Temporal reasoning over clinical text: the state of the art

Weiyi Sun,[1] Anna Rumshisky,[2] Ozlem Uzuner[3]

[1]Department of Informatics, University at Albany, SUNY, Albany, New York, USA
[2]Department of Computer Science, University of Massachusetts, Lowell, Massachusetts, USA
[3]Department of Information Studies, University at Albany, SUNY, Albany, New York, USA

Correspondence to
Weiyi Sun,
Department of Informatics, University at Albany, SUNY, 1400 Washington Ave, Draper 114B, Albany, NY 12222, USA;
wsun2@albany.edu, viviennesun7@gmail.com

Received 26 February 2013
Revised 17 April 2013
Accepted 20 April 2013
Published Online First 15 May 2013

Editor's choice
Scan to access more free content

To cite: Sun W, Rumshisky A, Uzuner O. *J Am Med Inform Assoc* 2013;**20**:814–819.

ABSTRACT

Objectives To provide an overview of the problem of temporal reasoning over clinical text and to summarize the state of the art in clinical natural language processing for this task.

Target audience This overview targets medical informatics researchers who are unfamiliar with the problems and applications of temporal reasoning over clinical text.

Scope We review the major applications of text-based temporal reasoning, describe the challenges for software systems handling temporal information in clinical text, and give an overview of the state of the art. Finally, we present some perspectives on future research directions that emerged during the recent community-wide challenge on text-based temporal reasoning in the clinical domain.

INTRODUCTION

This paper presents a high-level introduction to text-based temporal reasoning in clinical informatics, and provides an overview of the state of the art in clinical natural language processing (NLP) on this topic. Our aim is to familiarize clinical informatics researchers with the applications and challenges of text-based temporal reasoning in the clinical domain.

Temporal reasoning is an umbrella term that can refer to any time-related data processing task. It has been an active research area for about half a century in the field of artificial intelligence and has applications in many related research areas, including NLP,[1] data mining,[2] robotics,[3] and database design and query.[4] We limit the scope of this introduction to temporal reasoning in NLP, with a focus on applications in clinical informatics. For readers with limited experience in NLP, an NLP tutorial can be found in Nadkarni *et al*.[5]

In this paper, we define NLP-based temporal reasoning as a combination of: (1) temporal representation formalism; (2) extraction of temporal information from natural language text; and (3) temporal inference over the extracted information. Temporal representation formalism involves defining a machine-readable representation of the temporal dimension, including formalizing the notion of time, defining the temporal events, and specifying the possible temporal relations. Temporal information extraction refers to the automated mark-up and normalization of temporal information from natural language texts based on a formalized temporal representation. Temporal inference refers to the logical deductions performed on the extracted temporal information to enhance natural language understanding. Natural language text always carries within it a temporal interpretation. Whenever we describe a new occurrence, a change, or a progression of events, we introduce, explicitly or implicitly, a temporal dimension. NLP systems that aim to capture the information contained in natural language text must therefore be capable of accurately extracting, representing, and reasoning about temporality. Although the human brain is capable of processing temporal information very efficiently, temporal reasoning remains a difficult task for NLP systems. Identifying temporal relations between events in text is challenging due to the diversity of linguistic mechanisms for expressing temporal information, and due to the complex interplay of explicit and implicit inference required to understand such information.

A temporal dimension is essential for the interpretation of clinical narratives. The progression of illnesses and the events in a hospital course are typically recorded chronologically, and many clinically relevant events are only significant in a particular temporal context. The order in which the symptoms develop, the timing of different treatments, and the duration and frequencies of medications are all meaningful within a particular timeline.

In order to put this work in context, the reader should be aware of a body of research on temporal reasoning in the clinical domain that has used time-stamped structured data. Some of the applications for this work included assisted clinical decision making and improved patient care through database design.[6 7] Much of this work has been concerned with reasoning and inference based on absolute time points associated with explicitly coded time-stamped events, such as laboratory tests, doctor's visits, administered procedures, and records from medication logs.[8–15]

In this paper, we will focus on the temporal information represented in the *unstructured* narratives of clinical notes. We provide a high level overview of the problems, methods, and applications of temporal reasoning over clinical text. The rest of the paper is organized as follows. In the 'Applications of temporal reasoning in the clinical domain' section, we present clinical question answering, information extraction, information retrieval, and summarization as the most prominent applications of text-based temporal reasoning. In the 'Challenges of text-based temporal reasoning' section, we discuss the challenges and pitfalls of automating text-based temporal reasoning. Finally, in the 'Current state of the art in text-based temporal reasoning' section, we present the state of the art and summarize the research directions that have emerged as a result of a recent community-wide challenge on temporal information extraction over clinical text run under the aegis of Informatics for Integrating Biology and the Bedside (i2b2) project.[16]

APPLICATIONS OF TEMPORAL REASONING IN THE CLINICAL DOMAIN

Temporal reasoning techniques can empower a wide range of NLP application, and are readily transferrable between NLP applications. Information extraction, question answering, information retrieval, and summarization are the four most prominent tasks that benefit from temporal reasoning. Information extraction refers to the task of identifying key information in free text. Question answering aims to respond to natural language questions of humans with just the right information. Information retrieval searches for specific information in a free-text corpus. Finally, summarization condenses the contents of natural language text.

Medical information extraction

Electronic medical records (EMRs) contain an overwhelming amount of information. This information is often presented in expert shorthand and is stylistically quite distinct from the general domain English written text. As a result, NLP technologies developed in other domains do not readily transfer to the clinical domain.

Information extraction technologies designed on clinical narratives find and highlight the key information presented in unstructured medical records. Given the temporal organization of clinical information, the information extracted by information extraction systems also needs to be temporally organized. Such an organization would then allow us, for example, to determine whether an extracted drug is historically used, currently administered, or prescribed for future use; whether a patient is on a certain medication at a specific time; or when a particular intervention or test has been recommended.

Recently, there has been increased interest in incorporating temporal reasoning into information extraction systems. Denny *et al*[17] developed a system that extracts the time and status of colorectal cancer screening tests in EMRs. This system identifies clinical concepts in EMRs, finds and normalizes date information, and uses heuristics to assign the dates to clinical concepts. It can also support the process of determining a patient's colonoscopy status and examination time and save some of the time and effort required to manually establish this information. Similar ideas have been applied to determining patients' medication usage status. Liu *et al*[18] developed a framework to extract patient drug exposure histories from EMRs, that is, to determine if a patient was on a certain drug at a given time. In particular, they studied patients' warfarin exposure at the time of hospital admission. Their system achieved 87% precision and 79% recall. Irvine *et al*[19] developed a system that extracts and interprets temporal expressions from emergency department triage notes.

Clinical question answering

Question answering aims to answer natural language queries, such as:

- What medication was the patient on before the surgery?
- How often did the patient experience headache before that treatment?
- What symptoms did the patient experience after taking aspirin for 3 days?

Many clinically meaningful questions have a temporal dimension. In order to answer such questions, the system needs to: (1) identify the clinically significant events in the narratives and in the questions; and (2) perform temporal reasoning on these events, that is, determine the temporal relations between pairs of events from the narrative reports and the questions. Zhou *et al*[20] designed TimeText, which extracts time-related information and medical events from EMR narratives, and outputs a temporal relation constraint graph of the events. The authors applied TimeText to question answering and achieved an accuracy of 83.7% in answering 147 questions about 20 discharge summaries.[21] Tao *et al*[22] proposed a framework for applying semantic web technology to temporal reasoning question answering in clinical text. Clark *et al*[23] adopted this framework for analyzing reports of complaints about medical devices. In particular, the authors focused on adverse events related to late stent thrombosis. The CNTRO framework correctly answered 89.04% questions in the test set. Li *et al*[24] presented a temporal information extraction method, and demonstrated its incorporation into a question answering system to answer time-related questions. They obtained promising results on questions frequently posed by intensive care unit doctors.

Medical narrative summarization

Temporal reasoning is also used to create structured or visualized summaries of EMRs. Jung *et al*[25] developed a prototype system for extracting and graphically visualizing clinical timelines in EMRs. Their system relies on a context-free grammar, with hand-built core lexicon and semantics ontology. It is supported by off-the-shelf NLP components to generate semantic parses of sentences in the clinical narratives. The system then extracts events, temporal expressions, and temporal relations from the parsed sentences. Finally, it creates a visualization of the clinical timeline using the SIMILE tool (http://www.simile-widgets.org/timeline/).

Temporal relation-related information retrieval

Traditional information retrieval techniques based on keyword indexing usually do not support searches by time or by temporal relation. In the general domain, temporal information was found to be helpful in retrieval applications, such as temporal ranking of ad hoc search results, temporal clustering of retrieved documents, and exploratory search.[26] In the clinical domain where temporal information is more abundant, precise, and critical, temporal reasoning in relation-related information retrieval becomes more important and rewarding. For example, Crowley *et al*[27] presented a cancer tissue information retrieval system that supports time-related queries. However, this system relies on structured data, such as time stamps, to derive the temporal relations between documents. Future research may consider exploring unstructured data to expose the temporal relations between entities.

CHALLENGES OF TEXT-BASED TEMPORAL REASONING

Temporal reasoning in clinical narratives is a non-trivial task. Some of the causes of the complexity are described below.

- *Problems in temporal representation*: The philosophical concept of time has been debated ever since the ancient Greeks tackled the subject.[28] A number of distinct temporal representations have been developed in philosophy, linguistics, and computer science. We refer interested readers to Augusto[29] and Combi and Shahar[9] for details on temporal notions and their formalisms in computing. In short, Augusto[6] pointed out that to design a time ontology for computing, one needs to consider the many options for modeling the structure of time (eg, linear, branching, or circular); to determine whether time has start/terminal points; to choose between continuous or discrete representations of time; and to decide whether to use time points or intervals as temporal references. Each of these choices has merits and limitations. Finding a design that can describe all possible temporal information, supports temporal inference, and computes efficiently is not easy. In addition to the time ontology definition, a temporal representation scheme also needs to specify its concept primitives, that is, the primitives on which temporal reasoning will be performed. Some well-known choices include state,[30] events,[31] and changes.[32]

- *Complexity of temporal representation in natural language:* Temporal information is expressed in language by a variety of means, and often cannot be understood without the integration of multiple levels of linguistic processing, including, but not limited to, grammar, semantics, discourse, and inference. The information is often implicit and requires general conceptual knowledge. Below are some of the problematic aspects of the way temporal information is represented in language.
  - Underspecified temporal relations. Temporal relations in natural language are 'implicit and vague'[33]:
    - For example, we can express the same temporal relation by saying 'John fell from the stairs. He was sent to the hospital,' or 'John was sent to the hospital. He fell from the stairs.' As human readers, we understand the causal relation between 'falling from the stairs' and 'being sent to the hospital,' and as a result, we can reconstruct the temporal order of events even when the narrative order is reversed. However, this is a non-trivial task for machines and is especially complicated in clinical narratives where the deciphering of temporal relations can require medical knowledge.
    - The statement 'the patient complained about chest pain and fever' informs us that the 'chest pain' and 'fever' started before the current time, and both symptoms occurred roughly at the same time. But, we cannot tell whether the chest pain started before or after the fever, or whether both symptoms lasted until the current moment. Unlike humans who can make sense of such vague and fuzzy information, computer systems need to be specially designed to cope with temporal vagueness.
  - Vagueness of tense and aspect. Temporal information can be expressed by means of event tense (whether an event locates in the past, present, or future) and aspect (whether an event is completed or in progress). Even though English has grammatical tense markers, one cannot fully rely on the tense to locate an event on the temporal axis (or tree or circle). For example, the sentence 'Deb leaves tonight' is in the present tense but it indicates an event in the future. Vendler's work[34] categorizes events into several groups based on their lexical aspects (including states and processes, with the latter further divided into accomplishments and achievements). Each group of events has its own temporal properties. For example, a state (such as 'know') is expected to be continuous in every subinterval of the state, while an achievement (such as 'win') is usually instantaneous.
  - Relative times. Temporal expressions in natural language are frequently written in reference to other time points, for example, 'last Friday' (relative to the date when this phrase was written) and 'the day before surgery' (relative to the date of the surgery). A temporal reasoning system needs to determine the exact values (eg, the exact calendar dates) of relative temporal expressions. We refer the reader to Alonso *et al*[35] for more details.
  - Implicit event durations. Even when the information about event duration is not explicitly expressed, it is often a part of the general knowledge required for reasoning. For example, in the sentence, 'The patient saw a nutritionist,' we know that this event lasted more than a second, but less than 10 h. Events have typical durations that may be modified in context. Such vague durations are in part the reason that representing events in text as well-defined intervals often seems counterintuitive. We refer the reader to Pan *et al*[36 37] for details.
  - Temporal aggregates. Natural language has many expressions for collections of events, such as 'every Tuesday,' 'five business days,' 'the next three visits,' etc. In clinical notes, similar expressions are used for frequencies of medication-related events, and often present a challenge to temporal ordering of events.
- *Challenges in handling clinical narratives:* Performing natural language analysis on clinical narratives is difficult because clinical notes are usually ungrammatical, full of shorthand, abbreviations, and misspellings, and copy-and-paste text, as has been pointed out by Meystre *et al*.[38] As a result, some of the prerequisites of temporal reasoning, such as coreference resolution, parsing, or acronym disambiguation, are unsolved NLP problems themselves.[39]

CURRENT STATE OF THE ART IN TEXT-BASED TEMPORAL REASONING

Over the past decade, temporal reasoning research using both unstructured and structured data gained significant momentum. For structured data, Augusto[6] presented an extensive review of the literature on temporal database design that supports clinical decision making in diagnosis, prognosis, and treatment. Combi *et al*[9] provided an overview of temporal reasoning and temporal data maintenance in medical informatics. For unstructured data, such as clinical narratives, several reviews summarize and analyze temporal reasoning research literature.[38 40]

Temporal representation

A large body of literature has been published on temporal representation in artificial intelligence, including some extensive review pieces.[6 9 29 41 42] In this section, we introduce some of the better-known NLP temporal representation schemes.

The design of natural language temporal representation usually involves three steps: (1) defining an ontology of temporal expressions; (2) specifying concept primitives (eg, events, actions, statuses, or processes); and (3) defining temporal relations between the temporal expressions and concept primitives.

One of the early standards for representing temporal expression, developed for the Message Understanding Conference-7 (MUC-7), and subsequently adapted to newswire text,[43] was the TIMEX scheme that specified only the type of the temporal expression, which could be either TIME or DATE. Its successor, TIMEX2,[44 45] allowed the annotators to specify: (1) the normalized value of the temporal expressions; (2) their modifiers (eg, approximately, less than, more than); and (3) whether a temporal expression is a frequency, as well as the periodicity of the frequency. TIMEX2 could also describe the temporal granularity of expressions and indicate whether an expression is specific or generic (eg, 'last Monday,' vs 'a Monday'). Using TIMEX and TIMEX2 as a starting point, Pustejovsky *et al*[46] developed TIMEX3, which formed part of the TimeML specification language for temporal information. TIMEX3 defined a temporal function mechanism to represent relative times.

One of the most influential works on the representation of concept primitives and temporal relations was Allen's interval-based temporal logic.[47 48] Allen's theory assumes a linear model of time with intervals as temporal primitives, and events as concept primitives. Allen and Ferguson[48] defined six invertible relations (before/after, meets/met_by, overlaps/overlapped_by, starts/started_by, during/contains, finishes/finished_by) and one symmetric relation (equal) between temporal intervals, as well as transitive axioms to hold between these relations. This representation was used extensively in the TimeML specification,[46] which has been used to annotate a number of general domain corpora, including the TimeBank corpus[49] containing 183 news articles, the AQUAINT corpus (http://timeml.org/site/timebank/timebank.html) containing 73 news report documents, and the TempEval challenge corpora.

In the clinical domain, Galescu and Blaylock[50] annotated 40 discharge summaries selected from the i2b2 Challenge data[39] using an adapted version of the TimeML guidelines. Savova *et al*[51] described their intention to adopt TimeML to annotate clinical narratives. The annotation guidelines of the 2012 i2b2 NLP challenge in temporal reasoning were also adapted from TimeML[16 52]

Several temporal representation schemes tailored for clinical narratives have also been developed. Zhou *et al*[53] identified the main categories of temporal expressions in 100 discharge summaries and developed a clinical temporal constraint model based on simple temporal problems. Lai *et al*[54] extended this temporal constraint model to represent vague temporal relations. Tao *et al*[55] developed an OWL-based time ontology for clinical text, which supports semantic web techniques.

Temporal reasoning methods

Temporal reasoning operates over the specified temporal representation, that is, over temporal expressions and concept primitives, and involves the assignment of temporal relations. We review the current state of the art in each of these subtasks. Others have provided a thorough review of temporal reasoning methods but followed a different framework.[1 6 9 40]

Temporal expression extraction and normalization

Temporal expression recognition involves: (1) extraction, that is, determining the text span of an expression; and (2) normalization, that is, interpreting the meaning of the expression. For example, in the sentence 'the patient complained about a headache during the past 2 days,' the temporal expression is 'the past 2 days' and under the ISO 8601 standard, its normalized form is 'P2D' indicating a period of 2 days. Both rule-based and supervised machine learning methods have been explored to solve this task. Rule-based systems follow predefined rules to identify and interpret temporal expressions. Supervised machine learning methods derive latent patterns from labeled training data and apply them to unseen data.

The creation and publication of temporally annotated corpora, such as the TimeBank corpus, tremendously advanced research in temporal expression recognition, spurring several shared task challenges. MUC-7[56] addressed temporal expression recognition as part of the named entity task, although the normalization of temporal expression values was not targeted. ACE TERN[57] temporal expression tasks required the systems to create TIMEX2-style[44] temporal expressions which included normalization. TempEval-2[58] and TempEval-3,[59] two of the three temporal analysis shared tasks under SemEval,[58] used TIMEX3-style annotation from TimeML. The best performing systems in TempEval-2 included: (1) the HeidelTime[60] tool, a temporal expression tagger that uses four sets of hand-crafted rules to identify and classify temporal expressions; and (2) the TRIPS and TRIOS system,[61] which used a conditional random field (CRF)[62] machine learning classifier to detect temporal expressions, and a rule-based approach for temporal value normalization. Other temporal expression taggers that produce TIMEX3 annotations include: SUTime,[63] a rule-based temporal expression tagger in the Stanford CoreNLP pipeline which identifies and normalizes temporal expressions; and GUTime,[64] the temporal expression tagger in the TARSQI toolkit.[65] Results of the TempEval-2 competition suggested that temporal expression normalization was best performed by rule-based systems. Rule-based and machine learning, as well as hybrid systems, performed comparably in temporal expression recognition.

Identification of concept primitives

In the shared tasks based on the TimeML annotation guidelines, concept primitives are EVENTs, typically verbs or nominal forms of actions. Results of the shared tasks suggested that such primitives were best identified using statistical machine learning techniques. Thus, in TempEval-2, the best performing systems in EVENT detection included: TipSem,[66] which was based on a CRF,[62] a statistical modeling method for sequential data labeling; the TRIPS and TRIOS system,[61] which used a hybrid of rule-based and Markov logic network,[67] a statistical inference classifier; and Edinburgh-LTG,[68] which used a combination of rule-based and maximum entropy (MaxEnt) methods.[69]

Verb or action-based events are not the only choice for concept primitives. Bramsen *et al*[70] proposed ordering temporal segments instead of action-based events. The authors defined a temporal segment as 'a fragment of text that does not exhibit abrupt changes in temporal focus,' and used the BoosTexter classifier[71] to identify the boundaries of temporal segments. In the clinical domain, several studies chose to target medical concepts instead of action-based events as the temporal ordering primitives. Zhou *et al*[20] designed a temporal reasoning system over the medical language entities (eg, problems, tests, and treatments) extracted by MedLEE.[72] In the CNTRO ontology and its applications,[22 55] the temporal ordering primitives were defined to be clinical events, that is, any 'occurrence, state, perception, procedure, symptom or situation' in the clinical timeline. The 2012 i2b2 NLP challenge on temporal relations defined their temporal primitives to be clinical concepts (problems, tests, treatments, and clinical departments), occurrences, and evidential events that reveal the source of clinical information.[16]

Temporal relation classification

The two TempEval competitions revealed a trend that favors statistical machine learning methods compared to the rule-based methods for the task of temporal relation classification. In the TempEval-1 challenge,[73] statistical machine learning systems and hybrid method systems achieved similar scores, with the rule-based system lagging behind. The participants were required to classify local pairwise temporal relations between (1) events and temporal expressions in the same sentence, (2) events and the document creation time, and (3) main events in adjacent sentences. The methods chosen by the participants included support vector machine (SVM) classifiers,[74] hidden Markov model SVMs (HMM-SVM),[75] a hybrid method of hand-crafted rules and SVMs,[76] and rule-based systems.[77] TempEval-2 defined similar tasks in temporal relation classification. The chosen methods included MaxEnt,[78] Markov logic networks,[61 79] and CRFs.[66 80] Outside of the shared task challenges, several other studies also adopted statistical machine learning methods to classify temporal relations. SVMs[81 82] and MaxEnt[65 83] were among the popular choices.

Both TempEval-1 and TempEval-2 restricted the set of candidate event/time expression pairs in each subtask, making the choice of candidate pairs straightforward. Such restrictions are unrealistic in real-world NLP tasks. A useful end-to-end temporal reasoning system needs to determine which candidate event/time expression pairs can be usefully ordered. Due to the limited scope of the first shared tasks, the problem of selecting temporally related pairs has been given less attention until the 2012 i2b2 NLP challenge.[16] The following pair selection strategies are mentioned in the existing literature: (1) restricting the type of pairs to, for example, only intra-sentential pairs[82 84]; and (2) labeling all pairs of concept primitives and temporal expressions and then eliminating conflicted ones based on confidence scores.[65 81] The 2012 i2b2 Challenge required the participating systems to assign temporal relations to clinical events and temporal expressions without any pre-defined pair selection criteria. Pair selection strategy substantially affected the system evaluation scores. Strategies included using supervised learning methods to identify the linked event pairs, event and temporal expression pairs, and temporal expression pairs,[85 86] as well as using heuristics to select candidate pairs.[87–89]

CONCLUSION

Temporal reasoning is an active research area in artificial intelligence. Recent research suggests promising future applications of temporal reasoning in medical NLP. Nonetheless, temporal reasoning in natural language presents multiple challenges. These arise from the implicit nature of temporal representation in human language, often characterized by a considerable degree of temporal underspecification. Over the past decade, the shared task NLP challenges, along with the development of temporal representation standards, have led to significant advances in text-based temporal reasoning as it relates to representation, automated extraction, and inference over temporal information. With the i2b2 temporal corpus of medical narratives becoming available, we predict an even more fruitful future for temporal reasoning research in the medical NLP domain.

Contributors WS is the primary author. AR and OU offered insights and guidance during the preparation of the manuscript, and provided substantial edits to the manuscript.

Funding This project was supported by Informatics for Integrating Biology and the Bedside (i2b2) award No. 2U54LM008748 from the National Institutes of Health (NIH)/National Q5 Library of Medicine (NLM), by the National Heart, Lung, and Blood Institute (NHLBI), and by award No. 1R13LM01141101 from the NIH NLM. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NLM, NHLBI, or NIH.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Mani I, Pustejovsky J, Gaizauskas R. *The language of time: a reader*. Oxford University Press, 2005.
2 Bettini C, Jajodia S, Wang S. *Time granularities in databases, data mining, and temporal reasoning*. Springer, 2000.
3 Schulz R, Wyeth G, Wiles J. Are we there yet? Grounding temporal concepts in shared journeys. *Autonomous Mental Dev IEEE Trans* 2011;3:163–75.
4 Tansel AU. *Temporal databases*. Wiley Online Library, 2008.
5 Nadkarni PM, Ohno-Machado L, Chapman WW. Natural language processing: an introduction. *J Am Med Inform Assoc* 2011;18:544–51.
6 Augusto JC. Temporal reasoning for decision support in medicine. *Artif Intell Med* 2005;33:1–24.
7 Stacey M, McGregor C. Temporal abstraction in intelligent clinical data analysis: a survey. *Artif Intell Med* 2007;39:1–24.
8 Campbell KE, Das A, Musen MA. A logical foundation for representation of clinical data. *J Am Med Inform Assoc* 1994;1:218–32.
9 Combi C, Shahar Y. Temporal reasoning and temporal data maintenance in medicine: issues and challenges. *Comput Biol Med* 1997;27:353–68.
10 Keravnou ET. Temporal diagnostic reasoning based on time-objects. *Artif Intell Med* 1996;8:235–65.
11 Das AK, Musen MA. A foundational model of time for heterogeneous clinical databases. *Proceedings of the AMIA Annual Fall Symposium*. American Medical Informatics Association, 1997:106.
12 Nigrin DJ, Kohane IS. Temporal expressiveness in querying a time-stamp—based clinical database. *J Am Med Inform Assoc* 2000;7:152–63.
13 Kahn MG, Fagan LM, Tu S, *et al*. Extensions to the time-oriented database model to support temporal reasoning in medical expert systems. *Methods Inf Med* 1991;30:4.
14 Shahar Y. A framework for knowledge-based temporal abstraction. *Artif Intell* 1997;90:79–133.
15 Kohane IS, Haimowitz IJ. Hypothesis-driven data abstraction with trend templates. *Proceedings of the Annual Symposium on Computer Application in Medical Care*. American Medical Informatics Association, 1993:444.
16 Sun W, Rumshisky A, Uzuner O. Evaluating temporal relations in clinical text: 2012 i2b2 challenge. *J Am Med Inform Assoc* 2013;20:806−13.
17 Denny JC, Peterson JF, Choma NN, *et al*. Extracting timing and status descriptors for colonoscopy testing from electronic medical records. *J Am Med Inform Assoc* 2010;17:383–8.
18 Liu M, Jiang M, Kawai VK, *et al*. Modeling drug exposure data in electronic medical records: an application to warfarin. *AMIA Annual Symposium Proceedings*. Vol. 2011. American Medical Informatics Association, 2011:815.
19 Irvine AK, Haas SW, Sullivan T. Tn-ties: a system for extracting temporal information from emergency department triage notes. *AMIA Annual Symposium proceedings*. Vol. 2008. American Medical Informatics Association, 2008:328.
20 Zhou L, Friedman C, Parsons S, *et al*. System architecture for temporal information extraction, representation and reasoning in clinical narrative reports. *AMIA Annual Symposium Proceedings*. Vol. 2005. American Medical Informatics Association, 2005:869.
21 Zhou L, Parsons S, Hripcsak G. The evaluation of a temporal reasoning system in processing clinical discharge summaries. *J Am Med Inform Assoc* 2008;15:99–106.
22 Tao C, Solbrig H, Sharma D, *et al*. Time-oriented question answering from clinical narratives using semantic-web techniques. *The Semantic Web–ISWC 2010*. Springer Berlin Heidelberg, 2010:241–56.
23 Clark KK, Sharma DK, Chute CG, *et al*. Application of a temporal reasoning framework tool in analysis of medical device adverse events. *AMIA Annual Symposium Proceedings*. Vol. 2011. American Medical Informatics Association, 2011:1366.
24 Li M, Patrick J. Extracting temporal information from electronic patient records. *AMIA Annual Symposium Proceedings*. vol. 2012. American Medical Informatics Association, 2012:542.
25 Jung H, Allen J, Blaylock N, *et al*. Building timelines from narrative clinical records: initial results based-on deep natural language understanding. *Proceedings of BioNLP 2011 Workshop*. Association for Computational Linguistics, 2011: 146–54.
26 Alonso O, Gertz M, Baeza-Yates R. On the value of temporal information in information retrieval. *ACM SIGIR Forum*. Vol. 41. ACM, 2007:35–41.
27 Crowley RS, Castine M, Mitchell K, *et al*. caTIES: a grid based system for coding and retrieval of surgical pathology reports and tissue specimens in support of translational research. *J Am Med Inform Assoc* 2010;17:253–64.
28 Rau C. Theories of time in ancient philosophy. *Philos Rev* 1953:514–25.
29 Augusto JC. The logical approach to temporal reasoning. *Artif Intell Rev* 2001;16:301–33.
30 McCarthy J. *Situations, actions, and causal laws*. DTIC Document, 1963.
31 Kowalski R, Sergot M. A logic-based calculus of events. *New Gen Comput* 1986;4:67–95.
32 Gooday J, Galton A. The transition calculus: a high-level formalism for reasoning about action and change. *J Exp Theor Artif Intell* 1997;9:51–66.
33 Allen JF. An interval-based representation of temporal knowledge. *Proceedings of the 7th International Joint Conference on Artificial Intelligence*. Morgan Kaufmann, 1981:221–6.
34 Vendler Z. Verbs and times. *Philos Rev* 1957:143–60.
35 Alonso O, Strötgen J, Baeza-Yates R, *et al*. Temporal information retrieval: challenges and opportunities. *TWAW Workshop, WWW*, 2011:1–8.
36 Pan F, Mulkar R, Hobbs JR. Learning event durations from event descriptions. *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, 2006:393–400.
37 Pan F, Mulkar R, Hobbs JR. An Annotated Corpus of Typical Durations of Events. *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC)*, 2006:77–82.
38 Meystre SM, Savova GK, Kipper-Schuler KC, *et al*. Extracting information from textual documents in the electronic health record: a review of recent research. *Yearb Med Inform* 2008;35:128–44.

39 Uzuner O, Bodnari A, Shen S, *et al*. Evaluating the state of the art in coreference resolution for electronic medical records. *J Am Med Inform Assoc* 2012;19:786–91 (abstract).
40 Zhou L, Hripcsak G. Temporal reasoning with medical data–a review with emphasis on medical natural language processing. *J Biomed Inform* 2007;40:183.
41 Vila L. A survey on temporal reasoning in artificial intelligence. *AI Communications* 1994;7:4–28.
42 Chittaro L, Montanari A. Temporal representation and reasoning in artificial intelligence: issues and approaches. *Ann Math Artif Intell* 2000;28:47–106.
43 Setzer A, Gaizauskas R. Annotating events and temporal information in newswire texts. *Proceedings of the Second International Conference On Language Resources And Evaluation (LREC-2000)*, Vol. 31. Athens, Greece, 2000.
44 Ferro L, Mani I, Sundheim B, *et al*. *TIDES* Temporal Annotation Guidelines-Version 1.0.2. The MITRE Corporation, 2001.
45 Ferro L, Gerber L, Mani I, *et al*. *TIDES 2005 standard for the annotation of temporal expressions*. The MITRE Corporation, 2005.
46 Pustejovsky J, Castano J, Ingria R, *et al*. TimeML: robust specification of event and temporal expressions in text. *New Dir Ques Ans* 2003;3:28–34.
47 Allen JF. Towards a general theory of action and time. *Artifl Intell* 1984; 23:123–54.
48 Allen JF, Ferguson G. Actions and events in interval temporal logic. *J Logic Comput* 1994;4:531–79.
49 Pustejovsky J, Hanks P, Sauri R, *et al*. The TIMEBANK Corpus. In: *Proceedings of Corpus Linguistics*, 2003:647–56.
50 Galescu L, Blaylock N. A corpus of clinical narratives annotated with temporal information. *Proceedings of the 2nd ACM SIGHIT symposium on International health informatics*. ACM, 2012:715–20.
51 Savova G, Bethard S, Styler W, *et al*. Towards temporal relation discovery from the clinical narrative. *AMIA Annual Symposium Proceedings*. Vol. 2009. American Medical Informatics Association, 2009:568.
52 Sun W, Rumshisky A, Uzuner O. Annotating temporal information in clinical narratives. Forthcoming.
53 Zhou L, Melton GB, Parsons S, *et al*. A temporal constraint structure for extracting temporal information from clinical narrative. *J Biomed Inform* 2006;39:424–39.
54 Lai AM, Parsons S, Hripcsak G. Fuzzy temporal constraint networks for clinical information. *AMIA Annual Symposium Proceedings*. Vol. 2008. American Medical Informatics Association, 2008:374.
55 Tao C, Wei WQ, Solbrig HR, *et al*. CNTRO: a semantic web ontology for temporal relation inferencing in clinical narratives. *AMIA Annual Symposium Proceedings*. American Medical Informatics Association, 2010:787.
56 Proceedings of the 7th Message Understanding Conference (MUC-7), 1998. http://www.itl.nist.gov/iaui/894.02/related_projects/muc/proceedings/muc_7_proceedings/overview.html (accessed 17 Jan 2013).
57 The ACE 2004 Evaluation Plan: Evaluation of the Recognition of ACE Entities, ACE Relations and ACE Events. http://www.itl.nist.gov/iad/mig/tests/ace/2004/doc/ace04-evalplan-v7.pdf (accessed 17 Jan 2013).
58 Verhagen M, Sauri R, Caselli T, *et al*. SemEval-2010 task 13: TempEval-2. *Proceedings of the 5th International Workshop on Semantic Evaluation*. Association for Computational Linguistics, 2010:57–62.
59 UzZaman N, Llorens H, Allen J, *et al*. TempEval-3: Evaluating Events, Time Expressions, and Temporal Relations. arXiv preprint arXiv:12065333, 2012.
60 Strötgen J, Gertz M. HeidelTime: high quality rule-based extraction and normalization of temporal expressions. *Proceedings of the 5th International Workshop on Semantic Evaluation*. Association for Computational Linguistics, 2010:321–324.
61 UzZaman N, Allen JF. TRIPS and TRIOS system for TempEval-2: extracting temporal information from text. *Proceedings of the 5th International Workshop on Semantic Evaluation*. Association for Computational Linguistics, 2010:276–83.
62 Lafferty J, McCallum A, Pereira FCN. Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. In: *Proceedings of the Eighteenth International Conference on Machine Learning (ICML '01)*. Carla E Brodley and Andrea Pohoreckyj Danyluk, eds. San Francisco, CA: Morgan Kaufmann Publishers Inc, 282–89.
63 Chang AX, Manning CD. SUTIME: a library for recognizing and normalizing time expressions. *Lang Resources Eval* 2012.
64 Mani I, Wilson G. Robust temporal processing of news. *Proceedings of the 38th Annual Meeting on Association for Computational Linguistics*. Association for Computational Linguistics, 2000:9–76.
65 Verhagen M, Pustejovsky J. Temporal processing with the TARSQI toolkit. *22nd International Conference on Computational Linguistics: Demonstration Papers*. Association for Computational Linguistics, 2008:89–192.
66 Llorens H, Saquete E, Navarro B. TIPSem (English and Spanish): evaluating CRFs and Semantic Roles in TempEval-2. *Proceedings of the 5th International Workshop on Semantic Evaluation*, Association for Computational Linguistics, 2010: 84–291.
67 Richardson M, Domingos P. Markov logic networks. *Mach Learn* 2006;62: 107–36.
68 Grover C, Tobin R, Alex B, *et al*. Edinburgh-LTG: TempEval-2 system description. *Proceedings of the 5th International Workshop on Semantic Evaluation*. Association for Computational Linguistics, 2010:333–6.
69 Curran JR, Clark S. Language independent NER using a maximum entropy tagger. *Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003-Volume 4*. Association for Computational Linguistics, 2003:164–7.
70 Bramsen P, Deshpande P, Lee YK, *et al*. Inducing temporal graphs. *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 2006:189–98.
71 Schapire RE, Singer Y. BoosTexter: a boosting-based system for text categorization. *Mach Learn* 2000;39:135–68.
72 Friedman C, Hripcsak G, Shagina L, *et al*. Representing information in patient reports using natural language processing and the extensible markup language. *J Am Med Inform Assoc* 1999;6:76–87.
73 Verhagen M, Gaizauskas R, Schilder F, *et al*. SemEval-2007 task 15: TempEval temporal relation identification. *Proceedings of the 4th International Workshop on Semantic Evaluations*. Association for Computational Linguistics, 2007: 75–80.
74 Bethard S, Martin JH. CU-TMP: temporal relation classification using syntactic and semantic features. *Proceedings of the 4th International Workshop on Semantic Evaluations*. Association for Computational Linguistics, 2007:129–32.
75 Cheng Y, Asahara M, Matsumoto Y. NAIST.japan: temporal relation identification using dependency parsed tree. *Proceedings of the 4th International Workshop on Semantic Evaluations*. Association for Computational Linguistics, 2007:245–8.
76 Min C, Srikanth M, Fowler A. LCC-TE: a hybrid approach to temporal relation identification in news text. *Proceedings of the 4th International Workshop on Semantic Evaluations*. Association for Computational Linguistics, 2007:219–22.
77 Hagège C, Tannier X, *et al*. XRCE-T: XIP temporal module for TempEval campaign. *Proceedings of the 4th International Workshop on Semantic Evaluations*. Association for Computational Linguistics, 2007:492–5.
78 Derczynski L, Gaizauskas R. Usfd2: annotating temporal expressions and tlinks for TempEval-2. *Proceedings of the 5th International Workshop on Semantic Evaluation*. Association for Computational Linguistics, 2010:337–40.
79 Ha EY, Baikadi A, Licata C, *et al*. NCSU: modeling temporal relations with Markov logic and lexical ontology. *Proceedings of the 5th International Workshop on Semantic Evaluation*. Association for Computational Linguistics, 2010:341–4.
80 Kolya AK, Ekbal A, Bandyopadhyay S. JU_CSE_TEMP: a first step towards evaluating events, time expressions and temporal relations. *Proceedings of the 5th International Workshop on Semantic Evaluation*. Association for Computational Linguistics, 2010:15–16.
81 Chambers N, Jurafsky D. Jointly combining implicit constraints improves temporal ordering. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 2008:698–706.
82 Mirroshandel S Abolghasem, Ghassem-Sani G, Khayyamian M. Using syntactic-based kernels for classifying temporal relations. *J Comput Sci Technol* 2011;26:68–80.
83 Mani I, Verhagen M, Wellner B, *et al*. Machine learning of temporal relations. *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, 2006:753–60.
84 Cherry C, Zhu X, Martin J, *et al*. A la Recherche du Temps Perdu—extracting temporal relations from medical text in the 2012 i2b2 NLP Challenge. *J Am Med Inform Assoc* 2013;20:843–8.
85 Chang Y-C, Wu JC-Y, Chen J-M, *et al*. TEMPTing System: a hybrid method of rule and machine learning for temporal relation extraction in patient discharge summaries. *J Biomed Inform* 2013.
86 Grouin C, Grabar N, Hamon T, *et al*. Eventual situations for timeline extraction from clinical reports. *J Am Med Inform Assoc* 2013;20:820–7.
87 D'Souza J, Ng V. Classifying temporal relations in clinical data: a hybrid, knowledge-rich approach. *J Biomed Inform* 2013.
88 Tang B, Wu Y, Jiang M, *et al*. A hybrid system for temporal information extraction from clinical text. *J Am Med Inform Assoc* 2013;20:828–35.
89 Roberts K, Rink B, Harabagiu SM. A flexible framework for recognizing events, temporal expressions, and temporal relations in clinical text. *J Am Med Inform Assoc* 2013;20:867–75.