

# Temporal Feature Attachment over Linked Data Information Access

Md-Mizanur Rahoman<sup>1</sup> \* Ryutaro Ichise<sup>1,2</sup>

<sup>1</sup> The Graduate University for Advanced Studies

<sup>2</sup> National Institute of Informatics

**Abstract:** Over the linked data information retrieval, adaptation of temporal features such as date, time or time of an event, is paid little attention. Therefore, we propose a keyword-based linked data information retrieval framework, called TLDRet, that can incorporate temporal features and able to generate more concise results. Preliminary evaluation of our system shows promising performance.

## 1 TLDRet: Linked Data Retrieval Framework with Temporal Semantics

Temporal Linked Data Retrieval framework (TLDRet) [1] is our proposed system. We adapt temporal semantics on the top of a keyword-based QA system [2].

The QA system uses template which resembles graph-like structure of linked data and tries to subsume some part of the linked data to generate possible information. In general, a template is pre-defined structure which holds some position holders and accomplishes tasks by setting those holders with task specific parameters. Position holders of templates that are used in the QA system are either filled-up by the input keywords (or more precisely by linked data resources which represent input keywords) or they are kept by the variables considering variables could be filled-up by some linked data resources. The QA system utilize these kept variables to pick possible information.

On the other hand, on text, temporal semantics is indicated by *signal words* [3]. For example, question *Which US President born during World War I?* holds the word *during* as *signal word* which informs that question holds some temporal feature in it. Therefore, according to the *signal word*, if we can slice the question into two parts: *signal word* prior keywords which we call question focused keyword set (Q-FKS) and *signal word* and its follower keywords which we

call question restriction keyword set (Q-RKS), then annotate all associated temporal values to a common standard (e.g., TIMEX3), and answer the question by imposing time filter between the output of Q-FKS and the output of Q-RKS, we can able to adapt the temporal feature related semantics. For example, if the example question is presented by the input keywords *{US President, birthday, during World War I}*, then Q-FKS = *{US President, birthday}* and Q-RKS = *{during World War I}* and we can impose time filter so that we can pick US President whose birthday is during World War I that answers the question.

We apply two-phase-based processing: phase 1 - query text processing, phase 2 - semantic query. In phase 1, TLDRet orders input keywords and annotates temporal value of *temporal keywords* to TIMEX3. Then in phase 2, TLDRet imposes a time filter to produce the intended result.

In phase 1: query text processing, we perform pre-processing tasks before adapting the temporal semantics. Pre-processing includes slicing input keywords into Q-FKS and Q-RKS and annotating Q-RKS to TIMEX3 annotated temporal value. If Q-RKS hold event information such as *World War I*, QA system is executed for Q-RKS and annotate temporal part of Q-RKS to TIMEX3 annotated temporal value. Stanford parser helps us to annotate TIMEX3 annotated temporal value.

In phase 2: semantic query, we filter the output of Q-FKS by the TIMEX3 annotated value of Q-RKS. Over temporal feature related question, Saquete et al., introduce *ordering key* [3]. The *ordering key* preserves temporal semantics of input keywords by intro-

\*連絡先: The Graduate University for Advanced Studies  
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan.  
E-mail: mizan@nii.ac.jp

ducing some kind of information validity constraint. *Ordering key* defines constraint of information validity that is constructed for three parameters: i) *signal word*, ii) temporal feature related part of Q-FKS input keywords iii) temporal feature related part of Q-RKS input keywords. With this constraint, *ordering key* incorporates temporal semantics of input keywords which, eventually, gives option of information filtering. For every *signal word*, it introduces constraint of information validity. Such as, for the example question where *signal word* is *during*, temporal feature related part of Q-FKS input keywords is *birthday* and temporal feature related part of Q-RKS input keywords is *World War I*, so the constraint of information validity is  $start\ of\ World\ War\ I \leq birthday \leq end\ of\ World\ War\ I$ . For *signal word*, Saquete et al., devised corresponding *ordering key*. Therefore, in phase 2, we execute QA system over Q-FKS keywords to find Q-FKS keywords related result, then we parse the result by a parser, and we annotate temporal feature related part to TIMEX3 values, next according to the *signal word* corresponding *ordering key*, we filter the TIMEX3 annotated result of the Q-FKS keywords related result. This filtered result is considered as our final output.

In experiment, we use the Question Answering over Linked Data (QALD)<sup>1</sup> open challenge question sets in our experiment. The QALD open challenge includes natural language question sets from DBpedia and MusicBrainz datasets, which are divided into QALD-1 and QALD-2. TLDRet can able to answer all DBpedia temporal feature related questions, and suffer for some MusicBrainz temporal feature related questions. TLDRet also outperforms QALD-2 open challenge participant systems named SemSek, Alexandria, MHE, and QAKiS over DBpedia test questions<sup>2</sup>. We conclude that our proposed method successfully adapts *signal word*, *ordering key*, and explore all temporal values to a common annotation which filter out possible information efficiently.

## 参考文献

- [1] M.-M. Rahoman and R. Ichise. A temporal semantic facilitated linked data retrieval frame-

work. In *Proceedings of the 3rd Joint International Semantic Technology Conference*, 2013.

- [2] M.-M. Rahoman and R. Ichise. An automated template selection framework for keyword query over linked data. In *Proceedings of the 2nd Joint International Semantic Technology Conference*, pages 175–190, 2012.
- [3] E. Saquete, J. L. V. González, P. Martínez-Barco, R. Muñoz, and H. Llorens. Enhancing QA systems with complex temporal question processing capabilities. *Journal of Artificial Intelligence Research*, 35:775–811, 2009.

<sup>1</sup><http://greententacle.techfak.uni-bielefeld.de/~cunger/qald/>

<sup>2</sup><http://greententacle.techfak.uni-bielefeld.de/~cunger/qald/2/dbpedia-test-questions.xml>