Technical report

Temporal Representation and Reasoning for the Semantic Web

Ubbo Visser & Sebastian Hübner

TZI-Bericht Nr. 28, 2003

Organisation:

Ubbo Visser, TZI – Center for Computing Technologies, University of Bremen

Technologie-Zentrum Informatik

Technical Report 28

Temporal Representation and Reasoning for the Semantic Web

Ubbo Visser & Sebastian Hübner

TZI-Bericht Nr. 28
2003

Universität Bremen

TZI-Berichte

Herausgeber:
Technologie-Zentrum Informatik
Universität Bremen
Universitätsallee 21-23
28359 Bremen
Telefon: +49-421-218-7272
Fax: +49-421-218-7820
Email: info@tzi.de
http://www.tzi.de

Temporal Representation and Reasoning for the Semantic Web

Ubbo Visser and Sebastian Hübner

July 18, 2003

Abstract

This paper is a technical report about a new temporal representation scheme that is designed to serve the Semantic Web. It comes along with a temporal reasoning service. We discuss related work in this area, define our representation and inference mechanisms, and discuss implementation issues. We bring examples and conclude with a general discussion and some ideas for further developments.

1 Motivation

The number of web sites has increased drastically over the past few years. Currently, there are billions of web pages supplying information to users. Modern technology (e.g. multiagent systems) seem to support users with their requests while they are browsing through the web automatically, returning answers. However, the vast amount of web pages are unstructured or weakly structured, which makes it impossible for machines to understand the semantics of the content. The idea of the Semantic Web helps at this point: information sources should be annotated with metadata following some kind of formalization. Thus, machines are able to "understand" the meaning of the information sources and can deliver more accurate answers.

The Bremen Semantic Translator for Enhanced Retrieval (BUSTER) follows and supports this idea. It is an ontology-based prototype that helps applications or users to (a) *find* the needed information and (b) *integrate* and/or *translate* this information for further processes. Queries can be formulated to seek concepts using description logic-based reasoning services. This allows users to type in queries with some kind of "sloppiness", i.e., using common and everyday words to describe a concept (e.g., "forest"). The reasoning engine connected with ontologies can use the inherent inference mechanisms to derive appropriate answers. This kind of approach is not new but is included in the prototype. In addition, BUSTER allows for the search of *place names* such as "Weserbergland", which are commonly used in conversations but found nowhere in digital GIS. The combination of both concept and location queries lead to a new type of query, namely *concept@location*. Now, the user is able to formulate queries like "*Which hotels are in the Weserbergland?*". An appropriate reasoning engine based on connection graphs has been developed and partly implemented.

Another important part of a search is time-dependent: people are looking for hotels in areas for a certain period of time (e.g., during summer vacation) but do not want to specify time according to the user-unfriendly W3C standard. Therefore, we have developed a new time representation and a new reasoning engine based on Allen's time intervals and

Freksa's semi-intervals. This leads us to another type of query, namely *concept@location in time*.

This paper provides insight into a new temporal representation scheme that is designed for the use for the Semantic Web. The scheme comes along with a new reasoning service. This technical report gives also examples where necessary and concludes with a summary and references to future work.

2 Related Work

In this section, we will address related work that was completed in the area of qualitative temporal representation and reasoning.

2.1 Approaches for temporal representation and reasoning

Before we start presenting a picture about existing approaches in this line of research, we would like to discuss the basics about the presentation of time. A profound source of this is the catalog of temporal theories, which has been written by (Hayes 1996). The following is based in this compendium, except the summary of recent approaches.

Hayes introduces six meanings of time in his catalog of temporal theories. The first, and surely the most important one, sees time as a *physical dimension*, along with other physical dimensions such as voltage and length. The second meaning of time is what he called the *universe* of time, sometimes referred to as time line or time-plenum. The idea is that there is a endless discrete time stream. The third idea is based on pieces of time, also called *time-intervals*. An example of this is a time interval, which covers the rowing event at the last Olympic games. Another notion of time is that of a *point of time*. Here, we discuss a moment in the time continuum. While researchers still argue about the duration of a moment, we will postpone this discussion for now and go on to the fifth meaning of time: *duration*. An example of this is the amount of time needed to take a shower or get to work. The last notion of time is described as a position in a *temporal coordinate system*, such as June, 21st, 2003 or 5:15pm.

(Hayes 1996) argues that these time concepts have clear relationships to each other and can in fact be defined in various ways. Some theories follow the idea of taking time points as primitives, others are based on time intervals. The relation between points and intervals is important for the following, hence, we discuss this in more detail.

One view is that intervals *are* time points. These intervals are obviously as short as possible and thus, do not contain any sub-intervals (which is usually possible). They cannot overlay each other and do not have an internal structure. A colloquial term for this is the concept *moment*.

Another view is that there is an time continuum. This implies, that there is no such thing as a moment. The idea behind this is described in (Allen 1984), who also illustrates the problem of meeting intervals. If two intervals meet, which interval "inherits" the meeting point? In fact, is it possible at all to decide whether a point belongs to the first or second interval? This is a relevant topic, since a number of temporal approaches are based on points as primitive objects. These approaches further define intervals as a set of points. The other view is to use points to locate positions in or between intervals, which themselves are primitive objects.

(Hayes 1996) concludes that it is impossible to divide an interval exactly symmetrically in half following the first notion of time. This implies that there must be open and closed intervals. The second intuition does allow this, however, rejects the conclusion that the meeting (or split) point is contained in either half.

Language expressiveness

When describing time concepts, various *languages* can be used. These languages must cover temporal relations, allow propositions whose truth values might vary, and describe concepts whose properties might change over time.

One way to describe time is to use the concepts time themselves as objects. These objects can then be used in axioms depicting time to other things. An example for this is the following:

```
(submitted Ubbo_Visser PhD-Thesis 1995)
```

Another way to describe time ensures that sentences are 'true' at certain times. The following sentence states that it is true that I held a lecture on Artificial Intelligence 1 in Fall 2002.

```
(is_true (has_lecture Ubbo_Visser Artificial_Intelligence_1) Fall_2002)
```

Some theories use tenses. Tense logics extent usual logics by modal operators which allow to state that certain relations hold true in the past or in the future. Here is an example describing that I received my doctorate some time in the past (without saying when exactly).

```
(Past (has_received Ubbo_Visser Doctorate))
```

The final consideration with respect to language are temporal knowledge bases. The key behind this is that a language is imbedded in a temporal framework allowing to keep track of changes in the world and drawing inferences. The main problem here is to ensure consistency with the environment changing.

Following, we will give an overview about time point-based theories and interval-based theories. This subsection is partly based on (Hübner 2003).

2.2 Temporal theories based on time points

The temporal theories used in the approaches that we describe in the following are mostly consistent with the ideas stated by (Hayes 1996, p. 13). A time interval is a piece of the time line, has a unique temporal extent, consists of two end points and is uniquely determined by these. Also, a time point can be uniquely determined by the extent of the interval between this point and some temporal position which we call 'zero'.

However, it is also possible to use other structures, which also rely on time points. Using computers implies some restrictions on the temporal theory. In order to distinguish between variations of time point structures (discrete vs. continuous, bounded vs. unbounded, linear vs. branched), we need to define the used terms.

Therefore, the elementary time points and the existing precedence relation $\prec$ are formalized. This relation is partially ordered, hence, transitivity (2.1) and irreflexivity (2.2) hold true.

$$\forall x, y, z[(x \prec y \wedge y \prec z) \rightarrow (x \prec z)] \tag{2.1}$$

$$\forall x \, \neg(x \prec x) \tag{2.2}$$

A time point structure is therefore an ordered pair $\langle X, \prec\rangle$ based on a non-empty set of time points X and a precedence relation $\prec$.

The mentioned variations, which are based on point structures, can be defined through axioms. Whether the time is bounded or not, for instance, is dependent on the existence or non-existence of a start or end point (2.3-2.7). A combination (restricted or bounded in one direction only) is also possible and can be useful.

$$\exists x_a \neg \exists x (x \prec x_a) \tag{2.3}$$

$$\exists x_e \neg \exists x (x_e \prec x) \tag{2.4}$$

$$\forall x \, \exists x' (x' \prec x) \tag{2.5}$$

$$\forall x \, \exists x' (x \prec x') \tag{2.6}$$

A discrete time model allows us to determine the direct neighbors on both sides of a non-marginal point (2.7,2.8). This model is isomorphic to natural numbers $\mathbf{N}$. A dense time, on the other hand, is isomorphic to the rationals $\mathbf{Q}$ – where another point exists between pairwise disjunct time points (2.9)(cf. (Hayes 1996, p. 17)).

$$\forall x_1 \, [\exists x_2 (x_2 \prec x_1) \rightarrow \exists x_3 (x_3 \prec x_1 \wedge \neg \exists x_4 (x_3 \prec x_4 \wedge x_4 \prec x_1))] \tag{2.7}$$

$$\forall x_1 \, [\exists x_2 (x_1 \prec x_2) \rightarrow \exists x_3 (x_1 \prec x_3 \wedge \neg \exists x_4 (x_1 \prec x_4 \wedge x_4 \prec x_3))] \tag{2.8}$$

$$\forall x_l x_r \, [x_l \prec x_r) \rightarrow \exists x_m (x_l \prec x_m \wedge x_m \prec x_r)] \tag{2.9}$$

The notion of a one-dimensional, deterministic time line is described with the ordering axiom (2.10). There are no branches and the time points are totally ordered.

$$\forall x \, x' \, (x \prec x' \vee x = x' \vee x' \prec x) \tag{2.10}$$

Another notion is the one with a branching tree in one direction (e.g., future 2.11) Here, we only can compare time points if they are directly on the time line without being in the branch. The idea behind this is the indeterminism of potential future (or past) situations that can take place from the actual situation.

$$\forall xyz \, [(y \prec x \wedge z \prec x) \rightarrow (y \prec z \vee y = z \vee z \prec y)] \tag{2.11}$$

Point structures are therefore a model whose properties can be mathematically exactly defined.

2.3 Temporal theories based on intervals

Human beings tend to formulate time with the help of intervals. These time intervals to a certain extent have interval structures as their underlying models. It is not necessary to have intervals only with exact same lengths, however, they must be non-empty, which basically means that start and end point are not the same. Again, axioms can be used to define the properties of these structures. The precedence relation is also partially ordered, hence, transitivity (2.1) and irreflexivity (2.2) hold true. In addition, we need a part-of relation $\subseteq$, which includes the identity and is therefore not a real part-of relation. Hayes calls this relation *inclusion* that has the properties transitivity (2.12), reflexivity (2.13), and anti-symmetry (2.14).

$$\forall x, y, z[(x \subseteq y \wedge y \subseteq z) \rightarrow (x \subseteq z)] \tag{2.12}$$

$$\forall x\,(x \prec x) \tag{2.13}$$

$$\forall x, x'[(x \subseteq x' \wedge x' \subseteq x) \rightarrow (x = x')] \tag{2.14}$$

We can therefore define an interval structure with the ordered triple $\langle X, \subseteq, \prec \rangle$, with the interval X, the inclusion $\subseteq$, and the precedence $\prec$.

Whether the time described by intervals is bounded or unbounded, dense, discrete, continuous etc. is similar to the properties of time point structures. However, the axiom describing $before$ can be interpreted in different ways: a time interval (including end point) is fully before another time interval or it overlaps partially. This leads us to the definition of overlapping (2.15) which we can use to define the precedence relation (2.16).

$$\forall x, y[(x \cap y := \exists z(z \subseteq x \wedge z \subseteq y)] \tag{2.15}$$

$$\forall x, x'[(x \prec x') \rightarrow \neg(x \cap x')] \tag{2.16}$$

We can now transform the axioms 2.3 and 2.4 (earlier/later time point exists) and the axioms 2.6 and 2.7 (earlier/later time point do not exist) to interval structures. Because overlapping includes identity, we can define the ordering relation according to axiom 2.10, using $\cap$ instead of $=$.

$$\forall x\, x'\,(x \prec x' \vee x \cap x' \vee x' \prec x) \tag{2.17}$$

Considering the density or discreteness of the time model we have to take into account that intervals can include other intervals (inclusion) but no gaps. The latter needs another axiom which can be described as convexity axiom (2.18).

$$\forall x, y, z[(x \prec y \wedge y \prec z) \rightarrow \forall z'[(z' \subseteq x \wedge z' \subseteq z) \rightarrow (z' \subseteq y)]] \tag{2.18}$$

In summary, we can derive two demands with regard to the model: intervals can be infinitely divided into smaller intervals (time line is dense or continuous) or we have to deal with small but non-dividable intervals.

We can see that properties of time point structures and time interval structures can be described with similar axioms.

2.4 Summary of recent approaches

Temporal representation and reasoning is an essential feature in any activities that involve changes. This explains, why temporal representation and reasoning services are so important and appear in so many areas, including planning, natural language understanding, and knowledge representation.

Recent articles describe approaches in the area of *Temporal Constraint Programming*, an important area of temporal reasoning (Schwalb and Vila 1998; Gennari 1998). Gennari describes a temporal reasoning system as a temporal knowledge base. It also contains a procedure to check its consistency, and inference mechanisms, which are able to derive new information and get a solution or all solutions to queries. Temporal reasoning tasks are mainly formulated as constraint satisfaction problems; therefore, the constraint satisfaction techniques can be used to check consistency, to search for solutions or all solutions for the given problem.

Events are the primitive entities in the knowledge base. They are characterized in temporal constraint programming by means of their time of occurrence, which can be given by time points or intervals (see above).

Temporal information can constrain events to happen at a particular time (e.g., "Coffee time is at 3:30 pm") or to hold during a time interval (e.g., "A class lasts 90 minutes"); moreover it can state relations between events of a qualitative type (e.g., "$Event_1$ is before $Event_2$") or of a metric one (e.g., "$Event_1$ has started at least three hours before $Event_2$").

Constraints can be either extensionally characterized by real or rational numbers, or intensionally represented as (finite) sets or relations of some algebra (e.g., Allen's interval algebra (Allen 1984)). According to the formalization of constraints and the time unit chosen, the approaches can be classified into three main streams.[1]:

- *Temporal reasoning with metric information:* In the quantitative approach to temporal reasoning with constraints, variables $X_1, \ldots, Xn$ range over real or rational numbers. Originally finite sets of real intervals, constraints are lately represented by unions of interval-sets. A temporal constraint is explicitly given as a set of intervals $I_1 \cup \ldots \cup I_n$ where $I_i = [l_i, r_i]$. The constraints can be unary or binary and are represented by $\{I_1, \ldots, I_n\} = \{[l_1, r_1], \ldots, [l_n, r_n]\}$. An unary constraint T_i restricts the domain of a variable X_i to the given set of intervals. Thus, it is represented by the disjunction $(l_1 \leq X_i \leq r_1) \vee \ldots \vee (l_n \leq X_i \leq r_n)$. The binary constraint T_{ij} restricts the values for the distance of the variables $X_j - X_i$ and represents the disjunction $(l_1 \leq x_j - x_i \leq r_1) \vee \ldots \vee (l_n \leq x_j - x_i \leq r_n)$ (Dechter, Meiri, and Pearl 1991). The authors assume that all the intervals are pairwise disjoint.

  Constraint propagation algorithms are based on metric properties of the continuous variable domain. Since the satisfiability problem of general temporal constraints is NP-hard, research if focussed on particular classes of temporal constraint problems such as single temporal constraint problems, backtracking algorithms, and constraint propagation algorithms in order to achieve local consistency or at least a good approximation of local consistency (e.g., (Schwalb and Dechter 1993)).

  In principle, these methods can be used for reasoning services on the Semantic Web. However, the adaptation for their use implies a large modelling effort.

- *Qualitative approaches based on Allen's interval algebra:* The most fundamental and well-known theory about reasoning with time intervals has been formulated by (Allen 1984). This approach has been revised over the years and is based on interval structures, which are used as primitives.[2]

  Allen motivates his approach with the problem that much of our temporal knowledge is *relative*, and hence cannot be described by a date (or even a fuzzy date). As Allen further argues in his paper, his framework is particularly designed for these reasons:

---

[1] Other authors such as (Schwalb and Vila 1998) and (Vila 1994) describe these three main streams as *metric point* (for metric information), *qualitative point* and *qualitative interval* (for qualitative approaches based on Allen's interval algebra), and combinations (for mixed approaches).

[2] There is a difference to the intervals described above since those intervals are composed by time points. Here, time intervals are primitives.

- it allows "significant imprecision": much temporal knowledge is relative and sometimes it has no relation to absolute dates;
- "uncertainty of information" can be represented by means of disjunctions of relations between two intervals;
- because of the qualitative representation of constraints one has a certain freedom when modelling knowledge and can choose the grain of reasoning, for instance expressing time in terms of days, weeks or business-days;
- the reasoning engine allows for *default reasoning* of the type "If I parked my car in lot A this morning, then it should still be there now".

In Allen's framework, variables range over real or rational valued intervals. Constraints are specified as unions of *atomic* (*basic*) relations, which are pairwise disjoint. Variables represent time intervals and the basic temporal relations are

$$Temporal\ relations = \left\{ \begin{array}{c} before, after, meets, met\_by \\ overlaps, overlaps\_by, during, contains, equals \\ starts, started\_by, finishes, finished\_by \end{array} \right\}$$

The class of all possible unions of the atomic relations forms a boolean algebra, Allen's interval algebra. There are 13 atomic relations and thus 2^{13} relations in total. Checking consistency for this algebra turned out to be NP-hard. Allen introduces a path-consistency algorithm to deal with the problems that propagates relations between intervals by means of composition. The algebra consists of $2^{13} = 8192$ relations which means that there are 2^{8192} possible subsets in that algebra, which make them intractable. Therefore, research in that area is concentrating on *tractable* and recently *maximal tractable* subalgebras. Some of the most important subalgebras of Allen's interval algebra are obtained by "translating" metric point relations into Allen relations. This means that there have to be languages to describe sets of qualitative or quantitative relations between points, and that these have to be translated in tractable subalgebras.

An exhaustive search by computers is a key technique to prove the maximality of the algebras that up to now have been discovered; this machine case analysis was firstly introduced by (Nebel and Bürckert 1995). A different approach to this problem in a geometric and not a logic apparatus, is given in Ligozat's work (Ligozat 1998; Ligozat 1996). Some of the studied subalgebras are the Point Algebra (Vilain and Kautz 1986; Beek 1992) and the NB algebra (Nebel and Bürckert 1995). To compute a solution, backtracking search is used. It has been shown that the search gets more effective with the additional use of path-consistency checking such as a forward-checking method within the backtracking algorithm (Schwalb and Vila 1998).

These mentioned arguments hold true also for the Semantic Web. Thus, interval-based approaches are valuable when discussing methods and techniques for temporal reasoning on the Web.

- *Mixed approach based on metric and qualitative constraints:* In this framework, the other approaches are combined in order to gain expressiveness, while trying not to loose the tractability of the problem; however, the complexity results are not always optimal. The ontological entities in the first approach are time points only, and the primitive entities in the second approach are time intervals. This third approach involves both points and intervals as primitive objects of the language; therefore new relations are introduced in order to "relate" time points and time intervals.

Some authors have studied particular metric temporal constraint problems in order to find new sub-algebras of interval algebra. This can be seen as a qualitative approach because its main goal is an interval algebra. An approach is "mixed" when it aims at using both the expressive power of the qualitative and of the quantitative approaches to create "new" temporal frameworks, of which the satisfiability can be decided in polynomial time. The research in this direction is one of the most promising (Stock 1997), however, the relative literature is still scarce.

2.5 Evaluation of approaches

Most of the representation and reasoning approaches in this area are based on point or interval structures using either composition tables or constraint-based methods. Again, we believe, that, in analogy to the terminological and spatial part, temporal ontologies are needed to meet the requirements of the Semantic Web. The following statements underline this.

There is a need for intuitive temporal names, especially when people are involved querying the Semantic Web. As with spatial terms, people would like to use common words for temporal concepts such as 'Summer vacation 2003' rather than fill in a W3C temporal date format (cf. section 3.1). Further, none of the discussed approaches can meet this demand, therefore, we must develop new methods for this intuitive labelling and construct temporal ontologies.

The approaches that are based on temporal intervals are basically eligible for our purpose, however, the existing methods need an significant extension. One reason for this is that none of the approaches are able to express fuzzy boundaries. An example for a fuzzy boundary is the temporal concept 'middle-age'. Experts argue about the exact time interval belonging to the Middle Ages, however, it is clear that the latest beginning of the Middle Ages is the reign of Karl the Great. Further, another clear disadvantage of the existing approaches is the lack of references to other intervals. It is not possible, e.g., to state that the earliest begin of the Middle Ages was the end of the Westroman Empire, which itself can be dated precisely. Therefore, there is a need to develop more sophisticated tools based on the previously mentioned approaches.

3 Temporal Representation and Reasoning

This section describes the requirements which must to take into account with regard to the annotation and querying of temporal information sources. In the following, we discuss how our qualitative abstraction of time is represented. Temporal relevance is an important feature for the calculation of overlapping time periods with unknown boundaries. This is discussed in the following subsection. We will also describe the development and implementation of new reasoning components and demonstrate the performance of this approach with examples.

The representation and reasoning features described in this chapter are based on the results of a masters thesis (Hübner 2003).

3.1 Requirements

Annotation and retrieval of temporal information should be more flexible, comfortable, and improve situations in practise (e.g., with the help of colloquial terms such as Easter 2003). Both the knowledge engineer and the user should have several options to annotate or retrieve information for their purpose.

3.1.1 Intuitive labelling

The most important requirement is the option to label time intervals with intuitive names. These names should be published and can therefore act as reference intervals for further internal or external use. However, typical country-dependent characters and unusual features have to be considered. We therefore restrict these names using existing standards such as UNICODE (The Unicode Consortium 1996) for characters and the XML standard for names (W3C 2000).

3.1.2 Time interval boundaries

Boundaries of time intervals should be flexible and have therefore various specifications. It is necessary that the boundaries on both sides of a time interval can differ. These different types are *exact*, *fuzzy*, *persistent*, and *unknown*. All possible combinations should be possible.

Exact boundaries of time intervals Exact boundaries represent a known, exact beginning and end. They are therefore the most simple case. An example for an exact boundary is the summer break in school: the vacation in the city of Bremen in 2002 started on the 20th of June and lasted until the 31st of July. The W3C offers a known encoding scheme (W3C 1998), however, this scheme only considers time between the years 1 and 9999 of the Gregorian calendar. If we consider having information sources describing Julius Caesars moves in the years BC, we will have a problem. Therefore, the encoding scheme has to be extended.

Fuzzy boundaries of time intervals There are cases when a boundary is known but cannot be exactly determined. The beginning of an interval can then be described with the "earliest" and "latest" beginning. The same holds true for the end of an interval. This type of boundary can be chosen if more than one "official" opinion about a certain boundary, e.g., if recognized experts opinions differ. This can occur often when using common terms such as the "Middle Ages" and are therefore important. We usually have a good impression of time interval covering the Middle Ages but, we cannot exactly determine the beginning and the end.

Persistent boundaries of time intervals Persistent boundaries can appear if a given boundary is unrestricted, i.e., the interval still exists or the interval is already valid. This type of boundary is necessary for the end of an interval, when an end to the interval is not reached and cannot be determined or estimated. We see this phenomenon in scientific programs: a time interval with a defined beginning and an undefined end. Sending satellites or probes in the universe or carrying out a long-term observation is another typical example. When also note this for the beginning of an interval. We could have a time interval that begins before the *annotated* time period. Instead of using the minimal value for the lower boundary, we can use the persistent type.

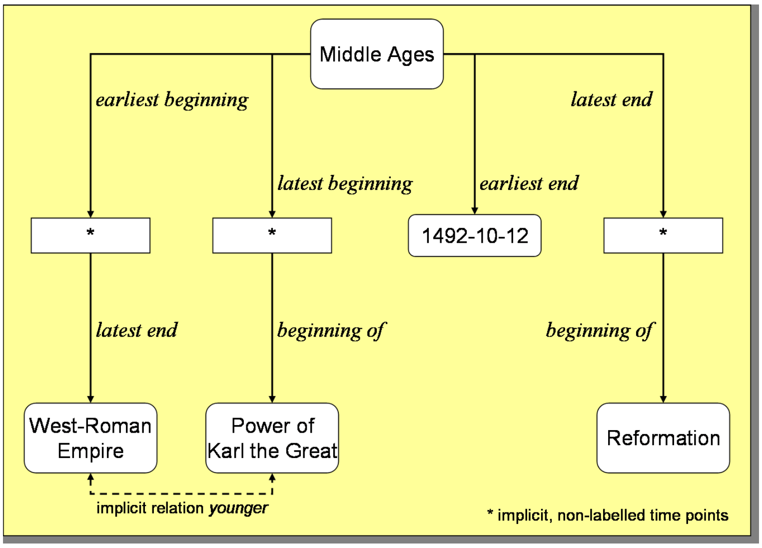


Figure 3.1: Interval structure, after Pitz (2002) and Giesenberg (2002)

Unknown boundaries of time intervals Unknown boundaries are necessary if no dates for the beginning or the end of a time interval are known. With this type of boundary it is also possible to define intervals where only one boundary (either the lower or upper boundary) is known. However, even if both sides are unknown, there is still the option to use this interval for statements about qualitative relations regarding other intervals. The delimitation to fuzzy or persistent boundaries is often not clear and is the discretion of the knowledge engineer. If we know the date of birth of a person but do not know the date of death, the use of an unknown boundary for the end of the time interval is obvious. If on the other hand existing documents (e.g., letters, official notifications) give proof at which time the person was alive and at which time that person died (also documents), we can use fuzzy boundaries. If that person is still alive, a persistent boundary could also be used. An interval with two unknown boundaries is a special case and states basically that there is a time interval only with a given name. If we use this interval with explicit relations (see below) we can make further statements.

3.1.3 Structures

An interval can be based on another interval, can be self-defined or imported. Exact and fuzzy boundaries for the beginning or the end of intervals for instance can be used to determine the exact end of an interval with the help of the beginning of another interval. Time points are used in order to carry out this operation. Therefore, functions are needed to extract these significant time points from the intervals. Examples for these functions are *beginning_of, end_of, earliest_beginning_of, latest_beginning_of, earliest_end_of, latest_end_of.*

An example for the different operations is the time interval "Middle Ages", which historically cannot be exactly determined. However, there are existing events that can be used for the beginning or the end (see figure 3.1). Implicit qualitative relations exist through structures which are build upon each other (see relation *younger* that holds between "West-Roman Empire" and "Reign of Karl the Great"). These implicit relations are at the users disposal, together with the explicit relations, and contain the same expressive power (e.g., transitivity).

3.1.4 Explicit qualitative relations

Making statements about relations between intervals when using persistent or unknown boundaries should also be possible. This can be of value when we do not focus on exact or fuzzy boundaries but need to use the interval for qualitative relations. Consider the following example: firstly, we describe and order historic epoches. Secondly, having described the other intervals such as government times, CVs, travel times etc. using the epoches intervals, we are able to derive temporal relations between the other intervals.

As already mentioned, the Dublin Core Metadata Initiative has made a suggestion for temporal annotation (DCMI Period). The required features however, are only partly covered when using their coverage.Temporal format. Therefore, new concepts and methods must be developed. When comparing qualitative temporal approaches that are based on intervals such as Allen's relations (see section 2.1 on page 2) we see that they require exact boundaries. Intervals with fuzzy, persistent or unknown boundaries are not considered. Also, structures are far more complex with Allen's approach because they can only be implicit and are therefore computational expensive. Allen's time logic can therefore act as a fundamental theory, which partly covers the mentioned requirements.

3.2 Representation

3.2.1 Period names

In the following, we present a new concept which we call *period names*. They allow the qualitative modelling of time and take the mentioned requirements for annotation and retrieval into account. Since we are dealing with annotation and retrieval for the Semantic Web, we use the XML notation to define the concepts and sub-concepts. XML as a description language offers the advantage to use its internal reference system, which is useful for both modelling and implementation. In particular, the construction of period name structures is easier and more comfortable. XML notation is also the basis for the reasoning components. However, we could also use other notations to show the representation (e.g., graphs).

The use of XML is not mandatory, however, we concentrate on this language with regard to the Internet. Therefore, we restrict the language and use the XML standard for names (Bray, Paoli, and Sperberg-McQueen 2000) for our underlying model. This standard requires that XML names consists only of letters and numbers. Special characters such as %, $, & or spaces are not accepted. However, the dot (.), the dash (-) and the underscore (_) are exceptions.

Definition 3.1 (PeriodName)
A period name consists of a header and a body. The header consists of the keyword ***periodName*** *and an attribute id, which labels the name of the period. The body consist of the definition of boundaries and relations.*

Here are two examples for the description of a periodName in XML notation.

Example 3.1

```
a)  <periodName id="Label">
      <!-- Definition of boundaries -->
      ...
```

```
        <!-- Definition of relations -->
        ...
    </periodName>

b)  <periodName id="AntiqueTime"/>
```

3.4 Boundaries

The most important property of a period name is its expansion. The model contains only intervals, which are non-empty and consist of more then one time point. Therefore, the start point must lie before the end point.

The basis of boundaries are period structures, which are constructed intervals using point structures (as described above). These point structures are bounded and discrete. We can assume a continuous time stream with discrete, ordered values. The minimal time unit is exactly one millisecond and all time points can be ordered and compared because of the linearity.

Issues about the accuracy of time intervals, which occur due to the discrete model, must be considered. For instance, we could have information that belongs to a century or year in historic time. Also, information such as months, days or hours that belong to daily news have to be taken into account. Computer interactions require even more accuracy, usually up to seconds or milliseconds. Our model represents time with millisecond accuracy which is also supported by ISO 8601 and W3C-DTF. Even though this level of accuracy is not always necessary, it is not a disadvantage. Fuzzy boundaries for example, can be used to define boundaries where we do not need exact time points based on milliseconds.

Definition 3.2
The temporal range $R^t = [B, E]$ consists of time points between the beginning B and the end E of the range. B is the time point 01.01.9999, 12:00am, 0 seconds and 0 milliseconds B.C. and in the following is denoted by -9999 and E is the time point 31.12.9999, 11:59pm, 59 seconds and 59 milliseconds and in the following is denoted by +9999. The year zero does not exist.

For our definitions, two additional sets are necessary:

Definition 3.3
$\mathcal{P}$ is a set of negative and positive persistent boundaries, $\mathcal{P} = \{\mathcal{P}^-, \mathcal{P}^+\}$.

Definition 3.4
$\mathcal{U}$ is a set of unknown boundaries.

3.4.1 Exact boundaries

Exact boundaries are used if a time interval has a known or exactly defined expansion. Starting points and ending points are defined by exactly one time point. The definition can be accomplished in four different ways:

Definition 3.5
Start and end points are defined explicitly by single time points $t_{begin} \in [B, E]$ and $t_{end} \in [B, E]$ with $t_{begin} < t_{end}$. A time point is defined by a millisecond. t_{begin} describes the start of a period and t_{end} the end of that period.

Lemma 3.1
Both time points are included, thus, the shortest time period is two milliseconds ($t_{begin} + 1 = t_{end}$).

The following example in XML notation describes a meeting on January 16, between 10 and 10.30am.

Example 3.2 (Meeting on January 16, between 10 and 10.30am)

```
<periodName id="Meeting">
  <begin>
    2003-01-16A10:00:00.000-00:00
  </begin>
  <end>
    2003-01-16A10:30:00.000-00:00
  </end>
</periodName>
```

Definition 3.6
Start and end points are defined by another existing time period. The start and end point can be single time points $t_{begin} \in [B, E]$ *and* $t_{end} \in [B, E]$ *or fuzzy boundaries. References and structures which are constructed from these, need the following keywords:* beginOf, endOf, beginfOf, endfOf.

This example denotes that the earliest begin of the Middle Ages is the end of the West-Roman empire.

Example 3.3 (Earliest begin of the Middle Ages is the end of the West-Roman Empire)

```
<periodName id="middle-ages">
  <beginf>
    <endfOf ref="West-Roman_Empire"/>
  </beginf>
</periodName>
```

The actual time is important, especially when formulating a query. Examples are: "the last two weeks" or "the next 24 hours".

Definition 3.7
The keyword **now** *is used for actual time points* $t \in [B, E]$. *'Now' is available with an accuracy of a millisecond and can be combined with the* begin/end-*attribute* offset *to define periods relative to the actual time.*

The following example shows the last minute from an actual time point.

Example 3.4 (Last minute from now on)

```
<periodName id="last_minute">
  <begin offset="-60000">
    <now/>
  </begin>
  <end>
    <now/>
  </end>
</periodName>
```

Relative periods from the actual time are important but are not sufficient enough to describe concepts such as "today" or "this year". Also, periods that occur regularly such as "Easter" or "Christmas" need to be considered. Formulas can be defined to describe these situations.

Definition 3.8
dformula *denotes a formula that returns a certain time point* $t \in [B, E]$*. The return value can be used directly for* begin *or* end*.*

Definition 3.9
pformula *denotes a formula that returns a time period* $t_{begin} < t_{end}$ *with* t_{begin} *and* $t_{end} \in [B, E]$*. pformula can be used only after reference keywords as they represent anonymous periods, which can be referenced as labelled periods.*

The example shows a time period from the beginning of this year until midnight today:

Example 3.5 (Time period from the beginning of this year until midnight today)

```
<periodName id="since_beginning_of_year">
  <begin>
    <beginOf>
      <pformula name="thisyear"/>
    </beginOf>
  </begin>
  <end>
    <dformula name="midnight"/>
  </end>
</periodName>
```

Fuzzy boundaries It is useful not to use exact boundaries while modelling common or colloquial terms. Therefore, we introduce fuzzy boundaries as an extension of exact boundaries and are able to use the already established means for these boundaries: explicit dates, references, now, and formulas.

Definition 3.10
Let $t_{begin} \in [B, E]$ *and* $t_{end} \in [B, E]$ *be the start and end point.* Fuzzy boundaries *consist of two boundaries for both the start and end point.* $t_{beginf} \in [B, E]$ *is the earliest beginning and* t_{begin} *is the latest beginning for that time period. Accordingly,* t_{end} *denotes the earliest ending and* $t_{endf} \in [B, E]$ *the latest ending.*

Lemma 3.2
In addition, the following order holds: $t_{beginf} < t_{begin} < t_{end} < t_{endf}$*.*

Lemma 3.3
The time difference a between t_{beginf} *and* t_{begin} *therefore has the minimum of 1 millisecond. The maximum is arbitrary. The same holds true for the time difference c between* t_{end} *and* t_{endf}*.*

The following example shows the fuzzy boundary "begin of the Middle Ages":

Example 3.6 (Earliest and latest begin of the Middle Ages)

```
<periodName id="begin-middle-ages">
  <beginf>
```

```
    <endfOf ref="West-Roman_Empire"/>
  </beginf>
  <begin>
    <beginfOf ref="Reign_of_Karl_the_Great"/>
  </begin>
</periodName>
```

An extension for references is also needed: we recall the known constructs *beginOf* and *endOf*. They denote the "inner" boundaries (latest begin or earliest end) of a time period. The extension is needed for the "outer" boundaries *beginfOf* and *endfOf* (earliest begin and latest end) of a period. The difference between two time periods, which are defined by exact boundaries and fuzzy boundaries that have the same extent, is the calculation with regard to relevance (see section 3.5).

Figure 3.2 shows a graphical notation of fuzzy boundaries. Three time periods, each with two fuzzy boundaries show that the extent of "fuzziness" (the tolerance or width of the boundaries) can vary arbitrarily. Also, we can see that the outer boundaries of time period A meet B's and C's latest begin. These outer boundaries have referenced boundaries form B and C.

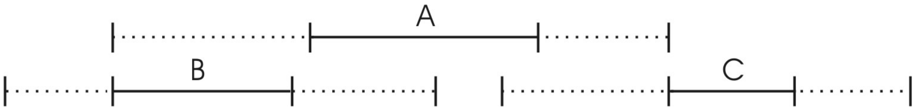


Figure 3.2: Graphical notation of fuzzy boundaries: three time periods with fuzzy boundaries

Persistent boundaries Persistent boundaries are necessary for two reasons: firstly, the start or endpoint of a time interval is before or after the range of the underlying model, i.e., before -9999 and after +9999. Secondly, a time interval could have a known exact or fuzzy beginning but an unknown end (or vice versa), e.g., the end of that time interval does have an open end in the future (long-term experiments). For both cases the keyword 'unlimited' is introduced.

Definition 3.11
P^- defines a boundary that is known or fuzzy, but before the beginning of the range, i.e., $P^- < B$. P^+ defines a boundary that is known or fuzzy but, after the end of the range, i.e., $P^+ > E$. The time point of a persistent boundary $P_t \in \{P^-, P^+\}$ consist of the keyword begin *or* end *followed by the keyword* unlimited *with the value* true *if the beginning or the end of the time interval is known but not in the valid range, i.e.,* $t_{begin} \notin [B, E]$.

The following example shows an interval with two persistent boundaries:

Example 3.7 (A time interval with two persistent boundaries)

```
<periodName id="Label">
  <begin unlimited="true"/>
  <end   unlimited="true"/>
</periodName>
```

A time interval with two persistent boundaries cannot be distinguished from another time interval with two persistent boundaries. Therefore, only combinations with other intervals with other types of boundaries is reasonable.

Unknown boundaries If no information about a time interval is known or the time points are too vague, i.e., even fuzzy boundaries are not reasonable, another type of boundary is necessary: the unknown boundary. It can help for a qualitative modelling and reasoning with regard to other (known) time intervals.

Definition 3.12
An unknown boundary consist of the keyword begin *or* end *followed by the keyword* unknown *with the value* true. *The time point of an unknown boundary* $t \in U$ *is not known. An unknown boundary could be in the valid range* $t \in [B, E]$ *or is part of a persistent boundary* $t \in \{P^-, P^+\}$, *it is simply not known. By default, the boundary is set to unknown.*

The following example shows an interval with two unknown boundaries:

Example 3.8 (A time interval with unknown boundaries)

```
a)  <periodName id="Label">
      <begin unknown="true"/>
      <end   unknown="true"/>
    </periodName>

B)  <periodName id="Label"> </periodName>
```

Figure 3.3 shows the reason for the integration of unknown boundaries: the boundaries of the three time intervals are not known but we can see that qualitative propositions between these intervals do exist. They can therefore be of value for reasoning processes.

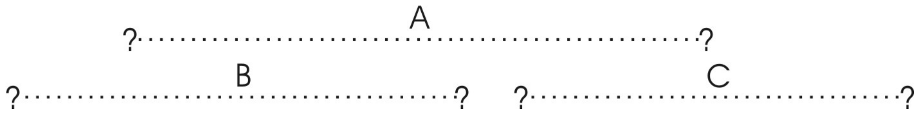


Figure 3.3: Graphical notation of unknown boundaries: three time periods

Combination of boundaries Using the same type of boundary for both start and end of a time interval is not useful. Time intervals with persistent boundaries especially develop their full potential in combination with time intervals having exact or fuzzy boundaries. Therefore, every possible combination of the described types of boundaries can be used while defining a period name. The user can also distinguish between subtypes of fuzzy boundaries such as explicit dates, references, ”now”, and formulas to combine them with the other mentioned options.

3.4.2 Relations

If we use exact boundaries only, implicit relations between time intervals can be defined. A time interval could be completely covered by another time interval, overlap partly or one time interval could lay before the other. (Allen 1984) identified 13 fundamental, distinguishable relations between time intervals. Freksa's critique that these are too exact and would imply too complicated models leads to the model of conceptual neighborhoods (Freksa 1992). He introduced new concepts, which aggregate subsets of Allen's relations. These concepts are not as accurate, but they are easier to calculate with.

We can calculate relations from exact boundaries. We also can do this with fuzzy boundaries if we neglect the transition areas and only consider the outer time points. Therefore, the addition of new relations using these types of boundaries does not provide more information. Furthermore, it can only lead to redundancies or, even worse, to inconsistencies.

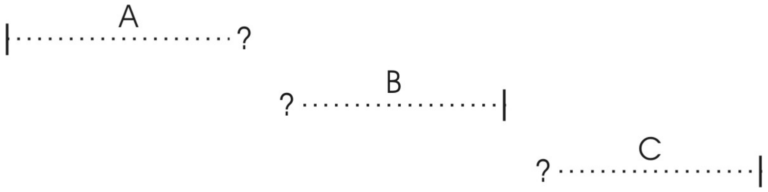


Figure 3.4: Explicit relations

However, new relations when dealing with single unknown boundaries or completely undetermined time intervals are important information sources. Consider the situation in figure 3.4: there are three time intervals, each with one known start or end time point. This leads to various sets of possible relations and we can assume that the relation between each pair is undetermined. Between A and B and A and C we can only eliminate $>$ (after) and mi (met-by) out of the 13 possible relations, the remaining 11 relations have to be considered:

$$\begin{array}{ll} A & \{=,<,m,o,oi,d,di,s,si,f,fi\}\ B \\ A & \{=,<,m,o,oi,d,di,s,si,f,fi\}\ C \\ B & \{<,m,o,s,d\}\ C \end{array}$$

If we would know that the end of time period A ends after the end of time period C (A *survives* C, $A\ sv\ C$) and add this piece of information to the system, the amount of possible relations could be reduced significantly:

$$\begin{array}{ll} A & \{oi,di,si\}\ B \\ A & \{sv\}\ C \\ B & \{<,m,o,s,d\}\ C \end{array}$$

Now, instead of 11 relations we only have three oi, di, si (overlapped-by, contains, started-by). According to Freksa, these remaining relations are also conceptual neighbors and can be aggregated into the concept "surviving contemporary of" (sc).

In order to specify a new explicit relation in a XML notation, the construct "relatedTo" is used. The attribute "ref" denotes another period name where the type of the relation is given by the attribute "type". Here is an example denoting the time period of the Middle Ages:

Example 3.9 (Middle Ages relations)

```
<periodName id="Middle-Ages">
  <!-- Definition of boundaries -->
    <begin unknown="true"/>
    <end   unknown="true"/>
  <!-- Definition of relations -->
    <relatedTo type="younger"     ref="Antiquity"/>
    <relatedTo type="survives"    ref="Antiquity"/>
    <relatedTo type="older"       ref="Modern_times"/>
    <relatedTo type="survivedBy"  ref="Modern_times"/>
</periodName>
```

In this example, both the starting and ending time points are defined as unknown. Then, we add the new relations (which are concerned with the relation of the outer time points) to those of other time intervals (e.g., younger as Antiquity).

3.5 Temporal relevance

When using the temporal model to both annotate and retrieve information from the Web, the following question arises: how do we determine which data or information sources fit the query and to which degree? This can be summarized in the term of temporal relevance. Usually, the relevance is drawn on a scale between 0 and 1. The degree of relevance then mirrors the percentage of ”fitness”, i.e., 0 means that the found data do not fit the query at all, whereas, 1 means that the data fit the query with 100%.

After a thorough study of Allen's relations, we can group these into two main areas. One group consists of relations that consider disjunct time intervals only, i.e., $before$ and $after$. The other group consists of relations that have an overlap of some kind (e.g., *during*, *contains*). However, there are two exceptions: *meets* and $met - by$. These can be seen as relations, which consider time intervals that are disjunct (by a millisecond) or overlapped (by one millisecond). For the following, we consider the latter and therefore group these two relations into the second area.

Furthermore, the temporal relevance can also be distinguished into two areas: (a) the distance and (b) the overlap of time periods. The latter can be refined to the consideration of distance between time points, namely the start and end time points of the considered time intervals.

3.5.1 Distance between time intervals

The calculation between two time intervals A and B where the relation $A\, before\, B$ holds true, is based on the distance between the end time point of A and the start time point of B. The length of the time interval is not relevant. Therefore, we can calculate the distance even if the other boundaries are unknown. Theoretically, 16 (4^2) combinations of two time intervals with different types of boundaries are possible. However, because we do not have to consider the types of boundaries that are at the start of A and the end of B we can reduce the number of combinations to ten (figure 3.5). The number of combinations from which we can draw conclusions is even lower if:

- one of the boundaries is unknown, we cannot make a comment about the relation and therefore we cannot calculate the distance. Four combinations out of the ten belong to this group (d,g,i, and j in figure 3.5);

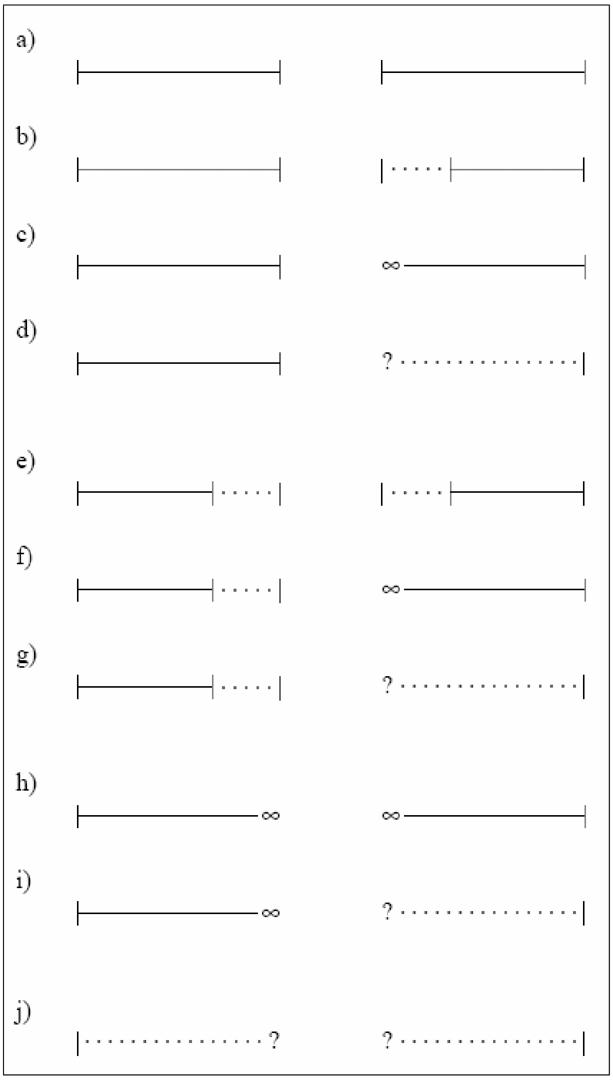


Figure 3.5: Distance: the possible combinations of boundary types

- at least one of the boundaries is persistent, no distance can be calculated because one time interval is overlapping the other. Four combinations out of the ten belong to this group (one belongs to both groups: i,c,f, and h in figure 3.5).

Thus, three combinations where we have exact and fuzzy boundaries are left and have to be further considered. In the case of two exact boundaries, the calculation is simple because we can use subtraction. In the case of at least one fuzzy boundary, we simply calculate the mean average value of the tolerance area, i.e., the mean average value between the inner and outer boundary and then use subtraction for the overall distance. Once we have the distance, we can norm this value in order to get a value between 0 and 1.

3.5.2 Overlapping of time periods

The calculation of relevance between two overlapping time intervals causes a new consideration: it is important to know which time interval is the reference time interval and which time interval is the comparer. Figure 3.6 gives us some insight into this problem: we can see that A and B as well as A and C are overlapping. However, from the viewpoint of B, A is more important because A covers B completely. On the other hand B is not as important for A because the degree of overlapping of B is smaller than the one of C.

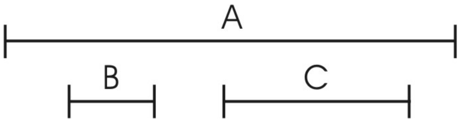


Figure 3.6: Relevance as extent of overlapping

In contrast to the process of calculation with regard to the distance where we could calculate with the two opposite time points, we have to consider all four boundary time points of the two time intervals. Theoretically, we have to consider 256 (4^4) combinations, which can be reduced due to symmetry drastically. None of the boundaries ought to be unknown since we cannot calculate any relevance. Also, persistent boundaries can be transformed into exact boundaries if the reference time interval has only exact or fuzzy boundaries.

The calculation of the relevance between two intervals with exact boundaries is straight forward: the length of the overlapping area can be related to the overall length of the reference time interval. If both intervals are identical, the relevance is 100%. The distance calculation with fuzzy boundaries must have a different result than the distance that would have been calculated using exact boundaries. Therefore, the width of the fuzzy area must have a significant influence on the result. Figure 3.7 shows the representation of fuzzy boundaries: the fuzzy area at the start of the time period (a) is the area between *beginf* and *begin*. The area between *begin* and *end* (b) is the area, which is certain, and the area between *end* and *endf* (c) is the fuzzy area at the end of the time period.

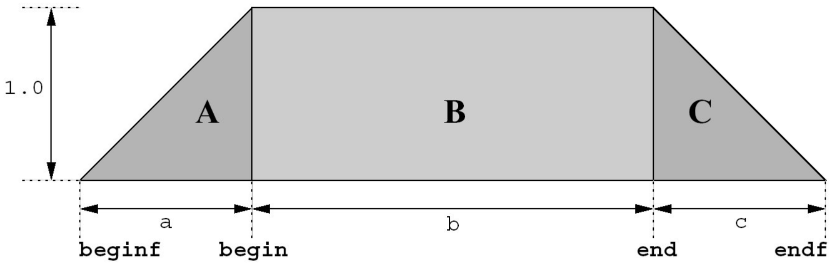


Figure 3.7: Representation for fuzzy boundaries

This representation follows the representation known as fuzzy set theory (Zadeh 1965) where we have a moving transition of elements belonging to a set or not. The membership is described by a function that maps onto values between 0 and 1. Thus, we are able to represent common terms with fuzzy boundaries such as "warm" or "tall". If we need to calculate the relevance of a time period (a possible answer to our query) with regard to a reference time period (our query), the overlapping area has to be determined. The overlapping area includes both fuzzy areas at the beginning and the end of the time periods and the "certain" area in the middle. After determining the overlapping area, we calculate the relation of the two time periods simply by dividing them: $\frac{A}{B}$ where A is the time period of the possible answer and B is the reference time period. The result is the temporal relevance for time period A with regard to time period B. Further details are described in (Hübner 2003).

3.6 Reasoning components

We have introduced the concept "time period" for the abstract representation of time and time-based relations. Also, a simple algorithm for the calculation of temporal relevance has been described. Both representation and relevance are necessary to develop reasoning components that are described in the following. However, some assumptions and restrictions must first be made.

We have seen 30 relations between time intervals in total, 13 have been introduced by Allen, another 17 by Freksa (17 semi-interval comparisons of Allen's disjunctive sets). The more temporal cohesions a reasoner is able to process, the more powerful and efficient it is. The development of a temporal reasoner is currently undertaken (a first prototype has been finished) and we have started with the most important temporal relations with regard to the Semantic Web:

- older
- younger
- contemporary
- survives
- survived-by

The selection of these five relations is described in (Hübner 2003) in more detail.

In order to get conclusions based on the temporal model, new algorithms have to be developed. Allen used a constraint-based system to reduce the set of possible relations when adding new information. The system is also able to detect inconsistencies, however, the system is very limited. Therefore, we extend and modify Allen's approach in order to tackle the new temporal model (e.g., for fuzzy, persistent, and unknown boundaries, references). A particular feature is the co-existence of quantitative descriptions of periods and qualitative relations of such periods. In addition, with regard to the Semantic Web, it is imperative to detect inconsistencies.

3.6.1 Relations between boundaries

Considering the boundaries of time periods we can derive implicit relations. First, we have to compare the time points of those boundaries. If these time points are exact, we can order these and get three relations: (a) a time point does lay *before* another time point ($<$), (b) a time point does lay *after* another time point ($>$), and (c) the time points are the *same* ($=$).

Considering at least one fuzzy boundary is sufficient to compare the outer time points. This way, we take the maximum expansion of the time period into account and therefore simulate a time period with exact boundaries. This is possible due to the fact that the relations identified in the two groups of relevance (distance and overlapping) do not distinguish between exact and fuzzy boundaries. Thus, we also have the three relations $<, >, =$ for fuzzy boundaries.

Persistent boundaries cannot be mapped onto concrete time points due to the concept of the point structure (see 3.2). Therefore, numerous situations must be distinguished.

First, persistent boundaries can appear in two ways: they are persistent with regard to the start of the time period (negative persistent) or persistent with regard to the end of the time period (positive persistent). For the comparison with exact time points, which are derived from exact boundaries or the transformation from fuzzy boundaries, and for the comparison with persistent boundaries the following theorems hold:

Theorem 3.1
A negative persistent boundary truly lays a) before all exact time points $t_n \in [B,E], n \in \mathbb{N}$ *of the time range and b) before all positive persistent boundaries.*

Theorem 3.2
A positive persistent boundary truly lays a) behind all exact time points $t_n \in [B,E], n \in \mathbb{N}$ *of the time range and b) behind all negative persistent boundaries.*

Theorem 3.3
The relation between two positive or two negative persistent boundaries is undetermined.

This corresponds to a intuitive notion of a time period, which is infinite far towards the past or the future. Therefore, we can determine three relations with regards to two persistent boundaries: $<, >$, and *unknown*. This is the basis for comparisons between time periods with regard to their position.

When considering time periods with unknown boundaries, only one relation with regard to another arbitrary boundary can be made: *unknown*.

Remark 3.1
No proposition can be made with regard to a position of an unknown, exact, fuzzy, or persistent boundary.

The proofs to the theorems 3.1-3.3 and the remark 3.1 can be done with the consideration of all possible cases. Figure 3.8 shows a time line with negative and positive persistent boundaries P^-, P^+, the temporal range denoted by $[B,E]$, and the unknown boundaries U.

Proof 3.1 (-3.3)

Figure 3.8: Time line

$$\forall s,t \in [B,E] \cup P^- \cup P^+ \cup U : s <^* t = \begin{array}{lll} s < t \text{ if} & s \in [B,E] \wedge t \in [B,E] & \\ \text{true if} & s \in P^- \wedge t \in [B,E] & \text{(Theorem 3.1)} \\ \text{true if} & s \in P^- \wedge t \in P^+ & \text{(Theorem 3.1)} \\ \text{true if} & s \in [B,E] \wedge t \in P^+ & \text{(Theorem 3.2)} \\ \text{false if} & s \in [B,E] \wedge t \in P^- & \\ \text{false if} & s \in P^+ \wedge t \in P^- & \\ \text{false if} & s \in P^+ \wedge t \in [B,E] & \\ \text{false else} & \text{(Theorem 3.3, Remark 3.1)} & \square \end{array}$$

Thus, we are able to define four relations between time periods: $<, >, =$, and *unknown*. With this help, we can compare two time periods to derive their position relatively to each other.

	necessary	sufficient positive	sufficient negative
older	$\alpha < A$	$\omega < A$	$\alpha \geq \Omega$
younger	$\alpha > A$	$\alpha > \Omega$	$\omega \leq A$
survives	$\omega > \Omega$	$\alpha > \Omega$	$\omega \leq A$
survived-by	$\omega < \Omega$	$\omega < A$	$\alpha \geq \Omega$
contemporary	$\alpha < \Omega \wedge \omega > A$	$\omega \leq \Omega \wedge \omega > A$ $\vee\ \alpha < \Omega \wedge \alpha \geq A$ $\vee\ \alpha \leq A \wedge \omega > A$ $\vee\ \alpha < \Omega \wedge \omega \geq \Omega$	$\alpha \geq \Omega \vee \omega \leq A$

Table 3.1: Necessary and sufficient conditions for the five selected relations

3.6.2 Relations between two time periods

(Freksa 1992) introduced a method to determine the relations between two time periods by comparing abstract semi-intervals. These semi-intervals lay at the beginning and the end of the involved time periods, and are denoted as α and ω (or A and Ω).

This way, Allen's fundamental relations as well as Freksa's conceptual neighborhoods can be described easily. The procedure can be adapted for periods with exact boundaries by replacing the comparisons between semi-intervals with comparisons between margin points. The three relations used by Freksa ($<, >, =$) are also defined in this scenario.

Because we do not have to distinguish between exact and fuzzy boundaries, a transformation to an exactly defined time period with a maximal expansion can be made. Also, we are able to perform the same comparisons: it is not important whether or not the overlapping area belongs to a fuzzy area, what counts is the *existence* of a time period that is covered by the two time periods. Therefore, Freksa's definitions can be transformed onto exact or fuzzy boundaries.

If one of the time periods has at least one persistent boundary, two points have to be considered. First, the relation *equal* ($=$) is not defined. However, since none of the selected time period relations *(older, younger, contemporary, survives, survived-by)* is dependent on that particular relation the importance is marginal. Second, the position of two time periods can be *unknown*. In that case, no further propositions can be made.

Positive and negative sufficient conditions One of the benefits of our approach is that we can draw conclusions about relations of time periods, even if we have to deal with incomplete information. However, we need the necessary (Hübner 2003, page 41pp.) and sufficient conditions in order to draw conclusions. Table 3.1 gives an overview about the necessary and sufficient conditions for the selected relations. Suppose, we have two time periods (A and B) with only one exactly defined margin point per time period. Time period A has a defined end point and time period B has a defined beginning. The other margin points are unknown. In this situation, the check of the relation A *older* B would return nothing since the relation between the beginning of $A(\alpha)$ and the beginning of $B(\alpha)$ is not defined (see remark 3.1 above). On the other hand, we do know that the end of $A(\omega)$ is truly before B. Together with our fundamental demand that a beginning of an interval lays always before the end ($\alpha < \omega; A < \Omega$) we can conclude that $\alpha < A$, i.e., the relation A *older* B is valid. The same holds true for B *younger* A.

This method checks a *positive sufficient* condition with $\omega < A$. However, there are also *negative sufficient* conditions. In the mentioned situation we can see that $\omega > \Omega$ does not hold: $A > \omega$ can be read from the position of the exact boundaries, $\Omega > A$ holds true by definition for all intervals. Also, due to the transitivity of $>$ the relation $\Omega > \omega$ holds true. Thus, the relation A *survives* B is therefore rejected by the negative sufficient condition $A > \omega$.

3.6.3 Relations between more than two time periods

The comparisons between boundaries and two time periods enable us to make statements about cohesions between more than two time periods. Allen's composition table is a known approach for the concatenation of two relations. However, the restriction of the selected relations *older, younger, contemporary, survives* and *survived-by* make the construction for another composition table unnecessary.

Theorem 3.4
The inverse relations rule themselves out (older and younger; survives and survived-by), *all other combinations are possible, e.g.,* A ol B; A ct B; A sv B. *Further, we can aggregate the relations into two groups: (a) reflexive (3.19) and symmetric (3.20)* (contemporary) *and (b) non-reflexive (3.21), anti-symmetric (3.22), and transitive (3.23)* (older, younger, survives, survived-by).

$$\forall p \in P \quad : \quad (p\ ct\ p) \tag{3.19}$$
$$\forall p_1, p_2 \in P \quad : \quad (p_1\ ct\ p_2) \longrightarrow (p_2\ ct\ p_1) \tag{3.20}$$

$$\forall p \in P \quad : \quad \neg(p\ ol\ p) \tag{3.21}$$
$$\forall p_1, p_2 \in P \quad : \quad \neg(p_1\ ol\ p_2 \wedge p_2\ ol\ p_1) \tag{3.22}$$
$$\forall p_1, p_2, p_3 \in P \quad : \quad (p_1\ ol\ p_2 \wedge p_2\ ol\ p_3) \longrightarrow (p_1\ ol\ p_3) \tag{3.23}$$

where P is the set of all time periods.

Also, we can show that a time period p_1 overlaps another time period p_2 if p_1 starts earlier and ends later (3.24); the invers relations hold correspondingly (3.25).

$$\forall p_1, p_2 \in P \quad : \quad (p_1\ ol\ p_2 \wedge p_1\ sv\ p_2) \longrightarrow (p_1\ ct\ p_2) \tag{3.24}$$
$$\forall p_1, p_2 \in P \quad : \quad (p_1\ yo\ p_2 \wedge p_1\ sb\ p_2) \longrightarrow (p_1\ ct\ p_2) \tag{3.25}$$

Proof 3.4
The statements 3.19-3.25 can be derived from the definitions of the relations about semi-interval comparisons and the implicit relations $\alpha < \omega$ (and $A < \Omega$) for each time period. This can be shown first for the relation contemporary*:*

$$\forall p \in P \quad : \quad (\alpha < \omega \wedge \omega > \alpha) \longrightarrow \forall p \in P : (p\ ct\ p) \quad \square \tag{3.26}$$

$$\begin{aligned} \forall p_1, p_2 \in P \quad : \quad & (p_1\ ct\ p_2) \\ & \longrightarrow (\alpha < \Omega \ \wedge \ \omega > A) \\ & \longrightarrow (A < \omega \ \wedge \ \Omega > \alpha) \\ & \longrightarrow (p_2\ ct\ p_1) \quad \square \end{aligned} \tag{3.27}$$

The relations older, younger, survives *and* survived-by *use semi-interval comparisons $<$ and $>$ exclusively, including their non-reflexivity, anti-symmetry, and transitivity:*

$$\forall p \in P \quad : \quad \neg(\alpha < \alpha) \longrightarrow \forall p \in P : \neg(p\ ol\ p) \quad \square \tag{3.28}$$

$$\begin{aligned} \forall p_1, p_2 \in P \quad &: \quad \neg(\alpha < A \ \wedge \ A < \alpha) \\ & \quad \longrightarrow \forall p_1, p_2 \in P : \neg(p_1 \ ol \ p_2 \ \wedge \ p_2 \ ol \ p_1) \quad \Box \end{aligned} \tag{3.29}$$

$$\begin{aligned} \forall p_1, p_2, p_3 \in P \quad &: \quad (p_1 \ ol \ p_2 \ \wedge \ p_2 \ ol \ p_3) \\ & \quad \longrightarrow (\alpha < \alpha'; \wedge \alpha' < A) \\ & \quad \longrightarrow (\alpha < A) \\ & \quad \longrightarrow (p_1 \ ol \ p_3) \quad \Box \end{aligned} \tag{3.30}$$

3.24 and 3.25 can be shown accordingly:

$$\begin{aligned} \forall p_1, p_2 \in P \quad &: \quad (p_1 \ ol \ p_2 \ \wedge \ p1 \ sv \ p_2) \\ & \quad \longrightarrow (\alpha < A \ \wedge \ \omega > \Omega) \\ & \quad \longrightarrow (\alpha < \Omega \ \wedge \ \omega > A) \\ & \quad \longrightarrow (p_1 \ ct \ p_3) \quad \Box \end{aligned} \tag{3.31}$$

$$\begin{aligned} \forall p_1, p_2 \in P \quad &: \quad (p_1 \ yo \ p_2 \ \wedge \ p1 \ sb \ p_2) \\ & \quad \longrightarrow (\alpha > A \ \wedge \ \omega < \Omega) \\ & \quad \longrightarrow (\omega > A \ \wedge \ \alpha < \Omega) \\ & \quad \longrightarrow (\alpha < \Omega \ \wedge \ \omega > A) \\ & \quad \longrightarrow (p_1 \ ct \ p_3) \quad \Box \end{aligned} \tag{3.32}$$

We can see that we can derive the cohesions between multiple time periods without a complex composition table. The most important means are symmetry of the *contemporary* relation and transitivity of the *older*, *younger*, *survives*, and *survived-by* relation.

3.7 Example

The following example gives us an impression of the reasoning performance of the mentioned temporal approach. As a basis we choose a structure with the period names ”antiquity”, ”Middle Ages”, and ”modern times”. We vary their boundaries in order to demonstrate the reaction of the underlying engine.

3.7.1 Qualitative statements

Suppose the boundaries of the three period names are completely undetermined and only a few qualitative statements with regard to the relations between them are known. Figure 3.9 shows this situation in a XML notation. The reasoner transforms the situation in an internal graphical structure and derives eight relations besides the known period names. Four of them are already given by the user (USER), another four can be derived by symmetry from *older* and *younger* as well as *survived* and *survived-by* (IMPLICIT). The reasoner shows the following output:

```
Parsing: ok.
-----
Transformation DOM->Internal Representation: ok.
-----
# of Periods found: 3 "antiquity"    [UNKNOWN,UNKNOWN]
```

```
<?xml version="1.0" encoding="ISO-8859-1"?> <periodNames
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:noNamespaceSchemaLocation="periodNames.xsd">

  <periodName id="antiquity">
    <relatedTo type="older" ref="middle-ages"/>
  </periodName>

  <periodName id="middle-ages">
    <relatedTo type="older"       ref="modern-times"/>
    <relatedTo type="survives"    ref="antiquity"/>
    <relatedTo type="survivedBy"  ref="modern-times"/>
  </periodName>

  <periodName id="modern-times"/>
</periodNames>
```

Figure 3.9: Example ”antiquity, Middle Ages, and modern times”

```
"middle-ages"   [UNKNOWN,UNKNOWN] "modern-times" [UNKNOWN,UNKNOWN]
-----
OLDER={
  "antiquity"   "middle-ages"    (? ,USER ,UNKNOWN)
  "middle-ages"  "modern-times"  (? ,USER ,UNKNOWN)
}
YOUNGER={
  "middle-ages"    "antiquity"   (? ,IMPLICIT,UNKNOWN)
  "modern-times"  "middle-ages"  (? ,IMPLICIT,UNKNOWN)
}
CONTEMPORARY={
}
SURVIVES={
  "middle-ages"    "antiquity"   (? ,USER ,UNKNOWN)
  "modern-times"  "middle-ages"  (? ,IMPLICIT,UNKNOWN)
}
SURVIVEDBY={
  "antiquity"   "middle-ages"    (? ,IMPLICIT,UNKNOWN)
  "middle-ages"  "modern-times"  (? ,USER ,UNKNOWN)
}
```

Each relation consists of origin (USER/IMPLICIT), the temporal relevance, and a validity status. The two latter are unknown at present and therefore, the ’?’ and the term UNKNOWN is given.

The next step is the expansion and verification of the internal model. The already known relations will be given again, however, if the verification process can verify the qualitative relations with the help of quantitative comparisons, these will be shown. In this case, the validity status is the same than the above, i.e., no quantitative comparisons can be made. Here is an extract of the output (note the expansion by the reasoner, e.g., the ”antiquity/modern-times”-relation in OLDER):

```
Expand and Verify: ok.
-----
OLDER={
  "antiquity"   "middle-ages"    (? ,USER ,UNKNOWN)
  "middle-ages"  "modern-times"  (? ,USER ,UNKNOWN)
  "antiquity"   "modern-times"  (? ,REASONER,UNKNOWN)
}
YOUNGER={
  "middle-ages"    "antiquity"   (? ,IMPLICIT,UNKNOWN)
  "modern-times"  "middle-ages"  (? ,IMPLICIT,UNKNOWN)
  "modern-times"  "antiquity"   (? ,REASONER,UNKNOWN)
}
CONTEMPORARY={
}
SURVIVES={
  "middle-ages"    "antiquity"   (? ,USER ,UNKNOWN)
  "modern-times"  "middle-ages"  (? ,IMPLICIT,UNKNOWN)
  "modern-times"  "antiquity"   (? ,REASONER,UNKNOWN)
}
SURVIVEDBY={
  "antiquity"   "middle-ages"    (? ,IMPLICIT,UNKNOWN)
  "middle-ages"  "modern-times"  (? ,USER ,UNKNOWN)
  "antiquity"   "modern-times"  (? ,REASONER,UNKNOWN)
}
```

At this point, the temporal model is expanded to its maximal extent (12 relations). So far, no inconsistencies have been found between the qualitative relations and quantitative boundaries. In addition, no inconsistencies have been detected between two or more qualitative statements. Thus, after verifying the consistency, queries can be formulated.

One example is the following: ”Which period names do have a known relation with *Middle-Ages*, what kind of relations are these, and which temporal relevance do they have?” Here is the outcome:

```
relatedTo middle-ages:
  older:        [antiquity(?)]
  younger:      [modern-times(?)]
  contemporary: []
  survives:     [modern-times(?)]
  survivedBy:   [antiquity(?)]
```

3.7.2 Quantitative statements

The second example consists of the same structure and periods but with determined boundaries at the beginning and the end. Figure 3.10 shows the details, please note that some of these boundaries reference already defined boundaries (e.g., endOf ref=”antiquity”).

After parsing and transforming the input, the following list of period names including their explicit relations is found:

```
# of Periods found: 3 "antiquity"     [-UNLIMITED,-46388592000000]
```

```
<periodName id="antiquity">
  <begin unlimited="true"/>
  <end>
    0500-01-01T00:00:00.000+00:00
  </end>
</periodName>

<periodName id="modern-times">
  <begin>
    1500-01-01T00:00:00.000+00:00
  </begin>
  <end unlimited="true"/>
</periodName>

<periodName id="middle-ages">
  <begin>
    <endOf ref="antiquity"/>
  </begin>
  <end>
    <beginOf ref="modern-times"/>
  </end>
</periodName>
```

Figure 3.10: Example "antiquity, middle ages, and modern times with determined boundaries"

```
"modern-times"  [-14830992000000,+UNLIMITED] "middle-ages"
[-46388592000000,-14830992000000]
-----
OLDER={}
YOUNGER={}
CONTEMPORARY={}
SURVIVES={}
SURVIVEDBY={}
```

Since the internal format of date consist of the number of milliseconds to or from the beginning of the "JAVA-epoche" (January 1st, 1970, 12.00am), the exact boundaries are shown as big negative numbers. The persistent boundaries differ in the sign according to the direction of "leaving" the range: negative sign for the past and positive sign for the future. The list of explicit relations is empty because there are no explicit relations given. After expanding and verifying the model, the output is the following:

```
OLDER={
  "antiquity"     "modern-times"  (? ,REASONER,VALID)
  "antiquity"     "middle-ages"    (? ,REASONER,VALID)
  "modern-times"  "antiquity"     (? ,REASONER,INVALID)
  "modern-times"  "middle-ages"    (? ,REASONER,INVALID)
  "middle-ages"    "antiquity"     (? ,REASONER,INVALID)
  "middle-ages"    "modern-times"  (1.0 ,REASONER,VALID)
```

```
}
YOUNGER={
  "modern-times"  "antiquity"     (? ,REASONER,VALID)
  "middle-ages"    "antiquity"     (? ,REASONER,VALID)
  "antiquity"     "modern-times"  (? ,REASONER,INVALID)
  "middle-ages"    "modern-times"  (? ,REASONER,INVALID)
  "antiquity"     "middle-ages"    (? ,REASONER,INVALID)
  "modern-times"  "middle-ages"    (? ,REASONER,VALID)
}
CONTEMPORARY={
  "antiquity"     "modern-times"  (? ,REASONER,INVALID)
  "modern-times"  "antiquity"     (? ,REASONER,INVALID)
  "antiquity"     "middle-ages"    (? ,REASONER,INVALID)
  "middle-ages"    "antiquity"     (? ,REASONER,INVALID)
  "modern-times"  "middle-ages"    (? ,REASONER,INVALID)
  "middle-ages"    "modern-times"  (? ,REASONER,INVALID)
}
SURVIVES={
  "antiquity"     "modern-times"  (? ,REASONER,INVALID)
  "antiquity"     "middle-ages"    (? ,REASONER,INVALID)
  "modern-times"  "antiquity"     (? ,REASONER,VALID)
  "modern-times"  "middle-ages"    (? ,REASONER,VALID)
  "middle-ages"    "antiquity"     (1.0 ,REASONER,VALID)
  "middle-ages"    "modern-times"  (? ,REASONER,INVALID)
}
SURVIVEDBY={
  "modern-times"  "antiquity"     (? ,REASONER,INVALID)
  "middle-ages"    "antiquity"     (? ,REASONER,INVALID)
  "antiquity"     "modern-times"  (? ,REASONER,VALID)
  "middle-ages"    "modern-times"  (? ,REASONER,VALID)
  "antiquity"     "middle-ages"    (? ,REASONER,VALID)
  "modern-times"  "middle-ages"    (? ,REASONER,INVALID)
}
```

All these relations are found by the reasoner. Those relations that could be proven within the process of expanding are marked as ”VALID”. On the other hand, those that could be proven ”INVALID” are marked as such. Please note that the invalid relations do not imply any inconsistencies. These are *implicit* relations and are therefore not inconsistent for the internal representation. The implicit relations are determined with the help of the theorems give in section 3.6. Theorem 1 for instance can be used to derive ”antiquity *older* middle-ages”.

The reasoner also found two relations where the temporal relevance could be determined (*middle-ages survives antiquity* and *middle-ages older modern-times*). In both cases, we compare the overlapping time interval of the actual time period with the time interval that is given by the significant points for the actual relation. *Older* uses the start points and *survives* uses the end points of the periods to compare. These time intervals are identical because the periods are standing in relation to *meets* or *met-by*. Therefore, a temporal relevance of 1.0 is calculated. The temporal relevance cannot be calculated if the time points are persistent or unknown.

The following examples are shorter and only the significant outcomes are shown.

3.7.3 Inconsistencies (quantitative/qualitative)

In order to demonstrate the behavior of the reasoner with regard to inconsistencies our former example will be extended by an explicit relation, which is in direct contradiction to the modelled boundaries: *middle-ages older antiquity*. The following demonstrates the output after parsing and transforming the given model:

```
# of Periods found: 3 "antiquity"     [-UNLIMITED,-46388592000000]
"modern-times"  [-14830992000000,+UNLIMITED] "middle-ages"
[-46388592000000,-14830992000000]
-----
OLDER={
  "middle-ages"  "antiquity"   (? ,USER ,UNKNOWN)
}
YOUNGER={
  "antiquity"   "middle-ages"  (? ,IMPLICIT,UNKNOWN)
}
CONTEMPORARY={}
SURVIVES={}
SURVIVEDBY={}
```

The validity value is unknown at this point. After expansion and verification inconsistencies are determined. Theorem 1 ,e.g., proves *antiquity older middle-ages* and therefore contradicts *antiquity younger middle-ages*, which implicitly can be derived with the help of the temporal model *middle-ages older antiquity*. The following outcome shows the inconsistencies, which make the overall model invalid (the invalid inverse relations are not shown for better understanding):

```
["Middle-Ages"---OLDER-->"Antiquity"]
KnowledgeBase contains 1 invalid and 0 contradictory relations!
Shutting down...
```

Once it is known that the temporal model is not consistent, queries cannot be made because the correctness of the results cannot be guaranteed.

3.7.4 Inconsistencies (reasoner implicit/qualitative)

Another example for inconsistencies is the contradiction between explicit qualitative relations and relations that are derived by the reasoner using quantitative knowledge. In order to demonstrate this, we modify our example slightly as shown in figure 3.11. The internal representation does not contain contradictions in the beginning between the boundaries and the modelled relations because they relate to an undetermined period:

```
# of Periods found: 3 "antiquity"     [-UNLIMITED,UNKNOWN]
"middle-ages"    [-46388592000000,UNKNOWN] "modern-times"
[UNKNOWN,UNKNOWN]
-----
OLDER={
  "middle-ages"    "modern-times"  (? ,IMPLICIT,UNKNOWN)
```

```
...
<periodName id="antiquity">
  <begin unlimited="true"/>
  <end unknown="true"/>
</periodName>

<periodName id="middle-ages">
  <begin>
    0500-01-01T00:00:00.000+00:00
  </begin>
  <end unknown="true"/>
</periodName>

<periodName id="modern-times">
  <begin unknown="true"/>
  <end unknown="true"/>
  <relatedTo type="younger" ref="middle-ages"/>
  <relatedTo type="older" ref="antiquity"/>
</periodName>
```

Figure 3.11: Example "antiquity, middle ages, and modern times creating an inconsistency"

```
  "modern-times"  "antiquity"     (? ,USER ,UNKNOWN)
}
YOUNGER={
  "modern-times"  "middle-ages"    (? ,USER ,UNKNOWN)
  "antiquity"     "modern-times"  (? ,IMPLICIT,UNKNOWN)
}
CONTEMPORARY={}
SURVIVES={}
SURVIVEDBY={}
```

We do not know the beginning or the end of "modern-times". Therefore, we can neither prove nor disprove *modern-times older antiquity* or *modern-times younger middle-ages* and the resulting inverse relations. Thus, the validity value stays unknown. During the expansion using the marginal points we can derive implicit relations such as *antiquity older middle-ages* using theorem 1 (because of the transitivity of the *older*-relation knowing *modern-times older antiquity*). Accordingly, we can prove the inconsistency *modern-times younger middle-ages*. Here is the outcome of the reasoning process:

```
["antiquity"---OLDER-->"modern-times",
 "middle-ages"---YOUNGER-->"modern-times"]
KnowledgeBase contains 0 invalid and 2 contradictory relations!
Shutting down...
```

The additional given relations are consistent in this case, however, combining those with quantitative statements can prove the contradictions.

3.7.5 Inconsistencies (qualitative/quantitative)

In our last example we demonstrate the appearance of contradictions having qualitative models only. We modify the above mentioned example accordingly showing cycles (see figure 3.12). While constructing the internal representation no inconsistencies between relations and boundaries were found because the latter are not defined. The expansion and verification process, however, finds contradictions within all three relations due to the asymmetry of *older*.

```
...
<periodName id="antiquity">
  <relatedTo type="older" ref="middle-ages"/>
</periodName>

<periodName id="middle-ages">
  <relatedTo type="older" ref="modern-times"/>
</periodName>

<periodName id="modern-times">
  <relatedTo type="older" ref="antiquity"/>
</periodName>
...
```

Figure 3.12: Example ”antiquity, middle ages, and modern times with qualitative relations only”

```
["antiquity"---OLDER-->"modern-times",
 "middle-ages"---OLDER-->"antiquity",
 "modern-times"---OLDER-->"middle-ages"]
KnowledgeBase contains 0 invalid and 3 contradictory relations!
Shutting down...
```

The reasoner identifies all inconsistencies, which can help to evaluate and modify the temporal model in order to eliminate the contradictions. In our case, the relation *modern-times older antiquity* could be eliminated or changed to *modern-times younger antiquity*.

We have shown that the reasoning process is able to detect all possible inconsistencies of a temporal model, which is based on a period names structure. Inconsistencies could appear (a) between qualitative statements and defined boundaries, (b) between qualitative statements and derived implicit relations, and (c) between qualitative statements containing cycles. In addition, inconsistencies are labelled to simplify the correction of the model.

4 System Demonstration

The prototype of the BUSTER systems is based on an open client/server architecture (cf. (Visser and Schuster 2002)) and can be divided into two main parts: the so-called BUSTER-cluster on the server side and a BUSTER client.

The cluster part contains all the relevant modules necessary to guarantee the functionalities described in the sections before. The following will include possible queries that can be made with the new temporal model. We do not include the terminological and spatial queries in this paper and therefore refer to (Visser 2003).

4.1 Simple queries

Here, we will only consider temporal queries. This part of the BUSTER system is currently under development. However, the temporal reasoning engine is already accessible by both the BUSTER server and the client. Although the system lacks comprehensive examples, one temporal model can be chosen by the user. The data we described consist of documents and information from the Bremen Senator for Construction and Environment (SBU), Referat 44. The temporal ontology contains the necessary knowledge and a reasonable differentiation for this case. Here is a part of the temporal model:

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
<periodNames xmlns:
    xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:noNamespaceSchemaLocation="http://www.tzi.de/buster/data/xsd/periodNames.xsd">
    <periodName id="Jahr_2001">
        <equals>
            <pformula name="wholeyear" year="2001" />
        </equals>
    </periodName>
    <periodName id="Jahre_1998_bis_2002">
        <begin>
            <dformula name="yearbegin" year="1998" />
        </begin>
        <end>
            <dformula name="yearend" year="2002" />
        </end>
    </periodName>
...
</periodNames>
```

The user chooses a temporal model and gets prompted with the possible templates. Suppose, he chooses the temporal concept 'Years_1998_until_2002'. The temporal reasoner expands and verifies the model as described in section 3 and calculates the temporal annotations within the CSDs of the information sources. As we can see, this temporal concept is modelled as a formula, hence, the reasoner is able to derive that a document or information source annotated with 'since_2001' fits the query. Figure 4.13 shows the result of that query.

4.2 Combined queries

Among the single terminological, spatial, and temporal queries, all possible combinations of queries can be made. We illustrate an additional type of query: "Spatio-temporal-terminological query", which we also call *(concept@location in time)*

4.2.1 Spatio-temporal-terminological Queries

The most sophisticated and interesting (from the Semantic Web point of view) type of query can be formulated as *concept@location in time*. Our example brings us in the area of tourism. We choose the application domain GeoShare for the terminological ontology, the North-Sea region as our spatial model and the temporal model from above, the SBU-Referat-44 model. Figure 4.14a shows the concepts we are looking for: we are interested in any information source or documents that contain something about fishing in the North-Sea region since 1990.

Figure 4.14b shows the result of our query. We can see that one of the found information source with the title ”Fischgewässer” Bremen contains the terminological concept ”angling” which is subsumed by fishing. The spatial reasoner found the location ”Bremen, Krfr.st.” (a suburb of the city Bremen), which clearly is part of the North-Sea region and the temporal reasoner proved that the document which has been annotated with ”seit Jahr 2002” also belongs to the class ”seit Jahr 1990”.

5 Conclusion and Future Work

We summarize the work we have done and also draw some line of research that needs to be done in the future.

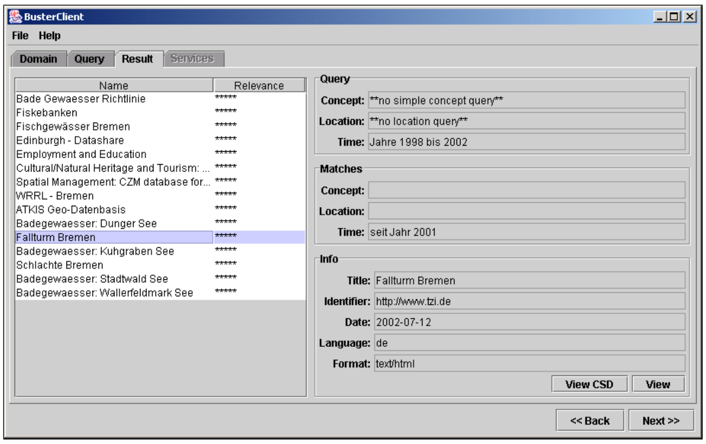
5.1 Conclusion

The most important result of our work is that our approach, both the conceptual and the implementation part, is operating the way we wanted it to operate. This includes all the requirements that have been defined before we started the work.

An important result is the type of queries that are possible. We are able to support the user (or other systems) with new types of queries because of the development of the spatial and temporal reasoners. These queries are *concept@location*, *concept in time*, or *concept@location in time*. This types of queries can help to support users or systems in finding what they are after in a more intelligent and accurate manner.

Another major result is the improvement of expressiveness. We called the requirement ”intuitive labelling” (e.g., place names, period names) and implemented this throughout our system. This is an important part of our approach enabling users to use colloquial terms while editing their search.

We showed that the existing temporal approaches are not satisfactory to serve the requirements of the modern Semantic Web. The major problem is the lack of expressiveness



Name	Relevance
Bade Gewaesser Richtlinie	*****
Fiskebanken	*****
Fischgewässer Bremen	*****
Edinburgh - Datashare	*****
Employment and Education	*****
Cultural/Natural Heritage and Tourism: ...	*****
Spatial Management: CZM database for...	*****
WRRL - Bremen	*****
ATKIS Geo-Datenbasis	*****
Badegewaesser: Dunger See	*****
Fallturm Bremen	*****
Badegewaesser: Kuhgraben See	*****
Schlachte Bremen	*****
Badegewaesser: Stadtwald See	*****
Badegewaesser: Wallerfeldmark See	*****

Query

Concept: **no simple concept query**

Location: **no location query**

Time: Jahre 1998 bis 2002

Matches

Concept:

Location:

Time: seit Jahr 2001

Info

Title: Fallturm Bremen

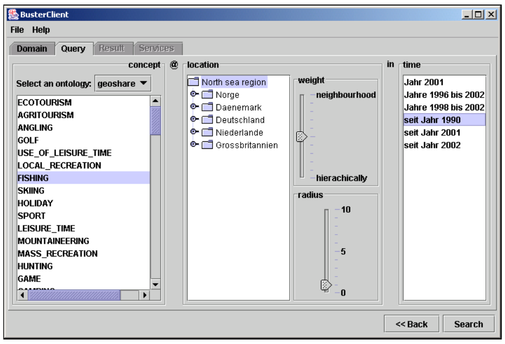
Identifier: http://www.tzi.de

Date: 2002-07-12

Language: de

Format: text/html

Figure 4.13: Result panel after querying the temporal part of BUSTER



Select an ontology: geoshare

ECOTOURISM
AGRITOURISM
ANGLING
GOLF
USE_OF_LEISURE_TIME
LOCAL_RECREATION
FISHING
SKIING
HOLIDAY
SPORT
LEISURE_TIME
MOUNTAINEERING
MASS_RECREATION
HUNTING
GAME

North sea region
Norge
Daenemark
Deutschland
Niederlande
Grossbritannien

Jahr 2001
Jahre 1996 bis 2002
Jahre 1998 bis 2002
seit Jahr 1990
seit Jahr 2001
seit Jahr 2002

Name	Relevance
Fischgewässer Bremen	*****
Spatial Management: CZM databas...	*****
WRRL - Bremen	*****

Query

Concept: FISHING
Location: North sea region
Time: seit Jahr 1990

Matches

Concept: ANGLING
Location: BREMEN, KRFR.ST.
Time: seit Jahr 2002

Info

Title: Fischgewässer Bremen
Identifier: harewms1.tzi.de/geosharewms1/htdocs/mapserv.phtml
Date: 2002-05-12
Language: de
Format: text/html

(a) Query panel

(b) Result panel

Figure 4.14: Query and result of BUSTER with a ”concept@location in time” type

and the non-existing solutions for intuitive labelling and annotation of data sources.

We developed a new representation scheme allowing us to define exact, fuzzy, persistent, and unknown boundaries. In addition, we are able to define internal relations or referrals which means that we can define a boundary of an interval with the help of a reference to the boundary of another interval. This leads to quite a number of possible combinations, which are supported as well.

Our developed and implemented temporal reasoning engine supports these requirements. The engine is a powerful tool to both check the underlying temporal model for consistency and derive new information hidden in the model. We think that this is an important step forward in the area of temporal annotation and reasoning with regard to the Semantic Web.

5.2 Future work

Future research concentrates on more relations that have to be integrated in the reasoning engine. We will also offer a small temporal reasoning service on the Web, which everybody is able to access to.

Another important step is to add more temporal relations and relax the restriction to *older*, *younger*, *contemporary*, *survives* and *survived-by*. A proper way to a solution would be using the conceptual neighborhoods *head-to-head* and *tail-to-tail* relations to declare the simultaneous beginning or end of time intervals.

References

Allen, J. F. 1984. “Maintaining Knowledge about Temporal Intervals.” *Communications of the ACM* 26 (11): 832–843.

Beek, Peter van. 1992. “Reasoning about qualitative temporal information.” *Artificial Intelligence* 58:297–326.

Bray, T., J. Paoli, and C.M. Sperberg-McQueen. 2000, 2000. “Extensible Markup Language (XML) 1.0 (second edition).” Technical Report REC-xml-20001006, W3C. http://www.w3.org/TR/2000/REC-xml-20001006 http://www.w3.org/TR/REC-xml.

Dechter, Rina, Itay Meiri, and Judea Pearl. 1991. "Temporal constraint networks." *Artificial Intelligence* 49:61–95.

Freksa, Christian. 1992. "Temporal Reasoning based on Semi-Intervals." *Artificial Intelligence* 54 (1): 199–227.

Gennari, Rosella. 1998. "Temporal Reasoning and Constraint Programming: A Survey." Masters Thesis, Universiteit van Amsterdam.

Giesenberg, Gabriele. 2002. "Kirchliches und städtisches Leben." In *Spuren der Jahrtausende - Archäologie und Geschichte in Deutschland*, edited by Uta von Freeden and Siegmar von Schnurbein, 390–417. Frankfurt am Main: Konrad Theiss Verlag Stuttgart.

Hayes, Patrick. 1996. "A Catalog of temporal theories." Technical report UIUC-BI-AI-96-01, University of Illinois 1995.

Hübner, Sebastian. 2003. "Qualitative Abstraktion von Zeit für Annotation und Retrieval im Semantic Web." Mastersthesis, Universität Bremen.

Ligozat, Gérard. 1996. "A new proof of tractability for ORD-Horn relations." *AAAI 96: 13th National Conference on Artificial Intelligence. IAAI 96: 8th Conference on Innovative Applications of Artificial Intelligence.* Portland, Oregon: AAAI-Press, 395–401.

———. 1998. ""Corner" relations in Allen's algebra." *Constraints* 3 (2/3): 165–177.

Nebel, Bernhard, and Hans-Jürgen Bürckert. 1995. "Reasoning about temporal relations: a maximal tractable subclass of Allen's interval algebra." *Journal of the ACM* 42 (1): 43–66.

Pitz, Ernst. 2002. "Mittelalter." In *Lexikon des Mittelalters*, Volume 6, 684–687. München: Deutscher Taschenbuch Verlag.

Schwalb, Eddie, and Rina Dechter. 1993. "Coping With Disjunctions In Temporal Constraint Satisfaction Problems." *The National Conference on Artificial Intelligence (AAAI-93).* Washington, D.C., July: AAAI Press, 127–132.

Schwalb, Eddie, and LLuís Vila. 1998. "Temporal Constraints: A Survey." *Constraints* 3 (2/3): 129–149.

Stock, Oliviero, ed. 1997. *Spatial and Temporal Reasoning.* Dordrecht, NL: Kluwer Academic Publishers.

The Unicode Consortium. 1996. "The Unicode Standard, Version 2.0." Technical paper, Addison-Wesley Developers Press.

Vila, LLuís. 1994. "A Survey on Temporal Reasoning in Artificial." *AI Communications* 7 (1): 4–28.

Vilain, Marc B., and Henry A. Kautz. 1986. "Constraint propagation algorithms for temporal reasoning." Edited by Tom Kehler and Stan Rosenschein, *AAAI.* Philadelphia, PA: AAAI Press, 377–382.

Visser, Ubbo. 2003. "Intelligent Information Integration for the Semantic Web." Habilitation - Overview article, University of Bremen.

Visser, Ubbo, and Gerhard Schuster. 2002. "Finding and Integration of Information - A Practical Solution for the Semantic Web -." Edited by Jérôme Euzénat, Asuncion Gomez-Perez, Nicola Guarino, and Heiner Stuckenschmidt, *ECAI 02, Workshop on Ontologies and Semantic Interoperability.* Lyon, France: ECCAI, 73–78.

W3C. 1998. “Date and Time Formats.” Technical Report, World Wide Web Consortium. W3C Note 27 August 1998, http://www.w3.org/TR/1998/NOTE-datetime-19980827, http://www.w3.org/TR/NOTE-datetime.

———. 2000, October. “Extensible Markup Language (XML) 1.0 (Second Edition).” Technical Report 6 October 2000, World Wide Web Consortium. W3C Recommendation, http://www.w3.org/TR/2000/REC-xml-20001006.

Zadeh, Lofti A. 1965. “Fuzzy Sets.” *Information and Control* 8:338–353.