Does GEM-Encoding Clinical Practice Guidelines Improve the Quality of Knowledge Bases? A Study with the Rule-Based Formalism

Gersende Georg, Brigitte Séroussi, and Jacques Bouaud

Mission Recherche en Sciences et Technologies de l'Information Médicale, DPA / DSI / AP-HP, Paris, France & INSERM ERM 202, UFR Broussais - Hôtel-Dieu, Université Paris 6, Paris, France

The aim of this work was to determine whether the GEM-encoding step could improve the representation of clinical practice guidelines as formalized knowledge bases. We used the 1999 Canadian recommendations for the management of hypertension, chosen as the knowledge source in the ASTI project. We first clarified semantic ambiguities of therapeutic sequences recommended in the guideline by proposing an interpretative framework of therapeutic strategies. Then, after a formalization step to standardize the terms used to characterize clinical situations, we created the GEM-encoded instance of the guideline. We developed a module for the automatic derivation of a rule base, $BR_{\text{-}GEM}$, from the instance. $BR_{\text{-}GEM}$ was then compared to the rule base, $BR_{\text{-}ASTI}$, embedded within the critic mode of ASTI, and manually built by two physicians from the same Canadian guideline. As compared to $BR_{\text{-}ASTI}$, $BR_{\text{-}GEM}$ is more specific and covers more clinical situations. When evaluated on 10 patient cases, the GEM-based approach led to promising results.

INTRODUCTION

Clinical practice guidelines (CPGs) have been elaborated to reduce practice variations among physicians and thus improve the quality of care. They are originally textual documents usually structured as a set of specific clinical situations for which evidence-based therapies are recommended. As the simple dissemination of guidelines had no impact on physician compliance with recommendations,[1] guideline knowledge is currently embedded within knowledge bases (KBs) of computer-based decision support systems (DSSs) that provide patient-specific recommendations at the point-of-care.

Original CPGs are expressed in natural language and usually suffer from incompleteness, ambiguities and imprecision. These drawbacks result in interpretation variations of guideline content during the formalization step of CPGs prior to the development of KBs.

ASTI[2] ("Aide à la Stratégie Thérapeutique Informatisée") is a French project which aim is to develop a guideline-based DSS to be used in primary care. It has been first applied to the management of hypertension. The KB used in the critic mode is modeled as a set of production rules that has been manually encoded by two physicians from the 1999 Canadian recommendations for the management of hypertension.[3] We have used the Guideline Elements Model[4] (GEM), proposed as a document-based model, to develop a new rule base from the same CPG. The aim of our work is to compare GEM-based production rules to those manually encoded by physicians to check whether the GEM-encoding step has an impact on the quality of the rule base produced.

BACKGROUND

The translation of medical knowledge, originally expressed in textual CPGs to KBs is currently manually processed. Once formalized, guideline knowledge may be easily represented. A variety of representation models have been published to facilitate computer-based implementation of guideline knowledge. The oldest one, and the most widely used, is the Arden Syntax[5] in which Medical Logic Modules (MLMs) support clinical decision by the generation of alerts and reminders. More recently, the Guideline Interchange Format[6] (GLIF) proposes to model guideline content as a flowchart of structured steps representing clinical actions and decisions.

However, the formalization step relies on a human interpretation of the guideline which is subject to variations according to the developer's experience, competence, and medical expertise.[7] A study using GLIF showed that representations encoded by different subjects were different both in content and structure. Intended to serve as a document model of CPGs, GEM[4] proposed to make direct use of the guideline document structure to improve guideline content interpretation. By describing pertinent concepts 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. However, substantial variation is still observed in the

creation of a GEM-encoded instance from a given CPG by different subjects.[8]

Few works have been published to propose a methodology to formally compare KBs. KBs are often simply analyzed in terms of coverage, and level of specificity, *e.g.* quantitative information.[9] For instance, Del Fiol *et al.*[10] proposed an evaluation of two drug KBs developed in different academic medical centers. The same inference module was applied to the two KBs to check for drug interactions in a database of drug prescriptions.

The aim of our work is to measure the impact of GEM-encoding. We thus compare two KBs represented as production rules and built from the same guideline document, *e.g.* the 1999 Canadian recommendations for the management of hypertension.[3] The first KB has been classically manually encoded to be used within the critic mode of ASTI. The second KB has been automatically derived from the GEM-encoded instance of the guideline document.

MATERIAL

ASTI project

The ASTI[2] project aims at designing a guideline-based DSS to enable general practitioners to avoid prescription errors and to improve compliance with best therapeutic practices. The "critic mode" operates as a background process and corrects the physician's prescription on the basis of automatically triggered rules that account for isolated guideline recommendations. The KB is formalized as "IF-THEN" production rules, and has been manually built from the Canadian CPGs[3] by two physicians of the project. IF-parts of the rules represent clinical situations descriptions. They are composed of a set of inclusion criteria, *e.g.* patient state, pathology, and current therapy, and, exclusion criteria, *e.g.* pathologies that the patient is not suffering from, as well as the current therapeutic level of intention, *e.g.* the rank of the current treatment step in the therapeutic strategy. THEN-parts correspond to the set of recommended actions and include the grade of the recommendation.

1999 Canadian recommendations for the management of hypertension

The 1999 Canadian recommendations for the management of hypertension[3] is the guideline chosen by the ASTI project as the knowledge source for the development of KBs. It is a textual guideline document, well structured in chapters that correspond to specific clinical situations for which an ordered sequence of therapeutic recommendations is proposed. As it is usually the case, the guideline suffers from incompleteness, *e.g.* no recommendation for complex poly-pathological patient conditions, and ambiguities, *e.g.* the terms used are imprecise or not defined, the chronological sequence of therapeutic recommendations is unclear.

GEM DTD

GEM is a guideline document model based on an XML DTD[4] that organizes the heterogeneous guideline knowledge according to a multi-level hierarchy of more than 100 discrete elements structured in nine major branches. Among them, the *knowledge components* element include *recommendation* (which in turn comprises *conditional* and *imperative*), *definition*, and *algorithm* elements. We only used the *conditional* element that represents recommendations applicable only under specific circumstances. It is composed of different sub-elements among which only few are actually used (*decision.variable*, *action*, *recommendation.strength*).

METHOD

Our approach is based on the derivation of production rules represented as "IF-THEN-WITH" statements. We first created a normalized GEM-encoded instance of Canadian CPGs. Then, we developed a module to automatically extract decision rules from the GEM-encoded instance. We compared the resulting GEM-based rule base to the one manually built by two physicians of the ASTI project according to two criteria: *(i)* descriptive, *e.g.* quantitative and qualitative evaluation of production rules, and *(ii)* operational, *e.g.* comparison of therapeutic recommendations proposed by both approaches on a sample of 10 cases.

Creation of the GEM-encoded instance

To facilitate the automated extraction of production rules, we first extended the GEM DTD to have a similar XML structure for *decision.variable* and *action* elements (figure 1). As a *value* sub-element is defined for *decision.variable* elements, we added a *value* sub-element to *action* elements.

```
<!ELEMENT decision.variable (#PCDATA | value | decision.variable.description |
test.parameter | decision.variable.cost | %block;)*>

<!ELEMENT action (#PCDATA | value | action.benefit | action.risk.harm |
action.description | action.cost | %block;)*>
```

Fig. 1: Extended GEM DTD with the *value* sub-element for the *action* element.

Then, we marked-up the original document to identify which parts of the guideline were matching *decision.variable*, *action*, and *recommendation. strength* elements. We performed a normalization step to standardize attribute *ids* introduced in each

sub-element *value* of *decision.variable* elements to describe patient clinical situations. A similar normalization process was performed to resolve guideline semantic ambiguities in the representation of the chronological steps of therapeutic strategies. We proposed a framework formalizing the therapeutic strategy S recommended in the guideline.[11] S is represented by an ordered sequence of therapeutic lines L_i, *e.g.* S={L_1, L_2, …}. Each therapeutic line L_i is made of a set of treatments ordered according to therapeutic levels of intention INT_{i_j}, *e.g.* L_i = {INT_{i_1}, INT_{i_2}, …}. According to a patient clinical situation and her response to the ongoing treatment, the new 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.

Derivation of production rules from the GEM-encoded instance

The construction of the rule base relies on the identification of *decision.variable, action*, and *recommendation.strength* elements from the GEM-encoded instance. The aim is to locate and extract the contents of these different elements to generate rules in the following format:

" IF *decision.variable* THEN *action*
WITH *recommendation.strength* "

The IF-part corresponds to the set of *decision.variable*, the THEN-part to the set of *action* elements, and the WITH-part to the *id* of the *recommendation.strength* element. We used the XML parser SAX[12] to extract elements related to the *id* of corresponding *values* from the GEM-encoded instance.

Comparison of rule bases

To compare the rule base derived from the GEM-encoded instance, denoted $BR_{\text{-GEM}}$, to the one manually built in the ASTI project, denoted $BR_{\text{-ASTI}}$, we have used both descriptive and operational criteria.

On the descriptive side, we compared both rule bases on a quantitative basis, *i.e.* the number of rules, the number of premises in IF-parts, and the number of actions in THEN-parts. A qualitative evaluation allowed to analyze both KBs in terms of coverage, *i.e.* the number of clinical situations which are taken into account by the two rule bases.

On the operational side, we first developed a simple inference engine working in forward chaining to exploit $BR_{\text{-GEM}}$. Then, the resulting GEM-based system and the critic mode of ASTI have been compared on the basis of the treatments recommended by both approaches on a sample of 10 patient cases. We distinguished the results when therapies recommended by both approaches were identical ("="), and when the treatments recommended were different but compatible ("≅"), *e.g.* the intersection of the therapies recommended with both approaches was not empty.

RESULTS

Rules of $BR_{\text{-ASTI}}$ were initially produced in a factorized form, *e.g.* with THEN-parts formalized as conjunction of therapeutic choices. The first step was then to develop rules of $BR_{\text{-ASTI}}$ to have a comparable structure for both rule bases. Once this development step was performed on the 34 initially factorized rules of $BR_{\text{-ASTI}}$, we obtained 98 rules in $BR_{\text{-ASTI}}$ to be compared to the 104 completly instanciated decision rules of $BR_{\text{-GEM}}$, derived from the GEM-encoded instance.

Descriptive criteria

Quantitative comparison

In both approaches, IF-parts correspond to patient clinical descriptions. For instance, the guideline concerning patients that suffer from hypertension and stable angina, is represented by the Canadian recommendations as illustrated by the figure 2.

1. For patients with stable angina and hypertension, β-adrenergic antagonists are preferred as initial therapy (grade D).
2. Alternative therapies would include long-acting calcium-channel blockers (grade B). Short-acting calcium-channel blockers should not be used (grade C).

Fig. 2: Therapeutic recommendations for hypertensive patients with ischemic heart disease.

The second recommendation of the previous example is represented in $BR_{\text{-ASTI}}$ as:

"**IF**
pathology = HT
and pathology = ST_ANG
and level_of_intention = 2
THEN
nature = C08C *// long-acting calcium channel blockers*
and grade = B"

In $BR_{\text{-GEM}}$, the rule corresponding to the same recommendation is represented as:

"**IF**
patient_state.pathology = HT
and patient.pathology=ST_ANG
and treatment.line=L1 and treatment.intention=INT1
and treatment.type=MONO
and treatment.nature = BAA
and treatment.response = INT

THEN
treatment.line = L1 and treatment.intention = INT2
and treatment.type=MONO
and treatment.nature= LA_CCB
// long-acting calcium channel blockers
WITH
recommendation.strength = B"

The therapeutic level of intention is encoded by a unique attribute in BR_{-ASTI}. On the contrary, following the interpretative framework we previously introduced, steps of the therapeutic strategy are characterized in BR_{-GEM} by a therapeutic line and a therapeutic level of intention. In addition, the level of drug combination of the ongoing treatment, *i.e.* MONO for monotherapy, is indicated as well as the nature of the treatment, *e.g.* the therapeutic class of drugs. The response to the current treatment is explicit in BR_{-GEM} rules by the instanciation of a specific attribute, *i.e. treatment.response* = INT (for intolerate), which is not the case in BR_{-ASTI}. As a consequence, the number of criteria in IF-parts of BR_{-GEM} should be higher than the one in BR_{-ASTI} which is confirmed by the computation (table 1).

THEN-parts are similarly formalized in both approaches and characterize the therapeutic class recommended by the guideline in the clinical situation described by the IF-part. Whereas therapeutic classes are expressed as ATC codes in BR_{-ASTI}, therapeutic classes are expressed according to the labels used in the CPGs in BR_{-GEM}. Like in IF-parts, the level of drug combination, *i.e.* mono, bi, or tritherapy, is also more precisely described in BR_{-GEM}, *e.g. treatment.type*. It is the same for the two other criteria used to position the treatment recommendation in the therapeutic history, *e.g.* the therapeutic line and the therapeutic level of intention. As foreseen, THEN-parts of rules are also more specific in BR_{-GEM} than in BR_{-ASTI}.

Tab. 1: Quantitative comparison of BR_{-ASTI} and BR_{-GEM}.

	BR_{-ASTI}	BR_{-GEM}
# of elementary rules	98	104
# of premises (mean value)	2.93	4.49
# of actions (mean value)	3.10	4.42

Qualitative evaluation

The differences observed between BR_{-GEM} and BR_{ASTI} come from the ambiguity of Canadian CPGs that allows for different interpretations of some parts of the textual document. BR_{-GEM} describes 30 clinical situations, whereas BR_{-ASTI} covers only 19 clinical situations. 15 clinical situations are common to BR_{-GEM} and BR_{-ASTI}, and are denoted S_{com}. For instance, the case of patients under 60 years, suffering of hypertension with diabetes and without overt nephropathy correspond to a clinical situation that is commonly represented by both BR_{-GEM} and BR_{-ASTI}. The corresponding textual recommendation is provided in figure 4.

> 3. Preferred therapy for patients with diabetes, hypertension and overt nephropathy (albuminuria greater than 300 mg/day) is an ACE inhibitor (grade A).

Fig. 4: Therapeutic recommendation for hypertensive patients with diabetes.

In BR_{-ASTI} it is represented as:
"**IF**
pathology = HT
and pathology = OVER_NEPH
and level_of_intention = 1
THEN
nature = C09A *// ACE inhibitor*
and grade = A"

In BR_{-GEM}, it is represented as:
"**IF**
patient_state.age = AM
and patient_state.pathology = HT
and patient.pathology = DIA
and patient.pathology = OVER_NEPH
THEN
treatment.line = L1
and treatment.intention = INT1
and treatment.type=MONO
and treatment.nature= ACE_in *// ACE inhibitor*
WITH
recommendation.strength = A"

15 clinical situations are specific to $BR_{-GEM,}$ and are denoted GEM-spe. Among the 15 GEM-spe situations, 8 correspond to clinical situations described as chapter headers of the CPG that have not been taken into account in BR_{-ASTI}. This concerns 2 situations of patients with cerebrovascular disease, 3 situations of patients with peripheral vascular disease, 2 situations of patients with hyperuricemia and gout, and 1 situation of patients with hyperlipidemia. For instance, the case of patients suffering from hypertension with a history of gout is covered by the recommendation provided in figure 5.

> 3. If a diuretic is essential for the control of hypertension in a patient with a history of gout, gout can be prevented by the concurrent use of allopurinol (grade D).

Fig. 5: Therapeutic recommendation for hypertensive patients with hyperuricemia and gout.

In BR_{-GEM}, it is represented as:
"**IF**
patient_state.pathology = HT
and patient.pathology = GOUT
and treatment.line = L1
and treatment.intention = INT1
and treatment.type=MONO

and treatment.nature= DIU // *diuretics*

THEN

treatment.line = L1
and treatment.intention = INT2
and treatment.type=BI
and treatment.nature= DIU // *diuretics*
and treatment.nature= allopurinol

WITH

recommendation.strength = D"

There is no correspondent rule in BR_{-ASTI}. The 7 remaining GEM-spe situations correspond to 5 clinical situations described by ASTI at a lower level of abstraction. This is due to the document-based approach used to produce BR_{-GEM}. The remaining 2 clinical situations concern specific therapy description.

The 4 clinical situations specific to BR_{-ASTI} , and denoted ASTI-spe, correspond to "particular" textual interpretation of the guideline.

Evaluation on real patient cases

We compared the GEM-based system and the critic mode of the ASTI project on the basis of the treatments recommended by both approaches on a sample of 10 patient cases reduced to 8 cases as 2 patient cases were not exploited by ASTI. From the 8 analyzed cases, therapies recommended by both approaches were identical in 37% of the cases (3/8), and compatible in 40% of the cases (2/5). When the recommended therapies were not identical, the GEM-based approach always provided more relevant recommendations.

CONCLUSION

Previous works have established that textual CPGs expressed in natural language are subject to variations of interpretation. This results in various formalizations of original documents when manually encoded, using any dedicated formalisms, and different instances when GEM-encoded. Apart from this variability of interpretation, the aim of our work was to measure the impact of the GEM-encoding step in the translation of guidelines as formalized KBs, and to check whether this step could improve the quality of resulting KBs.

We developed a system that automatically produced a rule base from a GEM-encoded instance. Compared to BR_{-ASTI}, this rule base, denoted BR_{-GEM}, is richer (more rules), more specific (more elements in both IF-parts and THEN-parts of rules), and covers a larger number of the clinical situations described in the guideline document. This can be interpreted by the positive effect of using GEM that relies on the logical structure of the document to cut and highlight relevant parts of guideline that physicians and computer scientists may discard or forget when manually elaborating KBs. The comparison of GEM approach and critic mode of ASTI led to very promising results that need to be confirmed on a larger scale evaluation.

REFERENCES

1. Matillon Y, Durieux P. *L'évaluation médicale, du concept à la pratique*. Médecine-Sciences, Flammarion, 2000;pp.43-54.
2. 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.
3. Feldman RD, Campbell N, Larochelle P, Bolli P, Burgess ED, Carruthers SG, *et al*. 1999 Canadian recommendations for the management of hypertension CMAJ 1999;161(12):1-17. URL: http://www.cmaj.ca/content/full/161/12_suppl/s1
4. Shiffman RN, Karras BT, Aagrawal 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.
5. Hripcsak G. Arden Syntax for Medical Logic Modules. *MD Comput* 1991;8(2):76-78.
6. Ohno-Machado L, Gennari JH, Murphy SN, Jain NL, Tu SW, Oliver DE, Pattison-Gordon E, Greenes RA, Shortliffe EH, Barnett GO. The Guideline Interchange Format: A model for representing guidelines. *J Am Med Inform Assoc* 1998;5(4):357-72.
7. 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.
8. 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.
9. Giuse NB, Giuse DA, Bankowitz RA, Miller RA. Comparing contents of a knowledge base to traditional information sources. *Annu Symp Comput Appl Med Care* 1993;(Suppl):626-30.
10. Del Fiol G, Rocha B, Kuperman GJ, Bates DW, Nohama P. Comparison of two knowledges bases on the detection of drug-drug interactions. *J Am Med Inform Assoc* 2000;7(Suppl):171-5.
11. Georg G, Séroussi B, Bouaud J. Interpretative framework of chronic disease management to guide textual guideline GEM-encoding. In: Baud R, Fieschi M, Le Beux P, Ruch P (eds). MIE 2003;IOS Press:pp.531-6.
12. URL: http://www.megginson.com/SAX/