Translating Arden MLMs into GLIF Guidelines – A Case Study of Hyperkalemia Patient Screening

Dongwen WANG
Department of Biomedical Informatics, Columbia University, New York, NY 10032, USA

Abstract. To re-examine the validity of the medical knowledge that are embedded in the legacy system, we translated a Medical Logic Module (MLM) for hyperkalemia patient screening into the GuideLine Interchange Format (GLIF). We used a set of guiding principles to direct the translation. In addition, we used the GLIF3 Guideline Execution Engine (GLEE) as a testing tool to validate the encoded GLIF guideline by applying it to 5 simulated patient cases. The result has shown that it is possible to translate Arden MLMs into GLIF guidelines. However, significant efforts are necessary to handle the problems arose during the translation process. Automatic translation could be a more generalizable approach for future work.

1. Introduction

Studies have shown that computer-based clinical decision support systems can improve clinician performance and patient outcomes [1-4]. Representation of medical knowledge is a critical issue for the success of such systems. A variety of formalisms have been used to represent the medical knowledge in clinical decision support systems. For example, the Arden Syntax has been developed to represent the modular medical logic [5]; the GuideLine Interchange Format (GLIF) has been developed to represent clinical practice guidelines [6]. Although there are previous studies to compare different medical knowledge representation formats [7-9], few of them performed direct translation of specific pieces of medical knowledge from one format to another and applied them to particular patient cases.

Columbia Presbyterian Medical Center (CPMC) is the birthplace of the Arden Syntax [5]. There are more than 200 Medical Logic Modules (MLMs) encoded in the Arden Syntax that have been developed over a period of 6 years [10]. In 1999, the clinical information system at CPMC underwent a major overhaul due to the Y2K requirement. As a result, the MLMs in the legacy system were no longer executable. Recently, in a new clinical event monitor project, we have been trying to re-examine the validity of the medical knowledge that are embedded in the existing MLMs. A possible approach is to use the GLIF3 Guideline Execution Engine (GLEE) [11,12] as a testing tool for this purpose. Accordingly, we need to re-encode the medical knowledge embedded in the existing MLMs into the GLIF format such that GLEE can be used to simulate the application of the medical knowledge to specific patient cases. In this paper, we describe our approach to the translation of Arden MLMs into GLIF guidelines. We use the translation of an MLM for hyperkalemia patient screening as an example to discuss the issues arose during the translation process.

2. Methods

Based on an existing MLM for hyperkalemia patient screening [13], a faculty-level clinician informaticist made a slight modification to it such that the revised MLM could reflect the most updated medical knowledge and avoid the decisions that were based on arguable criteria. The resulting MLM was then taken as the starting point of the translation. All the major representation features of the Arden Syntax, including most of the slots in the `maintenance`, `library`, and `knowledge` category, were used in this MLM.

The GLIF model used in this study contains different types of guideline step, such as the `action_step`, `case_step`, `choice_step`, and `patient_state_step`, which are used to represent specific tasks in guideline application. In addition, the model contains elements that are used to represent patient data, clinical events, guideline maintenance, and guideline documentation. This version of the GLIF model, represented in the RDF format, is available at: http://www.dbmi.columbia.edu/homepages/wandong/GESDOR/GLIF.rdfs. During the translation process, we took Protégé-2000 as the GLIF editing tool [14].

We used a set of guiding principles to direct the translation from Arden MLMs into GLIF guidelines. These principles include: (1) In general, an MLM is translated into a `Guideline` instance of GLIF. In case an MLM invokes another MLM, the second MLM is translated into a `Guideline` instance that is nested as a `subguideline` of the `Guideline` instance to which the first MLM is translated. (2) Most slot information in the `maintenance` category of an MLM is translated into a `Maintenance_Info` instance of GLIF, which in turn is attached as the value of the `maintenance_info` slot of the `Guideline` instance to which the MLM is translated. The mapping between the slots in the `maintenance` category of an MLM and the GLIF elements can be found in Table 1. (3) Most slot information in the `library` category of an MLM is translated into a `Supplemental_Material_List` instance of GLIF, which in turn is attached as the value of the `didactics` slot of the `Guideline` instance to which the MLM is translated. The mapping between the slots in the `library` category of an MLM and the GLIF elements can also be found in Table 1. (4) The definitions of clinical data in the `data` slot of an MLM are translated into a set of instances of the `Literal_Data_Item` class, the `Variable_Data_Item` class, or the `Event` class of GLIF, which are then referenced by other entities of the translated GLIF guideline in different context. (5) The procedure codes in the `logic` slot of an MLM are translated into a sequence of `Guideline_Step` instances that constitute the `algorithm` of the `Guideline` instance to which the MLM is translated. Specifically, the data assignment statements of an MLM are translated into the `Action_Step` instances of GLIF with the `Assignment_Action` or `Get_Data_Action` type of task; the if…then statements of an MLM are translated into the `Case_Step` instances of GLIF with `options` and associated `Decision_Conditions` specified; when necessary, instances of `Patient_State_Step` are inserted into the `algorithm` to facilitate the translation of specific flow control. Here the specification of the `Criterion` for `Decision_Condition` is a rewriting of Arden's logical expression using the Guideline Expression Language (GEL) [15]. (6) The definitions of clinical events in the `evoke` slot of an MLM are translated into the `triggering_events` of the `Guideline_Step` instance at the start of the `algorithm` of the GLIF guideline. (7) The statements in the `action` slot of an MLM are translated into a sequence of `Action_Step` instances with the `Medically_Oriented_Action_Specification` type of task, which is then attached to the sequence of `Guideline_Step` instances that are translated from the `logic` slot.

Finally, we used GLEE as a tool to test the validity of the translated GLIF guideline. For this purpose, we applied the encoded guideline to 5 simulated patient cases, which covered all the possible execution paths of the guideline as expected by the expert who created the cases.

3. Result

Based on the guiding principles described in the Methods Section, we translated the MLM for hyperkalemia patient screening into a GLIF guideline. The mapping between specific slots and statements of the MLM to the entities of the GLIF guideline is summarized in Table 1. The algorithm of the encoded GLIF guideline is shown in Figure 1.

Table 1: Translation of MLM slots and statements to GLIF entities

MLM Slots and Statements	GLIF Entities
MLM (1)	Guideline (1)
maintenance.title (1)	Guideline.maintenance_info.title (1)
maintenance.filename (1)	
maintenance.version (1)	Guideline.maintenance_info.encoded_guideline_version (1)
maintenance.institution (1)	Guideline.maintenance_info.developing_institution (1)
maintenance.author (1)	Guideline.maintenance_info.author (1)
maintenance.specialist (1)	Guideline.maintenance_info.author (1)
maintenance.date (1)	Guideline.maintenance_info.authoring_date (1)
maintenance.validation (1)	Guideline.maintenance_info.representation_status (1)
library.purpose (1)	Guideline.didactics.items.material (1)
library.explanation (1)	Guideline.didactics.items.material (1)
library.keywords (1)	Guideline.didactics.items.material (1)
knowledge.type (1)	
knowledge.data (1)	Event (1)
knowledge.data (4)	Variable_Data_Item (4)
knowledge.data (4)	(leave to local system)
knowledge.evoke (1)	Triggering_Event (1)
knowledge.logic.assignment_statement (6)	Action_Step (7), Variable_Data_Item (5)
knowledge.logic.if_then_statement (5)	Case_Step (4), Three_Valued_Criterion (8)
	Patient_State_Step (2)
knowledge.action (1)	Action_Step (1), Literal_Data_Item (1)

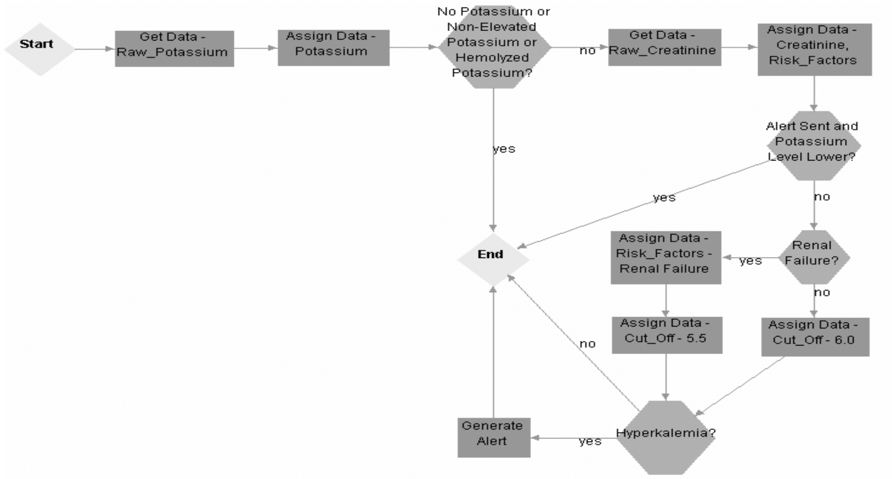

Figure 1. The algorithm of the GLIF guideline translated from the MLM for hyperkalemia patient screening. The diamond boxes are patient state steps; the rectangle boxes are action steps; the hexagon boxes are case steps. The arrows represent the flow control during guideline application.

To validate the encoding of the GLIF guideline, we used GLEE as a tool to examine the correctness of execution when the encoded guideline was applied to 5 simulated patient cases that covered all possible execution paths. The result has shown that the actual execution path for each of the 5 cases matched with its expected path.

4. Discussion

Several problems have been found in the translation of the MLM into the GLIF guideline. First, the general medical knowledge and the local policy to use the knowledge (such as the generation of alerts during specific clinical shifts) were mixed together in the MLM. During the translation, we separated these two types of statements in the MLM – only the general medical knowledge was translated and the policy issues were left to the local system to handle. Accordingly, we slightly modified the definitions of the clinical data in the original MLM that were used solely for policy issues, such that the values of these data could be directly used as the indicators of the local policy. Second, the flow control of decisions and actions in the MLM was embedded within the procedure codes. The order of the case steps and action steps in the resulting GLIF guideline was thus decided by the order of the statements in the MLM from which these steps were translated. The insertion of patient state steps depended on specific statements in the MLM, such as "`conclude false`" that implied the end of execution. Branch steps and synchronization steps were never used in the resulting guideline, which indicated the possible limitation of MLM to model concurrent tasks or tasks with undefined order. As the MLM for hyperkalemia patient screening did not invoke any other MLMs, there was no subguideline in the resulting GLIF guideline. Third, the definition of data in MLM was based on the local method for their retrieval; therefore the curly braces problem arose [16]. During the translation, we had to redefine these data using the local controlled medical terminology, the Medical Entity Dictionary (MED) [17], and the local data model [18]. Although this is a step toward the sharing of data definition, final solution to the curly braces problem depends on a widely accepted standard for controlled medical terminology and clinical data model [8]. Fourth, the clinical event that drives the execution of the MLM was encoded as the triggering event of the guideline step at the start of the algorithm. Although this approach worked well when applying the guideline to individual patient case, the performance could be a serious issue when a batch of patients is eligible for the guideline. An alternative approach is to use the triggering event to populate the patient list for batch-mode execution and then use GLEE to execute the guideline in the batch-mode [12]. Finally, we were unable to translate the information in some of the MLM slots, such as the `filename` slot in the `maintenance` category and the `type` slot in the `knowledge` category, to the GLIF guideline side. However, the information in these slots did not have direct effects to guideline execution, as shown in the result.

We have shown that it is possible to translate the medical knowledge embedded within the Arden MLM into the GLIF format. However, such translation is labor-intensive, error-prone, and difficult to be generalized to a larger scale. An alternative approach is to automatically perform this translation based on the mapping of the elements between these two knowledge representation formats. Considering that direct translation could be difficult, such translation can be realized through an intermediate layer that comprises the generalized elements derived from different models [12,19]. The guiding principles we used for the translation from the MLM to the GLIF guideline in this work can be taken as the starting point for this purpose.

5. Conclusion

It is feasible to translate the medical knowledge embedded in the Arden MLMs into the GLIF format. However, due to the different representation features of Arden Syntax and GLIF, significant efforts are necessary to handle the problems arose during the translation process. Automatic translation could be a more generalizable approach for future work.

Acknowledgment

This work is part of the Patient Health Monitor project conducted under the 528753/PO P417322 contract with Virginia Commonwealth University's Medical Informatics and Technology Applications Consortium (MITAC), a National Aeronautics and Space Administration (NASA) Commercial Space Center. We thank Drs. Peter D. Stetson, Yves A. Lussier, Eneida A. Mendonca, Vimla L. Patel, and Edward H. Shortliffe for their helps and supports to this work.

References

[1] Johnston ME, Langton KB, Haynes RB, Mathieu A. Effects of computer-based clinical decision support systems on clinician performance and patient outcome. A critical appraisal of research. Ann Intern Med. 1994;120(2):135-42.

[2] Lobach DF, Hammond WE. Development and evaluation of a Computer-Assisted Management Protocol (CAMP): improved compliance with care guidelines for diabetes mellitus. Proc Annu Symp Comput Appl Med Care. 1994;:787-91.

[3] Tierney WM, Overhage JM, Takesue BY, Harris LE, Murray MD, Vargo DL, et al. Computerizing guidelines to improve care and patient outcomes: the example of heart failure. J Am Med Inform Assoc. 1995;2(5):316-22.

[4] Balas A, Austin SM, Mitchell JA, Ewigman BG, Bopp KD, Brown GD. The clinical value of computerized information services: a review of 98 randomized clinical trials. Arch Fam Med. 1996;5:271-8.

[5] Hripcsak G, Ludermann P, Pryor TA, Wigertz OB, Clayton PD. Rationale for the Arden Syntax. Comput Biomed Res. 1994;27(4):291-324.

[6] Peleg M, Boxwala A, Tu S, Wang D, Ogunyemi O, Zeng Q. Guideline Interchange Format 3.5 technical specification. Available at: http://www.glif.org/. Accessed on May 21, 2003.

[7] Tu SW, Musen MA. Representation formalisms and computational methods for modeling guideline-based patient care. In: Proceedings of the First European Workshop on Computer-Based Support for Clinical Guidelines and Protocols. 2000;:125-142.

[8] Wang D, Peleg M, Tu SW, Boxwala AA, Greenes RA, Patel VL, et al. Representation primitives, process models and patient data in computer-interpretable clinical practice guidelines: a literature review of guideline representation models. Int J Med Inform. 2002;68:59-70.

[9] Peleg M, Tu S, Bury J, Ciccarese P, Fox J, Greenes RA, et al. Comparing computer-interpretable guideline models: a case-study approach. J Am Med Inform Assoc. 2003;10(1):52-68.

[10] Jenders RA, Huang H, Hripcsak G, Clayton PD. Evolution of a knowledge base for a clinical decision support system encoded in the Arden Syntax. Proc AMIA Annu Symp. 1998;:558-62.

[11] Wang D, Shortliffe EH. GLEE – a model-driven execution system for computer-based implementation of clinical practice guidelines. Proc AMIA Symp. 2002;:855-9.

[12] Wang D. A generic execution model for sharing of computer-interpretable clinical practice guidelines [dissertation]. New York (NY): Columbia Univ.;2003.

[13] MLM for screening of hyperkalemia patients. Available at: http://www.dmi.columbia.edu/resources/arden/mlm/potassium_and_hyperkalemia. Accessed on October 20, 2003.

[14] Musen MA, Fergerson RW, Grosso WE, Noy NF, Crubezy M, Gennari JH. Component-based support for building knowledge-acquisition systems. Conference on Intelligent Information Processing (IIP 2000) of the International Federation for Information Process World Computer Congress. (WCC 2000). 2000.

[15] Peleg M, Ogunyemi O, Tu S, Boxwala AA, Zeng Q, Greenes RA, et al. Using features of Arden Syntax with object-oriented medical data models for guideline modeling. Proc AMIA Symp. 2001;:523-7.

[16] Pryor TA, Hripcsak G. Sharing MLM's: an experiment between Columbia-Presbyterian and LDS Hospital. Proc Annu Symp Comput Appl Med Care. 1993;:399-403.

[17] Cimino JJ, Clayton PD, Hripcsak G, Johnson SB. Knowledge-based approaches to the maintenance of a large controlled medical terminology. J Am Med Inform Assoc. 1994;1(1):35-50.

[18] Johnson SB. Generic data modeling for clinical repositories. J Am Med Inform Assoc. 1996;3(5):328-39.

[19] Wang D, Peleg M, Bu D, Cantor M, Landesberg G, Lunenfeld E, et al. GESDOR – a generic execution model for sharing of computer-interpretable clinical practice guidelines. Proc AMIA Symp. 2003;:694-8.