Usability Study of Computer Support to Time-oriented, Skeletal Planning with the *Asgaard* Project

Klaus Hammermüller
Vienna University of Technology, Institute of Software Technology,
Favoritenstraße 9/E188, A-1040 Vienna, Austria, Europe
`http://www.ifs.tuwien.ac.at/~hammer`

April 10, 2000

Abstract

The *Asgaard* project stresses the issue of time-oriented, skeletal planning, primarily in the medical domain. We try to support therapy planning by adding computer-aided quality assessment, plan validation and other high-level tasks to the field of planning in real-world environment. Key component is a descriptive plan representation language, called *Asbru* to enable the acquisition of computer readable medical guidelines. The research question of the Ph.D. student thesis is to prove a basic assumption of the project, that the use of *Asbru* and computer support is helpful in a real-world, time-oriented planning situation. The idea behind is to connect scientific concepts to the intended real-world target environment. A comparison with the usefulness of related modeling techniques, like workflow-process modeling, will be performed.

1 Introduction

The development of *Asgaard*[1] [11] began with a data-oriented monitoring of artificially ventilated newborn infants [8]. It was not possible to extract overall goals without the knowledge of the physicians intention. So there was a lack on planning capability. From this point modeling medical guidelines got more and more into focus and lead to a general framework specification in the time-oriented planning domain, called "The *Asgaard* Project". It deals with different aspects of planning, by developing "problem solving methods" (PSM) to support human actors during planning work. These PSM are based on a plan description language called *Asbru* [10] to enable the acquisition of computer readable medical guidelines. Those guidelines are established in wide areas of medical care, but often only in a human readable format. *Asbru* shall make semantic information accessible to automated computation [9].

[1] In Norse mythology, *Asbru* (or *Bifrost*) was the bridge from our world to *Asgaard*, the home of the gods.

The *Asgaard* project

Figure 1 gives a general outline of the framework. A human actor shall be supported (not replaced) during planning work using PSM (e.g. "verifying and validating" a plan). Additionally the system may compute different views and projections of a plan (e.g. visualization of complex relations modeled in a plan). The PSM are using a sharable plan-specification library which is the implementation of skeletal plans written in *Asbru*. A skeletal plan is familiar to a typical case or situation and not to an individual case. It looks more like a generic guideline but a complete report, ready to be matched and fitted to real-world data of an individual case. This distinction allows the reuse of collected knowledge and efficient access to all linked information [6].

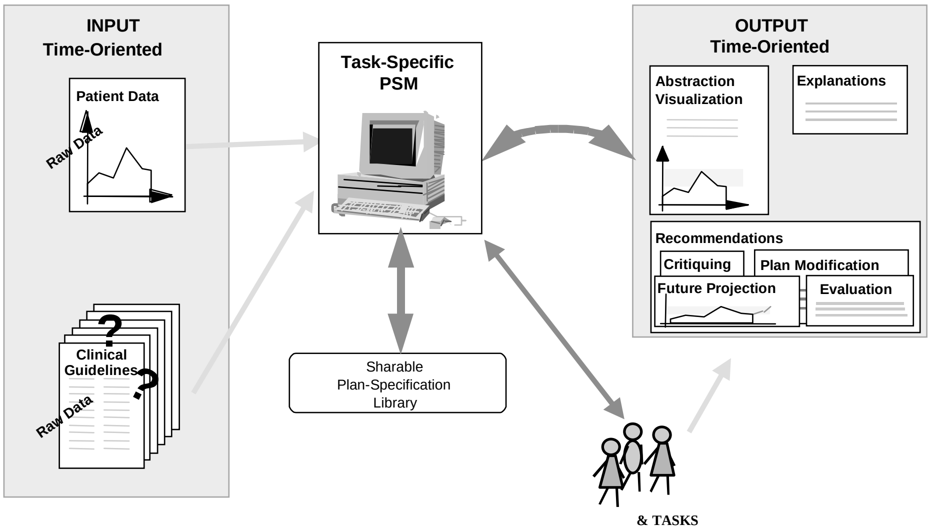

Figure 1: Overview of the Asgaard Framework

This concept is very powerful and may be applied not only in medicine but also in other environments like workout-planning in sports [7] or industrial planning processes. Currently the planning support is performed utilizing general purpose system on the one hand and domain specific problem-oriented system on the other hand. However, an approach dealing with no appropriate task-specific domain knowledge is limited. Additionally needed is a hierarchical decomposition from a very general model down to a specific scenario.

Section 2 lines out which topic shall be observed and section 3 gives a picture about the planned evaluation environment. Section 4 includes a short critiquing of the thesis. Last in section 5 the state of the work is discussed.

2 Ph.D. student thesis

Maybe it seems waste of time to have a look about the utility of the *Asgaard/Asbru* concept, it seems to be a trivial question. It is not. The concept of the *Asgaard* framework is very closely related to the idea to support planning in *reality*. It is not another toy problem planning system, it argues with the reality-impact and integrates the human actor. Therefore we need a usability study to evaluate the benefits and limitations. If it can be shown, that *Asgaard* is useful in real-world use, an implementation-guide is needed to figure out how such a system can be successfully integrated into a real-world environment.

The main hypothesis to prove is

“*Asbru* is useful to support the task of time-oriented plan management.”

There are different dimensions in analyzing this hypothesis:

- Modeling the planning environment in terms of software-engineering aspects (like in [4]) describing *how* to integrate a system into real-world planning work;
- The social impact thinking about “safety” in a system generating automated suggestions [12], a human actor may blindly rely on them and cause harm to others doing so;
- The economic analysis in a certain situation. A concept which doesn't have a valuable relation between cost and benefit for the intended user will never be used in practice;
- The knowledge modeling capabilities of the *Asbru* language have to be evaluated on a complex environment to ensure that necessary information for the PSM are expressible with *Asbru*.

These dimensions are from very different disciplines and do not seem to fit to the work of artificial intelligence concepts, but we have the problem that we need to acquire knowledge from domain experts and get real user to use the system. We collect real-world knowledge putting it into our formal models and have to prove how the intended impact evidence proving in reality. So both sides - acquisition and resulting proposition needs to face with this basic hypothesis.

Besides, we need a clear position in the field of practice. There are familiar modeling techniques, like workflow modeling, to compare. We also search for some more general patterns in planning, but at this point of work this is a very vague task.

To verify the hypothesis a well defined evaluation environment is needed, like described in Section 3. This is also the opportunity to gather extensive experience in knowledge modeling and acquisition.

3 Evaluation Environment

As the *Asgaard* project has its roots in the Artificial Intelligence, we want to proof *Asbru*'s capabilities in a real-world environment [7]. Therefore we have chosen a preventive domain, establishing support to workout plans at fitness seminars. A participant to these seminars want to check and improve health and fitness. Coaches do the education, but with the large number of participants a manual production of individual recommendations is hardly possible. Evaluation is much easier in this domain than in a hospital where the access to physicians and patients is much more limited.

Definition of the Evaluation Environment

In Figure 2 a detailed overview of our planning environment is given. The grayed boxes are the tasks which are performed by the seminar's participants, the white boxes are tasks which are supported by the *Asgaard* framework. In the horizontal dimension there are three blocks in time: Two days of the coached seminar, two months of unsupervised but monitored individual workout which closes with a half day retest to evaluate the workout performance.

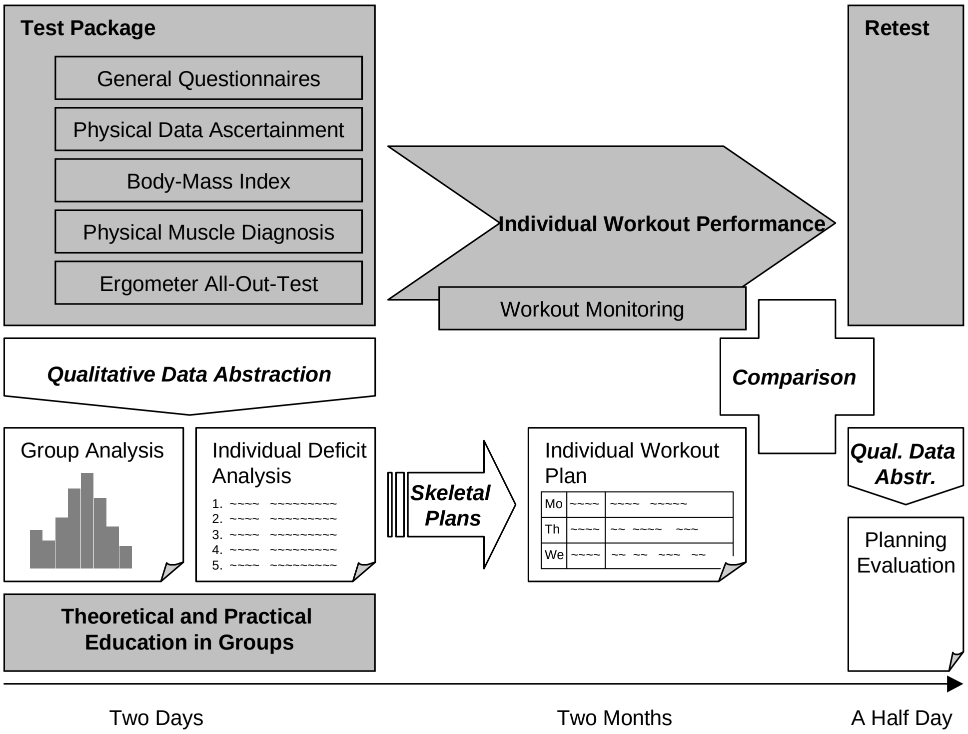

Figure 2: Evaluation Scenario of the Asgaard Framework

Starting with a well defined set of tests and questionnaires the first processing step is to derive a qualitative abstraction of the collected data to produce description about the participants' physical situation (e.g. to indicate if a measured heart-rate 140 is a "good" or "bad" condition). These computations are particularly statistical positions in a large reference group (the data is already available). This leads to an individual deficit list of an participant. In combination with the aims of the participant included in the questionnaire all information necessary are available for the following PSM:

- Selection of a skeletal plan matching to the individual needs of the participants;
- Instantiation and adaptation of this skeletal plan automatically in response to suggestions of the seminar's staff;
- Verification and validation of the (optional manual modified) plan;

- By monitoring the individually performed workout after the fitness seminar the ability to reach the intended aims may be supervised;
- The retest two month later enables an evaluation and critiquing of the whole process.

There is a large list of PSM defined and these are some of them. The focus of the work is not to have a complete *Asgaard* framework with all PSM implemented, but evaluating the use of *Asbru* itself. So knowledge modeling and acquisition is one main task and the production of a useful output for the seminar participants is the other. Some of the PSM (like verification and validation) first may be done manually as part of our research work.

Evaluation Implementation

To have a real evaluation we need three test groups, we are planning to run a large number of tests, at least greater than 50 per group:

- No planning at all;
- Planning performed by human expert;
- Planning supported by the *Asgaard* framework.

The hypothesis in this context is that the best plans are from an human expert, followed by a computer-aided plan. The result in the retest will not differ very much, because a low fitness-level allows a broad variation of the plans content. What will be very different is the time consumption of the planning work itself and the automated support monitoring unsupervised workout. No planning at all will lead to random results at the retest.

Even this is a large number for us, it is not for statistics. So we will perform black-box tests with external experts judging produces plans and results without the knowledge to wich group the participant belong. This will give us feedback about the quality of Asbru's output.

4 Problem Domain

Significant Problems in the Research

The Ph.D. student thesis is in a very early stage, so some of the problem fields which have to be observe may be somewhat weak defined yet, but the most significant problems may be clear even now:

- The interdisciplinary of the hypothesis. The different dimensions of the hypothesis may not only sound technically, but this is part of *Asgaard*s principle design. According this fact very well defined sub-questions building the hypothesis are needed.
- Extracting new scientific knowledge. The close connection to the *Asgaard* framework may a special concern, but more general pattern shall be explored too.

State of the Art

To make any comparison to existing solutions or approaches much more research work has do be done, structured by following areas:

- State of the art in planning is pretty good covered in [6];
- State of the art in workout planning has been examined in [3];
- Evaluation scenario, defining metrics and measurements building a model for the “usability” of the scenario [1];
- Examination of related approaches in workflow models like well structured processes vs. ad-hoc processes [2]

Unfortunately in this early stage of work the research on these topics is not very far, but we have to cover a broad field of related work in different disciplines.

Existing Solutions

There are partly existing solutions in different parts of the described scenario:

- There are workout-planning tools for special topics like endurance workout. Also automated diet analyses and recommendation are available;
- There are very general planning tools e.g. in project-management the network planning technique, or in the workflow management domain the process and data modeling technique;
- Also expert systems are nothing new, there is a broad variety available, mostly on a very formal and technical sound.

What is very special to the configuration with *Asbru* are two things:

- The model from the *Asbru* language tries to come as close as possible to the way how human are communicating plans. It is a formal language like other one too, but tries to stay human readable at all;
- *Asbru* can integrate very different topics into a unique planning scheme like workout and diet components. It is not a general concept but is structured about hierarchical refinements.

Technical Implementation

Asbru is defined in the XML (Extensible Markup Language) format, hence we are able to use a broad variety on existing tools, mainly written in Java. For simple transformation stylesheets are used, for more complex one Java tools are implemented. Our key tool is named *Pontifex*[2] producing from a simple XML-scheme as well as the DTD (Document Type Definition), corresponding Java classes, and a parser-extension to instantiate Java-objects of the XML document. This allows us to connect very complex and specialized functionality to a very simple and well-supported formal representation.

[2] latin: the bridge builder

In Figure 3, we give a simplyfied overview of the data processing to allow the generation of an individual deficit analysis based on quantitative check-up results.

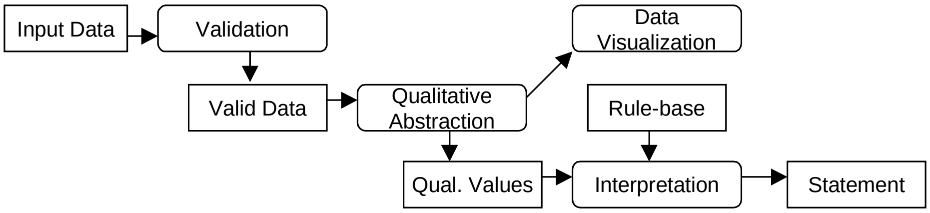

Figure 3: Process Model of the Qualitative Data Abstraction

The first step is the validation of the raw data to eliminate false input data and complete missing or wrong data if possible. This has been implemented in a previous project [5], even this is no high-frequency domain the basic model and intended output is the same. The second step is an abstraction of measure points to a description of the validated data over time. The last step is the interpretation of this description with a small domain-specific knowledge base containing the modeled rules from the coaches. The output from this process is the individual deficit analysis.

We do not try to create a plan from scratch, but we use a library of skeletal plans to select a fitting plan for an individual case, identified by the deficit analysis and general preferences given by the questionnaire. Skeletal plan means, it is a generic frame oriented on a typical problem pattern and written by a domain expert. In sports there are many patterns how to do workout available and these can be easily reused, this process is described in Figure 4.

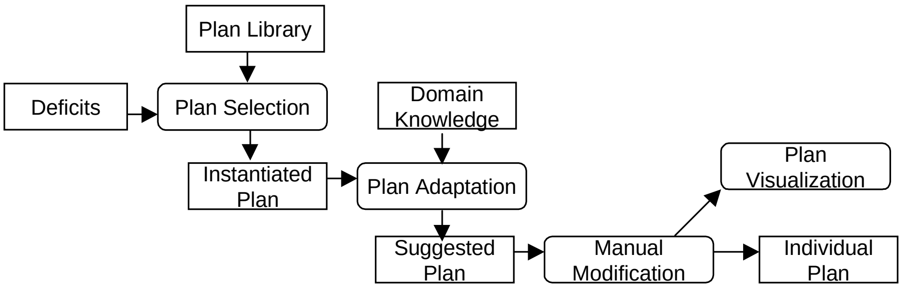

Figure 4: Generation of an Individual Workout Plan

In the first step we have the individual deficit lists and all the processed information. This information is used to select a useful plan for the current situation. The output is an instantiated plan holding the information from both sources. In the second step this plan is adapted to the real-world environment by considering the individual concerns (e.g., how much time is available or what exercises are possible). Therefore, additional domain-specific knowledge is necessary, e.g., to replace an exercise by another. The last step is the reviewing of the authored plans by the coach who may perform some refinements or select another skeletal plan. The output is a complete individual workout plan.

5 Summary

5.1 State of the Work

This work is on the very beginning. We have the general outline and selected all domain experts for the evaluation settings, sub-domains are sport science, nutrition science and medicine. As the setup of the evaluation involves a great number of staff and hundreds of seminar-participants it is not a small project. In fact we do this in a sub-project to *Asgaard* and have named it *Idun*[3], which is about the knowledge acquisition in the domain of sport and a prototype implementation for evaluation purpose. Actually we are in the middle of the work, with rapide progress.

We do this in cooperation with the Austrian Army Sport Science Service (HSWD) who had been arranging fitness seminars in a large number during the last years. Referential data about the test scenario is available. Actually the work is to build a knowledge model and acquire knowledge, what has already been started. The knowledge model is based on *Asbru* in XML (Extensible Markup Language) format. Tools for knowledge acquisition are available right now and also our own extensions are stable. PSM like selection of skeletal plans are provided but not implemented today. However the implementation is work in progress, the scientific work on some of the PSM has been initiated or finished. Technically we use a very lean design with a combination of XML and Java to avoid limitations by the complexity and limitations of propriety tools. Our experience in using powerful tools are as bad as writing powerful tools for ourselves. So we try to reduce software to the necessary minimum.

5.2 Conclusion

In the end there is a question about how useful planning is in real life and of cause, this is not a scientific question at all. But maybe this thesis leads to a better understanding what is possible in the support of planning itself.

For more information about this topic have a look at
`http://www.ifs.tuwien.ac.at/asgaard` or
`http://www.ifs.tuwien.ac.at/~hammer`.

Acknowledgements

I wish to thank Silvia Miksch (head of *Asgaard*) who has been doing great in supporting my ideas and giving valuable suggestions, as well as Manfred Zeilinger (head of HSWD) who gives these ideas the necessary practical impact. This project is supported by *"Fonds zur Förderung der wissenschaftlichen Forschung - FWF" (Austrian Science Fund)*, P12797-INF.

[3] *Idun* was the goddes of the spring. She had a box with golden apples which gave youth to the lords of *Asgaard*.

References

[1] Scenario-Based Design - Envisioning Work and Technology in System Design. 1995.

[2] A. Cichocki, Helal A. S., Rusinkiewicz M., and Woelk D. *Workflow and Process Automation: Concepts and Technology.* Kluwer Academic Publishers, Boston, 1998.

[3] K. Hammermüller and Zeilinger M. Implementation of Computeraided Workout Planning at HSNS (Army Sport Center). *2nd Symposium on Computer Science in Sport (CSS99)*, forthcoming, 2000.

[4] K. Hammermüller and S. Miksch. A Workflow Model for the Asgaard Project: A Time-oriented, Skeletal Planning Workbench in Medicine. *Workshop "Computers in Anaesthesia and Intensive Care: Knowledge-Based Information Management", in conjunction with the Joint European Conference on Artificial Intelligence in Medicine and Medical Decision Making (AIMDM'99)*, 1999.

[5] W. Horn, S. Miksch, G. Egghart, C. Popow, and F. Paky. Effective Data Validation of High-Frequency Data: Time-Point-, Time-Interval-, and Trend-Based Methods. *Computer in Biology and Medicine, Special Issue: Time-Oriented Systems in Medicine*, 27(5):389–409, 1997.

[6] S. Miksch. Plan Management in the Medical Domain. *AI Communications*, 4, 1999.

[7] S. Miksch and K. Hammermüller. Asbru, a Plan-Representing Language Modelling Time-oriented, Skeletal Plans in Sport. *2nd Symposium on Computer Science in Sport (CSS99)*, forthcoming, 2000.

[8] S. Miksch, W. Horn, C. Popow, and F. Paky. Utilizing Temporal Data Abstraction for Data Validation and Therapy Planning for Artificially Ventilated Newborn Infants. *Artificial Intelligence in Medicine*, 8(6):543–576, 1996.

[9] S. Miksch and R. Kosara. Communicating Time-Oriented, Skeletal Plans to Domain Experts Lucidly. In *10th International Conference of Database and Expert Systems Applications (DEXA99).* Springer, 1999.

[10] S. Miksch, Y. Shahar, and P. Johnson. A Task-Specific, Intention-Based, and Time-Oriented Language for Representing Skeletal Plans. *7th Workshop on Knowledge Engineering: Methods & Languages (KEML-97)*, 1997.

[11] Y. Shahar, S. Miksch, and P. Johnson. The Asgaard Project: A Task-Specific Framework for the Application and Critiquing of Time-Oriented Clinical Guidelines. *Artificial Intelligence in Medicine*, 14:29–51, 1998.

[12] L. J. Skitka, K. L. Mosier, and M. Burdick. Does Automation Bias Decision Making? *International Journal of Human-Computer Studies*, 50, 1999.