The Role of Model Checking in Critiquing based on Clinical Guidelines*

Perry Groot[a] Arjen Hommersom[a] Peter Lucas[a] Radu Serban[b]
Annette ten Teije[b] Frank van Harmelen[b]

[a] *Radboud Universiteit Nijmegen, Toernooiveld 1, 6525 ED Nijmegen,*
[b] *Vrije Universiteit, de Boelelaan 1081a, 1081 HV Amsterdam*

The full version of this paper appeared in: *Proceedings of the 11th Conference on Artificial Intelligence in Medicine*, number 4595 of LNAI, pages 411–420, Springer-Verlag, Berlin Heidelberg, 2007.

1 Introduction

There is an increasing interest amongst researchers to develop computerised versions of clinical guidelines [3], which at the moment are still just documents, using one of the specialised guideline representation languages. The resulting computer-based guidelines can then act as a basis for the development of decision-support systems, which, thus, allow computer-based deployment of guidelines in a clinical setting. One possible application of such clinical decision-support systems is *critiquing*, i.e., to spot and analyse differences between the proposed actions taken by a medical doctor, and a set of 'ideal' actions as prescribed by the computerised guideline. The common feature of a critiquing system is that the user of the system provides as input (1) a problem description (e.g., patient symptoms), and (2) a proposed solution (e.g., a treatment plan). This second input is what distinguishes critiquing systems from the more traditional expert systems, which only take a problem description as input [2]. The second input to a critiquing system, i.e., a proposed solution, is typically the output of an expert system.

One way to look upon a patient and a patient's disease logically is as a concurrent system, i.e., as a system described in terms of states and state transitions in time. Model checking technology offers methods that allow one to analyse concurrent systems for their consistency. One can rely on an extensive collection of tools and techniques readily available. It is a well investigated technique for verification of systems that can be modelled by a finite transition system. However, model checking has been mainly applied to technical systems, such as hardware, software-based communication protocols, concurrent programs, etc. Recently, it has been proposed to use model checking for the verification of clinical guidelines [1]. This raises the question whether model checking can also be used as a basis for critiquing. It is this question that is being explored in detail in this paper.

2 Approach

In the proposed method for critiquing medical treatment plans using model checking, the input to the system consists of patient data and a treatment plan (Figure 1). Patient data consists of patient symptoms and test outcomes measured for the patient, whereas the treatment plan consists of all actions (to be) performed by the practitioner. We will assume that these can be provided to the system as temporal logic formulas. Work in the Protocure project[1] and elsewhere shows that this is within the realms of the possible, given extensive methodological and tool support for the formalisation of both clinical guidelines and patient records.

*This work has been supported by the European Commission's IST program, under contract number IST-FP6-508794 Protocure-II. We acknowledge the help received from medical experts from the Dutch Institute for Healthcare Improvement, Utrecht, The Netherlands and from the Fundacio 'Josep Laporte', Barcelona, Spain.

[1] `http://www.protocure.org`

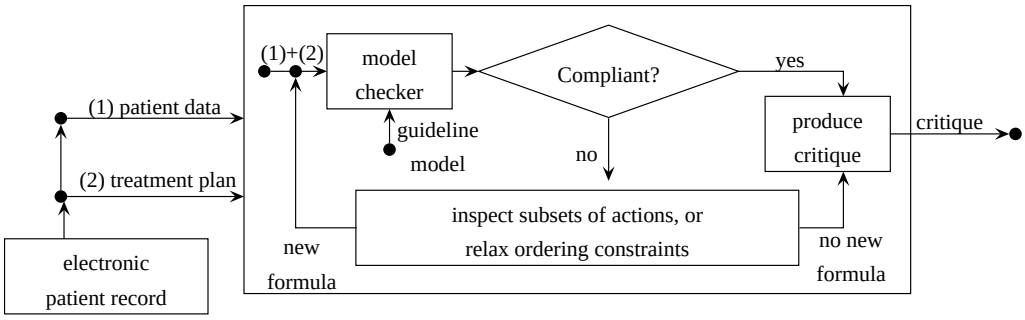

Figure 1: Critiquing approach using model checking. Given patient data and a treatment plan as input (temporal specifications), the critiquing system uses a model checker to verify consistency w.r.t. to a guideline model (state transition system) to generate a critique (empty in case of compliance).

The critiquing system uses the patient data and treatment plan as specifications that need to be checked against a formal model of the guideline, i.e., a state transition system. When the specifications are consistent with the guideline model, no critique needs to be generated as the proposed treatment plan conforms with the guideline. In case an inconsistency is found between the specification and the guideline model, the specification is weakened to get insight to which extent the treatment plan is consistent with the guideline. There are two possible reasons for the incompatibility:

Non-compliant order: It is possible that each of the actions in the treatment plan can be applied to this patient, but only in a different order than the treatment plan proposes. This can be established by removing the order between some of the actions in the treatment plan.

Non-compliant actions: Another possibility is that, according to the guideline, some of the actions cannot be prescribed at all for the patient in question. This can be investigated by considering a subset of the actions in the treatment plan.

These two approaches can be combined and lead to further insight into the nature of the detected inconsistency allowing the system to exploit these insights into a critique, which is then given to the practitioner.

3 Conclusions

In the full paper, we have investigated the feasibility of this approach by applying it to a medical guideline for breast cancer treatment. Compared to simulation-based critiquing of an operational version of the guideline, we found that model checking provides additional value. Critiquing based on running the operational guideline model through an interpreter only checks the consistency of a patient record against a single trace through the guideline (namely, the one chosen by the interpreter), while model checking compares the patient record against all possibilities allowed by the guideline. This difference is crucial when the guideline is under-specified, which is usually the case, and therefore contains non-deterministic choices between treatments. A general conclusion is that a correspondence is needed between the terminology of the guideline and the data. This is currently already being partially implemented by the Dutch Institute of Healthcare Improvement: newly constructed guidelines are currently being equipped with a data-collection dictionary.

References

[1] S. Bäumler, M. Balser, A. Dunets, W. Reif, and J. Schmitt. Verification of medical guidelines by model checking – a case study. In A. Valmari, editor, *Proceedings of 13th International SPIN Workshop on Model Checking of Software*, volume 3925 of *LNCS*, pages 219–233. Springer-Verlag, 2006.

[2] B.G. Silverman. Survey of expert critiquing systems: Practical and theoretical frontiers. *Communications of the ACM*, 35(4):106–127, 1992.

[3] S. H. Woolf. Evidence-based medicine and practice guidelines: an overview. *Cancer Control*, 7(4):362–367, 2000.