Information Retrieval: Ranking Results according to Calendar Criteria

Delphine Battistelli, Marcel Cori, Jean-Luc Minel, Charles Teissèdre

Information Retrieval: Ranking Results according to Calendar Criteria

Delphine Battistelli[1], Marcel Cori[2], Jean-Luc Minel[2], and Charles Teissèdre[2,3]

[1] STIH - Paris Sorbonne
28, rue Serpente, 75006 Paris France delphine.battistelli@paris-sorbonne.fr
[2] MoDyCo CNRS UMR 7114 - Université Paris Ouest
200 av. de la République, 92001 Nanterre, France mcori,jminel@u-paris10.fr
[3] Mondeca
3, cité Nollez, 75018 Paris, France charles.teissedre@mondeca.com

Abstract. Our work deals with calendar information as it is expressed in natural language (NL), that is to say through textual units such as prepositional phrases or noun phrases (e.g. *in the 90s, at the beginning of the XVth century*). We call these textual units Calendar Expressions (CE). Our work aims at showing how Information Retrieval systems can benefit from dealing with CE. In this paper we describe our overall approach which consists in a formal analysis of CEs that leads to a semantic representation. We then detail an algorithm that uses this representation to filter and rank CEs embedded in texts, according to a query containing a CE. The algorithm is integrated in an experimental search engine (called CaSE). Our representation of calendar information as it is expressed in NL and the function which computes the proximity between the two CEs, one in the text and the other in the query, provides a mean to process a query without any overlapping.

1 Introduction

Our work deals with the issue of searching for information according to calendar criteria, as they are expressed in natural language. Considering for example a corpus related to the history of the United States, a user could be interested in finding information such as "*prohibition at the beginning of the 30s*". There could be several more or less relevant answers to such a query in documents. One of the best answers could be "*In 1931, shortly before the end of Prohibition, Madden got out of bootlegging*", but an answer such as "*By the end of the 20s, Bureau of Prohibition agent Eliot Ness began an investigation of Capone and his business*" may also satisfy the user.

Leaving aside the issue of processing keywords such as "*prohibition*", our work focuses on the problem of linking the calendar information contained in the query with the calendar information we can find in the corpus. This requires a linguistic analysis of textual units conveying calendar information. It also requires being able to rank the set of units that are considered as relevant answers. We define a *Calendar Expression (CE)* as a more or less complex adverbial unit, including prepositions and/or other elements interacting directly with an explicit calendar reference. For example (the calendar references, which can be of different granularity levels, are underlined):

(E0) *in 1931*
(E1) *at the beginning of the 30s*
(E2) *three months before the beginning of the year 1985*
(E3) *until three months before the beginning of the 30s*
(E4) *between the end of the year 2007 and the beginning of March 2009*

One originality of our approach is that it is based on a theoretical modelling of CEs underpinned by a linguistic analysis. This approach is implemented and integrated in an experimental search engine (called CaSE). The CaSE system is divided into multiple steps for documents and query analysis (see figure 1): (1) CE annotation (parsing step); (2) a transduction step to transform the annotation outputs into Calendar Intervals; and (3) the core retrieval process.

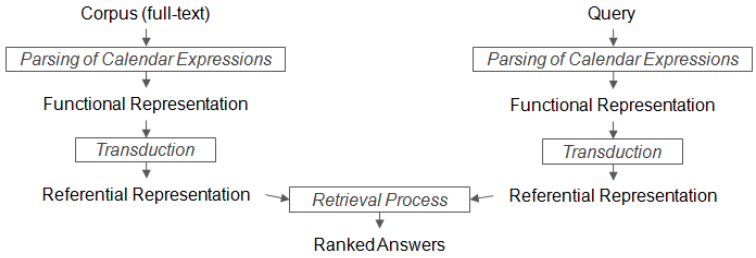

Fig. 1. Querying Process

Based on the resources for annotation described in [18], the first step of the processing chain annotates the CEs embedded in documents. As presented in Section 2, this first step delivers a linguistic analysis of these expressions, named *Functional Representation* (see section 2.1). Based on the formal transduction processing described in [5], the second step transforms this Functional Representation into what we name a *Referential Representation*, or Calendar Intervals (see sections 2.2 and 2.3). The retrieval algorithm used during the search process is described in section 3. It relies on an empirical function that computes a semantic distance evaluation between the Referential Representation of the CE embedded in the query and the Referential Representations of CEs found in documents. In section 4, we compare our approach to related work. We stress the fact that analysing calendar information within texts via textual units such as CEs (called ”temporal adverbials” in theoretical linguistics, see [12] for instance) is an innovative and intuitive way for IR systems that require some calendar criteria. It distinguishes our approach from all existing approaches in this area.

2 Formal Representation of Calendar Expressions

In this section, we describe our two-step approach to CE semantics. As detailed in [4], the first step provides a functional analysis of the different CE component units. The

second step transforms functional representations of CEs into referential ones. This transformation is useful for the last step of the processing chain: the ranking of CEs embedded in documents in response to a CE embedded in a user's query.

2.1 Functional Representation of Calendar Expressions

The Functional Representation of a CE is defined as a sequence of operations that act on a *Calendar Base (CB)* - see [4] for more details. The Calendar Base is the core calendar reference indicated by a textual unit (such as "*17th century*", "*June 6, 1897*", for instance). Several operators can successively apply upon the CB: (i) Zooming and Shifting Operators (corresponding to textual units such as "*mid*", "*three months before*"), (ii) Zoning Operators (corresponding to textual units such as "*at*", "*before*", "*until*") and (iii) Composition Operators (involved in CEs such as "*from the mid 80s to the mid 90s*"). More precisely, we consider four sets of operators:

OpZooming = {ZoomID, ZoomBegin, ZoomEnd, ZoomMiddle}
OpShifting = {Shift(u, n); $n \in \mathbb{Z}$ and u is a unit of time}
OpZoning = {ZoningID, Before, After, Until, Since}
OpComposition = {Between}

We define a *Functional Expression (FE)* as follows:
(i) if α is a CB or an FE, if $\Omega \in$ OpZooming $\cup$ OpShifting $\cup$ OpZoning, then $\Omega(\alpha)$ is an FE.
(ii) if α and β are two FEs, if $\Omega \in$ OpComposition, then $\Omega(\alpha, \beta)$ is an FE.

The parsing process associates an FE to each CE found in texts. For instance, the Functional Representation of the CEs *(E1)* to *(E4)* are:

(FE1) ZoningID (ZoomBegin(CB(decade: 1930)))
(FE2) ZoningID(Shift(month,-3) (ZoomBegin(CB(year: 1985))))
(FE3) ZoningUntil (Shift(month, -3) (ZoomBegin(CB(decade: 1930))))
(FE4) Between(ZoningID (ZoomEnd(CB(year: 2007))),
ZoningID(ZoomBegin(CB(month:3, year:2009))))

2.2 Referential Representation

In this section we describe a method, improved and augmented from [5], to transform the Functional Representation of CEs into its referential counterpart, also named *Calendar Intervals (CI)*. This step is useful to perform similarity comparisons between several CEs.

Calendar Units. We take a finite set of *units* $U = \{u, v, w, ...\}$, e.g. {*millennium, century, decade, year, month, day, ...* }. To each unit u is associated an infinite sequence:

$$S(u) = \langle \ldots, u_{-n}, \ldots, u_{-1}, u_0, u_1, \ldots, u_m, \ldots \rangle$$

which describes the succession of dates according to a given unit. For example, if u is *month*, $S(u)$ is a sequence such as $\langle \ldots$, 2010-11, 2010-12, 2011-1, 2011-2, $\ldots \rangle$.

We also take an order relation between units: we say that unit v is smaller than unit u, and we write $v \leq u$ (e.g. *day* $\leq$ *year*). When $v \leq u$, we define two mappings $b_{u\to v}$ and $e_{u\to v}$ such that:

(i) $\forall i \;\; b_{u\to v}(i) \leq e_{u\to v}(i)$

(ii) if $i_1 < i_2$ $e_{u\to v}(i_1) < b_{u\to v}(i_2)$

If $b_{u\to v}(i) = j$ and $e_{u\to v}(i) = k$, it means that v_j is the beginning of u according to v and that v_k is the end of u according to v. In particular, for each u and for each i, $b_{u\to u}(i) = i$ and $e_{u\to u}(i) = i$.

Calendar Intervals A *Calendar Interval* (CI) is given by an ordered pair of elements taken from one of the sequences $S(u) : \langle u_i, u_j\rangle$ (with $i \leq j$). We can also write: $\langle i, j, u\rangle$. u_i represents the date of the beginning of the CI, u_j represents the date of the end and u is the unit. Particular cases where $i = -\infty$ or $j = +\infty$ are admitted. The case of the empty CI, written $\emptyset$, is also admitted. For each CI $\langle i, j, u\rangle$ where the unit is u and for each v smaller than u we can associate its *image* according to v:

$$f_{u\to v}(\langle i, j, u\rangle) = \langle b_{u\to v}(i), e_{u\to v}(j), v\rangle$$

For instance, the image of the CI $\langle$1995-3, 1996-5$\rangle$ according to the *day* is the CI $\langle$1995-3-1, 1996-5-31$\rangle$.

Properties Given two CIs A and B whose units are respectively u and v, let w be the smallest unit among u and v, $f_{u\to w}(A) = \langle i, j, w\rangle$, $f_{v\to w}(B) = \langle k, l, w\rangle$.

The *intersection* of A and B is the CI $A \cap B = \langle max(i, k), min(j, l), w\rangle$ except if $max(i, k) > min(j, l)$. In this case $A \cap B = \emptyset$. We will say that A is *included* in B (or B *contains* A) iff $i \geq k$ and $j \leq l$. A equals B iff A is included in B and B is included in A.

The *relative length* of A and B ($B \neq \emptyset$) is the value: $rl(A/B) = \frac{j-i+1}{l-k+1}$

If $A = \emptyset$, $rl(A/B) = 0$ for each $B \neq \emptyset$.

If B is an infinite CI and $A \neq \emptyset$ we say that $rl(A/B) = \varepsilon$, a value greater than 0 but smaller than all other positive numbers.

If A and B are infinite CIs:

- if A is strictly included in B, $rl(A/B) = 1 - \varepsilon$
- if B is strictly included in A, $rl(A/B) = 1 + \varepsilon$
- if B equals A, $rl(A/B) = 1$

2.3 Calendar Intervals associated to Calendar Expressions

Given the Functional Representation of a CE, a translation process transforms it into a Calendar Interval. Remember that a Functional Expression is obtained by the successive application of four kinds of operators to a Calendar Base. Simultaneously, we consider four kinds of successive operations that are applied to the CI associated to a calendar base. This translation process enables us to associate a computed CI with each CE.

(1) To each CB we associate a CI for which the beginning is equal to the end, such as $\langle u_i, u_i\rangle$. For instance:

January 1985 : $\langle$1985-1, 1985-1$\rangle$
10th of January 1985 : $\langle$1985-1-10, 1985-1-10$\rangle$
80s : $\langle$198_ , 198_$\rangle$

(2) Let us consider an FE α with which a CI $\langle i, j, u\rangle$ is associated.

(2.1) If Ω is an operator of OpZooming we associate a CI to $\Omega(\alpha)$ for each unit v strictly smaller than u.

ZoomBegin	$\langle b_{u\to v}(i), b_{u\to v}(i) + \lfloor \tau(e_{u\to v}(j) - b_{u\to v}(i) + 1)\rfloor, v\rangle$
ZoomEnd	$\langle e_{u\to v}(j) - \lfloor \tau'(e_{u\to v}(j) - b_{u\to v}(i) + 1)\rfloor, e_{u\to v}(j), v\rangle$
ZoomMiddle	$\langle b_{u\to v}(i) + \lfloor \mu(e_{u\to v}(j) - b_{u\to v}(i) + 1)\rfloor,$ $e_{u\to v}(j) - \lfloor \mu(e_{u\to v}(j) - b_{u\to v}(i) + 1)\rfloor, v\rangle$

Here τ, τ' and μ are coefficients taken between 0 and 1 (or between 0 and $\frac{1}{2}$). In the following we will take $\tau = \tau' = \mu = 0.25$. Thanks to the floor function[4], we always obtain integers. Consequently, the results will differ depending on the unit taken into account. For instance, for the CE *(E1)* given in the Introduction, we obtain $\langle$1930, 1933$\rangle$ or $\langle$1930-1, 1933-4$\rangle$ or $\langle$1930-1-1, 1933-4-20$\rangle$.

(2.2) If Ω is an operator of OpShifting where v is a unit of U smaller than u (or equal to u), we associate a CI to $\Omega(\alpha)$.

ShiftBefore($v, -n$)	$\langle b_{u\to v}(i) - n, b_{u\to v}(i) - n, v\rangle$
ShiftAfter($v, +n$)	$\langle e_{u\to v}(j) + n, e_{u\to v}(j) + n, v\rangle$

For instance, for the CE *(E2)*, we obtain: $\langle$1984-10, 1984-10$\rangle$.

(2.3) If Ω is an operator of OpZoning we obtain for $\Omega(\alpha)$:

Before	$\langle -\infty, i-1, u\rangle$
After	$\langle j+1, +\infty, u\rangle$
Until	$\langle -\infty, j, u\rangle$.
Since	$\langle i, +\infty, u\rangle$.

For instance, for the CE *(E3)* we obtain $\langle -\infty$, 1929-10$\rangle$.

(3) Let us consider two FEs α and β with which two CIs $\langle i_1, j_1, u\rangle$ and $\langle i_2, j_2, v\rangle$ are associated. We take w, the largest unit smaller than u and v. To Between(α, β) we associate $\langle e_{u\to w}(j_1) + 1, b_{v\to w}(i_2) - 1, w\rangle$.

For instance, in order to represent *(E4)*, we obtain: $\langle$2008-1-1, 2009-2-28$\rangle$

3 Filtering and Ranking Answers for Retrieval Purposes

3.1 The Issue

We assume that we have a set $\mathcal{A} = \{A_1, A_2, \ldots, A_n\}$ of CEs translated into CIs. The CE used in a query is also translated into a CI, called Q. The goal is to extract a relevant subset $\mathcal{A}(Q) = \{A_{i_1}, A_{i_2}, \ldots, A_{i_p}\}$ from $\mathcal{A}$ and to order it from the most to the least relevant. To evaluate the relevance of CEs, we first consider an adequacy criterion, then, if necessary, an order criterion, in the event of equal adequacy.

[4] $\lfloor x\rfloor$ designates *floor*(x).

3.2 Adequacy Criteria

We take into account four kinds of criteria, from the best to the worst. These criteria can be described in terms of Allen relations [1].
(1) Equality: if any A_i is equal to Q, it is the best match of the query.
(2) Inclusion: if any A_i is included in Q, it also matches the query. It corresponds to A_i *during* Q or A_i *starts* Q or A_i *finishes* Q: for instance, if Q corresponds to *in 1980* and A_1 corresponds to *from March to May 1980*.
(3) Partial matching: if Q is included in any A_i, i.e. if Q occurs *during* A_i or Q *starts* A_i or Q *finishes* A_i we say that A_i contains Q (Containing criterion). For instance, if A_2 corresponds to *from 1978 to 1982*, A_1 is a better answer than A_2 to Q because all A_1 matches the query and not all A_2.
If A_i *overlaps* Q or Q *overlaps* A_i, we say that the Overlapping criterion holds: for instance, if A_3 corresponds to *from November 1979 to May 1980*. We do not consider that this kind of adequacy is better or worse than the Containing case.
(4) In all other cases, there is no matching between Q and A_i. Nevertheless we may keep some A_i according to proximity criteria.

3.3 Scoring the Answers

We first have to order several answers which satisfy the Inclusion criterion. The greater the coincidence area is, the better the answer will be. To do this, we take into account the value of $rl(A/Q)$ for an answer A to a query Q. For instance, A'_1 corresponding to *from February to November 1980* will be better than A_1 (see table 1).

To order answers A contained in Q, we take into account $rl(Q/A)$. For instance A'_2 corresponding to *from October 1979 to March 1981* is better than A_2.

In the case of Overlapping, we take into account two factors: the part of A included in Q relative to Q and the part of A included in Q relative to A. These two factors lead us to define two types of quantities: the *pertinence* of an answer A relative to a query Q, written $pert(A/Q)$, and the *precision* of an answer A relative to Q, written $prec(A/Q)$:

$$pert(A/Q) = rl((A \cap Q)/Q) \qquad prec(A/Q) = rl((A \cap Q)/A)$$

These two types of values are not of equal importance however. We therefore introduce a new coefficient, α, lower than 1 in order to compute a *score* for an answer relative to a query:

$$score(A/Q) = \frac{prec(A/Q) + \alpha\, pert(A/Q)}{1 + \alpha}$$

Several series of experiments led us to take $\alpha = 0.4$ but this value may be adjusted.

Note that in the Equality case we have $score(A/Q) = 1$. In the Inclusion case we have a value depending only on $rl(A, Q)$. In the Containing case we find a value depending only on $rl(Q, A)$. And in the No matching case we find $score(A/Q) = 0$.

This score can then be applied to all cases. This allows us to compare answers satisfying distinct criteria, and it can happen that an answer satisfying the Overlapping criterion or the Containing criterion has a better score than an answer satisfying the Inclusion criterion.

Table 1 shows some scores found for the query *in 1980*. We can see that an infinite CI is a possible answer, and its score can be compared to scores of finite CIs. We can also see that an answer satisfying the Inclusion criteria such as A_1'' can sometimes be assigned a lower score than an answer satisfying the Containing criterion such as A_2', because its coincidence area with the query is too small.

Table 1. Scores and distances for answers to two queries

	answers to the query *in 1980*	score
A_0	*in 1980*	1.
A_1'	*from February to November 1980*	0.952
A_1	*from March to May 1980*	0.785
A_2'	*from October 1979 to March 1981*	0.762
A_1''	*on May 25, 1980*	0.715
A_3	*from November 1979 to May 1980*	0.629
A_2	*from 1978 to 1982*	0.428
A_3''	*since January 1980*	0.285
A_3'	*since May 1980*	0.190
A_2''	*from July 1980 to June 2010*	0.154

answers to the query *since 1980*	prec	distance
since 1980	1.	0 year
year 1982	1.	2 years
since 1983	1.	3 years
from 1983 to 1986	1.	4 years
since 1978	$1-\varepsilon$	2 years
since 1975	$1-\varepsilon$	5 years
from 1979 to 1981	0.666	0 year
until 1984	ε	4 years
until 1975	0	5 years

3.4 Ordering the Answers

Let us define the *pole* of a CI. For $\langle i, \infty, u\rangle$ the pole is i, for $\langle -\infty, j, u\rangle$ the pole is j. If $\langle i, j, u\rangle$ is obtained by a ZoomBegin, the pole is i. If it is obtained by a ZoomEnd, the pole is j. Otherwise the pole is $\lfloor \frac{i+j}{2} \rfloor$.

The *distance* of two CIs A and B having the same unit u is defined as being the absolute value of the difference between the poles: $|pole(A) - pole(B)|$.

If two answers have the same score, we order them according to their distance from the query. The distance allows us to keep some answers with the score of zero if their distance from the query is not too great.

We also use the distance for ordering answers to an infinite query. If Q is an infinite CI, we do not consider the score but only the precision. If two answers have the same precision, those having the smallest distance from the query are preferred. The process is illustrated in table 1. The answers are presented from the best to the worst. ε is the symbol introduced in 2.2, representing a number smaller than all other numbers.

3.5 Implementation

This Retrieval model is integrated in a Lucene[5] based Search Engine with a standard Keyword Analyzer, for French and English documents. The system can handle queries

[5] `http://lucene.apache.org/`

that combine keywords and calendar expressions, such as *prohibition at the beginning of the 30s*, *Luther King around 1963* or *constitution 18th century*[6].

The processing chain annotates Calendar Expressions, computes Calendar Intervals, indexes the results and analyses queries during the retrieval process.

Calendar Expression annotation is performed thanks to the annotator described in [18]. It provides Functional Representations of CEs, as described in section 3.1 (for this task, the authors report a recall rate of 84.4% and a precision rate of 95.2% for French). The Calendar Intervals transducer module is fed with the annotator's output. It delivers Calendar Intervals as output. Once computed, Calendar Expressions are indexed along with the sentences in which they are embedded.

Documents in this context are seen as a set of sentences containing Calendar Expressions. The documents returned to a query are those that contain the best ranked sentences. Instead of the regular snippets provided by common search engines in the result list, the system displays the most relevant sentences for each document. The relevant extracts are sentences in order to ensure that the distance between the Calendar Expression and the keywords being searched is not too great. This is a provisional simplification of a complex linguistic problem, namely the scope of temporal adverbials, that is to say, how temporal adverbials are involved in discourse structuring and can thus contribute to the calendar anchoring of situations that are described in sentences following the one containing a temporal adverbial ([19], [13], [7]).

The query processor we implemented analyses queries submitted by users and separates the set of keywords (the thematic query) from the calendar criterion. The calendar criterion of the query is then used by the temporal retrieval module to provide a ranked list of documents.

4 Related Work

[2] highlight the importance of temporal information in Information Retrieval and note the scarcity of work in this area. Our work aims to contribute to filling this gap by showing two main points: 1) how IR systems (e.g. search engines) can benefit from taking into account calendar information, embedded both in documents and in queries; 2) how this applicative area can benefit from taking into account the way language expresses reference to calendar time via what is named in theoretical linguistics "temporal adverbials" (e.g. "*in 1998*", "*at the beginning of 1998*", "*since 1998*", "*two months before the end of 1998*").

Let us consider point 1). As mentioned in [6], calendar information as it is encoded in an expression such as "*in 1998*" is frequent across many kinds of documents and can be extracted with relative ease. However, it is not so immediately clear how it should be integrated into a retrieval model. Indeed, we can observe that almost all the approaches are based only on the publication dates of documents. For example, both [14] and [8] propose language models that take into account the publication date of documents, in order to favor, for instance, the most recent documents. [11] focus on

[6] The system can be tested at the following address:
`http://client1.mondeca.com/TemporalQueryModule/?locale=en`

constructing query-specific temporal profiles based on the publication date of relevant documents.

Among the very few approaches dealing with calendar information embedded in documents, we can first mention [10], which developed a temporal search engine supporting a Web search on temporal information embedded in Web pages. Secondly, and the closest to our approach, we can mention [3] which proposes a search engine capturing calendar information in documents. Their goal is to build clusters of documents and then rank documents in each cluster according to the calendar information that the documents contain. Like the other experimentations mentioned, they do not provide a means to express queries containing a Calendar Expression. In this sense, expressiveness is limited.

Our approach differs from all these approaches mainly by the fact that we consider temporal adverbial units, called here Calendar Expressions. Regarding point 2), with the exception of the approach adopted in [17], none of the approaches conducted in NLP (see [16]) and IR consider this kind of textual unit as a temporal expression in itself. This is mainly because the issue of analysing temporal information in texts by a named-entities approach has influenced (and is still influencing) a lot of studies. In a named-entities approach, only strict calendar reference is considered and analysed, e.g. "*1998*" in the previous examples. This is what is named a "Temporal Expression" (or a "Date") in the most popular annotation schemata (TIMEX2 [9] and TIMEX3 [15, 16]). All the approaches cited above use this kind of approach to calendar information in texts.

For a retrieval purpose (such as developing a search engine for instance), we believe that our approach is more intuitive and can lead to better performance from a user's point of view concerning the semantic relations between calendar information as it is expressed in a query and in a collection of documents. Indeed, our approach makes it possible, in a unified formal manner, to exploit (as pursued for example in [3]), the multi granularity of calendar information (e.g. "*on May, 25 1980*", "*in May 1980*"), and also the semantics of units that appear in a calendar expression (e.g. "*around/after/in May 1980*"). Let us imagine a user who is searching for events that occurred "*in May 1980*". Our system will establish a relation between the query and possible answers such as "*around May 1980*" and "*after May 1980*" given in this order. None of the above-mentioned approaches can currently deal with this kind of scenario.

5 Conclusion

We have described the main theoretical principles of our approach and a processing chain based on these principles which identifies Calendar Expressions in French or English texts and parses them in order to build functional representations. These representations are then transformed into referential representations. The whole process provides a way to compute a distance between a temporal query and these expressions. We have presented a heuristic function which provides a means to score and order all the answers. We have shown that this approach is able to process queries which contain different levels of calendar granularity.

We are currently extending Calendar Expression modelling so that the system can cope with expressions that refer to several areas on the Calendar system, such as iterative expressions (e.g.: "*every Monday in February 2011*") or aggregates (e.g. :"*April, 11, 12, 15 and 22, 2011*").

Acknowledgments. This project is partially granted by Chronolines ANR project (ANR-10-CORD-010).

References

1. Allen, J. F.: Maintaining Knowledge about Temporal Intervals. *Communications of the ACM*, 26 (1983) 832–843.
2. Alonso, O., Gertz, M., Baeza-Yates, R.: On the Value of Temporal Information in Temporal Information Retrieval. *SIGIR Forum*, 41(2) (2007) 35-41.
3. Alonso, O., Gertz, M., Baeza-Yates, R.: Clustering and Exploring Search Results using Timeline Constructions. *Proc. CIKM'09*, Hong Kong (2009).
4. Battistelli, D., Couto, J., Minel, J.-L., Schwer, S.: Representing and Visualizing Calendar Expressions in Texts. *Proc. STEP'08*, Venice, (2008).
5. Battistelli, D., Cori, M., Minel, J.-L., Teissèdre, C.: Semantics of Calendar Adverbials for Information Retrieval. *ISMIS'11*, *LNAI*, 6804 (2011) 622-631.
6. Berberich, K., Bedathur, S., Alonso, O., Weikum, G.: A Language Modeling Approach for Temporal Information Needs. *ECIR 2010*, *LNCS*, 5993 (2010) 13-25.
7. Charolles, M., Vigier, D.: Les adverbiaux en position préverbale : portée cadrative et organisation des discours. *Langue Française*, vol 2, no. 148 (2005) 9-30.
8. Dakka, W., Gravano, L., Ipeirotis, P.G.: Answering General Time-Sensitive Queries. *Proc. CIKM'08*, Napa Valley, California (2008).
9. Ferro, L., Gerber, L., Mani, I., Sundheim, B., Wilson, G.: TIDES 2005 Standard for the Annotation of Temporal Expressions. `http://fofoca.mitre.org/annotation_guidelines/2005_timex2_standard_v1.1.pdf`.
10. Jin, P., Lian, J., Zhao, X., Wan, S.: TISE: A Temporal Search Engine for Web Contents. In *IITA'08*, (2008).
11. Jones R., Diaz, F.: Temporal Profiles of Queries. *ACM Trans. Inf. Syst* (2007).
12. Klein, W.: Time in Language. *Routledge*, London (1994).
13. Le Draoulec, A., Péry-Woodley, M.-P.: Time Travel in Text: Temporal Framing in Narratives and Non-narratives. In *MAD 03*, Amsterdam (2003) 267-275.
14. Li, X., Croft, W.B.: Time-Based Language Models. *CIKM 2003* (2003).
15. Pustejovsky, J., Castano, J., Ingria, R., Saurí, R., Gaizauskas, R., Setzer, A., Katz, G.: TimeML: Robust Specification of Event and Temporal Expressions in Text. *IWCS-5* (2003).
16. Pustejovsky, J., Ingria, R., Saurí, R., Castano, J., Littman, J., Gaizauskas, R., Setzer, A., Katz, G., Mani, I.: The Specification Language TimeML. *The Language of Time. A Reader*, Oxford University Press Inc., New York (2005).
17. Schilder, F., Habel, C.: From Temporal Expressions to Temporal Information: Semantic Tagging of News Messages. *Proc. ACL-2001*, Workshop on Temporal and Spatial Information Processing, Toulouse, France (2001).
18. Teissèdre, C., Battistelli, D., Minel, J.-L.: Resources for Calendar Expressions Semantic Tagging and temporal Navigation through Texts. *Proc. LREC'10*, Malta (2010).
19. Van Raemdonck, D.: Est-il pertinent de parler d'une classe d'adverbes de temps ? *Actes de CLAC, 7* (2001).