

Extracting Events with Informal Temporal References in Personal Histories in Online Communities

Miaomiao Wen, Zeyu Zheng, Hyeju Jang, Guang Xiang, Carolyn Penstein Rosé

Language Technologies Institute, Carnegie Mellon University
{mwen, zeyuz, hyejuj, guangx, cprose}@cs.cmu.edu

Abstract

We present a system for extracting the dates of illness events (year and month of the event occurrence) from posting histories in the context of an online medical support community. A temporal tagger retrieves and normalizes dates mentioned informally in social media to actual month and year referents. Building on this, an event date extraction system learns to integrate the likelihood of candidate dates extracted from time-rich sentences with temporal constraints extracted from event-related sentences. Our integrated model achieves 89.7% of the maximum performance given the performance of the temporal expression retrieval step.

1 Introduction

In this paper we present a challenging new event date extraction task. Our technical contribution is a temporal tagger that outperforms previously published baseline approaches in its ability to identify informal temporal expressions (TE) and that normalizes each of them to an actual month and year (Chang and Manning, 2012; Strotgen and Gertz, 2010). This temporal tagger then contributes towards high performance at matching event mentions with the month and year in which they occurred based on the complete posting history of users. It does so with high accuracy on informal event mentions in social media by learning to integrate the likelihood of multiple candidate dates extracted from event mentions in time-rich sentences with temporal constraints extracted from event-related sentences.

Despite considerable prior work in temporal information extraction, to date state-of-the-art resources are designed for extracting temporally scoped facts about public figures/organizations from newswire or Wikipedia articles (Ji et al., 2011; McClosky and Manning, 2012; Garrido et

[11/15/2008] I have noticed some pulling recently and I won't start **rads** until *March*.
[11/20/2008] It is slowwwly healing, so slowly, in fact, that she said she HOPES it will be healed by *March*, when I am supposed to start **rads**.
[1/13/2009] I still have one last chemo to go on *the 19th* and then start **rads** in *5 wks*.
[1/31/2009] I go for my first meeting with the **rad** onc on *2/10* (my *50th birthday!*).
[2/23/2009] I had my first **rad** *today*.
[3/31/2009] *Tomorrow* will be my last full **rads**
[4/2/2009] I started **rads** in *Feb*, just did #29 *today*.
[4/8/2009] The **rad** onc wants to see me again *next week* for a skin check as I have had cellulitis twice since *August*.
[6/21/2010] My friend Lisa had her port put in *last week* and will begin *2 weeks* of **radiation** on *Tuesday*.

Figure 1: User posts containing keywords for the start of Radiation. Event keywords are in bold and temporal expressions are in italics.

al., 2012). When people are instead communicating informally about their lives, they refer to time more informally and frequently from their personal frame of reference rather than from an impersonal third person frame of reference. For example, they may use their own birthday as a time reference. The proportion of relative (e.g., “last week”, “two days from now”), or personal time references in our data is more than one and a half times as high as in newswire and Wikipedia. Therefore, it is not surprising that there would be difficulty in applying a temporal tagger designed for newswire to social media data (Strotgen and Gertz, 2012; Kolomiyets et al., 2011). Recent behavioral studies (Choudhury et al., 2013; Park and Choi, 2012; Wen et al., 2012) demonstrate that user-focused event mentions extracted from social media data can provide a useful timeline-like tool for studying how behavior patterns change over time in response to mentioned events. Our research contributes towards automating this work.

2 Task

Our task is to extract personal illness events mentioned in the posting histories of online community participants. The input to our system is

a candidate event and a posting history. The output is the event date (month and year) for the event if it occurred, or “unknown” if it did not occur. The process iterates through a list of 10 cancer events (CEs). This list includes breast cancer Diagnosis, Metastasis, Recurrence, Mastectomy, Lumpectomy, Reconstruction, Chemotherapy-Start, Chemotherapy-End, Radiation-Start and Radiation-End. For each of these target CEs, we manually designed an event keyword set that includes the name of the event, abbreviations, slang, aliases and related words.

For each of the 10 events, all sentences that mention a related event keyword are extracted from the user’s posting history. Figure 1 shows several sentences that were extracted for one user for the start date of Radiation. The task is to determine that the beginning of this user’s Radiation therapy was 2/2009. Note that the user began to post about Radiation before she started it. She first reported planning to start Radiation in March, but then rescheduled for February. Most of the TEs are non-standard and need to be resolved to calendar dates (year and month).

Once the full set of event mention sentences has been extracted for a user, all the temporal expressions (TEs) that appear in the same sentence with an event mention are resolved to a set of candidate dates. Besides a standard event-time classifier for within-sentence event-time anchoring, we leverage a new source of temporal information to train a constraint-based event-time classifier. Previous work only retrieves time-rich sentences that include both the query and some TEs (Ji et al., 2011; McClosky and Manning, 2012; Garrido et al., 2012). However, sentences that contain only the event mention but no explicit TE can also be informative. For example, the post time (usually referred to as document creation time or DCT) of the sentence “metastasis was found in my bone” might be labeled as being *after* the “metastasis” event date. These DCTs impose constraints on the possible event dates, which can be integrated with the event-time classifier, as a variant on related work (Chambers, 2012).

3 Related Work

Previous work on TE extraction has focused mainly on newswire text (Strotgen and Gertz, 2010; Chang and Manning, 2012). This paper presents a rule-based TE extractor that identifies

and resolves a higher percentage of nonstandard TEs than earlier state-of-art temporal taggers.

Our task is closest to the temporal slot filling track in the TAC-KBP 2011 shared task (Ji et al., 2011) and timelining task (McClosky and Manning, 2012). Their goal was to extract the temporal bounds of event relations. Our task has two key differences. First, they used newswire, Wikipedia and blogs as data sources from which they extract temporal bounds of facts found in Wikipedia infoboxes. Second, in the KBP task, the set of gold event relations are provided as input, so that the task is only to identify a date for an event that is guaranteed to have been mentioned. In our task, we provide a set of potential events. However, most of the candidate events won’t have ever been reported within a user’s posting history.

Temporal constraints have proven to be useful for producing a globally consistent timeline. In most temporal relation bound extraction systems, the constraints are included as input rather than learned by the system (Talukdar et al., 2012; Wang et al., 2011). A notable exception is McClosky et al. (2012) who developed an approach to learning constraints such as that people cannot attend school if they have not been born yet. A notable characteristic of our task is that constraints are softer. Diseases may occur in very different ways across patients. Recurring illnesses falsely appear to have an unpredictable order. Thus, there can be no universal logical constraints on the order of cancer events.

Our approach to using temporal constraints is a variant on previously published approaches. Garrido et al. (2012) made use of DCT (document creation time) as well, however, they have assumed the DCT is within the time-range of the event stated in the document, which is often not true in our data. Chambers (2012) utilized the within-sentence time-DCT relation to learn constraints for predicting DCT. We learn the event-DCT relations to produce constraints for the event date.

4 Corpus Annotation

We have scraped the posts, users, and profiles from a large online cancer support community. From this collection we extracted and then annotated two separate corpora, one for evaluating our TE retrieval and normalization, the other one for event date extraction.

For creating the TE extraction corpus, we ran-

domly picked one post from each of 1,000 randomly selected users. We used this sampling technique because each user tends to use a narrow range of date expression forms. From these posts, we manually extracted 601 TEs and resolved them to a specific month and year or just year if the month was not mentioned. Events not reported to have occurred were annotated as “unknown”. Our corpus for event date extraction consists of the complete posting history of 300 users that were randomly drawn from our dataset. Three annotators were provided with guidelines for how to infer the date of the events (Wen et al., 2013). We achieved .94 Kappa on identification of whether an event has a reported event date in a user’s history or not. In evaluation of agreement on extracted dates, we achieved a .99 Cronbach’s alpha. From this corpus, 509 events were annotated with occurrence dates (year and month). In our evaluation, we use data from 250 users for training, and 50 for testing.

5 Method

Now we explain on a more technical level how our system works on our task. Given an event and a user’s post history, the system searches for all of the sentences that contain an event keyword (*keyword sentence*) and all the sentences that contain both a keyword and a TE (*date sentence*). The TEs in the *date sentences* are resolved and then used as candidate dates for the event. For selecting among candidate dates, our model integrates two main components. First, the Date Classifier is trained from *date sentences* to predict how likely its candidate TE and the gold event date are to *overlap*. Then, because constraints over event dates can be informed by temporal relations between the event date and the DCT, the Constraint-based Classifier provides an indication of the plausibility of candidate dates. The integrated system combines the predictions from both classifiers.

5.1 Temporal Tagger

We design a rule-based temporal tagger that is built using regular expression patterns to recognize informal TEs. Similar to SUTime (Chang and Manning, 2012), we identify and resolve a wide range of non-standard TE types such as “Feb ’07 (2/2007)”. The additional types of TE we handle include: 1)**user-specific TEs**: A user’s age, cancer anniversary and survivorship can provide

temporal information about the user’s CEs. We obtain the birth date of users from their personal profile to resolve age date expressions such as “at the age of 57”. 2)**non-whole numbers** such as “a year and half” and “1/2 weeks”. 3)**abbreviations of time units** : e.g. “wk” as the abbreviation of “week”. 4)**underspecified month mentions**, we resolve the year information according to the DCT month, the mentioned month and the verb tense.

5.2 Date Classifier

We train a MaxEnt classifier to predict the temporal relationship between the retrieved TE and the event date as *overlap* or *no-overlap*, similar to the within-sentence event-time anchoring task in TempEval-2 (UzZaman and Allen, 2010). Features for the classifier include many of those in (McClosky and Manning, 2012; Yoshikawa et al., 2009): namely, event keyword and its dominant verb, verb and preposition that dominate TE, dependency path between TE and keyword and its length, unigram and bigram word and POS features. New features include the Event-Subject, Negative and Modality features. In online support groups, users not only tell stories about themselves, they also share other patients’ stories (as shown in Figure 1). So we add subject features to remove this kind of noise, which includes the governing subject of the event keyword and its POS tag. Modality features include the appearance of modals before the event keyword (e.g., may, might). Negative features include the presence/absence of negative words (e.g., no, never). These two features indicate a hypothetical or counter-factual expression of the event.

To calculate the likelihood of a candidate date for an event, we need to aggregate the hard decisions from the classifier. Let DS_u be the set of the user’s *date sentences*, let D_u be the set of dates resolved from each TE. We represent a MaxEnt classifier by $P_{relation}(R|t, ds)$ for a candidate date t in *date sentence* ds and possible relation $R = \{overlap, no-overlap\}$. We map the distribution over relations to a distribution over dates by defining $P_{DateSentence}(t|DS_u)$:

$$P_{DateSentence}(t|DS_u) = \frac{1}{Z(D_u)} \sum_{t_j \in D_u} \delta_{t_j}(t) P_{relation}(overlap|t_j, ds_j) \quad (1)$$

$$\delta_{t_j}(t) = \begin{cases} 1 & \text{if } t = t_j \\ 0 & \text{otherwise} \end{cases}$$

We refer to this model as the Date Classifier.

5.3 Constraint-based Classifier

Previous work only retrieves time-rich sentences (i.e., *date sentences*) (Ling and Weld, 2010; Ji et al., 2011; McClosky and Manning, 2012; Garrido et al., 2012). However, *keyword sentences* can inform temporal constraints for events and therefore should not be ignored. For example, “Well, I’m officially a Radiation grad!” indicates the user has done radiation by the time of the post (DCT). “Radiation is not a choice for me.” indicates the user probably never had radiation. The topic of the sentence can also indicate the temporal relation. For example, *before* chemotherapy, the users tend to talk about choices of drug combinations. *After* chemotherapy, they talk about side-effects.

This section departs from the above Date Classifier and instead predicts whether each *keyword sentence* is posted *before* or *overlap-or-after* the user’s event date. The goal is to automatically learn time constraints for the event. This task is similar to the sentence event-DCT ordering task in TempEval-2 (UzZaman and Allen, 2010). We create training examples by computing the temporal relation between the DCT and the user’s gold event date. If the user has not reported an event date, the label should be *unknown*.

We train a MaxEnt classifier on each event mention paired with its corresponding DCT. All the features used in the classifier component that are not related to the TEs are included. Let KS_u be the set of the user’s *keyword sentences*, let D_u be the set of dates resolved from each *date sentence*. We define a MaxEnt classifier by $P_{relation}(R|ks)$ for a keyword sentence ks and possible relation $R = \{before, overlap-or-after, unknown\}$. DCT is the post time of the *keyword sentence* ks . The $rel(DCT, t)$ function simply determines if the DCT is *before* or *overlap-or-after* the candidate date t . We map this distribution over relations to a distribution over dates by defining $P_{KeywordSentence}(t, KS_u)$:

$$P_{KeywordSentence}(t, KS_u) = \frac{1}{Z(D_u)} \sum_{ks_j \in KS_u} P_{relation}(rel(dct_j, t)|ks_j) \quad (2)$$

$$rel(dct, t) = \begin{cases} before & \text{if } dct < t \\ overlap-or-after & \text{if } dct \geq t \end{cases}$$

5.4 Integrated Model

Given the Date Classifier of Section 5.2 and the Constraint-based Classifier of Section 5.3, we create a *Integrated Model* combining the two with the following linear interpolation as follows:

$$P(t|posts_u) = \lambda P_{DateSentence}(t|DS_u) + (1 - \lambda) P_{KeywordSentence}(t|KS_u)$$

where t is a candidate event date. The system will output t that maximizes $P(t|posts_u)$ and *unknown* if DS_u is empty. λ was set to 0.7 by maximizing accuracy using five-fold cross-validation over the training set.

6 Evaluation Metric and Results

6.1 Temporal Expression Retrieval

We compare our temporal tagger’s performance with SUTime (Chang and Manning, 2012) on the 601 manually extracted TEs. We exclude user-specific TEs such as birthday references since SUTime cannot handle those. We first evaluate identification of the extent of a TE and then production of the correctly resolved date for each recognized expression. Table 1 shows that our tagger has significantly higher precision and recall for both.

| | | P | R | F1 |
|---------------|------------|-------------|-------------|-------------|
| Extents | SUTime | 97.5 | 75.4 | 85.0 |
| | Our tagger | 97.9 | 91.8 | 94.8 |
| Normalization | SUTime | 89.4 | 71.2 | 79.3 |
| | Our tagger | 91.3 | 85.5 | 88.3 |

Table 1: Temporal expression retrieval results

6.2 Event-date Extraction

6.2.1 Evaluation metric

The extracted date is only considered correct if it completely matches the gold date. For less than 4% of users, we have multiple dates for the same event (e.g., a user had a mastectomy twice). Similar to the evaluation metric in a previous study (Ji et al., 2011), in these cases, we give the system the benefit of the doubt and the extracted date is considered correct if it matches one of the gold dates. In previous work (McClosky and Manning, 2012; Ji et al., 2011), the evaluation metric score is defined as $1/(1 + |d|)$ where d is the difference between the values in years. We choose a much stricter evaluation metric because we need a precise event date to study user behavior changes.

6.2.2 Baselines and oracle

Based on our temporal tagger, we provide two baselines to describe heuristic methods of aggregating the hard decisions from the classifier

| | | Baseline1 | | | Baseline2 | | | Date | | | Integrated | | | Oracle |
|----------------|-------|-----------|-----|-----|-----------|-----|-----|------|-----|-----|------------|-----|-----|--------|
| CE | count | P | R | F1 | P | R | F1 | P | R | F1 | P | R | F1 | F1 |
| Diagnosis | 112 | .64 | .70 | .67 | .60 | .66 | .63 | .68 | .75 | .71 | .68 | .75 | .71 | .80 |
| Metastasis | 7 | .16 | .58 | .25 | .12 | .43 | .19 | .25 | .86 | .39 | .25 | .86 | .39 | .86 |
| Recurrence | 14 | .14 | .35 | .20 | .11 | .29 | .16 | .13 | .36 | .19 | .13 | .36 | .19 | .47 |
| Chemo-start | 54 | .49 | .61 | .54 | .42 | .52 | .46 | .52 | .66 | .58 | .58 | .74 | .65 | .76 |
| Chemo-end | 43 | .44 | .59 | .50 | .36 | .49 | .42 | .47 | .63 | .54 | .48 | .66 | .56 | .84 |
| Rad-start | 38 | .35 | .47 | .40 | .30 | .40 | .34 | .36 | .47 | .41 | .40 | .53 | .46 | .64 |
| Rad-end | 35 | .48 | .63 | .54 | .30 | .39 | .34 | .50 | .66 | .57 | .50 | .66 | .57 | .84 |
| Mastectomy | 68 | .58 | .71 | .64 | .52 | .62 | .57 | .62 | .76 | .68 | .62 | .76 | .68 | .77 |
| Lumpectomy | 33 | .49 | .71 | .58 | .43 | .76 | .46 | .46 | .79 | .58 | .46 | .79 | .62 | .91 |
| Reconstruction | 43 | .38 | .57 | .46 | .29 | .44 | .35 | .41 | .63 | .50 | .43 | .65 | .52 | .86 |

Table 2: Event-level five-fold cross-validation performance of models and baselines on training data.

learned in Section 5.3. The first baseline, Baseline1, is to pick the date with the highest classifier’s prediction confidence. The second baseline, Baseline2, is along the same lines as the Combined Classifier used in (McClosky and Manning, 2012). For example, if the candidate date is “6/2009” and we have retrieved two TEs that are resolved to “6/2009” and “4/2008”, then $P(\text{“6/2009”}) = P_{relation}(\text{overlap}|\text{“6/2009”}) \times P_{relation}(\text{no-overlap}|\text{“4/2008”})$.

To set an upper bound on performance given our TE retrieval system, we calculate the oracle score by considering an extraction as correct if the gold date is one of the retrieved candidate dates. The oracle score can differ from a perfect score since we can only use candidate temporal expressions if (a)the relation is known and (b)mentions of the event are retrievable, (c)the TE and event keyword appear in the same sentence, and (d)our temporal tagger is able to recognize and resolve it correctly.

6.2.3 Results

We present the performance of our models, baselines and the oracle in Table 2. Both the Date Classifier and Integrated model significantly outperform the baselines ($p < 0.0001$, McNemar’s test, 2-tailed). This shows the value of our approach to leveraging redundancy of event date mentions. Incorporating time constraints further improves the F1 of the Date Classifier by 3%. The Integrated model achieves 89.7% of the oracle result.

| Model | P | R | F1 |
|------------------|-------------|-------------|-------------|
| Baseline1 | 46.1 | 63.7 | 53.5 |
| Baseline2 | 39.3 | 54.4 | 45.6 |
| Date Classifier | 49.6 | 67.7 | 57.3 |
| Integrated Model | 51.0 | 69.3 | 58.8 |
| Oracle | 77.3 | 77.3 | 77.3 |

Table 3: Performance of systems on the test set.

Table 3 shows the performance of our systems and baselines on individual event types. The Joint

Model derives most of its improvement from performance related to the Chemotherapy/Radiation-start date. This is mainly because Chemotherapy and Radiation last for a period of time and there are more event-related discussions containing the event keyword. None of our systems improves on cancer Metastasis and Recurrence. This is likely due to the sparsity of these events.

7 Conclusion

We presented a novel event date extraction task that requires extraction and resolution of non-standard TEs, namely personal illness event dates, from the posting histories of online community participants. We constructed an evaluation corpus and designed a temporal tagger for non-standard TEs in social media. Using a much stricter standard correctness measure than in previous work, our method achieves promising results that are significantly better than two types of baseline. By creating an analogous keyword set, our event date extraction method could be easily adapted to other datasets.

8 Acknowledgments

We want to thank Dong Nguyen and Yi-chia Wang, who helped provide the data for this project. The research reported here was supported by National Science Foundation grant IIS-0968485.

References

- Javier Artiles, Qi Li, Taylor Cassidy, Suzanne Tamang, and Heng Ji. 2011. CUNY BLENDER TACKBP2011 Temporal Slot Filling System Description. In *Proceedings of Text Analysis Conference (TAC)*.
- Nathanael Chambers, Shan Wang, and Dan Jurafsky. 2007. Classifying temporal relations between events. In *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics*.
- Nathanael Chambers. 2012. Labeling documents with timestamps: Learning from their time expressions. *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics*.
- Angel X. Chang and Christopher D. Manning. 2012. SUTIME: A library for recognizing and normalizing time expressions. In *8th International Conference on Language Resources and Evaluation (LREC 2012)*.
- De Choudhury, M., Counts, S., and Horvitz, E. 2013. Major Life Changes and Behavioral Markers in Social Media: Case of Childbirth. In *Proc. CSCW 2013*.
- Guillermo Garrido, Anselmo Penas, Bernardo Cabaleiro, and Alvaro Rodrigo. 2012. Temporally Anchored Relation Extraction. In *Proceedings of the 50th annual meeting of the association for computational linguistics*.
- Heng Ji, Ralph Grishman, and Hoa Trang Dang. 2011. Overview of the TAC 2011 Knowledge Base Population track. In *Proceedings of Text Analysis Conference (TAC)*.
- Hyuckchul Jung, James Allen, Nate Blaylock, Will de Beaumont, Lucian Galescu, and Mary Swift. 2011. Building timelines from narrative clinical records: initial results based-on deep natural language understanding. In *Proceedings of BioNLP 2011*.
- Oleksandr Kolomiyets, Steven Bethard and Marie-Francine Moens. 2011. Model-Portability Experiments for Textual Temporal Analysis. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*.
- Oleksandr Kolomiyets, Steven Bethard, and Marie-Francine Moens. 2012. Extracting narrative timelines as temporal dependency structures. In *Proceedings of the 50th annual meeting of the Association for Computational Linguistics*.
- Xiao Ling and Daniel S Weld. 2010. Temporal information extraction. *Proceedings of the Twenty Fifth National Conference on Artificial Intelligence*.
- David McClosky and Christopher D. Manning. 2012. Learning Constraints for Consistent Timeline Extraction. *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP2012)*.
- Mike Mintz, Steven Bills, Rion Snow, and Dan Jurafsky. 2009. Distant supervision for relation extraction without labeled data. In *Proceedings of the 47th annual meeting of the Association for Computational Linguistics*.
- Heekyong Park and Jinwook Choi. 2012. V-model: a new innovative model to chronologically visualize narrative clinical texts. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems. ACM*.
- Catherine Plaisant, Brett Milash, Anne Rose, Seth Widoff, and Ben Shneiderman. 1996. LifeLines: visualizing personal histories. In *Proceedings of the SIGCHI conference on Human factors in computing systems*.
- James Pustejovsky, Jos M. Castao, Robert Ingria, Roser Sauri, Robert J. Gaizauskas, Andrea Setzer, Graham Katz, and Dragomir R. Radev. 2003. TimeML: Robust specification of event and temporal expressions in text. *TimeML: Robust specification of event and temporal expressions in text. In New Directions in Question Answering'03*.
- James Pustejovsky, Patrick Hanks, Roser Sauri, Andrew See, Robert Gaizauskas, Andrea Setzer, Dragomir Radev. 2003. The Timebank corpus. In *Corpus Linguistics*.
- Preethi Raghavan, Eric Fosler-Lussier, and Albert M. Lai. 2012. Learning to Temporally Order Medical Events in Clinical Text. In *Proceedings of the 50th annual meeting of the Association for computational Linguistics*.
- Jannik Strotgen and Michael Gertz. 2010. HeidelTime: High Quality Rule-Based Extraction and Normalization of Temporal Expressions. In *SemEval '10*.
- Jannik Strotgen and Michael Gertz. 2012. Temporal Tagging on Different Domains: Challenges, Strategies, and Gold Standards. In *LREC2012*.
- Partha Pratim Talukdar, Derry Wijaya, and Tom Mitchell. 2012. Coupled temporal scoping of relational facts. In *Proceedings of the fifth ACM international conference on Web search and data mining. ACM*.
- Naushad UzZaman and James F. Allen. 2010. TRIPS and TRIOS system for TempEval-2: Extracting temporal information from text. In *Proceedings of the 5th International Workshop on Semantic Evaluation*.
- Yafang Wang, Bing Yang, Lizhen Qu, Marc Spaniol, and Gerhard Weikum. 2011. Harvesting facts from textual web sources by constrained label propagation. In *Proceedings of the 20th ACM International Conference on Information and Knowledge Management*.

- Miaomiao Wen, Hyeju Jang, and Carolyn Rosé. 2013. Coding Manual for Illness Event Date Extraction. *Carnegie Mellon University, School of Computer Science, Language Technology Institute*.
- K.-Y. Wen, F. McTavish, G. Kreps, M. Wise, and D. Gustafson. 2012. From diagnosis to death: A case study of coping with breast cancer as seen through online discussion group messages. *Journal of Computer-Mediated Communication*, 16:331-361.
- Katsumasa Yoshikawa, Sebastian Riedel, Masayuki Asahara, and Yuji Matsumoto. 2009. Jointly identifying temporal relations with markov logic. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*.
- Li Zhou and George Hripcsak. 2007. Temporal reasoning with medical data—a review with emphasis on medical natural language processing. *Journal of biomedical informatics* 40.2 (2007): 183.