Extracting Temporal Constraints from Clinical Research Eligibility Criteria Using Conditional Random Fields

Zhihui Luo, PhD[1], Stephen B. Johnson, PhD[1], Albert M. Lai, PhD[2], Chunhua Weng, PhD[1]
[1]Department of Biomedical Informatics, Columbia University in the City of New York, NY;
[2]Department of Biomedical Informatics, The Ohio State University, Columbus, OH

Abstract

Temporal constraints are present in 38% of clinical research eligibility criteria and are crucial for screening patients. However, eligibility criteria are often written as free text, which is not amenable for computer processing. In this paper, we present an ontology-based approach to extracting temporal information from clinical research eligibility criteria. We generated temporal labels using a frame-based temporal ontology. We manually annotated 150 free-text eligibility criteria using the temporal labels and trained a parser using Conditional Random Fields (CRFs) to automatically extract temporal expressions from eligibility criteria. An evaluation of an additional 60 randomly selected eligibility criteria using manual review achieved an overall precision of 83%, a recall of 79%, and an F-score of 80%. We illustrate the application of temporal extraction with the use cases of question answering and free-text criteria querying.

Introduction

As of today, ClinicalTrials.gov archives the summaries for over 110,423 clinical trial protocols online and is a valuable public data set for knowledge discovery and reuse. An important section of each clinical trial protocol is the eligibility criteria section, which specifies the criteria a research volunteer must meet in order to be part of a study. According to a prior study[1], about 38% of eligibility criteria contain temporal constraints, such as "*Intake of ATT during the past 5 years.*" Efficient extraction and representation of such temporal constraints are important steps for computer-based eligibility query formulation and electronic patient screening. However, temporal constraints in clinical research eligibility criteria have not been well studied, partly due to their multi-dimensional complexities.

In clinical research eligibility criteria, temporal expressions show a wide range of choices of terms as well as syntactic and semantic structures. A temporal constraint can be linked to the subject, the object, or both in an eligibility criterion sentence. Relative time (e.g., *Antacids for 4 hours before and 4 hours after itraconazole*) and absolute timestamps (e.g., "*1968/7/4*") are both present in eligibility criteria. There can be multiple temporal relationships, events, or time anchors in one eligibility criterion, such as "*Chronic administration (defined as more than 14 days) of systemic high dose immuno-suppressant drugs during a period starting from six months prior to administration of the vaccine and ending at study conclusion.*" Furthermore, many medical concepts and their abbreviations are commonly used (e.g., *CT or MRI since the onset of lacunar infarct.*)

The complexities of these linguistic features suggest that simple text parsing methods, such as regular expressions, may not be an effective way to extract and structure the complex temporal constraints in eligibility criteria. Automatic extraction of temporal information from free-text documents is an active but still challenging research area[2, 3]. Our goal in this paper is to develop automated approaches for extracting the primary constructs of temporal constraints in clinical research eligibility criteria in order to formalize such temporal expressions.

Several models of time have been developed in the informatics community. Shahar and Musen defined the knowledge-based temporal abstraction (KBTA)[4] framework, which uses ontology-based approaches to support temporal abstraction in clinical domain. KBTA decomposes the task of temporal representation into parallel subtasks, such as *Temporal-context restriction*s for limiting the scope of inference, *Vertical temporal inference* for generalizing data into classes, and *Horizontal temporal inference* for inferring similar types of preposition. The acquired temporal knowledge is available as a Protégé Ontology[5, 6]. Another temporal model, Temporal Constraint Structure (TCS) [7], was proposed to represent temporal expression in discharge summaries. TCS contains a set of fields, including the representation of time internals, and the beginning and end of medical events. A recent temporal

model named CNTRO (Clinical Narrative Temporal Relation Ontology)[8] uses the Web Ontology Language (OWL) and represents links between temporal entities as RDF triples.

These methods provide varying degrees of expressive power to formalize temporal information in clinical narratives. However, there is a gap between these formalized models and free-text clinical narratives. It is a laborious process to manually encode free-text clinical narratives into formalized computable models. Furthermore, to our knowledge, the differences between temporal expressions in eligibility criteria and clinical narratives are understudied. Currently there is no temporal model or temporal extraction tool developed for clinical research eligibility criteria. It is also difficult to adapt temporal extraction tool designed for other text to clinical research eligibility criteria.

We developed a frame-based temporal ontology for clinical research eligibility criteria. In this paper, we describe an ontology-based approach for structuring free-text temporal constraints in clinical research eligibility criteria. We developed a Conditional Random Field (CRF)-based method to automatically annotate the elements of temporal constraints. This method extends previous research of automatic temporal annotation, such as TARSQI[9] for news articles and TimeText[10] for discharge summaries, by focusing on eligibility criteria using a machine learning approach, with the ultimate goal of facilitating automatic patient eligibility determination for clinical research.

Methods

Temporal Parsing Workflow

The classes and slots from our ontology were used as time tags. A CRF-based parser was trained to automate the process of temporal extraction. Figure 1 shows an overview of the system. We used a semantic lexicon[12] to find UMLS recognizable medical terms in free-text eligibility criteria. A Context Scanning Strategy (CSS)[11, 12]-based shallow annotator was used to generate the machine learning features, with which we trained a CRF parser to label temporal elements according to a temporal ontology.
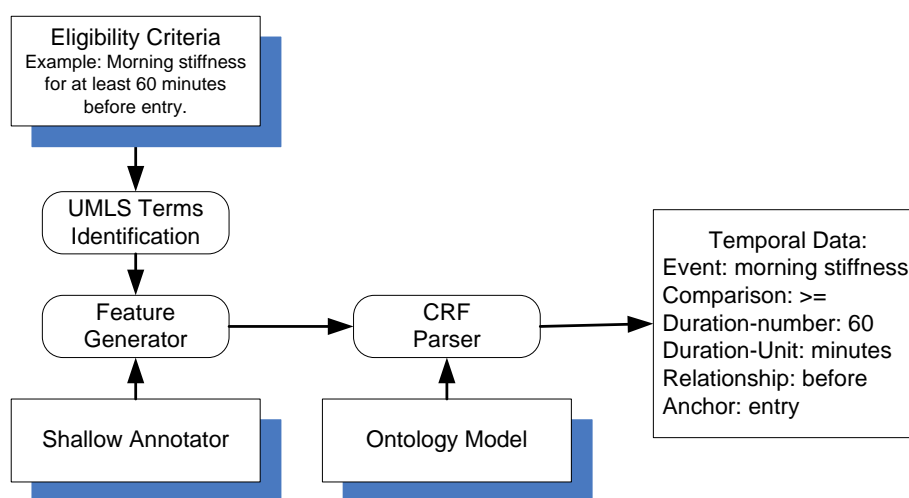

Figure 1: Overview of the temporal processing system

Temporal Ontology

We built a temporal ontology after reviewing the temporal entities of existing knowledge representations, such as the KBTA ontology[4], the TimeML ontology[13] and the TCS[7] model. Figure 2 shows the major entities in our temporal ontology. The class "Temporal Constraint" has five elements: EVENT, ANCHOR, REFERENCE INTERVAL (e.g., *for six months*), TEMPORAL RELATIONSHIP (e.g., *before*, *after*, *during*) and TEMPORAL PATTERN (e.g., *every two weeks*). An event may have zero or more temporal constraints. Each of these classes may contain one or more attributes with finer granularity organized hierarchically. For example, TIME INVETRVAL contains the attributes COMPARSION, DATE, TIME, START_POINT, END_POINT and DURATION. We used the attribute names as semantic labels for later annotation. If a temporal class did not contain any attributes, such as

EVENT, we used the class name as annotation labels. The attributes were organized hierarchically; we preferred to assign finer granularity of labels when possible. For example, the interval, "*six months,*" is a time duration. We assign DURATION-NUMBER to word "*six*" and DURATION-UNIT to word "*months.*" In addition to the semantic labels, we also included the two most common syntactic labels in temporal expressions: CONJUCTION (e.g., *and*, *or*) and MODIFIER (e.g., *last*, *next*). We describe the difference between our temporal ontology and existing temporal knowledge representation in the discussion section.

The coverage of our temporal ontology was evaluated by two independent labelers (MB, CW) using another set of 50 randomly selected eligibility criteria sentence. All temporal entities in the 50 criteria were successfully encoded into the temporal ontology. The human labelers identified 287 temporal entities, among which 76% were exact matches between two labelers, 18% were partial matches with trivial difference caused by the flexibility of our ontology that has more than one correct representation for the same temporal elements (e.g. joined event at different granularity), and 6% where partial matches with significant differences (e.g., for the temporal element "*weekly and biweekly*", one labeler annotated it as DURATION and the other as PATTERN). We consider the inter-rater agreement of the temporal entities within the representation was 94%.
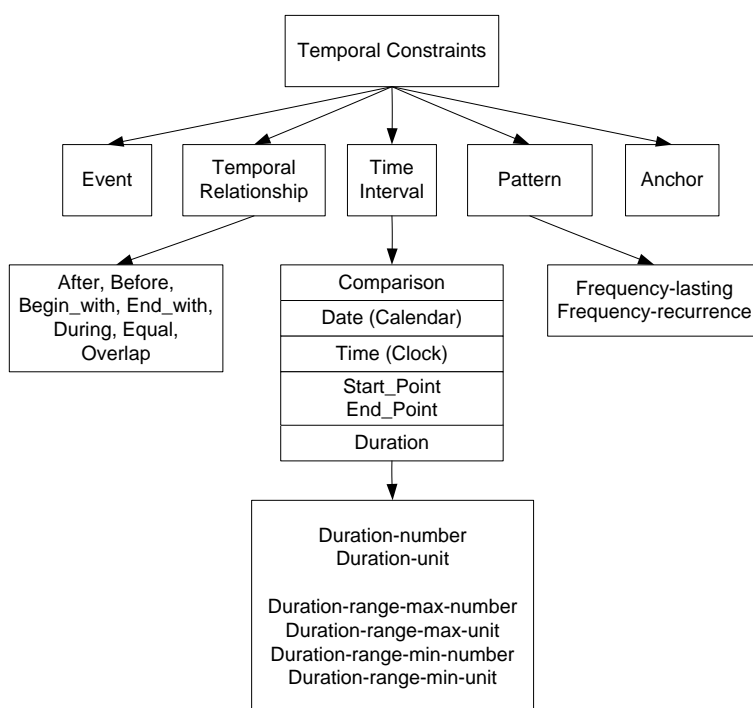

Figure 2: "Has-a" relationship in the temporal ontology for clinical research eligibility criteria.

Parser Development

We used a training corpus with 150 temporal eligibility criteria randomly selected from ClinicalTrails.gov. Each criterion was processed by a UMLS-based annotator[14], which identifies UMLS recognizable terms. The lead author (ZL) examined all temporal expressions in each criterion and annotated all criteria using the temporal tags. These data were used to train the machine-learning-based CRF parser.

We developed a parser to extract temporal information automatically from free-text eligibility criteria using a machine-learning approach. We defined this problem as a sequential tagging task. We used Conditional Random Fields (CRFs)[15] to train a temporal parser from manually-annotated criteria. The primary advantage of a CRF-based method over other probability models, such as hidden Markov models (HMMs), is that it takes into account conditional probabilities while relaxing the independence assumption required by HMMs. CRFs have been successfully applied to different NLP problems, such as part-of-speech tagging[15] and noun phrase chunking[16]. Given

a sequence of criterion features (e.g., the phrases of a criterion), $F = (f_1, f_2, f_3...f_n)$, we looked for the most likely correct sequence of temporal labels, $L = (l_1, l_2, l_3...l_n)$, based on the observation of the features. The possible sets of temporal labels were derived from the temporal ontology as discussed in previous section. We trained our CRF parser with a probability model M that correctly predicts a new label sequence, $L = M(F)$, based on feature set F. Formally, a CRF can be defined as an undirected graph model. The labels, L, are vertices representing random variables and edges representing conditional dependencies. The random variables are assumed conditionally dependent on the observed features, F. The CRF method normally assumes a first-order Markov property, which means the current label is dependent only on its predecessors.

The conditional probability of a temporal label l given the observed feature sequence f is defined as $P(l \mid f)$. The learning process is to find the weight parameter values λ_k that will maximize the likelihood of (l, f) given a training set. The labeling process for a new unseen sequence is achieved by using the Viterbi algorithm[17] which selects the sequence S that has highest likelihood $\arg\max_s P(l \mid f)$. We used the open source package MALLET[18] to implement a parser based on the CRF algorithm. We used linear CRF and used the default Gaussian variance to regularize the weight parameter.

Feature selection is an important component of CRF. We defined three types of features. The simplest features are the words and terms themselves. In the pre-processing step, we used a semantic annotator to identify UMLS recognizable concepts[14]. Clinical concepts can be recognized as terms instead of separate words. For example, *heart attack* can be recognized as a single entity. The second type of feature is common time-related terms. We developed a rule-based feature identification program using the Context Scanning Strategy (CSS)[11, 12] to recognize these words. The CSS is designed to identify and interpret surface temporal features from the eligibility criteria. The surface features include numerals (e.g., *one, 4*), months (e.g., *September, Jan*), week days (e.g.,. *Monday*), and time units (e.g., *minute, day*). The third type of feature is contextual information. For any given term, we used its previous term's features and successor term's features as context.

Results

Table 1 shows the results of temporal annotation and the extracted temporal constraints. As shown in Figure 2, we defined five common temporal elements: EVENT, TEMPORAL_INTERVAL, TEMPORAL_RELATIONSHIP, PATTERN, and ANCHOR. Normally, EVENT is the subject of a criterion sentence, which indicates a required clinical attribute of a patient. Classified using UMLS semantic types[19], EVENT most commonly includes *Therapeutic or Preventive Procedure* and *Disease or Syndrome*, *Pharmacologic Substance* and *Health Care Activity*. ANCHOR is a medical event (e.g., *study entry*) or a specific time point (e.g., *January 30, 2003*). Often, TIME_INTERVAL is used to modify the temporal properties of EVENT and ANCHOR.

We used an interval-based temporal knowledge representation so that everything is a time interval; the beginning or ending point of an interval can be treated as an instantaneous interval. Time duration is used to specify the length of an event or anchor (Table 1, Example 3.1). The time duration normally contains a comparison (e.g., *larger than*, *within*), the units (e.g., *days*, *months*) and the numeric range of time (e.g. *six days*). We also defined attributes to capture temporal expressions that have a minimum and a maximum limit (Table 1, Example 3.2). We labeled temporal relationships in eligibility criteria using Allen's temporal algebra[20] (Table 1, Example 3.3-3.4). The temporal relationship was often used to express relative position between events and anchors in the timeline, such as an event happened before an anchor or an event happening at a time point. The PATTERN attribute was used to capture temporal information that relates to recurring times (Table 1, Example 3.5). The most common event that was modified by a temporal pattern was the dosing information of a pharmaceutical substance (e.g., *take medication every two days, Insulin injection once per day*). In eligibility criteria, many temporal expressions contain conjunctions. (See Table 1.) Example 3.6 shows the use of a conjunction to link two different time interval constraints. One interval constraint was put before the anchor *administration of drug* and the other after this anchor.

Table 1: Annotation results and temporal constraint structure

Eligibility Criteria	Temporal Annotation	Temporal Constraint Structure
Time interval:		
Example 3.1: Morning stiffness for at least 60 minutes.	EVENT=morning_stiffness EQUAL=for DURATION-COMPARISON=at_least DURATION-NUMBER=60 DURATION-UNIT=minutes	Morning_ stiffness for at_least 60 minutes EVENT EQUAL COMPARISON NUMBER UNIT DURATION
Example 3.2: Females: 3 - 10 years past surgical menopause	DURATION-RANGE-MIN-NUMBER=3 DURATION='-' DURATION-RANGE-MAX-NUMBER=10 DURATION-RANGE-MAX-UNIT=years MODIFIER=past EVENT=surgical_menopause	Female: 3 – 10 years past surgical_menopause RANGE-MIN RANGE-MAX RANGE-UNIT MODIFIER EVENT DURATION
Temporal relationship:		
Example 3.3: Use of systemic corticosteroids within 3 months prior to screening	EVENT=use_of EVENT=systemic EVENT=corticosteroids DURING=within DURATION-NUMBER=3 DURATION-UNIT=months BEFORE =prior_to ANCHOR=screening	Use_of systemic corticosteroids within 3 months prior_to screening EVENT DURING NUMBER UNIT BEFORE ANCHOR DURATION
Example 3.4: Patients were followed until August 2003	EVENT=Patients EVENT= were EVENT =followed END_WITH =until DATE= August DATE= 2003	Patients were followed until August 2003 EVENT END_WITH DATE
Pattern:		
Example 3.5: Blood transfusion more than once per month	ENVET=blood_transfusion FREQUENCY-RECURRENCE=more_than FREQUENCY-RECURRENCE=once FREQUENCY-RECURRENCE=per_month	Blood_transfusion more_ than once per_ month EVENT FREQUENCY-RECURRENCE

Eligibility Criteria	Temporal Annotation	Temporal Constraint Structure
Time interval with conjunctions:		
Example 3.6: Excluded for up to 4 hours before and 4 hours after administration of drug	EVENT=excluded EQUAL=for DURATION-COMPARISON =up_to DURATION-NUMBER=4 DURATION-UNIT=hours BEFORE=before DURATION-CONJUCTION =and DURATION-NUMBER =4 DURATION-UNIT =hours AFTER =after ANCHOR=administration ANCHOR=of ANCHOR= drug	Excluded for up_to 4 hours before and 4 hours after administration of drug EVENT / EQUAL / COMPARISON / NUMBER / UNIT / BEFORE / CONJUNCTION / NUMBER / UNIT / AFTER / ANCHOR DURATION1 DURATION2
Example 3.7: Surgery occurred more than 28 days but within 12 weeks prior to enrollment	EVENT=Surgery EVENT=occurred DURATION-COMPARISON=more_than DURATION-NUMBER=28 DURATION-UNIT =days DURATION-CONJUCTION =but DURATION-COMPARISON= within DURATION-NUMBER =12 DURATION-UNIT=weeks BEFORE=prior_to EVENT=enrollment	Surgery occurred more_than 28 days but within 12 weeks prior_to enrollment EVENT / COMPARISON1 / NUMBER1 / UNIT1 / CONJUNCTION / COMPARISON2 / NUMBER2 / UNIT2 / BEFORE / ANCHOR DURATION

To evaluate the precision and recall of the developed temporal parser, 60 randomly selected eligibility criteria were manually labeled by an independent human annotator (CW) as a gold standard. There is no overlap between the 60 testing criteria and the original 150 training criteria. The output of the CRF parser was evaluated against manually labeled criteria. The overall F-score was 79.81%. Different labels have varying F-Scores. For instance, instances of temporal relationship (e.g., DURING, AFTER, BEFORE) have better performance than instances of CONJUCTION, MODIFIER and ANCHOR.

Table 2: Evaluation of temporal extraction

Temporal Elements	Precision	Recall	F-Score
DURATION-UNIT	100.00%	100.00%	100.00%
DATE(CALENDAR)	100.00%	100.00%	100.00%
DURATION-COMPARISON	96.97%	94.12%	95.52%
DURING	100.00%	91.30%	95.45%
DURATION-NUMBER	98.31%	90.63%	94.31%
AFTER	100.00%	82.76%	90.57%
BEFORE	100.00%	81.82%	90.00%
EVENT	85.30%	79.07%	82.07%
FREQUENCY-RECURRENCE	100.00%	66.67%	80.00%
CONJUNCTION	73.68%	58.33%	65.12%
MODIFIER	54.17%	50.00%	52.00%
ANCHOR	40.00%	58.82%	47.62%
DURATION	32.35%	73.33%	44.90%
Average	**83.14%**	**78.99%**	**79.81%**

Discussion

Error Analysis

We manually reviewed eligibility criteria instances that were used in the evaluation and we analyzed errors and their causes. For the EVENT instances, errors were usually caused by conjunction events. For example, in the criterion "At least 3 months since *prior radiotherapy except for palliative radiotherapy to bone lesions,*" the human annotator labeled the whole italic phrase as a medical event, but the machine only labeled the term *prior radiotherapy*. Similar errors also occurred when the several events were connected by complex conjunctions. The event labeling problem in itself is a combination of a Named Entity Recognition task[21] and a sentence chunking task[16]. One solution for improving the event labeling is to separate the event recognition task from the other temporal information parsing. A good event parser will also help to improve the extraction of associated temporal information, such as improving the low recall rate of CONJUCTION and MODIFIER labels. CONJUCTJION and MODIFIER are the only two syntactic labels used in the temporal annotation. One way to improve the accuracy of syntactic labels is to incorporate syntactic learning features to train the temporal parser, such as adding part-of-speech tagging as learning feature.

The performance of the ANCHOR label is lower than others, primarily due to the ambiguities of the EVENT label and ANCHOR label. In clinical research eligibility narratives, the ANCHOR can be an event itself or a time point. When the ANCHOR is an event, the parser is confused and does not know if the term should be labeled an EVENT or an ANCHOR. A solution to this problem is to provide additional syntactic relationships to train the parser. For example, the simple criterion "*heart attack before study entry*" contains two clinical events, "heart attack" and "study entry". The "study entry" is within a preposition chunk "before study entry" which modifies "heart attack." Hence, "study entry" can be identified as a time anchor. The sub-labels of time duration, including UNIT, NUMERAL and COMPARISON, have higher accuracy. However, the accuracy of the DURATION label is low because it is often used on the assisting words (e.g., within *the* last 24 hours) within a duration of time. We plan to improve the temporal parser by addressing these problems in future research.

Related Work

In the biomedical informatics field, several methods have been proposed to address the requirement of modeling free-text clinical narratives. For example, the ERGO annotation[1] is a formal representation of eligibility criteria based on rule grammars and an ontology. ERGO captures the semantics of eligibility criteria by annotating free-text criteria using a predefined ontology, which is similar to our temporal ontology-based approach. However, the ERGO model does not capture the temporal aspects of eligibility criteria[1]. Hyun et al.[22] developed a representation for marking up temporal information in discharge summaries. The markup language consists of five elements: reference point, direction, number, time unit and pattern. Zhou et al.[7] presented the Temporal Constraint Structure (TCS) model for formalizing temporal information in clinical narratives. The authors categorize temporal expressions into different categories, such as "Date and time," "Duration," and "Key events." The TCS used a set of fields (e.g., *event_point, anchor, relation*) to construct temporal constraint intervals and indicate the start and end of clinical events. The TCS fields are similar to our temporal labels. TCS tagger[10] was developed to implement the TCS model. TCS tagger is a component of a temporal reasoning system called TimeText[10]. However, TimeText was developed using discharge summaries and was not tested using clinical research eligibility criteria. In this paper, we used an ontology-based approach to represent temporal information from free-text criteria, which is similar to the ERGO annotation. However, we focused on formalizing the temporal aspects of clinical research eligibility criteria. The temporal representation is not addressed by the ERGO annotation. Our temporal ontology is an extension of previous work on modeling temporal information in clinical trials scheduling tasks[23]. We defined annotation labels for NLP according to the attributes within the temporal ontology. We complemented previous studies by training a CRF-based parser to extract temporal information automatically from free-text eligibility criteria.

Another related method is TimeML[13]. TimeML is a markup language for temporal expressions using an XML format. It marks explicit temporal expressions, such as time, date and duration, using a <TIMEX3> tag. We

provided labels with finer granularity, which can capture important sub-elements within temporal expressions, such as interval duration, comparison expression, date, time, and time units. These fine-grained elements are better aligned with the phrases in eligibility criteria and are crucial for generating computable representation of eligibility criteria. The <EVENT> tag in TimeML is similar to our EVENT label, which captures elements in text that express events modified by temporal constraints. The label <SIGNAL> in TimeML is used to annotate text that indicates relations between temporal prepositions (e.g., *before, after*) and temporal quantifiers (e.g., *two times*). The <LINK> tag specifies relationships between temporal objects in text, such as <ALink> for aspectual relationship, <TLink> for event type and modality, and <SLink> for event subordination. In our study, we labeled temporal relationships using Allen's temporal relations as described in Figure 2. TimeBank[24] is a corpus of 183 news articles annotated with the TimeML specification. However, news articles are significantly different from eligibility criteria narratives; hence we currently cannot compare their annotation results directly.

Temporal Extraction Application Use Case

Use case 1: Retrieving studies with similar temporal events

Figure 3 shows an example of retrieving clinical studies on the ClinicalTrials.gov website that have similar temporal events using formalized temporal information in eligibility criteria. The query is composed through a simple temporal template designed for time intervals. The query template contains four temporal slots: event, comparison, numeric range, and time unit. Each slot is underscored differently. A user can select or enter slot values to perform a temporal query. The query will be used to search a repository of structured eligibility criteria. Example query results are shown in Figure 3. The trial IDs on ClincalTrials.gov are shown together with the retrieved eligibility criteria. For each criterion, parsed temporal elements are underscored in a style corresponding to their temporal attributes in the query template. This use case illustrates the potential use of our temporal processing method for querying free-text eligibility criteria.

Figure 3: Example temporal query

Use case 2: Temporal question answering

Figure 4 shows examples of using temporal extraction to support a question answering system. Questions are constructed using templates with slots. The first two examples show how to use the question system to determine whether a patient is eligible to participate a given clinical trial. A user can type in a patient's clinical attributes and a target trial ID into the question template. The system will first find the selected event within the target clinical trial protocol. Then the event's temporal constraints will be parsed using the temporal extraction method. A comparison will be performed on the structured temporal information to determine if the patient can be included in the given clinical trial. For the first example, the question will be matched against the clinical research eligibility criterion "*Inclusion criteria: Life Expectancy > 6 months*" in clinical trial NCT00001133. The query is asking whether a patient with *life expectancy <= 3 months* will be excluded from this trial. The answer is YES for this question. The third example shows how to ask a question about a temporal fact from a given clinical trial protocol. The question concerns when an event happened for a patient to be included in a given clinical trial. The system will answer the question with a temporal constraint extracted from the eligibility criteria for the given clinical trial.

Much work remains to be done in order to develop a real world implementation of a temporal answering system, such as developing a complete natural language processing method[25] to align the same event concepts with different expressions. However, these simple use cases show that structured temporal information can be used in many applications. It has potential to be used in computational patient recruitment and patient filtering for clinical trial studies.

Figure 4: Example of temporal question answering

Conclusion

We developed a method to automatically structure temporal constraints in clinical research eligibility criteria using a Conditional Random Fields-based algorithm. The overall accuracy of temporal object recognition indicated by F-Score was 79.81%. This method is promising to support two use cases: querying free-text criteria and answering temporal questions with structured temporal expressions.

Acknowledgments

References

1. Tu SW, Peleg M, Carini S, et al. A practical method for transforming free-text eligibility criteria into computable criteria. *Journal of Biomedical Informatics.* 2011;44(2): 239-250.
2. Mani I. Recent Developments in Temporal Information Extraction. *Proceedings of the International Conference Recent Advances in Natural Language Processing (RANLP)* 2005; Borovets, Bulgaria: 401-419.
3. Hripcsak G, Butte AJ, Szolovits P, Kohane IS, Albers DJ. Open Challenges for Incorporating the Temporal Dimension of Medical Records for Phenotype Definitions. *In the Proceedings of AMIA Annual Symposium* 2010; Washington DC: 1364-1366.
4. Stein A, Musen MA, Shahar Y. Knowledge acquisition for temporal abstraction. *Proceedings of AMIA Annual Fall Symposium* 1996; Atlanta, USA: 204-208.
5. Musen MA, Eriksson H, Gennari JH, Tu S, Puerta AR. PROTEGE-II: computer support for development of intelligent systems from libraries of components. *In the Proceedings of Medinfo ('95) Symposium* 1995; Vancouver, Canada: 766-770.
6. Noy N, Grosso W, Musen M. Knowledge-Acquisition Interfaces for Domain Experts: An Empirical Evaluation of Protege-2000. *In the 12 International Conference on Software Engineering and Knowledge Engineering (SEKE)* 2000; Chicago, USA: 5-7.
7. Zhou L, Melton GB, Parsons S, Hripcsak G. A temporal constraint structure for extracting temporal information from clinical narrative. *Journal of Biomedical Informatics.* 2006;39(4): 424-439.
8. Tao C, Wei W-Q, Solbrig HR. CNTRO: A Semantic Web Ontology for Temporal Relation Inferencing in Clinical Narratives. *In the Proceedings of AMIA Annual Symposium* 2010; Washington DC: 787-791.
9. Verhagen M, Mani I, Sauri R, et al. Automating temporal annotation with TARSQI. *Proceedings of the ACL 2005 on Interactive poster and demonstration sessions* 2005; Ann Arbor, Michigan: 81-84.
10. 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(1): 99-106.

11. Vazov N. A system for extraction of temporal expressions from French texts based on syntactic and semantic constraints. *Proceedings of the workshop on Temporal and spatial information processing - Volume 13* 2001; Toulouse, France: 1-8.

12. Descles J-P, Cartier E, Jackiewicz A, Minel J-L. Textual Processing and Contextual Exploration Method. *CONTEXT* 1997; Brazil, University Federal do Rio de Janeiro: 189-197.

13. Pustejovsky J, Castano J, Ingria R, et al. TimeML: Robust Specification of Event and Temporal Expression in Text. *Fifth International Workshop on Computational Semantics (IWCS-5).* 2003; Tilburg University, Netherlands: 337-353.

14. Luo Z, Duffy R, Johnson S, Weng C. Corpus-based Approach to Creating a Semantic Lexicon for Clinical Research Eligibility Criteria from UMLS. *AMIA Summit on Clinical Research Informatics.* 2010; San Francisco, California: 26-31.

15. Lafferty J, McCallum A, Pereira F. Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. *ICML '01 Proceedings of the Eighteenth International Conference on Machine Learning* 2001; San Francisco: 282-289.

16. Sha F, Pereira F. Shallow parsing with conditional random fields. *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology - Volume 1* 2003; Edmonton, Canada: 134-141.

17. Forney GD, Jr. The viterbi algorithm. *Proceedings of the IEEE.* 1973;61(3): 268-278.

18. McCallum AK. MALLET: A Machine Learning for Language Toolkit. 2002.

19. McCray AT. The UMLS semantic network. *Annual Symposium on Computer Applications in Medical Care* 1989; Washington; DC 503-507.

20. Allen JF. Maintaining knowledge about temporal intervals. *Commun. ACM.* 1983;26(11): 832-843.

21. Settles B. Biomedical named entity recognition using conditional random fields and rich feature sets. *Proceedings of the International Joint Workshop on Natural Language Processing in Biomedicine and its Applications* 2004; Geneva, Switzerland: 104-107.

22. Hyun S, Bakken S, Johnson SB. Markup of Temporal Information in Electronic Health Records. *Consumer-Centered Computer-Supported Care for Healthy People.* 2006;122: 907-908.

23. Weng C, Kahn M, Gennari J. Temporal knowledge representation for scheduling tasks in clinical trial protocols. *Proceedings of the AMIA Symposium* 2002; Boston, MA: 879-883.

24. Pustejovsky J, Hanks P, Saurí R, et al. An Overview of the TIMEBANK Corpus. *In the Proceedings of Corpus Linguistics* 2003; Lancaster, United Kingdom: 647-656.

25. Weng C, Wu X, Luo Z, Boland MR, Theodoratos D, Johnson SB. EliXR: An Approach to Eligibility Criteria Extraction and Representation. *Journal of the American Medical Informatics Association (JAMIA).* 2011; Accepted.