Proceedings of the Twenty-First International FLAIRS Conference (2008)

Temporal Representation and Reasoning for the Semantic Web

Sebastian Hübner and **Ubbo Visser**
TZI - University of Bremen
Am Fallturm 1, 28359 Bremen, Germany
{shuebner|visser}@informatik.uni-bremen.de

Abstract

We introduce *period names*, a new approach for quantitative and qualitative temporal representation and reasoning of resources in the Semantic Web. Our work has been implemented into the Bremen University Semantic Translator for Enhanced Retrieval (BUSTER). This enables us to perform a query of the form *concept@location in time*.

Motivation and requirements

An important part of search in the Web is time-dependent. The W3C recommends a standard that is machine oriented, limited to single points in time, and offers neither meaningful labels nor qualitative relations. Humans however, make use of these features very often. There is thus a need for the development of a new temporal representation scheme and a reasoning engine. A small study of online news articles from BBC News and The New York Times revealed that the temporal terms can be classified into the following groups: (a) *Qualitative vs. quantitative*. The qualitative altervative covers terms like "now" and "for a long time", whereas the quantitative altervative covers terms like "on 21 November" and "in the last year". (b) *Absolute vs. relative*. The absolute altervative includes terms like "in 2001" and "beginning of November", whereas the relative altervative covers terms like "earlier" or "in two months".

Taking this into account, requirements for temporal annotation (and reasoning) in the Semantic Web can be formulated: a) *Intuitive labelling:* label time intervals with intuitive names. b) *Boundaries:* boundaries of time intervals should be flexible. The required types are *exact*, *fuzzy*, *persistent*, and *unknown*. c) *Structures:* an interval can be based on another interval, can be self-defined or imported. Functions (e.g. earliest_beginning_of) are needed to extract the significant time points from the intervals. d) *Explicit qualitative relations:* relations between intervals are needed when using persistent or unknown boundaries.

Related work

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

Articles in the knowledge representation area describe approaches about *Temporal Constraint Programming*, an area of temporal reasoning, e.g. (Artale, Parent, & Spaccapietra 2007). Temporal reasoning tasks are mainly formulated as CSPs. One of the most fundamental and well-known theories about reasoning with time intervals has been formulated by Allen (1984). A good summary concluding more than 1.600 research papers with respect to temporal issues in the *database* area has been published by Richard Snodgrass (2000). It contains a thorough description about timely issues and DBMS and introduces the concept of a 'period' (cf. (Rizzolo & Vaisman 2007)).

All of the approaches have deficiencies and most of them are also not capable of representing the concepts found in our motivation experiment.

Temporal representation with *period names*

We use XML notation to define the concepts and subconcepts. A *period name* consists of a header and a body. The header consists of the keyword *periodName* and an attribute called *id*, which labels the name of the period. The body consist of the definition of boundaries and relations. The most important property of a period is its expansion. The underlying model contains only intervals, which are non-empty and consist of more then one time point. The start point must lie before the end point. The basis of *boundaries* are period structures, which are constructed intervals using point structures. Point structures are bound and discrete. We assume a continuous time stream with discrete, ordered values. All time points can be ordered.

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., in the following denoted by -9999 and E is the time point 31.12.9999, 11:59pm, 59 seconds and 999 milliseconds, in the following denoted by +9999. As in usual calendars, the year zero does not exist.

There are four boundary types: *exact, fuzzy, persistent* and *unknown* and for definition purposis, two additional sets are necessary: (i) $\mathbf{P}$, a set of negative and positive persistent boundaries, $\mathbf{P} = \{P^-, P^+\}$, and (ii) U, a set of unknown boundaries $U = \{B^u\}$.

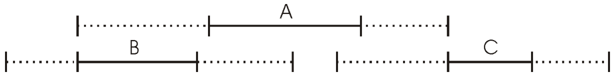

Figure 1: Graphical notation of fuzzy boundaries

Due to space limitations, we would like to give only one example: **Fuzzy boundaries:** They consist of two boundaries for both start and end point. An example from history would be the "beginning of the Middle Ages". The earliest begin is the end of the West-Roman Empire, the latest begin is the Reign of Karl the Great. Figure 1 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 and C's latest begin. These outer boundaries have referenced boundaries form B and C. We can calculate relations from exact and fuzzy boundaries.

Reasoning with *period names*

There are 30 relations between time intervals in total. We have started with the most important temporal relations with regard to the Semantic Web: *older* (ol), *younger* (yo), *contemporary* (ct), *survives* (sv), and *survived-by* (sb).

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 in this way. 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 etc.). A particular feature is the co-existence of quantitative descriptions of periods and qualitative relations of periods.

The reasoning engine considers relations between boundaries (e.g. before, after) and relations between two time periods (e.g. older younger, survives), We have shown (cf. (Visser 2005)) that relations between more than two time periods (e.g. older and younger, survives and survived-by) can be computed without Allen's composition table. The most important means are symmetry of the *ct* relation and transitivity of the *ol, yo, sv*, and *sb* relations. The inverse relations rule themselves out *(ol and yo; sv and sb)*, all other combinations are possible, e.g., $A\ ol\ B \land A\ ct\ B \land A\ sv\ B$. Also, we can aggregate the relations into two groups: (a) reflexive (1) and symmetric (2) *(ct)* and (b) non-reflexive (3), anti-symmetric (4), and transitive (5) *(ol, yo, sv, sb)*.

$$\forall p \in P \quad : \quad (p\ ct\ p) \tag{1}$$

$$\forall p_1, p_2 \in P \quad : \quad (p_1\ ct\ p_2) \longrightarrow (p_2\ ct\ p_1) \tag{2}$$

$$\forall p \in P \quad : \quad \neg(p\ ol\ p) \tag{3}$$

$$\forall p_1, p_2 \in P \quad : \quad \neg(p_1\ ol\ p_2 \land p_2\ ol\ p_1) \tag{4}$$

$$\forall p_1, p_2, p_3 \in P \quad : \quad (p_1\ ol\ p_2 \land p_2\ ol\ p_3) \longrightarrow (p_1\ ol\ p_3) \tag{5}$$

$$\forall p_1, p_2 \in P \quad : \quad (p_1\ ol\ p_2 \land p_1\ sv\ p_2) \longrightarrow (p_1\ ct\ p_2) \tag{6}$$

$$\forall p_1, p_2 \in P \quad : \quad (p_1\ yo\ p_2 \land p_1\ sb\ p_2) \longrightarrow (p_1\ ct\ p_2) \tag{7}$$

P is the set of all time periods. We can also show that a time period p_1 overlaps another time period p_2 if p_1 starts earlier and ends later (6); the inverse relations hold correspondingly (7).

Results

An example gives a notion of the representation and reasoning performance of our approach. We choose the term "Fall 2008" to show that *period names* are capable of integrating different temporal aspects into one new expression. There are various definitions of the term "Fall 2008". From the meteorological and climatic point of view, Fall starts on September 1st. This leads to a simple arrangement of the four seasons, e.g., "Winter 2008" starts on 1 December 2008, exactly 3 months later. From the astronomical point of view the definition of "Fall 2008" is different. The start depends on the date when night and day have the same length. This date differs from year to year. For example, in 2007 the astronomical fall started on 23 September, in 2008 on 22 September.

We satisfy the needs of both domains - meteorology and astronomy - and use the fuzzy boundary style for both beginning and end of "Fall 2008". For the earliest beginning, we choose 1 September (at midnight), and for the latest beginning we choose the exact astronomical start (at 16:45 MEZ). As our approach offers the opportunity to extract and reuse significant points in time arising from other period names, we take the definition of "Winter 2008" into account. Vice versa, we also use "Fall 2008" to define "Summer 2008". The concrete notation of the period names is shown on the poster due to space limitations.

We can show that the example covers the initially formulated requirements. Our temporal reasoning engine is able to draw conclusions based on the given information.

Conclusion and future work

We introduced our approach *period names* for the quantitative and qualitative temporal annotation of resources in the Semantic Web. Periods of time can be created of exact, fuzzy, persistent, and unknown boundaries.

Future research concentrates on the completion of our a temporal reasoning engine that already supports most of the needed features.

References

Allen, J. F. 1984. Towards a general theory of action and time. *Artificial Intelligence* 23(2):123–154.

Artale, A.; Parent, C.; and Spaccapietra, S. 2007. Evolving objects in temporal information systems. *Ann. of Mathem. and AI* 50:5–38.

Rizzolo, F., and Vaisman, A. A. 2007. Temporal XML: modeling, indexing, and query processing. *The VLDB Journal.* published online at Springer.

Snodgrass, R. T. 2000. *Developing Time-Oriented Database Applications in SQL.* Data Management Systems. Morgan Kaufmann.

Visser, U. 2005. *Intelligent Information Integration for the Semantic Web*, volume 3159 of *LNAI.* Heidelberg: Springer.