inference mechanism extension for ontology of medical records, which would yield sufficiently accurate localization in time.

Concerning people, the main temporal properties correspond to their lives and some events they encountered. Major time property is the duration of a time period (e.g. *convalescence period*, *pregnancy*), which could be expressed in terms of starting and ending time points. However, time durations may be relative as well (e.g. *for three months*) and thus having no exact starting or ending time. Time properties are not often expressed in a straightforward way as they may be inherent in the data. In this case, there are many expressions with different semantics (e.g. *tomorrow*, *at the beginning of the year*, *Monday*, or *June 5th*).

A categorization of temporal statements containing the most frequent expressions with respect to their accuracy with utility for the domain of our interest has been proposed in (Matoušek, 2004):

1. Precise statements. The whole data is available, maximum precision is reached (e.g. *January 12, 2006, 12:30:00*).
2. Statements with higher granularity. Data is available, but not so precise. It is necessary to distinguish instants and intervals (e.g. *January 12, 2006* can be seen either as an instant of higher granularity or as a 24 hour time interval).
3. Incomplete statements. Some information is missing for precise time identification. One may intentionally use this kind of statement for recurring temporal positions – regularly repeated instants (e.g. *January 12, 12:30:00*).
4. Uncertain statements with absolute specification of uncertainty (e.g. *between February 12 and February 13, 2006*).
5. Uncertain statements with relative specification of uncertainty (e.g. *around February 12, 2000, before 90s*).
6. Statements referencing other statements with temporal properties (e.g. *after the first surgery, during the last in-hospital treatment*).
7. Statements with unknown or missing information (e.g. *that time I was doing ...*).

Expressions related to the current time (e.g. yesterday, tomorrow) are supposed to implicitly belong to the category 6 above.

3. TEMPORAL SUPPORT

In order to support reasoning in the uncertain temporal domain some aspects including relations between time points and intervals, temporal granularity, temporal uncertainty and combining granularity values with other uncertainty factors have to be modeled. The following subsections describe them.

3.1 Time Relations

There are thirteen possible and mutually exclusive relations between two time intervals called *Allen's interval relations* introduced in (Allen, 1983). For graphical representation of the thirteen Allen relations see Figure 1 (moving from left to right corresponds to the passing of time).

For the intended inference capabilities, the supported time point (instant) relations are necessary. They are obvious but have to be stated. They include *before*, *equals* and *after* relations.

3.2 Time Granularity

Time granularity is a concept of the level of detail in which the time is considered (measured). Different time statements can refer to different time granularities. E.g. “May, 12, 2003” is a time statement with day granularity, while “In 2002” has the year granularity. With coarser time granularities their granularity values are defined, too. Time granularity defines its own unit scale for time positions.

Let us show an example of a relation between two time points with granularity (t^g). This is done as an extension of Allen's interval relationships for two time points with different granularity:

$$t_1^{g1} \text{ precedes } t_2^{g2} \quad \text{iff } \text{End}(t_1^{g1}) \text{ before } \text{Start}(t_2^{g2})$$

For time points with granularity common relations used to compare the finest time points were presented in (Matoušek, 2004).

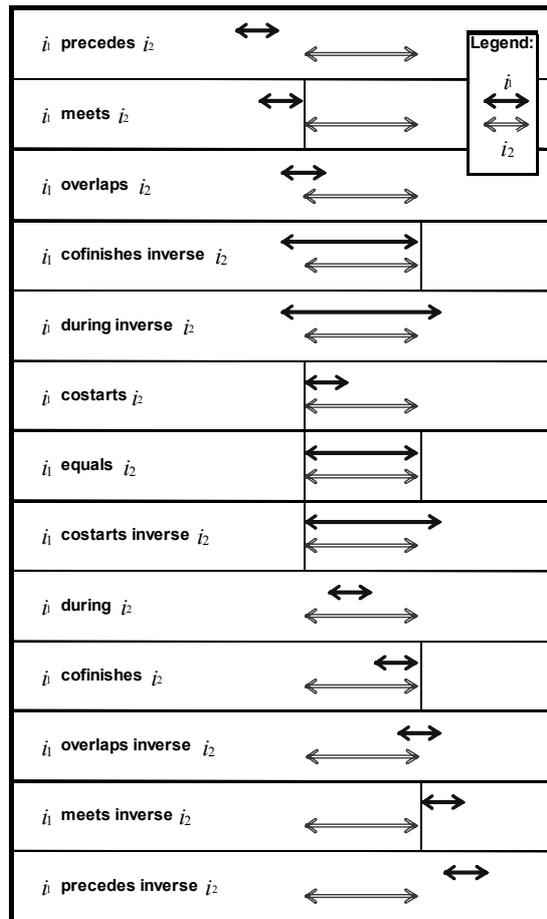


Figure 1 – Allen's temporal relations of time intervals

3.3 Temporal Position Uncertainty

As we have seen, there are various types of indeterminate expressions like: *now*, *(some time) before*, *(some time) after*, *at the break of a decade*, *during*, or even *unknown*.

We introduce *Time Uncertainty* as a concept, whose values represent individual uncertainty types falling into time intervals. Other approaches may consider e.g. also

uncertainty in recurring temporal entities. The properties of time uncertainty are *FromTimePoint*, *ToTimePoint*, *BeforeRelTime* and *AfterRelTime*. The supported representations of time uncertainty are:

For *absolute* specification of a time range of uncertainty, given by specified time points, we introduce uncertainty as a property pair of two time points: *FromTimePoint* and *ToTimePoint*.

For *relative* time range of uncertainty, there is another pair of relative temporal position properties and their respective granularities: *BeforeRelTime*, *AfterRelTime*, and *BeforeGranularity*, *AfterGranularity*.

To clarify the usage of this time uncertainty representation, let us give here some examples:

Temporal statements like “*around the year 2004*”, “*after 3rd of June 1991*”, will use in its representation the *BeforeRelTime* and *AfterRelTime* together with uncertain location properties. On the other hand if an event happened in some time between “*The first examination and the last manifestation of a specific symptom*”, then *FromTimePoint* and *ToTimePoint* will be used.

For uncertain time points we cannot easily define common relations used to compare the finest time points like before, equals, after. We can only express their sufficient conditions corresponding to certainty or necessary conditions corresponding to possibility. These possible and necessary bounds fulfill the rule that if one time point is *certainly before* another, then also the former is *possibly before* the latter.

Of course, an analogous property holds for the remaining pairs: *<certainly equal, possibly equal>* and *<certainly after, possibly after>*.

3.4 Combining Time Uncertainty and Granularity

As it is natural to combine uncertainty and granularity specifications, we support uncertain time points with granularity, which have specified both granularity and uncertainty.

An example of the possibility expression which is valid follows:

$$\begin{aligned} &{}^u t_1^{g1} \text{ possibly equals } {}^u t_2^{g2} \\ &\text{iff not} (\text{Start}({}^u t_2^{g2}) \text{ after } \text{End}({}^u t_1^{g1})) \\ &\quad \text{and } (\text{not } \text{Start}({}^u t_1^{g1}) \text{ after } \text{End}({}^u t_2^{g2})) \end{aligned}$$

where ${}^u t^g$ denotes an uncertain time point with uncertainty u and granularity g . The uncertainty types under consideration are the *absolute* and *relative* time ranges of uncertainty (see section 3.3) and time granularity has been introduced in section 3.2.

For uncertain time points with granularity common relations used to compare the finest time points i.e. *before* and *after* expressed again in terms of their sufficient and necessary conditions can be derived e.g.:

$$\begin{aligned} &{}^u t_1^{g1} \text{ certainly before } {}^u t_2^{g2} \\ &\text{iff } {}^u t_1^{g1} \text{ certainly precedes } {}^u t_2^{g2} \end{aligned}$$

The possible and necessary bounds of common comparison relations fulfill the rule that if one time point is *certainly after* another, then also the former is

possibly after the latter. The same property holds for the pairs *<certainly equal, possibly equal>* and *<certainly after, possibly after>*.

4 ONTOLOGICAL MODEL OF UNCERTAIN TIME

In this section we describe how the framework for temporal reasoning proposed in (Matoušek, 2004) can be practically used. We describe a calendar and time system that was used in our prototype of a temporal inference engine. We also provide a short description of a temporal reasoning system implemented in OCML language. It enables users to enter calendar data and to carry out temporal reasoning on a knowledge base containing temporal definitions. The temporal reasoning engine builds on the inference capabilities of OCML.

The basic class diagram of the time ontology in our temporal inference engine is depicted in the Figure 2.

The classes in the right frame correspond to the definitions in the framework for reasoning in the time domain. The main components of our temporal model are subclasses of *temporal-entity* – *time-point* and *time-interval*. The property *timeline-of* of *temporal-entity* is a sort of stereotype, which enables distinguishing different kinds of temporal entities. Any instance of temporal entity can have any number of timelines assigned to it. Query results can be constrained by including the name of a timeline that is the subject of interest in the query submitted to the temporal inference engine. Thus, timelines define a kind of namespaces. Beside classes, the model contains basic time point and time interval relations, rules, and functions that handle both granularity and uncertainty introduced as part of the framework.

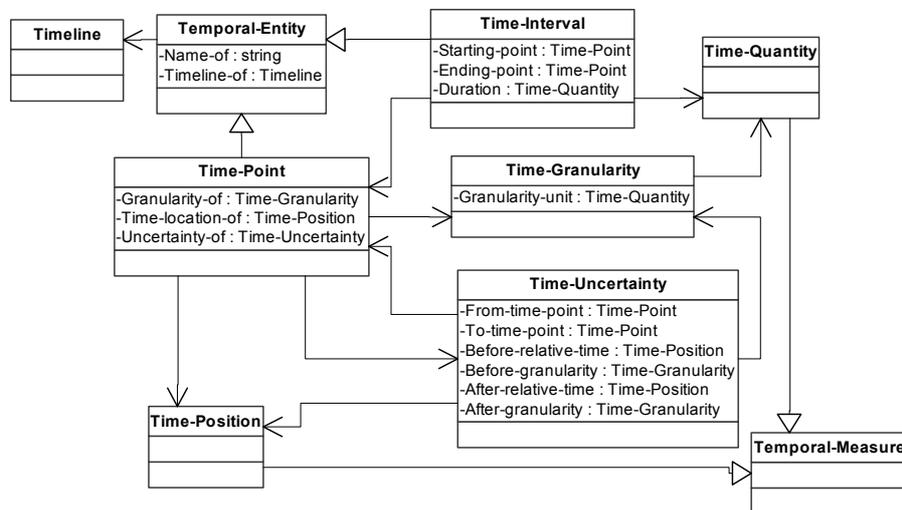


Figure 2 – Temporal ontology: core concepts

4.1 Calendar Dates and Time

In order to support calendar date and time specification, we introduce a subclass of *time-point*, denoted as *calendar-time-point* with the system of classes around it in the Figure 3.

The slots *century-of*, *year-of*, *month-of*, *date-of*, *week-day-of*, *hour-of*, *minute-of*, and *second-of* enable representation of individual parts of calendar date and time. The prototype inference engine provides internal routines, which are used to recalculate these slots into the internal representation of time locations and conversely back from the time location into these slots. When filling out only the slots with the respective information available or applicable, incomplete date specifications or those on a coarser granularity level can be used.

Some general constraints should always be satisfied, when working with temporal entities. One example is the property of transitivity of functions *before* and *equals*.

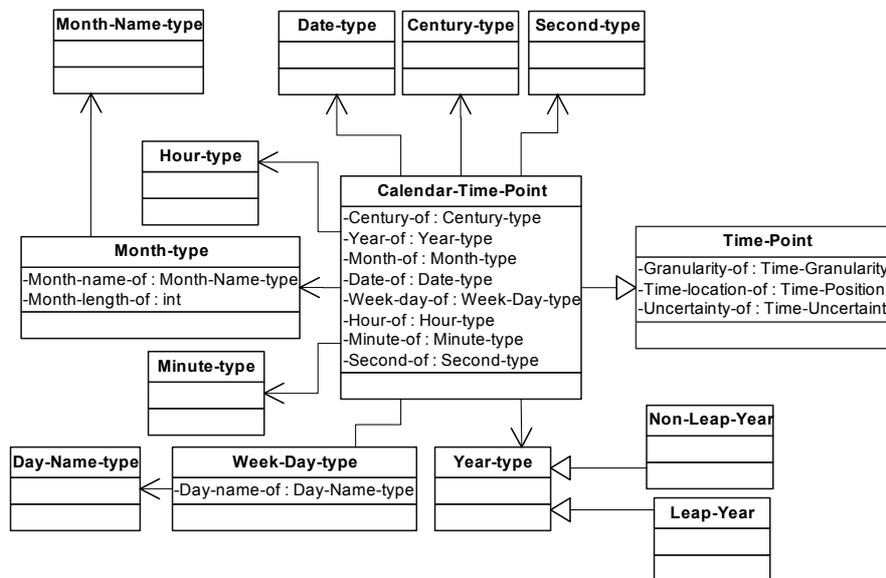


Figure 3 – Temporal ontology: calendar concepts

4.2 Domain Modeling and Inference Support

The ontology modeling style in OCML (Motta, 1999) enables to declaratively introduce classes and instances. More general relations can be also used as labeled n-ary relationships between OCML entities. Classes are considered as strictly unary relations (or predicates) and their slots (or attributes) as binary relations.

For example, the class *Time-point* is a subclass of *Temporal-entity*, which can have specified at most one time location, assigned at most one granularity type and one uncertainty type:

```
(def-class Time-Point (Temporal-Entity)
  ((Time-location-of :max-cardinality 1
    :type Time-Position)
   (Granularity-of :max-cardinality 1
    :type Time-Granularity)
   (Uncertainty-of :max-cardinality 1
    :type Time-Uncertainty)))
```

Moreover, procedural functions and backward chaining rules can be used to perform some necessary calculation or to introduce additional rules for inference. E.g. *get-location* function can be used in part of the *Before* relation definition:

```
(def-relation Precedes (?int1 ?int2)
  :constraint (and (Time-Interval ?int1)
                  (Time-Interval ?int2))
  :iff-def (and (Ending-Point ?int1 ?end)
                (Starting-Point ?int2 ?start)
                (Before ?end ?start)))
```

The proof system in OCML is capable to answer queries about the model using an algorithm for formula satisfaction with backward chaining like in Prolog (Derensart, 1996), i.e. depth-first search with chronological backtracking.

Practically, in our domain of interest, when comparing temporal entities, the relevant finest time points are being compared as necessary with respect to the given time point definitions including granularity and uncertainty. This enables the inference mechanism to properly reason about the individual temporal statements.

5. A SIMPLE PRACTICAL EXAMPLE

Let us show how the temporal inference engine can be used. We shall model the life and illness periods of a fictitious person, J. Hoffman. At the beginning the important time points have to be defined:

```
(def-instance John-Hoffman-birth Calendar-Time-point
  ( (date-of 14) (month-of 5) (year-of 1916)
    (granularity-of day-granularity)))

(def-instance John-Hoffman-illness-start Calendar-Time-point
  ( (date-of 26) (month-of 8) (year-of 1946)
    (granularity-of day-granularity)))

(def-instance John-Hoffman-illness-end Calendar-Time-point
  ( (date-of 29) (month-of 11) (year-of 1978)
    (granularity-of day-granularity)))
```

The following time intervals are related to the specified time points:

```
(def-instance Illness-John-Hoffman Time-interval
  ( (starting-point John-Hoffman-illness-start)
    (ending-point John-Hoffman-illness-end)))

(def-instance Life-John-Hoffman Time-interval
  ( (starting-point John-Hoffman-illness-start)
    (ending-point John-Hoffman-death)))
```

If a statement is inaccurate, we may need to create an instance of time uncertainty. Here, we model the moment John Hoffman's knee got injured, which happened sometime around the year 1970. It is possible to include uncertainty parameterization as in the following case e.g. using *param-around-unc* instance of a newly defined class *time-parameter*:

```
(def-instance param-around-unc time-parameter((value-of 10)))

(def-instance Around-a-Year Time-Uncertainty
  (
    (Before-relative-time param-around-unc)
    (Before-granularity year-granularity)
    (After-relative-time param-around-unc)
    (After-granularity year-granularity)))

(def-instance JH-Knee-Injury Calendar-Time-point
  (
    (timeline-of John-Hoffman)
    (year-of 1970)
    (granularity-of year-granularity)
    (uncertainty-of around-a-year)))
```

Having defined the necessary facts, we may be interested in particular results using the inference engine. For example, having consulted data concerning all patient examinations and surgeries performed, the following query will retrieve the important time points between the first and the second surgery, which took place in the years 1965 and 1980 respectively:

```
(ocml-eval
  (findall ?a
    (and (timeline-of ?a John-Hoffman)
         (and (before JH-first-surgery ?a)
              (after JH-second-surgery ?a))))))
```

The returned result then finds (possibly among others) the John Hoffman's knee injury:

```
(JH-Knee-Injury)
```

6. CONCLUSIONS

Current results are giving useful means for agent's representation of even uncertain temporal aspects in medical patient's records including the uncertain facts. Using the proposed solution the medical ontology used by intelligent software agents was augmented and the first example results have been shown.

As a recommendation we suggest to extend the existing means of agent communication (e.g. KQML language performatives) to include the OCML language linkage in order to enable the exploitation of the presented solution. Alternatively, re-implementation in other reasoning frameworks is also possible.

In future the presented temporal ontology could be extended to cover a broader variety of patient-relevant data. Additional effort can be concerned with medical standard concepts mapping.

7. ACKNOWLEDGMENTS

The research is supported by the research grant of the Czech Ministry of Education, Youth and Sport: *Transdisciplinary Research in the Area of Biomedical Engineering II*, No. MSM6840770012.

8. REFERENCES

1. Allen JF. Maintaining Knowledge about Temporal Intervals. *Communications of the ACM*, 26 (11) : 832-843, 1983.
2. Deransart P, Ed-Dbali A, Cervoni L. *Prolog: The Standard*. Springer, 1996.
3. Foster D, McGregor C, El-Masri S. A Survey of Agent-Based Intelligent Decision Support Systems to Support Clinical Management and Research. In *Proceedings of the 1st International Workshop, MAS*BIOMED'05*, 2005, pp. 16-34.
4. Matoušek K, Uhlíř J. On Representing Uncertain Historical Time. *Proceedings of DEXA 2004*. Los Alamitos: IEEE Computer Society Press:105-109, 2004.
5. Motta E. *Reusable Components for Knowledge Modeling*. IOS Press, 1999.