MHB - A Many-Headed Bridge between Informal and Formal Guideline Representations

Andreas Seyfang[1], Silvia Miksch[1], Cristina Polo-Conde[2], Jolanda Wittenberg[3],
Mar Marcos[2], Kitty Rosenbrand[3]

[1]Vienna University of Technology, Austria
[2]Universitat Jaume I, Spain
[3]Dutch Institute for Healthcare Improvement, the Netherlands

Abstract. Clinical guidelines are becoming more and more important as a means to improve the quality of care by supporting medical staff. Modelling guidelines in a computer-processable form is a prerequisite for various computer applications, to improve the quality of guidelines and to support their application. However, transforming the original text into a formal guideline representation is a difficult task requiring both the skills of a computer scientist and medical knowledge.

To bridge this gap, we have designed an intermediate representation called the MHB. It is a *Many-Headed Bridge* between informal representations such as free text and tables and more formal representations such as Asbru, GLIF, or PROforma. Obtaining an MHB representation from free text should be easier than modelling in a more formal representation because the vague expressions found in the guideline do not need to be replaced by precise information immediately.

1 Introduction

Clinical guidelines are "systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances" [1]. A guideline describes the optimal care for patients and therefore, when properly applied, it is assumed that they will improve the quality of care. Evidence-based guidelines are becoming an important and indispensable part of quality health care.

There are several formal guideline representations. Asbru, EON, GLIF, Guide, Prodigy and PROForma have been compared by [2]. Further representations are Glare [3] and GEM [4]. Although the degree of formalization varies between these approaches, the majority of them represent the guideline in a format which is precise enough to execute it (semi-)automatically.

While it is desirable to produce a formal model of a guideline, it is difficult and expensive. If a guideline is revised, the modelling effort is lost and it is not easy to detect which changes in the formal model are required by the changes in the original text. The main reason for this is that there is a large gap between natural language and the currently available formal representations. To close the above described gap in a versatile manner we designed an *intermediate representation* called MHB.

2   The Many-Headed Bridge MHB

As the name suggests, the MHB is not designed as a unidirectional bridge from a one type of natural language guideline to one formal representation. Instead, it is designed as a versatile device to improve guideline quality.

The overall structure of an MHB file is a series of *chunks*. Each chunk corresponds to a certain bit of information in the natural language guideline text, e.g., a sentence, part of a sentence, or more than one sentence. The information in a chunk is structured in various *dimensions*, e.g. control flow, data flow. Dimensions consist of optional *aspects* which contain attributes. The aspects of different dimensions are independent, i.e. a chunk can have any combination of aspects (i.e. several dimensions).

All aspects contain link elements which store the connection between each part of the MHB file and the original guideline text. Most aspects contain an attribute `degree-of-certainty` to explicitly show the confidence the original guideline authors expressed in this statement (e.g., *should*, *seems advisable*), independent of the formal grade of evidence which is derived from the quality of the studies this statement is based on. Chunks can also contain a `refer-to` element used to refer to other chunks placed elsewhere within the file to make the reuse of chunks possible.

The subsections below describe the representation of each of these dimensions. Due to space limitations, we only show the main features of this language. The full specification can be found in [5] at `www.protocure.org`.

2.1   Dimension Control Flow

One of the most prominent aspects of a guideline is: *when* to do *what*. MHB offers the following means to express this.

Decisions. In MHB the basic structure of a single decision is `if-then`. It consists of a `condition`, a `condition-modifier`, a `result`, a `result-modifier`. The `condition-modifier` supports various categories of choices, such as negative recommendations and strong and weak arguments pro and contra a certain option. The `result` designates the recommended action. In MHB the element `option-group` is used to group several `if-then` elements. The options can exclude each other or not as specified by the attribute `selection-type`.

Decomposition. A task can be decomposed into subtasks. The MHB element `decomposition` names a task as an attribute and the names of its subtasks. Often one task is performed more then once – either for a certain number of times or at certain times during another task. In these cases, `repetition` specifies a task ("envelope task") which continues during all the repetitions and a subtask ("repeated task") which is performed repeatedly.

Figure 1 shows an example for a contra-indication. The corresponding text in the guideline is

> *Absolute contra-indications for BCT: multicentricity (two or more tumors in different quadrants of the breast); . . .*

CHUNK-10041			
control	if-then	condition	multicentricity
		condition-modifier	strong rule-out
		result	BCT
data	usage	name	multicentricity
	definition	name	multicentricity
		description	two or more tumors in different quadrants of the breast
structure		status	recommendations

Fig. 1. Model of an absolute contra-indication.

2.2 Dimension Data Flow

Interwoven with control flow is the description of the data processing involved in the diagnosis and treatment of the patient. In MHB, we distinguish the following (compare Figure 1): The `definition` of a data item; The `usage` of a data item is made explicit to varying degrees in actions described in the guideline and calculation of other values. The `input` of a data item is sometimes explicitly described in the description of the patient interview or diagnosis. `abstraction`-rules describe the calculation or abstraction of one data item based on others.

2.3 Dimension Temporal Aspects

Both data and control flow may have temporal aspects which can be qualitative or quantitative. MHB covers the complexity of Asbru (which has the most complex means of modelling the temporal dimension) in modelling temporal aspects, but adds more standard concepts such as average or precise duration. For each of start, end, and duration, the minimum, maximum, estimate, and precise value can be given. The precise value excludes others, but the other three values can be combined. The difference between estimate and precise value lies in the semantics given in the guideline. If start or end are given relative to a certain starting point and it is not obviously the start of the plan described, then reference point must be noted together with the offset in the respective attribute.

In addition to the above, the temporal dimension also models qualitative temporal relations such as "A is started after the end of B". While this could be implemented using the above elements, we provide a distinct element for qualitative relations to improve the comprehensibility of the model.

2.4 Dimension Evidence

An evidence-based guideline builds a bridge from carefully examined pieces of evidence which are obtained for the problem to generally applicable recommendations. Evidence for a statement can appear in various forms:

For so called summary statements of the evidence (also called scientific conclusion), a `grade` is given to show the overall strength of the evidence supporting this

conclusion. In addition, every single literature reference that this statement is built on is graded by a `level` of evidence. This level depends on the quality of the study and the study design. Statements in the guideline can have a `literature-reference`.

2.5 Dimension Background Information

Background information describes various aspects of the topic. This may refer to a particular statement or group of statements or may only be loosely coupled to particular statements or recommendations. Also the potential for formal encoding can vary.

Intentions of the described actions or recommendations inform and motivate the reader about the reasons for certain steps. *Effects* are relations between data or phenomena and other phenomena which are not seen as events or actions. *Relations* are similar to effects, but do not postulate that one of the two named entities is the cause of the other. Other *educational information* and *explanations* give further information on details of the disease not directly related to guideline execution. *Indicators* are measurable elements of health care that give an indication about the quality.

2.6 Other Dimensions

Resources. Each action consumes resources of various nature: *Personal* resources such as the working time of clinical staff; *Devices* such as treatment facilities; and *Financial cost*.

Patient related aspects. While the resources dimension mostly represents the view of the care provider, there are several other general issues mentioned in a guideline which see treatment from the patient perspective: *Risk*, *patient discomfort*, and *health prospective* related a certain treatment option or diagnostic action.

3 Evaluation

Our evaluation of the MHB was performed both on a theoretical and on a practical level. On the *theoretical level*, the mapping between MHB and various formal representations for clinical guidelines and protocols was discussed and documented [5].

In a *practical evaluation* we modelled a significant part of the Dutch Guideline for the Treatment of Breast Carcinoma [6] in MHB. Our experience has shown that MHB is appropriate to model the statements found in the significant guideline parts. MHB not only provides constructs to express the essential knowledge we intended to model, but also allowed for a modelling with the degree of detail necessary for our purposes. An initial problem was the variation observed across the MHB models obtained initially. To solve this, we elaborated a series of basic MHB modelling recommendations. Thanks to these recommendations, the degree of variation was greatly decreased, regardless of the different background of the modellers.

Currently, we are translating the MHB model to Asbru. Although this is still work in progress, our initial experience shows that the MHB model together with the original guideline text forms a better basis for guideline formalization than the original guideline text alone. Already at this stage of modelling, various analyses have been performed, e.g. to detect the unintended use of synonyms and missing knowledge.

4   Conclusions

Our experiences in the Protocure [7] project led to the following conclusions. MHB is easier to understand than Asbru by those without computer background. However, a significant effort in training was necessary.

It is easier to create an MHB model from the original guideline text than an Asbru model. The main reason for this is that MHB does not demand complete information. Also, MHB can be structured like the guideline, while formal representations such as Asbru and others model a guideline as a hierarchy of plans or tasks which is not easy to detect in the original guideline text.

It is easier to create an Asbru model based on MHB than based on the original text alone. While missing knowledge and vague information in the guideline text still cause modelling problems, they are more efficiently handled since they are already displayed in the MHB model.

The major drawback of MHB compared to other, more strict representations such as Asbru or GLIF lies in the fact that the syntax of MHB does not impose strict rules for the usage of each attribute (or aspect). The usage is only described in a guidelines [5] and it is the author's responsibility to follow them. While this is an advantage in the early modelling phase, it takes considerable effort to arrive at a uniform naming scheme for the tasks and data items in the guideline. However, this is a known problem shared by all formal guideline representations.

Weighing the advantages and limitations of MHB, we conclude that MHB is a suitable solution for bridging the gap between the original guideline text and formal representations such as Asbru.

References

1. Field, M.J., Lohr, K.H.: Clinical Practice Guidelines: Directions for a New Program. National Academy Press (1990)
2. Peleg, M., Tu, S., Bury, J., Ciccarese, P., Fox, J., Greenes, R., Hall, R., Johnson, P., Jones, N., Kumar, A., Miksch, S., Quaglini, S., Seyfang, A., Shortliffe, E., Stefanelli, M.: Comparing Computer-Interpretable Guideline Models: A Case-Study Approach. Journal of the American Medical Informatics Association **10** (2003) 52–68
3. Terenziani, P., Molino, G., Torchio, M.: A Modular Approach for Representing and Execution Clinical Guidelines. Artificial Intelligence in Medicine **23** (2001) 249–276
4. Shiffman, R.N., Karras, B.T., Agrawal, A., Chen, R., Marenco, L., Math, S.: GEM: A Proposal for a More Comprehensive Guideline Document Model Using XML. Journal of the American Medical Informatics Association **7** (2000) 488–498
5. Seyfang, A., Miksch, S., Votruba, P., Rosenbrand, K., Wittenberg, J., von Croonenborg, J., Reif, W., Balser, M., Schmitt, J., van der Weide, T., Lucas, P., Homersom, A.: Specification of Formats of Intermediate, Asbru and KIV Representations. EU Project Protocure (2004)
6. Nationaal Borstkanker Overleg Nederland (NABON): Guideline for the Treatment of Breast Cancinoma. Van Zuiden Communications B.V. (2002)
7. Balser, M., Coltell, O., van Croonenborg, J., Duelli, C., van Harmelen, F., Jovell, A., Lucas, P., Marcos, M., Miksch, S., Reif, W., Rosenbrand, K., Seyfang, A., ten Teije, A.: Protocure: Supporting the Development of Medical Protocols Through Formal Methods. In Kaiser, K., Miksch, S., Tu, S., eds.: Computer-based Support for Clinical Guidelines and Protocols, IOS Press (2004) 103 – 107