The New Navigators: from Professionals to Patients
R. Baud et al. (Eds.)
IOS Press, 2003

Interpretative framework of chronic disease management to guide textual guideline GEM-encoding

Gersende Georg[a], Brigitte Séroussi[a], Jacques Bouaud[a]

[a]*Mission Recherche en Sciences et Technologies de l'Information Médicale, DSI, AP-HP, Paris, France*

Abstract

The aim of this work is to develop an XML-based application for the automated generation of decision rules from a textual guideline encoded using the Guideline Elements Model (GEM). A formalization of guideline-based chronological steps of treatment has been proposed to resolve the semantic ambiguities of the original document. The GEM DTD has been extended in order to standardize both decision variable and action representations in recommendations. Under these assumptions, the 1999 Canadian Recommendations for the management of hypertension have been marked-up as a GEM-encoded instance of the extended DTD. An XML parser has been used to extract the relevant elements as IF and THEN clauses of decision rules. This GEM application generated 104 rules to be compared to the 98 rules manually developed from the same guideline during the ASTI project.

Keywords:

clinical practice guidelines ; decision support systems ; knowledge modeling ; GEM.

1. Introduction

Clinical practice guidelines (CPGs) are originally textual documents usually structured as a set of clinical situations for which evidence-based therapeutic recommendations are provided. As the simple dissemination of guidelines has no impact on physician compliance [1], computer-based decision support systems (DSSs) embedding CPGs within their knowledge bases are currently developed and provide patient-based recommendations at the point of care. However, the formalization of CPGs expressed in natural language relies on a human interpretation step that may not capture all the nuances expressed in original documents. Textual guideline encoding is thus subject to variations according to the developer's experience, competence, and medical expertise [2]. The Guideline Elements Model (GEM) [3,4] has been proposed as a document-based model to structure guideline knowledge. In this paper, we present an experiment in using GEM as a preliminary step for the generation of a knowledge base made of decision rules. The domain of application is the management of hypertension. We first elaborated a formalization of the guideline content that clears up some of the semantic ambiguities on treatment chronology and then developed a normalized GEM-encoded instance that is automatically processed to generate decision rules.

2. Background

Translation of textual documents to any knowledge representation formalism is a complex task. Natural language expressiveness does indeed allow for contextual interpretation that formalization cannot afford. When developing a DSS knowledge base, the variability in interpreting guideline knowledge depends on the degree of prior experience and knowledge of the concerned medical domain a developer may have. A study using GLIF [2] showed that the representations encoded by different subjects with various types of computer science or clinical expertise were different both in content and structure. GEM [3] is intended to serve as a document model of CPGs. By describing concepts pertinent to guideline representation, attributes of these concepts and relationships among them, GEM aims at promoting translation of textual guidelines into a format that can be processed by computers [5,6,7]. For instance, an XML-based application that facilitates the automated generation of partially populated MLMs from GEM-encoded guidelines has been published [5]. However, GEM has several limitations. Although it has been found comprehensive enough to model the information content of CPGs, substantial variation is still observed in the creation of a GEM-encoded instance from a given CPG by different subjects [8]. GEM indeed does little to resolve ambiguities that are present in many textual guidelines: the model is simply an abstraction of the guideline document. Taking into account the ambiguities present in the Canadian CPGs for the management of hypertension [9], we propose a framework to represent chronological aspects of the therapeutic strategy for chronic diseases and a slight extension of GEM that enables a straight generation of executable decision rules.

3. Material

3.1. The 1999 Canadian recommendations for the management of hypertension

Like the ASTI project [10], we worked on the 1999 Canadian recommendations for the management of hypertension [9]. This guideline document is well structured in chapters that correspond to specific clinical situations. For instance, the case of diabetes as a complicating factor of hypertension is presented in figure 1. Within each chapter, an ordered sequence of therapeutic recommendations is proposed that needs to be interpreted.

VII Diabetes

Recommendations

1. Hypertension in people with diabetes (blood pressure greater than 140/90 mm Hg) should be treated to obtain target blood pressure lower than 130/80 mm Hg (grade C).
2. People with diabetes and hypertension with blood pressure of 130/80 to 139/89 mm Hg and target-organ damage should be treated to obtain a target blood pressure lower than 130/80 mm Hg (grade D).
3. For patients with diabetes who have hypertension without overt nephropathy and are under 60 years of age, preferred therapy is either an ACE inhibitor or a cardioselective β-adrenergic antagonist (grade A).
4. Second-line therapy includes low-dose thiazide diuretics (grade B), long-acting calcium channel blockers (grade B) and α-adrenergic antagonists (grade C). α-adrenergic antagonists and centrally acting antihypertensive agents should be used with caution in the presence of autonomic neuropathy (grade C).
5. Preferred therapy for patients with diabetes, hypertension and overt nephropathy (albuminuria greater than 300 mg/day) is an ACE inhibitor (grade A).
6. When an ACE inhibitor causes adverse effects, an angiotensin II receptor antagonist may be substituted (grade D).
7. Preferred therapy for patients with diabetes and isolated systolic hypertension who are over 60 years of age is either low-dose thiazide diuretics or long-acting dihydropyridine calcium-channel blockers (grade C).
8. If monotherapy with first-line agents is ineffective, contraindicated or associated with adverse side effects, the following should be considered:

i. A long-acting calcium-channel blocker may be combined with an ACE inhibitor (grade B). A low-dose thiazide diuretic may be added to an ACE inhibitor without adversely affecting microalbuminuria (grade B).
ii. For patients with renal insufficiency, a loop diuretic may be required to control volume and blood pressure (grade C).
iii. Indapamide may be substituted for low-dose thiazide as it may reduce microalbuminuria (grade C).

Figure 1 : Therapeutic recommendations for hypertensive patients with diabetes.

As it is usually the case, this textual CPG suffers from incompleteness (no recommendation for complex polypathological patient conditions). The terms used are imprecise ("with caution"), not defined ("autonomic neuropathy"), or vague ("causes adverse effects"). The chronological sequence of therapeutic recommendations is ambiguous starting in the case of diabetes (Figure 1) with "preferred therapy" (item 3), followed by "second-line therapy" (item 4), and then "preferred therapy" (item 5). The line of therapy corresponding to the 6th item is even not mentioned.

3.2. The GEM DTD

GEM is a guideline document model based on an XML DTD [3] that organizes the heterogeneous knowledge contained in practice guidelines. It is a multi-level hierarchy of more than 100 discrete elements structured in nine major branches. Among them, the *knowledge components* section represents the recommendation's logic and constitutes "the essence of practice guidelines". We only used the *conditional* element (cf. figure 3) that represents recommendations applicable only under specific circumstances. It is composed of different sub-elements among which only a few are actually used (*decision.variable, action, recommendation.strength, evidence.quality*).

4. Method

Following Tierney [11] who recommended that guideline developers structure recommendations as 'if-then-else' statements, the aim of this work is to automatically generate a set of decision rules from a GEM-encoded CPG. Decision rules are represented as IF-THEN-WITH statements where the IF-part corresponds to a set of *decision.variable* elements of the GEM DTD, the THEN-part corresponds to a set of *action* elements, and the WITH-part corresponds to the *id* of the *recommendation.strength* element. To enable the extraction of executable rules from a GEM-encoded instance, it was necessary to have the same structure for both decision variables and actions in the DTD. So we first extended the original DTD. In parallel, we proposed an interpretation framework to resolve the semantic ambiguities of the original guideline document with respect to the treatment strategies. Finally, a normalized GEM-encoded instance of the Canadian CPGs was developed (Figure2).

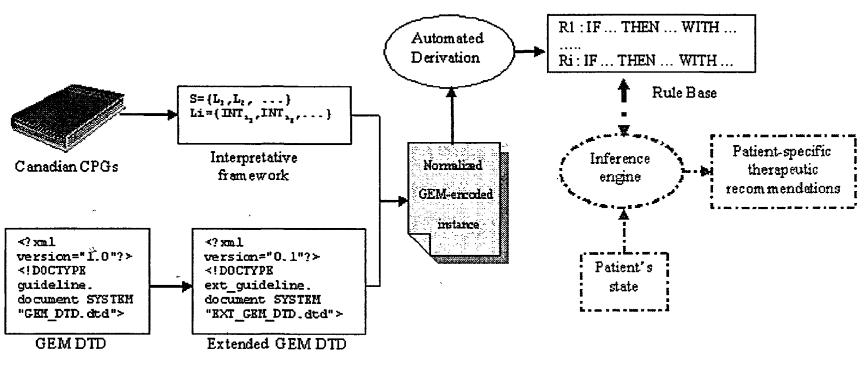

Figure 2 : Guideline knowledge processing from the textual document to the rule base.

4.1. Extension of the original GEM DTD

Conditional recommendations mainly rely on *decision.variable* and *action* elements. In the GEM DTD, decision variables are described by a value, a description, test parameters and a cost (Figure 3). Actions are described by a description, benefit and risk that can be gathered as sub-elements of an artificially constructed "action parameter" and a cost. Thus, to enable the automated derivation of rules, we needed a common data model for *decision.variable* and *action* elements. Then the *value* sub-element was added to the *action* element.

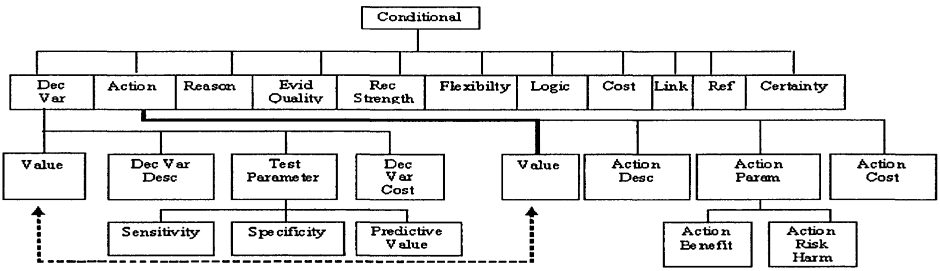

Figure 3 : Extension of the original GEM DTD.

4.2. Creation of a normalized GEM-encoded instance

The first step was to mark-up the textual guideline to match elements of the extended DTD. Then, we formalized the description of both patient clinical situation and treatments and we normalized the terms used in DTD elements.

4.2.1. Marking-up the Canadian CPGs

We marked-up the original document to identify which parts of the text were matching *decision.variable* and *action* elements. In the special case of the 3rd item of recommendations for diabetes, we linked the part of the sentence corresponding to the specification of the patient condition to *decision.variable* and what was related with the recommended therapy to *action* (Figure 4).

<decision.variable source= "explicit"> For patients with diabetes who have hypertension without overt nephropathy and are under 60 years of age *</decision.variable>*
<action source="explicit"> Preferred therapy is either an ACE inhibitor or a cardioselective β-adrenergic antagonist (grade A) *</action>*

Figure 4 : Mark-up of the text corresponding to the 3rd recommendation for diabetes.

4.2.2. Modeling and normalization of decision variables

We identified three classes of patient parameters, *i.e.* age, risk factors, and associated diseases represented by "state_patient.age", "state_patient.risk_factor" and "state_patient.pathology" in corresponding *ids* of decision variables. Previously marked-up textual portions identified as decision variables have been cut in sub-sentences corresponding to these classes. Normalized attribute *ids* have been finally introduced in each sub-element *value* of *decision.variable* elements.

<decision.variable source= "explicit" decision.variable.id="**state_patient.age**"> under 60 years of age
<value source= "implicit" value.id="**INF_60**"/></decision.variable>
<decision.variable source="explicit" decision.variable.id="**state_patient.pathology**"> hypertension
<value source= "implicit" value.id="**HT**"/></decision.variable>
<decision.variable source= "explicit" decision.variable.id="**state_patient.pathology**"> diabetes
<value source= "implicit" value.id="**DIA**"/></decision.variable>
<decision.variable source= "explicit" decision.variable.id="**state_patient.normality**"> no overt nephropathy
<value source= "implicit" value.id="**N_NEPH**"/></decision.variable>

Figure 5 : Normalization of notions expressed in decision variable values.

4.2.3. Interpretative framework of therapeutic lines and modeling of actions

In the follow-up of chronic diseases, therapeutic recommendations depend on the patient state and on his therapeutic history, *i.e.* prior prescriptions that either were not adequate or provided side effects. To resolve guideline semantic ambiguities in the presentation of the chronological steps of therapy, we proposed a framework formalizing the therapeutic strategy for a given patient profile.

A therapeutic strategy S is represented by an ordered sequence of therapeutic lines L_i, *i.e.*, $S=\{L_1, L_2, \ldots\}$. Each therapeutic line L_i is made of a set of treatments ordered according to therapeutic levels of intention $INT_{i,j}$, *i.e.* $L_i = \{INT_{i_1}, INT_{i_2}, \ldots\}$. According to a patient clinical situation and his response to the ongoing treatment, the recommended treatment may be either the next level of intention within the same therapeutic line or the first level of intention of the following therapeutic line. For instance with the 3rd recommendation for diabetes (Figure 1), "preferred therapy" is interpreted as the first level of intention of the first line of therapy. The type of therapy is specified (monotherapy) as well as the nature of the pharmacological drug class (ACE inhibitor). Another conditional recommendation is built with the same *decision.variable* element but the *treatment.nature* of the *action* element is instanciated by "β-adrenergic antagonist".

<action source="explicit" action.id="**treatment.line**"> first line treatment
<value source="implicit" value.id="**L1**"/></action>
<action source="explicit" action.id="**treatment.intention**"> first intention
<value source="implicit" value.id="**INT1**"/></action>
<action source="explicit" action.id="**treatment.type**"> monotherapy
<value source="implicit" value.id="**MONO**"/></action>
<action source="explicit" action.id="**treatment.nature**"> an ACE inhibitor
<value source="implicit" value.id="**ACE_IN**"/></action>

Figure 6 : Interpretative framework proposed for therapeutic lines.

The normalization phase, applied on decision variables, was also performed on actions leading to the instanciation of the *value* sub-element introduced in the extended DTD.

4.2.4. Rule extraction

Decision rules are represented as "IF *decision.variable* THEN *action* WITH *recommendation.strength*". Elements related to the *id* of corresponding values were extracted using an XML parser (SAX). For the previous example, we obtained the following rule:

IF	**THEN**	**WITH**
state_patient.age=INF_60	treatment.line=L1	recommendation.strength=A
and state_patient.pathology=DIA	and treatment.intention=INT1	
and state_patient.pathology=HT	and treatment.type=MONO	
and state_patient.normality=N_NEPH	and treatment.nature=ACE_IN	

5. Conclusion

GEM alone provides little help to resolve semantic ambiguities present in textual CPGs. So, we proposed an interpretative framework to model the ordered steps of therapy in the management of a chronic disease in therapeutic lines and levels of intention. A normalization of medical notions structuring the description of clinical conditions and actions has been performed. The automated process of the normalized GEM-encoded Canadian CPG on hypertension generated 104 completely instanciated decision rules. As compared to the 98 ASTI rules manually encoded by 2 physicians of the project on the basis of the same Canadian CPG, GEM-generated rules appear to cover a larger number of theoretical clinical situations. Closer to the original textual document, GEM-generated rules are also more specific, the number of decision variables in IF-parts is higher, in average, than in ASTI rules. On a first evaluation, GEM-generated knowledge base leads to clinical recommendations of a higher quality when processed on a sample of 10 patient cases. A further comparison of both rule bases is currently under progress.

References

[1] Matillon Y, Durieux P. *L'évaluation médicale, du concept à la pratique*. Medecine-Sciences, Flammarion, 2000;pp.43-54.

[2] Patel VL, Allen VG, Arocha JF, Shortliffe EH. Representing clinical guidelines in GLIF: individual and collaborative expertise. *J Am Med Inform Assoc* 1998;5(5):467-83.

[3] Shiffman RN, Karras BT, Agrawal A, Chen R, Marenco L, Nath S. GEM: a proposal for a more comprehensive guideline document model using XML. *J Am Med Inform Assoc* 2000;7(5):488-98.

[4] Karras BT, Nath SD,Shiffman RN. A preliminary evaluation of guideline content mark-up using GEM – An XML Guideline Elements model. *J Am Med Inform Assoc* 2000 ;7(Suppl):413-7.

[5] Agrawal A, Shiffman RN. Using GEM-encoded guidelines to generate Medical Logic Modules. *J Am Med Inform Assoc* 2001;8(Suppl):7-11.

[6] Shiffman RN, Agrawal A, Deshpande AM, Gershkovich P. An approach to guideline implementation with GEM. In: Patel VL, Rogers R, Haux R (eds). Medinfo 2001;84(1):271-5.

[7] Karras BT, Nath SD, Shiffman RN. A preliminary evaluation of guideline content mark-up using GEM – An XML Guideline Elements Model. *J Am Med Inform Assoc* 2000;7(Suppl):413-7.

[8] Gershkovich P, Shiffman RN. An implementation framework for GEM encoded guidelines. *J Am Med Inform Assoc* 2001;8(Suppl):204-8.

[9] Feldman RD, Campbell N, Larochelle P, Bolli P, Burgess ED, Carruthers SG, *et al.* Recommandations de 1999 pour le traitement de l'hypertension artérielle au Canada. *CMAJ* 1999;161(12):SF1-25. URL:http://www.cma.ca/cmaj/vol-161/issue-12/hypertension/hyper-f.htm

[10] Séroussi B, Bouaud J, Dréau H, Falcoff H, Riou C, Joubert M, Simon C, Simon G, Venot A. ASTI: A guideline-based drug-ordering system for primary care. In: Patel VL, Rogers R, Haux R (eds). Medinfo 2001;84(1):528-32.

[11] Tierney WM, Overhage JM, Takesue BY, Harris LE, Murray MD, Vargo DL, McDonald CJ. Computerizing guidelines to improve care and patient outcomes: the example of heart failure. *J Am Med Inform Assoc* 1995;2:316-22.

Address for correspondence

Gersende Georg
SIM, 91 Boulevard de l'Hôpital, 75634 Paris Cedex 13, France
gge@biomath.jussieu.fr