Building Timelines from Narrative Clinical Records: Initial Results Based-on Deep Natural Language Understanding

Hyuckchul Jung, James Allen, Nate Blaylock, Will de Beaumont,
Lucian Galescu, Mary Swift

Florida Institute for Human and Machine Cognition
40 South Alcaniz Street, Pensacola, Florida, USA

`{hjung,blaylock,jallen,wbeaumont,lgalescu,mswift}@ihmc.us`

Abstract

We present an end-to-end system that processes narrative clinical records, constructs timelines for the medical histories of patients, and visualizes the results. This work is motivated by real clinical records and our general approach is based on deep semantic natural language understanding.

1 Introduction

It is critical for physicians and other healthcare providers to have complete and accurate knowledge of the medical history of patients that includes disease/symptom progression over time and related tests/treatments in chronological order. While various types of clinical records (e.g., discharge summaries, consultation notes, etc.) contain comprehensive medical history information, it can be often challenging and time-consuming to comprehend the medical history of patients when the information is stored in multiple documents in different formats and the relations among various pieces of information is not explicit.

For decades, researchers have investigated temporal information extraction and reasoning in the medical domain (Zhou and Hripcsak, 2007). However, information extraction in the medical domain typically relies on shallow NLP techniques (e.g., pattern matching, chunking, templates, etc.), and most temporal reasoning techniques are based on structured data with temporal tags (Augusto, 2005; Stacey and McGregor, 2007).

In this paper, we present our work on developing an end-to-end system that (i) extracts interesting medical concepts (e.g., medical conditions/ tests/treatments), related events and temporal expressions from raw clinical text records, (ii) constructs timelines of the extracted information; and (iii) visualizes the timelines, all using deep semantic natural language understanding (NLU).

Our deep NLU system extracts rich semantic information from narrative text records and builds logical forms that contain ontology types as well as linguistic features. Ontology- and pattern-based extraction rules are used on the logical forms to retrieve time points/intervals, medical concepts/ events and their temporal/causal relations that are pieced together by our system's temporal reasoning component to create comprehensive timelines.

Our system is an extension to a well-proven general-purpose NLP system (Allen et al., 2000) rather than a system specialized to the clinical domain, and the temporal reasoning in our system is tightly integrated into the NLP system's deep semantic analysis. We believe this approach will allow us to process a broader variety of documents and complex forms of temporal expressions.

In the coming sections, we first present a motivating example, a real clinical record of a cancer patient. Next, we give an overview of our NLU system including how medical ontology is integrated into our system. The overview section is followed by detailed description of our information extraction and temporal reasoning approach. Then, we discuss our results and conclude.

2 Motivating Example

Our work is carried out as a collaboration with the Moffitt Cancer Center (part of the NCI Comprehensive Cancer Centers), who have provided us with access to clinical records for over 1500 patients. Figure 1 shows a (de-identified) "History of Present Illness" (HPI) section of a Thoracic Consultation Note from this data set.
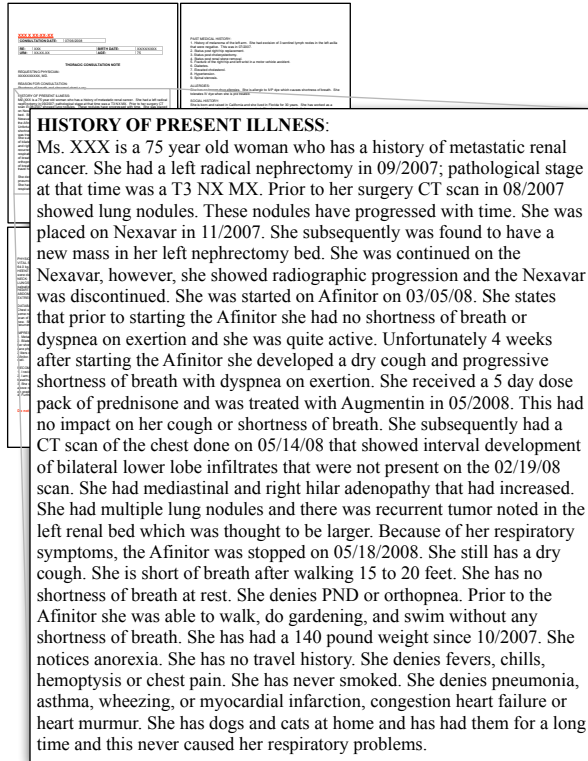
Figure 1: A sample medical record -- Thoracic Consultation Note[1]

HISTORY OF PRESENT ILLNESS:
Ms. XXX is a 75 year old woman who has a history of metastatic renal cancer. She had a left radical nephrectomy in 09/2007; pathological stage at that time was a T3 NX MX. Prior to her surgery CT scan in 08/2007 showed lung nodules. These nodules have progressed with time. She was placed on Nexavar in 11/2007. She subsequently was found to have a new mass in her left nephrectomy bed. She was continued on the Nexavar, however, she showed radiographic progression and the Nexavar was discontinued. She was started on Afinitor on 03/05/08. She states that prior to starting the Afinitor she had no shortness of breath or dyspnea on exertion and she was quite active. Unfortunately 4 weeks after starting the Afinitor she developed a dry cough and progressive shortness of breath with dyspnea on exertion. She received a 5 day dose pack of prednisone and was treated with Augmentin in 05/2008. This had no impact on her cough or shortness of breath. She subsequently had a CT scan of the chest done on 05/14/08 that showed interval development of bilateral lower lobe infiltrates that were not present on the 02/19/08 scan. She had mediastinal and right hilar adenopathy that had increased. She had multiple lung nodules and there was recurrent tumor noted in the left renal bed which was thought to be larger. Because of her respiratory symptoms, the Afinitor was stopped on 05/18/2008. She still has a dry cough. She is short of breath after walking 15 to 20 feet. She has no shortness of breath at rest. She denies PND or orthopnea. Prior to the Afinitor she was able to walk, do gardening, and swim without any shortness of breath. She has had a 140 pound weight since 10/2007. She notices anorexia. She has no travel history. She denies fevers, chills, hemoptysis or chest pain. She has never smoked. She denies pneumonia, asthma, wheezing, or myocardial infarction, congestion heart failure or heart murmur. She has dogs and cats at home and has had them for a long time and this never caused her respiratory problems.

The text of this section provides a very detailed description of what problems/tests/treatments an anonymous cancer patient went through over a period. Such narrative text is common in clinical notes and, because such notes are carefully created by physicians, they tend to have only relevant information about patient medical history.

Nonetheless, there are lots of challenges in constructing complete and accurate medical history because of complex temporal expressions/ relations, medical language specific grammar/ jargons, implicit information and domain-specific medical knowledge (Zhou and Hripcsak, 2007).

In this paper, as an initial step towards constructing complete timelines from narrative text, we focus on sentences with explicit temporal expressions listed below (tagged as *Line 1 ~ 11*) plus a sentence in the present tense (*Line 12*):

- *Line 1*: She had a left radical nephrectomy **in 09/2007**; pathological stage **at that time** was a T3 NX MX.

- *Line 2*: **Prior to** her surgery CT scan **in 08/2007** showed lung nodules.
- *Line 3*: She was placed on Nexavar **in 11/2007**.
- *Line 4*: She was started on Afinitor **on 03/05/08**.
- *Line 5*: She states that **prior to** starting the Afinitor she had no shortness of breath or dyspnea on exertion and she was quite active.
- *Line 6*: Unfortunately **4 weeks after** starting the Afinitor she developed a dry cough and progressive shortness of breath with dyspnea on exertion.
- *Line 7*: She received a 5 day dose pack of prednisone and was treated with Augmentin **in 05/2008**.
- *Line 8*: She subsequently had a CT scan of the chest done **on 05/14/08** that showed interval development of bilateral lower lobe infiltrates that were not present on the **02/19/08** scan.
- *Line 9*: Because of her respiratory symptoms, the Afinitor was stopped **on 05/18/2008**.
- *Line 10*: **Prior to** the Afinitor she was able to walk, do gardening, and swim without any shortness of breath.
- *Line 11*: She has had a 140 pound weight **since 10/2007**.
- *Line 12*: She denies fevers, chills, hemoptysis or chest pain.

In these 12 sentences, there are instances of 10 treatments (e.g., procedures such as "nephrectomy" and drugs such as "Nexavar"), 3 tests (e.g., CT-scan), 13 problems/symptoms (e.g., lung nodules) and 2 other types of clinical findings (e.g., the cancer stage level "T3 NX MX"). There are also 23 events of various types represented with verbs such as "had", "was", "showed", and "was started".

While there are simple expressions such as "on 03/05/08" in *Line 3*, there are also temporal expressions in more complex forms with time relations (e.g., "prior to"), time references (e.g., "at that time") or event references (e.g., "4 weeks after starting Afinitor"). Throughout this paper, we will use *Line 1 ~ 12* as a concrete example based on which we develop general techniques to construct timelines.
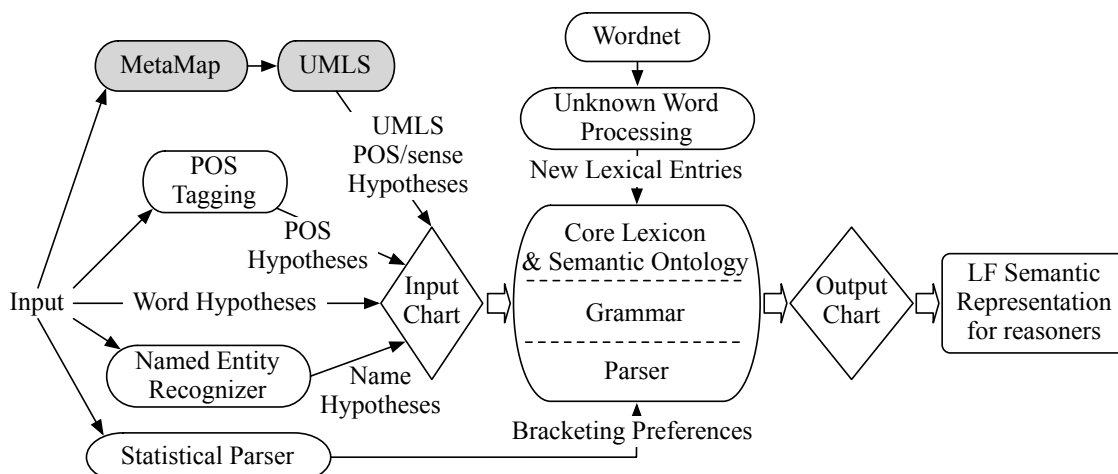

Figure 2: Front-end language processing components with MetaMap and UMLS

3 Natural Language Understanding (NLU) System

Our system is an extension to an existing NLU system that is the result of a decade-long research effort in developing generic natural language technology. The system uses a "deep" understanding approach, attempting to find a linked, overall meaning for all the words in a paragraph. An architectural view of the system is shown in Figure 2.

3.1 Core NLU Components

At the core of the system is a packed-forest chart parser which builds constituents bottom-up using a best-first search strategy. The core grammar is a hand-built, lexicalized context-free grammar, augmented with feature structures and feature unification. The parser draws on a general purpose semantic lexicon and ontology which define a range of word senses and lexical semantic relations. The core semantic lexicon was constructed by hand and contains more than 7000 lemmas. It can be also dynamically augmented for unknown words by consulting WordNet (Miller, 1995).

To support more robust processing as well as domain configurability, the core system is informed by a variety of statistical and symbolic preprocessors. These include several off-the-shelf statisical NLP tools such as the Stanford POS tagger (Toutanova and Manning, 2000), the Stanford named-entity recognizer (NER) (Finkel et al., 2005) and the Stanford Parser (Klein and Manning, 2003). The output of these and other specialized preprocessors (such as a street address recognizer) are sent to the parser as advice. The parser then can include or not include this advice (e.g., that a cer-

tain phrase is a named entity) as it searches for the optimal parse of the sentence.

The result of parsing is a frame-like semantic representation that we call the Logical Form (LF). The LF representation includes semantic types, semantic roles for predicate arguments, and dependency relations. Figure 3 shows an LF example for the sentence "*She had a left radical nephrectomy in 09/2007*". In the representation, elements that start with colons (e.g., :THEME) are semantic roles of ontological concepts, and role values can be a variable to refer to another LF term.

3.2 UMLS Integration

By far the most critical aspect of porting our generic NLU components to the task of understanding clinical text is the need for domain-specific lexical and ontologic information. One widely used comprehensive resource that can provide both is the National Library of Medicine's Unified Medical Language System (UMLS) (Bodenreider, 2004). UMLS was integrated into or system via MetaMap (Aronson and Lang, 2010), a tool also developed by NLM, that can identify and rank UMLS concepts in text.

Specifically, we added MetaMap as a special kind of named entity recognizer feeding advice into the Parser's input chart (see Figure 2). We run MetaMap twice on the input text to obtain UMLS information both for the maximal constituents, and for individual words in those constituents (e.g., "lung cancer", as well as "lung" and "cancer").

The lexicon constructs representations for the new words and phrases on the fly. Our general approach for dealing with how the corresponding concepts fit in our system ontology uses an ontol-

(F *V1* (:* ONT::HAVE W::HAVE) :AFFECTED *V2* :THEME *V3* :MOD *V4* :TENSE W::PAST)
(PRO *V2* (:* ONT::PERSON W::SHE) :PROFORM ONT::SHE :CO-REFERENCE *V0*)
(A *V3* (:* ONT::TREATMENT W::LEFT-RADICAL-NEPHRECTOMY) :DOMAIN-INFO (UMLS .....))
(F *V4* (:* ONT::TIME-SPAN-REL W::IN) :OF *V1* :VAL *V5*)
(THE *V5* ONT::TIME-LOC :YEAR 2007 :MONTH 9)

Figure 3: LF semantic representation for "*She had a left radical nephrectomy in 09/2007*"

(*?x1 ?y2* (*? type1* ONT::HAVE) :AFFECTED *?y2* :THEME *?y3* :MOD *?y4*) (*?x2 ?y2* (*? type2* ONT::PERSON))) (*?x3 ?y3* (*? type3* ONT::TREATMENT ONT::MEDICAL-DIAGNOSTIC) :DOMAIN-INFO *?!info*)	*List of LF patterns*
-extract-person-has-treatment-or-medical-diagnostic>	*Unique rule ID*
(EVENT :type ?type1 :class occurrence :subject ?y2 :object ?y3)	*Output Specification*

Figure 4: An example extraction rule

ogy specialization mechanism which we call ontology grafting, whereby new branches are created from third party ontological sources, and attached to appropriate leaf nodes in our ontology.

The UMLS Semantic Network and certain vocabularies included in the UMLS Metathesaurus define concept hierarchies along multiple axes. First, we established links between the 15 UMLS semantic groups and corresponding concepts in our ontology. Second, we selected a list of nodes from the SNOMED-CT and NCI hierarchies (27 and 11 nodes, respectively) and formed ontological branches rooted in these nodes that we grafted onto our ontology.

Based on these processes, UMLS information gets integrated into our LF representation. In Figure 3, the 3rd term has a role called :domain-info and, in fact, its value is *(UMLS :CUI C2222800 :CONCEPT "left nephrectomy" :PREFERRED "nephrectomy of left kidney (treatment)" :SEMANTIC-TYPES (TOPP) :SEMANTIC-GROUPS (PROC) :SOURCES (MEDCIN MTH))* that provides detailed UMLS concept information. Here, the semantic type "TOPP" is a UMLS abbreviation for "Therapeutic or Preventive Procedure". More details about complex issues surrounding UMLS integration into our system can be found in (Swift et al., 2010).

4 Information Extraction (IE) from Clinical Text Records

In this section, we describe how to extract basic elements that will be used as a foundation to construct timelines. We first describe our general approach to extracting information from LF graphs. Then we give details specific to the various types of information we extract in our system: various clinical concepts, temporal concepts (points as well as intervals), events and temporal relations.

4.1 LF Pattern-based Extraction

Given LF outputs from the NLU system described in Section 3, we use LF pattern-based rules for information extraction. The basic structure of an extraction rule is a list of LF patterns followed by a unique rule ID and the output specification.

Each LF-pattern specifies a pattern against an LF. Variables can appear anywhere except as role names in different formats:

- ?x - (unconstrained) match anything
- ?!x - match any non-null value
- (? x V1 V2 ...) - (constrained) match one of the specified values V1, V2, ...

As an example, the extraction rule in Figure 4 will match LFs that mean a person had a treatment or a medical-diagnostic with explicit UMLS information (i.e., part of LFs in Figure 3 matches). The output specification records critical information from the extraction to be used by other reasoners.

The extraction rules have all been developed by hand. Nevertheless, they are quite general, since a) LF patterns abstract away from lexical and syntactic variability in the broad class of expressions of interest (however, lexical and syntactic features may be used if needed); and b) LF patterns make heavy use of ontological categories, which provides abstraction at the semantic level.

4.2 Clinical Concept Extraction

Among various types of concepts included in clinical records, we focus on concepts related to problems/tests/treatments to build a medical his-

```
((?x1 ?y1 (? type1 ONT::SUBSTANCE) :domain-info
?info :quantifier ?quan)
-extract-substance>
(extraction :type substance :concept ?type1 :umlsinfo
?info :ont-term ?y1 :quantifier ?quan))
```

Figure 5: A rule to extract substances

```
(?x1 ?y1 (:* ont::event-time-rel w::until) :val ?val)
(?x2 ?val (? type2 ont::time-loc) :mod ?mod)
(?x3 ?mod (? type3 ont::event-time-rel) :displacement
?displacement)
(?x4 ?displacement (? type4 ont::quantity) :unit ?unit
:amount ?amount)
(?x5 ?amount ont::number :value ?num)
```

Figure 6: LF patterns to extract a time-span

tory and extract them using extraction rules as described above. Figure 5 shows a rule to extract substances by matching any LF with a substance concept (as mentioned already, subclasses such as pharmacologic substances, would also match).

The rule in Figure 5 checks the :quantifier role and its value (e.g., none) is used to infer the presence or the absence of concepts. Using similar rules, we extract additional concepts such as medical-disorders-and-conditions, physical-symptom, treatment, medical-diagnostic, medical-action and clinical-finding. Here, medical-action and clinical-finding are to extract concepts in a broader sense.[2] To cover additional concepts, we can straightforwardly update extraction rules.

4.3 Temporal Expression Extraction

Temporal expressions are also extracted in the same way but using different LF patterns. We have 14 rules to extract dates and time-spans of varying levels of complexity; for the example in Figure 1 six of these rules were applied. Figure 6 shows LF patterns for a rule to extract temporal expressions of the form "until X days/months/years ago"; for example, here is what the rule extracts for "until 3 days ago":

```
(extraction :type time-span :context-rel (:*
ont::event-time-rel w::until) :reference (time-position
:context-rel (:* ont::event-time-rel w::ago) :amount 3
:unit (:* ont::time-unit ont::day)))
```

From this type of output, other reasoners can easily access necessary information about given temporal expressions without investigating the whole LF representation on their own.

4.4 Event Extraction

To construct timelines, the concepts of interest (Section 4.2) and the temporal expressions (Section 4.3) should be pieced together. For that purpose, it is critical to extract events because they not only describe situations that happen or occur but also represent states or circumstances where something holds. Furthermore, event features provide useful cues to reason about situations surrounding extracted clinical concepts.

Here, we do not formally define events, but refer to (Sauri et al., 2006) for detailed discussion about events. While events can be expressed by multiple means (e.g., verbs, nominalizations, and adjectives), our extraction rules for events focus on verbs and their features such as class, tense, aspect, and polarity. Figure 7 shows a rule to extract an event with the verb "start" like the one in *Line 4*, "She was *started* on Afinitor on 03/05/08". The output specification from this rule for *Line 4* will have the :class, :tense, and :passive roles as *(aspectual initiation)*, *past*, and *true* respectively.

These event features play a critical role in constructing timelines (Section 5). For instance, the event class *(aspectual initiation)* from applying the rule in Figure 7 to *Line 4* implies that the concept "Afinitor" (a pharmacologic-substance) is not just something tried on the given date, 03/05/08, but something that continued from that date.

4.5 Relation Information Extraction

The relations among extracted concepts (namely, conjoined relations between events and set relations between clinical concepts) also play a key role in our approach. When events or clinical concepts are closely linked with such relations, heuristically, they tend to share similar properties that are exploited in constructing timelines as described in Section 5.

5 Building Timelines from Extracted Results

Extracted clinical concepts, temporal expressions, events, and relations (Section 4) are used as a

---

[2] While concept classification into certain categories is a very important task in the medical domain, sophisticated concept categorization like the one specified in the 2010 i2b2/VA Challenge (https://www.i2b2.org/NLP/Relations/) is not the primary goal of this paper. We rather focus on how to associate extracted concepts with other events and temporal expressions to build timelines.

```
(?x1 ?ev (? type1 ont::start) :affected ?affected :tense ?tense :passive ?passive :progressive ?progresive
  :perfective ?perfective :negation ?negation)
-extract-start-event>
(EVENT :type ?type1 :class (aspectual initiation) :subject ?affected :object null :tense ?tense :passive
  ?passive :progressive ?progresive :perfective ?perfective :negation ?negation :ont-term ?ev)
```

Figure 7: An event extraction rule example

foundation to construct timelines that represent patients' medical history. In this section, we present timeline construction processes (as shown in Figure 8), using example sentences from Section 2.

Step 1: We first make connections between events and clinical concepts. In the current system, events and clinical concepts are extracted in separate rules and their relations are not always explicit in the output specification of the rules applied. For instance, Figure 9 shows LFs for the sentence in *Line 7* in a graph format, using simplified LF terms for illustration. The clinical concept "prednisone" and the event "received" get extracted by different rules and the relation between them is not explicit in their output specifications.

To address such a case, for a pair of an event and a clinical concept, we traverse LF graphs and decide that a relation between them exists if there is a path that goes through certain pre-defined concepts that do not separate them semantically and syntactically (e.g., concepts of measure-units, evidence/history, development, and some propositions).

Step 2: Second, we find temporal expressions associated with events. This step is relatively straightforward. While temporal expressions and events get extracted separately, by investigating their LFs, we can decide if a given temporal expression is a modifier of an event. In Figure 9, the time-span-relation (i.e., "in") in the dotted-line box is a direct modifier of the event "was treated".

Step 3: Next, we propagate the association between events and temporal expressions. That is, when the relation between an event and a temporal expression is found, we check if the temporal expression can be associated with additional events related to the event (esp. when the related events do not have any associated temporal expression).

In Figure 9, the event "received" does not have a temporal expression as a modifier. However, it is conjoined with the event "was treated" in the same past tense under the same speech act. Thus, we let the event "received" share the same temporal expression with its conjoined event. Here, the con-

Inputs: Clinical concepts, Temporal Expressions, Events, Relations, LFs
Outputs: Clinical concepts with associated dates or timespans.

Steps:
1. Build links between events and clinical concepts
2. Find associated temporal expressions for events
3. Propagate temporal expressions through relations between events when applicable
4. Compute concrete time values for temporal expressions, taking into account the context of clinical records
5. Compute time values for clinical concepts based on their associated events

Figure 8: Pseudocode for Timeline Construction

joined relation was extracted with relation rules described in Section 4.5, which allows us to focus on only related events.

Step 4: When temporal expressions do not have concrete time values within the expressions, we need to designate times for them by looking into information in their LFs:

- *Event references*: The system needs to find the referred event and gets its time value. For instance, in "4 weeks after starting Afinitor" (*Line 6*), "starting Afinitor" refers to a previous event in *Line 4*. The system investigates all events with a verb with the same- or sub-type of ont::start and Afinitor as its object (active verbs) or its subject (passive verbs). After resolving event references, additional time reference or relation computation may be required (e.g., computation for "4 weeks after").

- *Time references*: Concrete times for expressions like the above example "N weeks after <reference-time>" can be easily computed by checking the time displacement information in LFs with the reference time. However, expressions such as "N days ago" are based on the context of clinical records (e.g., record creation
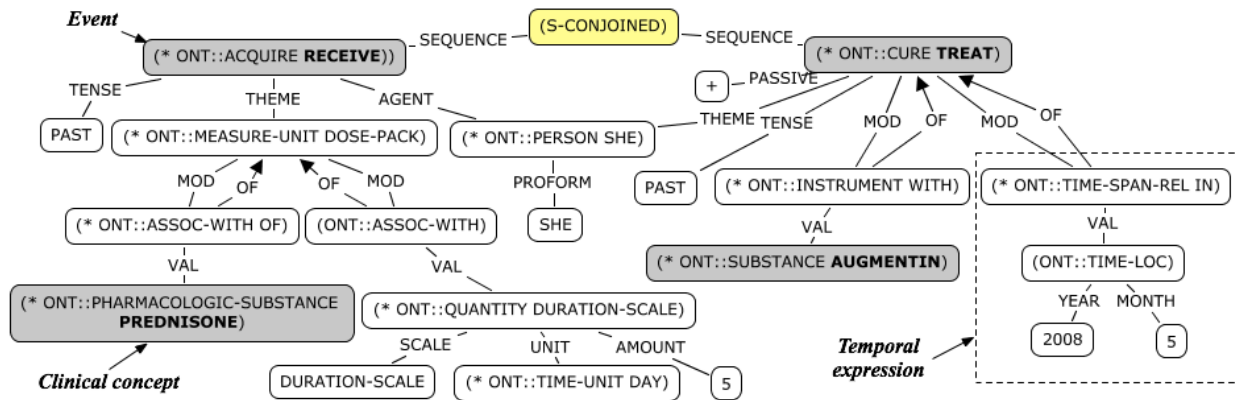
151

Figure 9: Graph format LFs of the sentence in *Line 7* -- "She received a 5 day dose pack of prednisone and was treated with Augmentin in 05/2008."

time). Document creation time is usually represented as metadata attached to the document itself, or it could be retrieved from a database where clinical records are stored. In addition, previously mentioned dates or time-spans can be referred to using pronouns (e.g., "at that/this time"). For such expressions, we heuristically decide that it refers to the most recent temporal expression.

- *Time relation*: Some temporal expressions have directional time relations (e.g., "until", "prior to", and "after") specifying intervals with open ends. When the ending time of a time span is not specified (e.g., "since 10/2007" in *Line 10*). We heuristically set it from the context of the clinical record such as the document creation time.

Step 5: Finally, we designate or compute times on or during which the presence or the absence of each clinical concept is asserted. Since temporal expressions are associated with events, to find time values for clinical concepts, we first check the relations between events and clinical concepts. When an event with a concrete time is found for a clinical concept, the event's class is examined. For classes such as state and occurrence, the concrete time value of the event is used. In contrast, for an aspectual event, we check its feature (e.g., initiation or termination) and look for other aspectual events related to the clinical concept and compute a time span. For instance, regarding "Afinitor", *Line 4* and *Line 9* have events with classes (aspectual initiation) and (aspectual termination) respectively, which leads to a time span between the two dates in *Line 4* and *Line 9*. Currently, we do not resolve conflicting hypotheses.

Assertion of Presence or Absence of Clinical Concepts: To check if a certain concept is present or not, we take into account quantifier information (e.g., none), the negation role values of events, and the verb types of events (e.g., "deny" indicates the absence assertion). In addition to such information readily available in the output specifications of the clinical concept- and event-extraction rules, we also check the path (as in Step 1) that relates the clinical concepts and the events, and the quantifiers of the concepts in the path are used to compute negation values. For instance, given "The scan shows no evidence of lung nodules", the quantifier of the concept "evidence" indicates the absence of the clinical finding "lung nodules".

6 Timeline Results and Discussion

For the example in Section 2 (*Line 1 ~ 12*), we extract all the instances of the clinical concepts and the temporal expressions. Out of 23 events, 17 were extracted. While we missed events such as *state/was* (*Line 5*), *done* (*Line 8*), and *walk/do/ swim* (*Line 10*), our event extraction rules can be extended to cover them if need be.

Figure 10 visualizes the extraction results of the example. We use a web widget tool called Simile Timeline (www.simile-widgets.org/timeline/). Some property values (that were also extracted by rules) are shown alongside some concepts (e.g., weight measurement). Note that not all extracted clinical findings are displayed in Figure 10 because we visualize clinical concepts only when they are associated with temporal expressions in our LFs. For instance, the CT-scan on 05/14/08 in *Line 8* is not shown because the date was not associated with it due to fragmented LFs from the Parser.
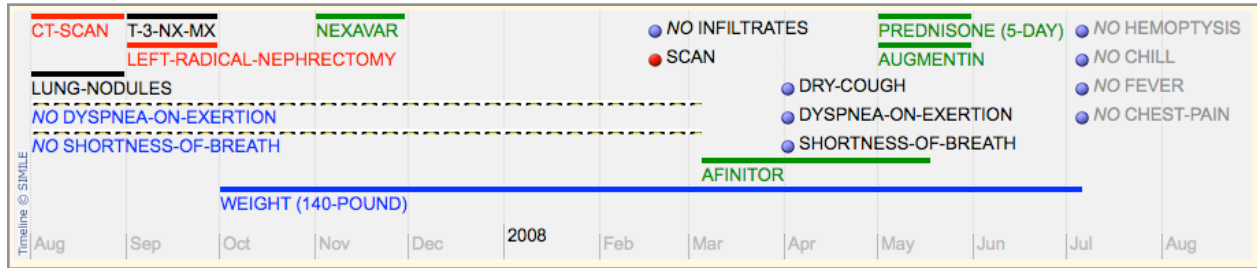
Figure 10: Visualization of timeline results

However, we were still able to extract "no infiltrates" and "scan" from a meaningful fragment.

In addition to the fragmented LF issue, we plan to work on temporal reasoning for concepts in the sentences without explicit temporal expressions, and the current limited event reference resolution will be improved. We are also working on evaluation with 48 clinical records from 10 patients. Annotated results will be created as a gold-standard and precision/recall will be measured.

7 Related Work

Temporal information is of crucial importance in clinical applications, which is why it has attracted a lot interest over the last two decades or more (Augusto, 2005). Since so much clinical information is still residing in unstructured form, in particular as text in the patient's health record, the last decade has seen a number of serious efforts in medical NLP in general (Meystre et al., 2008) and in extracting temporal information from clinical text in particular.

Some of this surge in interest has been spurred by dedicated competitions on extraction of concepts and events from clinical text (such as the i2b2 NLP challenges). At the same time, the evolution of temporal markup languages such as TimeML (Sauri et al., 2006), and temporal extraction/inference competitions (such as the two TempEval challenges, Verhagen et al., 2009) in the general area of NLP have led to the development of tools such as TARSQI (Verhagen et al., 2005) that could be adapted to the clinical domain.

Although the prevailing paradigm in this area is to use superficial methods for extracting and classifying temporal expressions, it has long been recognized that higher level semantic processing, including discourse-level analysis, would have to be performed to get past the limits of the current approaches (cf. Zhou and Hripcsak, 2007).

Recent attempts to use deeper linguistic features include the work of Bethard et al. (2007), who used syntactic structure in addition to lexical and some minor semantic features to classify temporal relations of the type we discussed in Section 4.3. Savova and her team have also expressed interest in testing off-the-shelf deep parsers and semantic role labelers for aiding in temporal relation identification and classification (Savova et al., 2009); although we are not aware of any temporal extraction results yet, we appreciate their effort in expanding the TimeML annotation schema for the clinical domain, as well as their efforts in developing corpora of clinical text annotated with temporal information.

The work of Mulkar-Mehta et al. (2009) also deserves a mention, even though they apply their techniques to biomedical text rather than clinical text. They obtain a shallow logical form that represents predicate-argument relations implicit in the syntax by post-processing the results of a statistical parser. Temporal relations are obtained from the shallow LF based on a set of hand-built rules by an abductive inference engine.

To our knowledge, however, our system is the first general-purpose NLU system that produces a full, deep syntactic and semantic analysis of the text as a prerequisite to the extraction and analysis of relevant clinical and temporal information.

8 Conclusion

In this paper, we presented a prototype deep natural language understanding system to construct timelines for the medical histories of patients. Our approach is generic and extensible to cover a variety of narrative clinical text records. The results from our system are promising and they can be used to support medical decision making.

9 Acknowledgement

153

References

James Allen, Donna Byron, Myroslava Dzikovska, George Ferguson, Lucian Galescu, and Amanda Stent. 2000. An architecture for a generic dialogue shell. Journal of Natural Language Engineering 6(3):1–16.

Mary Swift, Nate Blaylock, James Allen, Will de Beaumont, Lucian Galescu, and Hyuckchul Jung. 2010. Augmenting a Deep Natural Language Processing System with UMLS. Proceedings of the Fourth International Symposium on Semantic Mining in Biomedicine (poster abstract)

Alan R. Aronson and François-Michel Lang. 2010. An overview of MetaMap: historical perspective and recent advances. Journal of the American Medical Informatics Association. 17:229-236.

Juan C. Augusto. 2005. Temporal reasoning for decision support in medicine. Artificial Intelligence in Medicine, 33(1): 1-24.

Steven Bethard, James H. Martin, and Sara Klingenstein. 2007. Timelines from Text: Identification of Syntactic Temporal Relations. In Proceedings of the International Conference on Semantic Computing (ICSC '07), 11-18.

Olivier Bodenreider. 2004. The Unified Medical Language System (UMLS): integrating biomedical terminology. Nucleic Acids Research, Vol. 32.

Jenny Rose Finkel, Trond Grenager, and Christopher Manning. 2005. Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling. Proceedings of the Annual Meeting of the Association for Computational Linguistics.

Dan Klein and Christopher D. Manning. 2003. Fast Exact Inference with a Factored Model for Natural Language Parsing. In Advances in Neural Information Processing Systems 15 (NIPS 2002), Cambridge, MA: MIT Press.

S. M. Meystre, G. K. Savova, K. C. Kipper-Schuler, J. F. Hurdle. 2008. Extracting information from textual documents in the electronic health record: a review of recent research. IMIA Yearbook of Medical Informatics.

George A. Miller. 1995. WordNet: A lexical database for English. Communications of the ACM, 38(5).

R. Mulkar-Mehta, J.R. Hobbs, C.-C. Liu, and X.J. Zhou. 2009. Discovering causal and temporal relations in biomedical texts. In AAAI Spring Symposium, 74-80.

Roser Sauri, Jessica Littman, Bob Knippen, Robert Gaizauskas, Andrea Setzer, and James Pustejovsky. 2006. TimeML annotation guidelines. (available at http://www.timeml.org/site/publications/timeMLdocs/annguide_1.2.1.pdf)

G. Savova, S. Bethard, W. Styler, J. Martin, M. Palmer, J. Masanz, and W. Ward. 2009. Towards temporal relation discovery from the clinical narrative. Proceedings of the Annual AMIA Symposium, 568-572.

Michael Stacey and Carolyn McGregor. 2007. Temporal abstraction in intelligent clinical data analysis: A survey. Artificial Intelligence in Medicine, 39.

Kristina Toutanova and Christopher D. Manning. 2000. Enriching the Knowledge Sources Used in a Maximum Entropy Part-of-Speech Tagger. In Proceedings of the Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora (EMNLP/VLC-2000).

M. Verhagen, I. Mani, R. Sauri, R. Knippen, S.B. Jang, J. Littman, A. Rumshisky, J. Phillips, and J. Pustejovsky. 2005. Automating temporal annotation with TARSQI. In Proceedings of the ACL 2005 on Interactive poster and demonstration sessions (ACLdemo '05), 81-84.

M. Verhagen, R. Gaizauskas, F. Schilder, M. Hepple, J. Moszkowicz, and J. Pustejovsky. 2009. The TempEval challenge: identifying temporal relations in text. Language Resources and Evaluation 43(2):161-179.

Li Zhou, Carol Friedman, Simon Parsons and George Hripcsak. 2005. System Architecture for Temporal Information Extraction, Representation and Reasoning in Clinical Narrative Reports. Proceedings of the Annual AMIA Symposium.

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.