Logic and the Quality of Medical Guidelines[1]

Arjen Hommersom[a] Peter Lucas[a] Michael Balser[b]

[a]Dept. of Information and Knowledge Systems
University of Nijmegen, Email: {arjenh,peterl}@cs.kun.nl
[b]Institut für Informatik, Universität Augsburg
Email: balser@informatik.uni-augsburg.de

Abstract

Requirements about the quality of medical guidelines can be represented using schemata borrowed from the theory of abductive diagnosis, using temporal logic to model the time-oriented aspects expressed in a guideline. In this paper, we investigate how this approach can be mapped to the facilities offered by a theorem proving system for program verification, KIV. It is shown that the reasoning that is required for checking the quality of a guideline can be mapped to such theorem-proving facilities.

1 Introduction

Health-care is becoming more and more complicated at an astonishing rate. Medical doctors are increasingly expected to take decisions balancing benefits for the patient against financial costs. Guidelines are documents supporting health-care professionals in managing a disease in a patient to avoid substandard practices or outcomes. Their aim is to promote standards of medical care.

AI researchers see guidelines as good real-world examples of highly structured, systematic documents that are amenable to formalisation. The quality of medical guidelines can be studied by carrying out a formal analysis of the text itself, which is an *object-level analysis*. Here, we are concerned with the *meta-level analysis* of guideline quality, which consists of formalising general properties to which a guideline should comply, and then investigating whether this is the case. For example, a good-quality medical guideline regarding treatment of a disorder should preclude the prescription of redundant drugs. Carrying out such checks could be valuable, in particular during the process of *designing* medical guidelines.

In this paper we study the use of logical deduction using temporal logic to formally establish whether a guideline fulfils particular quality requirements. For this purpose use was made of the theorem prover KIV [1]. This is a somewhat unusual approach, as KIV and its underlying logics are especially targeted at the verification of parallel programs, whereas here we are concerned with a type of reasoning that comes from AI.

[1]This work has been partially supported by the European Commission's IST program, under contract number IST-FP6-508794 Protocure II.

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: insulin with optionally an oral antidiabetic

Figure 1: Tiny fragment of a clinical guideline on diabetes mellitus type 2.

2 Temporal Logic and Guideline Representation

The design of a medical guidelines is far from easy. Firstly, the gathering and classification of the scientific evidence underlying and justifying the recommendations mentioned in a guideline is time consuming, and requires considerable expertise in the medical field concerned. Secondly, medical guidelines are very detailed, and making sure that all the information contained in the guideline is complete for the guideline's purpose, and based on sound medical principles is hard work. Fig. 1 shows a tiny portion of the guideline for general practitioners about the treatment of diabetes mellitus type 2, which will be used as a running example in this paper.

As medical management is a time-oriented process, diagnostic and treatment actions described in guidelines are performed in a temporal setting. It has been shown previously that the step-wise, possibly iterative, execution of a guideline, such as the example in Fig. 1, can be described precisely by means of temporal logic [6]. This is a modal logic, where relationships between worlds in the usual possible-world semantics of modal logic is understood as time order, i.e. formulae are interpreted in a *temporal structure* $\mathcal{F} = (\mathbb{T}, <, I)$. We will assume that the progression in time is *linear*, i.e. $<$ is a strict linear order. For the representation of the medical knowledge involved it appeared to be sufficient to use rather abstract temporal operators as proposed in literature [5]. The language of standard logic, with equality and unique names assumption, is augmented with the modal operators G, H, P and F. Semantically, $t \vDash \mathsf{G}\varphi$ is defined as $\forall t' \geq t : t' \vDash \varphi$ (φ is true now and at all future times) and $t \vDash \mathsf{H}\varphi$ as $\forall t' < t : t' \vDash \varphi$ (φ has always been true in the past). The last two operators are simply defined in terms of the first two operators where $\mathsf{P}\varphi \equiv \neg\mathsf{H}\neg\varphi$ (somewhere in the past) and $\mathsf{F}\varphi \equiv \neg\mathsf{G}\neg\varphi$ (somewhere in the future). For a full axiomatisation of this logic, see Ref. [7].

Even though this logic was shown to be suitable for representation purposes, we had to map it to the temporal logic underlying KIV, which we had chosen as the system to be used for formal verification. The interactive theorem prover KIV offers support for future-time linear temporal logic [2]. Reactive systems can be described in KIV by means of state-charts or parallel programs; here we use parallel programs. A state of a system can be described by first-order logic. Furthermore, static variables v, which have the same values at each time point, are distinguished from dynamic variables V. A specialty of KIV is the use of primed and double-primed variables: a primed variable V' represents the value of this variable after a system transition, the double-primed variable V'' is interpreted as the value after an environment transition. System and environment transitions alternate, with V'' being equal to V in the successive state.

The supported future-time temporal operators include a $\Box$ operator and a $\Diamond$ operator which are equivalent to the G and F operators. Furthermore it includes an **until** operator such that $t \vDash \varphi$ **until** ψ iff $\exists t' \geq t(t' \vDash \psi \wedge \forall\, t \leq t'' < t' : t'' \vDash \varphi)$ (φ holds until ψ eventually holds).

In addition, parallel programs can be written in KIV. A temporal logic property for a parallel program is verified in KIV by symbolic execution with induction. Hence, there is a major difference between the temporal logic underlying KIV and the one discussed above for representing medical knowledge, both in intention and in expressive power.

3 Application to Medical Knowledge

It is assumed that two types of knowledge are involved in detecting the violation of good medical practice. Firstly, knowledge concerning the (patho)physiological mechanisms underlying the disease, and the way treatment influences these mechanisms. We call this *background knowledge*. The knowledge involved could be causal in nature, and is an example of *object-knowledge*. Secondly, knowledge concerning good practice in treatment selection; this is *meta-knowledge*. Below we present some ideas on how such knowledge may be formalised using temporal logic.

We are interested in the prescription of drugs, taking into account their mode of action. Abstracting from the dynamics of their pharmacokinetics, this can be formalised as $(\mathsf{G}\, d \wedge r) \rightarrow \mathsf{G}(m_1 \wedge \cdots \wedge m_n)$, where d is the name of a drug or possibly of a group of drugs indicated by a predicate symbol, r is a (possibly negative or empty) *requirement* for the drug to take effect, and m_k is a mode of action, such as decrease of release of glucose from the liver, which holds at all future times.

The modes of action m_k can be combined, together with an *intention* n (achieving normoglycaemia, i.e. normal blood glucose levels, for example), a particular patient *condition* c, and *requirements* r_j for the modes of action to be effective such that $(\mathsf{G}m_{i_1} \wedge \cdots \wedge \mathsf{G}m_{i_m} \wedge r_1 \wedge \cdots \wedge r_p \wedge \mathsf{H}c) \rightarrow \mathsf{G}n$.

Good practice medicine can then be formalised as follows. Let $\mathcal{B}$ be background knowledge, $T \subseteq \{d_1, \ldots, d_p\}$ be a set of drugs, C a collection of patient conditions, R a collection of requirements, and N a collection of intentions which the physician has to achieve. A set of drugs T is a *treatment* according to the theory of abductive reasoning if [4]:

(1) $\mathcal{B} \cup \mathsf{G}T \cup C \cup R \nvDash \bot$ (the drugs do not have contradictory effects), and

(2) $\mathcal{B} \cup \mathsf{G}T \cup C \cup R \vDash N$ (the drugs handle all the patient problems intended to be managed)

If in addition to (1) and (2) condition (3) $O_\varphi(T)$ holds, where O_φ is a meta-predicate standing for an optimality criterion or combination of optimality criteria φ, then the treatment is said to be *in accordance with good-practice medicine*. A typical example of this is subset minimality $O_\subset$:

$$O_\subset(T) \equiv \forall T' \subset T : T' \text{ is not a treatment according to (1) and (2)}$$

i.e. the minimum number of effective drugs are being prescribed.

4 Management of Diabetes Mellitus Type 2

4.1 Diabetes Type 2 Background Knowledge

Since diabetes mellitus is a very complicated disease one would expect that the diabetes mellitus type 2 guideline is also complicated. This, however, is not the case, as may already be apparent from the guideline fragment shown in Fig. 1. This indicates that much of the knowledge concerning diabetes mellitus type 2 is missing from the guideline, and that without this background knowledge it will be impossible to spot the sort of flaws we are after. Hence, the conclusion is that a deeper biological analysis is required, the results of which are presented below.

The protein hormone insulin, which is produced by the *B cells* in the Langerhans islets of the *pancreas*, has two major effects. Firstly, it increases the uptake of glucose by the liver, where it is stored as glycogen, and inhibits the release of glucose from the liver. Furthermore, it increases the uptake of glucose by insulin-dependent tissues, such as muscle and adipose tissue.

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 the decreased production of insulin by the B cells.

Treatment of diabetes type 2 consists of a combination of four types of drugs. The use of *sulfonylurea* (SU) drugs, such as tolbutamid, stimulate the B cells in producing more insulin, and if the cells are not completely exhausted, the hyperglycaemia can thus be reverted to normoglycaemia (normal blood glucose levels). Secondly, *biguanides* (BG) drugs, such as metformin, inhibit the release of glucose from the liver. Finally, injection of *insulin* can be prescribed, which is the ultimate causal treatment.

The background knowledge concerning the (patho)physiology of the glucose metabolism as summarised above is formalised using temporal logic, and kept as simple as possible. The specification is denoted by $\mathcal{B}_{\mathrm{DM2}}$:

(1) $\mathsf{G}\,\mathrm{Drug}(\mathit{insulin}) \rightarrow \mathsf{G}\,(\mathit{uptake}(\mathit{liver}, \mathit{glucose}) = \mathit{up} \wedge \mathit{uptake}(\mathit{peripheral\text{-}tissues}, \mathit{glucose}) = \mathit{up})$

(2) $\mathsf{G}(\mathit{uptake}(\mathit{liver}, \mathit{glucose}) = \mathit{up} \rightarrow \mathit{release}(\mathit{liver}, \mathit{glucose}) = \mathit{down})$

(3) $(\mathsf{G}\,\mathrm{Drug}(\mathrm{SU}) \wedge \neg \mathit{capacity}(\mathit{B\text{-}cells}, \mathit{insulin}) = \mathit{exhausted}) \rightarrow \mathsf{G}\,\mathit{secretion}(\mathit{B\text{-}cells}, \mathit{insulin}) = \mathit{up}$

(4) $\mathsf{G}\,\mathrm{Drug}(\mathrm{BG}) \rightarrow \mathsf{G}\,\mathit{release}(\mathit{liver}, \mathit{glucose}) = \mathit{down}$

(5) $(\mathsf{G}\mathit{secretion}(\mathit{B\text{-}cell}, \mathit{insulin}) = \mathit{up} \wedge \mathit{capacity}(\mathit{B\text{-}cells}, \mathit{insulin}) = \mathit{subnormal} \wedge \mathrm{QI} \leq 27 \wedge \mathsf{H}\,\mathrm{Condition}(\mathit{hyperglycaemia})) \rightarrow \mathsf{G}\,\mathrm{Condition}(\mathit{normoglycaemia})$

(6) $(\mathsf{G}\mathit{release}(\mathit{liver}, \mathit{glucose}) = \mathit{down} \wedge \mathit{capacity}(\mathit{B\text{-}cells}, \mathit{insulin}) = \mathit{subnormal} \wedge \mathrm{QI} > 27 \wedge \mathsf{H}\,\mathrm{Condition}(\mathit{hyperglycaemia})) \rightarrow \mathsf{G}\,\mathrm{Condition}(\mathit{normoglycaemia})$

(7) $((\mathsf{G}\mathit{release}(\mathit{liver}, \mathit{glucose}) = \mathit{down} \vee \mathsf{G}\mathit{uptake}(\mathit{peripheral\text{-}tissues}, \mathit{glucose}) = \mathit{up}) \wedge \mathit{capacity}(\mathit{B\text{-}cells}, \mathit{insulin}) = \mathit{nearly\text{-}exhausted} \wedge \mathsf{G}\mathit{secretion}(\mathit{B\text{-}cells}, \mathit{insulin}) = \mathit{up} \wedge \mathsf{H}\,\mathrm{Condition}(\mathit{hyperglycaemia}) \rightarrow \mathsf{G}\,\mathrm{Condition}(\mathit{normoglycaemia})$

(8) $(\mathsf{G}\mathit{uptake}(\mathit{liver}, \mathit{glucose}) = \mathit{up} \wedge \mathsf{G}\mathit{uptake}(\mathit{peripheral\text{-}tissues}, \mathit{glucose}) = \mathit{up}) \wedge$
$\mathit{capacity}(\mathit{B\text{-}cells}, \mathit{insulin}) = \mathit{exhausted} \wedge \mathsf{H}\mathrm{Condition}(\mathit{hyperglycaemia}))$
$\rightarrow \mathsf{G}(\mathrm{Condition}(\mathit{normoglycaemia}) \vee \mathrm{Condition}(\mathit{hypoglycaemia}))$

(9) $(\mathrm{Condition}(\mathit{normoglycaemia}) \oplus \mathrm{Condition}(\mathit{hypoglycaemia})$
$\oplus \mathrm{Condition}(\mathit{hyperglycaemia}))$

where $\oplus$ stands for the exclusive OR. Note that when the B-cells are exhausted, increased uptake of glucose by the tissues may not only result in normoglycaemia but also in hypoglycaemia (something not mentioned in the guideline).

4.2 Quality Check

The consequences of various treatment options were examined using the method introduced in Section 3. Hypothetical patients are considered, and treatment is selected according to the guideline fragments given in Fig. 1. Firstly, consider a patient with hyperglycaemia due to nearly exhausted B-cells:

$$\mathcal{B}_{\mathrm{DM2}} \cup \mathsf{G}\,T \cup \{\mathit{capacity}(\mathit{B\text{-}cells}, \mathit{insulin}) = \mathit{nearly\text{-}exhausted}\} \cup \{\mathsf{H}\mathrm{Condition}(\mathit{hyperglycaemia})\} \models \mathsf{G}\mathrm{Condition}(\mathit{normoglycaemia})$$

holds for $T = \{\mathrm{Drug(SU)}, \mathrm{Drug(BG)}\}$, which also satisfies the minimality condition $O_{\subset}(T)$. Prescription of treatment $T = \{\mathrm{Drug(SU)}, \mathrm{Drug(BG)}, \mathrm{Drug(insulin)}\}$ for a patient with exhausted B-cells, as is suggested by the guideline, yields:

$$\mathcal{B}_{\mathrm{DM2}} \cup \mathsf{G}\,T \cup \{\mathit{capacity}(\mathit{B\text{-}cells}, \mathit{insulin}) = \mathit{exhausted}\} \cup \{\mathsf{H}\mathrm{Condition}(\mathit{hyperglycaemia})\} \models \mathsf{G}(\mathrm{Condition}(\mathit{normoglycaemia}) \vee \mathrm{Condition}(\mathit{hypoglycaemia}))$$

In the last case, it appears that it is possible that a patient develops hypoglycaemia due to treatment; if this possibility is excluded, then the minimality condition $O_{\subset}(T)$ does not hold. In either case, good practice medicine is violated, which is to prescribe as few drugs as possible, taking into account costs and side-effects of drugs. Here, three drugs are prescribed whereas only two should have been prescribed (BG and insulin), and the possible occurrence of hypoglycaemia should have been prevented.

5 Quality Checking using KIV

In the previous section we have seen that temporal logic can be used to formally check a medical guideline, but so far only from a theoretical point of view. Here we will study how such proofs can be constructed semi-automatically in terms of symbolic execution with induction using the theorem prover KIV.

5.1 Specification

For the specification, we translate the constructs that were employed in the formalisation in the previous section. Firstly, the universal quantification of the axioms

over all points in time is made explicit. Secondly, the modal operators have to be translated. The only modal operators that were used were G and H. As stated, the operator G is semantically equivalent to KIV's $\Box$ operator. However, KIV does not support past-time operators, but it is possible to translate any temporal formula with past-time operators to an equivalent temporal formula with only future-time operators that includes 'until' [3]. This implies that after translation it is possible, at least in principle, to verify the temporal formulas introduced in sections 3 and 4.

In KIV, functions and predicates are static, i.e. they do not change over time. Therefore, for the formalisation in KIV functions and predicates were mapped to dynamic variables. For example, *secretion*(*B-cells*, *insulin*) was mapped to a dynamic variable named `BsecretionI`. KIV's specification is based on algebraic structures. Since variables in axioms of algebraic specifications are universally quantified, a procedure with name '`patient`' was used to bind these variables. This gives each relevant variable a context and prohibits instantiations of axioms with variables that have different names.

The axioms (3), (4) and (7) were selected and translated to KIV's syntax as described above. In addition, a number of variables were primed to deal with the consistency condition mentioned in Section 3, as will be discussed in Section 5.3. This yielded the following three sequents, denoted by $\mathcal{A}$:

$$\texttt{[patient(; Drugs, Condition, UptakeLG, UptakePG, ReleaseLG BcapacityI, BsecretionI, QI)]} \vdash$$
$$\Box\ (((\Box\ \texttt{SU} \in \texttt{Drugs}) \wedge \texttt{BcapacityI} \neq \texttt{exhausted}) \rightarrow \Box\ \texttt{BsecretionI}' = \texttt{up});$$

$$\texttt{[patient(; Drugs, Condition, UptakeLG, UptakePG, ReleaseLG BcapacityI, BsecretionI, QI)]} \vdash$$
$$\Box\ ((\Box\ \texttt{BG} \in \texttt{Drugs}) \rightarrow (\Box\ \texttt{ReleaseLG}' = \texttt{down}));$$

$$\begin{aligned}&\texttt{[patient(; Drugs, Condition, UptakeLG, UptakePG, ReleaseLG BcapacityI, BsecretionI, QI)]} \vdash\\ &\neg(\texttt{Condition} = \texttt{hyperglycaemia until}\\ &\quad \neg((((\Box\ \texttt{ReleaseLG}' = \texttt{down}) \vee (\Box\ \texttt{UptakePG} = \texttt{up}))\\ &\qquad \wedge(\texttt{BcapacityI} = \texttt{nearly-exhausted}) \wedge \Box\ \texttt{BsecretionI}' = \texttt{up})\\ &\qquad \rightarrow (\Box\ \texttt{Condition}' = \texttt{normoglycaemia})));\end{aligned}$$

Now define ${\mathcal{B}'}_{\mathrm{DM2}}$ as the conjunction of the right-hand-sides of $\mathcal{A}$. We will show how the meta-level properties follow from these right-hand-sides. The procedure `patient` only acts as a placeholder.

For the rest of the paper define Ψ as

$$\begin{aligned}\Psi(T) = \neg(&\texttt{Condition} = \texttt{hyperglycaemia until}\\ &\neg(((\Box\ \texttt{Drug} = T) \wedge \texttt{BcapacityI} = \texttt{nearly-exhausted})\\ &\quad \rightarrow (\Box\ \texttt{Condition}' = \texttt{normoglycaemia})));\end{aligned}$$

which is semantically equivalent to

$$\begin{aligned}&\mathsf{G}(\mathsf{G}\,T \wedge \mathit{capacity}(\mathit{B\text{-}cells}, \mathit{insulin}) = \mathit{nearly\text{-}exhausted} \wedge\\ &\quad \mathsf{H}\mathrm{Condition}(\mathit{hyperglycaemia}) \rightarrow \mathsf{G}\mathrm{Condition}(\mathit{normoglycaemia}))\end{aligned}$$

5.2 Proof

Again, consider a patient with hyperglycaemia due to nearly exhausted B-cells and $T = \{\text{Drug(SU)}, \text{Drug(BG)}\}$. The following sequent was proven by KIV in about 50 steps: $\texttt{[patient(...)]} \vdash \Psi(\{\texttt{SU}, \texttt{BG}\})$. The proof relies on the fact that the axioms can be inserted with the appropriate (program-)variables, after which the patient procedure can be removed from the sequent and the real work starts. Hence, the consequent of the sequent is deduced from the axioms ${\mathcal{B}'}_{\text{DM2}}$. This yields ${\mathcal{B}'}_{\text{DM2}} \vdash \Psi(\{\texttt{SU}, \texttt{BG}\})$.

An outline of the proof follows. The proof obligation $\Gamma \vdash \Delta,\ \neg(\varphi\ \textbf{until}\ \psi)$ is equivalent to $\Gamma,\ \varphi\ \textbf{until}\ \psi \vdash \Delta$. The sequent is proved by induction over the number of steps it takes to satisfy ψ. For this, introduce a fresh dynamic variable N and generalise the sequent to $(N = N'' + 1 \wedge \varphi)\ \textbf{until}\ \psi,\ \Gamma \vdash \Delta$. The equation $N = N'' + 1$ ensures that N decreases in each step. Now, we can perform induction with induction term N which yields $(N = N'' + 1 \wedge \varphi)\ \textbf{until}\ \psi,\ \Gamma,\ N = n,\ \Box(N < n \rightarrow \mathsf{IndHyp}) \vdash \Delta$, where $\mathsf{IndHyp} = ((N = N'' + 1 \wedge \varphi)\ \textbf{until}\ \psi) \wedge \bigwedge \Gamma \rightarrow \bigvee \Delta$ and n is a new static variable. We move to the next state by symbolically executing the temporal formulae. For example, $\varphi\ \textbf{until}\ \psi \Leftrightarrow \psi \vee (\varphi \wedge \circ(\varphi\ \textbf{until}\ \psi))$ is used to execute the **until** operator. In this case, the induction hypothesis can be applied in all possible successive states.

5.3 Disproofs

The final part of this section we will show disproofs of properties that do not follow from ${\mathcal{B}'}_{\text{DM2}}$ by using program verification techniques. In the previous section we reasoned with the given axioms $\mathcal{A}$, but here we use a more extensive implementation of the `patient` procedure as discussed in Section 5.1, which not only binds variables, but implements part of the therapeutic reasoning.

Now, define the theory $M = \{\texttt{[patient(...)]}\} \cup \bigcup_{x \neq \texttt{Drugs}} \{\Box\ x' = x''\}$ where the last term denotes that variables, except for `Drugs`, are not altered by the environment, but only by the program itself. In about 400 steps using KIV it was proved that $M \vdash {\mathcal{B}'}_{\text{DM2}}$, which implies $M \vDash {\mathcal{B}'}_{\text{DM2}}$ assuming KIV is sound. From this and the fact that M is consistent (since a program is consistent and the environment is not altered), we have shown that ${\mathcal{B}'}_{\text{DM2}} \nvDash \bot$. The number of steps shows that this proof was significantly harder. The reason is that in many cases an invariant could only be defined after an initial symbolic execution. This caused an explosion of states that had to be considered. Furthermore, the invariants that had to be formulated were less straightforward.

Now showing that this set of drugs is a minimal treatment, as discussed in Section 4, we construct for all $T' \in \wp\{\text{SU}, \text{BG}\},\ T' \neq \{\text{SU}, \text{BG}\}$:

$$\begin{aligned} M_{T'} = M \cup \{&\Box\ \texttt{Drugs}' = \texttt{Drugs}'', \texttt{Condition} = \texttt{hyperglycaemia}, \\ &\texttt{BsecretionI} = \texttt{down}, \texttt{BcapacityI} = \texttt{nearly-exhausted}, \\ &\texttt{ReleaseLG} = \texttt{up}, \texttt{UptakePG} = \texttt{down}, \texttt{Drugs} = T'\} \end{aligned}$$

Again, $M_{T'}$ is consistent. It was proved in about 25 steps with KIV that $M_{T'} \vdash \neg\Psi(\{T'\})$. Because of monotony of temporal logic and $M \vDash {\mathcal{B}'}_{\text{DM2}}$, we have

$M_{T'} \vDash \mathcal{B}'_{\mathrm{DM2}}$. Since $M_{T'}$ is consistent, we can conclude that $\mathcal{B}'_{\mathrm{DM2}} \nvdash \Psi(\{T'\})$. Hence, $T = \{\mathrm{Drug(SU)}, \mathrm{Drug(BG)}\}$ is a minimal treatment. As one might expect, it shows that after the construction of the appropriate countermodel, disproofs are fairly easy.

6 Discussion

The quality of guideline design is for the largest part based on its compliance with specific treatment aims and global requirements. To this purpose, use was made of the theory of abductive, diagnostic reasoning, i.e. we proposed to diagnose potential problems with a guideline using abduction [4]. This is a meta-level characterisation of medical guideline quality. What was diagnosed were problems in the relationship between medical knowledge, suggested treatment actions in the guideline text and treatment effects; this is different from traditional abductive diagnosis, where observed findings are explained in terms of diagnostic hypotheses.

In this paper, we have made use of the interactive theorem prover KIV [1] to actually quality check a medical guidelines using the theory of quality of guidelines developed previously [5]. This complements the earlier work on object-level verification of medical guidelines using KIV [6]. About half of the steps that were needed to complete the proofs had to be done manually. However, most of the interactive steps were rather straightforward, and thus, we expect that with more specific heuristics, the proposed meta-level approach can be almost fully automated.

References

[1] Balser M, Reif W, Schellhorn G, and Stenzel K. KIV 3.0 for provably correct systems. In: LNCS 1641, Springer-Verlag, Berlin, 1999.

[2] Balser M, Duelli C, Reif W, and Schellhorn G. Verifying Concurrent Systems with Symbolic Execution. J. Logic and Computation 2002; 12(4): 549–560.

[3] Gabbay DM. The declarative past and imperative future: executable temporal logic for interactive systems. In: H. Barringer (ed.). LNCS 398, Springer-Verlag, Berlin, 1989. pp. 409–448.

[4] Lucas PJF. Symbolic diagnosis and its formalisation. The Knowledge Engineering Review 1997; 12(2): 109–146.

[5] Lucas PJF. Quality checking of medical guidelines through logical abduction. In: F. Coenen, A. Preece, A.L. Mackintosh (eds.). Research and Developments in Intelligent Systems XX, Springer, London, pp. 309–321, 2003.

[6] Marcos M, Balser M, Ten Teije A, and Van Harmelen F. From informal knowledge to formal logic: a realistic case study in medical protocols. EKAW, 2002.

[7] Turner R. Logics for Artificial Intelligence. Ellis Horwood, Chichester, 1985.