GOALS - Modeling Clinical Guidelines Based on TimeML Concepts

Reinhardt Wenzina
Institute of Software Technology and Interactive Systems
Vienna University of Technology
Favoritenstrasse 9-11
1040 Vienna, Austria
reinhardt@wenzina.com

Katharina Kaiser
University of Applied Sciences St. Pölten
Matthias Corvinus-Straße 15
3100 St. Pölten, Austria
katharina.kaiser@fhstp.ac.at

ABSTRACT

Clinical practice guidelines aim at raising the quality of healthcare. They are written in a narrative style and have to be translated into a computer-interpretable guideline (CIG) to be usable in a clinical software application. In this project we present the GOALS methodology which defines a step-wise approach to support this modeling process. The methodology is specified independently from the target CIG language and uses a guideline's text annotated with temporal concepts provided by TimeML as a starting point. It describes step-by-step how parts of the guideline's model can be generated and finally assessed by means of an evaluation scheme. By means of a scenario-based evaluation we show the applicability of GOALS by translating temporally-related sentences of a clinical protocol into its semi-formal model. Thus, we conclude that this methodology indeed supports the translation process.

Categories and Subject Descriptors

I.2.1 [**Artificial Intelligence**]: Applications and Expert Systems—*Medicine and science*; I.2.4 [**Artificial Intelligence**]: Knowledge Representation Formalisms and Methods—*Representation languages, Temporal logic*; I.2.7 [**Artificial Intelligence**]: Natural Language Processing—*Language parsing and understanding, Text analysis*

General Terms

Theory, Algorithms

Keywords

Temporal knowledge representations; computer-interpretable clinical guidelines; medical guideline modeling; TimeML; temporal relations; Information Extraction (IE); Natural Language Processing (NLP)

1. INTRODUCTION

Clinical practice guidelines (CPGs) [6] are documents commonly published in a narrative format including recommendations describing appropriate care for the management of patients with a specific clinical condition, such as diabetes or chronic heart failure. CPGs are often used within clinical decision-support systems, but need to be represented in specialized languages (see [10] for comparison and overview) therefore.

One important characteristic within CPGs is temporal information that affects the execution of individual activities within clinical processes. Often, this temporal information is only vaguely or incompletely described, which increases the complexity of the temporal constraints. This is, amongst others, one aspect that makes it a very cumbersome and complex task to transform a narrative CPG into a computer-interpretable guideline (CIG).

In this work we present a methodology to support the CIG-modeling task using TimeML [11], a specification language for events and temporal expressions in natural language. As CPGs describe activities (corresponding to events) and contain (temporal) expressions for ordering and relating them, we propose to use this specification for annotation. We developed a methodology using these annotations and Natural Language Processing (NLP) techniques, which supports the guideline modeler by having the temporal aspects related to events and activities identified and having parts of the final CIG generated semi-automatically. This is especially important as dealing with temporal relations in CPGs are one of the major challenges in CIG modeling.

In the following section we give an overview of automated methods in CIG modeling and the application of TimeML. In section 3 we outline our motivation, followed by our approach introducing basic concepts of TimeML in the context of CPGs and by presenting our methodology including the proposed evaluation scheme in section 4. In section 5 we show the applicability of our methodology using a real-world example. Section 6 concludes our ideas and gives an outlook on future work.

2. BACKGROUND AND RELATED WORK

Various approaches based on NLP techniques have been developed to support the modeling process of CPGs. A rule-based information extraction method was presented by Kaiser et al. [4] which detected activities in control flow related aspects, however, to identify their chronological order was not part of the method. Thorne et al. [21] solved this problem by developing a supervised approach which identified process fragments and determined the temporal relations among the medical activities. The detection of temporal relations was limited to simple before/after relations and, therefore, was not appropriate to model the temporally complex control flow related aspects of a guideline.

Generally, the description of temporal relationships among activities is a domain independent challenge and is not restricted to the medical area only. Therefore, Pustejowsky et al. [11] specified the TimeML markup language to annotate events, time expressions, and temporal relations among them in narrative texts. Originally, TimeML was applied to articles of the newswire domain and in the following also transferred to other application areas. TimeML increasingly gained importance in the natural language processing community – not least because a revised and interoperable version of TimeML was published as an ISO standard [12, 2].

UzZaman et al. [22] implemented the TRIOS system containing a semantic parser to extract events, their linguistic features, and relations based on TimeML and showed the flexibility of the specification language by extending it to the authors' special needs. OntoTimeFL [7], another formalism based on TimeML, categorized complex events with regard to their specific features. The THYME[1] project was one of the first approaches which applied TimeML to the medical domain. It focused, both, on the development of a temporal relation annotation scheme and an annotation guideline for clinical free texts.

The automatic generation of TimeML compliant annotations is the goal of the TARSQI Toolkit (TTK). It was built to answer temporally based questions concerning events in news articles [23]. As temporal expressions in medical notes show a higher complexity than in news articles, Reeves et al. extended the TTK's time tagger and developed the MED-TTK [14]. The upcoming version of TTK contains extensions to the medical domain but its implementation is still in progress [24]. Another prototypical application to generate TimeML annotations was built by Gooch [3]. He used the UMLS – the Unified Medical Language System [5] – to categorize medical concepts as events according to the TimeML specification.

3. RESEARCH QUESTION, METHOD, AND EXPECTED RESULT

The TimeML specification language increasingly gains importance due to the possibility of annotating events and their temporal relations in natural language texts. TimeML focuses on (1) the identification of temporally anchored events, (2) their ordering by means of temporal reasoning, (3) the dealing with vaguely specified temporal expressions (e.g., post-dinner), and (4) the reasoning about the duration of an event [11].

As temporal aspects have great significance within clinical guidelines [20] (e.g., the description of care-paths), temporal reasoning methods may support the automatic modeling process of a guideline. These methods are based on the temporal representation of a document consisting of temporal expressions, concept primitives, and temporal relations [19] in order to handle vague and/or complex temporal dimensions. TimeML fulfills these requirements and, furthermore, has already been successfully applied to medical texts such as clinical narratives and discharge summaries. However, existing research results can only partially be adopted because clinical guidelines differ in many ways. Clinical narratives, for example, are full of abbreviations, contain explicit time information (e.g., laboratory tests, doctor's visits) and represent the patients's progression of illness, to name but a few. On the contrary, the language in guidelines is highly sophisticated, time and date information is only implicitly known (e.g., first trimester of pregnancy) and, moreover, often expressed vaguely.

Although there are differences between medical guidelines and narratives (as mentioned above), TimeML is a possible solution to annotate the various temporal information aspects of clinical guidelines, as it includes Allen's algebra of intervals [1], which is also implemented in CIG languages such as Asbru. Consequently, our research question is:

Can temporal annotations based on TimeML support the modeling process of a clinical practice guideline into its computer interpretable representation?

This support comprises the automatic generation of parts of the CIG model as well as the extension of authoring tools. Both objectives could save time and increase the quality of the translation process.

The following steps of our approach should lead us towards finding a satisfactory answer to our research question.

1. Discussion of the temporal concepts of TimeML in the context of guidelines.

2. Development of a multi-step methodology for an automatic translation process, independent from the target CIG language.

3. Elaboration of an evaluation scheme to define the levels of support of the methodology.

4. Application and evaluation of the multi-step methodology onto a non-trivial example of a guideline or protocol.

Step 4 will show the applicability of TimeML to our problem.

4. APPROACH

Following our steps described above, we discuss the basic concepts of TimeML in the context of clinical guidelines below. Thereafter, we present our methodology and an evaluation scheme to verify its benefits.

[1]https://clear.colorado.edu/TemporalWiki

4.1 Temporal Concepts of TimeML

4.1.1 Events & Timex:

The general definition of an *event* is described in the TimeML annotation guideline as *".. a cover term for situations that happen or occur"* [16]. It can be punctual or last for a period of time. For the usage of TimeML in medical records the term was redefined for an i2b2 challenge in 2012 to *".. anything that is relevant to the patient's clinical timeline"*[2]. This statement shows a certain fuzziness in the definition, nevertheless we will adopt it to clinical guidelines. As our approach starts from an already correctly annotated guideline, such a vague definition does not really have an effect on our general methodology.

Timex expressions are primarily used to represent explicit temporal expressions (e.g., times, dates). Although clinical guidelines hardly contain such time stamps, *Timex* expressions are used to represent durations, frequencies, etc.

4.1.2 Links

The relation between events and/or between events and timex expressions is defined as a link. TimeML distinguishes among

- TLink which defines a temporal relation in order to build up a chronology of events (e.g., to show a sequence of consecutive tasks).
- SLink which expresses a subordination relation of events, often found in conditional sentences. Thus, they play an important role for describing condition-based clinical activities.
- ALink which defines an aspectual relation, showing the progression of an event (start, finish, etc.).

The examples[3] below show selected SLinks (the introducing *event* is marked "in bold" and the consequence, which takes the place of the subordinated *event*, is "underlined").

(1) *If there is no **response** the drug should be <u>discontinued</u>.*
(2) ***Women with pain** should be <u>re-examined</u> after two hours.*
(3) *The partogram should be <u>used</u> once labour is **established**.*

4.1.3 Time Anchoring

Compliant with the original specification of TimeML, every temporal link references to the document creation time (DCT). As the DCT contains no important information for clinical guidelines, we use the concept of 'narrative time' introduced in [9] and [13]. It describes the current temporal anchor for events in a guideline and it changes during the reading process. This concept leads to fewer temporal links without losing temporal information. In Figure 1 the narrative time anchors for the clinical protocol *Management of active low-risk labour - Admission for Birth* [15] are shown. In this case the document structure of the protocol has a major influence on the definition of the narrative times.

[2]https://clear.colorado.edu/TemporalWiki
[3]The events are identified compliant with the guideline presented in [18].

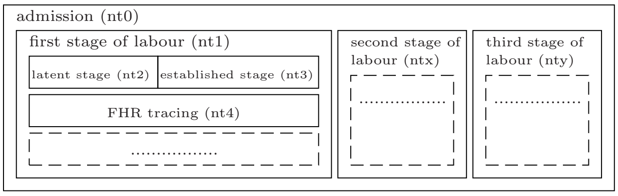

Figure 1: The narrative times (nt?) within the protocol.

4.2 The Multi-Step GOALS-Methodology

One of the objectives of our project was to develop a step by step process (see Figure 2) to model a guideline – annotated with TimeML – independent from the target language (e.g., Asbru). As TimeML only represents temporal concepts, only time related information dimensions of a guideline are focused on.

CIG languages, in general, are formal languages with a defined syntax and defined semantics. From the information extraction point of view we can interpret them as templates with information slots to fill. Consequently, in Step (1) of our methodology the mapping of time related concepts (TLinks, SLinks, and ALinks) to templates of the target language has to be defined and the information slots identified. Step (2) deals with locating of information directly kept in the attributes of the TimeML annotations and the information which can be derived through transitive chains of TimeML links. Step (3) linguistically preprocesses the original text (e.g., sentence splitting, co-reference resolution, identification of adverbial phrases, etc.) depending on the selected target language and the information needs. In Step (4), empty slots or incorrect values in the slots have to be evaluated. Furthermore, we have to check if an extension to the TimeML specification can compensate these deficits. In such a case, we extend the TimeML specification and adapt the annotations accordingly. Otherwise, missing information must be sought from lexical resources (e.g., medical vocabularies) in Step (5). If there are still any open slots, the process will have to be restarted at Step (1). This whole procedure is called the GOALS-methodology which is an acronym of the verbs defining the individual steps.

Finally, the outcome represents parts of a CIG model. Generally, every step of the process can lead to the generation of additional templates, which may also represent non-temporal aspects of a guideline.

In [26] we developed a method to identify *condition-action sentences* by using a heuristic-based information extraction method which led to a recall value of 75% and a precision value of 88%. The results indicate that Step (3) can be supported by means of natural language processing methods. Additionally, in [27] we demonstrated a method to automatically identify condition-based activities for control-flow related aspects in a guideline document based on temporal concepts. Therein, we proposed an extension to the TimeML specification and gained an accuracy of 68%. Consequently, this approach is a particular realization of Step (4) of GOALS.

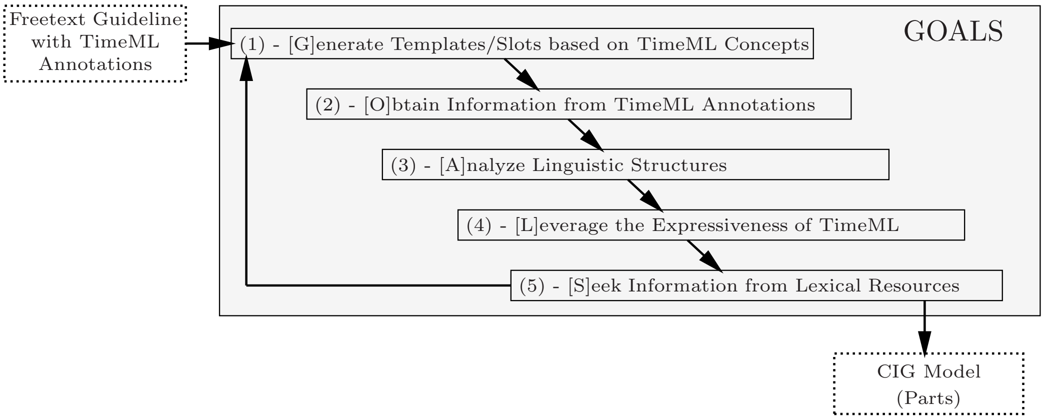

Figure 2: The GOALS-Methodology.

4.3 Evaluation Scheme

The manual process of modeling a guideline is supported by various software tools such as the Document Exploration and Linking Tool with Add-ons DELT/A [25]. This tool displays both the original guideline text and its corresponding (semi-)formal representation next to each other. The modeler marks a piece of the text in the original guideline and selects the appropriate structure (=template) of the target language, transfers the information manually into the slots and adds medical knowledge where it is necessary. These modeling steps can be simplified by our GOALS methodology in many ways. Consequently, we define different levels of support which can also be used as a kind of evaluation scheme.

Level A: The templates and information slots of the CIG language are identified correctly.

Level B: The content of attributes of TimeML annotations to the corresponding slot of the target language is transferred accurately.

Level C: Linguistic processing of the analyzed text provides the appropriate phrases – semantically correct but in different wording.

Level D: The information extraction methods deliver the phrases, words, or information entities in a standard notation format (e.g., *verb-object* notation for activities [8]).

Level E: The templates are completely filled - no slot is left open and the result is completely consistent with the gold standard.

Every achieved level leads to a significant reduction of workload for the guideline modelers.

5. SCENARIO-BASED EVALUATION

Knowledge engineers and medical experts have to work together to generate a computer-interpretable guideline. The modeling of CPGs can be done in different ways and is supported by various authoring/editing tools depending on the target language. Several approaches have been developed to ease this modeling process – one of them is the Many-Headed Bridge (MHB)[17]. MHB specifies an intermediate representation language which provides a more semi-structured and less formal format and bridges the gap between the free-text guideline and its corresponding formalized model.

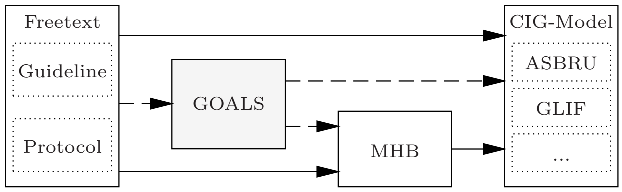

Figure 3: Standard modeling processes (solid arrows) - GOALS integrated processes (dashed arrows).

We have decided to show the applicability of our GOALS methodology by supporting the modeling process of a protocol into its MHB format (see Figure 3) because: (1) protocols are formulated in a more precise way than a guideline, (2) temporal relations in protocols are not as vaguely defined as those in a guideline, and (3) MHB is less formal but also contains all information dimensions for further translation into a formal language.

Supposing our methodology supports this individual modeling process, the conclusion to apply it successfully to more complex tasks can be drawn and would lead to a positive answer of our research question.

5.1 MHB Dimensions and Aspects

In MHB the knowledge of a CPG is represented in a series of chunks that correspond to a certain bit of information in the CPG (e.g., a sentence, part of a sentence, more than one sentence). These chunks are associated to predefined dimensions such as control flow, data flow, evidence, and temporal concepts. The aspects of every dimension are described by using natural language, partly copied from the original guideline text or subsumed and enriched with the knowledge of the modeler. As there is no explicit specification on how to formulate these aspects, their descriptions can syntactically vary among different modelers. Nearly every information dimension of MHB (all except 'evidence') can contain temporal concepts and therefore can be associated to TimeML concepts later on.

5.2 GOALS in Action

We selected a text passage of the guideline *Management of active low-risk labour - Admission for Birth [15]* according to the following criteria. It should (1) contain MHB-control dimensions, (2) express a condition-based activity, (3) consist of at least two consecutive sentences which are temporally related to each other in order to get a higher complexity involved, and (4) contain different kinds of subordination clauses (e.g., conditional and causal clauses).

The corresponding MHB model – which was developed by a knowledge engineer – was finally used as a Gold Standard.

The selected sentences and the contained TimeML concepts (annotated in **bold style** and marked up with '_*e?*' for events and '_*t?*' for timex expressions) are shown below.

(1) When **FHR tracing**_*e1* is reassuring at **admission**_*e2*, the woman should be **allowed**_*e3* to move freely, even if membranes are not **intact**_*e4*.
(2) When **FHR tracing**_*e5* is non reassuring it should be maintained and **reevaluated**_*e6* after a period of **20 minutes**_*t1* because the fetus could be in a quiet **period**_*e7*.

The verb 'maintained' in the second sentence is not tagged as an event, because if we consider 'FHR tracing' as a narrative container, 'maintain' expresses no event during the given time span. Table 1 shows the identified TimeML links and their relevance to the modeling process. In general, SLinks

Table 1: Selected sentences, their TimeML-Links (S=subordination link and T=temporal link – including their attribute 'relType') and their mapping to MHB dimensions (t0 describes the date of admission); column 'rel' describes the relevance for the model

Sent	#	Link	relType	MHB	rel
1	1	S(e1, e3)	conditional	control	yes
	2	S(e4, e3)	conditional	control	yes
	3	T(e1, e2)	simultaneous	time	yes
	4	T(e2, t0)	identity	time	yes
	5	T(e2, e4)	simultaneous	time	no
	6	T(e1, e4)	simultaneous	time	no
2	7	S(e5, e6)	conditional	control	yes
	8	S(e7, e6)	modal	backgr.	yes
	9	T(e5, e7)	simultaneous	time	no
	10	T(e6, t1)	after	time	yes
	11	T(t0, t1)	after	time	no
1+2	12	T(e1, e5)	identity		yes

of type 'conditional' describe condition/action sentences and therefore are relevant to the model (see Table 1: #1, #2, and #7). Link #8 shows a structurally-based subordination relation representing a reason clause. Furthermore, TLinks, which anchor events to specific Timex expressions, are relevant for the time dimension in MHB (links #4 and #10). The TLink #3 was also included because it is transitively linked to t0 via e2. Link #12 shows the identity e1 = e5 and in consequence, link #3 has to be modeled, too. Due to the expressivness of the target language MHB, the links #5, #6, #9, and #10 are marked as not relevant.

5.2.1 [G]enerate Templates and Slots

When going through the list of temporal relations of Table 1 the corresponding MHB structure is built (see Fig. 4 – step #1). Link #1 indicates a condition based activity, therefore a new control tag – containing an if-then tag and its attributes – is generated. As there is no container for the control tag so far, it is embedded within a surrounding chunk tag. As link #2 is also related to event *e3*, no new control tag is necessary. A new time tag is created for link #4 because it relates to the explicit timex expression *t0* in contrast to link #3. The subordination link #7 belongs to a new sentence, hence a new control tag with a subordinated if-then tag is created. The reason clause of link #8 shows an explanation for an event and therefore, is bound to the MHB background tag. The temporal link #10 mandatorily results in a time tag and the identity of *e1* and *e5* expressed in link #12 shows that although two sentences are analyzed together, only one chunk tag is needed.

5.2.2 [O]btain Information from TimeML

The TimeML annotations contain specific information about events which is transferred directly into the open information slots (see Fig. 4 – step #2). In our case we identify the modality attribute 'should' of the subordinated events in links #1 and #7 and map them to the degree-of-certainty attribute of the control tag.

5.2.3 [A]nalyze Linguistic Structures

As MHB aspects represent chunks of information – commonly as parts of sentences – these sentences have to be linguistically analyzed. The control tag, for example, needs the antecedent and the consequent of the condition clause for its attributes 'condition' and 'result' . The anaphora resolution in '.. it should be maintained' leads to 'FHR tracing should be maintained'. This resolution is necessary because of the splitting of the conditional sentence. The reason clause in the first sentence describes some background information and therefore, the identification of adverbial clauses is necessary. The resulting structure is shown in Fig. 4 – step #3.

5.2.4 [L]everage the Expressiveness of TimeML

The first sentence contains two different subordination links describing two antecedents. Hence, the condition attribute also contains both. Structurally, the second subordination clause is an SLink of type condition but semantically it only describes background information. In order to solve this problem, we introduce new 'relType' attributes replacing the original 'condition' attribute. Furthermore, we need to distinguish between conditional clauses and time clauses. The

	Step #1	Step #2	Step #3
chunk			
control			
if-then			
condition			FHR tracing is reassuring at admission Even if membranes are not intact [ERROR]
degree-of-cert.		should	should
result			the woman should be allowed to move freely
control			
if-then			
condition			FHR tracing is non reassuring
degree-of-cert.		should	should
result			FHR tracing [it] should be maintained and reevaluated [after a period of 20 minutes]
background			
explanation			
information			because the fetus could be in a quiet period

Figure 4: GOALS compliant modeling process: from protocol to MHB – steps #1 to #3.

first one is introduced by 'if' and the second one by 'when'. We propose the following extensions to the 'relType' in subordination links:

- *time_conditional_restrictive*: time clauses introduced by 'when', semantically describing a restriction, and
- *state_conditional_informative*: conditional clause introduced by 'if', semantically describing non-restrictive information.

According to the values of the new attributes, we have to reset the 'relType' to *'time_conditional_restrictive'* for the links #1 and #7 and to *'state_conditional_informative'* for the link #2. Following this extension, we are able to generate a new background tag and solve the problem of the wrongly assigned antecedent in the first control tag. Additionally, due to the relation chain S(e1,e3) - T(e1,e2) - T(e2,t0), we set the information slots for the time tag based on S(e1,e3) – subject = e1 and start = e3.

5.2.5 [S]eek Information from Lexical Resources

In order to complete the structure of the MHB chunk, we have to add the data tag describing the piece of information used in this chunk (e.g., FHR tracing). We use the Unified Medical Language System's Semantic Network (UMLS SN) [5] to assign clinical concepts to nouns or gerundive constructions in the sentence. Generally, if the semantic type of the clinical concept is 'finding', 'organism function', or 'qualitative concept', it can be assigned to the name attribute of the data tag. The results of steps #4 and #5 are shown in Fig. 5.

5.3 Analysis and Discussion

Knowledge engineers, when modeling a guideline in MHB, are allowed to change the wording of the text when conserving the original semantics. Consequently, an evaluation of our results with the gold standard on the basis of a 'word by word' matching technique does not seem adequate. Thus, a comparison based on semantics is the appropriate solution to handle this situation.

The analysis of the results of step #5 (see Figure 5) shows that the MHB chunk (=templates), its dimensions and aspects are correctly identified by our method. Moreover, all thirteen aspects (=information slots) are correctly filled. The 'result' slot of the 'if-then' aspect is set to 'woman should be allowed to move freely' instead of 'No limitations in movements of woman'. Semantically both phrases express the same statement but the wording is different. The second 'result' slot shows a very similar effect. As there is no specification in MHB about the exact formulation of the aspects in natural language (as already discussed above), both phrases are equivalent. The different wording in the 'precise-value' of the 'time' dimension does not miss important information as well as the 'subject' in the second 'time' dimension.

The application of our GOALS-methodology for this particular case reaches 'Level C' of our evaluation scheme and therefore shows that the usage of TimeML concepts can ease the modeling process in the given scenario.

An implementation of the results of this methodology into existing software modeling tools can lead to savings in time because

- the phrases in the guideline are already identified,
- the correct information extraction templates selected, and
- the information slots pre-filled.

The real raise of efficiency and effectiveness however, depends on the experience of the modeler and is part of a later evaluation project.

6. CONCLUSIONS AND FURTHER WORK

In this paper we presented the methodology GOALS, which proposes a step-by-step guidance to generate a computerized model of a clinical practice guideline. The methodology acts on temporal concepts provided by the TimeML specification language and works independently from the target CIG language. A 5-level evaluation scheme was designed in order to inspect the quality of the generated model. We showed the

				Step #4 + #5	Gold Standard
chunk					
	control				
		if-then			
			condition	FHR tracing is reassuring at admission	FHR tracing is reassuring at admission
			d-o-c	should	should
			result	woman should be allowed to move freely	No limitations in movements of woman
	background				
		explanation			
			info	Even if membranes are not intact.	Even if membranes are not intact.
	time				
		subject		FHR tracing	FHR tracing
			start-pv	At admission	At admission
	data				
		usage			
			name	FHR tracing	FHR tracing
	control				
		if-then			
			condition	FHR tracing is non reassuring	FHR tracing is non reassuring
			d-o-c	should	should
			result	FHR tracing should be maintained and reevaluated after a period of 20 minutes	Maintain AND reevaluate FHR tracing
	time				
		subject		FHR tracing	Reevaluate FHR tracing
			start-pv	After 20 minutes	After a period of 20 minutes
	background				
		explanation			
			info	because the fetus could be in a quiet ...	because the fetus could be in a quiet period

Figure 5: GOALS compliant modeling process: from protocol to MHB – steps #4 to #5 in comparison to the Gold Standard.

proper functioning of GOALS in a scenario-based evaluation where temporally related guideline sentences were transformed into the CIG language MHB. The result showed that our methodology supports the laborious translation task of the modelers and therefore, our research question was proven right.

Ongoing steps will be (1) the application of our methodology to an entire protocol or guideline, (2) the definition of target language specific modeling rules, (3) the implementation of these rules in a model authoring tool, and (4) a comprehensive evaluation of the time saving of the modeling experts.

As GOALS focuses on temporal concepts and the relations among them, certain information dimensions of a guideline (e.g., levels of evidence) have not been taken into consideration yet. However, our methodology will easily be able to be extended to future requirements due to its flexible structure.

7. ACKNOWLEDGMENTS

This research was carried out as part of project no. TRP71-N23 funded by the Austrian Science Fund (FWF).

8. REFERENCES

[1] J. F. Allen. Maintaining knowledge about temporal intervals. *Commun. ACM*, 26(11):832–843, Nov. 1983.

[2] Association francaise de normalisation and International Organization for Standardization. Language resource management. Semantic annotation framework (SemAf). Time and events SemAF-Time, ISO-TimeML, 2013.

[3] P. Gooch. A lightweight, pattern-based approach to identification and formalisation of TimeML expressions in clinical narratives. In *The Sixth Informatics for Integrating Biology and the Bedside (i2b2) Natural Language Processing Challenge for Clinical Records*, 2012.

[4] K. Kaiser, A. Seyfang, and S. Miksch. Identifying treatment activities for modelling computer-interpretable clinical practice guidelines. In *Knowledge Representation for Health-Care*, volume 6512 of *Lecture Notes in Computer Science*, pages 114–125. Springer Berlin Heidelberg, 2011.

[5] D. Lindberg, B. Humphreys, and A. McCray. The Unified Medical Language System. *Methods of Information in Medicine*, 32(4):281–291, 1993.

[6] K. N. Lohr and M. J. Field. *Clinical practice guidelines: directions for a new program.* National Academy Press Washington, D.C, 1990.

[7] F. Mele and A. Sorgente. A formalism for temporal annotation and reasoning of complex events in natural language. In *DART@AI*IA*, volume 771 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2011.

[8] J. Mendling, H. A. Reijers, and W. M. P. van der Aalst. Seven Process Modeling Guidelines (7PMG). *Inf. Softw. Technol.*, 52(2):127–136, Feb. 2010.

[9] T. Miller, S. Bethard, D. Dligach, S. Pradhan, C. Lin, and G. Savova. Discovering temporal narrative containers in clinical text. In *Proceedings of the 2013 Workshop on Biomedical Natural Language Processing*, Sofia, Bulgaria, August 2013. Association for Computational Linguistics, Association for Computational Linguistics.

[10] N. Mulyar, W. M. van der Aalst, and M. Peleg. A Pattern-based Analysis of Clinical

Computer-interpretable Guideline Modeling Languages. *Journal of the American Medical Informatics Association*, 14(6):781 – 787, 2007.

[11] J. Pustejovsky, J. M. Castaño, R. Ingria, R. Saurí, R. J. Gaizauskas, A. Setzer, G. Katz, and D. R. Radev. TimeML: Robust Specification of Event and Temporal Expressions in Text. In *New Directions in Question Answering*, pages 28–34. AAAI Press, 2003.

[12] J. Pustejovsky, K. Lee, H. Bunt, and L. Romary. ISO-TimeML: An International Standard for Semantic Annotation. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta, May 2010. European Language Resources Association (ELRA).

[13] J. Pustejovsky and A. Stubbs. Increasing informativeness in temporal annotation. In *Proceedings of the Fifth Linguistic Annotation Workshop, LAW 2011*, pages 152–160. Association for Computational Linguistics, 2011.

[14] R. M. Reeves, F. R. Ong, M. E. Matheny, J. C. Denny, D. Aronsky, G. T. Gobbel, D. Montella, T. Speroff, and S. H. Brown. Detecting temporal expressions in medical narratives. *International Journal of Medical Informatics*, 82(2):118–127, 2013.

[15] Remine. Documentation of formalized guidelines, 2010.

[16] R. Saurí, J. Littman, R. Gaizauskas, A. Setzer, and J. Pustejovsky. TimeML Annotation Guidelines, Version 1.2.1, 2006.

[17] A. Seyfang, S. Miksch, M. Marcos, J. Wittenberg, C. Polo-Conde, and K. Rosenbrand. Bridging the gap between informal and formal guideline representations. In *European Conference on Artificial Intelligence (ECAI-2006)*, volume 141, pages 447–451, Riva del Garda, Italy, 2006. IOS Press.

[18] W. Sun, A. Rumshisky, and O. Uzuner. Annotating temporal information in clinical narratives. *Journal of Biomedical Informatics*, 46, Supplement(0):S5–S12, 2013. 2012 i2b2 NLP Challenge on Temporal Relations in Clinical Data.

[19] W. Sun, A. Rumshisky, and O. Uzuner. Temporal reasoning over clinical text: the state of the art. *JAMIA*, 20(5):814–819, Sept. 2013.

[20] P. Terenziani, E. German, and Y. Shahar. The temporal aspects of clinical guidelines. *Stud Health Technol Inform*, 139:81–100, 2008.

[21] C. Thorne, E. Cardillo, M. Montali, C. Eccher, and D. Calvanese. Process fragment recognition in clinical documents. In *Proceedings of the 13th Conference of the Italian Association for Artificial Intelligence (AI*IA 2013)*, 2013.

[22] N. UzZaman and J. Allen. TRIOS TimeBank Corpus: extended timebank corpus with help of deep understanding of text. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta, May 2010. European Language Resources Association (ELRA).

[23] M. Verhagen and J. Pustejovsky. Temporal Processing with the TARSQI Toolkit. In *COLING (Demos)*, pages 189–192, 2008.

[24] M. Verhagen and J. Pustejovsky. The TARSQI Toolkit. In *LREC*, pages 2043–2048. European Language Resources Association (ELRA), 2012.

[25] P. Votruba, S. Miksch, and R. Kosara. Tracing the Formalization Steps of Textual Guidelines. In *Computer-based Support for Clinical Guidelines and Protocols. Proceedings of the Symposium on Computerized Guidelines and Protocols (CGP 2004)*, volume 101 of *Studies in Health Technology and Informatics*, page 172–176. IOS Press, 2004.

[26] R. Wenzina and K. Kaiser. Identifying Condition-Action Sentences Using a Heuristic-based Information Extraction Method. In *Proceedings of the Joint International Workshop: KR4HC'13+ProHealth'13*, pages 17–29, 2013.

[27] R. Wenzina and K. Kaiser. Using TimeML to Support the Modeling of Computerized Clinical Guidelines. *Studies in Health Technology and Informatics*, 205:8–12, 2014.