User Modelling for Live Help Systems

Johan Aberg, Nahid Shahmehri, and Dennis Maciuszek

Department of Computer and Information Science
Linköpings universitet, S-581 83 Linköping, Sweden
E-mail: {johab,nahsh}@ida.liu.se
Phone, Fax: +46-13-281465, +46-13-282666

Abstract. We have explored the role of user modelling in live help systems for e-commerce web sites. There are several potential benefits with user modelling in this context: 1) Human assistants can use the personal information in the user models to provide the users with efficient support tailored to their personal needs; 2) Assistants can be more comfortable in their supporting role; 3) Consultation resources can be saved, and thus, financial savings can be made for the e-commerce company. A user modelling approach has been implemented and deployed in a real web environment as part of a live help system. Following the deployment we have analysed consultation dialogue logs and answers to a questionnaire for participating assistants. This paper elaborates on these results, which show that assistants consider user modelling to be helpful and that consultation dialogues can be an important source for user model data collection. Based on lessons learned from the study, future directions for research and development are carefully analysed and laid out.

1 Introduction

It has been shown that customer service has a positive influence on e-commerce. For example, in [22] it is suggested that customer service has a positive effect on user attitudes toward Internet catalogue shopping. Further, it has been suggested that customer service is of great importance for a web site's credibility [11]. Still, the current state of practise in customer service for e-commerce is limited and in need of improvements [15, 21].

In our previous work we have introduced a general model for customer service for web sites [1], now referred to as a model for *live help*[1]. The model features a combination of human assistants and computer-based support. We propose a flexible user interface where users can select how they want to interact with the system. For example, users can choose whether they only want computer-based customer service or if they prefer to chat with human assistants via text chat, voice chat, or other means of interaction. In our model, we also aim at providing personalised customer service by employing user modelling.

[1] Originally we used the term *web assistant system*. However, similar system have recently begun to appear on e-commerce sites, commonly referred to as live help systems. To avoid future confusion we now refer to our work using this newly adopted terminology.

There are several potential benefits with user modelling for live help systems. Knowledge about the user can allow a human assistant to provide high quality and personalised support to the individual user [12]. User modelling can also allow human assistants to be more comfortable in their supporting role, simply because the information in the user model can make them feel familiar with the user. Further, user models can make help sessions more efficient and the dialogues smoother, because the assistants do not have to ask the user for the same information over and over. In [6] an example is presented illustrating the potential financial savings to be made for a company employing a kind of live help system, due to the shorter dialogue time: assuming a modest 20 second reduction per help session, a large company can save $1.5M per year, under realistic conditions.

We have studied our proposed model in a two-step project. In step 1 we explored the value of our model in an e-commerce setting, and we conducted an exploratory usability study based on a limited prototype implementation designed for communication between a user and an assistant [1]. In general, the user feedback was very positive, and we found indications that a user modelling tool would be of help for assistants. Thus, we decided to continue our study in a second step.

In step 2, our main aim was to test the technical feasibility of the live help model. To do this we implemented an instance of the full model and deployed it at an existing web site for a three-week period.

This paper is an extension of a previous short paper [4], presented at the ACM conference on electronic commerce[2]. The focus of this paper is on the study of the user modelling component that was part of step 2. We explore the value and feasibility of user modelling for live help systems. Apart from testing technical feasibility we focus on two main questions: 1) What are the subjective opinions of assistants towards the concept of such a user modelling tool? 2) What kind and amount of user model data can be collected from consultation dialogues, and what are the linguistic characteristics for the dialogues? We are also looking into future directions for research and development in some detail, based on the lessons learned from our study.

Positive feedback from assistants regarding question 1 means that this kind of user modelling can be a valuable component of a live help system. Negative feedback on the other hand means that we must question the value of user modelling for live help systems. The importance of acquiring user model data is highlighted by question 2. Consultation dialogues have the potential to be a rich source for user model data acquisition, and can be a complement or a replacement for other sources such as product ratings or registration forms. The linguistic characteristics of the dialogues are of importance for the automatic extraction of user data.

This paper is structured as follows. In section 2 we give a brief overview description of the live help system, and in section 3 we provide a detailed pre-

[2] The present paper provides a much more detailed presentation of the results. We have also added the treatment on future directions.

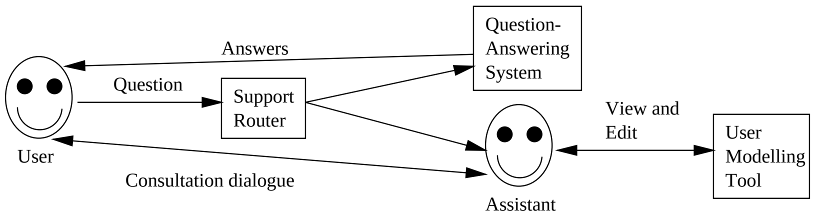

Fig. 1. Overview of the live help model

sentation of our user modelling approach. Section 4 describes the field study involving the user modelling system and section 5 presents the results. In section 6 we comment on limitations of the study, and in section 7 we present related work. Section 8 analyses three important directions for future work, and section 9 concludes the paper.

2 Live Help System

An overview illustration of our live help model is presented in Figure 1. The support router is responsible for deciding whether the user needs computer-based support or support by a human assistant. The computer-based support is a question-answering system. If the support router connects the user to a human assistant they can have a real-time consultation dialogue. A user modelling tool for supporting the assistant is also part of the model.

In our implementation of the model the support router always routes the user through the question-answering system before connecting to a human assistant. The question-answering system is implemented using an information retrieval approach with frequently asked questions (FAQs) [2]. The user modelling component is the focus of this paper and will be further described in the next section.

The user's support process is initiated when the user asks a question in natural language. The question is fed to the automatic question-answering system. FAQ items which closely match are returned as potential answers to the question. If the user indicates that the returned answers are not satisfactory, the support router will connect the user to a human assistant with expertise matching the topic of the question. If all the appropriate assistants are currently busy, the user can choose to wait in a queue. Once an assistant is available the user is connected to that assistant and can proceed with a consultation dialogue via textual chat.

The implemented live help system has been evaluated from the users' point of view in [3]. The findings are very encouraging, particularly when it comes to users' attitudes.

3 User Modelling Approach

Information about a user is stored in a predefined attribute hierarchy, in an overlay style. A user's model is displayed for an assistant as soon as a consultation dialogue begins. The assistant can then make use of the information in the model to tailor the consultation to that individual user. No automatic inference is made on the data in the user model, although the assistant is of course free to make inferences as a part of his or her interpretation of the user data. The assistant can also update the model by filling in attribute values based on what is learned from the consultation dialogue with the user. Further, some basic demographic information (age, gender, and country) is automatically inserted in the user model via questions in a registration phase for the live help system (not shown in Figure 1).

We have chosen a simple approach, and there are two reasons for this. First, we look into the general value of this kind of user modelling tool. If we get positive results we can continue to explore technical issues and more advanced designs in a next step. Second, our aim to evaluate the system in a field study requires a simple system that voluntary assistants can take up with minimal instructions and training.

To find out what kind of user attributes would be most useful for the assistants, we ran a user poll at the web site of our field study (the site is called Elfwood and is in the art and literature domain). In the poll, we asked what kind of questions users wanted to ask in a live help system. Most users wanted help with art creation or help with finding interesting art and literature.

Based on the poll results, we decided to let the detail level of the attribute hierarchy roughly correspond to the number of questions expected for that attribute category. Our assumption was that a detailed attribute structure would be most useful for categories where a large number of related questions was expected. The user model attribute hierarchy is illustrated in Figure 2. The bracketed numbers in the figure correspond to the number of times that user data occurred in the consultation dialogues. The relevance of these numbers is discussed in section 5.1.

The tool for viewing and editing a user model is shown in Figure 3. Each attribute is displayed as a rectangular button with the attribute name as a label. Attributes without a corresponding value are shown in grey in the figure. An attribute that has been given a value is shown in black, with the actual value written after the attribute name. The detail level of the display can be adjusted by the assistant by expanding or contracting branches in the tree. By clicking on an attribute button, an editor window is brought up, where the assistant can create a value or change an existing value. The value can be chosen from a predefined value set or be created as an arbitrary text string. Textual comments can also be associated to a value. This feature can be used for explaining a given value.

A user's skill or interests may change over time and therefore it is important for the system to be able to handle this temporal aspect of user modelling. In order to deal with this a history feature is used. An assistant can update an

- Personal data (10) - Age (5), Gender, Country (6), Occupation (1), Name (36), Conversation style
- Elfwood data (3) - Elfwood member (54), Link to art (41), Link to stories (10)
- Art skill (10)
  - Art media (2)
    - Wet - Ink, Oil paint, Watercolour (2), Acrylics
    - Dry - Pencil (7), Coloured pencil (1), Charcoal, Conte, Pastel
    - Digital (2) - Adobe Photoshop (8), MetaCreations Painter, Paintshop Pro (2), Graphics tablets (4), 3D programs
  - Art objects (6) - Humans (6), Animals, Buildings, Nature
  - Art styles (2) - Realism (1), Anime/Manga (7), Impressionist, Art nouveau
  - Art techniques (13) - Perspective, Sketching (3), Detail drawing
- Writing skill (9)
  - Writing styles (1) - Humour, Serious writing, Fantasy (3), Sci-fi (2), Horror
  - Writing technical (1) - Grammar, Characters, Setting, Plot, Point of view
- Elfwood skill
  - Site navigation - Pictures, Stories
  - Member functions - Intranet, Tour creation, Picture upload, FARP (creation)
  - User functions - Text search, Attribute search, FARP (usage), FantasyHoo
- Computer skill (1) - Internet (1), Scanners (3), MS Windows (1), Linux, Unix

Fig. 2. The complete user model attribute hierarchy

attribute that already has been assigned a value. The old value is then stored in a history file associated with that particular attribute. The history file is shown in the attribute editor.

A somewhat controversial design decision was to hide the user model from the users in the sense that users had no tool available for viewing or updating their own models. The reason was purely technical. We thought that adding such tools to the users' web clients would make the help system more complex and error prone, and thus risk that users lose interest in using the system. Still, we recognise the many advantages of making the user models public (e.g. [8]) and consider this as an important aspect of our future work. The design decision should not affect the results in this paper since our focus is on the assistants' subjective impressions of the tool concept.

Kass and Finin [17] analysed user modelling for natural language systems. Several dimensions for categorising user models were discussed. According to these dimensions our approach to user models can be classified as follows. Our models are *individual*, *dynamic*, and intended for *long term* and *descriptive* usage. A *single* user is modelled by a *single* user model. Our user models are mainly intended for *providing help and advice* for the user, and for *providing output to the user*. Further, the information in a user model can be classified into several categories. We model *capabilities* and *knowledge and belief*. We also model what we call personal data, such as name and age of a user.

4 Field study

The overall research objective in step 2 has been to test the technical feasibility of our live help model. Therefore, a field study, where the system is tested in a real environment, is a natural research method. Consequently, the user modelling

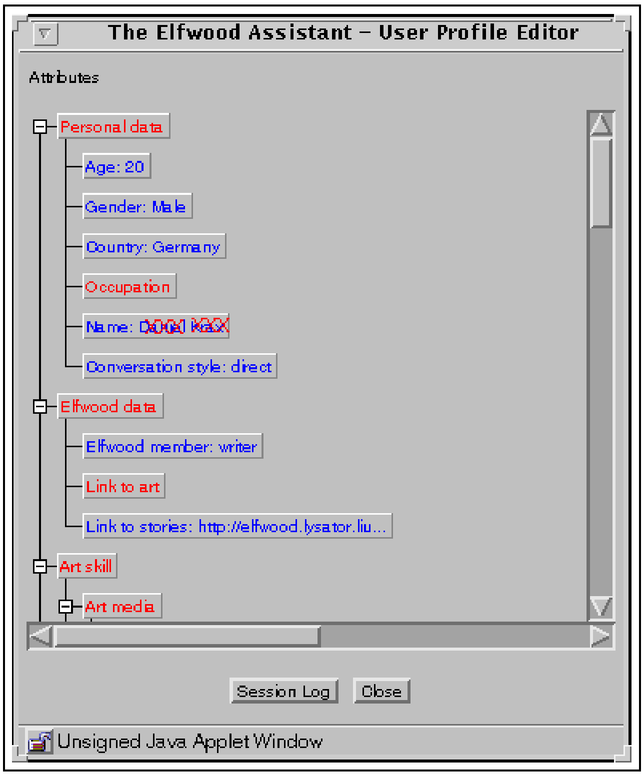

Fig. 3. Screen shot from the user model viewer

part of the live help system is also evaluated in this way. The field study consisted of two parts: system deployment for data collection, and data analysis.

4.1 Environment

The live help system has been attached to an existing web site for a period of three weeks. The site, called Elfwood, is a non-profit site with a focus on art and literature, where amateur artists and writers can exhibit their material. At the time of writing around 9400 artists and 1900 writers exhibit their work in the fantasy and science fiction genre. The artists and writers are members of the site with access to an extranet for updating their own exhibition galleries. A part of the site is devoted to teaching art, and offers a large number of feature articles on different topics.

Elfwood has around 14,500 daily visitor sessions (many are by non-members), where each session averages approximately 35 minutes. About 60% of the sessions are conducted by users from the US. The remaining users are mainly from Europe and Canada.

We mainly supported three types of user tasks. 1) Learning how to create art and literature related to fantasy and science fiction. 2) Searching for interesting art and literature at Elfwood. 3) Member activities, such as uploading new art and literature, and the management of each member's exhibition area.

We chose Elfwood as the environment for our study for two reasons. First, we wanted a site with a reasonable number of users and user traffic, and with a user community that would allow the recruitment of suitable assistants. Second, we wanted to test our system in a low risk environment where unexpected system problems would not have large financial consequences. This meant that we could not go for an e-commerce site at this early stage of the research. Still, we acknowledge the importance of continuing research in real e-commerce settings in the future.

4.2 Participants

Voluntary assistants participated in the study from their home or work environment. They were recruited some months before the field study began and they were all Elfwood members. In the end 30 persons with proper expertise served as assistants. The live help system was not designed to allow multiple simultaneous consultation dialogues for assistants, so no assistant helped more than one user at a time.

During the field study, 636 users registered with the system, and 129 of these users worked through the system to have consultation dialogues with assistants.

4.3 Data collection and analysis

In this study we have used two main data sources, namely the logs of the consultation dialogues and a questionnaire for the assistants. While our current focus is on the subjective opinions of assistants, it is also desirable to study the users' perspective. Such a study can best be pursued by conducting a controlled experiment including a control group without user modelling.

Dialogue analysis During the three weeks of the data collection period a total of 175 consultation dialogues took place. We have analysed the dialogue logs in order to answer the following questions. 1) *How much user model data, and what type of data can be collected from help dialogues?* 2) *In what conversational circumstances does the user provide user model data?* 3) *What are the chat language characteristics for the consultation dialogues?*

In investigating these questions we evaluate how useful consultation dialogues can be as a source of acquiring information about the user. We are also looking for conversational strategies to aid assistants in optimising the amount and quality of user model data that can be acquired. Further, knowledge about the different conversational circumstances in which user model data comes up, and knowledge about the chat language characteristics, is of importance for the automatic extraction of this data.

The following paragraph illustrates the start of the dialogue part of a log file. We logged the exact time that each utterance reached the chat server. We also logged the times when the assistant began a typing session on the keyboard. (This information was used as an awareness cue in the user interface.)

```
<time> 0:30:31 <chat starts>
<assistant id> *****
<time> 0:30:41 <assistant typing>
<time> 0:30:41 <assistant> hi
<time> 0:30:56 <assistant typing>
<time> 0:31:1 <user> Hi.
<time> 0:31:15 <assistant> May I ask your name first?
<time> 0:31:20 <user> I can't think of any good ideas for backgrounds for my
drawings.
<time> 0:31:24 <assistant typing>
<time> 0:31:39 <user> My user or name or my real first name?
```

The methods we have used for investigating each question are as follows. For question 1 we have simply counted the number of times that a user made a statement about himself or herself. We have also matched the type of the statement to the existing hierarchy of user model attributes. Note that only statements of potential long term interest for the tasks that we intended to support are counted. Issues of only short term interest need not be collected for the user model since they have no future value. Thus certain data, for example, data related to the user's short term goals, is not considered.

For question 2 we have followed the work of Elzer and colleagues [10]. They identified four different conversational circumstances for user model data collection from dialogues. *Reject-Solution* (Rej-Soln) is when a user rejects a proposed solution and motivates the rejection by giving a piece of personal information. *Volunteered-Background* (Vol-Back) occurs when the user provides some personal data as part of a problem description. *Volunteered* (Vol) happens when a user volunteers personal information in a conversation without being prompted to do so. *Question-and-Answer* (Q-A) is when the user gives some personal information in response to a question from the system (an assistant in this case). We classified each utterance containing user model data according to these circumstances and summarise the results quantitatively.

For question 3 we have considered the language characteristics mainly in terms of grammar, spelling, and message order.

Assistant questionnaire In order to also get a subjective view from the assistants, we employed a questionnaire. This questionnaire contained items related to various aspects of the user modelling system. All 30 assistants who had participated in the field study received the questionnaire by e-mail just after the data collection period was over. We received 22 answers. Unfortunately, only 14 of these were answered completely, which corresponds to a response rate of roughly 47%. The respondents were from North America (64%), Europe (29%), and Oceania (7%). The assistants were from different age groups: 10-19 (43%), 20-29 (36%), 30-39 (14%), and 40-49 (7%). A total of 57% were female. All respondents had at least 3 years of Internet experience. Most assistants used dial-up connection to the Internet (57%), while some used a cable modem (21%), and

Circumstance	Total no. occurrences	Dialogue average
Q-A	134 (50.8%)	0.77
Vol	67 (25.4%)	0.38
Vol-Back	51 (19.3%)	0.29
Rej-Soln	12 (4.5%)	0.07
All circumstances	264	1.51

Table 1. Conversational circumstances statistics

others had a direct cable access (21%). The responding assistants participated in 7.4 consultation dialogues on average.

There were two types of questions in the questionnaire. The first type asked the respondent to rate a statement using a 1 to 10 scale. The second type listed a number of alternative statements and asked the respondent to rank their relative importance. For each question, the respondent was asked to explain the answer.

5 Results

5.1 Dialogue analysis

How much user model data, and what type of data can be collected from help dialogues? We found a total of 264 user statements containing personal information that could be used for user modelling purposes. In this total we do not count information about a user that is out of the scope of the tasks for the live help system. There were 175 consultation dialogues in total, which means that each dialogue revealed on average 1.51 pieces of information about a user. Regarding the type of user model data that can be collected from consultation dialogues we present the distribution over the attribute hierarchy in Figure 2 (note that the figure appears a few pages back). The numbers in brackets that appear after some of the attributes correspond to the number of times that related user model data occurred in the dialogues. In some cases user model data did not fit into any of the most specific attributes. Then we placed the data in the least general attribute that fit with the data. For example, a user's e-mail address does not fit into any of the sub-attributes of "Personal data", and thus we placed the data in "Personal data".

In what conversational circumstances does the user provide user model data? We found no pieces of information given by a user that did not fit into the set of conversational circumstances previously described. This indicates that they are fairly complete. Statistics from the conversational circumstance analysis is summarised in Table 1. Note that more than 50% of the user data comes from a question by an assistant. This indicates that assistants have an important role to play in user model data collection.

What are the chat language characteristics for the consultation dialogues? We observed that the dialogue language is notably informal. The fact that messages

Statement	Mean	S. dev.
In general the user models were (no help=1, helpful=10)	6.93	2.95
The amount of time available to view user model data during a help session was (on average) (too limited=1, enough=10)	7.79	2.91
How did the value of a user model change as more data was added to it? (less helpful=1, more helpful=10)	8.29	2.09
The number of times that a user model gave me wrong assumptions about a user were (few=1, many=10)	1.57	0.85
Extracting information from a help conversation for insertion in the user model was (hard=1, easy=10)	5.79	2.94

Table 2. Questionnaire results about the user modelling system

are sent as a whole, and not character by character, means that there is a time lag between a message being written and a message being read. Also, a certain space between two related messages may be filled by other information. The answer to a question or the rejection of a solution can reach the recipient right away or several lines later, with a "gap" of unrelated text in between. Furthermore, answers and rejections need not occur in the same order as their questions and solutions. Another striking aspect of the dialogues is that they are generally full of misspelled words, grammar mistakes and incomplete sentences. These results have implications for automatic information extraction, and we discuss this issue further in section 8.1.

5.2 Assistant questionnaire

In Table 2 we present selected results from the rating statements in the assistant questionnaire. The statements left out of the presentation were all concerned with the usability of particular functions of the user modelling tool. Since our focus here is on the concept of this kind of user modelling tool and not on the actual implementation, they are of limited interest in this context.

Through ranking statements in the questionnaire, we learned that the most useful attributes were the Personal data attributes, the Elfwood data attributes, and the Art skill attributes. The Personal data attributes were useful in the sense that the assistants could know the user's name and thus have a more personal dialogue. One assistant also mentioned that he sometimes tailored his help based on the Age attribute by using the heuristic: old users are more experienced than young users.

An additional source of user information available at Elfwood was the members' own display of art and literature. Two Elfwood data attributes in the user models represented links to such user info. Links to a user's art or literature proved very valuable for the assistants as it gave a concrete indication of the user's art or writing skill.

It is important to analyse how the user modelling system was helpful to the assistants. For this purpose we consider the reasons given by the assistants who considered the system to be helpful. Not all these assistants elaborated on their answer, but five of them said that user modelling helped them tailor the help to the individual needs of the user. Another assistant said that the user model data made the dialogues smoother. Yet another assistant said that reviewing the user model reminded her to ask the user questions that were helpful.

6 Limitations

A main limitation of our study is that the data collection period only lasted for three weeks. Consequently, the number of users who used the system more than once was limited. Only 26 out of the 129 users used the system more than once, and thus only a small number of users had updated information in their user model. This means that the subjective opinions of the assistants must be interpreted with some care, since they have a somewhat limited experience with user models containing more information than the Age, Gender, and Country attributes that are present in all models. Another limitation is the low response rate for the assistant questionnaire. This could imply that the results are not representative for all the assistants.

Some of the assistants only got to assist first-time users with empty user models (except for the Age, Country, and Gender attributes). They gave low scores on the question about the helpfulness of user models, stating that the models never contained any helpful information. If we disregard the scores from these three assistants and recompute the mean and standard deviation for the corresponding questionnaire statement, we get a mean of 8.27 and standard deviation of 1.35 instead. This gives a stronger indication of the potential helpfulness of user modelling for assistants.

In summary, these limitations mean that we must consider our results as indications and not as proofs. A natural further step would be to use controlled experiments to statistically test the value of user modelling, both for users and for assistants.

7 Related Work

Since the initiation of our work, there have been commercial moves toward live help systems. Companies such as LivePerson, FaceTime, Cisco, and Blue Martini, to name a few, now offer commercial systems for human assistance in web sites. These systems have recently been adopted by hundreds of e-commerce sites. This trend confirms the importance of the kind of studies reported in this paper.

The commercial systems now available are clearly similar in spirit to our system. Also, the vendors have realised the potential of user modelling for live help systems. Still, to our knowledge there is limited prior research on user modelling in this context. One exception is the paper by Fridgen and colleagues [12]. They argue for the importance of customer models in the financial services industry, and suggest a process for establishing customer models and deducing user-specific actions. They also discuss different categories of user information that should be modelled. The paper thus complements our work in a nice way. The project reported is at an initial stage, and no evaluation is provided.

It is important that the user modelling system does not influence the dialogues too much. In [7] a study of a bank call centre is presented. It is shown that the assistants' interaction with the computer system shape the dialogue to a large extent. The assistants' task of smoothly integrating their computer support into the consultation dialogues requires skill.

Work has been done on user modelling in automated dialogue systems [17, 23, 16, 10]. The user modelling requirements for automated dialogue systems and for live help systems are somewhat different. For automated dialogue systems, data about a user's short term goals, for example, can be highly relevant in order to interpret the user's question. In contrast, for a live help system it is not important to model this data since the human assistant can interpret the user's question, using his or her human intelligence and domain knowledge. For live help systems, it is most important to model data about the user that is valid over several help sessions. Such data includes the user's preferences, knowledge, and beliefs, as well as personal data. Another difference is the necessity of visualisation of user model data for assistants in live help systems. Note however, that visualisation can also be important for automated dialogue systems using user models (at least for long term models) to allow the users to see their models.

8 Future Directions

8.1 Information extraction

Manual extraction of user model data can be cumbersome, as indicated in the assistant questionnaire results. Having the system relieve human assistants of this task would further improve live help in terms of efficiency and convenience for assistants. The company could save time and money, and the assistants could avoid unnecessary stress. When examining a consultation dialogue, either during chat or its log file, an assistant usually does the following: 1) Discover phrases that reveal information about the user. 2) Associate each phrase with attributes in the user model hierarchy. 3) Update those attributes' values in the individual user model.

Automation of this process is a natural language processing problem. Considering the unstructured language of chat dialogues though, in-depth grammatical and semantic understanding, as performed in automated dialogue systems [16, 10], is currently not feasible. Also considering that our sought-for long term, skill-related information is comparatively vague in nature, in-depth understanding is not necessary either. One solution is *information extraction* (IE). IE algorithms understand natural language texts only partially, but with the clear intention of discovering and storing specific information [9].

Consider the following sample dialogue lines taken from a log file.

```
<assistant> a digital drawing tablet would suit you very nicely..
 [...]
<user> I do have one but, but strangly I prefer a mouse it works better for me
for some reason.
```

Read by a human assistant, he or she could 1) discover the user model phrases `I do have one`, `I prefer a mouse`, and `it works better for me`, 2) associate the first phrase with the attribute Graphics tablets, and the others with Digital, 3) update the user's model with the information that he or she owns a graphics tablet, and that his or her preferred digital drawing instrument is the mouse. Altogether, this involves solving three *tasks* of IE: recognising the domain entity name `mouse` *(named entity recognition)*; replacing the anaphoric terms `one` and `it` by their referred-to entity names `digital drawing tablet` and `mouse` respectively *(coreference resolution)*; making out the relevant descriptions around each entity name, associating the phrases with attributes, and updating those attribute values with respect to their previous values *(template element construction)*.

From the dialogue analysis we observe: 1) Conversational circumstances play an important role in resolving coreferences. In the example above, replacing `one` with the entity name it refers to, shows how, in a Rej-Soln circumstance, reference and referred-to entity name appear in a "rejection/solution pair" of phrases. 2) The dialogue language style has implications for IE parsing. The example includes the misspelling `strangly`, the repeated word `but`, and a missing sentence separator between `mouse` and `it`. IE parsing needs to tolerate these.

All in all, we are enhancing our implemented user modelling tool by integrating a real-time, easily portable IE module. Its basic design will either follow Ragnemalm's *keyword based interpretation* approach [18], or the regular grammar driven "FASTUS" *partial parsing* approach [13]. An evaluation of two prototype implementations will reveal, if the user model acquisition task is best handled by focusing on keywords in the context of other keywords, or by recognising well specified grammar fragments.

8.2 Inference

Data collection for user models typically occurs incrementally, with little pieces of information added over several user interaction sessions. Therefore a user model usually contains limited information, and consequently inference would be valuable. Inference is a process of reasoning under uncertainty that aims at extending and generalising collected user data. For live help systems, inference can be done either by the human assistants or in some automatic way. Doing the inference manually has the disadvantage that it may be time consuming, and demand quite some effort from the assistants. Also, there is a risk of inconsistency. Thus, automatic inference should be investigated as an alternative.

Automatic inference on user model data has been studied previously to some extent in traditional user modelling research. However, before attempting to design an approach for this particular kind of application we should consider the special characteristics that may distinguish it from previous research:

– Explainability. Since the user models are being used by human assistants, explainability is of great importance. An assistant needs to be able to trust the user model data in order to feel comfortable in using it as a basis for

the consultation. The inference system needs to be able to explain the inferences to the assistant whenever the assistant is in doubt about some piece of information.
- Justifiability. The inference system needs to be able to justify inferences. For example, assume that an assistant recommends a user to purchase an expensive product, based on some piece of inferred information about the user. The assistant would then likely have to justify the recommendation to convince the user. Justifiability is somewhat different from explainability, as pointed out in [14]. Even if the inference system could explain some inferred data, the explanation may not be sufficient to justify an action taken based on the inferred data.
- Handling different value types. When assistants update a user model they may want to go outside the scope of predefined (numerical) attribute values by providing a new qualitative value. This can be of importance when special cases occur and the predefined values do not fit. Also, the assistants may want to explain an attribute's value by giving a textual description of why the value fits the user. The inference system needs to be able to handle these different types of qualitative and quantitative values.

One of the earliest attempts to apply inference to user models was the so-called *stereotype* approach [19, 20]. Later, more general approaches to dealing with reasoning under uncertainty have been applied to user modelling. In [14], Anthony Jameson considers three common approaches, namely *bayesian networks*, *Dempster-Shafer theory of evidence*, and *fuzzy logic*. Given the requirements presented above, the suitability of each possible approach to inference needs to be analysed with care, before a solution is attempted. Of course, the particularities of the application at hand must also be taken into consideration. We see this as an important direction for future work.

8.3 Privacy

User modelling in general, and perhaps for e-commerce in particular, raises privacy issues. Several studies have shown that users consider privacy to be of great importance for e-commerce (see e.g. [5]). We informed the users in our field study that they were being modelled, and how the collected data would be used. Of course, they were given the possibility to opt out. We believe that for user modelling to really be useful it is vital that users are informed of the ongoing data collection process, and how the data will be used. The users should also be allowed to be in control of the information that is kept about them. One idea for implementing this is to require web sites to have the user's digital signature on the user model, before being allowed to use it. We consider this an interesting issue for future work.

9 Conclusions

We conclude by considering the two initial questions raised in the introduction. First, the questionnaire results showed that the assistants in fact considered user modelling to be helpful for assisting users. Second, the analysis of the consultation dialogues showed that much important user information of various kinds can be gained through such dialogues. We also identified requirements for automatic extraction of user data.

In section 4.1 we argued for why we did not test our live help system in an e-commerce environment. Still, it is interesting to discuss the generalisability of our results on user modelling towards e-commerce web sites. A fundamental property for user modelling to work in any web site, is to have a large rate of users who frequently return to the web site. Further, our test site was a community oriented web site. Thus, our results are most likely to carry over to the kind of e-commerce web sites that gather a community of users that keep coming back. Amazon is an example of such an e-commerce site that has managed to get the users involved in providing feedback on the products, and created a community of reviewers. Also note that, as we have shown previously [1, 3], a live help system improves users' attitudes toward the site, and it alone may create a "community feeling" and encourage users to revisit the site.

Live help systems are here to stay. The questions we have considered in this paper represent the beginning of one branch of research on live help systems. In our analysis on future directions we have considered three issues that deserve thorough study. There are also other issues of importance: What are the conversational strategies that work best for acquiring user model data? What kind of user data is most important to model, for different kinds of web sites? How can the user models be integrated with the computer-based support?

References

1. Johan Aberg and Nahid Shahmehri. The role of human Web assistants in e-commerce: an analysis and a usability study. *Internet Research: Electronic Networking Applications and Policy*, 10(2):114–125, 2000.
2. Johan Aberg and Nahid Shahmehri. Collection and Exploitation of Expert Knowledge in Web Assistant Systems. In *Proceedings of the 34th Hawaii International Conference on System Sciences*, Maui, Hawaii, USA, January 3-6 2001.
3. Johan Aberg and Nahid Shahmehri. An Empirical Study of Human Web Assistants: Implications for User Support in Web Information Systems. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pages 404–411, Seattle, Washington, USA, 2001.
4. Johan Aberg, Nahid Shahmehri, and Dennis Maciuszek. User Modelling for Live Help Systems: Initial Results. In *Proceedings of the Third ACM Conference on Electronic Commerce*, Tampa, FL, USA, 2001. In press.
5. Mark S. Ackerman, Lorrie Faith Cranor, and Joseph Reagle. Privacy in E-Commerce: Examining User Scenarios and Privacy Preferences. In *Proceedings of the ACM Conference on Electronic Commerce (EC'99)*, pages 1–8, Denver, Colorado, USA, 1999.
6. Howard G. Bernett and Areg Gharakhanian. Call Center Evolution: Computer Telephone Integration and Web Integration. *The Telecommunications Review, MitreTek Systems*, pages 107–114, 1999.

7. John Bowers and David Martin. Machinery in the New Factories: Interaction and Technology in a Bank's Telephone Call Centre. In *Proceedings of the Conference on Computer-Supported Cooperative Work*, pages 49–58, Philadelphia, PA, USA, 2000.
8. R Cook and J Kay. The justified user model: a viewable, explained user model. In *Proceedings of the Fourth International Conference on User Modeling*, pages 145–150, 1994.
9. Hamish Cunningham. Information Extraction – a User Guide (updated version). Department of Computer Science, University of Sheffield, UK, April 1999.
10. Stephanie Elzer, Jennifer Chu-Carrol, and Sandra Carberry. Recognizing and Utilizing User Preferences in Collaborative Consultation Dialogues. In *Proceedings of the Fourth International Conference on User Modeling*, pages 19–24, 1994.
11. BJ Fogg, Jonathan Marshall, Othman Laraki, Alex Osipovich, Chris Varma, Nicholas Fang, Jyoti Paul, Akshay Rangnekar, John Shon, Preeti Swani, and Marissa Treinen. What Makes Web Sites Credible? A Report on a Large Scale Quantitative Study. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pages 61–68, Seattle, WA, USA, 2001.
12. Michael Fridgen, Jürgen Schackman, and Stefan Volkert. Preference Based Customer Models for Electronic Banking. In *Proceedings of the 8th European Conference on Information Systems*, pages 789–795, Vienna, Austria, 2000.
13. Jerry R. Hobbs, Douglas Appelt, John Bear, David Israel, Megumi Kameyama, Mark Stickel, and Mabry Tyson. FASTUS: A Cascaded Finite-State Transducer for Extracting Information from Natural-Language Text. In *Finite State Devices for Natural Language Processing*, pages 383–406. MIT Press, Cambridge, MA, USA, 1996.
14. Anthony Jameson. Numerical Uncertainty Management in User and Student Modeling: An Overview of Systems and Issues. *User Modeling and User-Adapted Interaction*, 5:193–251, 1996.
15. Sirrka L. Jarvenpaa and Peter A. Todd. Consumer Reactions to Electronic Shopping on the World Wide Web. *International Journal of Electronic Commerce*, 1(2):59–88, 1997.
16. Robert Kass. Building a User Model Implicitly from a Cooperative Advisory Dialog. *User Modeling and User-Adapted Interaction*, 1(3):203–258, 1991.
17. Robert Kass and Tim Finin. Modeling the User in Natural Language Systems. *Computational Linguistics*, 14(3):5–22, September 1988.
18. Eva L. Ragnemalm. *Student Modelling based on Collaborative Dialogue with a Learning Companion.* PhD thesis, Department of Computer and Information Science, Linköpings universitet, S-581 83 Linköping, Sweden, 1999.
19. Elaine Rich. Users are individuals: individualizing user models. *Int. J. Man-Machine Studies*, 18:199–214, 1983.
20. Elaine Rich. Stereotypes and User Modeling. In W. Wahlster and A. Kobsa, editors, *User Models in Dialog Systems*, pages 35–51. Springer Verlag, 1989.
21. Peter Spiller and Gerald L. Lohse. A Classification of Internet Retail Stores. *International Journal of Electronic Commerce*, 2(2):29–56, 1998.
22. Leo R. Vijaysarathy and Joseph M. Jones. Print and Internet catalog shopping. *Internet Research: Electronic Networking Applications and Policy*, 10(3):191–202, 2000.
23. W. Wahlster and A. Kobsa, editors. *User Models in Dialog Systems.* Springer Verlag, 1989.