Iterative TempoWordNet

Mohammed Hasanuzzaman, Gaël Dias, Stéphane Ferrari, and Yann Mathet

Normandie University
UNICAEN, GREYC UMR 6072, France
first.last@unicaen.fr

Abstract. TempoWordNet (TWn) has recently been proposed as an extension of WordNet, where each synset is augmented with its temporal connotation: past, present, future or atemporal. However, recent uses of TWn show contrastive results and motivate the construction of a more reliable resource. For that purpose, we propose an iterative strategy that temporally extends glosses based on TWn^t to obtain a potentially more reliable TWn^{t+1}. Intrinsic and extrinsic evaluation results show improvements when compared to previous versions of TWn.

1 Introduction

There is considerable academic and commercial interest in processing time information in text whether that information is expressed explicitly or implicitly. Recognizing such information can significantly improve the functionality of Natural Language Processing and Information Retrieval applications.

Whereas most of the prior Computational Linguistics and Text Mining temporal studies have focused on temporal expressions and events, there has been a lack of work looking at the temporal orientation of word senses. Towards this direction, [1] developed TempoWordNet (TWn), an extension of WordNet [5], where each synset is augmented with its temporal connotation (past, present, future or atemporal). The authors show that improvements can be reached for temporal sentence classification and temporal query intent classification when unigrams are temporally expanded using the new constructed TWn.

In order to propose a more reliable resource, [3] recently defined two new propagation strategies, Probabilistic and Hybrid, respectively leading to two different TempoWordNet resources called TWnP and TWnH. Although some improvements are evidenced in their experiments, no conclusive remarks can be reached as TWnP evidences highest results in terms of temporal sentence classification but human judgment tends to prefer TWnH.

In this paper, we propose an iterative strategy to build an accurate TWn both in terms of human judgment and classification results. Our underlying idea is simple. Taking into account that (1) synsets are temporally classified based on their gloss content and (2) temporal sentence classification is boosted by TWn, temporally expanding glosses based on a given TempoWordNet at step t (TWn^t) may allow to obtain a more accurate TempoWordNet at step $t+1$ (TWn^{t+1}) when propagating temporal connotations. Our methodology is intrinsically and extrinsically evaluated in a similar way as [3]. Results show steady improvements of the current iterative TWn when compared to previous versions of TWn.

2 Methodology

To build a more reliable temporal lexical resource, we propose to rely on two previous findings. First, [1, 3] evidenced that synset time-tagging can be achieved by gloss classification with some success. The underlying idea is that the definition of a given concept embodies its potential temporal dimension. Second, [1, 3] also mention that temporal sentence classification can be improved when temporal unigrams are expanded with their synonyms stored in a TempoWordNet version.

From these two assumptions, a straightforward enhancement maybe proposed to improve the reliability of the temporal expansion process within WordNet. Indeed, glosses are sentences defining the concept at hand. As a consequence, temporally expanding glosses based on some TempoWordNet version may improve the performance of the classifier used to propagate the temporal connotations and as a consequence meliorate the intrinsic quality of the obtained temporal resource.

Note that this process can be iterated. If a better TempoWordNet can be obtained at some step t, better gloss expansion may be expected and as a consequence a more accurate temporal resource may be obtained at step $t+1$, which in turn may be reused for gloss expansion, and so on and so forth. This iterative strategy is defined in algorithm 1, where the process stops when a stopping criterion is satisfied.

Algorithm 1 Iterative TempoWordNet algorithm

1: $sl \leftarrow$ list of temporal seed synsets
2: $ps \leftarrow$ propagation strategy
3: $i \leftarrow 0$
4: $\mathrm{Wn}^0 \leftarrow \mathrm{Wn}$
5: **repeat**
6: $\quad \mathrm{TWn}^i \leftarrow \mathrm{PropagateTime}(\mathrm{Wn}^i, sl, ps)$
7: $\quad \mathrm{Wn}^{i+1} \leftarrow \mathrm{ExpandGloss}(\mathrm{Wn}, \mathrm{TWn}^i)$
8: $\quad i \leftarrow i + 1$
9: **until** stopping criterion
10: return TWn^i

Let sl be a set of manually time-tagged seed synsets, ps a propagation strategy (e.g. hybrid) and a working version of WordNet Wn^0 (e.g. WodNet 3.0). The iterative construction of TempoWordNet is as follows. The first TempoWordNet TWn^0 is constructed in a similar manner as proposed in [3]. Based on sl, an initial two-class (temporal vs atemporal) classifier is learned over a 10-fold cross validation process, where synsets are represented by their gloss constituents (i.e. unigrams) weighted by gloss frequency. The propagation of the temporal connotations is then run based the learned classifier and the given propagation strategy ps. New expanding synsets are included in the initial seeds list and the exact same learning process iterates N times so to ensure convergence in terms of classification accuracy. At the end of this process, the synsets of the working version of WordNet Wn^0 are either tagged as temporal or atemporal depending on classification probability, thus giving rise to two distinct partitions Wn^0_t (temporal Wn^0) and $\mathrm{Wn}^0_{\bar{t}}$ (atemporal Wn^0). In order to fine tune Wn^0_t into past, present and future synsets, the exact same procedure is run exclusively based on the past, present and future synsets from sl over Wn^0_t. When classification convergence is reached[1], all synsets from Wn^0_t are time-tagged as past, present or future giving rise to Wn^0_{ppf}. Note that a synset can neither be exclusively past, present nor future (e.g. "sunday") although it is temporal. In this case, near equal class probabilities are evidenced. At the end of the construction process, $\mathrm{TWn}^0 = \mathrm{Wn}^0_{\bar{t}} \cup \mathrm{Wn}^0_{ppf}$.

Once TWn^0 has been constructed, it can be used to temporally expand Wn glosses giving rise to the second working version of WordNet Wn^1. In particular, within each gloss of Wn, all temporal words from TWn^0 are searched for. If one temporal concept is found in the gloss, then its synonyms are added to the gloss thus enlarging the possible lexical overlap between temporal glosses. So, each gloss is now represented by its unigrams plus the synonyms of its temporal constituents as a bag of words and is noted Wn^1. Note that word sense disambiguation is performed based on the implementation of the Lesk algorithm of the NLTK toolkit[2]. This process refers to line 7 of algorithm 1.

So, from the new Wn^1, the next TempoWordNet version TWn^1 can be processed, which in turn can give rise to a new WordNet working version Wn^2, which will lead to TWn^2 and so on and so forth. This iterative process stops when some stopping criterion is reached. Many ideas

[1] N iterations are performed.
[2] `http://www.nltk.org/`. Last accessed: 15.08.2015

can be presented but within the scope of this paper, we propose that the final TempoWordNet version is obtained when the difference between TWn^i and TWn^{i-1} is marginal in terms of temporal sets, i.e. $\text{TWn}^i \setminus \text{TWn}^{i-1} \leq \epsilon$, which means that the set of distinct temporal synsets converges.

TempoWordNet version	TWnL	TWnP	TWnH^0	TWnH^1	TWnH^2	TWnH^3
# temporal synsets	21213	53001	17174	2020	2804	2832
# past synsets	1734	2851	2547	305	120	1308
# present synsets	16144	19762	842	1247	2181	765
# future synsets	3335	30388	13785	468	503	759
# atemporal synsets	96402	64614	100441	115639	114855	114827
$\text{TWn}^i \setminus \text{TWn}^{i-1}$	-	-	-	15154	784	28
Fixed-marginal κ	0.507	0.520	0.420	0.625	0.616	0.599
Free-marginal κ	0.520	0.520	0.440	0.850	0.700	0.774

Table 1. Comparative features of different TempoWordNet versions.

TempoWordNet version	without TWn	TWnL	TWnP	TWnH^0	TWnH^1	TWnH^2	TWnH^3
Sentence classification	64.8	66.7	69.3	68.6	68.4	69.7	71.4
Tweet classification	39.7	49.1	51.5	49.8	51.9	52.5	53.1
Query intent classification	75.3	78.0	78.8	75.9	78.3	79.0	80.1

Table 2. F_1-measure results for temporal sentence, tweet and query intent classification with different TempoWordNet versions performed on 10-fold cross validation with SVM.

3 Evaluation

Our methodology is evaluated both intrinsically and extrinsically. The underlying idea being that a reliable resource must evidence high quality time-tagging as well as improved performance for some application. As for experimental setups, we used (1) the *sl* seeds list provided by [1], (2) the *ps* hybrid propagation strategy[3] proposed in [3], (3) version 3.0 of WordNet[4] for Wn and (4) $\epsilon = 100$. With respect to the learning procedures, the SVM implementation of Weka[5] was used with parameters tuning and convergence was ensured by iterating the temporal propagation $N = 50$ times. Note that all experimental results as well as produced resources are freely available[6] for the sake of reproducibility.

3.1 Intrinsic Evaluation

In order to assess human judgment about the temporal parts of TempoWordNet, inter-rater agreement (free-marginal and fixed-marginal multirater kappa[7]) with multiple raters is performed. Three annotators are presented with 50 temporal synsets and respective glosses, and must decide upon their correct classification i.e. temporal or atemporal. Note that past, present and future connotations are only indicative of the temporal orientation of the synset but cannot be taken as a strict class. Indeed, there are many temporal synsets, which are neither past, present nor future (e.g. "monthly"). Results are reported in Table 1 and assess moderate agreement for previous versions of TempoWordNet (TWnL [1], TWnP [3] and TWnH^0 [3]) while substantial agreement is obtained for the successive iterative versions. Table 1 also presents figures about

[3] Note that lexical propagation does not spread over a wide range of concepts [1] and probabilistic propagation shows semantic shift problems [3].
[4] `https://wordnet.princeton.edu/`. Last accessed: 17.04.2016
[5] `http://www.cs.waikato.ac.nz/ml/weka/`. Last accessed: 15.08.2015
[6] `https://tempowordnet.greyc.fr/`. Last accessed: 17.04.2016
[7] `http://justusrandolph.net/kappa/`. Last accessed: 17.04.2016

the distribution of temporal synsets of each TempoWordNet version. Interestingly, the iterative versions tend to time-tag a much smaller proportion of synsets when compared to previous ones[8].

3.2 Extrinsic Evaluation

In order to produce comparative results with prior works, we test our methodology on the balanced data set produced in [1], which consists of 1038 sentences equally distributed as past, present and future. Moreover, we propose to extend experiments on a corpus of 300 temporal tweets[9]. This corpus contains 100 past, 100 present and 100 future tweets, which have been time-tagged by annotators of the crowdflower[10] platform[11]. For both experiments, each sentence/tweet is represented by a feature vector where each attribute is either a unigram[12] or a synonym of any temporal word contained in the sentence/tweet and its value is tf.idf.

In order to strengthen comparative evaluation, we also propose to tackle the TQIC task of NTCIR-11 Temporalia [4], where a given search engine query must be tagged as past, recency, future or atemporal. For that purpose, we use the 400 queries provided by the organisers, which are equally distributed by temporal class. In particular, a query is represented in the same way as sentences and tweets[13] plus the additional time-gapped feature proposed in [2]. Comparative classification results are reported in Table 2. For all experiments, TWnH^3 produces highest classification results with respectively 2.1%, 1.6% and 1.3% improvements for sentence, tweet and query temporal classification over the second best (non-iterative) TempoWordNet version. Note that all improvements are statistically relevant and steadily occur between TWnH^0, TWnH^1, TWnH^2 and TWnH^3.

4 Conclusions

We proposed an iterative strategy to produce a reliable temporal lexical resource called TempoWordNet. The underlying idea is that based on the temporal expansion of glosses with some version of TempoWordNet at step t, a more accurate resource can be obtained at step $t + 1$. Intrinsic and extrinsic evaluations evidence improved results when compared to recent versions of TempoWordNet. However, deeper experiments must be performed with respect to (1) the stopping criterion of the iterative process, (2) the hard (past, present, future) classification of temporal synsets and (3) the integration of TempoWordNet into temporal classifcation tools.

References

1. G. Dias, M. Hasanuzzaman, S. Ferrari, and Y. Mathet. Tempowordnet for sentence time tagging. In *Companion Publication of the 23rd International Conference on World Wide Web Companion (WWW)*, pages 833–838, 2014.
2. M. Hasanuzzaman, G. Dias, and S. Ferrari. Hultech at the ntcir-11 temporalia task: Ensemble learning for temporal query intent classification. In *NTCIR-11 Conference (NTCIR)*, pages 478–482, 2014.
3. M. Hasanuzzaman, G. Dias, S. Ferrari, and Y. Mathet. Propagation strategies for building temporal ontologies. In *14th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pages 6–11, 2014.
4. H. Joho, A. Jatowt, R. Blanco, H. Naka, and S. Yamamoto. Overview of ntcir-11 temporal information access (temporalia) task. In *NTCIR-11 Conference (NTCIR)*, pages 429–437, 2014.
5. G. A. Miller. Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41, 1995.

[8] Note that the number of past, present and future synsets is based on the highest probability given by the temporal classifier, which does not necessarily imply that the synset belongs to the given class (e.g. almost equal probabilities can be evidenced).
[9] `https://tempowordnet.greyc.fr/`. Last accessed: 17.04.2016
[10] `http://www.crowdflower.com/`. Last accessed: 17.04.2016
[11] Annotation details are out of the scope of this paper.
[12] Stopwords removal is performed so to better access the benefits of TempoWordNet.
[13] Stopwords are not removed as queries are small.