# A Document Engineering Environment for Clinical Guidelines

Gersende Georg[1,2,3] and Marie-Christine Jaulent[1,2,3]
[1]INSERM U 872, Eq. 20, SPIM Paris, F-75006 France; [2]Université Paris, Descartes, UMR S 872, Paris, F-75006 France; [3]Centre de Recherche des Cordeliers, Université Pierre et Marie Curie - Paris 6, UMR S 872, Paris, F-75006 France
15, rue de l'Ecole de Médecine
75006 Paris, France
{Gersende.Georg,Marie-Christine.Jaulent}@spim.jussieu.fr

## ABSTRACT
In this paper, we present a document engineering environment for Clinical Guidelines (G-DEE), which are standardized medical documents developed to improve the quality of medical care. The computerization of Clinical Guidelines has attracted much interest in recent years, as it could support the knowledge-based process through which they are produced. Early work on guideline computerization has been based on document engineering techniques using mark-up languages to produce structured documents. We propose to extend the document-based approach by introducing some degree of automatic content processing, dedicated to the recognition of linguistic markers, signaling recommendations through the use of "deontic operators". Such operators are identified by shallow parsing using Finite-State Transition Networks, and are further used to automatically generate mark-up structuring the documents. We also show that several guidelines manipulation tasks can be formalized as XSL-based transformations of the original marked-up document. The automatic processing component, which underlies the marking-up process, has been evaluated using two complete clinical guidelines (corresponding to over 300 recommendations). As a result, precision of marker identification varied between 88 and 98% and recall between 81 and 99%.

## Categories and Subject Descriptors
J.3 [**Life and Medical Sciences**]: Medical information systems; I.7.2 [**Document and Text Processing**]: Document Preparation - *Markup languages* - *Hypertext/hypermedia*; I.2.7 [**Artificial Intelligence**]: Natural Language Processing - *Text analysis*

## General Terms
Algorithms, Documentation, Languages.

## Keywords
Clinical Guidelines, XML, deontic operators, GEM.



## 1. INTRODUCTION
Document processing is of particular importance in several areas of medical information systems, from patient records [17] to clinical guidelines [14]. Clinical guidelines are medical documents that contain best practice recommendations aimed at doctors, based on the concept of *Evidence-Based Medicine* [12]. These are complex documents which require significant amounts of specialized knowledge for their production. There is a growing interest in their computerization which should facilitate both their production, their standardization and their dissemination.

One of the early approaches to guideline computerization, referred to as the "document-based" approach, adopted traditional methods of document engineering by developing specific XML encoding model [2] such as the Guideline Elements Model (GEM) [13]. GEM is an XML framework based on a hierarchy of concepts describing the guidelines' contents, information, as well as meta-information for their use (such as guidelines objectives, intended audience and authors). Each GEM element corresponds to specific labels, some of which are normalized through a controlled medical vocabulary (for instance, the one defined by the National Guidelines Clearinghouse)[1]. The GEM framework consists of structuring the guideline document simply by using the set of XML mark-ups. This can however be a complex process if performed manually, as it requires an in-depth analysis of the guideline contents and, simultaneously, a constant reference to the GEM framework. A pilot study has shown that the complexity of manual analysis can affect the quality of GEM-based document encoding. [8]. To tackle these problems and support the process of manual marking-up of guidelines documents, the GEM-Cutter [14], application has been introduced. It is essentially an XML editor developed to facilitate the marking-up of textual guidelines. GEM-Cutter decreases the cognitive load of the user by offering on-line information on GEM categories and supporting an incremental process of document marking-up.

Another medical document engineering approach has been introduced by Svatek et al. [16] who developed the "Stepper" system, which supports the encoding of clinical guidelines through the marking-up of text based on a stepwise formalization process. The purpose of Stepper is to produce knowledge representations from text [15].

[1] http://www.guideline.gov/

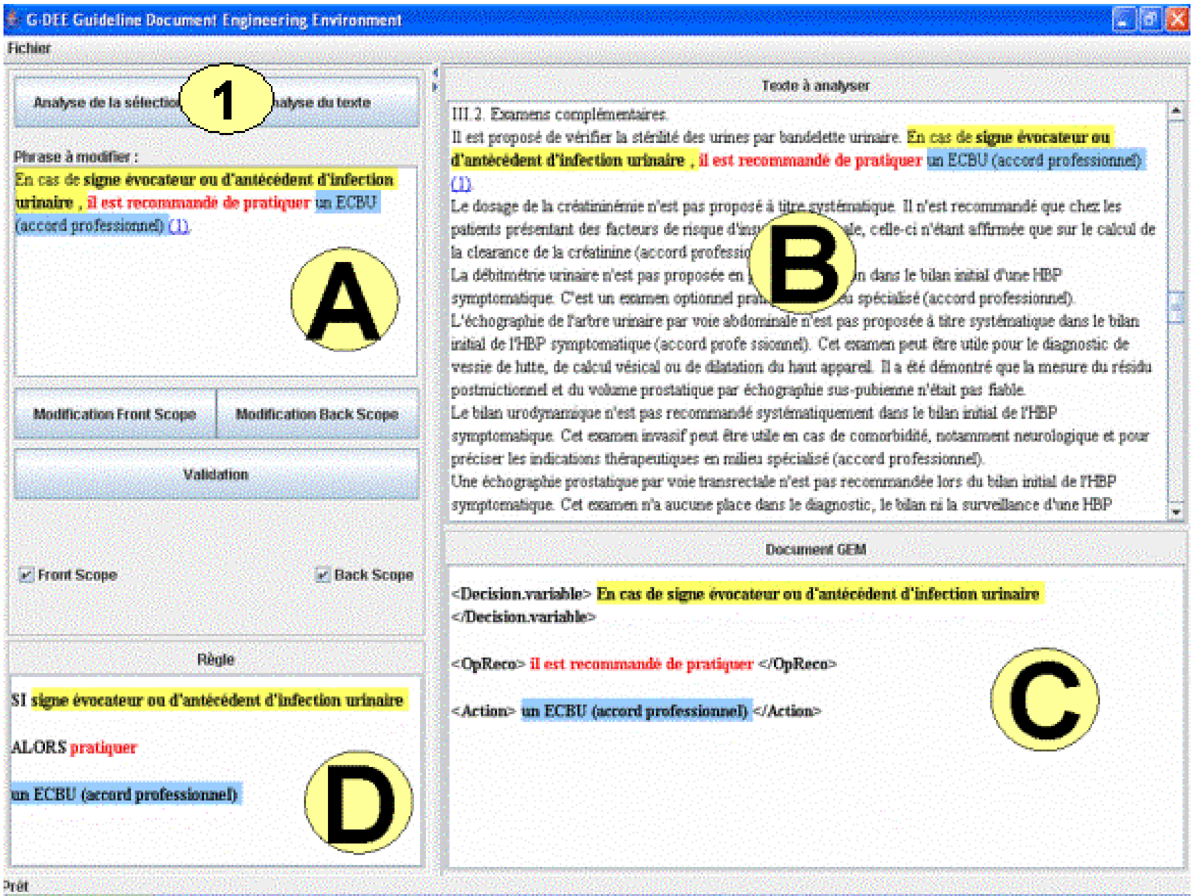


**Figure 1. An overview of the G-DEE environment.**

It is also supported by an interface similar to GEM-Cutter, whose purpose is to "minimize information loss during the encoding process". However, both Stepper and GEM-Cutter interfaces still rely on an entirely manual encoding process; they do not provide tools to assist such encoding on a content basis.

Several studies in document engineering have described applications for assisting document authoring, its structuring and the marking-up of specific information. The Multilingual Document Authoring (MDA) [3] enables to control biological experiment reports for the production of multilingual documents syntactically, stylistically and semantically. The Universal Parsing Agent (UPA) [19] used the GATE architecture (incorporating the Connexor™ dependency parser) to extract important information from documents, and enhance text with semantic tags. Previous research focused on the use of syntactic information to improve the performance of Information Retrieval systems, and analyzed the performance of different approaches for managing the syntactic variation of texts. Vilares et al. [18] used shallow parsing to identify word pairs related through the most significative syntactic dependencies such as noun-modifier or subject-verb. Focusing on the type of information contained in texts, previous research has introduced an approach to characterize three principal types of biomedical papers: reviews, research and clinical papers. Zerida et al. [20] defined a set of descriptors based on two concepts: rank in the hierarchy and salience, these descriptors being identified through linguistic markers. Bayerl et al. [1] have investigated the type of semantic information associated with mark-ups in scientific articles. They introduced XML mark-ups representing two different semantic levels: the thematic level (i.e. topics mentioned in the article) and the functional or rhetorical level.

In this paper, we present the G-DEE environment (for *Guidelines Document Engineering Environment*) dedicated to the study of

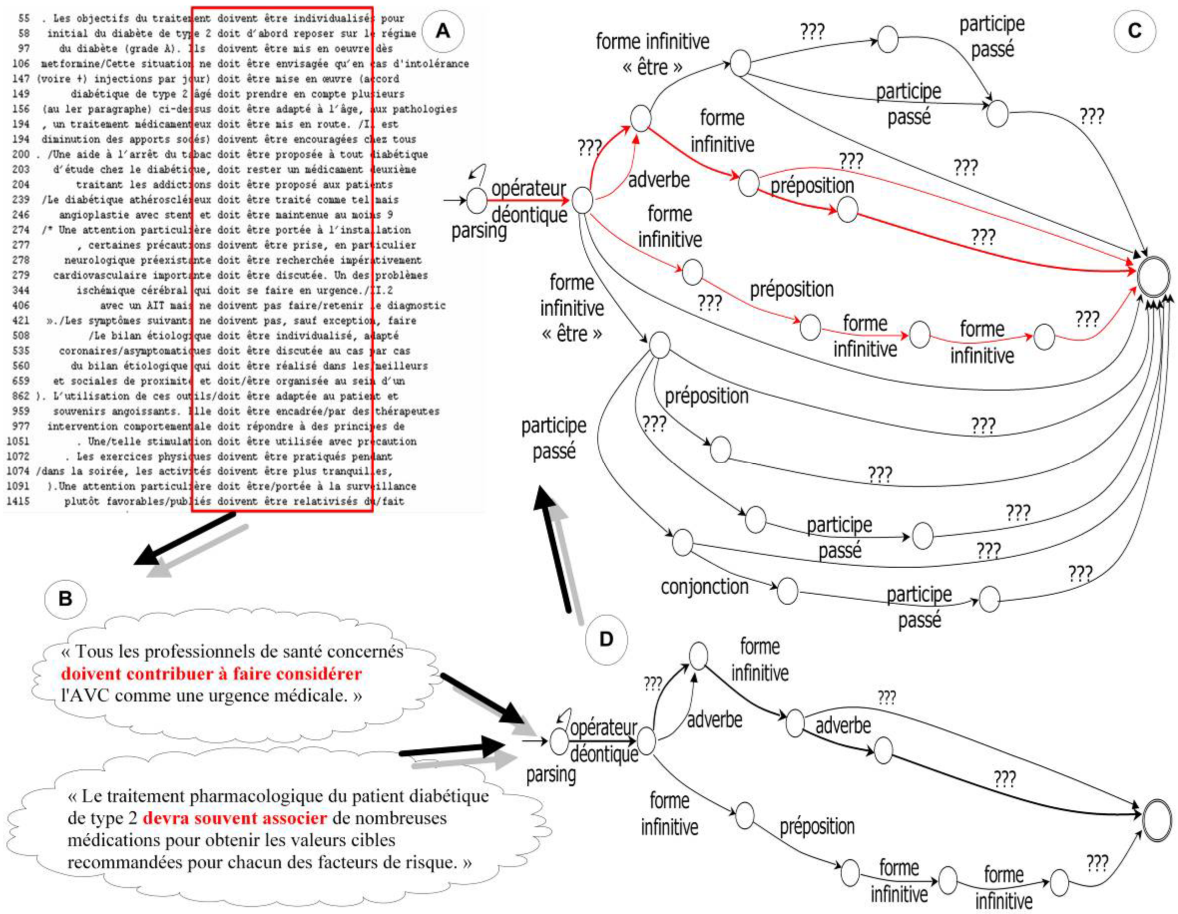


**Figure 2. Deriving a FSA for the recognition of the "should" deontic operator from a study of occurrences in context.**

clinical guidelines, which incorporates text processing functions to support encoding through a first level of automatic structuring. This document engineering environment also supports different transformations of the encoded document, based on XSL style sheets, to extract and visualize specific knowledge.

This article is organized as follows: the next section introduces our approach, which is based on the automatic recognition of specific linguistic markers. The 'System overview' section describes the G-DEE document engineering platform that incorporates shallow parsing techniques. Automatic content processing is then described in three parts: the recognition of deontic operators, the identification of operators' scopes and the processing of conditional connectors. We present details of XSLT transformations integrated into G-DEE to support document visualization and presentation. Finally, we give a preliminary evaluation of system performance.

## 2. APPROACH

Central to our approach is the fact that automatic content processing should support document structuring by generating mark-ups each time specific linguistic markers are recognized (for instance linguistic markers signaling recommendations, see below). Recommendations are the essence knowledge of the clinical guidelines [13] and are taken into account to elaborate knowledge bases and decision support systems. Clinical guidelines belong to the generic category of normative texts, to which much research has been dedicated. For instance, Moulin and Rousseau [9] have described a method to automatically extract knowledge from legal texts based on the hypothesis that these texts are naturally structured through the occurrence of specific linguistic expressions, known as "*deontic operators*" [9]. These operators manifest themselves through such verbs as "*pouvoir*" ("to be allowed to or may"), "*devoir*" ("should or ought to"), "*interdire*" ("to forbid"). These verbs correspond to traditional deontic modalities: permission, obligation and prohibition, which have been found by Kalinowski [7] to be the most characteristic linguistic structures of normative texts. Because clinical guidelines can also be categorized as normative texts, we have adapted Moulin and Rousseau's approach to the context of clinical guidelines, by identifying equivalent deontic elements specific to clinical recommendations. We carried out a lexicometric analysis on a corpus of 20 clinical guidelines (in French) published by the French National Authority for Health[2]. We first studied the frequency of deontic verbs for the set of 20 clinical guidelines collected (composed of 83 997 word occurrences). We used the statistical text analysis software

[2] http://www.has-sante.fr

Tropes™ [3] to analyze these documents, particularly words occurrences and lemmatized verbs. We considered "to recommend" (*recommander*) as the reference verb for the deontic modality, due to the fact that in medical texts it always expresses recommendations. We studied the lexical context of each verb of the corpus and identified those which are similar to the reference verb in terms of distribution. We also investigated the distribution of deontic operators throughout the text. Clinical guidelines being a set of structured recommendations, one would expect deontic operators to be distributed in a way which is consistent with these documents' style. By analyzing the distribution of the principal verbs constitutive of deontic operators (i.e. "*recommander*" (to recommend)) in each guideline, we obtained several distribution pattern. All these patterns share two common features. The first one is the scope of distribution, which spans across the entire text. The second one is the recurrence of groupings of deontic verbs. The latter finding is an indicator of textual structure, namely the repetition of deontic operators within specific sections.

While free text understanding is beyond the state-of-the-art of Natural Language Processing (NLP), it is possible to use shallow NLP techniques (such as Finite-State Automata (FSA) [11]) to recognize specific expressions. These techniques will specifically target the recognition of appropriate markers of textual structure, relieving the user from the early steps of document structure recognition (such as the identification of specific markers). We have developed an ad hoc parsing technology based on FSA which parses the document and generates mark-ups corresponding to deontic operators and their scopes [5]. In terms of document structure, the text segments structured by deontic expressions are called scopes. A scope that precedes a deontic operator is called front-scope, whereas the back-scope corresponds to a scope which follows the operator [9]. The marked-up document can subsequently be the object of various XSL-based transformations.

## 3. SYSTEM OVERVIEW

An overview of the G-DEE interface is presented in Figure 1. This environment supports various document processing features, some dealing with specific text display, and others triggering text analysis functions to extract knowledge or information (e.g. in a rule-based format). The interface supports the selective processing of text fragments, which are analyzed for deontic operators (Figure 1, interface button 1) and marked-up accordingly (Figure 1, window B). In the figure's example the sentence selected is: "*En cas de signe évocateur ou d'antécédent d'infection urinaire, il est recommandé de pratiquer un ECBU.*" (In case of symptoms of urinary tract infection, it is recommended to perform urinalysis.). The resulting marking-up can be validated interactively by the user (button 1 of the interface). In addition, G-DEE enables to automatically display contents of specific GEM elements, as well as deontic operators in window C. Window D displays decision rules automatically derived from the marked-up text, which can be used for knowledge extraction or analysis of text coherence. In a similar approach to that of Amaya Web Editor, for which generic and specialized views of documents are integrated [10], our XSL transformations are tuned to the display of structured document, and to the selective visualization of document information.

[3] http://www.acetic.fr

## 4. AUTOMATIC CONTENT PROCESSING

Text processing is based on a cascade of FSA (see figure above), and uses a customized parsing algorithm that we have developed, in particular for the efficient handling of shared patterns between FSA. The set of FSA constituting the grammar of our analyzer has been derived from the manual analysis of a corpus of medical texts, which includes 9 consensus conferences, 6 chapters from course material and 7 clinical guidelines (in the field of diabetes, hypertension, asthma, dyslipidemias, epilepsy, renal disease). This corpus contains a sufficient number of occurrences of deontic operators in different contexts to include a large number of syntactic variants ensuring sufficient coverage of the grammar drawn from it.

*Deontic Operators in Context*

To that effect, we used the "*Simple Concordance Program (release 4.08)*[4]" to analyze our corpus. This program provides context of the occurrences (part A – Figure 2) that can be analyzed to formalize syntactic construct as a FSA (part D). For example "*doivent contribuer à faire considerer*" (should contribute to have considered) or "*devra souvent associer*" (should often associate) are occurrences of the deontic verb "devoir" (should). The set of individual automata (part D) is aggregated to obtain generic automata (part C), which integrates the entire set of patterns built around the verb "should". The example automaton described on the Figure 2 - D recognizes the deontic operator "should" (present indicative), followed by the infinitive verb (contribute), a preposition (to) and two infinitive verbs (*faire considerer* "to have considered"). It can be noted that we do not use a separate POS tagger due to the fact that morphological as well as syntactic information is included in the terminal nodes of the FSA variants.

We also defined monadic and dyadic forms (following Moulin and Rousseau [9]). For instance, monadic deontic operators refer to pronominal forms in which the subject is not identified "*il est interdit de, il est possible de*" (it is forbidden, it is possible to), while dyadic deontic operators "*doit, peut, ne doit pas*" (should, can, not should) include their subject. Deontic monadic and dyadic forms can be used to characterize the role of front-scope and back-scope in terms of knowledge representation, in particular when associated to conditional markers.

We also consider the active or passive voice for a deontic operator, which has similar implication to the operator's arity (monadic or dyadic) in terms of knowledge representation. The active or passive voice of deontic operators associated to the monadic or dyadic form may for instance play a role in structuring the text when extracting an IF-THEN decision rule. The recognition of passive or active voice mainly aims at properly locating actions elements within recommendations. We studied connectors occurrence frequency (obtained with Tropes™) in a corpus composed of around 20 000 sentences extracted from clinical guidelines (Table 1).

[4] http://www.textworld.com/scp/

Table 1. Connectors distribution in Clinical Guidelines.

| Connectors | Percent of distribution |
|---|---|
| Purpose | 1% |
| Time | 3% |
| Comparison | 4% |
| Condition | 7% |
| Opposition | 8% |
| Cause | 21% |
| Disjunction | 21% |
| Conjunction | 52% |

Because the conjunction and disjunction connectors are not significant, we defined a set of additional automata to recognize conditions connectors (7%) and time connectors (3%), which both identify conditions within recommendations. The use of conditional markers will be further described in section 4.3.

## 4.1 Document Structuration through Recognition of Deontic Operators

When using automatic content processing functions to structure documents, each sentence in the document is parsed sequentially for the occurrence of deontic operators. Parsing is a two-step process comprising (i) FSA selection, (ii) actual sentence parsing and FSA instantiation.

A pre-processing step is used to identify whether the sentence contains more than one deontic operator. In this case, the pre-processing module segments the sentence based on punctuation rules [4] that supporting the recognition of different occurrences of deontic verbs. For example, "*En cas de ganglion pédiculaire envahi, si la résécabilité est de classe I, la chirurgie avec curage* ***ne peut être contre-indiquée****, mais cette décision* ***doit néanmoins s'intégrer*** *dans une approche multidisciplinaire.*" (In case of extension to pedicle lymph nodes, if surgical accessibility falls into Class I, surgery **cannot be contraindicated,** but this decision **should nevertheless be part of** a multi-disciplinary consultation). The first step determines which FSA should be selected for activation. Its aim is to handle conflicts between several syntactic patterns, due to shared sub-patterns, and to select the most relevant pattern on heuristic grounds (i.e. maximum number of matching tokens). It reduces the number of applicable FSA and ensures that most specific FSA are used in the first instance.

Parsing itself proceeds through a standard algorithm for FSA instantiation, which matches FSA categories to word occurrences. Its characteristic is to recognize specific lexico-syntactic patterns within longer segments of texts. This is why our FSA often include tokens corresponding to intervening sequences (some of which may be limited in length corresponding to "search windows").

FSA are stored in text files under the form of patterns of syntactic categories. Let us consider the following pattern [[aux_plur][pp_plur adv verbe_inf]]. Its first element "aux_plur" represents the plural auxiliary verbs, which can be matched to occurrences such as "are" or "will often have" (Figure 4). For the following sentence no word match is encountered for "aux_plur" and parsing with the above pattern exits on failure: "*La radiothérapie n'est pas non plus recommandée chez les sujets de moins de 60 ans, comme traitement des CBC sclérodermiformes, sur certaines zones (oreilles, mains, pieds, jambes, organes génitaux)*" (The radiotherapy is not either recommended to patients less than 60 years, like treatment of the sclerodermiforms CBC, on certain zones (ears, hands, feet, legs, genitals organs).).

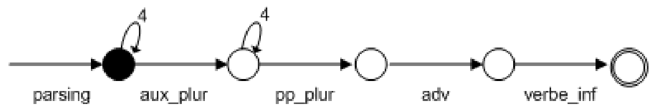


Figure 4. Parsing deontic expression with a FSA.

Another compatible pattern is thus selected, in this case "[[neg][negp pp_sing]]". In a similar way, the processor scans the sentence for a word matching the "neg" element (that corresponds to occurrences of negation, i.e. "is not" or "are not"). This is a successful match (Figure 5) and the parser analyses the remainder of the sentence "*pas non plus recommandée chez les sujets de moins de 60 ans, comme traitement des CBC sclérodermiformes, sur certaines zones (oreilles, mains, pieds, jambes, organes génitaux).*".

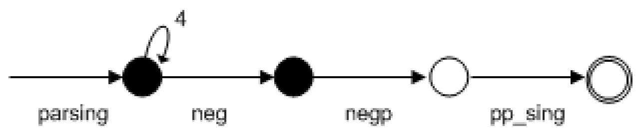


Figure 5. Progressive instantiation of a FSA during parsing of a deontic expression.

The following element (negation "*ne*" (english not)) is recognized in the sentence (Figure 6), and the processor analyzes the rest of the sentence.

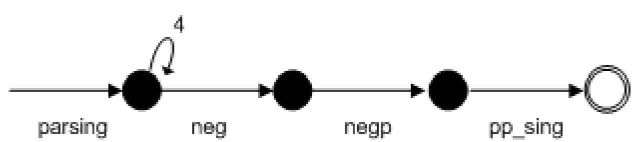


Figure 6. Successful instantiation of a FSA during parsing.

Patterns allow for intervening sequences in the expressions to be recognized, enabling the correct recognition of an operator despite the occurrence of adverbial locutions. In the above example, the blank between two brackets "][" actually corresponds to an analysis window of 4 words. The next category in the FSA is "pp_sing" (corresponding to a past participle, for example "recommended") and the remainder of the sentence is tested for that category, by considering each verb of the grammar's terminal vocabulary (corresponding to deontic verbs like "to advise" or "to recommend"). In the example considered, the verb "*recommander*" (to recommend) is recognized, and this successfully completes the instantiation of the pattern (Figure 6).

## 4.2 Recognition of Operators's Scopes: Structuring Sentences and Paragraphs

After the text has been tagged for deontic operators, a second step uses a specialized FSTN to properly delimitate (and mark-up) the corresponding scopes of the deontic operator using previously recorded information about the operator's voice, as shown below.

"***<Front-Scope>*** *La radiothérapie* ***</Front-Scope> <OpReco> n'est pas non plus recommandée </OpReco> <Back-Scope>*** *chez les sujets de moins de 60 ans, comme traitement des CBC sclérodermiformes, sur certaines zones (oreilles, mains, pieds, jambes, organes génitaux)* ***</Back-Scope>***."

Example in English: ***<Front-Scope>*** The radiotherapy ***</Front-Scope> <OpReco>*** is not either recommended to ***</OpReco> <Back-Scope>*** patients less than 60 years, like treatment of the sclerodermiforms CBC, on certain zones (ears, hands, feet, legs, genitals organs).) ***</Back-Scope>***.

## 4.3 Conditional Connectors Marking-up

Besides deontic operators, which signal recommendations, another type of linguistic marker plays an important role in structuring a document, namely the conditional. In the following example, the conditional "*si*" (if) introduces a condition whose identification is an important step for different tasks of information and knowledge extraction from text.

From our perspective of content-based structuring, this leads to a further structuring of the *front-scope*, using two new tags *<cond>* (for the marker) and *<condition>* (for the conditional proposition). This can be illustrated by the following marked-up example.

"***<Front-Scope> <cond>*** Si ***</cond> <condition>*** le diabète est diagnostiqué chez un patient âgé ***</condition> <SubScope> , </SubScope> <ScopeSec>*** un objectif de HbA1c comprise entre 6,5% et 8,5% ***</ScopeSec> </Front-Scope> <OpReco>*** **peut servir** ***</OpReco> <Back-Scope>*** de référence mais il est essentiel d'individualiser cet objectif en fonction du contexte médical et social (accord professionnel) ***</Back-Scope>***."

Example in English: ***<Front-Scope> <cond>*** If ***</cond> <condition>*** the diabetes is diagnosed in an elderly patient, ***</condition> <SubScope> , </SubScope> <ScopeSec>*** an objective of HbA1c ranging between 6.5% and 8.5% ***</ScopeSec> </Front-Scope> <OpReco>*** **can be used** ***</OpReco> <Back-Scope>*** as reference but it is essential to individualize this objective according to the medical and social context (professional agreement) ***</Back-Scope>***.

We introduced new tags, i.e. *<SubScope>* and *<ScopeSec>* to characterize each fragments of the sentence as well as the segments between punctuation sign and deontic operator tags.

# 5. XSL TRANSFORMATIONS

Content-based document structuring serves as a starting point for further transformations using more traditional techniques of document engineering. However, these two aspects are connected by the definition of XSL-based transformations relying on the mark-up categories defined for the first step of content-based document structuring. We describe in this section different XSL transformations that support visualizations of document structure, as well as the selective extraction and visualization of information (Figure 7). We present in this figure the process of the following recommendation marking-up: "*En cas de signe évocateur ou d'antécédent d'infection urinaire, il est recommandé de pratiquer un ECBU (accord professionnel).*" (In case of symptoms of urinary tract infection, it is recommended to perform urinalysis (professional agreement).).We developed an XML environment integrating a XSLT processor which controls these different transformations. We used this technique to define a set of XSL style sheets, based on the mark-up categories we defined for linguistic markers, from deontic operators to connectors and conditional markers (Figure 8). These determine the selective visualization or the extraction of specific textual information structured by these linguistic markers.

## 5.1 Document Structure and Visualization

The original functionality of G-DEE consists of structuring the text around linguistic markers, in particular recommendations. It is thus natural that the first use of encoding would be to visualize document structure so as to support its consultation, or its analysis, by users. To that effect, we have defined an XSL style sheet that automatically highlights recommendations. In this style sheet, we also specify the file structure after the XSLT processor is executed, i.e. HTML files, which are the format supported by the G-DEE visualization interface.

We defined an XSL style sheet that contains the rules executed by the XSLT processor. We defined three actions types: *(i)* the *<front-scope>* tags are interpreted by highlighting the text in yellow (attributes "*background-color:#FFFF66*"); *(ii)* the *<back-scope>* by highlighting the text in blue; *(iii)* and *<OpReco>* in red (for the deontic operator). As a result the textual elements corresponding to recommendations, and thus of particular significance can be visualized in the document, while immediately giving access to their different components: the type of deontic operator, which may be interpreted in terms of the "strength" of the recommendation as well as its conditions and actions. It is this type of visualization that will be used for the expert evaluation presented below.

The XSL style sheet header explicitly identifies the set of variables used in structuring the document (such as condition in front-scope, and back-scope in Figure 9)). This structure can be used to extract specific information of the marked-up text, and apply them a specific layout, for example the use of bold fonts to emphasize conditional element in a sentence.

Figure 8 illustrates the process of extracting specific information. The system tests for the occurrence of a *<FrontScope>* tag and, if successful, tests whether a condition marker is included using the *<xsl:when test= …>* tag. In this case, it will display a different layout, i.e. the condition will be emphasized using bold fonts in addition to the yellow highlighting.

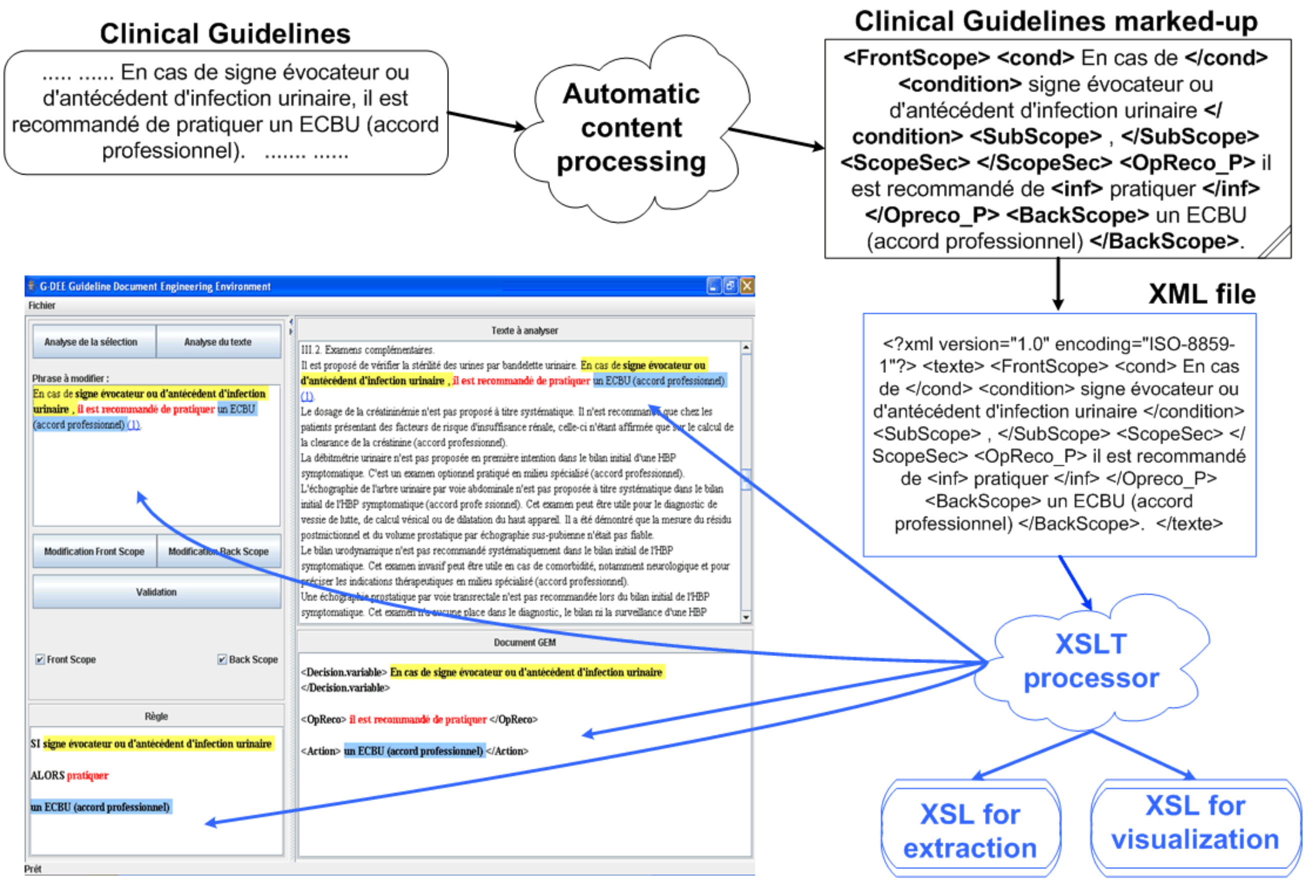


**Figure 7. Relation between XSL Transformations and the G-DEE visual interface.**

```
<xsl:template match="FrontScope">
<xsl:choose>
...
 <xsl:if test="condition">
  <xsl:if test="true()">
   <xsl:element name="font">
    <xsl:attribute name="style">background-color:#FFFF66;</xsl:attribute>
     <b><xsl:value-of select="condition"/></b>
    </xsl:element>
   </xsl:if>
  </xsl:if>
 ......
</xsl:template>
```

**Figure 8. Excerpt representing layouts for conditional elements.**

## 5.2 Visualizing GEM Encoding and Decision Rules

In this section, we illustrate the use of G-DEE to automatically assist the extraction of information relevant in a clinical context. This will be based on two examples. Firstly, the extraction of IF-THEN decision rules which are typical of medical knowledge-based systems; in that sense the environment can assist the process of knowledge acquisition from text. Secondly, in line with previous work in document engineering in Medicine, we show how G-DEE can support document encoding in the GEM format.

The first step is to define a style sheet containing the set of variables that characterize each recommendation's mark-ups (Figure 9).

```
<xsl:variable name="var_condition_fs">
        <xsl:value-of select="//FrontScope/condition"/>
</xsl:variable>
<xsl:variable name="var_cond_fs">
        <xsl:value-of select="//FrontScope/cond"/>
</xsl:variable>
....
```

**Figure 9. Example style sheet header describing variables.**

The textual segments that surround deontic operators (and from a content perspective, are structured by them), i.e. front-scope and back-scope contents, indicate the conditions and actions of a recommendation. We consider that each recommendation may be encoded as GEM elements [13] and/or represented as decision rules [6] depending on the application at hand. We defined a set of procedures to extract specific information from the content-based marked-up document in order to automatically generate these representations. These can be illustrated by considering different excerpts from clinical guidelines and their corresponding representations.

The first excerpt is dedicated to sentences in the passive voice, such as:

Example 1 (in French): "*La recherche d'une hypotension orthostatique (chute de la PAS de plus 20 mmHg et/ou de la PAD de plus de 10 mmHg, lors du passage en position debout),* ***est conseillée chez*** *tout hypertendu, en particulier chez le sujet de plus de 65 ans et le patient diabétique.*"

Example 1: "The search for an orthostatic hypotension (systolic pressure drops by more than 20 mmHg and diastolic pressure by more than 10 mmHg, when moving to a standing position), **is advised for** hypertensive patients, in particular for patients older than 65 years and diabetic patients."

The front-scope segments correspond to conditions part in this first example. The rules to identify which scopes actually correspond to decisions or actions (versus conditions) may be defined as:

- When a textual marker indicating a condition occurs in the front-scope, for instance "in case of", this textual segment corresponds to the condition part, while the contents of the back-scope correspond to the action.
- When a textual marker indicating a condition occurs in the back-scope, then the back-scope corresponds to the condition and the front-scope to the action.
- When no textual marker indicates a condition, the front-scope corresponds to the action, and the back-scope to the condition.

These rules are then incorporated into XSL style sheet and are associated to specific layouts. Figure 10 shows such a style sheet for visualizing GEM elements according to rules described above.

The second excerpt concern active voice sentences, such as:

Example 2 (in French): "*Le traitement pharmacologique du patient diabétique de type 2* ***devra souvent associer*** *de nombreuses médications pour obtenir les valeurs cibles recommandées pour chacun des facteurs de risque.*"

Example 2: "The pharmacological treatment of the diabetes 2 patient **will often have to associate** many medications to obtain the target values recommended for each risk factor."

In this example, the front-scope corresponds to conditions. We can define additional rules to extract text segments that correspond to decision and action elements:

- When a textual marker indicating a condition occurs in the front-scope, the front-scope contains the condition, and the back-scope the action.
- When a textual marker indicating a condition occurs in the back-scope, then the back-scope corresponds to the condition, and the front-scope to the action.
- When no textual marker indicates a condition, the front-scope corresponds to the condition and the back-scope to the action.

The presence of a condition marker within a scope (the verb taking part in the deontic operator being in the active or passive voice) implies that this scope contains a condition, and conversely the opposite scope (front/back) contains the action. For those sentences that do not contain such markers, we defined the following rule: "for an active voice sentence, the front-scope corresponds to the condition and the back-scope to the action, and conversely for the passive voice."

Figure 10 represents the transformation of the textual recommendation into GEM encoding elements (<*decision.variable*> and <*action*>) following rules that we described above to identify decision and action elements. This excerpt corresponds to rules to be activated when sentences contain a condition marker within the front-scope, in this case the GEM decision variable occurring in the front-scope and the action in the back-scope.

We proceeded in a similar way to extract decision and action elements for decision rules (Figure 1, window D). The XSL style sheet header is the same for variables definition. We then defined in a style sheet the different rules enabling to extract information corresponding to decision variables and actions elements.

```
<xsl:choose>
 <xsl:when test="$var_cond_fs">
  <xsl:if test="true()">
   <h3>&lt;Decision.variable&gt;
    <xsl:element name="font">
     <xsl:attribute name="style">background-color:#FFFF66;</xsl:attribute>
     <b><xsl:apply-templates select="$var_cond_fs"/></b>
     <b><xsl:apply-templates select="$var_condition_fs"/></b>
    </xsl:element>
    &lt;/Decision.variable&gt;
   </h3>
   <h3>&lt;OpReco&gt;
    <xsl:element name="font">
    <xsl:attribute name="style">color:#FF0000;</xsl:attribute>
    <b><xsl:apply-templates select="$var_op_reco"/></b>
    </xsl:element>
    &lt;/OpReco&gt;
   </h3>
```

**Figure 10. Visualizing condition in recommendations using the GEM format.**

We also integrated specific layouts to represent these decision rules in their respective G-DEE dedicated windows, i.e. IF *decision variables* THEN *action elements* (Figure 1 – window D).

These different transformations based on XSLT techniques can successfully structure clinical guidelines around recommendations and represent them in those formats typical of Medical Document Engineering (GEM) or Medical Knowledge-based systems (decision rules, albeit in textual format).

## 6. EVALUATION

In this preliminary evaluation, we do not consider usability aspects of the overall G-DEE visual interface, but limit ourselves to a performance analysis of the text processing tools that support content-based structuring. We tested the system on 276 sentences extracted from 5 randomly selected clinical guidelines. None of these clinical guidelines had been used for the definition of our deontic operators' grammar, which guarantees the validity of the test suite. For this evaluation, we mainly focused on the correct identification of the deontic expressions based on the following verbs: "*recommander*" ("to recommend"), "*devoir*" ("should or ought to"), "*pouvoir*" ("to be allowed to or may") and "*convenir*" ("to be appropriate") and their scopes. To evaluate system performance, we compared the system's output for automatic encoding to the manually encoded benchmark (an overview of such marking-up is shown in Figure 1). This benchmark contains 304 deontic operators, which had been previously identified manually together with their respective front-scope and back-scope.

As a preliminary result, our automatic structuring system correctly marked up 97% of the occurrences of deontic operators and their associated scopes on this test set.

To evaluate the global performance of G-DEE as a Document Engineering environment, we asked four medical experts involved in the development or evaluation of clinical guidelines to evaluate the marking-up generated by G-DEE on two entire guidelines, using a scoring sheet similar to these used in the evaluation of Information Extraction systems.

This evaluation compares the document structuring performed by G-DEE on an entire document to the spontaneous recognition of recommendations by experts in context. The work of each expert consists to check that each sentence is correctly marked-up and corresponds or not to an actual recommendation. The tables below represent the results obtained for the stroke (AVC) and hypertension (HTA) clinical guidelines. We observed that the percentage of sentences correctly structured by G-DEE varies significantly according to experts [81-99%] (Table 2).

**Table 2 – Evaluation results of G-DEE for AVC guidelines.**

| | Expert 1 | Expert 2 | Expert 3 |
|---|---|---|---|
| **Recall** | 0.99<br>(IC 95% :<br>0.96 ; 1) | 0.93<br>(IC 95% :<br>0.88 ; 0.98) | 0.81<br>(IC 95% :<br>0.74 ; 0.88) |
| **Specificity** | 0.89<br>(IC 95% :<br>0.83 ; 0.94) | 0,93<br>(IC 95% :<br>0.89 ; 0.98) | 0.82<br>(IC 95% :<br>0.76 ; 0.89) |
| **Precision** | 0.92 | 0.96 | 0.92 |
| **Noise** | 0.11 | 0.07 | 0.18 |
| **F-measure** | 0.95 | 0.94 | 0.86 |

We also analyzed sentences considered as false positives (FP) and false negatives (FN). We observed few FP [3 – 6], but we noticed disagreements between experts for FN [1 – 17] that correspond to recommendations that are not marked-up or incorrectly marked-up.

For the HTA clinical guideline, the percentage of sentences correctly marked-up as recommendations by G-DEE also varies between experts [84-96%] (Table 3). We also observed a few FP [2-11] for HTA clinical guidelines and disagreements between experts for FN [8 – 18].

It appears that the origin of certain disagreements between experts rests outside of the context of these experiments, in the authoring process of clinical guidelines. Rather than to propose strict norms for authoring, it seems better to explore the basis for disagreements and thus to propose specific authoring rules to the most frequent constructs considered as ambiguous.

**Table 3 – Evaluation results of G-DEE for HTA guidelines.**

| | Expert 1 | Expert 2 | Expert 3 | Expert 4 |
|---|---|---|---|---|
| **Recall** | 0,91<br>(IC 95% :<br>0,86 ; 0,95) | 0,86<br>(IC 95% :<br>0,80 ; 0,91) | 0,83<br>(IC 95% :<br>0,77 ; 0,89) | 0,85<br>(IC 95% :<br>0,80 ; 0,91) |
| **Specificity** | 0,84<br>(IC 95% :<br>0,78 ; 0,90) | 0,93<br>(IC 95% :<br>0,89 ; 0,97) | 0,96<br>(IC 95% :<br>0,93 ; 0,99) | 0,88<br>(IC 95% :<br>0,83 ; 0,93) |
| **Precision** | 0,88 | 0,95 | 0,98 | 0,92 |
| **Noise** | 0,16 | 0,07 | 0,04 | 0,12 |
| **F-measure** | 0,89 | 0,90 | 0,90 | 0,89 |

As a result of these experiments, we dissociated errors that concern G-DEE, more specifically due to syntactic coverage problems, and interpretation problems due to the quality of authoring. These experiments also highlighted several problems with the structure of clinical guidelines, which was precisely one of the objectives of this research. The experts have different point of views on what constitutes a recommendation, although those documents were already the result of a consensus within the working group in charge of authoring clinical guidelines.

## 7. CONCLUSIONS

Medical Informatics offers significant opportunities for the use of Document Engineering techniques due to the importance of document processing in clinical care, from patient records to clinical guidelines. The latter aspect focuses on medical texts that can be the object of several processes, such as knowledge diffusion, extraction and visualization. We have recently seen the mergence of specific Document Engineering research applied to Medicine, for instance through the GEM standard, which is one of the first approach proposing a XML model to structure clinical guidelines.

The extension we proposed consists in supporting document structuring using content-based automatic tools, while leaving the user in the loop. In that sense G-DEE as a Document Engineering environment assists the user in various consultation or analysis tasks, but is not meant to substitute itself to her for complete tasks such as GEM encoding of documents. Because the whole structuring process is performed from the automatic recognition of a limited number of linguistic markers, scalability of the approach would be achieved, within the limits of the state-of-the-art of document processing techniques. The current limitation of the approach lies in the syntactic coverage required to identify deontic operators. Although extensive coverage can be achieved from corpus analysis (because of the specific nature of deontic operators), occasionally new texts will introduce variants not previously encountered, which require extension of the grammar. A valuable extension of this approach would consist in further processing of the textual contents of a deontic operator's scopes, which would identify relevant content such as pharmacological treatments. Such processing can be based on terminological recognition or information extraction methods such as named entity recognition.

Finally, this semi-automatic approach to document structuring developed for clinical guidelines can potentially be applied to others types of normative texts which would share similar properties in terms of contents and life cycle.

## 8. ACKNOWLEDGMENTS

This work has been partly funded through a post-doctoral fellowship from "Region Ile-de-France".

## 9. REFERENCES


[1] Bayerl, P., Lüngen, H., Goecke, D., Witt, A. and Naber, D., Methods for the semantic analysis of document markup. in *Proceedings of the 2003 ACM Symposium on Document Engineering*, (Grenoble, France, 2003), ACM Press, New York, NY, 161-170.

[2] Bray, T., Paoli, J., Sperberg-McQueen, C., Maler, E. and Yergeau, F. Extensible Markup Language (XML) 1.0 (Third Edition). Recommendation, W.C. ed. *available at: http://www.w3.org/TR/REC-xml/*, 4 February 2004.

[3] Brun, C., Dymetman, M., Fanchon, E., Lhomme, S. and Pogodalla, S., Semantically-based text authoring and the concurrent documentation of experimental protocols. in *Proceedings of the 2003 ACM Symposium on Document Engineering*, (Grenoble, France, 2003), ACM Press, New York, NY, 193-202.

[4] Drillon, J. *Traité de la ponctuation française*, Paris (in French), 1991.

[5] Georg, G., Colombet, I. and Jaulent, M.-C., Structuring Clinical Guidelines through the Recognition of Deontic Operators. in *Proceedings of Medical Informatics Europe 2005*, (Geneva, Switzerland, 2005), IOS Press Amsterdam, 151-156.

[6] Georg, G. and Jaulent, M.-C., An Environment for Document Engineering of Clinical Guidelines. in *Proceedings of the American Medical Informatics Association*, (Washington, DC, 2005), 276-280.

[7] Kalinowski, G. *La Logique Déductive*. Presses Universitaires de France (in French), 1996.

[8] Karras, B., Nath, S. and Shiffman, R., A Preliminary Evaluation of Guideline Content Mark-up Using GEM--An XML Guideline Elements Model. in *Proceedings of the American Medical Informatics Association*, (2000), 413-417.

[9] Moulin, B. and Rousseau, D., Knowledge acquisition from prescriptive texts. in *Proceedings of the 3rd international conference on Industrial and engineering applications of artificial intelligence and expert systems*, (Charleston, South Carolina, United States, 1990), 1112-1121.

[10] Quint, V. and Vatton, I., Techniques for authoring complex XML documents. in *2004 ACM Symposium on Document Engineering*, (2004), ACM Press, 115–123.

[11] Roche, E. and Schabes, Y. *Finite-State Language Processing*. MIT Press, 1997.

[12] Sackett, D., Rosenberg, W., Gray, J., Haynes, R. and Richardson, W. Evidence-based medicine: what it is and what it isn't. *BMJ*, *312* (7023). 71-72, 1996.

[13] Shiffman, R., Karras, B., Agrawal, A., Chen, R., Marenco, L. and Nath, S. GEM: A proposal for a more comprehensive guideline document model using XML. *J Am Med Informatics Assoc*, *7*. 488-498, 2000.

[14] Shiffman, R., Michel, G. and Essaihi, A. Bridging the guideline implementation gap: a systematic approach to document-centered guideline implementation. *J Am Med Inform Assoc*, *11*. 418–426, 2004.

[15] Svatek, V., Kroupa, T. and Ruzicka, M., Guide-X – a Step-by-step, Markup-Based Approach to Guideline Formalisation. in *First European Workshop on Computer-based Support for Clinical Guidelines and Protocols*, (2001), 97-114.

[16] Svatek, V. and Ruzicka, M. Step-by-step mark-up of medical guideline documents. *International Journal of Medical Informatics*, *70* (2–3). 329–335, 2003.

[17] Takeda, H., Matsumura, Y., Kuwata, S., Nakano, H., Sakamoto, N. and Yamamoto, R. Architecture for networked electronic patient record systems. *International Journal of Medical Informatics*, *60*. 161–167, 2000.

[18] Vilares, J., Gómez-Rodríguez, C. and Alonso, M., Managing syntactic variation in text retrieval. in *Proceedings of the 2005 ACM Symposium on Document Engineering*, (Bristol, United Kingdom, 2005), ACM Press, New York, NY, 162-164.

[19] Whiting, M., Cowley, W., Cramer, N., Gibson, A., Hohimer, R., Scott, R. and Tratz, S., Enabling massive scale document transformation for the semantic web: the universal parsing agent. in *Proceedings of the 2005 ACM Symposium on Document Engineering*, (Bristol, United Kingdom, 2005), ACM Press, New York, NY, 23-25.

[20] Zerida, N., Lucas, N. and Crémilleux, B., Combining linguistic and structural descriptors for mining biomedical literature. in *Proceedings of the 2006 ACM Symposium on Document Engineering*, (Amsterdam, The Netherlands, 2006), ACM Press, New York, NY, 62-64.