Verification of Medical Guidelines using Task Execution with Background Knowledge

Arjen Hommersom and **Perry Groot** and **Peter Lucas**[1]
Michael Balser and **Jonathan Schmitt**[2]

Abstract. The use of a medical guideline can be seen as the execution of computational tasks, sequentially or in parallel, in the face of patient data. It has been shown that many of such guidelines can be represented as a 'network of tasks', i.e., as a number of steps that have a specific function or goal. To investigate the quality of such guidelines we propose a formalization of criteria for good practice medicine a guideline should comply to. We use this theory in conjunction with medical background knowledge to verify the quality of a guideline dealing with diabetes mellitus type 2 using the interactive theorem prover KIV. Verification using task execution and background knowledge is a novel approach to quality checking of medical guidelines.

1 INTRODUCTION

The trend of the last decades has been to base clinical decision making more and more on sound scientific evidence, i.e., *evidence-based medicine*, which has led medical specialists to develop medical guidelines, i.e., structured documents providing detailed steps to be taken by health-care professionals in managing the disease in a patient, for promoting standards of medical care. For their employment, computer-oriented languages are being developed. Examples are PRO*forma* [2] and Asbru [7], which model complex clinical processes as a 'network of tasks', i.e., a number of steps that have a specific function or goal [2]. However, guidelines should not be considered *static* objects as new scientific knowledge becomes known on a continuous basis. Rapidly changing evidence makes it difficult to adjust guidelines in such a way as to keep them up to date.

In this article, we approach this problem by applying formal methods to quality checking of medical guidelines. Here, we are mainly concerned with the meta-level approach [4] which we use to formalize general quality criteria of good practice medicine a guideline should comply to. A solid foundation can already be found in literature. In [2, 5] logical methods have been used to analyze properties of guidelines. In [4], it was shown that the theory of abductive diagnosis can be taken as a foundation for the formalization of quality requirements of a medical guideline in temporal logic. This result has been used in verifying quality requirements of good practice medicine of alternative treatments [3].

The contribution of this paper is twofold. First, we formalize quality requirements of medical guidelines which include, besides treatments, also the temporal relations between treatments. Second, us-

ing our quality requirements and medical background knowledge, we interactively verify a guideline dealing with the management of diabetes mellitus type 2. More specifically, we model the guideline as a 'network of tasks' using the language Asbru and, additionally, verify meta-level properties for this model in KIV, an interactive theorem prover. To the best of our knowledge, verification of a fully formalized guideline, as a network of tasks, using medical background knowledge has not been done before.

2 FORMALIZATION OF GUIDELINES

An example of a fragment of a guideline is shown in Figure 1, which is part of the guideline for general practitioners about the treatment of diabetes mellitus type 2, and is used as a running example in this paper. The guideline gives recommended interventions to be used for the control of the glucose level.

- Step 1: diet.
- Step 2: if quetelet index (QI) ≤ 27, prescribe a sulfonylurea drug; otherwise, prescribe a biguanide drug.
- Step 3: combine a sulfonylurea drug and biguanide (replace one of these by a α-glucosidase inhibitor if side-effects occur).
- Step 4: one of the following:
    - oral antidiabetic and insulin
    - only insulin

Figure 1. Guideline fragment on diabetes mellitus type 2 management. If one of the steps k is ineffective, the management moves to step $k + 1$.

It has been shown previously that the step-wise, possibly iterative, execution of a guideline can be described precisely by means of temporal logic [5]. In this paper we will use the variant of this logic supported by KIV [1], which is based on linear temporal logic, where time points are linearly ordered. For the verification of medical guidelines we assume at least three types of knowledge involved: medical background knowledge, order information from the guideline, and quality requirements.

Background knowledge: In diabetes mellitus type 2 various metabolic control mechanisms are deranged and many different organ systems may be affected. Glucose level control, however, is the most important mechanism. At some stage in the natural history of diabetes mellitus type 2, the level of glucose in the blood is too high (hyperglycaemia) due to decreased production of insulin by the B

cells. Oral anti-diabetics either stimulate the B cells in producing more insulin (sulfonylurea) or inhibit the release of glucose from the liver (biguanide). Effectiveness of these oral diabetics is dependent on the condition of the B cells. Furthermore, as a causal treatment, insulin can be prescribed. The mechanisms have been formalized in terms of a first-order predicate knowledge:

$$knowledge : patient \times patient$$

This predicate has been axiomatized with knowledge concerning the mechanism described above. For example

```
knowledge(pre, post) →
   (insulin ∈ pre['treatment'] →
   post['uptake(liver,glucose)'] = up ∧
   post['uptake(peripheral-tissue,glucose)'] = up)
```

denotes the physiological effects of insulin treatment. In this way, the predicate knowledge can be used to reason about the state transitions that occur during verification with temporal logic.

Medical guidelines in Asbru: The guideline fragment (Figure 1) was modelled in Asbru, as shown in Figure 2. This model consists of seven plans, which are ordered in a hierarchy. The top level plan Treatments_and_Control sequentially executes the four subplans Diet, SU_or_BG, SU_and_BG, and Insulin_Treatments. The latter is further refined by the two subplans Insulin_and_Antidiabetics and Insulin, which can be executed in any order.
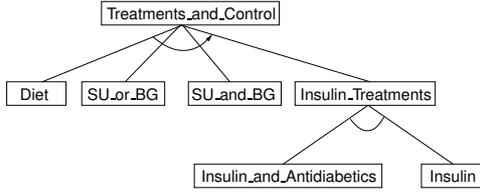

Figure 2. Asbru plan hierarchy of the diabetes mellitus type 2 guideline

Quality requirements: Here, we extend the formalization of good practice medicine of treatment choice [4, 3] to medical guidelines. Let $\mathcal{B}$ be medical background knowledge, T be a treatment, P be a patient group, N be a collection of intentions, which the physician has to achieve, and M be a medical guideline. Then M is called a *proper* guideline according to the theory of abductive reasoning, i.e., $M \in \mathsf{Pr}_P$, if:

(M1) $\mathcal{B} \cup M \cup P \not\models \bot$ (the guideline does not have contradictory effects), and

(M2) $\mathcal{B} \cup M \cup P \models \Diamond N$ (the guideline eventually handles all the patient problems intended to be managed)

Furthermore, let $\prec_\varphi$ be a partial order denoting a preference relation with $T \prec_\varphi T'$ meaning that T' is *more preferred* to T given criterion φ. If, in addition to (M1) and (M2), (M3) holds, with

(M3) $O_\varphi(M)$ holds, where O_φ is a meta-predicate standing for an optimality criterion or combination of optimality criteria φ defined as: $O_\varphi(M) \equiv \forall M' \in \mathsf{Pr}_P : \neg(M \prec_\varphi M')$,

then the guideline is said to be *in accordance with good-practice medicine*, denoted as $\mathsf{Good}_\varphi(M, P)$.

3 VERIFICATION USING KIV

Using the interactive verification tool KIV, we have verified the quality requirements discussed above, written as a sequent with the initial state of a patient group, the initial state of the guideline, the medical guideline, effects of treatment plans, the background knowledge, and environment assumptions as assumptions. For example, we verified that the order of any two treatments in the guideline was consistent with the preference order $\leq$, which minimizes drugs and number of insulin injections:

$$\Box \, \forall_T (Tick \wedge T = Patient['treatment']$$
$$\to \Box \, (\mathbf{last} \vee (Tick \to \neg(T \leq Patient['treatment']))))$$

Verification of such quality requirements could be done with a high degree of automation of upto 90%.

4 CONCLUSIONS

In our study we have setup a general framework for the verification of medical guidelines, consisting of a medical guideline, medical background knowledge, and quality requirements. We developed a model for the background knowledge of glucose level control in diabetes mellitus type 2 patients and a theory for quality requirements of good practice medicine based on the theory of abductive diagnosis. This model of background knowledge and theory of quality requirements were then used in a case study in which we verified several quality criteria of the diabetes mellitus type 2 guideline used by the Dutch general practitioners. In the case study we use Asbru to model the guideline as a network of tasks and KIV for the formal verification.

In the course of our study we have shown that the general framework that we have setup for the formal verification of medical guidelines with medical background knowledge is feasible and that the verification of quality criteria can be done with a high degree of automation. We believe both the inclusion of medical background knowledge and semi-automatic verification of the quality of decisions advised by the medical guideline are necessary elements for adequately supporting the development and management of medical guidelines.

REFERENCES

[1] M. Balser, C. Duelli, W. Reif, and G. Schellhorn, 'Verifying concurrent systems with symbolic execution', *Journal of Logic and Computation*, **12**(4), 549–560, (2002).

[2] J. Fox and S. Das, *Safe and Sound: Artificial Intelligence in Hazardous Applications*, AAAI Press, 2000.

[3] A.J. Hommersom, P.J.F. Lucas, and M. Balser, 'Meta-level Verification of the Quality of Medical Guidelines Using Interactive Theorem Proving', in *JELIA'04*, volume 3229 of *LNCS*, pp. 654–666, Lisbon, Portugal, (2004). Springer-Verlag.

[4] P.J.F. Lucas, 'Quality checking of medical guidelines through logical abduction', in *Proceedings of AI-2003, the 23rd SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence*, eds., F. Coenen, A. Preece, and A.L. Mackintosh, volume XX, pp. 309–321, London, (2003). Springer.

[5] M. Marcos, M. Balser, A. ten Teije, and F. van Harmelen, 'From informal knowledge to formal logic: A realistic case study in medical protocols', in *Proceedings of EKAW*, pp. 49–64. Springer, (2002).

[6] M. Peleg et al., 'Comparing computer-interpretable guideline models: a case-study approach', *Journal of the American Medical Informatics Association*, **10**(1), 52–68, (2003).

[7] A. Seyfang, R. Kosara, and S. Miksch, 'Asbru's reference manual, asbru version 7.3', Technical Report Asgaard-TR-20002-1, Vienna University of Technology, Institute of Software Technology, (2002).