

Tool Support for Reuse of Analysis Patterns - a Case Study

Petia Wohed

Department of Computer and Systems Sciences
Stockholm University/Royal Institute of Technology
Electrum 230, 164 40 Kista, Sweden
petia@dsv.su.se

Abstract: The size and complexity of modern information systems together with requirements for short development time increase the demands for reuse of already existing solutions. The idea of reuse itself is not novel and the ability of reuse is even a part of the learning process. However, not much support for reuse can be found for the analysis phase of information systems design. Collecting reusable solutions, called patterns, in a library and supporting the search for an appropriate pattern within such a library is one approach addressed in this area. A tool for this purpose has been partly implemented and the results from a case study testing this tool are reported here.

1 Introduction

The demands put on the new software systems increases continuously with the massive introduction of IT in our society. This line of development influences in turn the requirements during the development of these systems, which often have both comprehensive functionality as well as manage large volumes of data. To meet the needs of the market, short development time has become a necessity and also one of the most competitive instruments for software vendors. This development together with the human ability of reusing old, well tested solutions over and over again has forced modern development strategies into the approaches of object-orientation and component based development: approaches advocating the formalisation of diverse solutions in the shape of objects and components aimed to facilitate reuse.

However, there is a considerable difference on how long the adoption of these approaches has come within the different phases of information systems development. The contrast is clear between the implementation phase, supported by modern environments for standard programming languages offering large libraries of reusable objects, and the analysis phase where no such libraries have yet been implemented. The development of a tool supporting particularly the analysis phase of information systems development and more precisely supporting the conceptual modelling has therefore been the subject of our work.

The terminology of patterns is adopted. The concept was first established by Christopher Alexander. In the beginning of his book “A Pattern Language” [1] he

writes: “Each pattern describes a problem which occurs over and over again in our environment, and then describes the core of the solution to that problem, in such a way that you can use this solution a million times over, without ever doing it the same way twice”. Even if Alexander worked in, and developed his patterns for the field of architecture, his concept of pattern was applicable for, and inspired people working in the area of object-oriented information systems development.

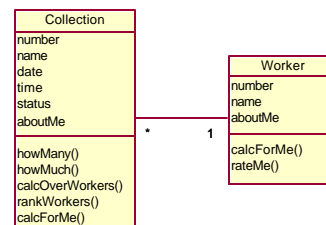
The aims of our work is to collect a library of patterns and automate the support for finding the most appropriate one for a certain situation, hoping that it is a suitable way for supporting the reuse of the huge amount of knowledge saved in the patterns. As a first step a pattern library including patterns from one domain only, the booking domain, has been collected and a modelling wizard tool for the same domain has been constructed and implemented [10]. As a second step a case study evaluating the usability of the tool has been performed and reported in this paper.

The next section summarises the work on analysis patterns related to our work. Section 3 describes the tool. The goal and the design of the study are described in Sections 4 and 5. Sections 6 and 7 document the outcome of the study. Finally, section 8 summarises the paper and gives directions for further work.

2 Related Work on Analyses Patterns

One of the earliest contributions on patterns in information systems analysis was provided by Coad, who in 1995 published “Object Models: Strategies, Patterns, and Applications” [2]. The book presents 148 strategy steps for building systems and 31 object-model patterns for building object models, and demonstrates their applicability by using them for building five different systems. Even if Coad dissociates from Alexander’s definition of a pattern claiming that “A pattern is a template, not a specific solution”, his main idea is still reuse.

The characteristic for Coad’s 31 patterns is that they are very small and generic, usually consisting of no more than two object types and a relationship between them. Coad divides 30 of his patterns in the following four categories: transaction patterns; aggregate patterns; plan patterns; and interaction patterns, and all of these patterns follow a template, the first pattern called the fundamental patterns (redrawn on the right).



Coad’s patterns do not address a particular problem, but are rather general and may be considered as building blocks of a schema.

Later, in 1996 and 1997 Hay and Fowler, respectively, independently published two more books on analysis patterns ([5] and [4]). Both these books collect a number of patterns addressing different problem domains. They do not pretend to be complete solutions, but rather consider parts of solutions, which can be reused in different situations. The main difference from Coad is that Hay’s and Fowler’s patterns could represent directly or with small modification at least some domains, whereas to model any domain with Coad’s patterns a number of patterns need to be synthesized.

However, neither Coad, Hay, nor Fowler provides any automated support for

extracting the most suitable pattern for a particular situation, but all these approaches assume that the systems analyst selects the necessary pattern by herself. A step in the direction for supporting the search of a pattern is provided by Purao and Storey. Their initial work in this area [8],[9], is based on the patterns defined by Coad, but it focuses on automating the systems design by intelligent pattern retrieval and synthesis. In order to provide such pattern retrieval and synthesis a sentence, of maximum of 24 words, describing the system is required. Based on this system description, patterns (from the library suggested by Coad) are retrieved and synthesised by the usage of techniques for natural language processing, automated reasoning and learning heuristics. When applying some of these techniques extra knowledge is used. The required system description and the retrieval of the solution are the most important differences between Purao's and Storey's approach and the approach in our work.

The modelling wizard tool described in the next section does not require any description of the system as a starting point. Instead it gathers information from the user and suggests the most suitable pattern from the pattern library. Besides the information received from the user and the pattern library no other information is used.

3 The Modelling Wizard Tool

The conceptual modelling wizard is based, as can be seen from its name, on the concept of a wizard tool, i.e., a tool gathering information from the user by asking him a number of questions and suggesting a solution tailored to the set of answers. Applied to the area of conceptual modelling, this wizard poses questions about a domain and suggests a conceptual schema for this domain (a detailed description of the tool can be found in [10]). So far the wizard supports the booking domain only.

Six questions were collected and implemented in the tool (see Figure 1). When collecting the questions, different booking solutions - both such found in Fowler [4] and Hay [5], and such constructed to capture as different booking situation as possible - were analysed. The questions were aimed to cover the differences between distinct booking situations so that a pattern, which provides a satisfactory solution, shall be identified. A brief description of the questions is given below.

The first question is about the cardinality of a booking, i.e., where a booking may consist of several booking lines (e.g. when booking a trip, both tickets and hotel rooms may be reserved), or not (usually when booking a rental car). The second question is about whether a booking concerns a (number of) concrete objects (like when booking a time to dentist) or not (e.g. when booking a book in the library, not a particular exemplar from the book, but rather the title is booked). The third question investigates whether the bookings made within a system have the same character (as is the case for cinema tickets bookings) or whether they may differ. (In the trip booking example above, tickets bookings require departure and destination places and time for departure, whereas hotel room bookings require arrival and departure times and number of beds.) The fourth question is aimed to clarify whether the request for a booking is necessary to keep information about. (For instance, when scheduling room

<p>1. Does a booking consist of</p> <p><input type="radio"/> one object, or</p> <p><input type="radio"/> may it consist of several objects?</p> <p>2. Does a booking concern a (number of)</p> <p><input type="radio"/> concrete object(s), or</p> <p><input type="radio"/> does it rather specify the character of the object(s)?</p> <p>3. Do all the bookings</p> <p><input type="radio"/> have the same character</p> <p>or may they be divided into several categories?</p> <p><input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4</p> <p><input type="radio"/> larger than 4</p> <p>4. Is it necessary to keep information about the request for a booking, before making the booking?</p> <p><input type="radio"/> no</p> <p>If yes , does the booking request concern</p> <p><input type="radio"/> a concrete object(s)</p> <p><input type="radio"/> a specification of an object(s)</p> <p>5. Does a motivation need to be given for a booking?</p> <p>5'. Does a motivation need to be given for a booking request? (depending on the answer to question 4)</p> <p><input type="radio"/> yes</p> <p><input type="radio"/> no</p> <p>6. May a booking be done on the behalf of someone else?</p> <p>6'. May a booking request be done on the behalf of someone else? (depending on the answer to question 4)</p> <p><input type="radio"/> no,</p> <p>If yes, is it important to keep information about the party who made it?</p> <p><input type="radio"/> no</p> <p><input type="radio"/> yes.</p>
--

Figure 1 The questions implemented in the tool

bookings for university courses, booking requests are collected from the head teachers for each course and used as input in the scheduling work.) The fifth and sixth questions depend on the answer to the fourth question. If booking requests are necessary to keep information about these questions are posed for the booking requests otherwise they are posed for the bookings. The fifth question asks if a motivation (a purpose) for a booking/booking request is necessary. (For the university scheduling work each head teacher gives the course he/she will make the bookings for.) The sixth question asks whether a booking/booking request may be done on behalf of someone else (e.g. a secretary may book the business trips for her/his boss).

For the sake of user-friendliness the questions are placed in sequence by showing a new question first when the previous one has been answered. A conceptual schema solution is gradually built and refined, and graphically presented after each new

answer. For instance the first alternative in the first questions should result in the schema in Figure 2a) whereas the second alternative should result in the solution presented in Figure 2b). Choosing further the first alternative from the second question should result in expanding the schema into the one shown in Figure 3a) or 3c), depending on the point of the departure, which is the result from the first question. Choosing the second alternative should give one of the solutions in Figures 3b) and 3d).

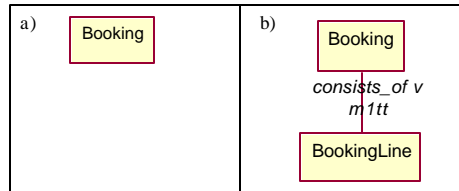


Figure 2 The alternative solutions after the first question¹

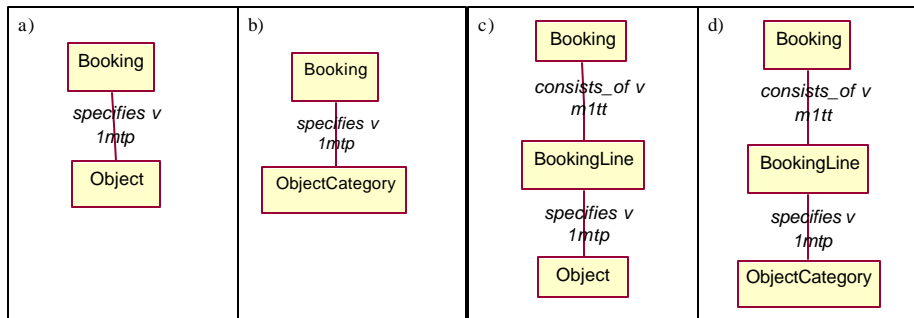


Figure 3 The alternative solutions after the second question

In this way, the wizard interviews the user and piece by piece build a schema depending on the answers. A very basic help function for clarifying the last unanswered question is implemented. Facilities for changing the schemas will also be implemented with the development of the tool. In Figure 4, two different solutions, a minimal contrasted to a more complex solutions, are shown to illustrate the tool.

4 The Motivation and the Goal of the Study

When reasoning about a tool, two aspects are necessary to investigate namely, the quality and the usability of the tool. Furthermore, two necessary criteria for the

¹ The notation is UML similar, except the direction and the cardinality constraints for a relation, which follows Johannesson's notation [6] and are represented by an error in the name of the relation and a quadruple $\langle 1m, 1m, tp, tp \rangle$ under the name of the relation, respectively. The first position in the quadruple shows whether a relation is single (1)- or multy (m)- valued. The third position shows where a relation is total (t) or partial (p). The second and the fourth positions show the same properties for the inverse of a relation.

quality of this kind of tools are the quality of the proposed solutions and the complete-

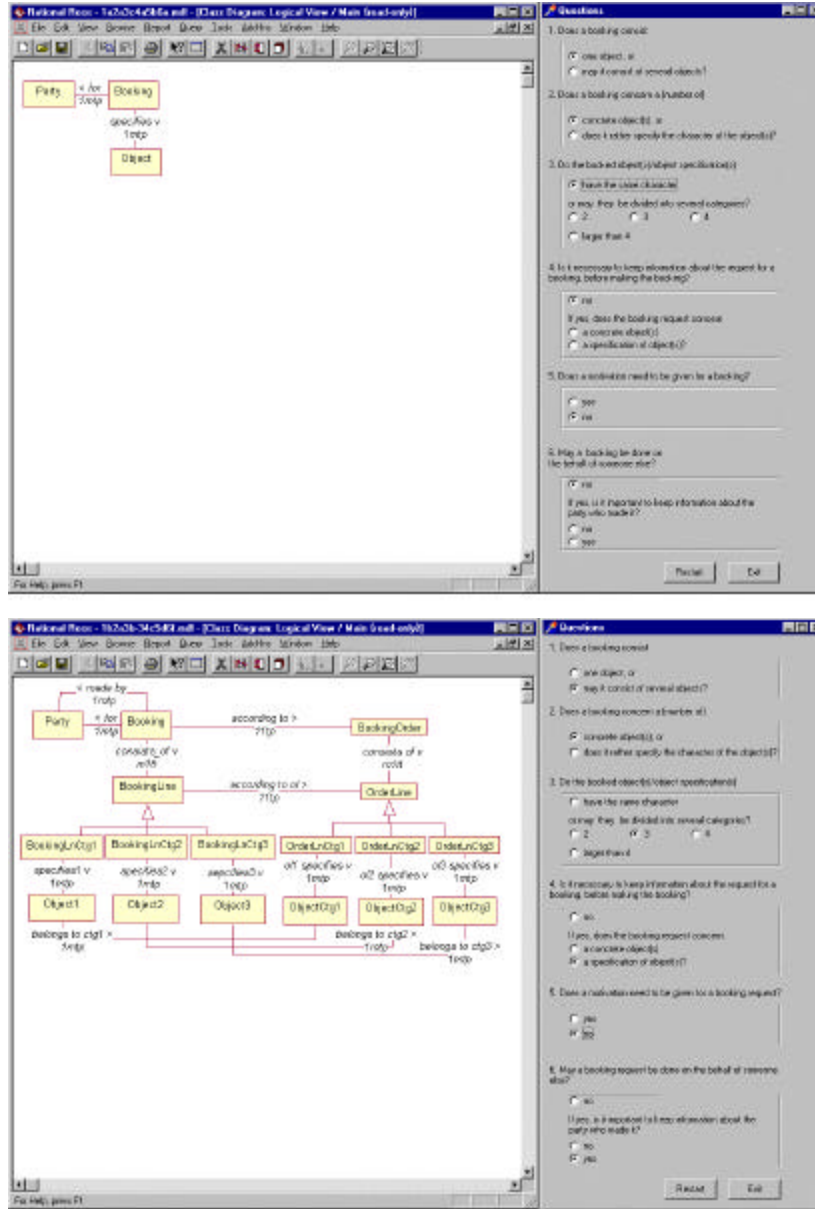


Figure 4 Two answer sets and the contrasting solutions corresponding to them

ness of the questions, where the fulfilment of the second one is much more difficult to guarantee. In the works provided in [10] and [11] the completeness of the questions is discussed based on Fillmore's case grammar theory [3] and Coad's library of

modelling patterns [2], respectively. However no one of these approaches can guarantee the completeness of the questions. In order to do that a huge work has to be done empirically, testing the tool on real cases.

According to the usability, the tool may be used for at least two different purposes: by expert software engineers as a starting point in data analysis and reuse processes; and, in the education during the self-study process for improving students' modelling skills. Different case studies testing the usability for these two purposes have to be provided. Due to the easy access to students, our work has initially been concentrated on the usability of the tool as a pedagogical aid.

A first study has been provided. The general hypothesis was that the wizard might successfully be used in the education process. The main expectation from the tool as a teaching aid was that it should stimulate the students to reflect on the different solutions depending on the differences in describing the corresponding situations to be modelled. In this way the students should be pushed to catch different modelling patterns with an underlying understanding about them. The intention was to compare the learning process between a test and a reference group, as well as to analyse the usage of the modelling wizard by the test group. Due to the small number of the subjects in the test group no quantitative, but only qualitative analysis has been done for the second matter.

5 Method and Design of the Study

37 subjects participated and worked in groups of two or three. Only one subject worked alone. The groups were divided into one test group and one reference group. The test group included nine groups organising 20 subjects and the reference group included seven groups of 17 subjects. All subjects were computer and systems sciences students attending an advanced course on logical and physical database design. The study was performed as an optional assignment of the course and gave a bonus point on the final exam to the students who participated in it. Approximately 80% of the class attended it.

The study consisted of three parts. Firstly, after the usual questions about the age and gender the students were given an initial modelling task for a hospital booking system. The time for this task was limited to 20 minutes and the task was deliberately constructed to be difficult.

Secondly, the students were given four small tasks to be done in a sequence. The domain was chosen to be a university booking system in order to ensure the students' knowledge about it and reduce the uncertainty about the domain. The test groups were asked to use the modelling wizard when performing the tasks, whereas the reference groups were asked to model by pen on paper. This sequence of tasks was constructed to exemplify four different booking situations starting with small and uncomplicated modelling cases and moving into more complex situations. The intention was to push the students to reflect on the similarities and differences between the four different situations. For the test groups, this task also intended to exemplify the capacity of the tool. During this task the test groups were expected to learn the tool, as well as to reflect on the solutions they were proposed by it. The only introduction to the tool was the information that it consisted of a number of questions and about the help

facilities it offered. For the sake of learning the tool, the groups were offered an example solution for each task, as a feed back so the subjects should be able to verify whether they correctly understood the questions placed by the tool or not. The groups with deviating solutions from the example solutions were asked to identify the differences between both solutions in order to force them to further discuss the meaning of each question. To ensure that the reference group had access to the same kind of information as the test group, the example solution were even offered to them as well as a book on conceptual modelling which they were allowed to use during the whole study.

During the third part of the study, the students were asked to reconsider their solutions from the initial task and document improvements, where they had any. The time for this task was limited to 15 minutes. Finally the subjects were asked to individually fill in a questionnaire which for the test group included questions about the reflection on the tool.

The time for the study was limited to two hours. The subjects were asked to evaluate their modelling skills within a scale of one to five² three times: in the beginning of the study, after the first part of the study, and finally before the final questionnaire. The study was documented by recording all groups on video.

6 Results

The schemas produced by the groups during the initial task as well as those produced in the final phase were evaluated by two independent judges. The judges used the same scale as the one used by the students (1-5) for evaluating their modelling skills. The results from these evaluations are shown in Figure 5. The minus in some cells

		T1	T2	T3	T4	T5	T6	T7	T8	T9	R1	R2	R3	R4	R5	R6	R7
judge1	Initial	2	3	3	3	3	3	3	2	3	2	2	2	2	2	3	2
	Final	3	3	3	2	4	-	-	3	-	2	2	2	(3)	-	-	(2.5)
judge2	Initial	1	2	4	2	3	4	3	3	2	2	3	3.5	2	2.5	3	1
	Final	2.5	2.5	4.5	3	3	-	-	4	-	2	3	4	(3.5)	-	-	(2)

Figure 5 The marks for the initial and the final models produced by the test (T1,...,T9) and the reference (R1,...,R7) groups given by the two judges³.

stands for an unchanged schema. Two of the groups did not modify their schemas because they did not thought they gained any new knowledge during the study, whereas three of the groups reconsidered, but were satisfied with their initial schemas. Generally, the changes improved the schemes. In the cases where it was not noticed

² The scale 1 to 5, where 1 stands for bad and 5 for excellent, was until some years ago the national scale within high school in Sweden. The evaluation within high schools followed the normal dispersion. This scale was chosen in the study, due to the population's familiarity with it.

³ Two groups, R4 and R7, documented the changes they wanted to perform in text only. They did not had time to draw the changes graphically. Therefore, the judgements of their final solutions as well as the differences from the initial solution are presented in parentheses.

with the mark is due to the changes was not big enough or due to in changing parts of the schema some other parts failed outside consideration and was forgotten in the new schema. However, no big difference between the test and the reference group can be noticed on the modelling skills progress during the study.

However, a behaviour difference is that the test groups generally suggested much more comprehensive changes when improving their original models, whereas all the reference groups, except one, performed small changes only. The changes are measured by the number of object types introduced or reduced, the number of relations introduced, reduced or changed, and the number of generalization constraints (Isa) introduced to the initial solution (see Figure 6). The introduction/reduction of an object type is considered as a bigger change compared to the changes of a relation, since they have a bigger influence on the stability of a schema [7].

		T1	T2	T3	T4	T5	T6	T7	T8	T9	R1	R2	R3	R4	R5	R6	R7
Object T	Introduced	2	1	2	2	2	-	-	2	-	6	-	-	(-)	-	-	(2)
	Reduced	1	-	3		3	-	-	-	-	9	-	-	(-)	-	-	(-)
Relation	Introduced	5	2	1	3	5	-	-	1	-	8	-	-	(1)	-	-	(1)
	Reduced	-	1	5	6	10	-	-	-	-	8	-	-	(-)	-	-	(-)
	Changed	1	5	2	3	2	-	-	3	-	-	1	1	(1)	-	-	(-)
Isa	Introduced	-	-	1	2	1	-	-	-	-	-	-	-	(-)	-	-	(-)

Figure 6 The differences between the initial and the final schemas³

Beside the evaluation of the schemas done by the judges, the students' own evaluations on their modelling skills were considered. The subjects were asked to evaluate their modelling skills three times during the study:

- firstly in the beginning of the study;
- secondly after the difficult initial modelling task; and
- thirdly at the end of the study after the four small tasks were performed and the reconsideration of the initial schema with possible changes was done.

The reason with the second evaluation was to confirm the first one, whereas the reason with the third evaluation was to measure the subjects' own impression of their progress during the study. The results are presented in Figure 7 and Figure 8. Most of the subjects kept their evaluations unchanged (shown by minuses in the tables) and no significant differences can be seen between the test and the reference groups. One reason for which the subjects did not increase their marks in the third evaluation may lie in the limited time for which the study was performed. One hour, which the subject approximately had for performing the sequence of the four modelling tasks, was not long enough to give the students confidence to evaluate their own skills with a higher mark. Instead some subjects decreased their marks. Notable, even if not surprisingly, is that the five subjects who decreased in their evaluation were all female. Notable is also that one subject, subject number 34, evaluated his skills with five plus, a mark outside the given scale.

gr	T1		T2		T3			T4		T5	T6		T7			T8			T9	
nr	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
gen	f	f	m	m	f	m	f	m	m	f	m	m	f	f	f	f	m	f	m	m
ev1	2	2	3	3	2	3	3	3	1	4	3	4	4	2.5	3	4	4	4	3	3
ev2	-	-	-	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ev3	-	-	-	3	-	-	-	-	-	3	-	-	-	-	-	-	-	-	-	-

Figure 7 The test subjects own evaluations of his/her modelling skills (gr = the number of the group, nr = the number of the subject, gen = the gender of the subject, ev = evaluation)

gr	R1		R2		R3			R4		R5		R6			R7		
nr	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	35	37
gen	f	f	f	f	m	f	m	m	f	m	m	f	f	m	f	f	f
ev1	2	2	3	3	3	3	3	4	3	3	3	2.5	2.5	5+	3	3	3
ev2	-	-	-	-	-	-	-	-	-	-	-	-	3	-	2	-	-
ev3	1	-	-	-	-	-	-	-	2.5	-	-	-	2.5	-	-	-	-

Figure 8 The references subjects own evaluations of his/her modelling skills

Finally, an interesting observation is that no one from the reference groups used the book they were offered. Most of them just forgot, immediately after reading the instructions, that a book was available and they did not even note the presence of it beside other necessary material like paper and pens.

7 Evaluation

In this section some reflection made from the test groups' work with the modelling tool are presented. Since the number of the test subjects is only 20 no reliable quantitative measures may be done. Instead qualitative analysis has been performed. It is based on the questions posed to the subjects after their work with the tool. The analysis was supported by the video records from the study. In the questionnaire, the subjects were asked to answer about the questions, the examples, and the solutions from the tool as well as whether they reflected on the schemas, and finally how they assessed the tool for teaching aims. Six alternatives: 'not at all', 'very little', 'little', 'much', 'very much', and 'very very much' and place for comments were given to each question.

7.1 Were the questions implemented in the wizard clear?

The questions: 1) "Did you understand the questions which the tool placed?" and 2) "How difficult did you find the questions placed by the tool?" were placed to gather information for the formulations of the questions and their impact on the domain analysis process. The subjects' impressions from the question were mixed. Some of them found the questions clear where other found them unclear. Some of the subjects found the questions difficult whereas others did not. However, generally the meaning of the second question in the tool "Does a booking concern a (number of) a concrete object(s), or b) does it rather specify the character of the objects(s)?" was difficult to grasp by the subjects. Even the fourth question, which includes the same distinction,

i.e., the distinction between an object and an object specification, caused problems. This distinction needs therefore to be better clarified.

Furthermore, the fifth question, asking whether a motivation is needed for a booking, was usually misinterpreted, even if it did not confuse the subjects, since they did believe they correctly understood it. Unfortunately, the choice of the word motivation was the failure since it did not get the subjects to associate with the action, for which a booking is done.

7.2 Did the students use the help examples?

Most of the subjects did use the help examples. Most of them thought that the examples were clearly written. Still, some subjects thought that the examples were difficult to transfer to the problems they had to solve and had therefore difficulties in using them immediately. Even the opinion that all the examples should follow one and the same case was uttered. The reason for not giving the examples from the same case now was the intention to find the most suitable examples rather than to pick them up from the same case.

Furthermore, more sophisticated help was required. Some subjects wanted to see in advance what the result of choosing different answer alternatives should be. This effect is for the moment obtained by picking different answer alternatives and considering the suggested solutions. However, in this way only one schema is shown at the time, which does not for instance support comparison of schemas resulting from different answers. Comparing two or three solutions may, therefore, be facilitated by extending the help function in the way proposed by the subjects.

7.3 Did the tool support reflection?

On the question “Did the tool help you to reflect on your schemas?” answered only one subject ‘not at all’. This was the same subject who worked alone. Ten subjects (of 20) answered very positively by choosing one of the alternatives ‘much’, ‘very much’ or ‘very very much’. Almost all of these ten subjects have also answered positively on the first question whether they understood the questions or not. Therefore, a possible relation may be that students who understood the questions were also able to understand and reflect on the solutions suggested by the tool.

7.4 Shall such a tool be used for teaching aims?

Fifteen (of 20) subjects answered positively on this question. One subject did not have any opinion on this subject. Most of those who judged the tool positively as a teaching aim thought they had learned ‘a little’ or ‘much’ from it and/or that it was ‘motivating’ or ‘very motivating’ to use it. Also all the subjects who thought that the tool supported reflection answered positively on this question.

The four subjects who were negative for using the tool within the education motivated it in the following way: “The general object names were not good. Explanation on why the tool models in a certain way is missing”, “The answer alternatives were too limited”, “The tool does not give chance to experiment, which often teaches more”, “I better like to work for myself instead of getting questions”. The last two comments are difficult to do anything about and they are therefore left aside. The generality, both of the questions and of the suggested solutions, is necessarily for such a tool. It is indeed desirable to be able to change the names in the suggested solutions with application specific words, which is not yet supported by the

prototype, but is planned to be added to the tool's functionality. According to the limitation of the answer alternatives it is important to guarantee the completeness of the alternatives for each question, which we consider they are. However, much more difficult is to guarantee the completeness of the questions and a lot of empirical studies are necessary before doing this. Finally, since the comment on missing explanations from the tool was even left from other subjects we are discussing it in the following subsection.

7.5 General Observations

Some subjects did not get the connection between the questions and the presented solutions. One suggestion was to improve the help function by explaining how a particular answer influences the development of a schema and why. Another possible solution is to implement verbal feedback explaining impact of the answers on the suggested solutions. Beside, a different reflection that may have a big influence on the understanding of a schema, is that all compulsory relations should be drawn from the very start. For instance, a booking is always made for a party. Introducing an object type Party connected to the object type Booking already in the beginning should help on understanding the intermediate solutions considerably. For the moment the Party object type is introduced very late in the intermediate solutions and it turns out from the study that this is inappropriate.

Surprisingly, some subjects found difficulties with the language. Two of those who were negative for using the tool as a teaching aim pointed out this as a problem. The language in the tool is English. However, the instructions for the study and the exercises were given in Swedish. The language difficulty was explicitly noted by some subjects when they had to answer on how well they understood the questions and the examples. Other had difficulties with interpreting some parts of the suggested solution due to problems with understanding some words. It was surprising since a large part of the literature for the university students is in English and the subjects were not expected to have problems with the language. However, the language used in a conceptual schema is really limited and not recognising a single word may cause problems for understanding a whole schema.

Furthermore, a limitation of the tool was identified when preparing the study. This limitation appears in the case when different booking categories are distinguished (identified by question 4) and they differ in matters identified by the rest of the questions. For the moment, the rest of the questions are posed for all the different bookings together, instead of posing them for each one category separately. The subjects were asked to note any limitations they discovered. The initial modelling task was deliberately constructed to include the situation described above. However, no one of the subjects found this limitation. Only one of the test groups had a suggestion about extension of the tool namely, that it should even ask whether a withdrawn booking should be taken care of, or not.

8 Conclusions and Further Work

The subject for this work has been reuse in the information systems analysis process. It is one in a series of works on documentation and collection of analysis patterns in order to expose them and make them easily available for the systems engineers. The

approach is however slightly different from the preceding approaches [2],[4], and [5] by its attempt for automating the search for the most suitable pattern when a library has been collected. It also differs from [9] by its attempt of gathering the necessary information directly from the user, instead of using meta-knowledge for the retrieval of an appropriate solution. The goal is to realize these properties by a modelling wizard tool. Such a tool should be useful in some of the following situations:

- To support novices in the process of domain analysis and conceptual schema construction;
- To be used by expert software engineers as a starting point in data analysis and reuse processes;
- With some extensions to be tailored for novice database developers with no formal knowledge, of database design, developing their own databases for use in the home;
- In the education during the self-study process for improving students' modelling skills.

As a necessary step of the development of a tool, a prototype developed for one domain has been implemented [10]. Before any work on generalising the prototype was provided [11], a case study investigating the usability of the tool has been performed. The study aimed to evaluate the tool as a pedagogical aid, as well as to give input about the opinion about such a tool. It was performed on students. The hypothesis of the study was that the tool could successfully be used in the education process. The outcome of the study is that, for the short time the subjects had, there was not a considerable difference in the learning capacity between subjects who used it and those who did not. However, a positive result from the study was that the subjects who used the tool differed from the subjects who did not, in the number of changes they wanted to perform in a schema constructed by them earlier in the study. Furthermore, the subjects who used the tool were positively for using it within the education. Some of them explicitly noted that the usage of such tools should be on voluntary bases and as a complement to, rather as a reduction of, the teachers.

The future work will go in three main directions carried out in parallel.

- The generalisation of the tool into domain independent wizard, continuing the work provided in [11]. It is necessary to provide research on how a general tool shall be constructed in order to support the modelling of complex UoD ranging over, not only one but, several different domains. One possibility should be to equip the modelling wizard with schema integration facilities. Still a number of questions should need to be considered in order to coordinate the behaviour of the tool with the behaviour of a user. One such question is to investigate the human's techniques when modelling complex domains and her way of dividing complex problems into smaller parts, either to make them easily to grasp and solve, or to distribute the responsibilities between different parties;
- The evaluation of the completeness of the wizard, which have to be done by empirical work either by using the wizard in a number of real cases or by letting a number of experts use and evaluate the tool; and

- The evaluation of the usability of the tool for the different categories of users mentioned above. The work on evaluating the wizard as a pedagogical aid can be extended to also consider the usage of the tool for a longer time in the education process, i.e., the impact of the tool when using it during a whole course, rather than using it for one hour only.

Acknowledgements

I would like to thank Paul Johannesson, Maria Bergholtz, Eva Fähræus, Sari Hakkarainen, Klas Karlgren, Robert Ramberg and Jakob Tholander, for their help with the case study. I would also like to thank the anonymous reviewers for the valuable comments on an earlier draft of the paper.

References

- [1] C. Alexander et al. *A Pattern Language*, NY: Oxford University Press, 1977.
- [2] P. Coad et al. *Object Models: Strategies, Patterns, & Applications*, Yourdon Press, Prentice Hall, 1995.
- [3] C.H. Fillmore, “The case for case”, in Bach and Harms (eds.) *Universals in Linguistic Theory*, Holt, Rinehart and Winston, New York, 1968.
- [4] M. Fowler, *Analysis Patterns: Reusable Object Models*, Addison-Wesley, 1997.
- [5] D.C. Hay, *Data Model Patterns: Conventions of Thought*, Dorset House Publishing, 1996.
- [6] P. Johannesson, *Schema Integration, Schema Translation, and Interoperability in Federated Information Systems*, Dissertation at the Dept. of Computer and Systems Sciences, Stockholm University/KTH, Sweden, 1993.
- [7] P. Johannesson and P. Wohed, “Improving Quality in Conceptual Modelling by the Use of Schema Transformations”, in Thalheim, B. (Ed.), *Proc. of 15th Int. Conference on Conceptual Modeling ER96*, LNCS 1157, Springer-Verlag, 1996
- [8] S. Purao, “APSARA: A Tool to Automate Systems design via Intelligent Pattern Retrieval and Synthesis”, *The Data Base for Advances in Information Systems – Fall*, vol. 29, no. 4, 1998.
- [9] S. Purao and V.C. Storey, “Intelligent Support for Retrieval and Synthesis of Patterns for Object-Oriented Design”, in D.W. Embley and R.C. Goldstein (eds.) *Conceptual Modeling – ER’97*, LNCS 1331, Springer, pp. 30-42, 1997.
- [10] P. Wohed, “Conceptual Patterns for Reuse of Information Systems Design”, in B. Wangler and L. Bergman (Eds.), *Proc. of 12th Int. Conf. on Advanced Information Systems Engineering – CAiSE 2000*, LNCS 1789, pp 157-175, Springer-Verlag, 2000.
- [11] P. Wohed, “Conceptual Patterns – a Consolidation of Coad’s and Wohed’s Approaches”, to appear in M. Bouzeghoub, Z. Jedad and E. Métails (eds.), *Proc. of the 5th International Conference on Application of Natural Language to Information Systems – NLDB’2000*, LNCS, Springer-Verlag, 2000.